Essays on school choice and child development

Gennaro Rossi

Presented in fulfilment of the requirements
for the degree of Doctor of Philosophy

Department of Economics
University of Strathclyde
September 2022

Declaration

Acknowledgements

First of all, I am grateful to my supervisors and co-authors Markus Gehrsitz and Stuart McIntyre for their support over the last four years. I am deeply indebted to them for going through my drafts and for contributing to my personal and professional development, especially during the years of lockdown, when we would meet weekly.

I also thank my other co-author Daniel Borbely for being my *de facto* third supervisor, for his continuous feedback on research ideas and for having transmitted me so much knowledge, in terms of coding as well as econometrics. I sincerely hope that these were just the first few of many collaborations.

I am extremely grateful to Andy Dickerson, for his help and mentorship during a crucial time of my career. A special thanks goes also to Ellen Greaves, for providing extensive feedback on my work ahead of ESPE 2022 and for the many chats we had on the Calabrian roads in those few days of conferences. I also thank Angus Holford for his feedback.

I would also like to thank my PhD colleagues and friends Ali, Ben, Chirsty, Orion, Piero and Ross for the conversations and shared experiences. I am also grateful to all members of the Strathclyde Economics Department for their help, support, and encouragement.

Moreover, special thanks are due to my parents Andreina and Pasquale and my siblings Giovanbattista, Antonio and Rita for their continuous love and support. I am also grateful to my partner Sara for the endless encouragement and for putting up with my evenings and weekends at work, and to my son Noam, for unintentionally pushing me to be more time efficient. I also thank my friend Andrea for our constant exchanges about politics, economics and football.

Finally, I am also grateful to Mick Wilson and his team at the Scottish Government, including Gary Sutton and Beverley Ferguson for their huge help with my several queries. I also thank the Registers of Scotland for providing the data used in this work, as well as to Andrew McHugh and Heather Sinclair at the Urban Big Data Centre at the University of Glasgow and to Julian Augley, Fiona James and Suhail Iqbal from the eDRIS Team (Public Health Scotland) for their involvement in obtaining approvals, provisioning and linking data and the use of the secure analytical platform within the National Safe Haven. I also thank Nick Groome from Ordnance Survey and James Reid from EDINA Research and GeoData. Complete data acknowledgements follow.

This thesis uses data from the Registers of Scotland provided through the Urban Big Data Centre at the University of Glasgow. (1) © Crown copyright. Material is reproduced with the permission of the Keeper of the Registers of Scotland. (2) Registers of Scotland. Economic and Social Research Council. Registers of Scotland All Sales Data, 2019 [data collection]. University of Glasgow - Urban Big Data Centre.

This thesis uses data that were shared by the Scottish Government via Data Sharing Agreement No. EDA2001. The second chapter of this thesis was supported by the Nuffield Foundation through grant EDO/43743.

This thesis uses Ordnance Survey © Crown Copyright and/or database right 2019 OS. Ordnance Survey © Crown Copyright 2019. All rights reserved. Licence number 100034829. This data contains data created and maintained by the Scottish Local Government.

This thesis also uses Scotland Census 2011 © Crown Copyright, as well as data on Scottish Index of Multiple Deprivation, © Crown copyright 2016, supplied by National Records of Scotland.

This thesis uses Improvement Service data, which is © Crown copyright and database rights 2019 Improvement Service (licence number 100050367).

Additional education data were provided by the Scottish Government, some of which as part of the pre-existing Freedom of Information request 201900010285

Statement of Co-Authorship

While **Chapter 1** is single-authored, **Chapter 2** and **Chapter 3** are joint work. **Chapter 2** is co-authored with my colleagues Daniel Borbely, Markus Gehrsitz, Stuart McIntyre and Graeme Roy, and is available as an IZA Discussion Paper (No. 14678). **Chapter 3** is co-authored with my colleagues Daniel Borbely, Markus Gehrsitz and Stuart McIntyre, and is available as a Strathclyde Working Paper (No. 22- 5). The undersigned hereby certify that:

1. They meet the criteria for co-authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the working paper in their field of expertise;
2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the working paper;
3. There are no other authors of the publication according to these criteria;
4. There are no potential conflicts of interest that require to be disclosed to (a) granting bodies, (b) the editor or publisher of any journal or other publication.

In the case of **Chapter 2** contributions to the work involved the following:

Contributor	***Main areas of contribution***	***Publication title, date of publication or status***
Daniel Borbely [redacted]	Co-writing, Methodology, Coding, Estimation, Data Curation	Borbely, D., Gehrsitz, M., McIntyre, S., Rossi, G. and Roy, G. (2021) Early-Years Multi-Grade Classes and Pupil Attainment. *IZA Discussion Paper No. 14678* https://www.iza.org/publications/dp/14678/early-years-multi-grade-classes-and-pupil-attainment
Markus Gehrsitz [redacted]	Co-writing, Methodology, Coding, Estimation, Data Curation	
Stuart McIntyre [redacted]	Co-writing, Methodology, Coding, Estimation	
Gennaro Rossi [redacted]	Co-writing, Methodology, Coding, Estimation, Data Curation	
Graeme Roy [redacted]	Co-writing, Methodology, Coding, Estimation, Data Curation	

In the case of **Chapter 3**, contributions to the work involved the following:

Contributor	***Main areas of contribution***	***Publication title, date of publication or status***
Daniel Borbely [redacted]	Co-writing, Methodology, Coding, Estimation, Data Curation	Borbely, D., Gehrsitz, M., McIntyre, S. and Rossi, G. (2021) Does the Provision of Universal Free School Meals Improve School Attendance and Behaviour?. *Strathclyde Discussion Papers Economics No. 22 - 05.* https://www.strath.ac.uk/media/1newwebsite/departmentsubject/economics/research/researchdiscussionpapers/22-05_Free_School_Meals_Manuscript_Strathclyde_DP.pdf
Markus Gehrsitz [redacted]	Co-writing, Methodology, Coding, Estimation, Data Curation	
Stuart McIntyre [redacted]	Co-writing, Methodology, Estimation	
Gennaro Rossi [redacted]	Co-writing, Methodology, Coding, Estimation, Data Curation	

Synopsis

This thesis is a collection of three distinct empirical essays on school choice and child development. We employ applied econometric techniques to estimate the causal effect of two aspects of the Scottish educational system on pupils' behaviour and academic attainment: multi-grade classes and the expansion of free school meals' provision. In addition, we examine which secondary schools' characteristics drive households' residential choices. These three studies contribute to an extensive literature as well as to a series of heated debates.

Chapter 1 is titled "*School Performance, Non-Cognitive Skills and House Prices.*" This chapter is a single-authored study of the drivers of secondary school choice in Scotland. In many educational systems, children attend their closest school, with little room for choice. This implies that households decide where to live, at least in part, to secure their children a place at their preferred school. Advocates of school choice motivate its importance in light of improved efficiency, but also better matching effect for single students. Conversely, other experts believe that letting parents decide which school their children should attend may create distortions whereby schools improve under aspects which are preferred by parents, i.e. a more favourable selection, as opposed to improved effectiveness. A wealth of studies shows the complexity of school choice, which lends itself to strategic behaviour. For this reason, studying the drivers of school choice is imperative.

Most of the literature to date focuses on school-level average test scores or socio-economic composition as a factor determining school choice. This, however, might reveal very little about the inherent quality of a school. Another strand of the literature focused on measures of value-added (or effectiveness). To date, the evidence points towards parents not caring about objective measures of effectiveness, but only about peers' performance, even if this is mainly driven by selection. One characteristics of these studies, however, is that they focus mostly on measure of value-added which are unknown to parents. Therefore, the question arises whether this lack of an effect is driven by a lack of information, rather than a lack of interest. Furthermore, most of the studies to date focus on a limited number of school performance indicators, thus overlooking multidimensionality of school outputs and whether parents care about these. Finally, despite the extensive evidence on the effects of spillovers in peers' non-cognitive skills, little is known on whether this dimension of school quality is valued by parents.

In this study I examine secondary school choice by estimating house price capitalisation for an array of school characteristics. To overcome the endogeneity of school 'quality' I leverage Scotland's strict residence-based attendance system and employ a boundary regression discontinuity design, whereby I compare prices of property which are located in the same neighbourhood but on opposite sides of catchment area boundaries. This study contributes to the relevant literature in a number of ways. First, it uses a vast array of school outcomes, including post-school destinations, as potential drivers of school choice; second, it attempts to overcome the debate in the literature about the lack of interest in school value-added by using a highly salient and easy-to-understand measure of value-added; third, it uses peers' non-cognitive skills' measures as a potential driver of school choice; fourth, it provides nation-wide evidence.

I find that peers' academic performance and socio-economic composition command a house price premium of about 4% for a one-standard-deviation increase. On the other hand, school value-added and non-cognitive skills do not appear to be capitalised into house prices.

Chapter 2 is titled "*Early-Years Multi-Grade Classes and Pupils Attainment*". This chapter is a co-authored study of the educational gains from sharing the classroom with older, more experienced peers, and focuses on Scotland. There is an extensive literature studying different dimensions of peer effects, e.g. gender, ability, socio-economic background. On the other hand, evidence on age peer effect is relatively scarce. This is despite multi-grade (composite, henceforth) classes, i.e. the practice of grouping pupils from adjacent year groups, being commonly employed in many educational contexts.

In this paper we examine whether primary school first-graders experience cognitive gains by sharing the classroom with older peers (second-graders). We match data from the Scottish Pupil Census to academic attainment data obtained from Curriculum for Excellence records, i.e. teacher-assessed attainment in literacy, numeracy, reading, listening and speaking. This empirical exercise carries several methodological challenges. The most relevant is given by the fact that neither are composite classes randomly generated, nor are pupils randomly assigned to them. For instance, selection into composite classes can happen on the account of abilities or to preserve existing networks of friends. To disentangle this endogeneity issue, we make use of an algorithm employed by Scottish Local Authorities, which predicts the creation of composite classes both at the extensive and intensive margin, based on enrolment projections. The algorithm leverages class size caps set by the central government, and it predicts the creation of composite classes any time this enables a saving in the number of classes to be created, and therefore teachers to be appointed. We show how as little as one additional pupil can trigger the creation of at least one composite class across the entire school. Therefore, the creation of composite classes is, in some circumstances, as good as random. We use the prediction from this algorithm as an instrumental variable for the actual composite status.

This study contributes to an emerging, small, literature by adding to its external validity. The existing evidence predominantly pertains to rural areas, whereby small cohort sizes dictate the necessity to create fewer classes by pooling different grades. By contrast, composite classes are widespread in Scotland, in rural as well as urban areas.

We find that first-graders who share a classroom with second-graders experience educational gains of about 0.2-0.3 of a standard deviation in literacy and numeracy, in line with previous studies. We show that the effects are not driven by gender, socio-economic or urban/rural status. Finally, we do not find evidence of a detrimental effect on older peers (second-graders).

Chapter 3 is titled *"Does the Provision of Free School Meals Improve School Attendance and Behaviour?"* and is also a co-authored work. This chapter is an evaluation of the expansion of free school meals' provision which took place in Scotland in January 2015. Free school meals (FSM, hereafter) provision is the subject of a long-standing debate. In many contexts, FSM are a means-tested benefits, meaning that their eligibility is contingent to households financial circumstances. Advocates of FSM expansion argue that FSM improve attendance, academic performance, behaviour, diet as well as uptake among previously eligible pupils by reducing stigma. Opponents, on the other hand, point to the high costs associated with such policies, as well as noting the potential detrimental effect on diets. Scientific evidence largely supports the claims of the advocates of such policies. In addition, some studies find that FSM provision leads to improvement in households' finances. Apart from these studies, most of the evidence is confined to short-term outcomes such as test-scores and overlooks the effects on non-cognitive components, which are equally likely to drive future attainment.

In this study we estimate the impact of FSM provisions' expansion on students' attendance, health-related absences, and misbehaviour. We do so by leveraging a change in policy which extended FSM eligibility to all students in the first, second and third grades of primary school (P1 to P3). We estimate a difference-in-differences (DiD) model with variation in treatment intensity. This is determined by the pre-policy share of FSM-takers. Therefore, schools with relatively fewer pupils taking FSM one year before the policy are indeed more exposed to it. The importance of our study can be assessed from a twofold perspective. First, Scotland, in line with UK, reports among the highest rates of child obesity worldwide. Second, outcomes such as attendance and misbehaviour are correlated with the 'Big Five' personality traits and therefore non-cognitive skills.

Overall, we find precisely estimated null effects of the policy on all dimensions we consider. The absence of effects is not driven by school characteristics which could have plausibly affected the success of the policy, i.e. deprivation, rurality and school resources.

Contents

List of Figures

List of Tables

Chapter 1

SCHOOL PERFORMANCE, NON-COGNITIVE SKILLS AND HOUSE PRICES

1.1 Introduction

Parents choose where to live, at least in part, in order to secure a place at the local school for their children. The main argument in favour of school choice relies on its impact on school productivity (Hoxby, 2003) but also on the possibility that it allows pupils to choose the path that best suits their needs (Hoxby, 2000b), thus improving their attainment (see e.g. Lavy, 2010; Deming et al., 2014). However, school choice can be a complex matter, which lends itself to strategic behaviour (see e.g. Abdulkadiroğlu and Sönmez, 2003; Burgess et al., 2015; Calsamiglia and Güell, 2018; Calsamiglia et al., 2020). Therefore, examining drivers of school choice is paramount. Not only does it provide insights on the extent to which parents are knowledgeable about their children's learning experience, it might also shed light on school performance-enhancing policy mechanisms.[1] For instance, if choice is driven by school-level average outcomes, schools might attempt to 'improve' pupil selection rather than effectiveness (MacLeod and Urquiola, 2015; Abdulkadiroğlu et al., 2020). Therefore, this offers clear scope for policy intervention aiming to provide parents with tailored guidance in the choice of schools that best suits their children, as well as to advise schools on how to improve their performance.

The aim of this paper is to assess which secondary school characteristics parents seek the most, by appraising the capitalisation of school attributes into house prices, looking at Scotland as a case study. I mostly focus on two measures of school-level performance, i.e. average academic achievement and value-added. In the remainder of this paper I will refer to school-level average attainment as peers' (academic) performance, and interchangeably employ the terms 'value-added' and 'effectiveness' to identify the schools' contribution towards pupils' academic progress.

