Museums Around:
Cultural Infrastructure and Educational Outcomes

Mengwei Lin*

Job Market Paper

This version: November 15, 2026

Click for the latest version.

Abstract

Cultural infrastructure is part of the environment in which people live, and for the young it may shape what they achieve and what they study. Does a museum nearby change educational attainment and choices? Combining administrative records on 8.5 million students in England with geocoded museum openings, I answer this question using a stacked difference-in-differences design. A museum opening within 2 km of students' homes raises their percentile rank in upper-secondary exams by 1.5 points, the effect fading with distance. Subject choices also respond to museum themes: history, social-life, and industry museums raise the share of the cohort entering the related subjects by around 1 percentage point, and science museums raise entry into academic STEM by 2.5 percentage points. Cultural infrastructure in the adolescent environment thus shapes what young people achieve and what they study.

JEL Codes: I21, J24, Z11, R23
Keywords: Museums, Educational attainment, Subject choice, Cultural infrastructure, Neighbourhood effects

*University of Barcelona & Barcelona Institute of Economics (IEB). Email: mengwei.lin@ub.edu.
I am grateful to Andreu Arenas, Jenifer Ruiz-Valenzuela, and Olmo Silva for their guidance and support throughout the project. I thank Sandra McNally and Henry Overman for detailed comments on the draft, and Guglielmo Ventura for generous help in setting up the project. This paper also benefited from conversations at different stages with Sara Bagagli, Hector Blanco, Lex Borghans, Tommaso Colussi, Thomas Cornelissen, Ana Costa-Ramón, Bart de Koning, Ciprian Domnisoru, Alex Eble, Nicolas Gendron-Carrier, Bart Golsteyn, Jordi Jofre-Monseny, Steffen Künn, Jaime Millán-Quijano, Caterina Muratori, Lorenzo Neri, Fabien Petit, Giorgio Pietrabissa, Ana Rodríguez-González, and Rosa Sanchis-Guarner, and from the suggestions of seminar, workshop, and conference participants. This work was undertaken in the Office for National Statistics Secure Research Service using data from ONS and other owners and does not imply the endorsement of the ONS or other data owners. All errors are my own.

1 Introduction

"Real museums are places where Time is transformed into Space." So writes Orhan Pamuk in *The Museum of Innocence*, whose narrator turns years of love into a museum.[1] Societies do the same on a larger scale, building cultural institutions like museums to keep what is known and valued and to pass it on. There are now more than 100,000 of them around the world (UNESCO, 2021), and England, the setting of this paper, committed £760 million to museums in January 2026 (Department for Culture, Media and Sport, 2026). Once built, such institutions become part of the environment of those living around them. For young people, that environment might shape later outcomes, such as educational ones (Chetty et al., 2016; Chyn and Katz, 2021; Laliberté, 2021).

Yet whether museums affect young people around them is mostly unknown.[2] Museums nearby could shape both how well students perform and what they study. First, unlike libraries, parks, or theatres, a museum presents knowledge through original objects, which may stimulate students' intellectual curiosity and improve their academic performance. Second, museums are organised around themes, from science to arts, and thus may steer students towards studying related subjects, a choice they often make with little information about what a field involves. Each effect would speak to two questions that remain open: which features of that environment affect what students achieve, and what shapes their study choices. Evidence is so far scarce because data linking museums to students' outcomes are hard to obtain, and because museums stand where particular families live, making it difficult to identify their effects.

In this paper, I examine whether local cultural infrastructure, in particular museums, affects how well students nearby perform in upper-secondary education and what they choose to study. This adds to what is known about whether the cultural environment shapes educational outcomes, in three ways. It isolates one feature of a place, the cultural input, rather than the neighbourhood as a whole; it shows that the adolescent environment shapes not only how well students do, but also what they study; and it extends the evaluation of cultural investment from local economies to

[1] Orhan Pamuk, novelist and the 2006 Nobel laureate in literature. *The Museum of Innocence*, translated from the Turkish by Maureen Freely (Pamuk, 2009).

[2] No study observes what students choose to study. On attainment, evidence comes from a few small experiments on school visits in the US, finding effects on critical thinking, behaviour and test scores (Greene et al., 2014; Bowen et al., 2014; Erickson et al., 2024), and from surveys, showing family-reported visits not associated with exam results (Sullivan, 2001; Stopforth and Gayle, 2022).

the young people who live nearby. I ask three closely connected questions. Does a museum opening near a student's home change whether the student continues to upper-secondary study?[3] Does it raise the exam results of those who do progress? And does a museum's theme steer students towards the subjects it presents? Each of these margins has lasting consequences for the degrees students enter and what they earn, as discussed below.

To answer these questions, I use the administrative records on the population of students in state-funded schools in England, drawn from the Longitudinal Education Outcomes (LEO) Standard Extract. The student records cover 8.5 million students finishing lower-secondary education from 2002/03 to 2016/17, including each student's home neighbourhood, prior attainment, and the upper-secondary qualifications they enter and achieve. This is combined geographically with the Mapping Museums database, a novel census of museums in the UK open to the public between 1960 and 2020. England has a large and growing museum sector. By 2005, the number of museums had more than doubled, to 2,542 from 1,043 in 1960. The database records each museum's location, opening year, and characteristics such as size and theme.

I measure exposure by the distance between a museum and a student's home, because museum-going by teenagers is common, mostly happens outside school, and mostly involves short trips.[4] I estimate the effects of museum openings using a stacked difference-in-differences design. It compares students whose home neighbourhood gets a new museum within a given distance of home, from 1 to 10 km, to students in neighbourhoods that get one within the same distance several years later, across the cohorts before and after the opening. Using stack-wise not-yet-treated students as controls restricts the comparison to places that all get a museum and share similar characteristics (Section 3), and stack-by-Local-Authority and stack-by-calendar-year fixed effects compare students within the same area and cohort (Wing et al., 2024).

The main threat to this design is that museums may open in places that are changing in other ways. In my setting, this concern is limited by the comparison I use and by what the data show. First, the identifying variation comes from the

[3]Upper-secondary study here refers to qualifications at Level 3 of the national qualifications framework in England, such as academic A-levels and vocational qualifications like BTECs, which give access to university (Section 2.1).

[4]In the DCMS Taking Part survey, between 55% and 64% of children aged 11–15 visited a museum in the previous year and three in four child visitors went in their spare time rather than only with their school (Department for Culture, Media and Sport, 2010–2024), and in the 2008/09 wave where such questions are included, nearly half of all visits involved a journey of at most 30 minutes (Department for Culture, Media and Sport, 2010). See Section 2.3 and Appendix C.

timing of openings between places that get one, rather than from the type of place that gets a museum. Besides, most openings are smaller museums set up when a building, a collection, or funding became available (Appendix B). Additionally, I also check directly that the students in treated areas do not change at the opening. The population and composition of treated neighbourhoods do not evolve differently from those of control neighbourhoods, and neither neighbourhood deprivation nor students' eligibility for free school meals changes after an opening.

I find that a museum opening near students' homes raises upper-secondary attainment but leaves progression to upper-secondary study largely unchanged on average. As I do not observe who visits, the estimates capture the effect of having a museum nearby, whether through visits, through schools' use of the museum or through the wider environment it brings to a neighbourhood, rather than the effect of a visit. Within 2 km, students' average score in upper-secondary exams rises by 1.48 percentile points in the ranking, about 0.05 standard deviations. This effect is comparable to those of other local facilities, as research shows that access to a Sure Start children's centre within 2.5 km in early childhood raises GCSE results by 0.031 standard deviations (Carneiro et al., 2024), and investment in public libraries raises reading scores by 0.01 to 0.04 standard deviations (Gilpin et al., 2024). The effect fades with distance, to 0.80 percentile points within 5 km and 0.15 within 10 km, which fits with how far people travel to museums (Department for Culture, Media and Sport, 2010–2024).

The gains are larger for students eligible for free school meals and for those with lower prior attainment, and progression to upper-secondary study rises only for students in the middle of the prior-attainment distribution, where the margin of continuing lies. Attainment effects also depend on museum size. For small museums, which make up most of the openings, the effects are more locally concentrated, while for larger museums their impacts stay farther away. This dimension of effect heterogeneity also resonates with the finding on fading effects.

Museums also shape students' subject choices. I group museum themes into families and pair each family with the subject group it is most likely to promote: science and nature, and industries and trades, with STEM subjects; history and humanities with humanities subjects; arts and media with the arts; and social life with the social sciences. I measure the upper-secondary subject choices, which students make at the end of the previous stage, as the share of the cohort entering each subject group. Students choose several subjects at once, usually three in the case of A-levels, so a

student can enter several groups, and a rise in one group need not come at the expense of another. With an opening within 5 km, museums of social life and social history raise entry into the social science subjects by over 1 percentage point, about 9% of its pre-opening mean of 13.9%, and history and humanities museums raise entry into the humanities by 0.90 percentage points. Science museums, though with few openings, raise academic STEM entry by 2.51 percentage points. Arts museums do not change entry into arts subjects, possibly because students already encounter the arts in many other ways. Cultural infrastructure in the adolescent environment thus shapes what young people achieve and what they study.

Existing research shows that these outcomes matter well beyond school. Progression to upper-secondary study is the main gateway to higher education and to better jobs: students more likely to progress also become more likely to enter tertiary education (Machin et al., 2020), and completing upper-secondary education raises wages even after accounting for cognitive and non-cognitive skills (Heckman et al., 2018). Attainment in upper-secondary education decides which universities and programmes students can enter, with lasting effects on earnings (Campbell et al., 2022; Canaan and Mouganie, 2018). Subject choices shape what students study and earn afterwards: more science and mathematics at school increases the probability of a STEM degree and later earnings (De Philippis, 2023; Joensen and Nielsen, 2009), earnings differ across the fields chosen in upper-secondary school by as much as the return to two additional years of education (Dahl et al., 2023), and in vocational education the field matters more for earnings than the school (Aucejo et al., 2025).

This paper first contributes to the literature on how the places where children grow up shape their outcomes. In general, children whose families move to better neighbourhoods do better as adults, and exposure to better places continues to matter in adolescence (Chetty and Hendren, 2018; Chetty et al., 2016; Chyn, 2018; Deutscher, 2020; Chyn and Katz, 2021). However, moving changes the whole bundle of a neighbourhood at once and can disrupt children's lives (Chetty et al., 2016; Sanbonmatsu et al., 2006). Therefore, which features of places matter, and whether they can be improved for the children who stay, remains an important open question. In England, neighbours themselves matter little for attainment (Gibbons et al., 2013, 2017), and much of the neighbourhood effect comes from amenities and runs through schools (Laliberté, 2021; Neri, 2024). Recent work has turned to specific local inputs, such as new types of schools (Machin et al., 2025), home broadband (Sanchis-Guarner et al., 2025), youth clubs (Villa, 2024), fast-food restaurants (Abrahamsson et al., 2023)

and, closest to this paper, public libraries, which have been shown to affect children's reading scores (Gilpin et al., 2024; Hanzl and Gilpin, 2025). Evidence on cultural provision is either experimental and limited to school visits (Bowen et al., 2014; Kisida et al., 2014) or correlational, with cultural capital predicting reading performance (Breinholt and Jæger, 2020) but not exam grades (Stopforth and Gayle, 2022). This paper provides causal, population-scale evidence that changing one feature of a place, a museum, affects the educational attainment and choices of adolescents living around it.

Second, this paper contributes to the literature on subject choice. Choices of college majors are well studied (Altonji et al., 2016). They matter for earnings (Kirkeboen et al., 2016), first exposure to a field and role models change them (Fricke et al., 2018; Porter and Serra, 2020), and they depend more on students' tastes than on expected earnings, although where such tastes come from is little understood (Wiswall and Zafar, 2015). Subject choices in secondary school come earlier and constrain these later choices (Delaney and Devereux, 2019). The evidence on what shapes them comes mostly from inside the school, where peers, rank, and teachers matter (Legewie and DiPrete, 2014; Mouganie and Wang, 2020; Murphy and Weinhardt, 2020; Carlana, 2019), and from families and role models (Dahl et al., 2024; Joensen and Nielsen, 2018; Breda et al., 2023), and most of the literature concerns STEM and the gender gap in it (for example, see Cavaglia et al., 2020, for England). Evidence on the impact of the environment outside side has only begun to appear. For instance, girls assigned to a female family doctor become more likely to choose a STEM track (Riise et al., 2022).[5] To my knowledge, no study links local environment to what students choose to study. This paper provides such evidence across subject groups, in the direction of what museums present, mirroring the content-specific exposure effects that Bell et al. (2019) document for innovation.

Third, this paper contributes to the evaluation of place-based policies and cultural amenities. Place-based investments seek to improve the welfare of the people who live there, but evaluations in the UK often find that such investments change places more than they change residents' outcomes. For instance, in the UK, the Single Regeneration Budget raised the number of jobs located in target neighbourhoods but did little for the employment of the people living there (Gibbons et al., 2021), and support for local businesses displaced employment from nearby areas (Einiö and Over-

[5]For college majors, Kim et al. (2025) find that children who grow up in counties with more STEM workers are more likely to major in STEM.

man, 2020). Kline and Moretti (2014), Neumark and Simpson (2015), and Ehrlich and Overman (2020) provide broader international evidence and welfare arguments. Culture is a common object of such investments, yet existing evaluations concern local economies rather than residents. Cultural amenities attract high-skilled workers (Falck et al., 2011) and raise nearby property values (Been et al., 2016), and the European Capital of Culture programme raised the GDP of host regions (Gomes and Librero-Cano, 2018). Of the 36 evaluations of sport and culture projects reviewed by the What Works Centre for Local Economic Growth, only three concerned culture (What Works Centre for Local Economic Growth, 2016). This paper evaluates cultural infrastructure by its effects on young people, and shows that what an amenity offers matters for these effects.

The rest of the paper proceeds as follows. Section 2 describes the institutional setting and the data, Section 3 presents the empirical strategy, Section 4 reports the results, and Section 5 concludes.

2 Institutional Setting and Data

To study how museum openings affect students' upper-secondary outcomes, I combine two data sources: administrative records on the population of students in state-funded schools in England, drawn from the Longitudinal Education Outcomes (LEO) Standard Extract, and the Mapping Museums UK dataset, a census of museums in the UK. The two are linked through geography, by the distance between a student's home neighbourhood and each museum. This section first describes upper-secondary education in England and the student-level dataset (Section 2.1), then the museum sector and the museum data, including the openings that generate the treatment (Section 2.2), how young people visit museums and how I link museums to students (Section 2.3), and finally the analytic sample (Section 2.4).

2.1 Upper-Secondary Education and Student Records

Education in England is organised into "Key Stages", of which two important checkpoints are the national tests at the end of Key Stage 2 (KS2), or primary school, and the General Certificate of Secondary Education (GCSE) exams in Year 11, at the end of Key Stage 4 (KS4) when students are 16 (Hupkau et al., 2017; Machin et al., 2020). After Year 11, students decide whether and what to study in Years 12 and 13 in Key Stage 5 (KS5): the best-known Level-3 qualifications are the A-levels, academic qual-

ifications that students usually take in three subjects, chosen from over 80 (Ofqual, 2018), to study over two years before entering university; vocational Level-3 qualifications, such as the BTECs, provide learning in various vocational areas such as applied science and technology, business, and health and social care, and students can also take courses at lower levels or start apprenticeships (Hupkau et al., 2017; Campbell et al., 2022).[6] A-levels are mostly taken in school sixth forms and sixth-form colleges, whereas further-education colleges provide most vocational courses (Hupkau et al., 2017).

The LEO Standard Extract is an administrative database that links the National Pupil Database (NPD) to further-education, higher-education, tax and benefit records (Department for Education et al., 2026). For this paper, I use its NPD component, which covers all pupils in state-funded schools in England. The NPD tables are linked through an anonymised pupil identifier: the spring School Census, which records each pupil's home Lower Layer Super Output Area (LSOA) and characteristics; the KS4 tables, based on which the population of a cohort is defined and which record KS4 student characteristics and prior attainment at KS2; and the KS5 tables, which record the Level-3 qualifications entered and achieved as well as the point scores.

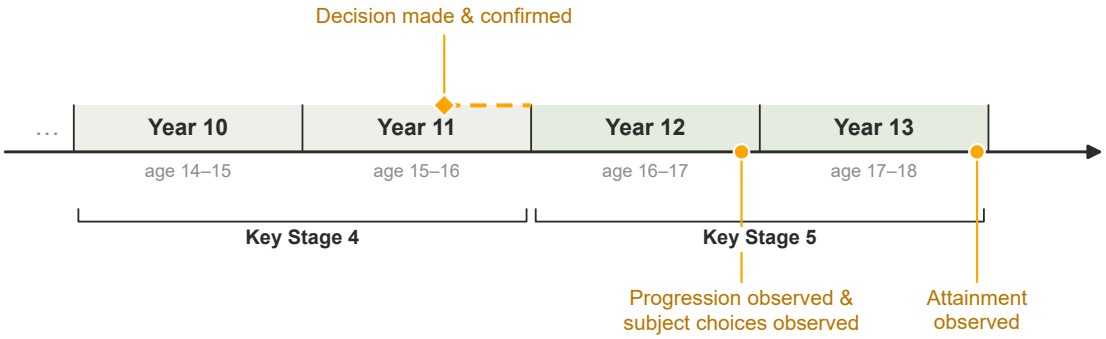
Students make decisions about their upper-secondary education by the end of KS4, which are confirmed upon entering KS5 (Figure 1). Because choices are made during Year 11, I organise the data around Year-11 cohorts. The analysis covers the Year-11 cohorts 2002/03 to 2016/17, that is, students who were in the spring semester of Year 10 in the calendar years 2002 to 2016. The 2001/02 cohort is excluded because its Year-10 School Census predates the first census in the extract and its KS4 records carry no KS2 prior attainment. The 2017/18 and 2018/19 cohorts are excluded because their KS5 examinations were replaced by centre- and teacher-assessed grades during the COVID-19 pandemic (see also Section 3 for an explanation of the sample period from the perspective of the method).

I locate each student at the LSOA of residence recorded in the spring School Census of Year 10, the calendar year before KS5. LSOAs are small census geographies with about 1,500 residents each; England has 32,844 LSOAs on 2011 boundaries, and I convert LSOAs recorded on 2001 boundaries to 2011 boundaries with the ONS lookup. After dropping students whose Year-10 home LSOA is missing and residents

[6]Counting distinct subjects across awarding bodies, Ofqual's market statistics record A-level certificates in 82 subjects in England in both 2012/13 and 2016/17, which official statistics group into around 40 subject groups; after the reform of A-levels from 2015, the number fell to 57 subjects by 2020/21 (Ofqual, 2018, 2026).

of Wales or Scotland who attend schools in England, the sample contains 8,477,187 students. Of these, 7.57 million have KS2 records, and 4.62 million enter Level 3 and therefore have a KS5 point score.

Figure 1. Timing of Decisions and Outcomes in Key Stages 4 and 5



Notes: This figure shows school years 10–13 (ages 14–18). Pupils sit GCSEs at the end of Year 11 and A-levels or other Level-3 qualifications at the end of Year 13. Applications for Level-3 courses are made in the spring of Year 11 (diamond) and confirmed at enrolment in September (dashed window). Progression and subject choices are measured from the qualifications entered in the first Key Stage 5 year, and attainment at the end of Year 13.

Progression to Level 3. Progression to Level 3 is easiest for students with good GCSE results, in particular with at least a grade C in English and maths (Hupkau et al., 2017; Machin et al., 2020). Students who finished Year 11 in 2013 were the first required to remain in some form of education or training, which can be part-time, until age 17, and the requirement was extended to age 18 from 2015 (Machin et al., 2020; Hupkau et al., 2017; Matthewes and Ventura, 2022). Because the requirement can also be met through an apprenticeship or through work combined with part-time study (Department for Education, 2024), studying at Level 3 remains a choice throughout the period.

I measure progression with two indicators defined over the whole Year-11 cohort, so that students without any KS5 record count as zero. The first measures *whether* the student proceeds to Level 3, which equals one if the student has any Level-3 entry in the KS5 records. The second measures whether the student achieves a *full* Level-3 programme, which equals one if they achieve passes equivalent in size to at least two A-levels; because the DfE indicator of this outcome changes definition in 2015/16, I also count students with at least two Level-3 passes, following Machin et al. (2020).

In the sample, 54.5% of students enter Level 3 and 45.8% complete a full Level-3 programme (Panel B of Table 1).

Attainment. Grades at KS5 are converted into points for the DfE performance tables. Until 2014/15, an A-level grade between A and E was worth between 270 and 150 points (Campbell et al., 2022); from 2015/16, the points range from 60 for an A* to 10 for an E (Department for Education, 2018), which reduced average points per entry roughly sevenfold.

I measure KS5 attainment by the average point score per Level-3 entry, the total Level-3 points divided by the number of Level-3 entries, in the student's KS5 outcome year. The outcome year is the academic year in which the student's total Level-3 point score is highest, or the later year in case of a tie. This measure is defined only for students who enter Level 3. To ensure comparability across years, and especially before and after 2015/16, I convert the score into a percentile rank. All percentile ranks in the paper are ranks within a cohort, but the cohort is defined at the stage at which the variable is measured: KS2 attainment and the IDACI rank of the home LSOA are ranked within the Year-11 (KS4) cohort, while KS5 scores are ranked within the KS5 cohort, defined by the academic year of the KS5 outcome.

Subject choice. KS5 students are allowed to mix academic and vocational subjects, without having to specialise in "tracks", although most students take mainly one type of qualification (Hupkau et al., 2017; Matthewes and Ventura, 2022). Each qualification is classified into one of the 15 Sector Subject Areas (SSAs) of the English qualification system, which I group into four subject groups: STEM (Science and Mathematics; Engineering and Manufacturing Technologies; Information and Communication Technology), Humanities (History, Philosophy and Theology; Languages, Literature and Culture), Arts (Arts, Media and Publishing), and Social Sciences.[7]

I measure subject choices from the Level-3 qualifications entered in the first KS5 year, the year in which the choice made in Year 11 takes effect. For each subject group, I define two indicators over the whole Year-11 cohort: *any entry*, which equals one if the student enters any Level-3 qualification in the group, and *academic*, which equals one if the student enters such a qualification conditional on having at least one academic qualification, conceptually similar to being on an academic track. I

[7]Here the STEM group differs slightly from the conventional STEM definition, as the Science and Mathematics SSA also includes psychology-related subjects.

also use indicators for holding any academic qualification and any vocational Level-3 qualification, which 44.2% and 20.7% of students do; the two are not mutually exclusive for the reason specified above.

Covariates. The School Census and the KS4 tables provide the student characteristics used as controls: gender, eligibility for free school meals (FSM), special educational needs (SEN), ethnicity, and whether English is the student's first language (EAL), all recorded in Year 11. I also control for prior attainment at KS2, and for the Income Deprivation Affecting Children Index (IDACI) of the home LSOA, both expressed as percentile ranks within the Year-11 cohort, and for whether the home LSOA is urban according to the ONS 2011 Rural–Urban Classification (Bibby and Brindley, 2013).