Whilst I find evidence of neighbourhood sorting plausibly based on school choice (Epple and Romano, 2003; Bayer et al., 2007; Greaves and Turon, 2021), I manage to isolate the effect of school quality on house prices by focusing on instances whereby catchment area boundaries separate *de facto* the same neighbourhood, generating a credible, quasi-random source of variation in school quality. I compare

[1]There is a wealth of studies of the effects of gaining (or missing) a place in a preferred, or high quality, school. See for instance Cullen et al. (2006), Clark (2010), Kirabo Jackson (2010), Deming (2011), Pop-Eleches and Urquiola (2013), Abdulkadiroğlu et al. (2014), Lucas and Mbiti (2014), Dee and Lan (2015), Clark and Del Bono (2016), Abdulkadroğlu et al. (2017), Dustan et al. (2017), Hoekstra et al. (2018), Gorman and Walker (2021), Ovidi (2021) and Beuermann and Jackson (2022).

houses within census blocks, therefore sharing the same observables, yet in different catchment areas. Therefore, price divergence across boundaries is causally determined by changes in school quality.[2] I link residential property transactions to local schools and use a rich set of school-level secondary data provided by the Scottish Government, alongside neighbourhood characteristics as controls. Given the large number of variables at hand, I use factor analysis and hence examine whether school quality is multidimensional.

I find that peers' academic performance and socio-economic composition command a house price premium of about 4% for a one-standard-deviation increase. On the other hand, school value-added and non-cognitive skills do not appear to be capitalised into house prices.

Long-standing evidence from various contexts suggests that for a one-standard-deviation increase in school average test scores, parents pay a house price premium of about 3-4% on average. Research in this area normally proxies school performance using high and low-stakes test scores (see, for example Black, 1999; Gibbons and Machin, 2003, 2006; Bayer et al., 2007; Fack and Grenet, 2010; Beuermann et al., 2018) or school socio-economic composition (Clapp et al., 2008; Fack and Grenet, 2010).[3] Another extensive, and in part more recent, strand of the literature focuses on measures of school effectiveness (Brasington and Haurin, 2006; Hastings et al., 2009; Gibbons et al., 2013; Imberman and Lovenheim, 2016; Beuermann et al., 2018; Abdulkadiroğlu et al., 2020). One common finding is that parents still choose schools based on their average peers' performance rather than value-added (see e.g. Brasington and Haurin, 2006; Abdulkadiroğlu et al., 2020). Part of the reason might be a lack of information on school effectiveness (Ainsworth et al., 2020). Imberman and Lovenheim (2016) exploit a set of school value-added statistics which were newly promulgated by newspapers. Despite the fact that parents response to newly released school information can be quick and sizeable (Figlio and Lucas, 2004; Hastings and Weinstein, 2008; Fiva and Kirkebøen, 2011; Greaves et al., 2021), Imberman and Lovenheim (2016) find no effect of value-added on house prices. The question therefore remains whether parents

[2] A strand of research in this area tackles this issue by exploiting reforms of the attendance system or re-drawing of catchment areas (see for example Bogart and Cromwell, 2000; Ries and Somerville, 2010; Machin and Salvanes, 2016). A similar approach is adopted by Calsamiglia and Güell (2018), yet in a different context.

[3] Further research in this area includes Kane et al. (2003), Reback (2005), Davidoff and Leigh (2008), Chung (2015), Lee (2015), Agarwal et al. (2016), Orford (2018), Harjunen et al. (2018), Bonilla-Mejía et al. (2020), Chan et al. (2020), Batson (2021), Han et al. (2021), Park et al. (2021) Wong and Deng (2021), Cheung et al. (2022) Ismail et al. (2022), Zheng (2022) and Zou (2022).

value school effectiveness at all.

This project carries two main contribution. First, it contributes to the school value-added literature by using a highly salient measure of value-added. I isolate peers' performance from school effectiveness by means of the 'virtual comparator', an index developed by the Scottish Government which compares actual school performance with attainment predicted by the demographic composition of the school. The virtual comparator is publicly available and is reported alongside actual school performance. This means that parents have a good idea of how well a school is performing with respect to a counterfactual. This is an important difference between my analysis and the existing literature in which school value-added is estimated by the researcher (Brasington and Haurin, 2006; Beuermann et al., 2018; Abdulkadiroğlu et al., 2020) – despite the internal validity of their estimates, it is uncertain whether parents are in fact aware of school effectiveness. This paper adds to this strand of the literature in which value-added information is widely available to parents (see for example Gibbons et al., 2013 and Imberman and Lovenheim, 2016), and it extends it in two significant ways. Firstly, it does so by using a more contextualised measure of value-added which takes into account pupils' characteristics which are correlates of their progress; secondly, by examining the capitalisation of value-added over a longer post-release time span.[4] One limitation of this measure of value-added is that is based on a statistical matching procedure on observables, and ignores unobservables-based sorting and match-specific effects. While this is undoubtedly a disadvantage with respect to previous work which were able to identify the causal effect of schools, I argue that this matching procedure still takes into account an extensive number of predictors of attainment. In addition, for the purpose of this analysis this measure should serve well as a signal of school quality to parents, after accounting for schools' intakes.

The literature on school choice finds that parents value peers effects above school effectiveness. The body of evidence on peer effects in primary and secondary education argues that while the presence of better peers improve one's outcomes (see for instance Sacerdote (2011) for a detailed review) such effects are highly non-linear, i.e. the extent to which one benefits from high-ability peers depends both on one's own position in the ability distribution and the position of one's peers. The evidence is rather mixed. For instance, Burke and Sass (2013) find that low-ability pupils are harmed by the presence of

[4] Imberman and Lovenheim (2016)'s main focus is on seven months post-release.

high-ability peers, while middle-ability pupils suffer (benefit) from the presence of low-ability (high-ability) classmates. Finally, the authors suggest that high-ability pupils benefit from the presence of lower-ability peers. In contrast, Lavy et al. (2012b) find that students at the bottom (5%) of the ability distribution impact negatively on their high-ability classmates, and that top (5%) students are beneficial mostly for girls (especially if low- middle-ability ones) while boys suffer the presence of good peers. In addition, Lavy et al. (2012a) argue that low-ability students have a detrimental effect on attainment of students across the entire distribution, but this effect is larger for low-ability students. One of the suggested mechanism is a lack of interaction between students of different types, with low-ability students (especially boys) suffering the presence of strong peers, with consequent reduction in motivation and self-esteem. Another channel could be teachers, whose focus can either be diverted from the top-middle to the bottom of the pupils' ability distribution, or perhaps they could target the top pupils, leaving the rest of the class behind. A similar mechanism is also identified by Duflo et al. (2011) who advocate that homogenous groups resulting from tracking can make teaching more effective. The heterogeneity of these findings can be reconciled under one policy recommendation: students tend to do better when grouped with peers of similar characteristics. Such a conclusion was previously reached by Hoxby and Weingarth (2005) and recently –yet in part– confirmed by (Booij et al., 2017), who focus on tutorial groups in higher education.[5]

For this reason, parents need to have a reliable prior on their own children's ability and how this fits within the ability distribution of the school they want. However, parents' perception of their children's relative ability might be flawed, leading to the wrong human capital investment (Kinsler and Pavan, 2021). While peers' performance and value-added can be the result of concurring factors –i.e. sorting, resources, teachers– the virtual comparator could potentially provide information on the counterfactual distribution of abilities within school, thus guiding parents' choice. One scenario is that by observing a positive, large gap between peers' performance and the virtual comparator (large value-added) parents' of low-ability children might opt for such a school, rather than simply seeking schools with high peers' performance.

[5]For instance, there might be no benefits for low-ability students from sharing a classroom with high-ability peers unless the classroom is heterogeneous enough (Booij et al., 2017).

However, school value-added plausibly increases non-linearly relative to pupils ability. There are a number of mechanism for this. First, there might be more room for improvement at the bottom of the ability distribution. Second, non-linear peer effects means that some students might experience additional educational gains thanks to an *ad hoc* classroom composition. Third, value-added might also be the result of better resources, i.e. funding, better teachers. However, if high-achieving students sort into schools with better teachers, whom in turn might be better at leveraging non-linear peer effects, then the importance of peers' quality in the value-added production function attenuates. Hence, it might be difficult for parents to isolate each different mechanisms driving value-added.

If parents valuation of school-level peers' achievement is linear, then parents of low- or middle-ability students will over-demand schools with strong peers. This will likely have adverse consequences in presence of non-linear peer effects. For instance, grouping low-ability students with high-achieving peers might lead to outcomes according to an 'invidious comparison model' (Hoxby and Weingarth, 2005). Alternatively, parents might use peers' performance (actual or relative to virtual comparator) to infer other school traits, i.e. resources and teachers. If this were the case, such over-demand might even disappear once school resources and teachers are accounted for. However, teaching quality is hardly observable to parents, and it might also have little independent variation relative to peers' performance. Therefore, to parents, peers' performance/value added might be a better signal of school "quality".

The second contribution of this paper is within the non-cognitive skills space. While some recent work explores school outcomes beyond academic scores, parental preferences for non-cognitive skills have been mostly overlooked.[6] Non-cognitive skills are soundly proved to be important determinants of various outcomes.[7] A very important feature of non-cognitive skills inheres to the long-lasting effects of early-years interventions (Chetty et al., 2011; Del Bono et al., 2016; Fort et al., 2020) but also to their malleability later on in life (Heckman and Mosso, 2014; Norris, 2020). For instance, Norris (2020) finds evidence of peer effects in learning attitudes in adolescence, and that parents' expectations can meliorate peers' influence. By virtue of its focus on secondary schools, my study builds on this evidence

[6] Beuermann et al. (2018) look at parental choice in terms of, among other things, school-level incarceration rate, teen births and future employment.

[7] See Kautz et al. (2014) for a review.

and closely relates to the very recent work on non-cognitive peer effects in secondary education. Shure (2021) documents positive and sizeable effects of peers' non-cognitive skills on individuals academic attainment. Similarly, Dickerson et al. (2018) show that academic aims and abilities affect secondary school peers' aspirations. I proxy non-cognitive skills by means of behavioural indicators, strongly correlated with the 'Big Five' (Lleras, 2008; Bertrand and Pan, 2013; Jackson, 2018), at the secondary school level, and examine whether potential spillovers in non-cognitive skills play a role in secondary school choice. To the best of my knowledge, this is the first work attempting to identify and quantify parents preferences for peers' non-cognitive skills, in particular in the context of their capitalisation into house prices. In addition, to the best of my knowledge this is to date the only work of this sort in Scotland and one of the few works in this field focusing on a nation-wide context, thus providing a contribution in terms of external validity.[8]

Finally, this work broadly contributes towards the empirical literature studying school demand via assignment mechanism (Harris and Larsen, 2015; Fack et al., 2019; Abdulkadiroğlu et al., 2017; Glazerman and Dotter, 2017; Kapor et al., 2020; Agarwal and Somaini, 2018) as well as the link between school performance and neighbourhood characteristics (Battistin and Neri, 2017; Neri et al., 2020; Greaves, 2022). Overall, my findings suggest a lack of multidimensionality in school outcomes and that parents seem to value schools with high-performing pupils. This preference persists also conditionally on other indicators of school performance.

The remainder of the paper is structured as follows. In the next section I outline the Scottish school system and its residence-based attendance system. In Section 1.3 and Section 1.4 I describe the data and the identification strategy, providing parametric evidence of its validity. In Section 1.5 I present and discuss the results. The final section concludes.

[8] see for example Gibbons and Machin (2003), Gibbons et al. (2013) or Burgess et al. (2019)

1.2 Institutional Background

The Scottish education system is predominantly composed of state-funded, non-religious schools, which constitute the primary focus of this paper.[9] Private (or independent) schools are mostly clustered in the *Central Belt* of the country, and according to the Scottish Council of Independent Schools (SCIS) census, in September 2015 they enrolled approximately 30,238 pupils, which amounts to 4.3% of the total school age population.[10] For secondary schools, the figure is 18,159, which is roughly 6% of the corresponding sub-population. These schools often follow a curriculum which is different from the national one, e.g. International Baccalaureate or the English A-levels. Religious schools, which account for roughly 15% of the total number of pupils in the country, are part of the state-funded school sphere and as such follow the national curriculum, but have different catchment areas.

Scottish pupils typically access their first year of primary school between their fifth and sixth year of age. Primary schools include seven stages (or grades), i.e. P1-P7, after which pupils move to the first of the six stages of secondary school, namely S1-S6. In 2010 a new national curriculum, the *'Curriculum for Excellence'*, is implemented and it consists of five levels: Early, First and Second levels take place within the seven years of primary school; Third/Fourth, during the first three years of secondary school (S1-S3) and finally the last one, Senior Phase, from S4 to S6. As pupils turn sixteen, usually in S4, they can either leave to pursue further education or go into employment (or unemployment), or progress to S5 (and perhaps later to S6) to attain further qualifications within the Scottish Qualification Authority (SQA), i.e. the national exam board.

Each of these qualifications has an equivalent level in terms of Scottish Credit and Qualification Framework (SCQF). This is a scheme which aims to facilitate credit transfer by organising degrees and qualifications in twelve levels, 12th being the most challenging one and corresponding to a PhD or a professional apprenticeship. Table A.2 outlines the possible qualifications that can be attained during 'Senior Phase' (S4-S6) alongside their corresponding SCQF levels. In particular, up to S4, pupils can attain a number of 'National 5' qualifications, corresponding to SCQF level 5. For example, if a pupil passes tests in Mathematics and English at the end of S4, she will be awarded two 'National

[9] These schools are also known as 'non-denominational'.
[10] http://www.scis.org.uk/facts-and-figures/

5' qualifications, or similarly, two awards at SCQF level 5. The decision to progress further to S5 is contingent on a pupil's intention to achieve further qualifications. Within the SQA classification, these are typically 'Higher' and 'Advanced Higher' qualifications, which correspond to SCQF levels 6 and 7 respectively and are roughly the equivalent of the A-levels.

As part of the admission criteria, Universities in Scotland (and rest of the UK) require a certain number of 'Higher' (or above) to be attained with certain grades and within specific subjects, depending on the course and institution the pupil is applying to. As a result, we can consider 'Higher' and 'Advanced Higher' as high-stakes qualifications. One important caveat is that Scottish Government's official statistics and data refer to the number of qualifications in terms of SCQF levels, rather than in SQA nomenclature. This is a detail I come back to later, in Section 1.3.

To gauge numbers, between school year 2013/2014 and 2015/2016 roughly 39% of pupils attained four or more SCQF level 6 awards, about 36% of leavers went to higher education while 66% obtained at least four awards at SCQF level 5 (see Table A.1 for summary statistics of the variables used in this work).