2.2 Museums in England and the Mapping Museums Data

The information on museums that I use comes from the Mapping Museums database, a census of UK museums open to the public between 1960 and 2020 (Candlin et al., 2020). I use the September 2021 release (Poulovassilis et al., 2021), restricted to the 3,163 museums located in England. It defines a venue as a museum if it holds a permanent collection, displays items from it, opens to the public regularly, and occupies a defined space (Candlin and Larkin, 2020).[8] To my knowledge, this paper is the first to use the database in economics; previous quantitative uses describe the geography and growth of the sector (Ballatore and Candlin, 2023; Candlin et al., 2024) or count museums for official statistics (Office for National Statistics, 2024).

England has a large and growing museum sector. By 2005, right before the period of treatment events used in the identification, the number of museums had increased to 2,542 from 1,043 in 1960, more than a doubling.[9] At this point in time, 71% of the museums operating in England were independent, most of them run by not-for-profit trusts, 24% were run by governments, mainly by local authorities, and 53% were small local museums, with at most 10,000 visits a year (Appendix Table B3

[8]The project assembled the census from many sources, including the lists of the Arts Council England accreditation scheme, the membership lists of the Association of Independent Museums, earlier national surveys, historic directories and guidebooks, regional lists, and online sources, and cross-checked the entries against each other, by telephone and e-mail, and with regional museum bodies (Candlin et al., 2020; Poulovassilis et al., 2019).

[9]The number for 1960 is an upper bound, since 173 museums only have an opening range starting from 1960, the earliest year covered by the database. If I instead spread each uncertain year evenly over its range, there were about 850 museums in 1960 and 2,450 in 2005, nearly a tripling.

presents detailed information on the museums at this point). Museums also differ widely in what they present, from local society to arts and history, and from industry and manufacturing to the natural world. Whether a museum is nearby, and of which kind, is therefore part of the environment in which a student grows up, and this environment changes as new museums open.

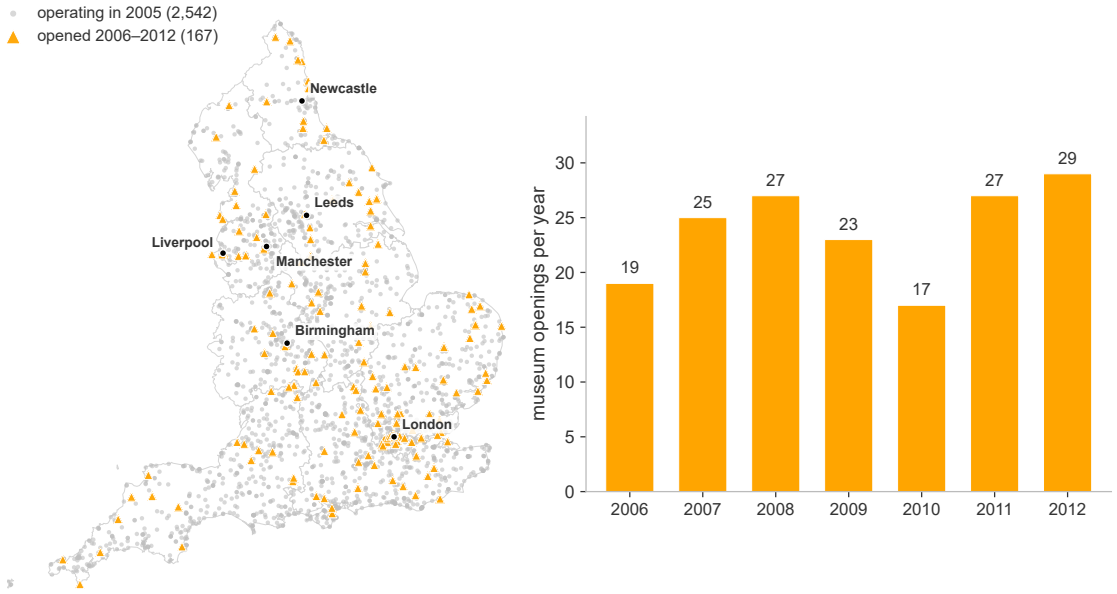
For each museum, the database records its location, the years in which it opened and closed (if so), its governance, its size, and its subject matter. Locations are geocoded from postcodes and checked by hand by the Mapping Museums project team (Ballatore and Candlin, 2023). Governance distinguishes government, university, and independent museums, with independent museums split into not-for-profit, private, National Trust, and English Heritage. Size is banded by annual visits into small (up to 10,000), medium (10,001 to 50,000), large (50,001 to 1 million), and huge (over 1 million); where visitor numbers are unavailable, the project imputes the band from the other attributes (Ballatore and Candlin, 2023). It is worth noting that small, independent museums account for most of the treatment variation I use later on for identification. Subject matter is classified hierarchically into 21 top-level categories, which I group into 7 broader categories so that each has enough openings for estimation. To avoid confusion with the educational subjects, I refer to the subject matter of a museum as the "theme" hereafter. These attributes describe each museum's status available at the time of the data collection and do not vary over time, so I use them as fixed characteristics of a museum.

Most importantly for this paper, the Mapping Museums database records an exact opening year when it can be pinned down, or a range of years if not (Candlin and Poulovassilis, 2020). Of the 3,163 museums in England, 2,806 have an exact opening year. I treat an opening as an event if its year is exact, and all 167 openings from 2006 to 2012, the opening years that form the stacks of the design (Section 3), have an exact year.

Figure 2 maps the museums operating in 2005 and the 167 openings of 2006–2012. With between 17 and 29 museums opening in each of these years across places, the design can exploit the staggered variation, as explained later in Section 3. Specifically, the openings took place in all nine regions of England, two-thirds of them in urban LSOAs. Compared with the 2005 stock, the new museums are more often small (73% against 53%) and less often run by government (9% against 24%), and 89% of them opened in an existing building rather than a new one. Many openings brought a museum to places that had none nearby: in its opening year, 40.1% of the new

museums were the first museum within 2 km and 15.6% within 5 km. Appendix B provides a closer look at the museum openings that generate the treatment variation.

Figure 2. Museum Stock in 2005 and Museum Openings from 2006 to 2012, England



Notes: The left panel maps the 2,542 museums operating in 2005 (grey) and the 167 museums opened in 2006–2012 (orange). The right panel shows these openings by year. They are the treatment events: the opening cohorts for which a full four-year pre- and post-opening window fits within the outcome data. Appendix Table A1 reports the yearly counts.

To relate museums to subject choices, I group the themes into seven theme families according to the field of knowledge a museum mainly presents. Figure 3 shows the grouping and summary statistics. (Detailed numbers are in Appendix Table A2.) Among all the museums in England, history and humanities is the largest family, with 32.0% of museums, followed by social life (23.8%), industries and trades (20.6%), and arts and media (9.2%); science and nature museums are rare, only 2.8% of museums and five of the 167 openings. The remaining museums are mixed or fall outside these fields.

I pair each family with the subject group it is most likely to promote: science and nature, and industries and trades, with STEM subjects, the former presumably more towards academic ones and the latter to non-academic ones; history and humanities with Humanities subjects; arts and media with subjects in Arts; and social life with Social Sciences. In the results, these theme families are labelled Science, Industry, Hist & Lit, Arts, and Society. For the heterogeneity analysis, I also group the openings

by size, into small museums (122 openings) and medium or larger ones (28), and by governance, according to ownership, into museums run by government, a university or a not-for-profit organisation, including those of the National Trust and English Heritage (113), and privately owned museums (38); openings of unknown size or governance are not assigned to these arms.

Figure 3. Museums in England by Theme Family



Notes: This figure shows the 3,163 museums in England by the paper's seven theme families (headers, with counts and shares) and the theme sub-categories of the Mapping Museums database (cells). Areas are compressed for legibility. Treatments in the subject-choice analysis are defined accordingly. Appendix Table A2 reports the counts.

2.3 Museum-Going and Linking Museums to Students

The two data sources are linked by distance. For each LSOA, I compute the straight-line distance from its population-weighted centroid (Office for National Statistics, 2023) to the location of each museum. The population-weighted centroid lies where most residents of the LSOA live, typically in the main village or town of a rural LSOA, and therefore reduces the measurement error that a geographic centroid would introduce in LSOAs that cover a large area.

I then count the museums operating and record the year of the first museum opening, for every LSOA, calendar year between 2001 and 2019, and distance $r \in \{1, 2, 5, 10\}$ km. The resulting panel covers the 32,570 LSOAs in England that lie

within 10 km of at least one museum, together with 58 Welsh LSOAs near the border; the remaining 274 LSOAs in England lie more than 10 km from any museum. Each student is matched to this panel through their home LSOA and the calendar year of the spring semester of Year 10.

The four distances span the range over which a museum can plausibly be part of a teenager's environment. In the DCMS Taking Part survey of 2008/09, 48.5% of respondents' museum visits of up to three hours' travel involved a journey of at most 30 minutes and 79.0% a journey of at most one hour (Panel A of Appendix Table A13; Department for Culture, Media and Sport, 2010), so that museums within walking distance (1–2 km) and within a short trip (5–10 km) are both relevant.

Around students' homes, museums are common but not dense: the average student has 0.26 museums within 1 km of home, 0.97 within 2 km and 4.96 within 5 km (Panel B of Appendix Table A13), while the median LSOA had no museum within 2 km and three within 5 km in 2016. Openings reach a large share of students: 14.7% of students were within 2 km, and 45.0% within 5 km, of a museum opening at some point between 2001 and 2019. Section 3 uses these opening dates to define treated and control students.

Museum-going is common among teenagers. In every wave of the DCMS Taking Part survey between 2005/06 and 2019/20, between 55% and 64% of children aged 11–15 had visited a museum in the previous twelve months (Department for Culture, Media and Sport, 2010–2024). Visits are occasional: pooling the waves, 63% of child visitors went once or twice a year and fewer than 5% at least once a month. Most child visitors go to museums in their own time: pooling the waves of 2010/11 to 2019/20, three in four child visitors (76%) made at least one museum visit in their spare time, and only 24% visited only with their school. Exposure to museums therefore runs largely through the places where students live, which is why I measure it around their homes (see Appendix C for details).

2.4 Descriptive Statistics

Table 1 reports summary statistics for the sample. Columns (1) and (2) describe all students. Half of them are female, 13.6% are eligible for free school meals, 19.5% have special educational needs, and 82.1% are White. These shares are close to the figures the DfE publishes for state-funded secondary schools over the same period; the FSM share is somewhat lower because eligibility declines with age (Appendix Table A12). Panel B shows the outcomes defined above, and Panel C shows the

Table 1. Summary Statistics

	All students		Ever near a museum opening	
	Mean (1)	SD (2)	≤2 km (3)	≤5 km (4)
A. Student characteristics				
Female (%)	49.1	50.0	49.3	49.2
FSM eligible (%)	13.6	34.3	20.0	17.2
SEN, any (%)	19.5	39.6	21.9	21.1
English first language (%)	88.9	31.4	78.7	82.9
White (%)	82.1	38.3	69.0	73.9
Asian (%)	7.4	26.2	13.6	11.4
Black (%)	4.1	19.8	8.5	6.8
Mixed (%)	3.1	17.3	4.5	4.0
Chinese (%)	0.4	6.1	0.5	0.5
Other ethnic group (%)	1.0	10.2	2.1	1.6
Ethnicity unclassified (%)	1.9	13.7	1.8	1.8
KS2 prior attainment (percentile)	50.0	28.8	47.9	48.6
IDACI (percentile)	49.9	28.9	60.6	57.0
Urban (%)	83.2	37.4	93.9	92.6
B. Outcomes				
Entered Level 3 (%)	54.5	49.8	54.1	53.9
Full Level-3 programme (%)	45.8	49.8	45.3	45.2
Any academic qualification (%)	44.2	49.7	43.5	43.5
Any vocational Level-3 qualification (%)	20.7	40.5	20.8	20.7
KS5 average point score (percentile)	49.9	28.9	49.6	49.5
C. Museums near the Year-10 home				
Operating museums within 1 km	0.26	0.76	0.79	0.36
Operating museums within 2 km	0.97	2.05	3.16	1.47
Operating museums within 5 km	4.96	9.52	12.83	8.39
Any museum within 5 km (%), by theme:				
History & humanities	57.4	49.4	75.8	74.6
Society	53.5	49.9	64.3	63.6
Industry	47.8	50.0	63.8	62.4
Arts	28.9	45.3	53.7	49.5
Science	10.2	30.3	22.5	18.6
Students	8,477,187		1,245,027	3,818,084
Share of all students (%)	100.0		14.7	45.0

Notes: This table reports summary statistics for pupils in the Year-11 cohorts 2002/03–2016/17. Columns (3) and (4) report means for students who were ever within 2 km and 5 km of a museum opening over 2001–2019, whether before or after their decision year. Binary variables are in percent; percentile ranks are within the Year-11 cohort, except the KS5 score, ranked within its outcome-year cohort (Section 2.1). The KS5 average point score is observed for Level-3 entrants (4.62 million students), and KS2 prior attainment and Panel C for the 7.57 million students with KS2 records. Columns (3) and (4) count all students ever near an opening; the regression samples are smaller, as they keep only the 2006–2012 opening stacks with a balanced window of four cohorts before and after the opening (Section 3.1). Panel C counts museums operating within each distance of the Year-10 home location. Appendix Table A12 compares student characteristics with national figures, and Appendix Table A3 compares students with and without a museum opening nearby.

museum environment of the students in the calendar year of the Year 10 spring semester: most students have a history and humanities museum (57.4%) or a social-life museum (53.5%) within 5 km of home, whereas only 10.2% have a science museum around.

Columns (3) and (4) describe the students within 2 km and 5 km of a museum opening in the panel. These students live in more urban and more deprived areas than the average student: 93.9% of those near an opening within 2 km live in an urban LSOA, against 83.2% of all students, and their average IDACI rank is the 61st percentile, against the 50th. They are also more often from minority ethnic backgrounds and less often native English speakers, but their KS2 prior attainment is similar to that of the average student (the 48th against the 50th percentile). As these comparisons pool students before and after the openings, they are descriptive; Section 3 further compares treated and never-treated students within the design (Figure 5).

3 Empirical Strategy

3.1 Student-Centric Stacked DiD Design

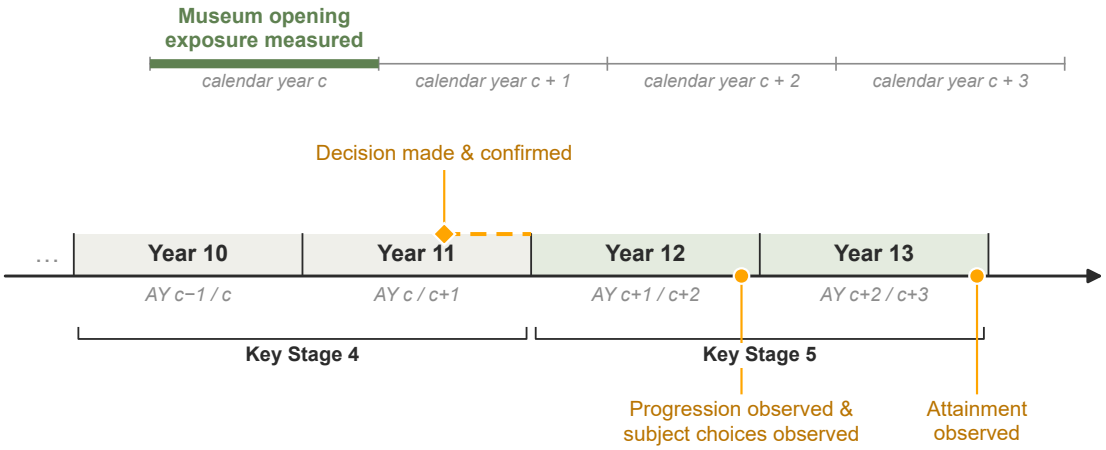
I identify the effects of museum openings on student outcomes using a student-centric stacked difference-in-differences (DiD) design. It compares students whose home neighbourhood gets a new museum within a certain distance to students in neighbourhoods where the first new museum opens within the same distance several years later, before and after the opening.

To explain the design clearly, I begin with the definition of key concepts and the corresponding notation. Let i index students, $j(i)$ the LSOA where student i lives in the Year 10 spring semester, $L(i)$, or more precisely $L(j(i))$, the Local Authority district this LSOA belongs to, and $c(i)$ the calendar year of the spring semester of Year 10, which identifies the student's cohort and the year in which treatment exposure is measured.

Figure 4 then depicts the time dimension of treatment exposure and KS5 outcomes. For a given radius r, g_j^{op} denotes the year in which the first museum within r km of LSOA j opens, with $g_j^{\text{op}} = \infty$ if no museum opens within r km over the analysis period. Whereas $c(i)$ dates the student's cohort, g_j^{op} dates the opening in the student's area. A student counts as treated if this first opening takes place no later than the calendar year of the Year 10 spring semester, that is, if $c(i) \geq g_{j(i)}^{\text{op}}$. I use the radii of $r \in \{1, 2, 5, 10\}$ km, with the treatment at each band referring to the first

opening within that distance.

Figure 4. Timing of Museum Exposure, Decisions and Outcomes



Notes: This figure visualises the timeline of treatment exposure and KS5 educational outcomes. A student is treated in year c if a museum within r km of the recorded home LSOA has opened by c. Level-3 decisions are made in the spring of Year 11 (year c+1) and confirmed in September.

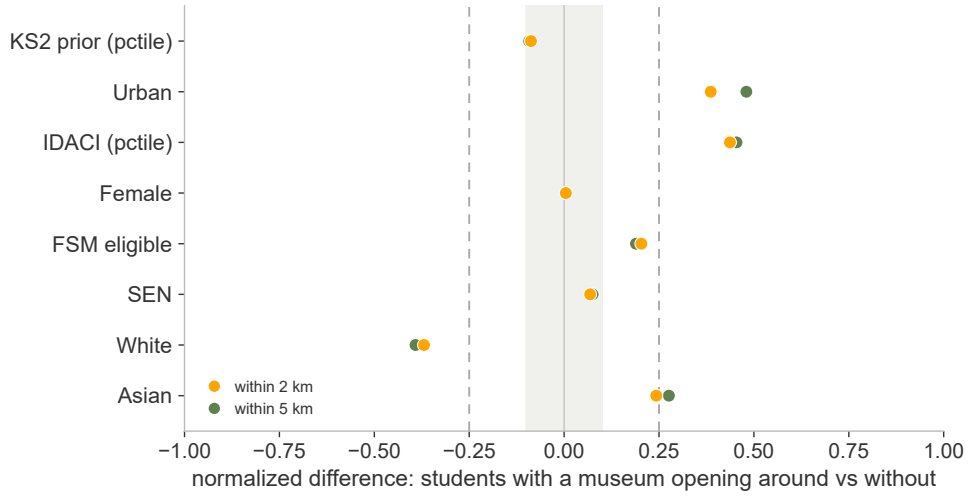
The design is implemented as follows. First, for every opening year (or treatment cohort) g, I build a separate dataset, or a "stack". Every stack consists of two groups of students: (1) the treated group, those whose home LSOA has its first museum opening within r km in year g ($g^{\text{op}}_{j(i)} = g$), and (2) the control group, those whose home LSOA has its first opening within r km only after year $g+4$ ($g^{\text{op}}_{j(i)} > g+4$). In the stacked DiD, students in each treatment cohort are compared to their own control group within the stack, hence avoiding "forbidden comparisons" and negative weighting in staggered DiD designs with a two-way fixed effects model (Goodman-Bacon, 2021; Wing et al., 2024; Blanco and Neri, 2025).

In particular, in the language of DiD designs, the control group includes stack-wise not-yet-treated students who are treated more than four years after the opening year of this stack, rather than students in areas that are never treated. This keeps the comparison between students in places with similar characteristics, since students near an opening live in more urban, poorer, and more diverse areas than those not, even though they do not systematically differ in personal characteristics such as prior attainment, gender, and having special needs, as shown in Figure 5.

Next, I construct the event-year variable, $\tau = c(i) - g$, to express year $c(i)$ relative to the opening year of the stack. I include only the stacks with treated students

for which LEO covers a balanced window of records of four years before and after the event, $\tau \in \{-4, \ldots, 4\}$, to avoid biasing the pre-trend estimates and post-event effects (Wing et al., 2024). Besides, the attainment used as an outcome variable is typically obtained after two years of Key Stage 5; due to COVID-19, the exams were cancelled in academic years 2019/2020 and 2020/2021, and grades were centre-assessed and teacher-assessed. Therefore, I drop these two cohorts ($g = 2017$ and $g = 2018$) because the outcomes are not comparable to previous ones. Together, these restrictions result in the use of museum openings between 2006 and 2012. Figure A1 shows included (green) and dropped (yellow) treatment cohorts, as well as the data missing (dashed cells) for the dropped ones. Last, I append all stacks, with each of the seven stacks covering its full event window, to obtain the analytic sample for the stacked DiD.

Figure 5. Normalized Differences in Characteristics of Students with and without a Museum Opening Nearby



Notes: This figure shows normalized differences in characteristics between students with a museum opening within 2 km (orange) or 5 km (green) of their Year-10 home at any point in the panel and students without one (N = 8.48 million). The normalized difference is $(m_1 - m_0)/\sqrt{(s_0^2 + s_1^2)/2}$ (Imbens and Rubin, 2015). The shaded area marks |0.10|, below which imbalance is conventionally negligible (Austin, 2009), and the dashed lines mark the Imbens–Rubin threshold of |0.25| (Imbens and Wooldridge, 2009); this rule of thumb was stated for differences scaled by $\sqrt{s_0^2 + s_1^2}$, so it is stricter on the scale used here. Appendix Table A3 reports the group means.

The key identification assumption for this design is a Conditional Parallel Trends Assumption that, in the absence of a museum opening, the average outcomes of students in treated and control LSOAs, conditional on the control variables and fixed effects specified below, would have evolved in parallel. Under this assumption, β_τ

for $\tau \geq 0$ identifies the average effect of a first museum opening within r km on the outcomes of treated students τ years after the opening, while the pre-trend coefficients for $\tau \leq -2$ can be interpreted as placebo tests on the cohorts before the opening.

3.2 Estimation and Weighting Scheme

I estimate the following equation of the event study at the level of student i in stack g:

$$Y_{ig} = \sum_{\substack{\tau=-4 \\ \tau \neq -1}}^{4} \beta_\tau \, D_{ig} \times \mathbf{1}\{c(i) - g = \tau\} + \delta D_{ig} + \alpha_{g,L(i)} + \lambda_{g,c(i)} + \mathbf{X}_i'\gamma + \varepsilon_{ig}. \quad (1)$$

β_τ is the effect of interest, capturing the impacts on the outcomes of treated students over event time, measured against the last cohort before the opening ($\tau = -1$). Y_{ig} is an upper-secondary outcome of student i: progression to Level 3, the KS5 average point score or a subject choice (Section 2). The treatment indicator D_{ig} equals one if student i belongs to the treated group of stack g and zero if they belong to its control group. ε_{ig} is the error term; standard errors are clustered at the LSOA level, the level at which treatment is assigned.