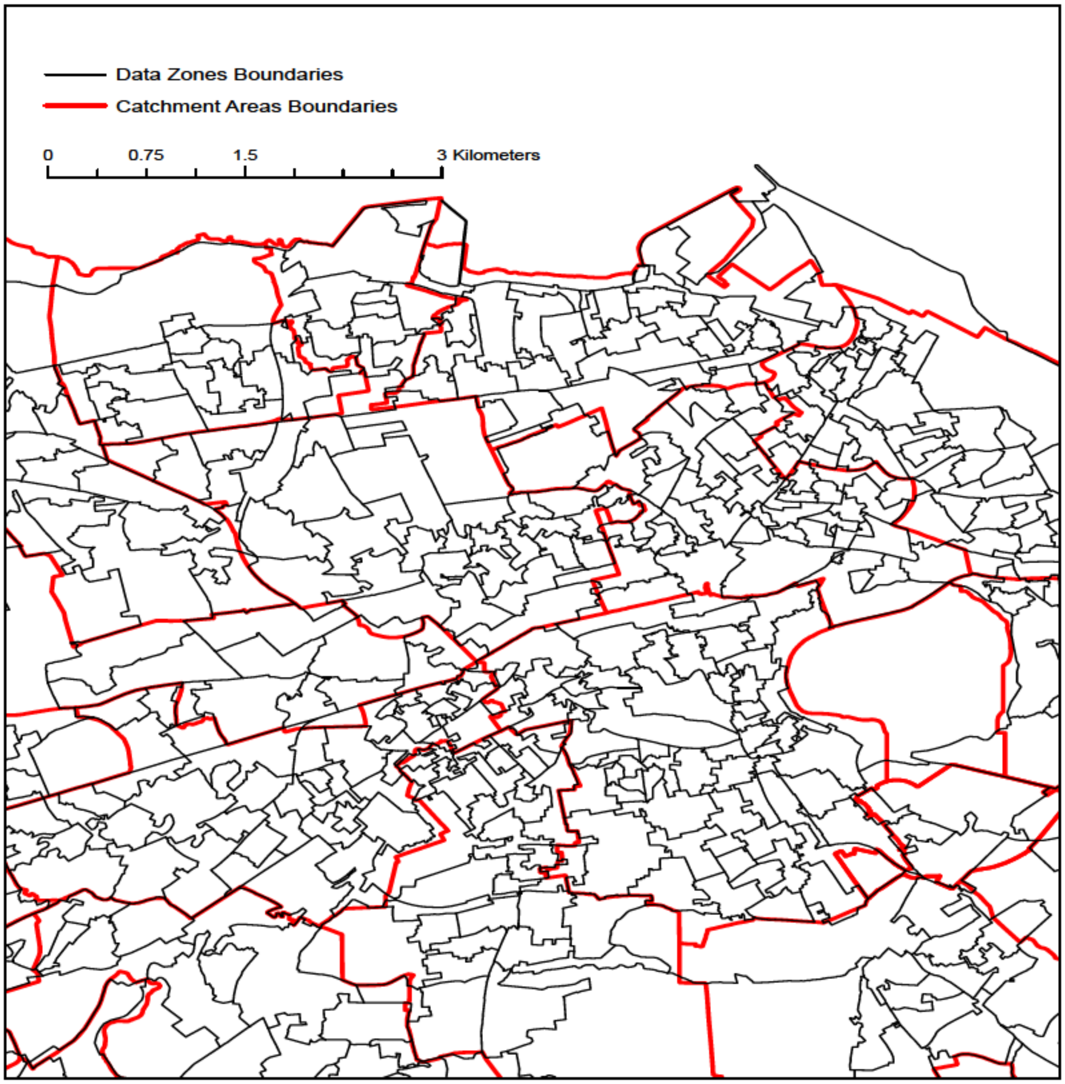
It is worth mentioning that in 2007 the Scottish Government waived university tuition fees for Scottish Students. In addition, these were largely unaffected by the higher education fees and financing changes introduced by the 2010 Higher Education Act (see Hassani-Nezhad et al., 2021).

Schools with higher performance in high-stakes tests might not necessarily be what parents want for their children, and they might instead opt for schools which are stronger at preparing their children for different paths such as employment after school, or further education via a college or vocational training provider. Hence, I also look at SCQF level 5 qualifications as these constitute the highest qualification prior to attaining school-leaving age.

Each of these qualifications can be achieved within a fairly large range of subjects alongside English and Mathematics. In order to frame attainment within literacy and numeracy skills, SCQF levels in literacy and numeracy are awarded to pupils who achieve National 5, Higher and above in specific subjects or course/units. For example, a pupil will achieve SCQF level 6 in literacy if she obtains a Higher qualification in all five unit groups for English.

As per school choice, Local Authorities are usually divided into catchment areas.[11] Figure 1.1 provides evidence on the size of catchment areas, taking Edinburgh City as an example. For non-

Figure 1.1: Catchment Areas and Data Zones



Note: This figure shows how data zone blocks nest within secondary schools catchment areas. In particular, the figure refers to central and northern areas in Edinburgh.

denominational schools, the main focus of this work, parents are advised to enrol their children in the designated local school by the Local Authority, which is required to enrol kids from the catchment area. However, parents can submit a 'placing request' for a school different from the designated one (Learning Directorate, 2016). Councils prioritise children living in the catchment area, and are not obliged to

[11]Scottish Local Authorities are roughly equivalent to US counties.

grant a placing request by parents residing outside its boundaries. In fact, they are more likely to grant such a request where this would not alter the pupils/teacher allocation, e.g. if the school needed to hire an additional teacher or set up a new classroom; or there were reasons to believe the placing request would not constitute a good fit, either for the children already in school or for the pupil whose parents have submitted the request.[12] Successful placing requests are rare and often pertain to children with additional support needs or who have sibling(s) already in the school. Moreover, Borbely et al. (2021) argue that catchment areas overlapping is rare. For instance, in 2017/18 and 2018/19 only about 13% of families submitted a placing request, with approximately 80% acceptance rate (Bhattacharya, 2021).

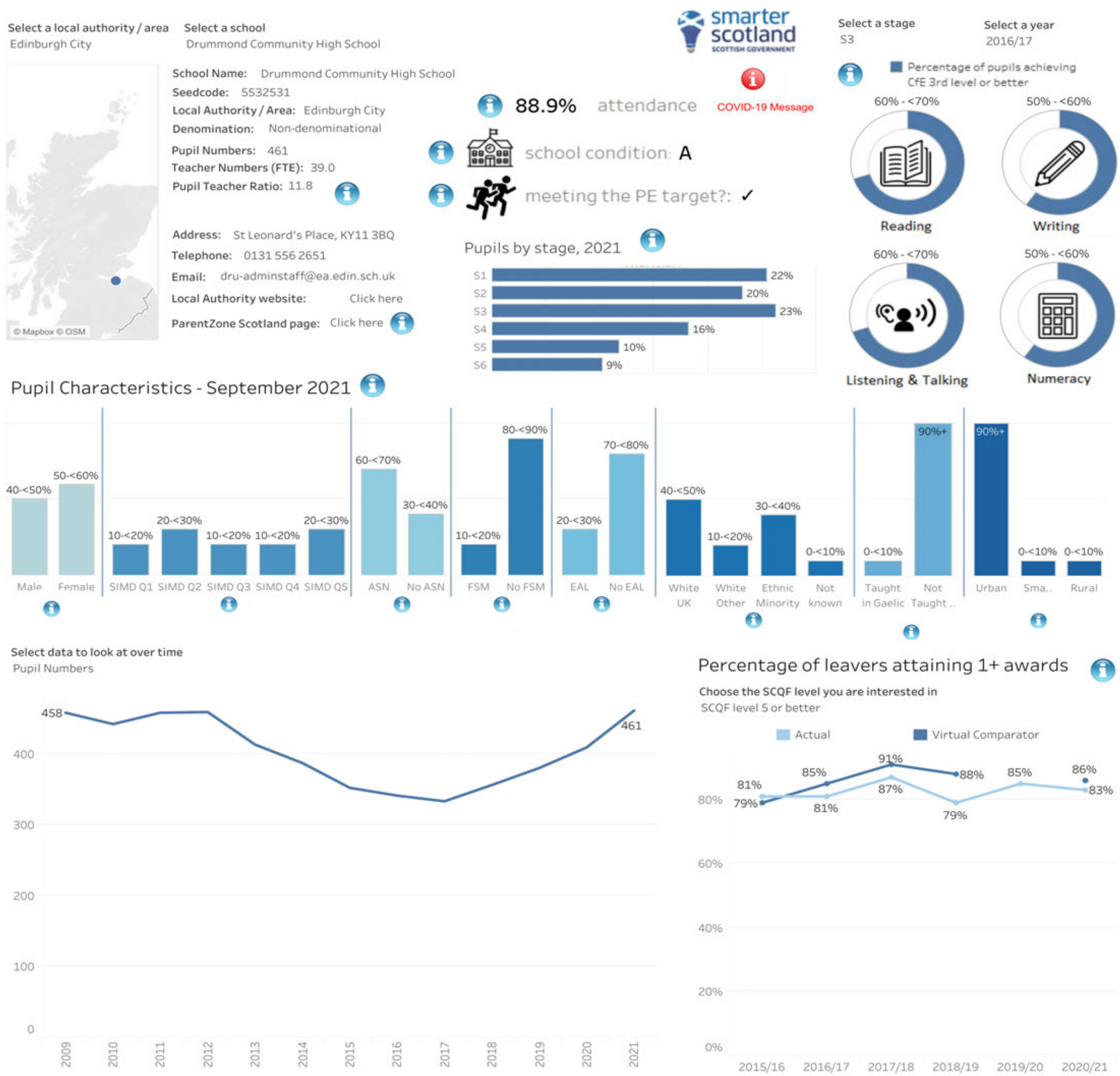
Denominational (mostly Roman-Catholic) schools' catchment areas are generally larger and stretch across those of non-denominational schools. They do not require pupils to be from a particular denomination or faith but intakes criteria vary by Local Authority.

Information plays a key role in school choice. Parents have access to a large number of school-level data. Figure 1.2 displays a sample of the interactive School Information Dashboard provided by the Scottish Government. While the current format was first published in August 2018, most of this school-level information, including the virtual comparator, was previously available in a variety of formats on the Scottish Government's website, as well as on Parentzone Scotland. This is a portal within the Education Scotland websites and it constitutes the main source of information for parents as it also contains links to specific schools' websites. From conversations with people at the Scottish Government, it emerged that the webpage received about 1,483,758 visualisations in 2015. Considered that parents would choose secondary schools when their children are in P7, whose entire cohort in state schools was about 52,000 pupils between school years 2014/15 and 2016/17, these figures suggest a reasonable engagement.

In the next section, I describe the data I use in this study and describe the school-level data which is accessible to parents.

[12] see, for more details, https://www.gov.scot/publications/choosing-school-guide-parents-nov-16/

Figure 1.2: Secondary School Information Dashboard



Note: Here is an example of secondary school information dashboard. This format was first released in August 2018, but previously the same information, yet in a different format, was available on *ParentZone Scotland* on the Education Scotland website.
Source: SG Education Analytical Services: Learning Analysis (2018).

1.3 Data

I use housing data provided by the Registers of Scotland via the Urban Big Data Centre at University of Glasgow.[13] The data contain detailed information about the sale price, full address, postcode as well as geographic coordinates for every residential property transaction that occurred in Scotland from January 2008 to April 2018. Missing coordinates are partly integrated by using the AddressBase® Plus dataset provided by Ordnance Survey.[14] In order to match each transactions to a specific school I employ catchment areas shapefiles covering all of Scotland, provided by the Improvement Service.[15] Schools' information is obtained from a publicly available database of the Learning Directorate at the Scottish Government. In particular, I collect school-level data on the number of pupils broken down by stage, percentage of pupils registered for free school meals, pupil-teacher ratio, information about the estates' capacity as well as location.

Attainment data are collected from the interactive school information dashboards elaborated by the Learning Directorate. They contain, among others, the following school-level variables: *i)* The percentage of leavers attaining at least four awards at SCQF level 6 (or above). *ii)* The percentage of leavers attaining at least four awards at SCQF level 5.[16] Regarding point *i)*, the goal is to use a measure of high-stakes attainment. A minimum of four 'Higher' qualifications is a standard admission criteria applied by many universities. However, observing the number of awards at SCQF level 6, as opposed to the number of 'Higher' qualifications, conceals the risk of dealing with a noisy measure of academic-oriented qualifications. In fact, a SCQF level 6 award could in principle correspond to a 'Higher' qualification as well as to a Scottish Vocational Qualification (SVQ) (see Table A.2). However, Figure A.1 shows that the school-level percentage of leavers achieving a minimum of four awards at SCQF level 6 is highly correlated with the percentage of leavers in higher education,[17] as well as with

[13] 1. © Crown copyright. Material is reproduced with the permission of the Keeper of the Registers of Scotland. 2. Registers of Scotland. Economic and Social Research Council. Registers of Scotland All Sales Data, 2019 [data collection]. University of Glasgow - Urban Big Data Centre.

[14] Ordnance Survey © Crown Copyright 2019. All rights reserved. Licence number 100034829

[15] © Crown copyright and database rights "2019" Improvement Service (licence number 100050367)

[16] Additional information, i.e. percentage of leavers in higher or further education, employment and S6 leavers is obtained from Freedom Of Information Request no. 201900010285

[17] After nine months from secondary school graduation.

the percentage of S6 leavers.[18] This shows that the 'four-award' threshold for SCQF level 6 captures variation in high-stakes attainment, alongside ensuring sample size gains.[19]

A school with many high-performing peers' is not necessarily a 'good' one in the sense of value-added, i.e. contribution to pupils' learning. This is why I also construct an indicator of school 'effectiveness' (or value-added) for each of the measures described in *i)* and *ii)* as well as for literacy and numeracy. These indicators consist of the difference between each of the actual measure of attainment at the school level and its virtual comparator. The latter is developed by the Learning Directorate at the Scottish Government. In order to 'adjust' school performance for its composition, a virtual comparator creates for each school a synthetic control. For each pupil in the school in question, a counterfactual is created by randomly selecting ten other pupils from other Local Authorities in Scotland, conditional on the same demographics for the 'treated' pupil such as gender, free school meal eligibility, additional support needs, school-leaving stage and Scottish Index of Multiple Deprivation (SIMD) of the area they live in.[20] The average attainment of these ten pupils constitutes the counterfactual score. The difference between the actual school performance and its virtual comparator should provide information about how well this school performed, given its composition. I assess whether these measures are capitalised into house prices.

Furthermore, I use the attendance and absence survey to glean information on within-school behaviour. The survey collects four main aggregates, namely attendance, authorised absences, unauthorised absences and exclusions (in terms of number of half-day openings, henceforth sessions) and breaks down the first three by type. For example, 'in school' attendances accounts for almost the entirety of overall attendance.[21] Authorised absences occur mostly on the account of illnesses, but also authorised family holidays and lateness, whereby a pupils would miss the majority of a session. Finally, unauthorised absences are largely driven by truancies.[22] I proxy for non-cognitive skills using school-

[18] Hence those leavers attaining two extra years of schooling.

[19] The corresponding school-level percentage is normally neither too high nor too low and thus its values are rarely suppressed for statistical disclosure control.

[20] This is an index developed by the Scottish Government and is meant to classify areas based on their degree of deprivation across seven domains, i.e. income, employment, education, health, housing, crime and access to services.

[21] Pupils doing an approved work experience or being off sick with educational provision also counts toward overall attendance, but to a far lesser extent.

[22] For more information, see https://www.gov.scot/publications/summary-statistics-schools-scotland-no-pages/7/.

level truancies, lateness, attendance and exclusions. In doing this I follow previous studies drawing from the Psychology literature (Lleras, 2008; Heckman and Kautz, 2012; Bertrand and Pan, 2013; Heckman et al., 2016; Jackson, 2018). In fact, the above behaviours are strong correlates of the 'Big Five' and their related traits. Agreeableness, conscientiousness, neuroticism are long-proved to be associated with absences, tardiness and anti-social behaviour (see for example, John et al., 1994; Barbaranelli et al., 2003; Carneiro et al., 2007; Duckworth et al., 2007; Lleras, 2008; Jackson, 2018). Table A.7 provides a brief description of school-level variables.

Finally, I use Census and Scottish Index of Multiple Deprivation (SIMD) variables, both at the data zone level, as controls.[23] There are 6,976 data zones in Scotland, each containing between 500 and 1,000 residents. According to the 2011 census, the median population is approximately 750 inhabitants, with a median (average) extension of 21 (1117) hectares. The SIMD information pertains to the 2016 edition, which uses the data zones as redefined in the 2011 Census. Figure 1.1 provides an example of how data zones nest within secondary school catchment areas. There are, on average, about 27 data zones per catchment area.

1.3.1 School 'quality' indicators and multidimensionality

When discussing school quality as perceived by parents, a distinction needs to be made between schools with high-performing peers, and 'effective' schools, defined as those best able to actively contribute to improve pupils' learning. These two scenarios do not necessarily coincide (Doris et al., 2022). Second, even when this distinction is clear, schools might differ in terms of a vast array of other characteristics, both within the cognitive and non-cognitive spheres. While it might seem obvious that schools that are 'good' under a certain profile (either because they have high-performing pupils, or because they are effectively enhancing their performance) should also be so under alternative profiles, there is evidence of school multidimensionality (see e.g., Beuermann et al., 2018).

In this paper I want to assess whether: *i)* Scottish secondary schools' output is 'multidimensional'; *ii)* parents perceive multidimensionality. Table 1.1 reports a factor analysis run on sixteen potential

[23] Scotland Census 2011, © Crown copyright. Data supplied by National Records of Scotland. Scottish Index of Multiple Deprivation, © Crown copyright 2016.

indicators of school 'quality', for all 275 schools contained in the full sample. The aim is to identify subsets of variables which might be indicative of some specific school latent characteristic, as expressed by each factor. Factor loadings below .6 have been marked with blanks.

Table 1.1: Factor Analysis

Variables	Factor 1	Factor 2	Factor 3
% Achieving SCQF level 6[a]	0.945		
% Achieving SCQF level 5[b]	0.871		
% Attaining Literacy & Numeracy SCQF level 5	0.821		
% of S6 leavers	0.881		
% of Leavers in Higher Education	0.912		
% of Leavers in Further Education	−0.633		
% of Leavers Working			
% of Leavers in Positive Destination	0.636		
Attendance Rate (In School)			0.669
Lateness Rate			−0.690
Truancy Rate			
Exclusions Rate			
% Achieving SCQF level 6 Value-Added		0.702	
% Achieving SCQF level 5 Value-Added		0.800	
Literacy & Numeracy - Value-Added		0.808	
% of Pupils not on FSM	0.630		0.640
Proportion	0.479	0.189	0.188
Cumulative	0.479	0.668	0.857

Note: This table shows the factor analysis run on the full sample of schools. The rotated factors being retained are those with eigenvalue ≥ 1 and they account for nearly 86% of the overall variance, 48% of which is in the first one.

[a] At least 4 awards
[b] At least 4 awards

As is evident from the first column, factor 1 loads higher on indicators of academic performance, in particular on the fractions of SCQF level 6 awardees, fraction of S6 leavers and those in higher education. This suggests some commonality in variables which might reflect the academic orientation of a school. Not surprisingly, factor 1 also features a proxy of school composition, yet with a lower loading, and it loads negatively on the fraction of leavers in further (vocational) education. Factor 2 meanwhile presents high loading on the value-added measures, while Factor 3 loads highly, though with a different sign, on attendance and lateness rates as well as on the school composition, which is consistent with studies linking socio-economic status to school absenteeism (see Sosu et al., 2021).[24]

[24] Results from a Principal Component Analysis are available on request and suggest essentially the same patterns.

Results in Table 1.1 can be summarised as follows. First, school performance might be bi-dimensional, i.e. value-added and average (peers') academic performance are two separate latent factors. Second, non-cognitive components and academic outcomes might be treated as two separate indicators. However, the third factor captures commonality in only a subset of potential measures of non-cognitive skills. Third, school composition enters both the academic performance and non-cognitive skills factors. Based on this, I build three different performance indexes by running principal component analysis (PCA) on three subset of variables separately, and extract the first component from each of these. These first components, alongside their variables and loadings (weights) are reported in Table 1.2.

Table 1.2: Indexes

Variables	Weights
Academic Index	
% Achieving SCQF level 6[a]	0.469
% Achieving SCQF level 5[b]	0.463
% Attaining Literacy & Numeracy SCQF level 5	0.440
% of Leavers in Higher Education	0.456
% of S6 Leavers	0.405
Value-Added Index	
Achieving SCQF level 6 Value-Added	0.587
Achieving SCQF level 5 Value-Added	0.606
Literacy & Numeracy - Value-Added	0.537
Non-cognitive Skills Index	
Attendance Rate (In School)	-0.595
Lateness Rate	0.509
Truancy Rate	0.402
Exclusion Rate	0.474
Socio-Economic Composition	
% of Pupils not on FSM	1

Note: This table shows how the main indexes used in this analysis have been built. Except for *Socio-Economic Composition*, which is fully proxied by the school-level % of pupils who are not eligible for free school meals, the other indexes are the first components of three different Principal Component Analysis run on the relevant groups of variables. The "Weights" of each variable correspond to the PCA loadings of each variable at hand.

[a] At least 4 awards

[b] At least 4 awards

Despite the third factor in Table 1.1 loading on attendance and lateness only, I follow the standard approach in the literature (e.g. Jackson, 2018) and build my non-cognitive skills index by running PCA on all four proxies of non-cognitive skills, i.e. attendance, lateness, truancies and exclusions. The rationale is to build a more encompassing measure. The non-cognitive skills first component loads negatively on attendance rate and positively for the remaining two variables. This implies that higher values of the component correspond to higher exclusion, truancy and lateness rates, and lower attendance. Hence, to build the index I re-coded the component so that higher scores correspond to higher level of non-cognitive skills. Finally, each of these indexes have been standardised, within each school year, to have mean zero and unit-variance. In addition to the above-mentioned three indicators, I also use a proxy for schools' socio-economic composition, which is merely the standardised value of the percent of pupils not registered for free school meals.[25] While this variable features in two of the three factors in Table 1.1, suggesting a strong correlation with both non-cognitive skills and academic performance, I still use it as a stand-alone potential predictor of school choice.

In summary, the four indicators are: *i)* an *Academic* index, which is meant to capture school-level average peers' performance; *ii)* a *Value-Added* index, as a measure of school effectiveness; *iii)* a *Non-cognitive Skills* index, as a measure of school-level average peers' behaviour; *iv)* a proxy for *Socio-Economic Composition*. These four indicators constitute my main (endogenous) regressors, whose mutual correlations are presented in Figure 1.3. The scatter plots are produced using the 275 schools present in the full sample, and each dot is the 2014-2016 average school-level indicator.

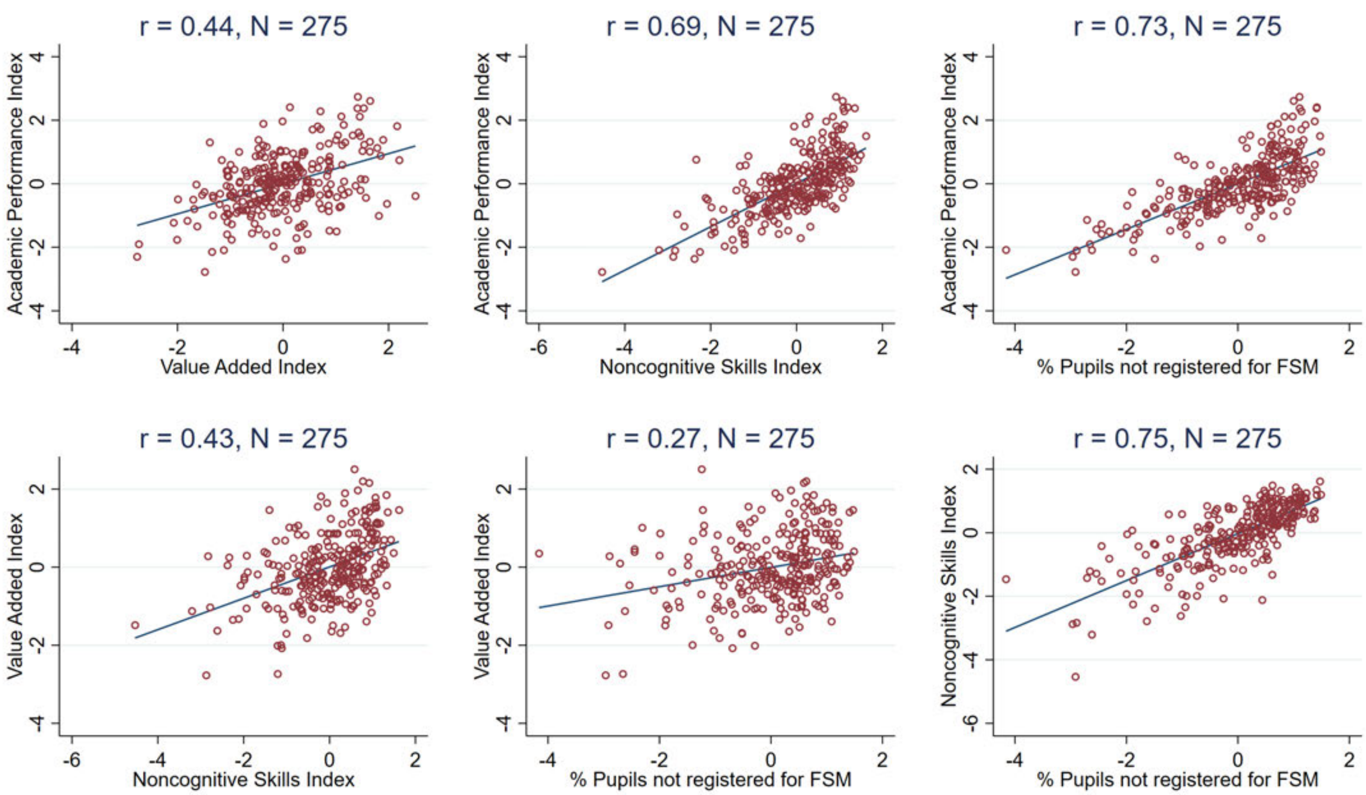
Peers' academic performance and value-added are only weakly correlated (r=0.44), which is indicative of the fact that schools with high peers' performance are not, on average, those which boost pupils attainment.[26] Not surprisingly, academic performance is also positively and strongly associated with the fraction of pupils from an advantaged socio-economic background (r=0.73), and similarly with peers' non-cognitive skills (r=0.69). Furthermore, value-added weakly correlates with non-cognitive skills (r=0.43) and school composition (r=0.27).[27] These results largely support the finding from Table 1.1 against school outcomes' multidimensionality, whereby value-added is the only dimension dis-

[25]Please note that, unlike for some stages in primary school, free school meals in secondary school have not been made universal, therefore this variable should still capture meaningful variation in socio-economic background.

[26]Note that value-added index relates to the same qualifications the academic performance index is based upon.

[27]While it can be argued that correlation coefficients within the order of 0.27 and 0.44 still imply strong associations, these are of significantly lower order relative to 0.73 or 0.69.

Figure 1.3: Correlation Between Indicators



Note: In this set of scatter plots is reported the correlation between performance indicators at the school level. In particular, each dot corresponds to a school whose value has been standardised to have mean zero and unit-variance and averaged across all school years (2013/14, 2014/15 and 2015/16).

cernable from peers' performance, behaviour and school composition. The remainder of this paper investigates whether this bi-dimensionality is perceived by parents and whether within the same dimension some indicators constitute a better signal of 'quality'.

1.4 Empirical Strategy

A viable way to determine households' preferences for schools is via the capitalisation of their characteristics into house prices, using a standard hedonic pricing model (Sheppard, 1999):

$$p = \beta Q + \gamma X + u$$

where house prices (p) are modelled as a function of school quality (Q) and house and neighbourhood features (X). However, obtaining a consistent and plausibly causal estimate for β, the marginal willingness to pay (WTP) for a move closer to a school perceived to be 'better' (i.e. higher quality/performance), is not straightforward. Schools are not randomly assigned to neighbourhoods and estimating their valuation in a simplistic fashion like the model above can lead to biased estimates if confounding factors and sorting are not properly accounted for (Epple and Romano, 2003; Bayer et al., 2007; Greaves and Turon, 2021). For instance, wealthier residents tend to buy more expensive houses. The main consequence, at least in this context, is that more affluent parents might devote more resources to their children's learning, thus affecting the observed performance of the local institution. Moreover, unobserved heterogeneity at the neighbourhood level in factors which affect property prices and are also correlated with school quality might induce further bias. Some previous research disentangles this endogeneity issue by exploiting quasi-random variation generated by reforms of the attendance system, or re-drawing of attendance areas (see for example Bogart and Cromwell, 2000; Ries and Somerville, 2010; Machin and Salvanes, 2016). In absence of such natural experiments, I follow a standard approach in the literature (see e.g. Black, 1999; Gibbons and Machin, 2003, 2006; Gibbons et al., 2013), by comparing house prices in close vicinity of catchment area borders, according to the following equation:

$$ln(p_{isnbt}) = \beta Q_{snb,t-1} + \gamma X_{nb} + f(d_{isb}) + \lambda_t + \underbrace{\delta_b + \varepsilon_{isnbt}}_{u_{isnbt}} \tag{1.1}$$

where p_{isnbt} is the price of house sale i, in school catchment area s, in data zone n, in proximity of catchment area boundary b, sold in year t, whose main predictor of interest is $Q_{snb,t-1}$, namely a proxy of school 'quality' at least one period prior to the sale. In particular, I use transaction prices from years 2015 to 2018, whereas school performance refers to years 2014 to 2016. For instance, properties sold in 2015 are matched to school performance in school year 2013/14. Not only does this method find support in the literature (e.g. Gibbons et al., 2013) but it is also driven by school statistics being published nearly a year after the school year has concluded. X_{nb} contains a set of data zone-specific controls at boundary b, while λ_t includes year of sale indicators. Additionally, $f(d_{isb})$ consists of location-specific features, such as polynomials of distance from boundary and catchment school. Finally, the (composite) error term u_{isnbt} breaks down into an idiosyncratic term, ε_{isnbt}, as well as a boundary-specific unobserved component, namely δ_b.