In terms of controls, the stack-by-Local-Authority fixed effects ($\alpha_{g,L(i)}$) account for time-invariant differences across the Local Authorities where students live, such as spending on culture and education, socio-economic structure and the school system. $\lambda_{g,c(i)}$ denotes stack-by-calendar-year fixed effects, or equivalently, stack-by-cohort fixed effects, since each student is observed once, and they control for time patterns common to all students of a cohort in a stack, such as national educational reforms. Further, $\mathbf{X}_i$ is a vector of covariates on student characteristics: gender, FSM eligibility, SEN, ethnicity, EAL, KS2 prior attainment (within-cohort percentile), the IDACI rank of the home LSOA (within-cohort percentile), and whether the home LSOA is urban. The identifying variation thus comes from comparing, within the same stack and Local Authority, the change in outcomes of students with similar characteristics in treated LSOAs with that of students in control LSOAs. I report results from three specifications: without covariates, with covariates in linear terms (the main specification), and with stack-interacted covariates.

I also report a pooled version of Eq. (1) that collapses the post-opening event

years into a single period:

$$Y_{ig} = \beta\, D_{ig} \times \text{Post}_{ig} + \delta D_{ig} + \alpha_{g,L(i)} + \lambda_{g,c(i)} + \mathbf{X}_i'\gamma + \varepsilon_{ig} \tag{2}$$

where $\text{Post}_{ig} = \mathbf{1}\{c(i) \geq g\}$ indicates the cohorts who are in the spring of Year 10 in or after the opening year ($\tau \geq 0$). The pooled ATT, β, compares the change in outcomes of treated students between the cohorts before and after the opening with that of control students. Appendix Table A5 reports the pooled ATT.

Since every fixed effect is specific to a stack and the cohort composition of the stacks is balanced, the stacked regression is a weighted average of the stack-by-stack comparisons: the pooled coefficient equals $\beta = \sum_g \omega_g \beta_g$, where β_g is the DiD estimate from stack g alone and the stack weights ω_g are non-negative, sum to one and depend only on the design, with $\omega_g \propto N_g^D N_g^C/(N_g^D + N_g^C)$ for N_g^D treated and N_g^C control students in stack g (Appendix D).

These are the inverse-variance stack weights, so the regression is the most precise combination of the stack-level estimates, and it is a convex average of causal effects whenever each stack identifies one. The concern raised by Wing et al. (2024), that an OLS stacked regression averages treated and control trends with different weights across stacks, applies to specifications without stack-specific time effects; here, the stack-by-year fixed effects difference out the trends within each stack before the stacks are combined. The choice of stack weights therefore matters only if the effect differs across opening years, and the stack weights lie between 7% and 23% across opening years and radii (Appendix Table D1).

As a robustness check, I re-estimate the main specification with the corrective observation weights of Wing et al. (2024).[10] Treated students keep a weight of one, and each control student in stack g receives the weight $(N_g^D/N^D)/(N_g^C/N^C)$, the ratio of the stack's share of all treated students to its share of all control students, where N^D and N^C are the totals across stacks. These observation weights thus rescale the control group of each stack so that every stack has the same treated share, and the pooled coefficient then equals $\sum_g (N_g^D/N^D)\,\beta_g$, the average effect over treated students, in which each opening year counts in proportion to the number of students it treats. Compared with OLS, they shift stack weight towards the later opening years, where the not-yet-treated control pool is smaller, at the cost of some precision. Section 4.1 reports the results.

[10] My code is adapted from Blanco and Neri (2025). I would like to acknowledge Hector Blanco for generously sharing the code and for providing helpful technical explanations.

I use the same design in three further exercises. First, to study which museums matter, I replace D_{ig} with D_{ig}^{k}, which equals one if the first museum opening within r km in year g is of type k (theme, size, or type of governance), and keep the same not-yet-treated control group. Second, to study which students benefit, I estimate the main specification separately on subsamples of students, defined by eligibility for free school meals and by tercile of KS2 attainment within the cohort; each subsample keeps its own fixed effects and its own not-yet-treated control group. Third, as a robustness check, I replace the control group with never-treated students, whose home LSOA has no museum opening within r km over the sample period ($g_{j(i)}^{\text{op}} = \infty$).

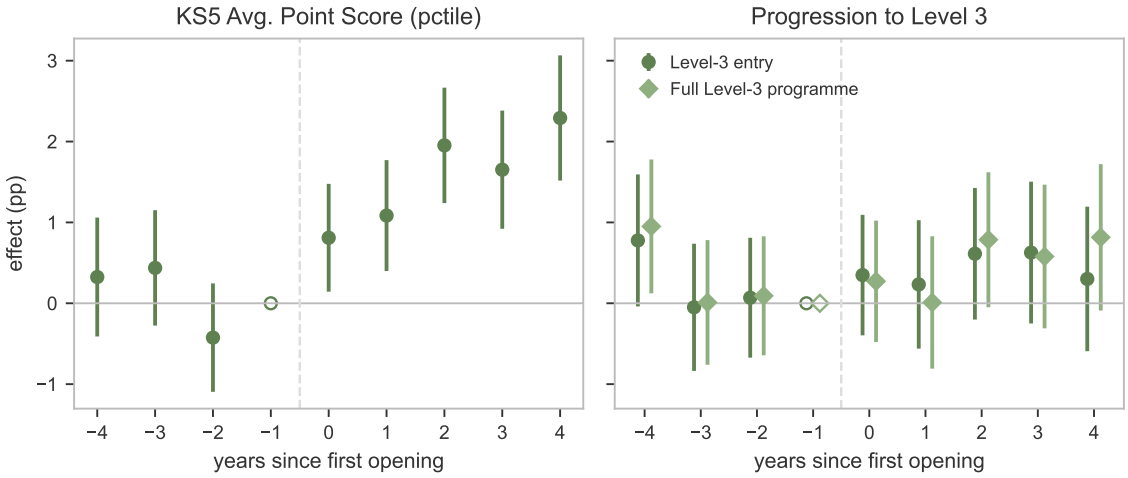
4 Results

A museum opening near students' homes raises their attainment at the end of upper-secondary education. Within 2 km, the KS5 average point score rises by 1.5 percentile points, and the effect fades with distance, while the share of students who progress to Level 3 does not change systematically. These effects do not seem to be explained by changes in who lives around the new museums. They are larger for students from poorer families and with lower prior attainment, and they depend on the kind of museum, with medium and large museums reaching farther than small ones and privately owned museums having slightly larger effects than public and not-for-profit ones. Subject choices, in turn, do not respond to museum openings in general but to what the museum presents, especially in the case of social-life museums, which move students choosing academic qualifications into social sciences.

4.1 Attainment and Progression

Figure 6 presents the main result within 2 km of home for the two outcomes of upper-secondary education, the KS5 average point score and progression to Level 3. (Appendix Table A4 reports the event-study estimates corresponding to this figure.) Before the opening, the attainment of treated and control students evolves in parallel, as pre-trend coefficients lie between -0.42 and 0.44 percentile points, not statistically different from zero. After the opening, attainment rises, with treated students in the first cohort after the opening scoring 0.81 percentile points higher than control students, and the gap widening to around 2 percentile points in later cohorts. The pooled ATT at this band is 1.48 percentile points and fades with the radius as shown later.

Figure 6. Event-Study Estimates within 2 km: KS5 Average Point Score and Progression to Level 3



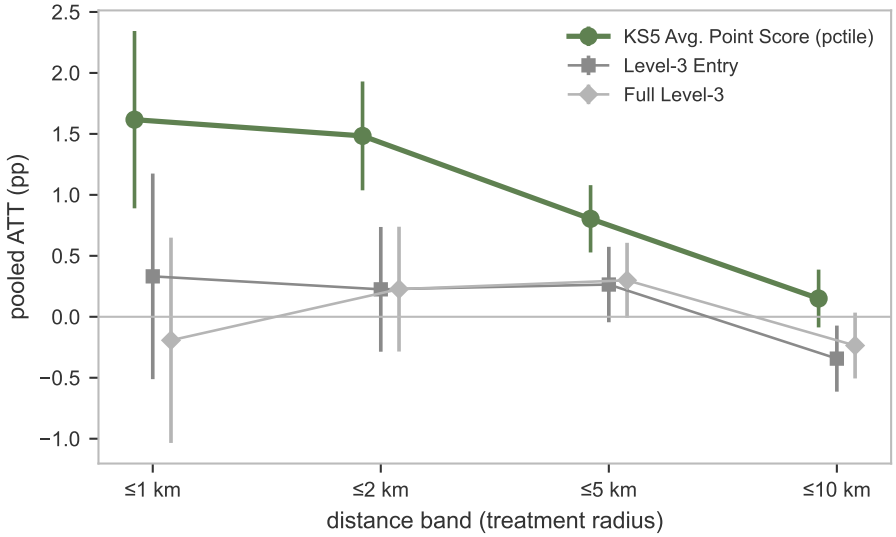
Notes: This figure shows event-study coefficients for the KS5 average point score (percentile points, left panel) and for progression to Level 3 (percentage points, right panel), measured by Level-3 entry (darker) and by the completion of a full Level-3 programme (lighter), around a first museum opening within 2 km of the home LSOA population centroid, from the design of Appendix Table A5. $\tau = -1$ is the omitted year; 95% confidence intervals are based on standard errors clustered by LSOA. Appendix Table A4 reports the coefficients.

Progression to Level 3, in contrast, does not seem to respond to museum openings. The coefficients for Level-3 entry and for completing a full Level-3 programme stay below one percentage point before and after the opening. Pooling the post-opening cohorts, Level-3 entry changes by 0.23 percentage points (s.e. 0.26) on a base of 49.6%, and the share of students completing a full Level-3 programme by 0.23 percentage points (s.e. 0.26) on a base of 41.8%. Therefore, even though museums do not increase progression to upper-secondary education, they improve the attainment of the students who would continue to study at Level 3 anyway.

Figure 7 depicts the spatial gradient of the effects. The attainment effect is about 1.5 percentile points within 1 km and within 2 km of home, and falls to 0.80 within 5 km and to 0.15 within 10 km (Appendix Table A5). Since each band contains the students of the nearer bands, the decline also implies that the effect within 5 km is smaller because of the smaller effects on students between 2 and 5 km from an opening. Progression effects are mostly flat. The effects on Level-3 entry and on completing a full Level-3 programme are all below 0.4 percentage points in absolute value.

The decline of the attainment effect with distance fits with how people travel to

Figure 7. Effects of a Museum Opening on Attainment and Level-3 Progression, by Distance Band



Notes: This figure shows pooled ATTs of a museum opening within each distance band of the Year-10 home LSOA, from student-centric stacked difference-in-differences with stack-wise not-yet-treated controls and student covariates in linear terms (the main specification). Effects are in percentile points for the KS5 average point score (within-cohort percentile among Level-3 entrants) and in percentage points for Level-3 entry and full Level-3 programme (full Year-11 cohort). 95% confidence intervals are based on standard errors clustered by LSOA. Appendix Table A5 reports the estimates and specification details.

museums. In the Taking Part survey, 48.5% of museum visits involve a journey of at most 30 minutes and 79.0% a journey of at most an hour, as reported in Section 2.3, and a museum within 2 km of home is within walking distance. This is also the reason for using 2 km as the main radius for attainment and progression.

To interpret the estimates, we first need to understand the magnitude of these effects. An effect of around 1.48 percentile points means raising the treated student's percentile ranking within the KS5 cohort by 1.48 in the upper-secondary exam. To compare with effects in other studies, this corresponds to about 0.05 standard deviations of the KS5 score[11], comparable to the effects of other local amenities on education. For instance, investment in public libraries in the US raises reading scores of third- to eighth-graders by 0.01 to 0.04 standard deviations (Gilpin et al., 2024), and the regeneration of public housing into mixed-income neighbourhoods in Lon-

[11]The standard deviation of a percentile rank is $100/\sqrt{12} \approx 28.9$ percentile points. Alternatively, if the underlying score is normally distributed and a museum shifts it by δ standard deviations for every treated student, the average percentile rank rises by $100\,\delta/(2\sqrt{\pi}) \approx 28.2\,\delta$ points, which gives $\delta \approx 0.052$, a result similar to the previous conversion.

don raises the maths scores of children in nearby primary schools by 0.064 standard deviations (Neri, 2024).

Second, as attainment is observed only for students who continue with upper-secondary education (Level 3), in principle, the effect could reflect a change in who enters rather than in how well students do. This concern is mitigated by two facts. As shown in Figure 7, first, progression in general is not affected by museum openings, especially within 1 and 2 km, where the attainment effect is largest. Second, within 5 km, the additional entrants are likely to come from the lower part of the attainment distribution, so a change in the composition of entrants would, if anything, result in estimating a lower bound of the attainment effect.

Last, the timing of the effects is also informative. In the design, the cohort at $\tau = -1$ spends most of KS5 after the opening, and the cohorts at $\tau = -2$ and $\tau = -3$ are already in KS5 when the museum opens. The fact that their attainment does not differ from that of earlier cohorts suggests that exposure during KS5 alone does not raise scores. The effect appears in the cohorts who are exposed before they make the decisions, and it grows with the number of years of exposure before the decision, as evidenced by the increasing effects over event time. An alternative reading is that new museums develop their programmes over their first years, which the data do not allow me to separate.

Finally, the effects are robust to how the regression is specified and weighted. The event-study estimates without covariates, with covariates, and with stack-interacted covariates are nearly identical at every distance (Appendix Figure A2). Giving every treated student the same weight, with the corrective observation weights of Wing et al. (2024) described in Section 3, gives a pooled ATT of 1.60 percentile points (s.e. 0.23) within 2 km, against 1.48 under OLS in the main specification, and event-study coefficients of the same shape (Appendix Figure A3). Using never-treated students as the control group gives effects in the same direction but about half as large. Within 2 km, the post-opening coefficients range from 0.34 to 1.02 percentile points, against 0.81 to 2.29 in the main specification (Appendix Table A15). Never-treated areas, which have no opening within the band over the whole period, are less urban and less deprived than the areas around openings (Figure 5), which is why I use them only as a robustness check.

4.2 Alternative Explanation of the Effects

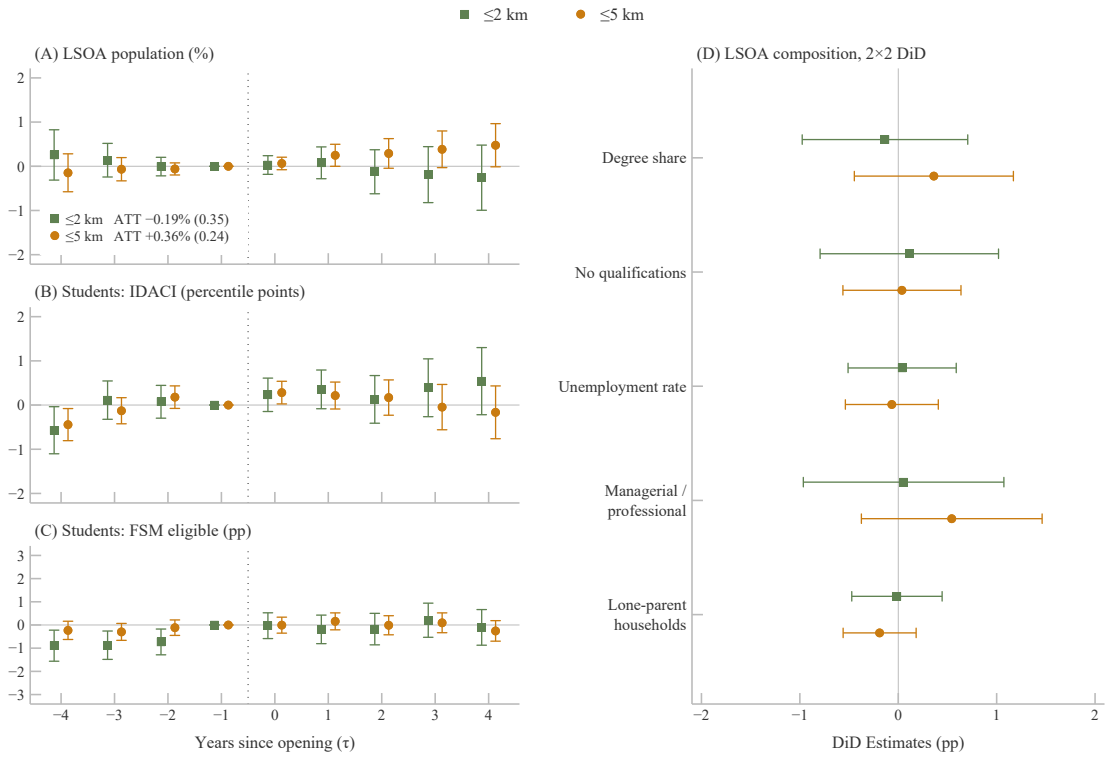
The main threat to the interpretation above is that museums might open in places that are changing in other ways, so that the students living in treated areas after an opening differ from those living there before. For instance, a new museum may be part of a wider regeneration that attracts more educated families, or it may raise house prices and push poorer families out. Therefore, what the attainment effect reflects could be a change in the composition of student families or neighbours rather than the impact of museums. Figure 8 examines this threat in four ways. Additionally, Appendix B sketches typical stories of a museum opening, which also help mitigate this threat.

First, I estimate the same stacked DiD at the level of LSOAs, with the log of the LSOA population as the outcome (panel A). The population of treated areas does not change relative to control areas: the pooled effect is −0.19% (s.e. 0.35) within 2 km and 0.36% (s.e. 0.24) within 5 km, and the pre-opening coefficients are jointly insignificant. (See Appendix Table A6 for detailed results.) Within 5 km, population grows slightly in the later post-opening years, but only by less than half a percent.

Second, I estimate a 2×2 DiD on the LSOA characteristics, using the 2001 and 2011 Censuses in LSOAs with a first opening in 2002–2010 and in LSOAs whose first opening came after 2013 (panel D). The shares of residents with a degree and with no qualifications, the unemployment rate, the share in managerial or professional occupations and the share of lone-parent households either all changed in the same way in the two groups, or there was no change. None of the effects is statistically significant (Appendix Table A7).

Third and fourth, I estimate the student-level design on two family background outcomes: the IDACI rank of the home LSOA and eligibility for free school meals (panels B and C). Neither changes at the opening. Yet within 2 km, pre-trend coefficients for FSM eligibility are negative and significant, and close to zero afterwards, which suggests that the share of FSM-eligible students in treated areas rose by about 0.8 percentage points between the cohorts at $\tau = -2$ and $\tau = -1$ and remained at that level. Three observations suggest that this shift does not drive the attainment effect. First, it goes in the opposite direction. Treated areas became slightly poorer, which, if anything, would be more likely to lower attainment. Second, controlling for student characteristics, including FSM eligibility and IDACI, raises the attainment effect slightly, from 1.40 to 1.48 percentile points within 2 km (Appendix Table A14). Third, the shift takes place one year before the opening and is not

Figure 8. Composition Checks: Area Population and Student Composition around Museum Openings



Notes: This figure shows composition checks for openings within 2 km (green squares) and 5 km (orange circles). Panel (A) reports event-study coefficients for log LSOA population (×100) from the stacked design estimated at the area level, with the pooled ATT in the legend. Panels (B) and (C) use students' IDACI percentile and FSM eligibility as outcomes of the student-level design, controlling only for gender and urban residence. Panel (D) reports two-by-two difference-in-differences estimates of 2001–2011 changes in LSOA characteristics, comparing LSOAs with a first opening in 2002–2010 with LSOAs whose first opening came after 2013. Event coefficients are relative to $\tau = -1$. 95% confidence intervals are clustered by LSOA in (A)–(C) and by Local Authority District in (D). Appendix Tables A6 and A7 report the estimates.

Source: ONS LSOA mid-year population estimates (panel A); Censuses 2001 and 2011 (panel D).

followed by any change after it. Within 5 km, the FSM coefficients are small and insignificant throughout.

4.3 Which Students Benefit

Students also differ in how much a new museum adds to their environment. Panel (a) of Figure 9 shows the attainment effect within 2 km separately by tercile of KS2 attainment within the cohort, by eligibility for free school meals in Year 11, and by the number of museums already operating near home before the opening (Appendix Table A8). The effect declines with prior attainment: students in the bottom tercile gain 2.12 percentile points, those in the middle tercile 1.57 and those in the top tercile 1.14.

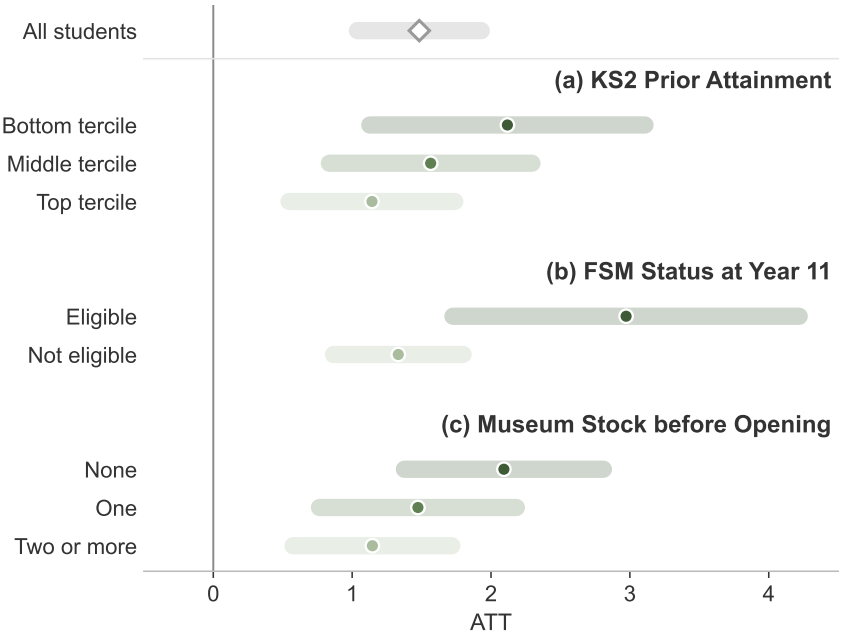
Progression, in contrast, rises only in the middle tercile, where Level-3 entry increases by 1.14 percentage points (column (2) of Appendix Table A8). This is where the margin of progression lies, since most students with high prior attainment continue anyway and most with low prior attainment do not. The additional entrants are likely to be weaker students, which, if anything, pulls the attainment effect of the middle tercile down.

Students eligible for free school meals gain more than other students (Panel (b) of Figure 9). Their attainment rises by 2.97 percentile points, against 1.33 for students who are not eligible, a difference of 1.64 percentile points. This pattern is consistent with museums lowering information costs. Students from poorer families have fewer sources of such knowledge, so a museum nearby may lower these costs more for them than for richer students who already encounter it in other ways.

Panel (c) of Figure 9 splits treated students by the number of museums operating within 2 km of their home in the year before the opening. Where the new museum is the first within 2 km, attainment rises by 2.09 percentile points, compared to 1.47 where one museum already existed and 1.15 where there were two or more. An additional museum adds less in areas where students already encounter more such sources of information and knowledge, which is also consistent with museums lowering information costs.[12]

[12]This pattern is not driven by deprivation. Areas that received their first museum within 2 km are, if anything, less deprived than areas that already had museums: the average IDACI score of the treated LSOAs is 0.18, against 0.24 where one museum already existed and 0.26 where there were two or more, computed from the English Indices of Deprivation 2010 with the pre-opening stock rebuilt from the Mapping Museums data. Since poorer students gain more from a museum, differences in deprivation alone would predict larger effects where museums already existed, the opposite of what panel (c) shows.