Estimating the coefficient of interest, β, requires the identifying assumption that u_{isnbt} needs to be orthogonal to $Q_{snb,t-1}$. This might not be the case if boundary-specific unobserved characteristics (δ_b) are correlated with school performance. Therefore, I compare sales which are close enough to each other to share similar spatial characteristics, yet are located on different sides of a catchment area boundary. By controlling for δ_b, I thus exploit within-boundary variation in school performance.

I visualise the identification strategy in Figure 1.4 taking as an example Leith Academy's catchment area and its southern boundary separating it from Portobello High School. The border on that section of Portobello Road separates two sides of the same street into different school catchment areas. Hence, any price variation across the border can be attributed to a difference in the performance between the two schools. Finally, to the best of my knowledge catchment area boundaries have remained essentially unchanged for years. After carefully inspecting the websites for Glasgow and Edinburgh Councils, the two largest Local Authorities, I find no evidence of secondary school catchment areas' boundaries being re-designed for the period this analysis refers to.

It is worth specifying that a 'boundary' is defined as the area within a certain distance from either side of a catchment 'border'.[28] Each boundary refers to two schools (one on each side) and might contain more than one data zone. Hence, it is possible that there could be variation in socio-demographic characteristics within boundaries, i.e. along and across the border, since these features are collected at the data zone level. This is an aspect I fully address in Subsection 1.4.1.

1.4.1 Validity

The validity of the above design crucially depends on two conditions. First, there needs to be enough within-boundary variation in school performance to explain house prices, i.e. there is on average suf-

[28] Note the subtle difference in the meaning of the words 'boundary' and 'border' for the purpose of this paper. While the 'border' is the actual line separating two catchment areas, the 'boundary' is instead defined as the locality in proximity (and on each side) of the 'border'.

Figure 1.4: Catchment Area Example - Leith Academy

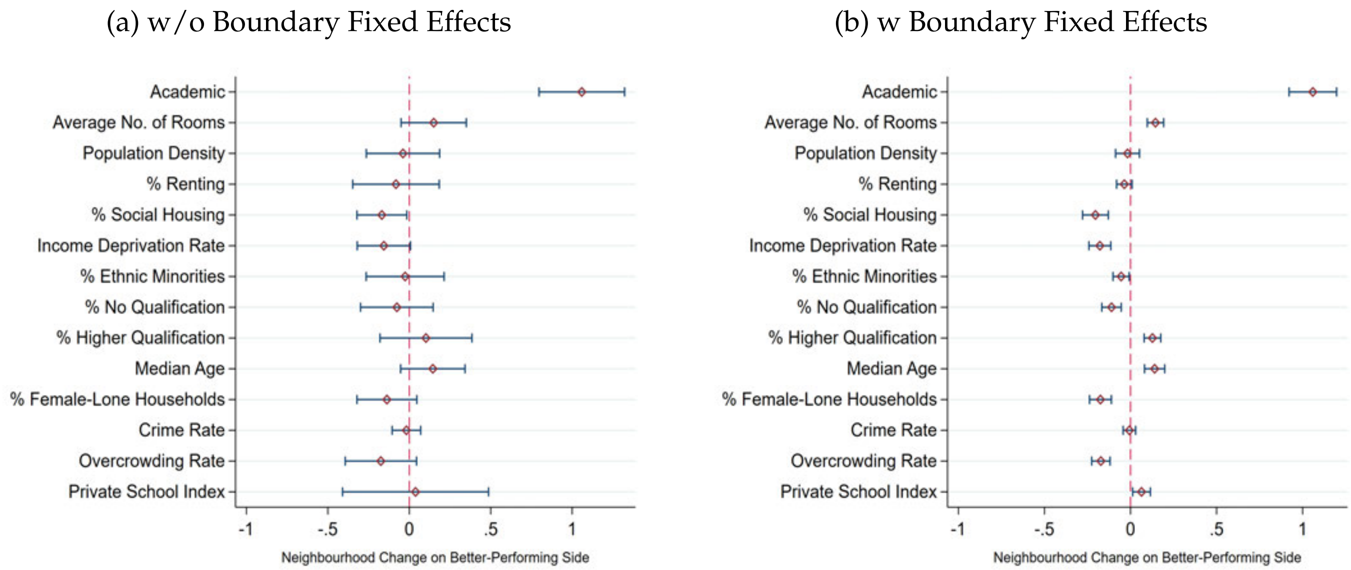
(a) Leith Academy Catchment Area (b) Leith Academy/Portobello High Boundary

Note: This figure shows the catchment area of Leith Academy, alongside a portion of its lower boundary, which separates Leith Academy catchment area from the one of Portobello High School.
Source: The Edinburgh Council (2019). Reproduced by the Edinburgh Council from Ordnance Survey material on behalf of the controller of Her Majesty Stationery Office © Crown copyright. Licence number 100023420. City of Edinburgh Council 2019.

ficient difference in performance between schools with bordering catchment areas. Second, any neighbourhood characteristic which is plausibly correlated with school attributes and is a valid predictor of house prices must not systematically differ across the two sides. This is the standard validity assumption of a regression discontinuity design (see for example Lee and Lemieux, 2010). In other words, the side of the boundary which is 'treated' with the higher-performing school must share similar housing and neighbourhood characteristics as the opposite side of the boundary which is used as its counterfactual.

Both conditions are examined in Figure 1.5. This reports estimated coefficients (alongside their 95% level confidence intervals) for each of the variables presented, regressed on a binary indicator for the 'good' side of the boundary, as the one with a higher value of the school-level *Academic* index. The sample is restricted to sales within 350 metres from their closest school area border. In addition, I exclude those boundaries coinciding with Local Authority borders. Standard errors are clustered at the secondary school catchment. To better gauge the size of the change, the variables have all been standardised to have mean of zero and unit-variance. Figure 1.5a reports results from simple univariate

Figure 1.5: Balancing Tests



Note: Each coefficient results from a regression of the variable at hand on a binary indicator of whether the property is located on the better-performing side of the catchment area boundary, based on the *Academic* performance indicator. The whiskers indicate 95% level confidence intervals.

regressions, whereas to produce Figure 1.5b I add boundary fixed effects to the models. We can see that, on average, the better-performing side of the boundary is characterised by a value of the *Academic* index which is approximately one-standard deviation greater than the opposite side. While this difference occurs by construction, it is still sizeable and a common finding in the literature (see for example, Black, 1999; Bayer et al., 2007 and also Gibbons et al., 2013), and is a similar order of magnitude when we compare sales within-boundaries as in Figure 1.5b. However, alongside this considerable difference in academic performance, there is no systematic difference in a series of socio-demographic characteristics recorded at the data zone-level. For instance, looking at Figure 1.5a, houses are not significantly different in size on either side of the boundary, nor are they located in substantially more or less densely populated areas.[29] A similar result applies to the percentage of people with a higher education degree, without qualification, female-headed households with children. Not only are these differences not statistically significantly different from zero (most of the confidence intervals include zero), but they also account for far less than half of a standard deviation within their own distributions.[30] Furthermore, the

[29]Unfortunately the data do not allow me to observe the characteristics of individual houses, therefore I use as a proxy the average number of rooms in the neighbourhood.

[30]See Panel B in Table A.1 for descriptive statistics

percentage of income deprived people changes only slightly, alongside the percentage of households in social housing, and they are barely statistically significantly different from zero at the 5% level. In particular, the side of the boundaries with higher academic performance has a percentage of households in social housing which is on average approximately .20 of a standard deviation lower than the lower-performing counterpart (22.38×.20 ≈ 4.4pp). This is arguably a small difference.

As I compare transactions within boundaries, i.e. by adding boundary fixed effects, differences in observable neighbourhood features appear more marked. Taken at face value, Figure 1.5b suggests that the higher-performing side of the boundary is characterised by larger and less overcrowded dwellings, more educated households, as well as fewer households in social housing and a smaller fraction of income deprived individuals. While these point estimates are largely unchanged relative to those in Figure 1.5a, they are now statistically significant, and this might support the hypothesis of sorting.[31] In the next section I propose a strategy to overcome this issue.

1.4.2 Data zone fixed effects

I enhance the model in Equation 1.1 by focusing only on those transactions which are in the same data zone (same n), but on opposite sides of the catchment area boundary. Figure 1.1 shows how catchment area borders often are coincident with data zones borders. However, there are instances in which the catchment area borders cut through data zones. By limiting the sample to transactions occurring within these data zones, I can compare houses which share the same neighbourhood characteristics. That is, I modify Equation 1.1 to be

$$ln(p_{isnbt}) = \beta Q_{snb,t-1} + f(d_{isb}) + \lambda_t + \delta_{nb} + \varepsilon_{isnbt} \tag{1.2}$$

whereby δ_{nb} represents data zones fixed effects. Note that as the analysis now focuses entirely on data zones spanning across the catchment area borders, by controlling for data zone fixed effects I automatically employ a boundary fixed effects strategy, but to a more localised extent. Importantly, I

[31] Figs. A.2 to A.4 repeat the same exercise using the other three performance indicators (value-added, non-cognitive skills and composition). The results of these analysis are analogous to the findings presented for the *Academic* index.

therefore automatically control for all the socio-demographic observables collected at the data zone-level and also all time-invariant unobservables. This motivates why the term X_{nb} does not feature in Equation 1.2 as this contains the data zone-level covariates listed in Figure 1.5. For this same reason, I do not repeat the exercise in Figure 1.5 within data zones. In addition, these covariates are time-invariant within the time window of this analysis. This technique has been widely employed in the literature (see e.g. Gibbons and Machin, 2003; Gibbons et al., 2013). While δ_{nb} is assumed to be boundary-specific unobserved and time-invariant component, I demonstrate in a robustness check that controlling for boundary-specific trends does not alter the results.

While observable characteristics are by construction fixed within data zones, sorting might still occur, i.e. wealthier, more educated households living on the higher-performing side of the boundary, yet within the same data zone. Unfortunately, lack of dwelling-specific information prevents me from addressing this methodological concern. For instance, if I observed dwellings' size or type, I could infer the type of household living in it and appraise any discontinuity across boundaries. Nevertheless, I argue that this extent of sorting is highly unlikely. First, data zones are fairly homogenous by construction. In fact, they nest within Local Authorities and were built based on a series of criteria such as: number of residents (between 500 and 1,000), compactness of shape, similarity of households' socio-economic characteristics and accordance with other boundaries and environmental features, e.g. wards, postcodes, lakes, rivers, roads, railways etc. (Flowerdew et al., 2004). Second, I provide some descriptive evidence of this by appraising within-data zone variation in: *i)* number of rooms; *ii)* council tax value. All households (with due exemptions) in UK are liable to pay council tax, which is levied on domestic properties. Council tax entails a banding system going from A (lowest rate) to H (highest rate) determined by the value of the property recorded in April 1991, alongside other characteristics such as size, layout, character and location (Valuation Office Agency, 2022).

Figure A.5 plots the distributions of the within-data zone standard deviations of the afore-mentioned variables. Based on the data zone-level fraction of dwellings, I calculate the average number of rooms and council tax values, and plot their standard deviations. The dotted red line represents the 75th percentile. Figure A.5a suggests that within 75% of data zones we have a variation of about 1.5 rooms. Similarly, Figure A.5b implies that within 75% of data zones the variation in council tax amounts to less than £300 (which corresponds to roughly a gap of two bands, e.g. from A to C).

Altogether, this is suggestive of data zones being fairly homogenous and therefore is unlikely that sorting might occur within them on the higher-performing side of the boundary.

1.5 Results

In Table 1.3 I present the results from estimating some variations of the models presented in Equation 1.1 and Equation 1.2. Each estimated coefficient corresponds to $\widehat{\beta}$, obtained from regressing $ln(price)$ on one performance indicator at the time. In other words, I propose estimates of valuation of each single school indicator unconditionally on the other. All specifications contains by default indicators for the year of sale, total number of pupils in school alongside school area's classification.[32] Standard errors are clustered at the catchment area level and reported in parenthesis.[33] Finally, regardless of the specification, each indicator is standardised within school year based on the full sample (275 schools).

I first focus on the models using the peers' academic performance index as the main regressor. In column (1) I run a 'naive' version of the main model by using the full sample of houses and conditioning on data zone covariates and Local Authority fixed effects. Results suggest that a one-standard-deviation increase in academic performance index is associated with a 3.1% rise in house prices. In column (2) I estimate Equation 1.1 by focusing on houses within 350 metres from their closest catchment area borders. Taken at face value, a one-standard-deviation increase in academic performance index corresponds to approximately 2.7% rise in house prices. In this specification I control for the set of neighbourhood characteristics listed in Figure 1.5, without which $\widehat{\beta}_{Academic}$ would likely suffer of serious omitted variable bias. However, these predictors are likely to be endogenous.

For this reason, from column (3) I focus on a sample of sales located within the same data zones, yet on opposite sides of an attendance boundaries.[34] Hence, by employing data zone fixed effects, any

[32] In particular, I include a dummy variable for urban schools, whereby "urban" is defined as any settlement above 10,000 inhabitants.

[33] Clustering at the data zone level only changes standard errors slightly, without altering the main results. Results are available on request.

[34] Although they are very rare, I still exclude sales in proximity of catchment boundaries overlapping with Local Authorities ones.

Table 1.3: Main Results

Dependent Variable: ln(House Prices)	(1) Full Sample	(2) 350 Metres	(3) Same DZ	(4) Same DZ	(5) 300 Metres	(6) 250 Metres
Academic	0.031***	0.027***	0.037***	0.038***	0.038***	0.039***
	(0.008)	(0.006)	(0.014)	(0.014)	(0.014)	(0.014)
Value-Added	−0.005	0.000	0.006	0.006	0.005	0.005
	(0.006)	(0.006)	(0.010)	(0.010)	(0.010)	(0.010)
Non-cognitive	0.010	0.002	0.021*	0.022*	0.024*	0.023*
	(0.008)	(0.007)	(0.013)	(0.012)	(0.012)	(0.013)
Socio-Economic Composition	0.036***	0.043***	0.038**	0.038**	0.042**	0.044**
	(0.011)	(0.011)	(0.019)	(0.019)	(0.019)	(0.018)
Observations	221,073	52,471	13,368	13,368	12,659	11,668
No. of Schools	275	188	155	155	149	146
Mean	11.81	11.84	11.89	11.89	11.89	11.88
SD	0.68	0.69	0.71	0.71	0.71	0.72
Neighbourhood Controls	✓	✓				
Local Authority FE	✓					
Boundary FE		✓				
Data Zone FE			✓	✓	✓	✓
School Distance				✓	✓	✓
Boundary Distance Squared				✓	✓	✓

Notes: Each Column represents a specification in which $ln(HousePrices)$ is regressed on one indicator at the time. Each model contains a set of year of transaction dummies, school size and an indicator of geographic location (urban). Neighbourhood controls include the covariates from Figure 1.5. Data zone fixed effects make neighbourhood covariates redundant, hence these are not included from column (3) onward. Column (4) includes distance to catchment school (in 100s metres) and distance-from-boundary polynomials. Columns (5) and (6) pertain to sample of houses within 300 and 250 metres from catchent borders and within the same data zones. Adjusted R^2 range between .56 and .59. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

variation in property prices is solely attributed to variation in school characteristics, conditional on year of sale and school's area. Note that being neighbourhood characteristics collected at the data zone level, as I employ data zone fixed effects I do not need to control for these covariates.

In column (3) I observe that a one-standard-deviation increase in academic performance commands a property price increase by 3.7%. This result is substantially unchanged after I control for a function of location-specific features in column (4). This is $f(d_{isb})$, and it is proxied by distance to catchment school and a polynomial (squared) of distance to boundary.[35] Columns (5) and (6) show how the results are not sensitive to the choice of alternative distances to boundaries. As the focus narrows down to houses within 300 and 250 metres to boundary, and within the same data zones, neither the precision of the estimates nor their magnitude are affected. Overall, these results translates into a 3-4% house price

[35] An alternative specification features linear and cubic distance. These results are not featured in the paper as they are essentially identical to those in column (4), however they are available on request.

premium as a result of a 13-14 percentage points (approximately one standard deviation) increase in the percentage of pupils attaining four or more SCQF level 6 or enrolling in higher education. This is shown in Table A.4, which presents results obtained from the same specifications as in Table 1.3, but using the single variables the *Academic* index is obtained from.

Results are fairly similar for $\widehat{\beta}_{Composition}$, whose magnitude hover around a 3.8-4.4% house price premium for a one-standard-deviation increase in the school-percentage of pupils not eligible for free school meals. For my measures of value-added and non-cognitive skills, however, point estimates are considerably smaller and not statistically significant, despite being precisely estimated.

The above estimates appear consistent with the relevant literature. In particular, for one-standard-deviation increase in academic performance, these results are in line with the 2.5% reported by Black (1999), 2% by Fack and Grenet (2010), and 3% by Gibbons et al. (2013).

1.5.1 Robustness

Results so far suggest that parents value peers' academic performance and schools' socio-economic composition, but not effectiveness or peers' behaviour. Furthermore, by exploiting within-neighbourhood variation, my identification strategy is clean and provides results which are in line with previous work. Nevertheless, there might still be threats to their validity.

In this section I explain these and set out how I explore them in more detail. First, it is necessary to isolate the effects of school resources from one purely driven by school outcomes. Second, while the main focus of this project is on state funded non-denominational schools, the Scottish education system features also a private sector -independent schools- as well as state funded denominational schools, the vast majority of which are Roman Catholic (henceforth, RC). While the latter follow the same curriculum as the non-denominational schools, I do not observe private schools' performance.[36] Third, I want to control for the possible presence of spatial trends in house prices, as well as unobserved heterogeneity across time and space. For this purpose, I use as a baseline my preferred specification, namely the one in column (3) in Table 1.3, and assess the sensitivity of my results to any departure from

[36] Some of which follow the English curricula or the International Baccalaureate.

this benchmark specification.

Table 1.4: Robustness Checks - Academic

Dependent Variable: ln(House Prices)	(1) Baseline	(2) Extra Controls	(3) Data Zones × Years	(4) Data Zones × Boundary Distance	(5) RC Sample	(6) RC Sample
Academic	0.037***	0.039**	0.034**	0.030**	0.041**	0.041**
	(0.014)	(0.015)	(0.016)	(0.013)	(0.016)	(0.016)
PT ratio		0.014				
		(0.014)				
(Roll/Capacity)×100		−0.003				
		(0.002)				
Private School Index		0.041				
		(0.033)				
RC Higher						0.028
						(0.023)
Observations	13,368	13,368	13,368	13,368	12,390	12,390
No. of Schools	155	155	155	155	124	124
Mean	11.89	11.89	11.89	11.89	11.89	11.89
SD	0.71	0.71	0.71	0.71	0.72	0.72
Datazone FE	✓	✓	✓	✓	✓	✓
Adjusted R-squared	0.594	0.594	0.607	0.637	0.599	0.599

Notes: All of the above specifications refer to the sample of houses within those data zones which stretch across two catchment areas. Every specification includes data zone fixed effects, alongside indicators for year of transaction, urban area and school size. Data zone fixed effects make neighbourhood covariates redundant, hence these are not included. In columns (3) and (4) data zone dummies are interacted with year indicators and distance from boundary respectively. Columns (5) and (6) focus on sales which are assigned to a Roman Catholic school. In column (6) I control for the local Roman Catholic school's level of 'Higher'. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.4 runs a battery of robustness checks for the academic performance indicator. The first column reports the baseline estimate. Column (2) addresses the first three points outlined at the beginning of this section by controlling for pupil-teacher ratio, number of pupils relative to capacity and private schools presence index similar to the one used by Fack and Grenet (2010).[37]

Columns (3)-(4) control for potential heterogeneity in 'data zone effects' over time and distance from boundaries. Columns (5)-(6) addresses whether property prices pick up variation in performance of the local RC school, proxied by its score in Higher qualification exams. One thing to notice is that RC schools are not present in every Local Authority. Therefore, in column (5) I first restrict the sample to all those transactions which are assigned to an RC school and check whether the results 'survive' this restriction. Subsequently, in column (6), I control for the local RC schools' performance, proxied by their share of 'Higher' qualifications.

We can see that in general there are only small changes to the main coefficient, $\widehat{\beta}_{Academic}$, when I try to verify its robustness to a series of factor. The conclusion from these robustness checks is that

[37] For each transaction I measure the distance to the closest 5 private schools and take the median of the inverse distance.

these changes to the main specification do not alter the central conclusion from Table 1.3. Table A.3 in appendix reports the same exercise but for the school socio-economic composition indicator.

1.5.2 Mechanism

My previous results examine preferences for some potential indicators of school quality, one at the time, via their house price capitalisation. Table 1.5 together with Figure 1.3 can shed light on the WTP of 'quality' indicators conditional on one another, alongside their correlations, thus providing further insights –after Table 1.1– on the extent to which school performance is multidimensional and whether parents are aware of multidimensionality (or absence of it).

Table 1.5: Multidimensionality

Dependent Variable: ln(House Prices)	(1)	(2)	(3)	(4)	(5)
	ln(Price)	ln(Price)	ln(Price)	ln(Price)	ln(Price)
Academic	0.037***	0.043***	0.035**	0.030*	0.035*
	(0.014)	(0.016)	(0.016)	(0.017)	(0.020)
Value-Added		−0.009			−0.009
		(0.012)			(0.012)
Non-cognitive			0.006		0.003
			(0.014)		(0.015)
Socio-Economic Composition				0.017	0.015
				(0.022)	(0.023)
Observations	13,368	13,368	13,368	13,368	13,368
No. of Schools	155	155	155	155	155
Mean	11.89	11.89	11.89	11.89	11.89
SD	0.71	0.71	0.71	0.71	0.71
Data Zone FE	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.594	0.594	0.594	0.594	0.594

Notes: All of the above specifications refer to the sample of houses within those data zones which stretch across two catchment areas. Every specification includes data zone fixed effects, alongside indicators for year of transaction, urban area and school size. Data zone fixed effects make neighbourhood covariates redundant, hence these are not included. Standard errors (in parenthesis) are clustered at the school catchment area level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Schools with high peers' performance do not coincide with those with high value-added. Nevertheless, from Table 1.3 emerges that parents prefer the previous over the latter. There are usually two possible interpretations of this: *i)* Parents are not aware of which schools are good at raising their pupils performance - however, I rule out this scenario as virtual comparators are publicly available and easily accessible alongside the other attainment statistics; *ii)* Parents simply do not consider value-added as a valuable information about the quality of a school. Perhaps, despite value-added information being widely available, this might not provide a great deal of insights to parents. Another explanation relies on the upper bound in attainment's improvement. For instance, parents of high-achieving pupils might not find school effectiveness an appealing feature, especially if this is only weakly correlated with other more sought after dimensions such as socio-economic composition and peers' performance. Further evidence of lack of interest in value-added comes from column (2) of Table 1.5 - even conditional on peers' academic performance, value-added capitalisation is still non-distinguishable from zero (it even turns negative) and non-statistically significant, whilst average academic performance increases by nearly 14%.

In addition, peers' academic performance is highly correlated with school socio-economic composition and peers' non-cognitive skills. While the previous is strongly capitalised into house prices, with approximately the same premium of academic performance, from Table 1.3 it also appears that non-cognitive skills' coefficient is still small in size and hardly ever statistically significant. A plausible explanation is that school-level behaviour is not primarily a factor of choice for parents, but might instead matter conditional on peers' academic performance. However, from columns (3) of Table 1.5 it can be noticed that non-cognitive skills still present no capitalisation while keeping academic performance constant. The latter, on other hand, remains well-capitalised and statistically significant.

Furthermore, looking at column (4), school socio-economic composition ceases to be an important factor of school choice conditional on peers' academic performance, which in turn shrinks slightly and it is not statistically significant at the 5% level. While this is likely a consequence of the strong correlation between these two indicators, it is further evidence of lack of interest in value-added - in other words, average academic performance matters less conditional on school intakes.

Finally, in column (5) I use all indicators within the same specification. The main outcome is that peers' academic performance is the only coefficient maintaining its original magnitude, yet losing statistical significance. However, like in column (4), this might be due to imperfect multicollinearity - standard error is significantly larger than in column (1) - as little independent variation is left.

The above findings can be summarised in two points. First, school socio-economic composition and peers' academic achievement might simply reflect the same school dimension, thus providing the same information to parents. However, the latter might be perceived as a better signal of 'quality'. Second, despite non-cognitive dimension being highly correlated with peers' academic performance, it does not appear to be capitalised into house prices. This points to the fact that parents do not value this dimension per se, and might be seeing peers' academic performance and free meals registrations as more comprehensive and informative school's traits.

An additional explanation lies in the salience of the information available to parents. Tables A.4 to A.6 repeat the same exercise as in Table 1.3 but using the single variables the indexes are obtained from. For instance, Table A.6 presents four different sets of specifications, one for each non-cognitive component, i.e. attendance, lateness, truancy and exclusions. While attendance exhibits statistically significant coefficients, whose magnitude aligns with previous results (approximately 2.7% house price premium as a result of a one-standard-deviation increase in attendance), lateness, truancy and exclusions present no house price capitalisation. This is likely due to attendance being the most salient non-cognitive skills component, featuring prominently within publicly available school information, while not much is known about the rest.

The information factor also plays a role in relation to the other indexes. Table A.4 shows that, taken individually, the percentage of pupils attaining SCQF level 6, Literacy and Numeracy as well as those enrolling in higher education, command a premium in line with the one of the overall *Academic* index. However, the percentage of pupils attaining SCQF level 5 and S6 leavers are less decisively capitalised. Coefficients are in the order of a 2% house price premium and are not statistically significant at the 5% level. While the previous can be due to SCQF level 5 being regarded as low-stakes qualifications, the number of S6 leavers is also not easily accessible to parents. On the other hand, SCQF level 6 qualifications are highly visible within the set of information available to parents.

1.6 Conclusion

Strategic residential choices based on school quality is widely understood, anecdotally even before it was demonstrated in the academic literature. This paper contributes to the sizeable literature exploring the extent to which school quality is capitalised into house prices, and extends this literature in a number of ways. First, it overcomes the issue of lack of information about value-added by using an indicator which has been salient and publicly available for years; second, the indicator can be seen as a contextualised measure of value-added, as opposed to the average gain in attainment by the end of school cycle; third, it examines whether parents are aware of school multidimensionality; fourth, it sheds light on the role of peers' non-cognitive skills in school choice.

I reach three main conclusions: First, school performance in secondary schools is not strictly multidimensional, but perhaps bi-dimensional, with high performing schools not necessarily being also the most effective ones. Second, school effectiveness does not appear to be a key factor in school choice. This result confirms recent findings from MacLeod and Urquiola (2015) and Abdulkadiroğlu et al. (2020) which show that school composition and peers' performance might be a valid signal of quality when information on effectiveness might not be available, provided that these two dimensions are correlated with effectiveness. In this paper however, not only is performance benchmark publicly available, but it is also only weakly correlated with school-level average performance. Therefore this finding suggests that school effectiveness alone does not appear to be a critical factor. One important implication is that parents might not necessarily value schools that offer learning improvements, but instead attach more value to schools with larger fractions of high-performing peers, even if this is driven by selection. Nonetheless, the increase in WTP for a place with high-performing peers conditional on value-added might provide a new outlook on the role of value-added in school choice. While this does not carry independent capitalisation, it might still constitute an important signal for parents when screening schools with strong pupils.

This leads to the third conclusion, namely the fact that school-level peers' academic attainment and school socio-economic composition are equally capitalised into house prices, when considered individually, and strongly (and positively) correlated. A possible conclusion is that these are in fact two sides of the same coin, and that they might simply contain the same signal of 'school quality', and that their individual effects are hard to isolate. Finally, despite the non-cognitive component correlating strongly with academic performance and composition, no independent effect was found. While this might not be part of households' utility function altogether, it could simply mean that parents infer peers' non-cognitive skills from composition or average attainment, especially if information on non-cognitive skills is only available to a limited extent.