Figure 9. Effects on the KS5 Average Point Score within 2 km, by Student Group



Notes: This figure shows pooled ATTs of a first museum opening within 2 km on the KS5 average point score, estimated separately on groups of students from the main specification of Appendix Table A5. The top row is the full-sample estimate. Panel (a) splits students by tercile of KS2 prior attainment within the cohort, panel (b) by eligibility for free school meals in Year 11, and panel (c) by the number of museums operating within 2 km of the home LSOA in the year before the opening. Bands are 95% confidence intervals based on standard errors clustered by LSOA. Appendix Table A8 reports the estimates, together with the effects on progression to Level 3.

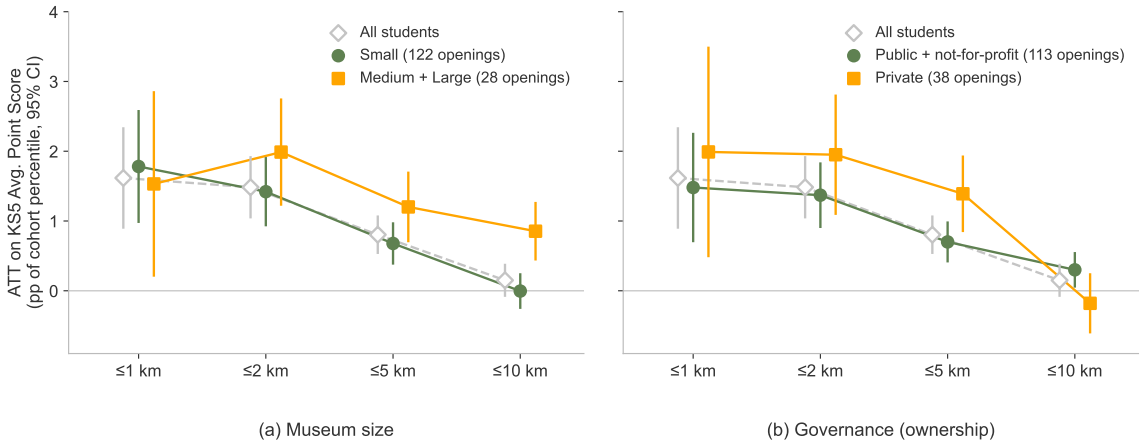
4.4 Which Museums Matter

Museums differ widely in size and in how they are run, and these differences may shape how far and how strongly they affect students. Figure 10 estimates the attainment effect separately by the size and the governance of the first museum opening in each band. Small museums, which make up 122 of the 150 openings of known size, raise attainment by 1.78 and 1.42 percentile points within 1 and 2 km, but their effect falls to 0.68 within 5 km and disappears within 10 km. Medium and large museums, in contrast, keep their effect at a distance, with 1.99 percentile points within 2 km, 1.20 within 5 km and 0.85 within 10 km (Appendix Table A9). This is what one would expect if larger museums attract visitors from a wider area and act as destinations rather than as neighbourhood amenities.

Governance also matters. Openings of museums run by government, a university or a not-for-profit organisation, which make up 113 of the 151 openings of known

governance, have effects almost identical to those of all openings, 1.37 percentile points within 2 km and 0.70 within 5 km. Openings of privately owned museums have slightly larger effects at every distance up to 5 km, 1.95 percentile points within 2 km and 1.39 within 5 km, where the difference between the two groups is 0.69 percentile points (s.e. 0.32), and no effect within 10 km. These differences should be read with caution. There are only 38 private openings, and governance is correlated with other features of a museum, such as its theme, its size and the activities it offers, so why governance matters remains an open question.

Figure 10. Effects on the KS5 Average Point Score, by Museum Size and Governance



(a) Museum size (b) Governance (ownership)

Notes: This figure shows pooled ATTs of a museum opening on attainment, by characteristics of the museum opening in each band (main specification, with student covariates in linear terms). Panel (a) splits openings into small (122) and medium or large museums (28). Panel (b) splits openings by ownership into museums run by government, a university or a not-for-profit organisation, including the National Trust and English Heritage (113), and privately owned museums (38). Open diamonds are the full-sample estimates. 95% confidence intervals are based on standard errors clustered by LSOA. Appendix Table A9 reports the estimates.

4.5 Museum Themes and Subject Choices

Museums differ from most other local amenities in that they present different branches of knowledge. Parks, libraries and theatres offer places to spend time or material on any subject, whereas a museum is organised around a theme, from science and nature to industries and trades, history and the humanities, arts and media, and social life (Section 2.2).

Many of these themes correspond closely to the subjects that students can choose in upper-secondary education. A science museum presents the content of STEM subjects, a history museum exhibits that of history, and a museum of social life and

social history shows how people lived and worked, which is the material of social science subjects. Museums also present these fields differently from school, through original objects, places and people.

A museum nearby may therefore not only motivate students and affect how well students do, as Section 4.1 shows, but also affect what they choose to study. When students face the choice upon progression to the next level of education, they often have little information about what studying a field involves, and exposure to a field has been shown to change the fields that students choose (Wiswall and Zafar, 2015; Fricke et al., 2018; Breda et al., 2023). In this section, I ask whether museums move students' subject choices towards the subjects that their themes present.

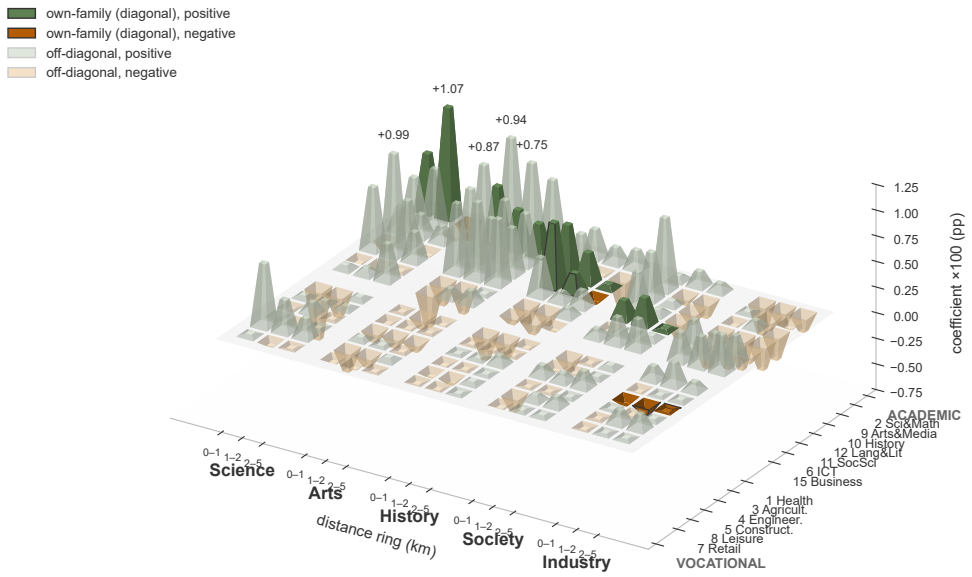
4.5.1 Descriptive Evidence

Before turning to the openings, I first examine the correlation between the theme of museums around a student's home and the subjects the student goes on to choose, using the stock of over 2,000 museums. Therefore, the treatment variation is richer and allows for finer distinctions, compared to the causal results using openings reported later. Figure 11 depicts the coefficients (height being magnitude, colour representing the sign) from regressions of Level-3 entry in each of the 15 Sector Subject Areas (SSAs) on indicators for a museum of each theme (Science, Arts, History, Society, and Industry) within 0–1, 1–2, and 2–5 km of home, with Local Authority × cohort fixed effects and student controls.

Two descriptive patterns stand out. First, the associations are mostly positive across the academic subject areas, whereas the associations with vocational subject areas are close to zero. Second, among academic subjects, the own-theme pairs are positive in the inner rings: science museums with science and mathematics (0.57 and 1.07 percentage points within 0–1 and 1–2 km), arts museums with arts (0.57 and 0.39), history and humanities museums with history (0.51 and 0.29) and languages (0.62 and 0.18), and social-life museums with the social sciences (0.18 and 0.26). Detailed estimates are reported in Appendix Table A10.

Some of these own-theme (or "diagonal") associations stand out against those of other themes. Science museums have the largest association with science and mathematics within 2 km, and history and humanities museums the second largest with history and with languages within 1 km. Arts museums, on the other hand, are positively associated with most academic subjects and the estimates are similar across all academic SSAs. An explanation could be that students living around arts

Figure 11. Museum Stock by Theme and Level-3 Subject Choice: Cross-Sectional Associations



Notes: This figure shows coefficients (percentage points) from regressions of Level-3 entry in each sector subject area on indicators for having a museum of each theme within 0–1, 1–2 and 2–5 km of the home LSOA, with Local Authority × cohort fixed effects and student controls. Academic subject areas are at the back and vocational ones at the front. Darker bars are own-family (diagonal) pairs; light bars are off-diagonal pairs. The five largest coefficients are labelled. Appendix Table A10 reports the coefficients.

museums, in areas rich in cultural capital, are more likely to study academic subjects. The cross-sectional version of the question is shot through with sorting: a family living next to the British Museum today has, in a meaningful sense, already chosen to live there, so a raw correlation between proximity and subject choice confounds exposure with that prior choice. The pattern is therefore suggestive rather than causal, and it is this pattern that motivates the design on openings.

4.5.2 Openings and Subject Choices

The descriptive evidence shapes the analysis of openings in two ways. First, I pay more attention to academic entries, where the associations concentrate, and report any entry, which also counts vocational students, alongside. As students can combine academic and vocational qualifications, I do not classify each subject entry by its type. Instead, I combine entry into a subject group with the information on holding academic qualifications, capturing the choices by students "on the academic route". Having this as an outcome variable also allows me to separate the two margins:

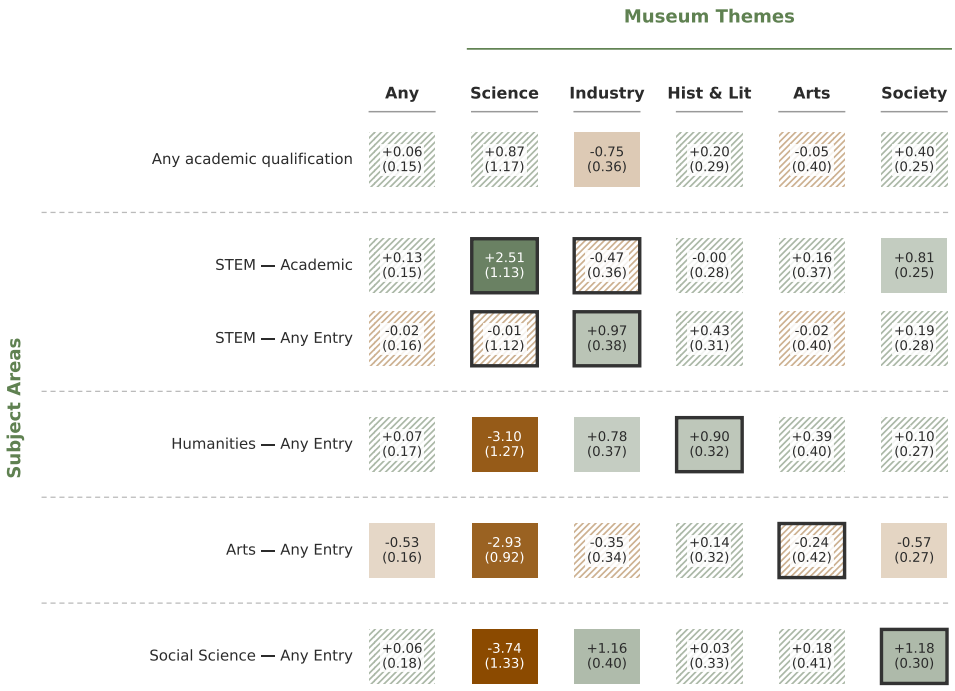
academic entry into a subject group can rise either because more students hold an academic qualification, or because students who hold one choose the subject group more often. The first channel can add at most as much as the effect on holding an academic qualification, so any effect on academic entry beyond that reflects a change in subjects.

Second, openings of a given theme are far rarer than museums of that theme in the stock, so I estimate the effects by theme family and subject group rather than by theme and SSA, and use 5 km as the main radius for this analysis, since the 5 km band contains about three times as many treated students as the 2 km band. Figure 12 reports the pooled ATTs within 5 km, with the theme families in columns and the subject outcomes in rows; the boxed cells are the ex ante pairs of Section 2.2.

The first column of the figure shows that museum openings in general do not change what students choose to study. Within 5 km of an opening, the effects on holding any academic qualification and on any academic entry into STEM, the humanities, and the social sciences are all close to zero and statistically insignificant. The only exception is the arts, where entry falls by about half a percentage point, by 0.53 for any entry (s.e. 0.16) and by 0.45 for academic entry (s.e. 0.13). This aggregate null, however, masks how students respond to what a museum presents, since museums of different themes may move students in different directions.

Turning to the ex ante pairs, the boxed cells of Figure 12 show that subject choices respond to museum themes mostly in an expected way. Science museums raise academic STEM entry by 2.51 percentage points (s.e. 1.13), about 9% of its pre-opening mean, and reduce any entry into the humanities, the arts and the social sciences by about 3 percentage points. Since a student can enter more than one subject group, a rise in one group does not mechanically lower the others, and for the other themes the falls in other groups are smaller, by around half a percentage point, and mostly confined to the arts; science museums are the exception, where the share holding any academic qualification does not change (0.87, s.e. 1.17) while STEM gains and every other group loses. Science museums therefore reallocate students from other subjects towards academic STEM rather than add to the academic route. It is worth noting that, with only five science openings, whether this is a generalisable pattern requires further investigation with more science museum openings available in the future or in other contexts. Industry museums raise any entry into STEM by 0.97 percentage points (s.e. 0.38) but not academic STEM entry (-0.47, s.e. 0.36), and they reduce the share of students with any academic qualification by 0.75 percentage

Figure 12. Effects of Museum Openings within 5 km on Level-3 Subject Choices, by Museum Theme



Notes: This figure shows pooled ATTs, in percentage points of the Year-11 cohort, of the first opening of a museum of each theme (columns) within 5 km of the Year-10 home on Level-3 subject outcomes (rows), from the stacked design used for attainment with stack-wise not-yet-treated controls of any theme. Cells report the estimate and its standard error in parentheses; solid cells are significant at the 5% level, hatched cells are not, and boxed cells are the ex ante hypothesis pairs. "Any Entry" is any Level-3 entry in the subject group, "Academic" is such an entry together with an academic qualification, and "Any academic qualification" is holding at least one A-level-type qualification. For STEM, I also report the "Academic" row, as it shows the hypothesised effect of science and nature museums. Appendix Table A11 reports the estimates, including academic entry into the other subject groups, and Appendix Table A16 those within 2 km. Appendix Figure A4 shows the event-study estimates for the ex ante pairs.

points (s.e. 0.36). This pattern can be read as industry museums presenting applied rather than academic knowledge.

History and humanities museums raise entry into the humanities by 0.90 percentage points within 5 km (s.e. 0.32), while their effect on academic entry is smaller and imprecise (0.36, s.e. 0.26), although it is larger within 2 km (1.24, s.e. 0.40). Arts museums do not change entry into the arts, either any entry (−0.24, s.e. 0.42) or academic entry (−0.33, s.e. 0.35). One reason may be that students already encounter the arts in many ways, so an arts museum adds less information to what students know about arts subjects than, say, a social-life museum adds about the social sciences. This result also resonates with the descriptive evidence where the associations between arts museums and all academic subjects are similar, further pointing to a story of a place-based correlation.

Finally, the strongest evidence comes from museums of social life and social history. An opening of this type of museum within 5 km raises entry into the social sciences by 1.18 percentage points (s.e. 0.30) for any entry and by 1.24 (s.e. 0.21) for academic entry, with a pre-opening mean of 13.9% for the latter. They are as large within 2 km (1.54 and 1.66; Appendix Table A16), and their event studies show a clear rise after the opening (Appendix Figure A4). Social-life museums also raise academic STEM entry by 0.81 percentage points, which suggests a broader shift towards academic study around these openings. Appendix Figure A4 reports the event studies for all ex ante pairs.

5 Conclusion

Where young people grow up shapes how they do at school, but moving to a better neighbourhood changes the whole bundle of a place at once and can disrupt children's lives (Chetty et al., 2016). A museum opening, in contrast, changes one feature of the place where students already live. A new museum within 2 km of home raises students' percentile rank in upper-secondary exams by 1.5 points, about 0.05 standard deviations, and the effect appears in the cohorts exposed before making their upper-secondary choices and grows over the cohorts after the opening. The gains are largest for students eligible for free school meals and for those with lower prior attainment. Museums also shape what students choose to study: much as children who grow up around inventors are more likely to become inventors (Bell et al., 2019), students living within 5 km of a new museum become more likely to study the subjects its

theme presents, by about 1 to 3 percentage points for most of the theme and subject pairs specified before estimation. Cultural amenities such as museums thus shape how well adolescents perform and what they choose to study, though on average not whether they continue.

For policy, the results suggest that cultural investment should also be judged by its effects on the young people who live nearby, not only by visitor numbers or its contribution to the local economy; the OECD notes that students' educational performance is one of the outcomes most commonly sought from cultural participation (OECD, 2022). Public budgets remain central to this investment, which governments are currently expanding or cutting. In January 2026, the UK government committed £1.5 billion to cultural infrastructure in England, £760 million of it for museums (Department for Culture, Media and Sport, 2026), while governments around the world have recently expanded such support, like Australia, or cut it, like Germany and the US (Prime Minister of Australia, 2023; Euronews, 2024; The White House, 2025). Private and not-for-profit organisations matter in these investments as well. Most of the museums that opened in England over this period were run by trusts or private owners, and almost all OECD countries give tax relief for philanthropic giving, including gifts to cultural institutions (OECD, 2020). The results also indicate what to build and where: medium and large museums raise attainment up to 10 km away, whereas the effect of small museums halves within 5 km, and what a museum presents shapes which subjects nearby students choose.

These findings leave several questions open for future research. First, I do not observe who visits the museums, so the estimates capture the effect of having a museum nearby rather than of a visit. Linking administrative records to data on museum visits, or on the learning programmes that museums run with schools, would allow for estimating the effects through visits, schools' use of new museums, and the wider environment that a museum brings to a neighbourhood. Second, museum characteristics are recorded once rather than over time. Information on what museums offer over time, such as exhibitions, fees, and learning programmes, would further show which features of a museum matter. Third, openings of some themes, such as science and nature, are rare. Settings with more openings of this kind would show whether the findings hold more widely. Finally, and more broadly, it remains open whether these effects persist. Following these students beyond school would show whether the gains in attainment and the shifts in subject choice carry over into tertiary education and the labour market.

References

Abrahamsson, Sara Sofie, Aline Bütikofer, and Krzysztof Karbownik, "Swallow This: Childhood and Adolescent Exposure to Fast Food Restaurants, BMI, and Cognitive Ability," NBER Working Paper 31226, National Bureau of Economic Research 2023.

Altonji, Joseph G., Peter Arcidiacono, and Arnaud Maurel, "The Analysis of Field Choice in College and Graduate School," in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 5, Elsevier, 2016, pp. 305–396.

Aucejo, Esteban M., Claudia Hupkau, and Jenifer Ruiz-Valenzuela, "Where versus What: College Value-Added and Returns to Field of Study in Further Education," *Journal of Human Resources*, 2025, *60* (2), 607–652.

Austin, Peter C., "Balance Diagnostics for Comparing the Distribution of Baseline Covariates between Treatment Groups in Propensity-Score Matched Samples," *Statistics in Medicine*, 2009, *28* (25), 3083–3107.

Ballatore, Andrea and Fiona Candlin, "A Geography of UK Museums," *Transactions of the Institute of British Geographers*, 2023, *48* (1), 213–229.

Been, Vicki, Ingrid Gould Ellen, Michael Gedal, Edward Glaeser, and Brian J. McCabe, "Preserving History or Restricting Development? The Heterogeneous Effects of Historic Districts on Local Housing Markets in New York City," *Journal of Urban Economics*, 2016, *92*, 16–30.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen, "Who Becomes an Inventor in America? The Importance of Exposure to Innovation," *Quarterly Journal of Economics*, 2019, *134* (2), 647–713.

Bibby, Peter and Paul Brindley, "Urban and Rural Area Definitions for Policy Purposes in England and Wales: Methodology (v1.0)," Published on GOV.UK, 28 August 2013. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/239477/RUC11methodologypaperaug_28_Aug.pdf.

Blanco, Hector and Lorenzo Neri, "Knocking It Down and Mixing It Up: The Impact of Public Housing Regenerations," *Review of Economics and Statistics*, 2025. Forthcoming.

Bowen, Daniel H., Jay P. Greene, and Brian Kisida, "Learning to Think Critically: A Visual Art Experiment," *Educational Researcher*, 2014, *43* (1), 37–44.

Breda, Thomas, Julien Grenet, Marion Monnet, and Clémentine Van Effenterre, "How Effective are Female Role Models in Steering Girls towards STEM? Evidence from French High Schools," *Economic Journal*, 2023, *133* (653), 1773–1809.

Breinholt, Asta and Mads Meier Jæger, "How Does Cultural Capital Affect Educational Performance: Signals or Skills?," *British Journal of Sociology*, 2020, *71* (1), 28–46.

Campbell, Stuart, Lindsey Macmillan, Richard Murphy, and Gill Wyness, "Matching in the Dark? Inequalities in Student to Degree Match," *Journal of Labor Economics*, 2022, *40* (4), 807–850.

Canaan, Serena and Pierre Mouganie, "Returns to Education Quality for Low-Skilled Students: Evidence from a Discontinuity," *Journal of Labor Economics*, 2018, *36* (2), 395–436.

Candlin, Fiona and Alexandra Poulovassilis, "Understanding and Managing Patchy Data in the UK Museum Sector," *Museum Management and Curatorship*, 2020, *35* (4), 446–459.

_ and Jamie Larkin, "What Is a Museum? A New Approach," *Museum and Society*, 2020, *18* (2), 115–131.

_ , Andrea Ballatore, Jamie Larkin, Mark Liebenrood, Alexandra Poulovassilis, and Valeri Katerinchuk, "The UK Museum Boom: Continuity and Change 1960–2019," *Cultural Trends*, 2024, *33* (5), 505–524.

_ , Jamie Larkin, Andrea Ballatore, and Alexandra Poulovassilis, "Mapping Museums 1960–2020: A Report on the Data," Report, Birkbeck, University of London, London 2020.

Carlana, Michela, "Implicit Stereotypes: Evidence from Teachers' Gender Bias," *Quarterly Journal of Economics*, 2019, *134* (3), 1163–1224.

Carneiro, Pedro, Sarah Cattan, and Nick Ridpath, "The Short- and Medium-Term Impacts of Sure Start on Educational Outcomes," IFS Report R307, Institute for Fiscal Studies 2024.

Cavaglia, Chiara, Stephen Machin, Sandra McNally, and Jenifer Ruiz-Valenzuela, "Gender, Achievement, and Subject Choice in English Education," *Oxford Review of Economic Policy*, 2020, *36* (4), 816–835.

Chetty, Raj and Nathaniel Hendren, "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects," *Quarterly Journal of Economics*, 2018, *133* (3), 1107–1162.

_ , _ , **and Lawrence F. Katz**, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 2016, *106* (4), 855–902.

Chyn, Eric, "Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children," *American Economic Review*, 2018, *108* (10), 3028–3056.

_ **and Lawrence F. Katz**, "Neighborhoods Matter: Assessing the Evidence for Place Effects," *Journal of Economic Perspectives*, 2021, *35* (4), 197–222.

Dahl, Gordon B., Dan-Olof Rooth, and Anders Stenberg, "High School Majors and Future Earnings," *American Economic Journal: Applied Economics*, 2023, *15* (1), 351–382.

_ , _ , **and** _ , "Intergenerational and Sibling Spillovers in High School Majors," *American Economic Journal: Economic Policy*, 2024, *16* (3), 133–173.

De Philippis, Marta, "STEM Graduates and Secondary School Curriculum: Does Early Exposure to Science Matter?," *Journal of Human Resources*, 2023, *58* (6), 1914–1947.

Delaney, Judith M. and Paul J. Devereux, "Understanding Gender Differences in STEM: Evidence from College Applications," *Economics of Education Review*, 2019, *72*, 219–238.

Department for Culture, Media and Sport, "Taking Part: The National Survey of Culture, Leisure and Sport, 2008–2009; Adult and Child Data," Data collection, UK Data Service 2010. SN 6530, https://doi.org/10.5255/UKDA-SN-6530-1.

_ , "Taking Part: The National Survey of Culture, Leisure and Sport, 2005–2006 to 2019–2020; Adult and Child Data," Data collections, UK Data Service 2010–2024. SN 5717, 6272, 6273, 6530, 6579, 6855, 7222, 7371, 7709, 7872, 8070, 8378, 8442, 8631 and 8745. The depositor is the Department for Digital, Culture, Media and Sport for 2019–2020, and the Department for Culture, Media and Sport with NatCen Social Research for 2016–2017.

_ , "Government Announces Bumper £1.5 Billion Package to Restore National Pride," Press release, 21 January 2026. https://www.gov.uk/government/news/government-announces-bumper-15-billion-package-to-restore-national-pride.

Department for Education, "16 to 18 Accountability Measures: Technical Guide. For Measures in 2018," Department for Education, October 2018. https://dera.ioe.ac.uk/id/eprint/32266/1/16-18_Accountability_Measures_Technical_Guide_Oct18.pdf.

_ , "Participation of Young People in Education, Employment or Training: Statutory Guidance for Local Authorities," Department for Education, April 2024. https://www.gov.uk/government/publications/participation-of-young-people-education-employment-and-training.

_ , HM Revenue and Customs, Department for Work and Pensions, and Higher Education Statistics Agency, "Longitudinal Education Outcomes SRS Iteration 2.1 Standard Extract – England," Dataset, ONS Secure Research Service; data owner: Department for Education 2026. https://doi.org/10.57906/pzfv-d195.

Deutscher, Nathan, "Place, Peers, and the Teenage Years: Long-Run Neighborhood Effects in Australia," *American Economic Journal: Applied Economics*, 2020, *12* (2), 220–249.

Ehrlich, Maximilian v. and Henry G. Overman, "Place-Based Policies and Spatial Disparities across European Cities," *Journal of Economic Perspectives*, 2020, *34* (3), 128–149.

Einiö, Elias and Henry G. Overman, "The Effects of Supporting Local Business: Evidence from the UK," *Regional Science and Urban Economics*, 2020, *83*, 103500.

Erickson, Heidi H., Angela R. Watson, and Jay P. Greene, "An Experimental Evaluation of Culturally Enriching Field Trips," *Journal of Human Resources*, 2024, *59* (3), 879–904.

Euronews, "Berlin Moves Ahead with €130 Million Cut to Culture Budget amid Protests," Euronews Culture, 24 December 2024. https://www.euronews.com/culture/2024/12/24/berlin-moves-ahead-with-130-million-cut-to-culture-budget-amid-protests.

Falck, Oliver, Michael Fritsch, and Stephan Heblich, "The Phantom of the Opera: Cultural Amenities, Human Capital, and Regional Economic Growth," *Labour Economics*, 2011, *18* (6), 755–766.

Fricke, Hans, Jeffrey Grogger, and Andreas Steinmayr, "Exposure to Academic Fields and College Major Choice," *Economics of Education Review*, 2018, *64*, 199–213.

Gibbons, Stephen, Henry G. Overman, and Matti Sarvimäki, "The Local Economic Impacts of Regeneration Projects: Evidence from UK's Single Regeneration Budget," *Journal of Urban Economics*, 2021, *122*, 103315.

_ , Olmo Silva, and Felix Weinhardt, "Everybody Needs Good Neighbours? Evidence from Students' Outcomes in England," *Economic Journal*, 2013, *123* (571), 831–874.

_ , _ , **and** _ , "Neighbourhood Turnover and Teenage Attainment," *Journal of the European Economic Association*, 2017, *15* (4), 746–783.

Gilpin, Gregory, Ezra Karger, and Peter Nencka, "The Returns to Public Library Investment," *American Economic Journal: Economic Policy*, 2024, *16* (2), 78–109.

Gomes, Pedro and Alejandro Librero-Cano, "Evaluating Three Decades of the European Capital of Culture Programme: A Difference-in-Differences Approach," *Journal of Cultural Economics*, 2018, *42* (1), 57–73.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

Greene, Jay P., Brian Kisida, and Daniel H. Bowen, "The Educational Value of Field Trips," *Education Next*, 2014, *14* (1), 78–86.

Hanzl, Lisa and Gregory Gilpin, "Unequal Access: How Public Library Closures Affect Educational Performance," EdWorkingPaper 25-1140, Annenberg Institute at Brown University 2025.

Heckman, James J., John Eric Humphries, and Gregory Veramendi, "Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking," *Journal of Political Economy*, 2018, *126* (S1), S197–S246.

Hupkau, Claudia, Sandra McNally, Jenifer Ruiz-Valenzuela, and Guglielmo Ventura, "Post-Compulsory Education in England: Choices and Implications," *National Institute Economic Review*, 2017, *240*, R42–R57.

Imbens, Guido W. and Donald B. Rubin, *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, Cambridge: Cambridge University Press, 2015.

_ **and Jeffrey M. Wooldridge**, "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 2009, *47* (1), 5–86.

Joensen, Juanna Schrøter and Helena Skyt Nielsen, "Is There a Causal Effect of High School Math on Labor Market Outcomes?," *Journal of Human Resources*, 2009, *44* (1), 171–198.

_ **and** _ , "Spillovers in Education Choice," *Journal of Public Economics*, 2018, *157*, 158–183.

Kim, Jeonghyeok, Rohit Munshi, and Brian Murphy, "Neighborhood Effects on STEM Major Choice," EdWorkingPaper 25-1263, Annenberg Institute at Brown University 2025.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad, "Field of Study, Earnings, and Self-Selection," *Quarterly Journal of Economics*, 2016, *131* (3), 1057–1111.

Kisida, Brian, Jay P. Greene, and Daniel H. Bowen, "Creating Cultural Consumers: The Dynamics of Cultural Capital Acquisition," *Sociology of Education*, 2014, *87* (4), 281–295.

Kline, Patrick and Enrico Moretti, "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs," *Annual Review of Economics*, 2014, *6*, 629–662.

Laliberté, Jean-William, "Long-Term Contextual Effects in Education: Schools and Neighborhoods," *American Economic Journal: Economic Policy*, 2021, *13* (2), 336–377.

Legewie, Joscha and Thomas A. DiPrete, "The High School Environment and the Gender Gap in Science and Engineering," *Sociology of Education*, 2014, *87* (4), 259–280.

Machin, Stephen, Sandra McNally, and Jenifer Ruiz-Valenzuela, "Entry through the Narrow Door: The Costs of Just Failing High Stakes Exams," *Journal of Public Economics*, 2020, *190*, 104224.

_ , _ , Camille Terrier, and Guglielmo Ventura, "Closing the Gap between Vocational and General Education? Evidence from University Technical Colleges in England," *Journal of Human Resources*, 2025. Published online 8 July 2025.

Matthewes, Sönke Hendrik and Guglielmo Ventura, "On Track to Success? Returns to Vocational Education against Different Alternatives," CVER Discussion Paper 038, Centre for Vocational Education Research, London School of Economics 2022.

Mouganie, Pierre and Yaojing Wang, "High-Performing Peers and Female STEM Choices in School," *Journal of Labor Economics*, 2020, *38* (3), 805–841.

Murphy, Richard and Felix Weinhardt, "Top of the Class: The Importance of Ordinal Rank," *Review of Economic Studies*, 2020, *87* (6), 2777–2826.

Neri, Lorenzo, "Moving Opportunities: The Impact of Mixed-Income Public Housing Regenerations on Student Achievement," *Journal of Public Economics*, 2024, *230*, 105053.

Neumark, David and Helen Simpson, "Place-Based Policies," in Gilles Duranton, J. Vernon Henderson, and William C. Strange, eds., *Handbook of Regional and Urban Economics*, Vol. 5, Elsevier, 2015, chapter 18, pp. 1197–1287.

OECD, "Taxation and Philanthropy," OECD Tax Policy Studies 27, OECD Publishing, Paris 2020.

_ , "The Culture Fix: Creative People, Places and Industries," Local Economic and Employment Development (LEED), OECD Publishing, Paris 2022.

Office for National Statistics, "Lower Layer Super Output Areas (December 2011) EW Population Weighted Centroids," ONS Open Geography Portal 2023.

_ , "Number of Museums across Local Authority Districts (LAD) in the United Kingdom," Dataset, released 7 March 2024. https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/datasets/numberofmuseumsacrosslocalauthoritydistrictsladintheunitedkingdom.

Ofqual, "Annual Qualifications Market Report: Academic Year 2016 to 2017," Official statistics Ofqual/18/6350/1 with data tables for England, Ofqual, 2018. Published 22 February 2018, updated 2 March 2018. https://www.gov.uk/government/statistics/annual-qualifications-market-report-academic-year-2016-to-2017.

_ , "Annual Qualifications Market Report: Academic Year 2024 to 2025," Official statistics Ofqual/26/7309/1 with data tables, Ofqual, 2026. Published 19 March 2026. https://www.gov.uk/government/statistics/annual-qualifications-market-report-academic-year-2024-to-2025.

Pamuk, Orhan, *The Museum of Innocence*, London: Faber and Faber, 2009. Translated from the Turkish by Maureen Freely.

Porter, Catherine and Danila Serra, "Gender Differences in the Choice of Major: The Importance of Female Role Models," *American Economic Journal: Applied Economics*, 2020, *12* (3), 226–254.

Poulovassilis, Alexandra, Fiona Candlin, Andrea Ballatore, Valeri Katerinchuk, Jamie Larkin, Nick Larson, and Mark Liebenrood, "Mapping Museums Data, September 2021," Dataset, Birkbeck, University of London 2021. https://doi.org/10.18743/DATA.00171.

_ **, Nick Larsson, Fiona Candlin, Jamie Larkin, and Andrea Ballatore**, "Creating a Knowledge Base to Research the History of UK Museums through Rapid Application Development," *Journal on Computing and Cultural Heritage*, 2019, *12* (4), 1–27.

Prime Minister of Australia, "Protecting Australia's Cultural and Historical Institutions," Media release, 5 April 2023. https://www.pm.gov.au/media/protecting-australias-cultural-and-historical-institutions.

Riise, Julie, Barton Willage, and Alexander Willén, "Can Female Doctors Cure the Gender STEMM Gap? Evidence from Exogenously Assigned General Practitioners," *Review of Economics and Statistics*, 2022, *104* (4), 621–635.

Sanbonmatsu, Lisa, Jeffrey R. Kling, Greg J. Duncan, and Jeanne Brooks-Gunn, "Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment," *Journal of Human Resources*, 2006, *41* (4), 649–691.

Sanchis-Guarner, Rosa, José Montalbán, and Felix Weinhardt, "Home Broadband and Human Capital Formation," *Economics of Education Review*, 2025, *108*, 102679.

Stopforth, Sarah and Vernon Gayle, "Parental Social Class and GCSE Attainment: Re-reading the Role of 'Cultural Capital'," *British Journal of Sociology of Education*, 2022, *43* (5), 680–699.

Sullivan, Alice, "Cultural Capital and Educational Attainment," *Sociology*, 2001, *35* (4), 893–912.

The White House, "Continuing the Reduction of the Federal Bureaucracy," Executive order, 14 March 2025. https://www.whitehouse.gov/presidential-actions/2025/03/continuing-the-reduction-of-the-federal-bureaucracy/.

UNESCO, "Museums around the World in the Face of COVID-19," UNESCO Report, United Nations Educational, Scientific and Cultural Organization, Paris April 2021. Annex 3 estimates 103,842 museums worldwide (about 104,000), based on the Museums of the World directory, IMLS data, and Member State surveys in 2019, 2020 and March 2021.

Villa, Carmen, "The Effects of Youth Clubs on Education and Crime," IFS Working Paper 24/51, Institute for Fiscal Studies 2024.

What Works Centre for Local Economic Growth, "Evidence Review 3: Sports and Culture," Evidence Review 3, What Works Centre for Local Economic Growth 2016. Updated June 2016. https://whatworksgrowth.org/wp-content/uploads/16-06-15_Culture_and_Sport_Updated.pdf.

Wing, Coady, Seth M. Freedman, and Alex Hollingsworth, "Stacked Difference-in-Differences," Working Paper 32054, National Bureau of Economic Research 2024.

Wiswall, Matthew and Basit Zafar, "Determinants of College Major Choice: Identification using an Information Experiment," *Review of Economic Studies*, 2015, *82* (2), 791–824.

A Appendix

A.1 Tables for the Main-Text Figures

Appendix Table A1. Museum Openings by Year, England, 2001–2019

Year	Openings	Treatment event	Year	Openings	Treatment event
2001	35	No	2011	27	Yes
2002	28	No	2012	29	Yes
2003	32	No	2013	27	No
2004	26	No	2014	33	No
2005	31	No	2015	19	No
2006	19	Yes	2016	25	No
2007	25	Yes	2017	20	No
2008	27	Yes	2018	16	No
2009	23	Yes	2019	18	No
2010	17	Yes			

Openings 2001–2019	477
Openings 2006–2012 (treatment events)	167
Museums operating in 2005	2,542

Notes: This table reports museum openings in England by year. The openings of 2006–2012 are the treatment events, for which a full four-year window before and after the opening fits within the data. Figure 2 plots the counts.

Appendix Table A2. Museum Stock and Openings by Theme Family

	Museums (stock)		Openings
Theme family / group	*N*	Share (%)	2006–2012
Science & nature	**88**	**2.8**	**5**
Natural world	58	1.8	2
Science & technology	20	0.6	2
Personality	10	0.3	1
Industries & trades	**652**	**20.6**	**29**
Industry & manufacture	144	4.6	6
Transport	260	8.2	9
Rural industry	113	3.6	4
Sea & seafaring	83	2.6	9
Utilities	35	1.1	0
Communications	17	0.5	1
History & humanities	**1,011**	**32.0**	**37**
Archaeology	85	2.7	3
Personality	61	1.9	2
Arts	5	0.2	0
Belief & identity	78	2.5	8
War & conflict	316	10.0	16
Buildings	466	14.7	8
Social life	**752**	**23.8**	**49**
Services	43	1.4	5
Medicine & health	35	1.1	7
Local histories	612	19.3	32
Rural industry	27	0.9	0
Personality	13	0.4	2
Mixed	13	0.4	2
Industry	9	0.3	1
Arts & media	**291**	**9.2**	**27**
Arts	253	8.0	24
Personality	20	0.6	0
Leisure & sport	18	0.6	3
Other	**205**	**6.5**	**17**
Leisure & sport	86	2.7	10
Other	67	2.1	5
Personality	31	1.0	2
Food & drink	21	0.7	0
Mixed	**164**	**5.2**	**3**
Mixed	164	5.2	3
Total	3,163	100.0	167

Notes: This table reports the museums in England by theme family and by the Mapping Museums theme group within each family. Some groups appear in more than one family because their sub-categories are split (for example, literary figures under History & humanities and art figures under Arts & media). Figure 3 plots the stock.

Appendix Table A3. Characteristics of Students with and without a Museum Opening Nearby

	Within 2 km				Within 5 km			
	With	Without	Diff.	Norm. diff.	With	Without	Diff.	Norm. diff.
KS2 prior (pctile)	47.9	50.4	−2.5	−0.09	48.6	51.2	−2.7	−0.09
Urban	93.9	81.4	12.5	0.39	92.6	75.5	17.1	0.48
IDACI (pctile)	60.6	48.1	12.5	0.44	57.0	44.2	12.8	0.45
Female	49.3	49.0	0.2	0.00	49.2	49.0	0.2	0.00
FSM eligible	20.0	12.5	7.5	0.20	17.2	10.7	6.5	0.19
SEN, any	21.9	19.1	2.8	0.07	21.1	18.2	3.0	0.08
White	69.0	84.3	−15.4	−0.37	73.9	88.8	−15.0	−0.39
Asian	13.6	6.4	7.2	0.24	11.4	4.1	7.3	0.28
Black	8.5	3.3	5.2	0.22	6.8	1.9	5.0	0.24
Mixed	4.5	2.8	1.6	0.09	4.0	2.3	1.7	0.10
English first language	78.7	90.7	−12.0	−0.34	82.9	93.9	−10.9	−0.35
Students	1,245,027	7,232,160			3,818,084	4,659,103		

Notes: This table reports the mean characteristics of students with and without a museum opening within each band of their Year-10 home at any point in the panel. Binary variables are in percent; IDACI and KS2 prior attainment are within-cohort percentile ranks. The normalized difference is $(m_1 - m_0)/\sqrt{(s_0^2 + s_1^2)/2}$ (Imbens and Rubin, 2015). Figure 5 plots the normalized differences.

Appendix Table A4. Event-Study Estimates within 2 km: KS5 Average Point Score and Progression to Level 3

	(1) KS5 average point score	(2) Level-3 entry	(3) Full Level-3 programme
$\tau = -4$	0.32	0.78*	0.95**
	(0.38)	(0.42)	(0.42)
$\tau = -3$	0.44	−0.05	0.01
	(0.36)	(0.40)	(0.39)
$\tau = -2$	−0.42	0.07	0.09
	(0.34)	(0.38)	(0.38)
$\tau = -1$	0	0	0
	(ref.)	(ref.)	(ref.)
$\tau = 0$	0.81**	0.35	0.27
	(0.34)	(0.38)	(0.38)
$\tau = +1$	1.08***	0.23	0.01
	(0.35)	(0.41)	(0.42)
$\tau = +2$	1.95***	0.61	0.79*
	(0.36)	(0.42)	(0.43)
$\tau = +3$	1.65***	0.63	0.58
	(0.37)	(0.45)	(0.45)
$\tau = +4$	2.29***	0.30	0.82*
	(0.39)	(0.46)	(0.46)
Pooled ATT	1.48***	0.23	0.23
	(0.23)	(0.26)	(0.26)
Pre-period mean	50.27	49.60	41.83
Treated N per τ	28,228	28,228	28,228
Control N per τ	123,569	123,569	123,569

Notes: This table reports event-study coefficients of a first museum opening within 2 km, from the main specification of Appendix Table A5, in percentile points for the KS5 average point score and in percentage points for the progression outcomes. $\tau = -1$ is the omitted year, and the pooled ATT is β in Eq. (2). Treated and control N are medians across event years. Standard errors clustered by LSOA are in parentheses. Figure 6 plots the coefficients. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table A5. Effects of a Museum Opening on Attainment and Level-3 Progression, by Distance Band

	(1) ≤1 km	(2) ≤2 km	(3) ≤5 km	(4) ≤10 km
		A. KS5 average point score (percentile)		
Pooled ATT	1.62***	1.48***	0.80***	0.15
	(0.37)	(0.23)	(0.14)	(0.12)
Pre-period mean	50.42	50.27	50.21	50.50
N	282,413	750,488	1,822,229	2,102,098
		B. Level-3 entry		
Pooled ATT	0.33	0.23	0.26*	−0.34**
	(0.43)	(0.26)	(0.16)	(0.14)
Pre-period mean	51.16	49.60	50.31	50.12
N	498,497	1,360,467	3,280,929	3,809,296
		C. Full Level-3 programme		
Pooled ATT	−0.19	0.23	0.30*	−0.24*
	(0.43)	(0.26)	(0.16)	(0.14)
Pre-period mean	43.49	41.83	42.71	42.47
N	498,497	1,360,467	3,280,929	3,809,296

Notes: This table reports pooled ATTs (Eq. (2)) of a first museum opening within the column radius of the Year-10 home LSOA. Each regression uses stack-wise not-yet-treated controls, stack × Local Authority and stack × calendar-year fixed effects, and student covariates in linear terms (the main specification). Effects are in percentile points for the KS5 average point score, observed for Level-3 entrants, and in percentage points of the Year-11 cohort for the progression outcomes. N counts student–stack observations of treated and stack-wise not-yet-treated control students in the 2006–2012 stacks over event years −4 to 4 (Section 3.1), so a control student can appear in several stacks, and N is not comparable to the counts of students ever near an opening in Table 1. Standard errors clustered by LSOA are in parentheses. Figure 7 plots the estimates. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table A6. Composition Checks: Area Population and Student Composition around Museum Openings

	(1)	(2)	(3)	(4)	(5)	(6)
	(A) LSOA population (%, log points)		(B) Students: IDACI (percentile points)		(C) Students: FSM eligible (pp)	
	≤2 km	≤5 km	≤2 km	≤5 km	≤2 km	≤5 km
$\tau = -4$	0.26	−0.15	−0.57**	−0.44**	−0.89***	−0.23
	(0.29)	(0.22)	(0.27)	(0.19)	(0.34)	(0.20)
$\tau = -3$	0.14	−0.07	0.11	−0.13	−0.87***	−0.30
	(0.19)	(0.13)	(0.22)	(0.15)	(0.31)	(0.19)
$\tau = -2$	−0.01	−0.06	0.07	0.18	−0.73***	−0.12
	(0.11)	(0.07)	(0.19)	(0.13)	(0.28)	(0.17)
$\tau = -1$	0	0	0	0	0	0
	(ref.)	(ref.)	(ref.)	(ref.)	(ref.)	(ref.)
$\tau = 0$	0.03	0.06	0.23	0.28**	−0.03	−0.01
	(0.11)	(0.07)	(0.19)	(0.13)	(0.28)	(0.18)
$\tau = +1$	0.08	0.25*	0.35	0.22	−0.19	0.16
	(0.18)	(0.13)	(0.22)	(0.16)	(0.31)	(0.19)
$\tau = +2$	−0.12	0.29*	0.13	0.17	−0.18	−0.01
	(0.25)	(0.17)	(0.28)	(0.20)	(0.35)	(0.21)
$\tau = +3$	−0.19	0.38*	0.39	−0.05	0.21	0.09
	(0.32)	(0.21)	(0.33)	(0.26)	(0.38)	(0.22)
$\tau = +4$	−0.26	0.48*	0.54	−0.16	−0.10	−0.26
	(0.38)	(0.25)	(0.39)	(0.30)	(0.39)	(0.23)
Pooled ATT	−0.19	0.36				
	(0.35)	(0.24)				
Pre-trend test (p)	0.41	0.55				
Treated LSOAs	1,885	4,720				
Control LSOAs	2,074	4,340				
N	87,444	188,118				

Notes: This table reports event-study coefficients relative to $\tau = -1$. Columns (1)–(2) estimate the stacked design at the LSOA level, with log population (×100) as the outcome; the pre-trend test is the joint test of $\tau = -4, \ldots, -2$. Columns (3)–(6) estimate the student-level design with IDACI and FSM eligibility as outcomes, controlling only for gender and urban residence; their pooled ATTs are not available. Standard errors clustered by LSOA are in parentheses. Figure 8 plots the coefficients. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
Source: ONS LSOA mid-year population estimates (columns 1–2).