In summary, the findings of this paper suggest approximately a 4% house price premium for a one-standard-deviation increase in school-average academic performance. A back of the envelope calculation results in parents paying on average a £7,000 premium at the mean to secure their children a place next to high-performing peers.[38] This valuation is in complete isolation from school effectiveness. In addition, parents completely overlook potential spillovers in non-cognitive skills, which have recently been proved to be important in shaping peers' attitudes (Norris, 2020) but also directly affecting academic success (Shure, 2021) and aspirations (Dickerson et al., 2018). This raises numerous concerns in terms of how school policy might adjust to these preferences but also, from an individual perspective, whether considering school 'quality' on a single metric might undermine the success and complexity of the learning process.

A number of additional areas remain for future work. Primarily, future research may identify what makes a school 'effective', and subsequently, examine whether 'better' peers matter more than these factors, and in particular which pupils benefit more from what. For instance, if the strength of peers' effect is a function of one's ordinal rank within class/cohort (Bertoni and Nisticò, 2023), it might be possible that pupils down the rank might benefit particularly from the influence of academically stronger class mates, rather than, perhaps, pedagogy or smaller class size. This might not necessarily be true, however, for the 'big fish in a small pond' (Elsner and Isphording, 2017), who might perhaps necessitate different stimuli and resources. The extent to which parents are aware of this and make the choice accordingly is crucial, and future research should aim at informing policies helping parents in the choice of the best path for their children.

[38] see Table A.1

Chapter 2

EARLY-YEARS MULTI-GRADE CLASSES AND PUPIL ATTAINMENT

2.1 Introduction

Classroom composition and peer effects have been shown to be important determinants of pupil achievement. Several studies have documented the benefits of classroom exposure to high-ability peers (Hanushek et al. 2003; Lefgren 2004; Ding and Lehrer 2007; Neidell and Waldfogel 2010; Lavy et al. 2012a,b), to female classmates (Hoxby 2000a; Lavy and Schlosser 2011; Black et al. 2013; Anelli and Peri 2019) and to classmates with college-educated mothers (Bifulco et al. 2011, 2014) as well as the adverse effects of disruptive peers (Figlio 2007a; Aizer 2008; Carrell and Hoekstra 2010, 2012; Carrell et al. 2018a). The ethnic makeup of classrooms (Angrist and Lang 2004; Hoxby and Weingarth 2005; Hanushek et al. 2009; Hanushek and Rivkin 2009; Fruehwirth 2013) and the effect of immigrant peers on natives (Gould et al. 2009; Ballatore et al. 2018) have also received attention. However, little is known about a widespread classroom structure that explicitly creates and harnesses peer effects: multi-grade classes. These are classes comprised of pupils from adjacent grades. For instance, first-graders being taught alongside second-graders, and thus being exposed to older, more experienced peers.[1]

Multi-grade classes are widely used. About 28% of schools in the US use a mixed class setup and more than a third of primary school pupils in France attend multi-grade classes (Leuven and Rønning 2014). Yet, multi-grade classes have not been widely studied. A notable exception is Sims (2008) who documents that multi-grade classes were an unintended consequence of California's Class Size Reduction Program as schools pooled pupils from adjacent grades in order to lower average class size and thus qualify for additional funding. He shows that this had a detrimental impact on the test scores of pupils in multi-grade classes. Recent studies of rural areas of Norway (Leuven and Rønning 2014) and Italy (Checchi and De Paola 2018; Barbetta et al. 2019) have built on this work. They exploit that in these rural settings cohorts are often so small that pooling several year-groups is done out of necessity. With the exception of Checchi and De Paola (2018), they find that pupils in these schools actually benefit from attending multi-grade classes.

In shaping policy, decision makers need to know whether the benefits documented by this nascent literature translate outside of a rural context or whether - consistent with Sims (2008) - multi-grade

[1] A related strand of both the education (Slavin, 1987) and economics literature (Betts, 2011) has explored the effects of ability grouping and academic tracking.

groupings may even have a detrimental impact on pupil performance. In our study we are able to examine this issue directly because in Scotland, the subject of this study and a constituent nation of the United Kingdom, multi-grade classes feature in virtually all primary schools. In fact, they are consciously created in both rural and urban schools, which allows our study to investigate their impact on attainment in settings in which the majority of pupils are educated. As such, our study holds important lessons for both policy makers and education practitioners.

In order to identify the causal effect of multi-grade classes, we exploit that in Scottish primary schools, an algorithm ("class planner") determines the most cost-efficient number, size, and composition of classes, subject to nationwide minimum and maximum class size rules. Specifically there are class size limits for single-year classes which vary by grade, and separate caps for multi-grade classes. The class planner is set up to minimize the number of classrooms a school needs to create. Combined with fluctuations in enrolment counts across years, this generates variation in the composition of classes within and across schools.[2] In effect, small and random variations in enrolment counts trigger the creation of multi-grade classes in some grades, in some schools and in some years, but not in others.

Enrolment in Scottish primary schools is, in turn and on the whole, determined by random population variation. Every primary school has a catchment area and pupils within a school's catchment area are entitled to attend their catchment area school. Small changes in enrolment in any primary school grade can lead to a re-shuffling of pupils into multi-grade and non-multi-grade classes across all grades of the school. The ramifications of this reshuffling are particularly pronounced in first grade. This renders it all but impossible for parents or school administrators to manipulate the overall school enrolment count to either trigger or prevent the creation of a multi-grade class.

We exploit this natural experiment by instrumenting each pupil's class status (multi-grade or single-year-group) with the class planner's recommendation for whether the pupil's year-group should contribute pupils to a multi-grade class. Note that the class planner only makes a recommendation on how many pupils in a grade should be put into a multi-grade class, but not *which* pupils. We there-

[2] For instance, the maximum class size for fourth and fifth grade in Scotland is 33, while multi-grade classes are capped at 25. Therefore, for an enrolment count of 45 fourth-graders and 46 fifth-graders, the class planner would recommend the creation of one 33 pupil fourth and fifth-grade class each, and one 25 pupil multi-grade class. Yet with the addition of just one fourth-grade pupil (i.e. 46 pupils in both grades), class size maxima would force the creation of two fourth-grade and two fifth-grade classes.

fore identify a local average treatment effect (LATE). We document that the compliers tend to be older members of cohorts who form the lower-grade part of a multi-grade class. They typically share their multi-grade classroom with the youngest and low-attainment members of the preceding cohort who have an additional year of primary school experience.

We combine our instrumental variable approach with novel, individual-level administrative data collected from successive waves of the Scottish Pupil Census (SPC) from 2007/08 to 2018/19. We link these data with assessment information and observe the exact classroom type and composition in each school-year. However, the predictive power of the class planner is strongest in first grade, whereas analyses of later grades may at times suffer from "weak instrument" issues (see Bound et al. (1995) and Lee et al. (2020)). This article, therefore, focuses its conclusions on the attainment effects of exposure to older, more school-experienced peers in first grade.

We find that exposure to second-graders in the first year of primary school by way of a multi-grade class leads to large improvements in literacy and numeracy. In fact, gains created by multi-grade classes are roughly equivalent to the attainment gap between the average pupil and a pupil in one of the 20% most deprived data zones in Scotland.[3] Boys and pupils from deprived neighbourhoods appear to benefit more from sharing a classroom with more experienced peers, although neither gender nor socio-economic differences are significant in a statistical sense. We also find little in the way of an urban/rural differential. We find no evidence that the achievement gains for school-starters come at the expense of learning progress of second-graders who shared a multi-grade classroom with first-graders. However, we also document that the benefits for first-graders are short-lived.

Ours is the first study to document the benefits of multi-grade classes in a setting where they are not a niche phenomenon but a staple of the education system. In Scotland, multi-grade classes are used by schools in more affluent and less affluent areas alike, as well as in urban and rural schools. As such, our study pushes a nascent literature on multi-grade groupings forward and adds to its external validity. We also contribute to a growing literature on early years learning, from which we know the disadvantages of early school start and low age rank (Bedard and Dhuey 2006; Black et al. 2011; Crawford et al.

[3]Data zones are small area statistical geographies constructed by the Scottish Government comprising areas of approximately equal population size.

2014; Cascio and Schanzenbach 2016; Ballatore et al. 2020). We find that multi-grade classes help the youngest pupils in these classes at least as far as attainment is concerned – this underlines a distinction between absolute and relative age. Finally, we show that multi-grade classes save classrooms – and thus costs – while at the same time accruing net benefits in terms of pupil performance. Indeed, our results suggest that multi-grade classes are a viable way to better reconcile policymakers' goals of promoting higher-achieving pupils and pursuing value–for–money in education spending.

2.2 Data and Background

Pupils in Scotland typically start school in August of the year in which they turn five. They attend primary school from first grade (P1) to seventh grade (P7) before transferring into secondary schools. Government-funded public schools are free for the approximately 700,000 pupils aged 5-19. There is only a small private school sector, accounting for about 4% of pupils, which is mostly clustered in the populous *Central Belt* of the country. The Scottish education system has always been separate from that of the rest of the UK, education is devolved to the Scottish Government. In contrast to England where parental school rankings are solicited and pupils then matched to schools with open slots, school choice in Scotland resembles the system that is in place in most of the United States. That is, school choice is largely contingent on non-overlapping catchment areas which are drawn up by Local Authorities (roughly equivalent to school districts), and rarely ever change. Each primary school has a catchment area and any pupil whose main residence is within this boundary is entitled to a place in that school. Parents may also ask for their children to attend a school other than their catchment area school via so-called "placing requests". These are applications to the local council to transfer a child to a specified school. However, these requests are not automatically approved and, overall, only 5% of pupils in our sample attend a school different from the one of their catchment area.[4] Therefore, sorting into catchment areas of schools that are perceived to be desirable is a strictly dominant strategy for parents. Rossi (2021), for instance, documents that housing prices on two sides of catchment border areas in

[4]Councils are under no obligation to grant these requests and will not do so if a school is at capacity. Places are allocated based on criteria decided by each Local Authority, typically children with additional support needs and/or with siblings in the specified schools get priority.

Scotland differ on average by as much as 4%.

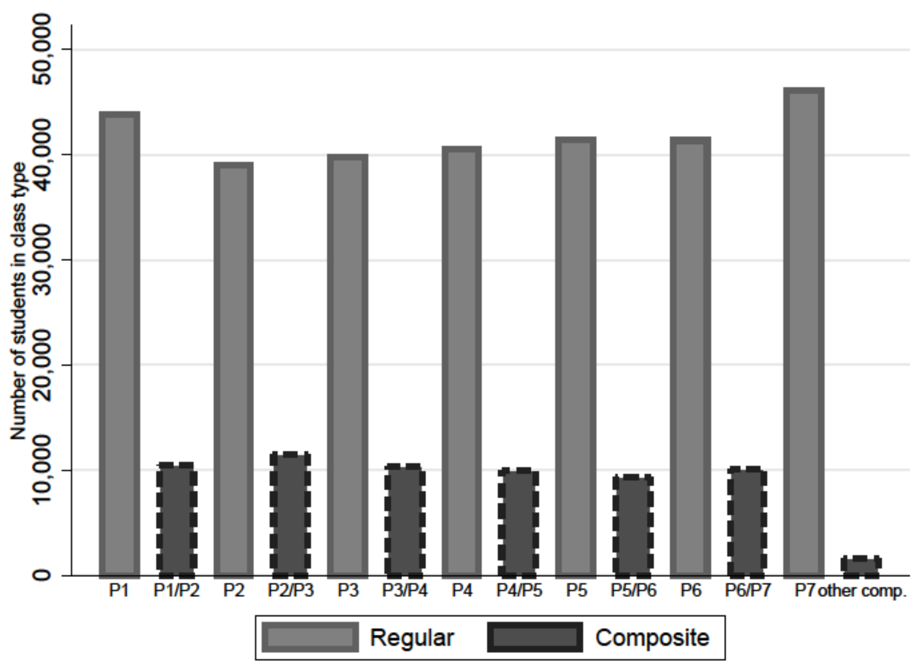
The Scottish Government centrally sets maximum class size rules in primary school which apply to the entire nation: class size in P1 must not exceed 25 pupils, the maximum for P2 and P3 is 30, and classes in P4-P7 are formed as multiples of, at most, 33 (see Table 2.1). A widespread feature of Scottish primary education are multi-grade classes, known as "composite classes" in Scotland (we use the two terms interchangeably throughout this paper). These are classes comprised of pupils from adjacent grades. The maximum class size for multi-grade classes is 25 and each grade needs to contribute a minimum of five pupils.

Table 2.1: Maximum Class Size Rules

Grade	Max. Size
Primary 1 (P1)	25
Primary 2 and 3 (P2, P3)	30
Primary 4 to 7 (P4-P7)	33
Composites (all grades)	25

Notes: Maximum class size rules in Scotland as of 2019. P1 cutoff was 30 prior to 2011.

Figure 2.1: Pupils by Grade and Class Type (2018)



Notes: This bar chart shows the distribution by class type (single-year vs multi-grade) of pupils in Scottish primary schools in 2018

Figure 2.1 provides an illustration of the distribution of pupils across single-year and multi-grade classes in 2018. Composite classes typically stretch across two grades and more than one in six Scottish primary school pupils attend a multi-grade class. In contrast to most of the examples in the literature to date, multi-grade classes are by no means a rural phenomenon in Scotland. For example, in 2018, 84% of primary schools in the City of Glasgow - the fourth largest city in the UK - featured at least one composite class.

Our data are drawn from the Scottish Pupil Census (SPC) for school years 2007/08 to 2018/19. The SPC takes place every year in September and collects information on every individual pupil and the schools they attend. Upon entering the Scottish school system, every pupil is assigned a unique ID, the so-called Scottish Candidate Number (SCN). We use the SCN to link pupils' records across years and to assessment data. Since 2015/16, every pupil's progress is assessed in both numeracy and literacy as either "Below Early Level", "Early Level", and at "1st/2nd/3rd/4th" level. These assessments are teacher-based but informed by standardized test scores to ensure consistency. Assessments are made at the end of P1 when pupils are expected to perform at early level, and at the end of P4 and P7 when students are expected to perform at the first and second level, respectively. We use the SCN to link each pupil to their assessments and create indicators for whether a pupil performs at the expected level in a given stage.