Appendix Table A7. Composition Checks: Changes in LSOA Characteristics, 2001–2011

	(1) ≤2 km	(2) ≤5 km
Degree share	−0.13	0.36
	(0.43)	(0.41)
No qualifications	0.11	0.04
	(0.46)	(0.31)
Unemployment rate	0.04	−0.07
	(0.28)	(0.24)
Managerial / professional	0.05	0.54
	(0.52)	(0.47)
Lone-parent households	−0.01	−0.19
	(0.23)	(0.19)
Treated / control LSOAs	2,989 / 1,377	8,234 / 2,737
N	4,350	10,961

Notes: This table reports two-by-two difference-in-differences estimates of the 2001–2011 change in LSOA characteristics, in percentage points. Treated LSOAs have a first museum opening within the column radius in 2002–2010; control LSOAs have their first opening after 2013. All regressions include Local Authority District fixed effects. Standard errors clustered by Local Authority District are in parentheses. Panel (D) of Figure 8 plots the estimates. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Source: Censuses 2001 and 2011.

Appendix Table A8. Effects within 2 km, by Student Group

	(1) KS5 average point score	(2) Level-3 entry	(3) Full Level-3 programme
All students	1.48***	0.23	0.23
	(0.23)	(0.26)	(0.26)
N	750,488	1,360,467	1,360,467
A. KS2 prior attainment			
Bottom tercile	2.12***	0.30	−0.11
	(0.51)	(0.42)	(0.39)
N	148,021	467,712	467,712
Middle tercile	1.57***	1.14**	1.03**
	(0.37)	(0.47)	(0.46)
N	238,972	400,698	400,698
Top tercile	1.14***	−0.07	0.00
	(0.30)	(0.36)	(0.40)
N	313,064	382,565	382,565
B. FSM status at Year 11			
Eligible	2.97***	0.29	0.38
	(0.64)	(0.59)	(0.57)
N	104,423	262,594	262,594
Not eligible	1.33***	0.29	0.19
	(0.24)	(0.28)	(0.28)
N	646,037	1,097,866	1,097,866
C. Museum stock before opening			
None	2.09***	0.24	0.29
	(0.37)	(0.41)	(0.42)
N	647,781	1,180,488	1,180,488
One	1.47***	0.38	0.43
	(0.36)	(0.39)	(0.38)
N	645,878	1,181,797	1,181,797
Two or more	1.15***	0.35	0.15
	(0.29)	(0.37)	(0.36)
N	665,717	1,212,978	1,212,978

Notes: This table reports pooled ATTs of a first museum opening within 2 km, estimated separately on each group of students from the main specification of Appendix Table A5, in percentile points for the KS5 average point score and in percentage points of the Year-11 cohort for the progression outcomes. Panel A splits students by tercile of KS2 prior attainment within the cohort, panel B by eligibility for free school meals in Year 11, and panel C by the number of museums operating within 2 km of the home LSOA in the year before the opening; in panel C only the treated group is split and the control students are shared. Standard errors clustered by LSOA are in parentheses. Figure 9 plots the estimates in column (1). * $p < 0.1$,

Appendix Table A9. Effects on the KS5 Average Point Score, by Museum Size and Governance

	(1) ≤1 km	(2) ≤2 km	(3) ≤5 km	(4) ≤10 km
A. Museum size				
Small (122 openings)	1.78***	1.42***	0.68***	0.00
	(0.41)	(0.25)	(0.15)	(0.13)
N	269,817	713,411	1,707,389	1,941,114
Medium + large (28)	1.53**	1.99***	1.20***	0.85***
	(0.68)	(0.39)	(0.26)	(0.21)
N	239,561	628,834	1,485,635	1,608,751
B. Governance (ownership)				
Public + not-for-profit (113)	1.48***	1.37***	0.70***	0.30**
	(0.40)	(0.24)	(0.15)	(0.13)
N	267,994	706,313	1,701,815	1,902,507
Private (38)	1.99***	1.95***	1.39***	−0.18
	(0.77)	(0.44)	(0.28)	(0.22)
N	237,504	626,127	1,461,906	1,581,426

Notes: This table reports pooled ATTs (percentile points) on the KS5 average point score by the size and governance of the first museum opening within the column radius, from the main specification of Appendix Table A5. Public and not-for-profit museums are run by government, a university or a not-for-profit organisation, including the National Trust and English Heritage; private museums are privately owned. Standard errors clustered by LSOA are in parentheses. Figure 10 plots the estimates. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table A10. Museum Stock by Theme and Level-3 Subject Choice: Cross-Sectional Associations

	Science			Arts			History			Society			Industry		
SSA subject area (km ring)	0–1	1–2	2–5	0–1	1–2	2–5	0–1	1–2	2–5	0–1	1–2	2–5	0–1	1–2	2–5
Academic subject areas															
2 Science & mathematics	**0.57**	**1.07**	**−0.17**	0.94	0.75	0.64	0.28	0.16	0.18	0.05	0.15	0.09	−0.30	−0.16	−0.17
9 Arts, media & publishing	0.43	0.57	−0.08	**0.57**	**0.39**	**0.31**	0.33	−0.03	−0.30	0.70	−0.09	−0.16	0.02	−0.11	−0.04
10 History, philosophy & theology	0.24	0.22	0.16	0.87	0.58	0.50	**0.51**	**0.29**	**0.04**	0.46	0.05	−0.13	−0.03	0.08	−0.25
12 Languages & literature	−0.11	0.45	−0.04	0.73	0.49	0.46	**0.62**	**0.18**	**−0.06**	0.36	0.06	−0.17	−0.13	−0.31	−0.26
11 Social sciences	0.62	0.99	0.29	0.68	0.74	0.63	0.37	0.10	0.16	**0.18**	**0.26**	**0.02**	−0.26	−0.21	−0.10
6 ICT	−0.03	−0.06	−0.00	−0.22	−0.25	0.00	−0.19	−0.14	0.09	−0.02	0.02	0.05	0.31	0.24	0.27
15 Business, admin & law	0.05	0.07	0.36	−0.34	0.09	0.13	−0.09	−0.13	−0.05	0.09	0.26	0.29	0.28	0.30	0.37
Vocational subject areas															
1 Health & public services	−0.50	−0.23	0.03	−0.00	−0.12	−0.03	−0.25	−0.18	−0.03	−0.13	−0.00	0.04	0.14	0.22	0.09
3 Agriculture & animal care	0.03	0.03	0.03	−0.03	−0.02	−0.01	−0.00	−0.01	−0.00	−0.05	−0.01	0.02	−0.02	0.00	0.01
4 Engineering & manufacturing	−0.02	0.11	−0.00	−0.11	−0.12	−0.10	0.00	0.01	0.09	0.11	0.09	0.16	**−0.06**	**−0.09**	**−0.01**
5 Construction	0.01	−0.04	0.01	0.04	0.01	−0.00	−0.03	−0.01	0.01	−0.01	−0.00	0.00	0.00	−0.01	0.02
8 Leisure & tourism	0.62	0.32	0.19	−0.22	−0.12	0.04	−0.05	−0.06	0.03	−0.16	0.07	0.10	−0.00	0.14	0.15
7 Retail & commercial	0.02	−0.04	−0.01	−0.04	−0.04	−0.01	−0.02	−0.02	0.00	−0.03	0.01	0.02	−0.01	0.01	0.01
Students	7,568,465														

Notes: This table reports coefficients (percentage points) from cross-sectional regressions of Level-3 entry in each sector subject area (rows) on indicators for a museum of each theme family within 0–1, 1–2 and 2–5 km of the Year-10 home (columns), entered jointly, with Local Authority × cohort fixed effects and student controls. Standard errors are not available. Bold cells are own-theme pairs. Figure 11 plots the coefficients.

Appendix Table A11. Effects of Museum Openings within 5 km on Level-3 Subject Choices, by Museum Theme

	(1) Any	(2) Science	(3) Industry	(4) Hist & Lit	(5) Arts	(6) Society
Any academic qualification	0.06	0.87	−0.75**	0.20	−0.05	0.40
	(0.15)	(1.17)	(0.36)	(0.29)	(0.40)	(0.25)
STEM — any entry	−0.02	**−0.01**	**0.97****	0.43	−0.02	0.19
	(0.16)	(1.12)	(0.38)	(0.31)	(0.40)	(0.28)
STEM — academic	0.13	**2.51****	**−0.47**	−0.00	0.16	0.81***
	(0.15)	(1.13)	(0.36)	(0.28)	(0.37)	(0.25)
Humanities — any entry	0.07	−3.10**	0.78**	**0.90*****	0.39	0.10
	(0.17)	(1.27)	(0.37)	(0.32)	(0.40)	(0.27)
Humanities — academic	0.15	−0.10	−0.38	**0.36**	0.19	0.12
	(0.14)	(1.24)	(0.34)	(0.26)	(0.38)	(0.25)
Arts — any entry	−0.53***	−2.93***	−0.35	0.14	**−0.24**	−0.57**
	(0.16)	(0.92)	(0.34)	(0.32)	(0.42)	(0.27)
Arts — academic	−0.45***	−0.97	−1.07***	−0.52**	**−0.33**	−0.67***
	(0.13)	(0.87)	(0.31)	(0.24)	(0.35)	(0.22)
Social science — any entry	0.06	−3.74***	1.16***	0.03	0.18	**1.18*****
	(0.18)	(1.33)	(0.40)	(0.33)	(0.41)	(0.30)
Social science — academic	0.16	−0.74	0.06	−0.51**	−0.01	**1.24*****
	(0.13)	(1.19)	(0.28)	(0.23)	(0.35)	(0.21)
Openings 2006–2012	167	5	29	37	27	49
Treated N per τ (median)	80,149	1,035	12,037	16,024	13,141	25,447

Notes: This table reports pooled ATTs (percentage points of the Year-11 cohort) of the first opening of a museum of the column theme within 5 km of the Year-10 home on Level-3 subject choices, from the main specification of Appendix Table A5 with not-yet-treated controls of any theme. The "Any" column pools all themes, and bold cells are the ex ante pairs (Section 2.2); the outcomes are defined in Section 2.1. Standard errors clustered by LSOA are in parentheses. Figure 12 plots the estimates. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.2 Supplementary Figures and Tables

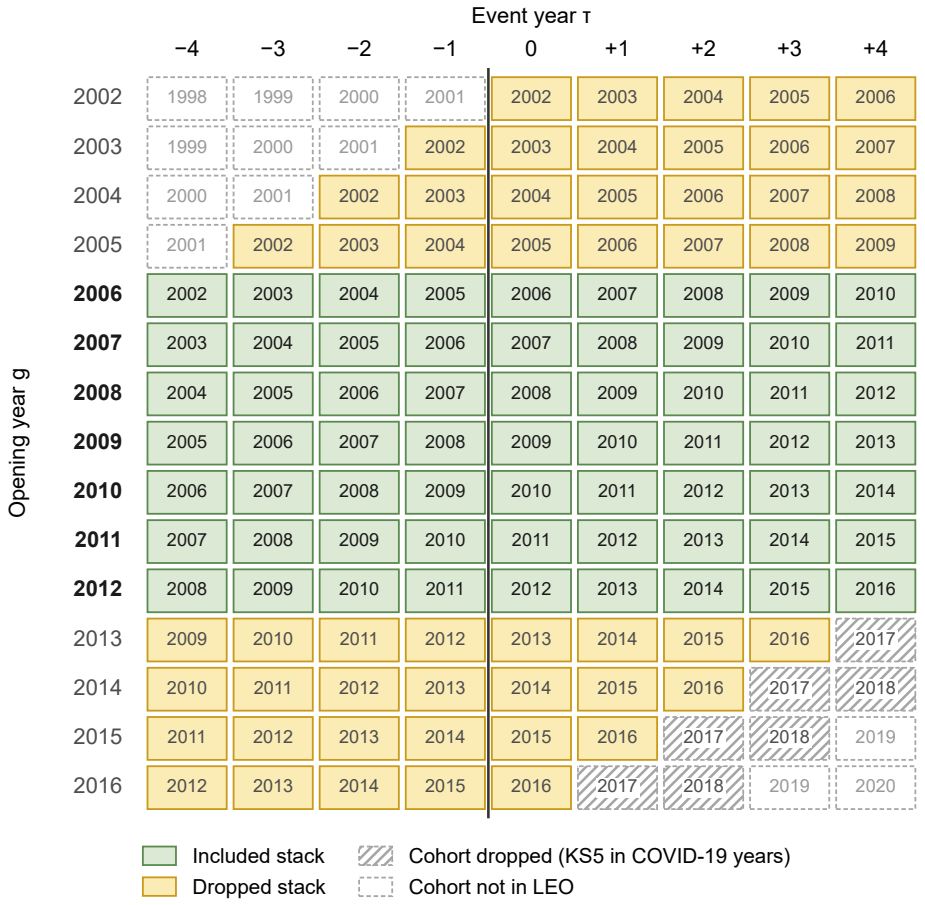
Appendix Table A12. Student Characteristics and National Benchmarks

	Mean	SD	DfE 2006–17
FSM eligible (%)	13.6	34.3	15.0
IDACI (pctile)	49.9	28.9	–
KS2 prior (pctile)	50.0	28.8	–
English first language (%)	88.9	31.4	86.9
Female (%)	49.1	50.0	49.4
SEN, any (%)	19.5	39.6	18.0
Urban (%)	83.2	37.4	–
White (%)	82.1	38.3	79.9
Asian (%)	7.4	26.2	8.6
Black (%)	4.1	19.8	4.7
Mixed (%)	3.1	17.3	3.7
Other ethnicity (%)	1.0	10.2	1.3
Observations	8,477,187		

Notes: This table compares the characteristics of the students in the sample with national figures. Binary variables are in percent; IDACI and KS2 prior attainment are within-cohort percentile ranks. The DfE column averages the published figures for state-funded secondary schools in England over the January censuses 2006–2017, which cover all secondary year groups.

Source: Department for Education, "Schools, Pupils and Their Characteristics" and "Special Educational Needs in England" (DfE column).

Appendix Figure A1. Treatment Cohorts with a Complete Event Window in the Stacked Design



Notes: This figure shows the treatment cohorts g (museum opening years; rows) and event years τ (columns) of the stacked design. Each cell gives the year $c(i) = g + \tau$ of the Year-10 spring census of the student cohort observed in that event year; the vertical line marks the opening. LEO covers Year-10 cohorts 2002–2018. Cohorts 2017 and 2018 are dropped (hatched) because their Key Stage 5 falls in the academic years 2019/2020 and 2020/2021, when examinations were replaced by centre- and teacher-assessed grades. A treatment cohort is included only if all student cohorts in $\tau \in \{-4, \ldots, 4\}$ are available, which holds for $g \in \{2006, \ldots, 2012\}$ (green). Dashed cells are cohorts without data.

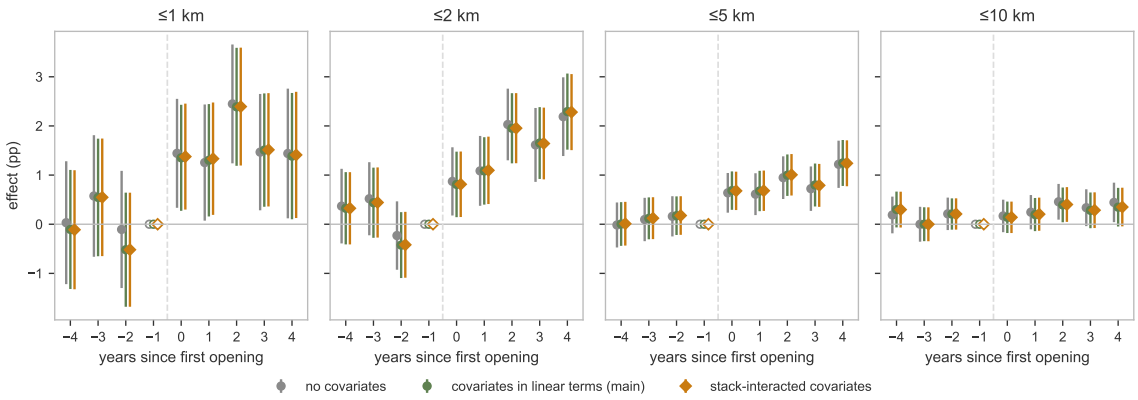
Appendix Table A13. Travel Time to Museums, Museum Stock around Students, and Reach of Openings

	Value
A. Travel time to the last museum visit (share of visits, %)	
≤30 min	48.5
31–60 min	30.5
61–90 min	10.5
91–120 min	6.9
121–150 min	1.7
151–180 min	1.9
Visits (unweighted N)	4,322
B. Mean number of operating museums around the Year-10 home	
Within 1 km	0.26
Within 2 km	0.97
Within 5 km	4.96
Students	7,570,468
C. Students ever within the band of a museum opening (%)	
Within 2 km	14.7
Within 5 km	45.0
Students	8,477,187

Notes: This table reports the survey-weighted distribution of adults' museum visits of up to three hours by travel time, with bins including their upper bound (Panel A), the mean number of museums operating within each distance of students' Year-10 homes (Panel B), and the share of students ever within each distance of a museum opening over 2001–2019 (Panel C).

Source: DCMS Taking Part survey 2008/09 (Panel A; Department for Culture, Media and Sport, 2010).

Appendix Figure A2. Event-Study Estimates for the KS5 Average Point Score, by Distance Band and Specification



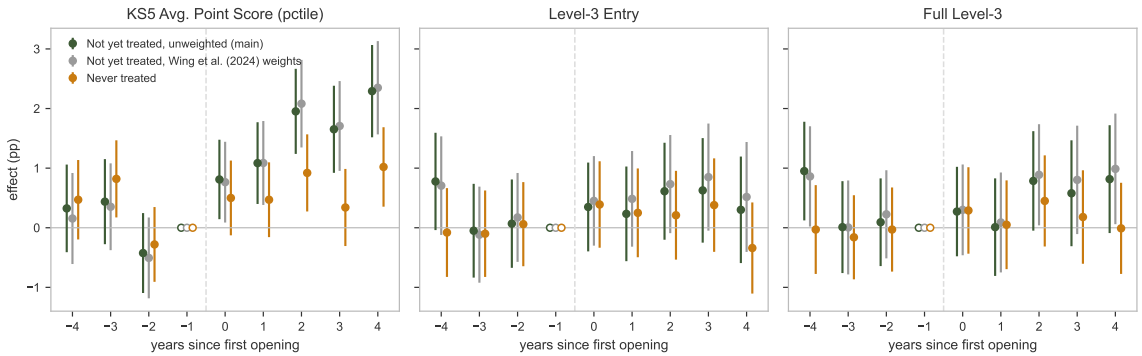
Notes: This figure shows event-study coefficients (percentile points) for the KS5 average point score around a museum opening within each distance band, from the stacked design of Figure 7. $\tau = -1$ is the omitted year (open markers). The three specifications are without covariates, with student covariates in linear terms (the main specification), and with stack-interacted covariates. 95% confidence intervals are based on standard errors clustered by LSOA. Appendix Table A14 reports the coefficients.

Appendix Table A14. Event-Study Estimates for the KS5 Average Point Score, by Distance Band and Specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
		≤1 km			≤2 km			≤5 km			≤10 km	
Covariates	None	Linear	By stack	None	Linear	By stack	None	Linear	By stack	None	Linear	By stack
$\tau = -4$	0.03	−0.11	−0.11	0.37	0.32	0.32	−0.01	0.00	0.01	0.19	0.30	0.30
	(0.64)	(0.62)	(0.62)	(0.39)	(0.38)	(0.37)	(0.23)	(0.23)	(0.23)	(0.19)	(0.19)	(0.19)
$\tau = -3$	0.57	0.54	0.55	0.52	0.44	0.44	0.10	0.12	0.12	0.00	0.00	0.00
	(0.63)	(0.61)	(0.61)	(0.38)	(0.36)	(0.36)	(0.23)	(0.22)	(0.22)	(0.18)	(0.18)	(0.18)
$\tau = -2$	−0.11	−0.52	−0.52	−0.23	−0.42	−0.42	0.16	0.17	0.18	0.21	0.21	0.21
	(0.61)	(0.59)	(0.59)	(0.35)	(0.34)	(0.34)	(0.21)	(0.20)	(0.20)	(0.17)	(0.16)	(0.16)
$\tau = -1$	0	0	0	0	0	0	0	0	0	0	0	0
	(ref.)	(ref.)	(ref.)	(ref.)	(ref.)	(ref.)	(ref.)	(ref.)	(ref.)	(ref.)	(ref.)	(ref.)
$\tau = 0$	1.44**	1.35**	1.38**	0.87**	0.81**	0.81**	0.64***	0.68***	0.68***	0.17	0.14	0.14
	(0.57)	(0.55)	(0.55)	(0.35)	(0.34)	(0.34)	(0.21)	(0.20)	(0.20)	(0.17)	(0.16)	(0.16)
$\tau = +1$	1.25**	1.30**	1.33**	1.09***	1.08***	1.10***	0.61***	0.68***	0.68***	0.25	0.20	0.20
	(0.60)	(0.58)	(0.58)	(0.36)	(0.35)	(0.35)	(0.22)	(0.21)	(0.21)	(0.18)	(0.17)	(0.17)
$\tau = +2$	2.45***	2.39***	2.39***	2.03***	1.95***	1.95***	0.95***	1.00***	1.01***	0.46**	0.39**	0.40**
	(0.62)	(0.61)	(0.61)	(0.37)	(0.36)	(0.36)	(0.22)	(0.22)	(0.22)	(0.19)	(0.18)	(0.18)
$\tau = +3$	1.47**	1.51**	1.51**	1.61***	1.65***	1.64***	0.72***	0.80***	0.79***	0.34*	0.28	0.29
	(0.60)	(0.59)	(0.59)	(0.38)	(0.37)	(0.37)	(0.23)	(0.22)	(0.22)	(0.19)	(0.19)	(0.19)
$\tau = +4$	1.44**	1.39**	1.41**	2.19***	2.29***	2.28***	1.22***	1.25***	1.24***	0.44**	0.35*	0.35*
	(0.67)	(0.66)	(0.65)	(0.41)	(0.39)	(0.39)	(0.25)	(0.24)	(0.24)	(0.20)	(0.20)	(0.20)
Pooled ATT	1.49***	1.62***	1.64***	1.40***	1.48***	1.48***	0.76***	0.80***	0.80***	0.23*	0.15	0.15
	(0.39)	(0.37)	(0.37)	(0.23)	(0.23)	(0.23)	(0.14)	(0.14)	(0.14)	(0.12)	(0.12)	(0.12)
Treated N per τ	9,715	9,715	9,715	28,228	28,228	28,228	80,439	80,439	80,439	118,845	118,845	118,845
Control N per τ	45,955	45,955	45,955	123,569	123,569	123,569	285,417	285,417	285,417	308,998	308,998	308,998

Notes: This table reports event-study coefficients (percentile points) for the KS5 average point score around a first museum opening within the column radius, from the design of Appendix Table A5 without covariates (None), with covariates in linear terms (Linear; the main specification) or with stack-interacted covariates (By stack). $\tau = -1$ is the omitted year. N are medians across event years. Standard errors clustered by LSOA are in parentheses. Figure A2 plots the coefficients. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Figure A3. Robustness to the Control Group and Weighting: Event Studies within 2 km



Notes: This figure shows event-study coefficients (×100) of a museum opening within 2 km under alternative control groups and weights, with student covariates in linear terms. Dark green markers are the main specification (stack-wise not-yet-treated controls, OLS), grey markers use the same controls with the corrective observation weights of Wing et al. (2024), and orange markers use never-treated controls. The two sets of estimates are close for these outcomes. $\tau = -1$ is the omitted year (open markers); 95% confidence intervals are based on standard errors clustered by LSOA. Appendix Table A15 reports the coefficients of the main specification and of never-treated controls.

Appendix Table A15. Robustness to the Control Group: Event Studies within 2 km

	(1)	(2)	(3)	(4)	(5)	(6)
	KS5 avg. point score		Level-3 entry		Full Level-3	
	Not yet treated	Never treated	Not yet treated	Never treated	Not yet treated	Never treated
$\tau = -4$	0.32	0.47	0.78*	−0.08	0.95**	−0.03
	(0.38)	(0.34)	(0.42)	(0.38)	(0.42)	(0.38)
$\tau = -3$	0.44	0.82**	−0.05	−0.10	0.01	−0.16
	(0.36)	(0.33)	(0.40)	(0.37)	(0.39)	(0.36)
$\tau = -2$	−0.42	−0.28	0.07	0.06	0.09	−0.03
	(0.34)	(0.32)	(0.38)	(0.36)	(0.38)	(0.36)
$\tau = -1$	0	0	0	0	0	0
	(ref.)	(ref.)	(ref.)	(ref.)	(ref.)	(ref.)
$\tau = 0$	0.81**	0.50	0.35	0.39	0.27	0.29
	(0.34)	(0.32)	(0.38)	(0.37)	(0.38)	(0.37)
$\tau = +1$	1.08***	0.47	0.23	0.25	0.01	0.05
	(0.35)	(0.32)	(0.41)	(0.38)	(0.42)	(0.38)
$\tau = +2$	1.95***	0.92***	0.61	0.21	0.79*	0.45
	(0.36)	(0.33)	(0.42)	(0.38)	(0.43)	(0.39)
$\tau = +3$	1.65***	0.34	0.63	0.38	0.58	0.18
	(0.37)	(0.33)	(0.45)	(0.40)	(0.45)	(0.40)
$\tau = +4$	2.29***	1.02***	0.30	−0.34	0.82*	−0.01
	(0.39)	(0.34)	(0.46)	(0.39)	(0.46)	(0.39)

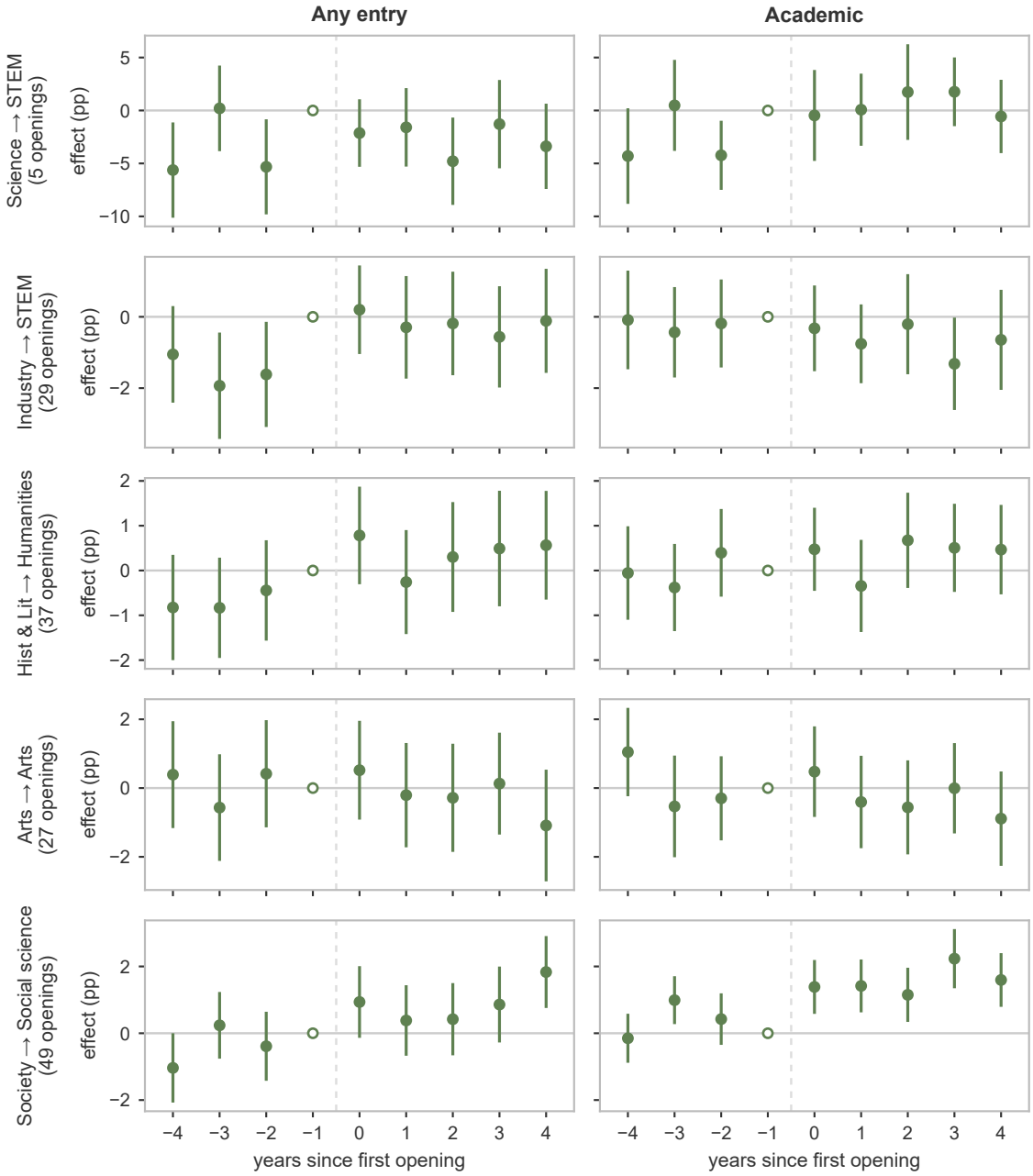
Notes: This table reports event-study coefficients of a museum opening within 2 km from the main specification, with not-yet-treated controls, and with never-treated controls, that is, areas with no opening within 2 km over the sample period. Units are as in Appendix Table A4, and $\tau = -1$ is the omitted year. Standard errors clustered by LSOA are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table A16. Effects of Museum Openings within 2 km on Level-3 Subject Choices, by Museum Theme

	(1) Any	(2) Science	(3) Industry	(4) Hist & Lit	(5) Arts	(6) Society
Any academic qualification	0.28	2.00	−1.14**	1.00*	0.67	0.60
	(0.25)	(1.47)	(0.55)	(0.53)	(0.69)	(0.45)
STEM — any entry	0.18	**−0.34**	**0.49**	0.08	1.23	0.41
	(0.26)	(1.37)	(0.58)	(0.49)	(0.95)	(0.53)
STEM — academic	0.33	**1.33**	**−1.21***	0.49	1.63**	1.14***
	(0.24)	(1.40)	(0.63)	(0.47)	(0.83)	(0.42)
Humanities — any entry	0.19	−1.23	0.03	**0.68**	1.62**	0.27
	(0.27)	(2.03)	(0.68)	(0.55)	(0.71)	(0.46)
Humanities — academic	0.35	0.21	−1.70***	**1.24*****	1.58***	0.36
	(0.22)	(1.46)	(0.58)	(0.40)	(0.58)	(0.46)
Arts — any entry	−0.23	−3.84**	1.26*	−0.33	**−0.57**	−0.52
	(0.26)	(1.73)	(0.72)	(0.50)	(0.63)	(0.51)
Arts — academic	−0.12	−1.87	−0.10	−0.03	**−0.40**	−0.63
	(0.20)	(1.98)	(0.59)	(0.39)	(0.49)	(0.42)
Social science — any entry	0.30	−0.70	0.62	0.22	0.63	**1.54*****
	(0.28)	(1.89)	(0.67)	(0.55)	(0.75)	(0.53)
Social science — academic	0.49**	0.81	−1.12**	0.72*	0.62	**1.66*****
	(0.21)	(1.11)	(0.50)	(0.41)	(0.50)	(0.38)
Openings 2006–2012	167	5	29	37	27	49
Treated N per τ (median)	28,195	643	4,013	5,989	4,903	8,104

Notes: This table reports the estimates of Appendix Table A11 for openings within 2 km of the Year-10 home LSOA; see the notes to that table for definitions. Standard errors clustered by LSOA are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Figure A4. Event-Study Estimates for Level-3 Subject Choices within 5 km, by Museum Theme



Notes: This figure shows event-study coefficients (percentage points of the Year-11 cohort) behind the ex ante pairs of Figure 12: the first opening of a museum of the row theme within 5 km of the Year-10 home, for any Level-3 entry in the matching subject group (left) and for such an entry together with an academic qualification (right). Estimates come from the design of Appendix Table A11, with stack-wise not-yet-treated controls of any theme. $\tau = -1$ is the omitted year (open markers). Row labels give the number of openings of each theme in 2006–2012. 95% confidence intervals are based on standard errors clustered by LSOA.

B Appendix: A Closer Look at the Openings

This section provides detailed information on the 167 museum opening events that generate the treatment variation for identifying the main effects. I first document the spatial distribution of the museum openings during this period and characterise their surroundings. Then, I provide some typical stories of why these museums opened. The information collected for this section comes from both the Mapping Museums UK dataset and publicly available sources, including public registries, museum websites, Wikipedia pages, and OpenStreetMap (OSM).

B.1 Where Are the Museum Openings?

In the identification, treatment variation comes from 167 museums opened between 2006 and 2012. These museums are distributed across all nine regions of England, broadly aligned with the regional distribution of the museum stock in 2005, as shown in Panel A of Appendix Table B1 (see Figure 2 for a map). According to the ONS 2011 Rural–Urban Classification, 66% of the openings happened in urban LSOAs (Panel B of Appendix Table B1).

Out of the 167 newly opened museums, 40.1% are the first within 2 km and 15.6% within 5 km (see Appendix Table B2). In the opening year, 27.5% have 1 or 2 museums within 2 km, 12.6% have 3–4, and the remaining 19.8% have 5 or more. At 5 km, 28.1% have 1–2 museums around, while 40.1% have 5 or more museums around.

Regarding the buildings where these museums are located, 89% opened in an existing building or venue, while 11% opened in new buildings. The geographical distribution of these two types is across all regions and does not show a clustering pattern, as shown in Appendix Figure B1.

Appendix Table B1. Location of Museum Openings and the 2005 Museum Stock

	Museum stock, 2005		Openings, 2006–2012	
	N	Share (%)	N	Share (%)
A. Region				
North East	101	4.0	12	7.2
North West	261	10.3	20	12.0
Yorkshire and The Humber	234	9.2	16	9.6
East Midlands	228	9.0	12	7.2
West Midlands	230	9.0	13	7.8
East of England	329	12.9	29	17.4
London	230	9.0	22	13.2
South East	484	19.0	24	14.4
South West	445	17.5	19	11.4
Total	2,542	100.0	167	100.0
B. Urban–rural classification of the LSOA (ONS 2011)				
Urban	1,560	61.4	111	66.5
Major conurbation	480	18.9	39	23.4
Minor conurbation	52	2.0	3	1.8
City and town	1,011	39.8	69	41.3
City and town, sparse setting	17	0.7	0	0.0
Rural	982	38.6	56	33.5
Town and fringe	368	14.5	22	13.2
Town and fringe, sparse setting	46	1.8	3	1.8
Village and dispersed	489	19.2	24	14.4
Village and dispersed, sparse setting	79	3.1	7	4.2
Total	2,542	100.0	167	100.0

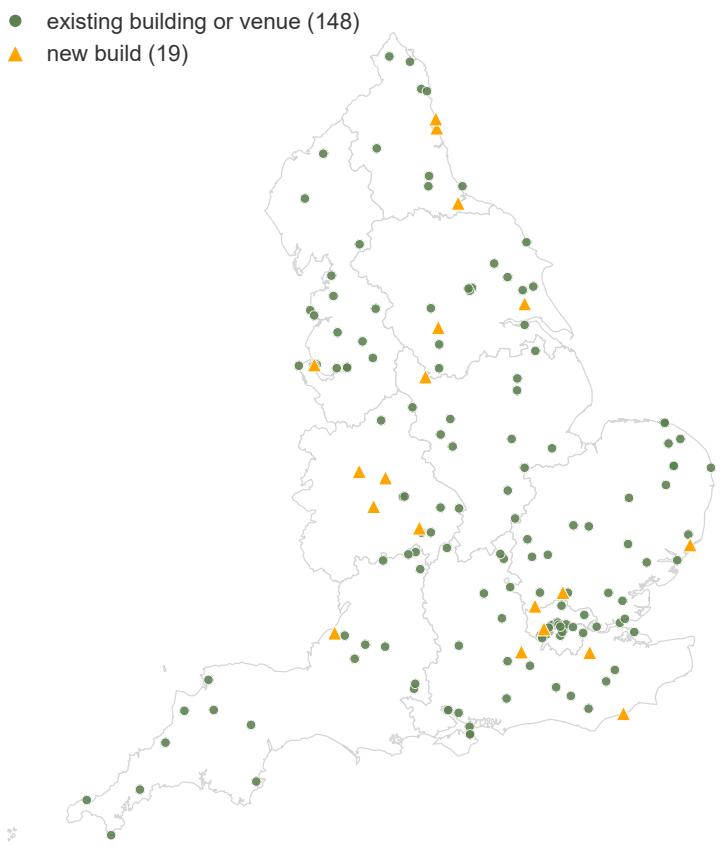
Notes: This table compares the location of the 167 museums opened in 2006–2012 with that of the 2,542 museums operating in England in 2005. Panel B uses the ONS 2011 Rural–Urban Classification of each museum's LSOA. Pearson χ^2 goodness-of-fit tests of the openings against the stock shares give $p = 0.031$ for regions, $p = 0.18$ for urban versus rural, and $p = 0.55$ for the eight classes.

Appendix Table B2. Museums Operating within 2 km and 5 km of Each Opening

Other operating museums	Within 2 km			Within 5 km		
	Openings	Share (%)	Cumulative (%)	Openings	Share (%)	Cumulative (%)
0	67	40.1	40.1	26	15.6	15.6
1	29	17.4	57.5	28	16.8	32.3
2	17	10.2	67.7	19	11.4	43.7
3–4	21	12.6	80.2	27	16.2	59.9
5–9	14	8.4	88.6	28	16.8	76.6
10–19	11	6.6	95.2	20	12.0	88.6
20–49	6	3.6	98.8	10	6.0	94.6
50 or more	2	1.2	100.0	9	5.4	100.0
Total	167	100.0		167	100.0	
Mean		4.2			11.1	
Median		1			3	
75th percentile		3			8	
90th percentile		11			21	
Maximum		56			118	

Notes: This table reports the distribution of the 167 museums opened in 2006–2012 by the number of other museums operating within 2 km and 5 km in the opening year, including other museums opening in the same year. Distances are straight-line distances between museums.

Appendix Figure B1. Museum Openings in Existing Buildings and New Builds



Notes: This figure maps the 167 museums opened in 2006–2012 by the building in which each museum opened. Existing buildings or venues comprise buildings converted to museum use (60), space within a building that keeps its main use, such as a library, town hall or stadium (50), historic buildings or sites opened in situ (31), and outdoor sites (7). New builds are buildings or wings built for the museum. New builds are not associated with region (permutation χ^2 test, $p = 0.28$) and are no closer to one another than 19 locations drawn at random from the 167 openings ($p = 0.61$).

B.2 Some Typical Opening Stories

Compared to the stock of 2,542 museums in 2005, the newly opened ones are more likely to be small and less likely to be government-run, according to Appendix Table B3. Most of them were opened by non-profit organisations. Based on the sources above, volunteer societies and trusts opened 72 of the 167 museums, whereas local authorities opened only 12. The reasons for opening are mostly specific to each museum, such as a community heritage project (76 openings), a building becoming available (73), a collection that an institution or a private collector wanted to put on display

(55 and 45), or an anniversary (12), with many openings having more than one of these reasons. For instance, Thame Town Council bought the town's former courthouse in 2005 to house a museum, which opened in 2007 and is run by volunteers; the Museum of British Surfing, after years of touring exhibitions, found a permanent home in 2012 in a railway building in Braunton owned by the parish council; and Benjamin Franklin House in London opened in 2006, on the 300th anniversary of Franklin's birth.

Only 20 openings were linked to a local regeneration programme, the case closest to the concern in Section 4.2 that a new museum comes with wider changes in its neighbourhood. They include large new buildings such as the Middlesbrough Institute of Modern Art (2007), the Hepworth Wakefield (2011) and the Museum of Liverpool (2011), each of which succeeded a museum or gallery that had closed nearby. For most openings, the timing thus reflects when a building, a collection, volunteers or funding became available, rather than changes in the neighbourhood around the museum.

Appendix Table B3. Size and Governance of Museum Openings and the 2005 Museum Stock

	Museum stock, 2005		Openings, 2006–2012	
	N	Share (%)	*N*	Share (%)
A. Size (annual visits)				
Small (up to 10,000)	1,347	53.0	122	73.1
Medium (10,001–50,000)	730	28.7	13	7.8
Large (50,001–1 million)	412	16.2	15	9.0
Huge (over 1 million)	10	0.4	0	0.0
Unknown	43	1.7	17	10.2
Total	2,542	100.0	167	100.0
B. Governance				
Government	611	24.0	15	9.0
Local authority	563	22.1	11	6.6
National	47	1.8	3	1.8
Other	1	0.0	1	0.6
Independent	1,812	71.3	145	86.8
Not-for-profit	1,111	43.7	89	53.3
Private	384	15.1	38	22.8
National Trust	158	6.2	3	1.8
English Heritage	49	1.9	2	1.2
Type unknown	110	4.3	13	7.8
University	68	2.7	4	2.4
Unknown	51	2.0	3	1.8
Total	2,542	100.0	167	100.0

Notes: This table compares the size and governance of the 167 museums opened in 2006–2012 with those of the 2,542 museums operating in England in 2005.

C Appendix: Museum-Going in England

This appendix presents a detailed descriptive analysis for understanding the museum-going behaviour of students in England, based on individual-level data from the Taking Part Survey (Department for Culture, Media and Sport, 2010–2024). I use 15 waves of this survey, from 2005/06 to 2019/20. This annually-run survey primarily collects data on cultural participation of a representative sample of English residents, including a section devoted to museum visits.

The survey consists of, first, an adult questionnaire for respondents over 16 years old when interviewed. Second, it has a child part for children between 5 and 15 years old, further divided into two sections targeting children aged from 5 to 10 and from 11 to 15, respectively, with the former introduced in 2008 and answered by parents. In the analysis, I focus on a sub-sample of the 11-to-15 year-olds in the child survey and of 16-to-18 year-olds in the adult survey, as they fall in the age range most relevant to the main analysis in this paper. Due to privacy reasons, from 2016 onwards, the survey no longer reports the exact age of the respondents but only age bands, and thus, for the adult subsample, I focus on 16-to-19 instead of 16-to-18 as it is the most comparable age group among those available.

Across different waves, the survey preserves basic information on museum-visit behaviour of the respondents, including whether they have visited any museum in the past year, how frequent their visits have been, why they have made the visits, etc. In some waves, this information is complemented by the travel time spent for the visits. Below, I provide a sketch of museum-going behaviour mainly based on these outcomes. Three facts stand out for the main analysis: most museum visits by school-age children take place outside school, museum-going is not confined to an affluent minority, and most visits involve short trips, consistent with treating museums as a local amenity.

C.1 Data Processing

All estimates reported in this appendix are weighted using the survey sampling weights provided with each wave. The Taking Part Survey applies a multi-stage probability sampling design with post-stratification rim weighting to adjust for differential non-response. The naming convention and construction of the weight variable changed several times across the fifteen waves; to ensure comparability across waves, I normalise all weights to a mean of one within each year, so that weighted sample

sizes are approximately equal to raw sample sizes.[13]

To characterise the museum-going behaviour, I make some choices regarding variables and their processing. First, the question on whether the respondent visited a museum or gallery in the past twelve months and the follow-up question on visit frequency are asked consistently across all fifteen waves, making them the primary variables for longitudinal analysis. Second, a question distinguishing whether children's visits took place during school lessons or in spare time is available for ten waves (2010/11 through 2019/20). The response categories for visit frequency were restructured between the 2007/08 and 2008/09 waves; the pooled frequency analysis in Appendix Figure C2 therefore uses only 2008/09 onward to maintain comparability (2010/11 onward for children aged 11 to 15). Third, exact age is recorded in the survey from 2005/06 through 2015/16, allowing precise identification of 16-to-18-year-olds in the adult sample. From 2016/17 onward, only banded age groups are available, and the narrowest band containing this age range is 16-to-19; I use this broader band for the later waves, as noted in the introduction to this appendix.

C.2 Museum Visit Rate

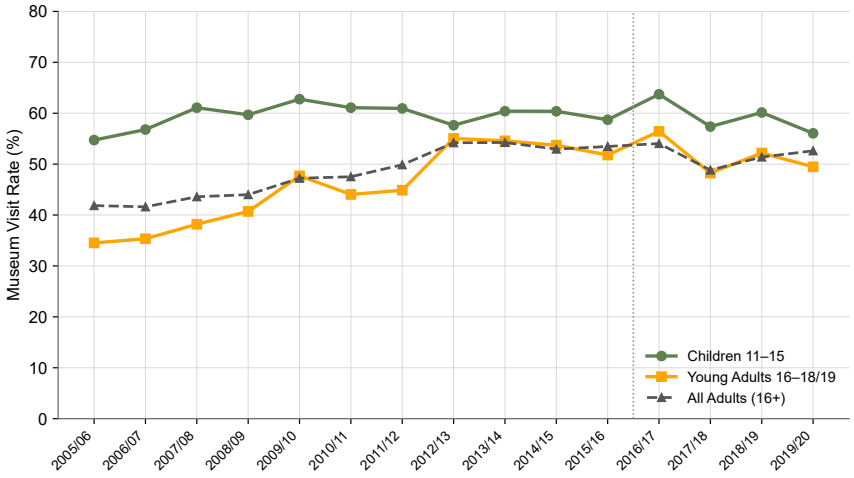
Appendix Figure C1 presents the percentage of respondents who reported visiting a museum in the past twelve months, from 2005/06 to 2019/20, by age group. Children aged 11 to 15 exhibit the highest visiting rate throughout the period, fluctuating between 55% and 64% with no pronounced trend. The rate for all adults rises steadily from approximately 42% in 2005/06 to above 50% by 2012/13, and stays between 49% and 54% thereafter. Young adults aged 16 to 18 (16 to 19 from 2016/17 onward) start at a considerably lower rate of around 35% in 2005/06, converge toward the all-adult rate by 2012/13, and then fluctuate between 48% and 56% in the later waves. The greater volatility in the young adult series possibly reflects the small sub-sample size in the adult survey. The gap between children and adults suggests the importance of visits as part of school activities, consistent with the descriptives presented in Appendix Figure C3 below, where about one in five child visitors visit only through school.

Visit intensity conditional on participation does not differ markedly between the two age groups of interest. Appendix Figure C2 shows the distribution of visit fre-

[13]The 2006/07 adult survey does not include a full rim-weighted variable. For that wave, I use the best available approximation, a design weight with non-response adjustment and capping, which may not fully correct for demographic imbalances. This limitation affects only one of the fifteen waves and has negligible impact on the overall trends.

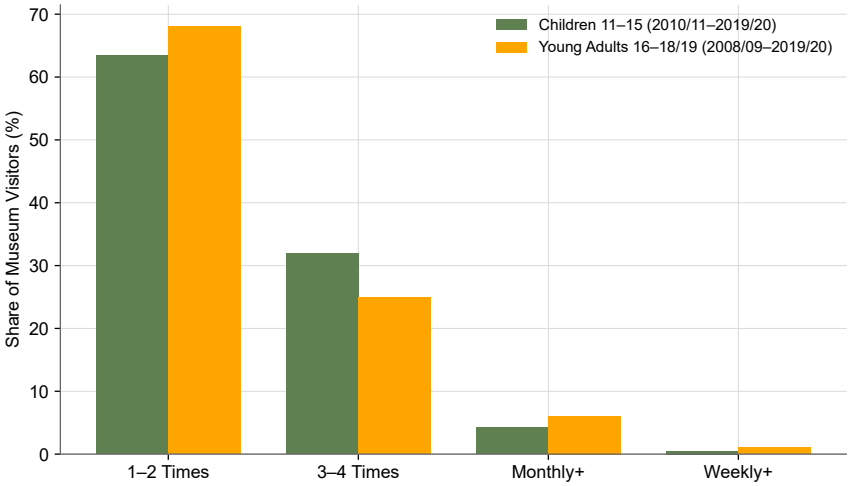
quency among those who reported at least one museum visit, pooling across years to increase precision. Over 60% of young adult visitors (aged 16 to 18/19) and a similar share of child visitors (aged 11 to 15) report visiting once or twice in the past twelve months. Approximately 25% to 32% visit three or four times a year, and fewer than 10% visit monthly or more often.

Appendix Figure C1. Museum Visit Rate by Age Group



Notes: This figure shows the share of respondents who visited a museum or gallery in the previous twelve months, by survey wave and age group. From 2016/17 (dotted line), the young-adult group covers ages 16–19 instead of 16–18, because the survey no longer records exact age. Estimates use the survey weights.

Appendix Figure C2. Frequency of Museum Visits, by Age Group



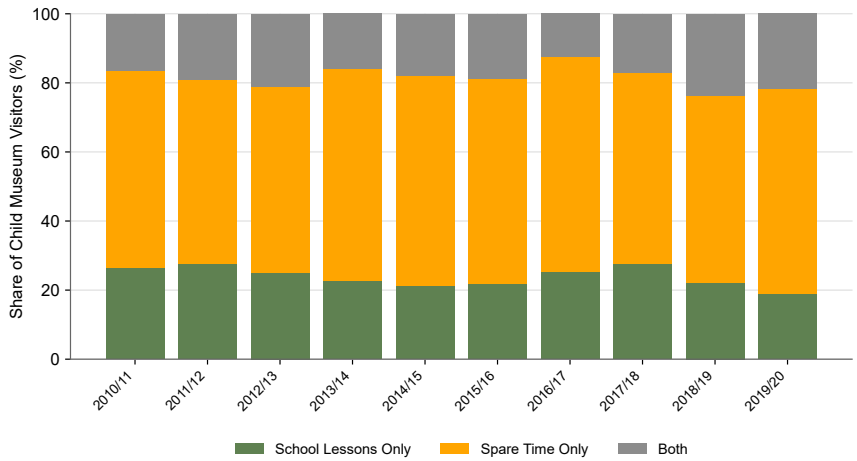
Notes: This figure shows the distribution of the number of museum visits in the previous twelve months among respondents who visited at least once, in a four-category scheme harmonised across waves. Young adults aged 16–18/19 are pooled over 2008/09–2019/20 and children aged 11–15 over 2010/11–2019/20. Estimates use the survey weights.

C.3 Contexts of Visits for Children

School-age children typically visit museums either on school trips as part of their curriculum or outside of school with family and friends. Appendix Figure C3 decomposes the visiting population of children aged 11 to 15 into three mutually exclusive categories: those who visited only during school lessons, those who visited only in their spare time, and those who visited in both contexts. Pooling the ten waves, 58% of child museum visitors report visiting only outside of school, another 18% report visiting both outside of school and through school, and only 24% report visiting only through school. This pattern is remarkably stable across the ten waves for which the question is available (2010/11 to 2019/20), with no discernible trend in the relative importance of either channel.

Taken together, three in four child visitors (76%, pooling the ten waves) made at least one museum visit in their spare time. Measured against all children aged 11 to 15, 46% visited a museum in their spare time in the past year, compared with 25% during school lessons. Museum-going at this age is therefore mainly a leisure activity organised around the home rather than the school, which is why the main analysis measures exposure by the distance between museums and students' home locations.

Appendix Figure C3. Museum Visits in School Lessons and in Spare Time among Children

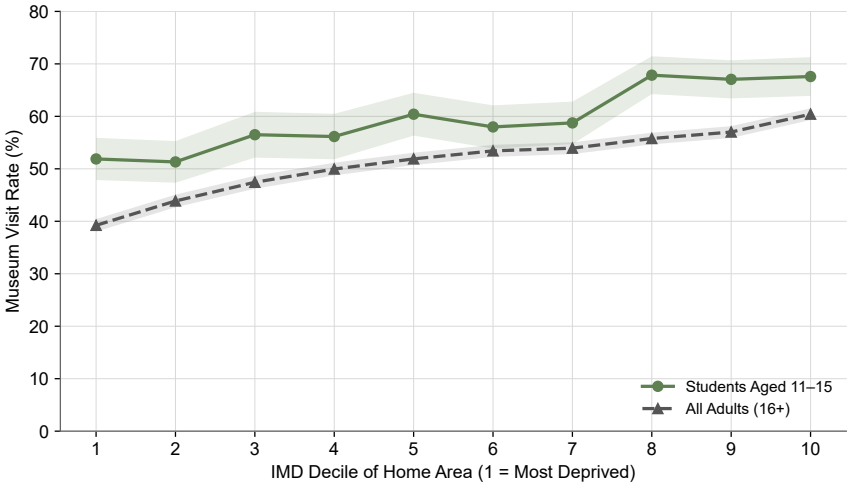


Notes: This figure shows, for each wave from 2010/11 to 2019/20, the shares of children aged 11–15 who visited a museum in the previous twelve months only during school lessons, only in their spare time, or in both. Estimates use the survey weights.

C.4 Who Visits Museums

Appendix Figure C4 shows the museum visiting rate by the deprivation of the respondent's home area, measured by its decile on the Index of Multiple Deprivation (IMD), for all adults and for students aged 11 to 15. Visiting rises with area affluence, but much less among students than among adults: from 52% to 68% for students, against 39% to 60% for all adults. Students living in the most deprived tenth of areas visit about as often as the average adult (52%). Museum-going has a social gradient, but it is not confined to an affluent minority.

Appendix Figure C4. Museum Visit Rate by Deprivation of the Home Area

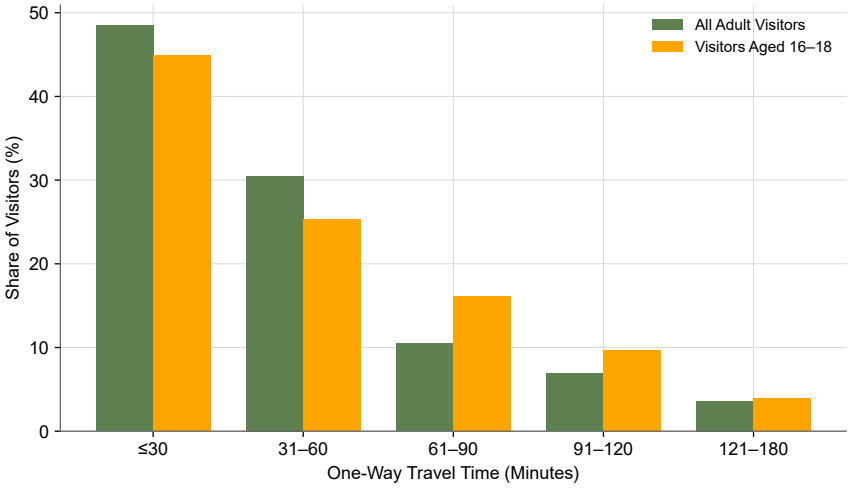


Notes: This figure shows the share of respondents who visited a museum in the previous twelve months, by decile of the Index of Multiple Deprivation of their home area (1 = most deprived), for all adults pooled over 2009/10–2019/20 and for students aged 11–15 pooled over 2010/11–2019/20. Estimates use the survey weights. Shaded areas are 95% confidence intervals.

C.5 Travel Time

The Taking Part Survey does not include a direct measure of the geographic distance between respondents and the museums they visit. However, the 2008/09 wave asked museum visitors to report the time spent travelling to the museum they last visited, recorded in hours and minutes (Department for Culture, Media and Sport, 2010). This allows an indirect assessment of how far visitors typically travel. Appendix Figure C5 displays the distribution of one-way travel time (for trips of up to 180 minutes) for all adult visitors and for the 16-to-18 sub-sample. Among all visitors, the median travel time is approximately 30 minutes, with 48% reaching the museum within 30 minutes and 79% within one hour. Among visitors aged 16 to 18, the distribution is broadly similar, 45% within 30 minutes and 70% within one hour, though the small sample size warrants caution in interpreting the tail of the distribution. Museums are thus mostly visited as a local amenity. Travel time does not map one-to-one into distance, as the survey does not record the mode of travel, but a trip of up to half an hour on foot, by public transport or by car typically covers a few kilometres rather than a few hundred metres or several tens of kilometres. This is the scale of the radii used in the main analysis (1 to 10 km), although the figures do not single out any specific cut-off.

Appendix Figure C5. Travel Time to Museums, by Age Group



Notes: This figure shows the distribution of one-way travel time to the museum last visited, in the 2008/09 wave, for all adult visitors (N = 4,322) and visitors aged 16–18 (N = 95). Trips of more than 180 minutes are excluded, and bins include their upper bound. Estimates use the survey weights.

D Appendix: Decomposing the Stacked Regression Estimates

This appendix shows that the OLS stacked regression of Section 3 estimates a convex average of the stack-specific effects, derives the stack weights it implies, and shows how the corrective observation weights of Wing et al. (2024) fit in.

Consider the pooled regression of Eq. (2), estimated by weighted least squares with observation weights $w_i > 0$, where $w_i = 1$ (OLS) in the main specification, without covariates. Write Post_{ig} as P_{ig}, and let $\mathbf{Z}$ be the vector of regressors except $D_{ig}P_{ig}$. For stack g, let N_g^D and N_g^C be the numbers of treated and control students, let $M_g^D = \sum_{i \in g} w_i D_{ig}$ and $M_g^C = \sum_{i \in g} w_i (1 - D_{ig})$ be their weighted counterparts, with $M_g = M_g^D + M_g^C$, let $p_g = M_g^D / M_g$ be the weighted treated share, and let q_g be the weighted share of post-opening cohorts. Let ATT_g be the effect identified by the DiD from stack g.

The event window is the same in every stack (Section 3). In addition, here I call the design "balanced" if the distribution of students over the event years is the same for treated and control students, in every Local Authority, and in every stack. Balance implies that the share of treated is the same, p_g, in every event year of stack g, and q_g takes a common value q. It holds exactly in a balanced panel and approximately in the repeated cross-section of cohorts used here (Appendix Table D1).

Proposition 1 (Stack weights)**.** *Under balance, let $\tilde{D}_{ig}$ be the residual from the weighted projection of $D_{ig}P_{ig}$ on $\mathbf{Z}$, and define*

$$S_g = \sum_{i \in g} w_i \tilde{D}_{ig}^2, \qquad \hat{\omega}_g = \frac{S_g}{\sum_h S_h}, \qquad \hat{\beta}_g = \frac{\sum_{i \in g} w_i \tilde{D}_{ig} Y_{ig}}{S_g}. \tag{D1}$$

Then (i) $\hat{\beta} = \sum_g \hat{\omega}_g \hat{\beta}_g$, where $\hat{\beta}_g$ is the coefficient obtained by estimating the regression on stack g alone and

$$\hat{\omega}_g \propto M_g \, p_g (1 - p_g) = \frac{M_g^D M_g^C}{M_g^D + M_g^C}, \tag{D2}$$

which in OLS equals $N_g^D N_g^C / (N_g^D + N_g^C)$; and (ii) each event-study coefficient of Eq. (1) is a weighted average of the coefficients from all stacks with the same stack weights.

Proof. Write $D_{ig}P_{ig} = (D_{ig} - p_g)(P_{ig} - q) + p_g P_{ig} + q D_{ig} - p_g q$. The first term is

orthogonal to $\mathbf{Z}$, since within every cell defined by stack, treatment status, and Local Authority, its weighted sum is $(D_{ig} - p_g)$ times the weighted sum of $P_{ig} - q$, which is zero under balance, and within every cohort of stack g it is $(P_{ig} - q)$ times the weighted sum of $D_{ig} - p_g$, which is zero because the cohort's treated share is p_g. The second term is a function of stack and year, the third is q times the treated-group dummy with the same q in every stack, and the fourth is constant within a stack. All of the latter three lie in the span of $\mathbf{Z}$. Therefore, $\tilde{D}_{ig} = (D_{ig} - p_g)(P_{ig} - q)$.

By the Frisch–Waugh–Lovell theorem, $\hat{\beta} = \sum_i w_i \tilde{D}_{ig} Y_{ig} / \sum_i w_i \tilde{D}_{ig}^2$, and splitting both sums by stack gives $\hat{\beta} = \sum_g \hat{\omega}_g \hat{\beta}_g$. The regression on stack g alone has the same form, with $\mathbf{Z}$ restricted to the columns of stack g, and the argument above applies to it unchanged because every sum in it is taken within a stack. Therefore, its residual is still $\tilde{D}_{ig}$, and its coefficient is $\hat{\beta}_g$. Squaring and summing within stack g gives $S_g = M_g \, p_g(1 - p_g) \, q(1 - q)$, and the factor in q cancels in $\hat{\omega}_g$.

For the coefficient at event year t in Eq. (1), the residual is $(D_{ig} - p_g) \, (q_{-1}\mathbf{1}\{\tau = t\} - q_t \mathbf{1}\{\tau = -1\})/(q_t + q_{-1})$, with q_t the weighted share of event year t. It is zero outside event years t and -1, so it is orthogonal to the interactions for the other event years, and the cell sums above vanish as before because its second factor has weighted mean zero in every cell. What remains of $D_{ig}\mathbf{1}\{\tau = t\}$ after subtracting it equals $q_t(q_t + q_{-1})^{-1} D_{ig}(\mathbf{1}\{\tau = t\} + \mathbf{1}\{\tau = -1\})$ plus a function of stack and year. The first part is a combination of the treated-group dummy and the other interactions with the same coefficient in every stack. Squaring and summing then gives $S_{g,t} = M_g \, p_g(1 - p_g) \, q_t q_{-1}/(q_t + q_{-1})$, so $\hat{\omega}_{g,t} = \hat{\omega}_g$. □

If parallel trends hold within every stack, each $\hat{\beta}_g$ converges to ATT_g, so $\hat{\beta}$ converges to the convex combination $\sum_g \omega_g \mathrm{ATT}_g$ of identified causal effects, whatever the treated shares across stacks. The stack weights depend only on the design, not on the outcome. A pre-opening coefficient is $\sum_g \hat{\omega}_{g,\tau} \hat{\beta}_{g,\tau}$, so if parallel trends fail in one stack, reweighting rescales the violation but cannot remove it. Because q is common to all stacks, the single pooled treated-group dummy of Eq. (2) absorbs the same component of $D_{ig}P_{ig}$ as seven stack-specific ones would, and a level gap between treated and control students that differs across stacks, $\delta_g \neq \delta$, is orthogonal to $\tilde{D}_{ig}$ within every stack and leaves $\hat{\beta}$ unchanged.

Eq. (D2) is a harmonic-type mean of the treated and control counts: when controls are abundant, $\omega_g \approx N_g^D$, the treated-share weight. But when controls are scarce, as in the later stacks under stack-wise not-yet-treated controls, $\omega_g \approx N_g^C$ and the stack is down-weighted. Under homoskedastic, independent errors and $w_i = 1$, $\mathrm{Var}(\hat{\beta}_g) \propto$

$1/S_g$, so these are inverse-variance stack weights, and OLS is the minimum-variance combination of the stack-level estimates when the effect is homogeneous.

Observation weights. By Proposition 1, any target stack weights π with $\pi_g > 0$ are reached by observation weights that keep every stack balanced and give the two arms of stack g weighted sizes $a\pi_g$ and $b\pi_g$, with a and b common to all stacks, since then p_g is the same in every stack and $\hat{\omega}_g = \pi_g$. The corrective observation weights of Wing et al. (2024) are the case $\pi_g = N_g^D/N^D$: treated students keep a weight of one and control students in stack g receive the weight $N_g^D N^C/(N_g^C N^D)$, so that $M_g^D = N_g^D$, $M_g^C = N_g^D N^C/N^D$, the treated share is the same in every stack, and $\hat{\beta}$ estimates the average effect over treated students.

Wing et al. (2024) show that the saturated stacked event study *without* stack-specific fixed effects, estimated by OLS, averages the untreated trends of treated students across stacks with weights N_g^D/N^D and those of control students with weights N_g^C/N^C, and is therefore biased by $\sum_g (N_g^D/N^D - N_g^C/N^C)\, E[\Delta Y_{ig}(0) \mid g]$ whenever treated shares and untreated trends both differ across stacks. In Eq. (2), the stack-by-year fixed effects absorb $E[\Delta Y_{ig}(0) \mid g]$ within each stack before the stacks are combined: the stack indicator lies in the column space of $\mathbf{Z}$, so $\sum_{i \in g} w_i \tilde{D}_{ig} = 0$ for every g, and the cross-stack terms in their decomposition vanish identically. Constant treated shares are needed only to make ω_g coincide with the treated-share target, not to make the estimate a convex combination of the stack-level effects.

How much the stack weights can matter. For any two weight vectors, $|\theta_\omega - \theta_\pi| = |\sum_g (\omega_g - \pi_g)\, \mathrm{ATT}_g| \leq \mathrm{TV}(\omega, \pi)\, (\max_g \mathrm{ATT}_g - \min_g \mathrm{ATT}_g)$, where $\theta_\pi = \sum_g \pi_g \mathrm{ATT}_g$ and $\mathrm{TV}(\omega, \pi) = \frac{1}{2} \sum_g |\omega_g - \pi_g|$ is the total variation distance between the two weight vectors, which depends only on the design. Appendix Table D1 reports ω_g for the main specification and the treated shares N_g^D/N^D.

Under stack-wise not-yet-treated controls, the treated share rises from 10% in the 2006 stack to 54% in the 2012 stack within 5 km, so treated shares are far from constant, yet the total variation distance is only 0.06 to 0.10 across radii: even if the effect differed by 2 percentile points across opening years, the two estimands would differ by at most 0.2 percentile points.

Covariates. With $\mathbf{X}_i'\gamma$ added, $\tilde{D}_{ig}$ becomes $(D_{ig} - p_g)(P_{ig} - q) - \tilde{\mathbf{X}}_i'\hat{\phi}$, where $\tilde{\mathbf{X}}_i$ are the covariates residualised on $\mathbf{Z}$ and $\hat{\phi}$ is the coefficient from regressing the first term on $\tilde{\mathbf{X}}_i$, whose cross-moment under balance is $\sum_g S_g\, \Delta\bar{\mathbf{X}}_g$, the S_g-weighted sum

of the difference-in-differences in mean covariates within each stack. The identity $\hat{\beta} = \sum_g \hat{\omega}_g \hat{\beta}_g$ and the properties of $\hat{\omega}_g$ are unchanged, while $\hat{\beta}_g$ and the closed form in Eq. (D2) are perturbed by terms in $\hat{\phi}$, that is, by the extent to which the composition of treated relative to control students changes at the opening. Section 4.2 shows that it changes little, and the estimates with and without covariates are nearly identical, 1.40 against 1.48 percentile points within 2 km (Appendix Table A14).

Appendix Table D1. Stack Weights of the OLS Stacked Regression, by Opening Year

	≤1 km		≤2 km		≤5 km		≤10 km	
Opening year g	ω_g	N_g^D/N^D	ω_g	N_g^D/N^D	ω_g	N_g^D/N^D	ω_g	N_g^D/N^D
2006	0.176	0.149	0.171	0.141	0.132	0.103	0.160	0.131
2007	0.163	0.139	0.170	0.144	0.188	0.160	0.227	0.224
2008	0.119	0.100	0.140	0.119	0.153	0.131	0.153	0.139
2009	0.119	0.104	0.121	0.103	0.117	0.099	0.139	0.129
2010	0.115	0.107	0.125	0.122	0.158	0.168	0.126	0.138
2011	0.165	0.189	0.153	0.185	0.153	0.187	0.125	0.159
2012	0.144	0.211	0.120	0.186	0.099	0.152	0.070	0.081
Treated share p_g, range	0.11–0.49		0.11–0.53		0.10–0.54		0.16–0.46	
Total variation distance	0.092		0.098		0.097		0.056	

Notes: This table reports, for each opening year g and radius, the stack weight ω_g that the OLS stacked regression with stack-wise not-yet-treated controls places on the comparison within stack g (Eq. (D2)), and the share of treated students in stack g, N_g^D/N^D, which is the stack weight under the corrective observation weights of Wing et al. (2024). p_g is the share of treated students in stack g. The total variation distance is $\frac{1}{2}\sum_g |\omega_g - N_g^D/N^D|$, computed from the LSOA panel of Section 2.3 with each LSOA weighted by its population as a proxy for the number of students.
Source: Census 2011 population by LSOA.