The SPC also documents the school and name of the class that each pupil attends as well as each pupil's grade or cohort. Since ours is individual level data, we can easily identify multi-grade classes and calculate class sizes which we cross-checked with official aggregates published by the Scottish Government. Table 2.2 presents summary statistics for about 190,000 first-graders who between 2015/16 and 2018/19 attended one of the 1,437 primary schools in our sample. Eighty-five and seventy-six percent of first-graders perform at level in numeracy and literacy respectively. The average class size is 21.8, about half the sample is female and the average school starting age is 5.2 years. We use the so-called Scottish Index of Multiple Deprivations (SIMD) as a proxy for socio-economic background. The SIMD ranks 6,976 'data zones' from most to least deprived in terms of income, employment, education, health, access to services, crime and housing. Unsurprisingly, about 20% of pupils come from households located in areas ranking in the bottom quintile.[5]

Table 2.2: Summary Statistics

	First-Graders (P1)		Fourth-Graders (P4)		Seventh-Graders (P7)	
	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.
Numeracy - Performing at level	0.851	0.356	0.759	0.428	0.731	0.444
Literacy - Performing at level	0.759	0.428	0.690	0.463	0.679	0.467
Reading - Performing at level	0.819	0.385	0.777	0.416	0.775	0.417
Writing - Performing at level	0.791	0.406	0.721	0.449	0.708	0.455
Listening & Talking at level	0.871	0.335	0.844	0.363	0.829	0.377
Class Size	21.813	3.265	26.635	3.955	26.413	4.323
Grade Enrolment	46.168	19.381	46.650	18.788	44.333	17.801
Female	0.491	0.500	0.493	0.500	0.491	0.500
White	0.828	0.377	0.855	0.352	0.878	0.327
Free Meal	0.339	0.473	0.179	0.384	0.167	0.373
Native English Speaker	0.926	0.262	0.924	0.265	0.937	0.243
Bottom 20% SIMD	0.226	0.418	0.217	0.412	0.216	0.411
Age (in Years)	5.210	0.307	8.205	0.308	11.209	0.313
% Female in School	0.490	0.032	0.490	0.032	0.490	0.032
% White British	0.848	0.123	0.852	0.116	0.853	0.119
% Free School Meals	0.247	0.199	0.246	0.197	0.250	0.198
% Native English Speakers	0.922	0.098	0.925	0.092	0.925	0.095
% in Bottom 20% SIMD	0.223	0.265	0.217	0.262	0.217	0.261
No. of Students in School	317.454	126.694	319.516	127.876	317.994	128.655
Observations	190,704		194,804		186,082	
No. of Schools	1,437		1,428		1,435	

Notes: All data stem from Scottish Pupil Census (SPC) 2015/16 - 2018/19, with assessment data added by matching via Scottish Candidate Number (SCN).

[5] Our sample also differs marginally from the original population data. We excluded about 1% of pupils who are either in special education classes, receive a Gaelic Medium education, or are in classes in which non-English speakers (e.g. refugees) were grouped together regardless of age/grade.

2.3 Empirical Strategy

Our aim is to compare attainment between pupils who attend multi-grade classes and those in single-year classes. We model attainment of pupil i in classroom c and grade g of school s in year t as a function of class type, observable student and school socio-economic characteristics as well as unobservable attributes. The following equation describes this education production function in its simplest form:

$$A_{icgst} = \beta_0 + \beta_1 Comp_{cgst} + \gamma X_{igst} + \delta_s + \tau_t + \varepsilon_{icgst} \tag{2.1}$$

Where A_{icgst} is achievement, in particular student competency in numeracy and/or literacy; $Comp_{cgst}$ is either a dummy that is equal to one for a multi-grade class and zero for a single-grade class, or a continuous variable equal to the number of older (younger) peers from preceding (succeeding) cohorts; X_{igst} is a vector of observed student characteristics such as age, gender, ethnicity, and socio-economic background, school-level fractions of the same characteristics, as well as a control for grade enrolment and class-size. δ_s and τ_t are sets of school and school-year fixed-effects, respectively.

Our main empirical concern relates to the endogeneity of $Comp_{cgst}$. Pupils who are placed in multi-grade classes are not randomly selected. In fact, both unobservable and observable pupil characteristics determine multi-grade status. For instance, head teachers might be inclined to select high ability students as the bottom part of a multi-grade class who are then pooled with low attainment pupils from the stage above. They are also encouraged to take social bonds into account, so as to keep groups of friends together. Maturity and age are also important considerations. Table 2.3 shows that older first graders are more likely to be placed in a P1/P2 multi-grade class whereas the opposite is true for second graders. While age and other demographic characteristics are observable, ability and social networks are not. As a result a simple OLS estimation of Equation 2.1 is likely to be severely biased.

To overcome this endogeneity problem, we use exogenous variation created by a class planning algorithm. Local Authorities use this tool to calculate the cost-minimizing number and type of classes, using a school's enrolment counts for each grade as inputs. In particular, the class planner takes into account that multi-grade classes can be used as means of reducing the number of classes that a school

Table 2.3: Self-Selection of Composite Class Pupils

	Prob(CompP1/P2) - First Graders			Prob(CompP1/P2) -Second Graders			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	0.004^{***}	0.004^{***}	0.003^{**}	-0.004^{**}	-0.004^{***}	-0.003^{**}	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
White	0.006^{*}	-0.004	-0.004	0.010^{***}	-0.002	-0.002	0.000
	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
Native English Speaker	0.015^{***}	0.016^{***}	0.015^{***}	-0.010^{*}	-0.009^{**}	-0.008^{**}	-0.002
	(0.005)	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.004)
Bottom 20% SIMD	-0.001	-0.003	-0.003	0.004	0.006^{***}	0.006^{***}	0.002
	(0.004)	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)	(0.003)
Age (in Years)	0.132^{***}	0.135^{***}		-0.103^{***}	-0.105^{***}		-0.104^{***}
	(0.006)	(0.006)		(0.006)	(0.006)		(0.007)
1st Age Quartile			-0.013^{***}			0.051^{***}	
			(0.002)			(0.004)	
3rd Age Quartile			0.027^{***}			-0.020^{***}	
			(0.003)			(0.003)	
4th Age Quartile			0.098^{***}			-0.028^{***}	
			(0.005)			(0.003)	
Low Literacy							0.029^{***}
							(0.004)
Low Numeracy							0.036^{***}
							(0.005)
Observations	190,704	190,704	190,704	203,139	203,139	203,139	139,198
R-squared	0.018	0.179	0.181	0.010	0.162	0.163	0.175
School FE	No	Yes	Yes	No	Yes	Yes	Yes

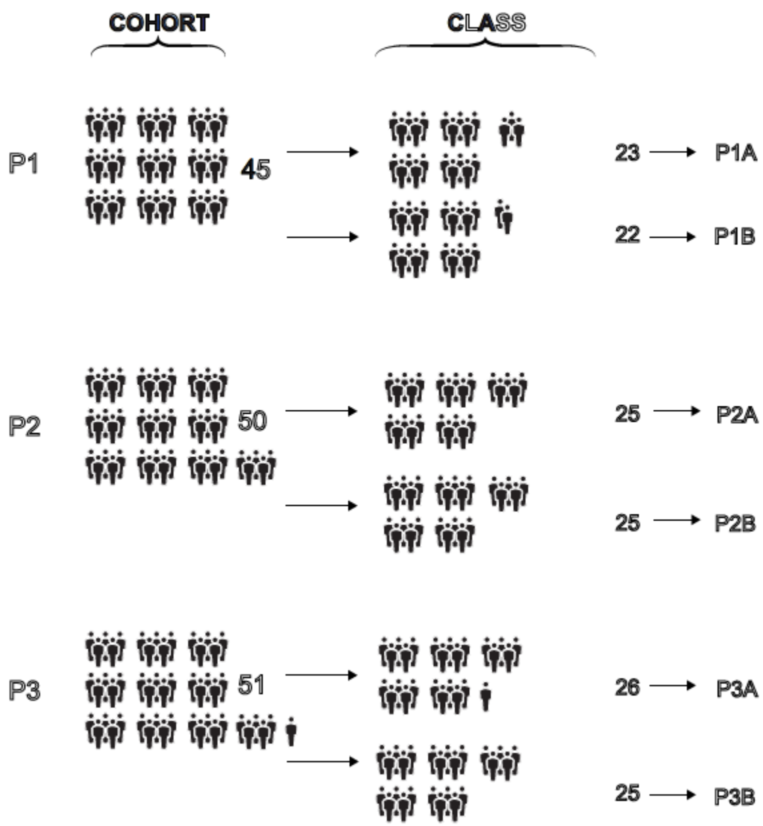
Notes: $***/**/*$ indicate significance at the 1%/5%/10%-level. Heteroscedasticity-robust standard errors adjusted for clustering at the school and year level are reported in parentheses.
This table regresses a dummy indicator for whether a pupil is part of a P1/P2 composite class on pupil characteristics. The first three columns show the results for first-graders who form the bottom component of a P1/P2 composite class. Columns (4) through (7) show our results for second graders who form the top component of a P1/P2 composite class.
Note that only P1 pupils from our main sample (with valid assessment data) are used. In column (7) only P2 pupils for whom P1 assessments (from previous year) were available, are part of the sample.
Low Literacy and Low Numeracy, respectively, indicate that P2 pupils scored below early level when in first grade.

needs to create, considering maximum class-size rules and ensuring that each grade contributes at least five pupils to a multi-grade class (if it is optimal to create one).

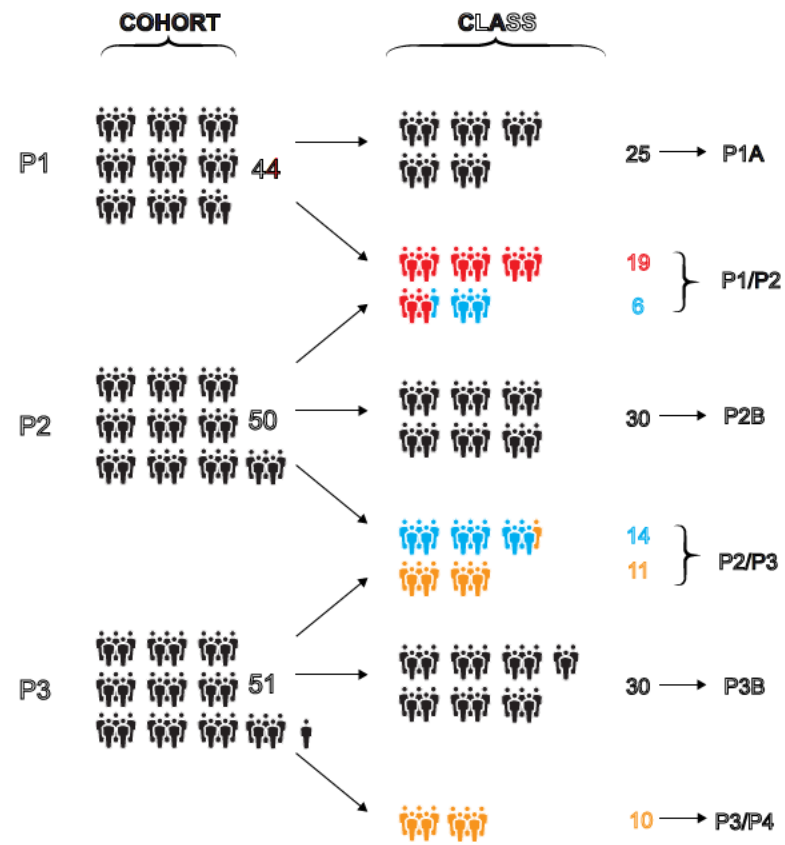
To illustrate our source of identifying variation, Figure 2.2a shows the optimal allocation – as predicted by the class planner – for one of the schools in our sample. Enrolment counts for all seven grades are in the high 40s or low 50s, as is typical in the average school. For illustrative purposes, we zoom in on the bottom three grades. The class planner here determines that the optimal allocation is to create two single-year classes for each grade. Figure 2.2b, on the other hand, shows the optimal allocation, as calculated by the class planner, for a case which is identical to the one in Figure 2.2a except that there are now 44 instead of 45 pupils enrolled in first grade. This marginal change triggers several multi-grade classes across different stages, and the suggested reallocation ultimately saves one classroom in a higher grade. This example illustrates that marginal changes in enrolment counts in any grade may trigger multi-grade classes and reshuffle pupils into different class types across all grades. As a result, pupils are quasi-randomly exposed to peers from either the same or older/younger age groups. We use the predictions of the algorithm as an instrument for the class status of each pupil. In its simplest form, we instrument $Comp_{cgst}$ with an indicator for whether the class planner suggests that grade g should contribute to a multi-grade class.

Figure 2.2: Class Planner Examples

(a) Class Planner Example - Scenario 1



(b) Class Planner Example - Scenario 2



Notes: This is an illustration of the allocations suggested by the class planner. In reality, enrolment counts for all seven primary school grades are fed into the class planner, for ease of interpretation we focus here on the bottom three grades of an anonymized primary school. We show two scenarios. The only difference between both scenarios is that in scenario 1 (on the left) this school has an enrolment count of 45 first graders, whereas in scenario 2 (on the right), there are 44 first graders enrolled. As is apparent from the figure, this marginal difference leads to fundamentally different class planner predictions. In scenario 1, none of the pupils is assigned to a composite class (i.e. $Comp^{pred}_{gst} = 0$), in scenario 2 all grades are assigned to treatment.

One key identifying assumption in our empirical setup is that of a strong first stage. Local Authorities use the class planner tool to allocate teaching resources to schools based on enrollment counts. Head teachers are not obliged to exactly follow the class allocation suggested by the class planner. However, given that they only receive the resourcing commensurate to the number of classes predicted by the class planner their ability to deviate from class planner suggestions is limited. We analytically assess compliance and thus the strength of our instrument by running a standard first stage regressions corresponding to the following equation:

$$Comp_{icgst} = \alpha_0 + \alpha_1 Comp_{gst}^{pred} + \gamma X_{igst} + \delta_s + \tau_t + \varepsilon_{icgst} \tag{2.2}$$

Where $Comp_{icgst}$ is a dummy indicator for whether class c in grade g which contains pupil i, is a multi-grade class whereas $Comp_{gst}^{pred}$ is an indicator for whether, according to the class planner, grade g should contribute to a multi-grade class, thus exogenously boosting the probability that pupils in this grade end up in a multi-grade class. As the main focus of our analysis are pupils in first grade of primary school (P1), we estimate the following variation of Equation 2.2:

$$CompLow_{ic,1,st} = \alpha_0^1 + \alpha_1^1 CompLow_{1,st}^{pred} + \gamma X_{i,1,st} + \delta_s + \tau_t + \varepsilon_{ic,1,st} \tag{2.3}$$

Where $CompLow_{ic,1,st}$ indicates whether class c in grade 1 which contains pupil i, constitute the bottom part of a multi-grade class, whereas $CompLow_{1,st}^{pred}$ indicates whether grade 1 should contribute, as the 'bottom grade', to a multi-grade class based on class-planner prediction. Our analysis of first graders allows us to isolate the effects of exposure to more experienced P2 peers. In our main specification, we therefore redefine our treatment dummy variable, $Comp_{icgst}$, as a continuous variable that measures the number of peers from the preceding cohort of second graders, $P2Peers_{icgst}$, who share multi-grade classroom c with pupil i. Therefore, alongside Equation 2.3, we estimate:

$$P2Peers_{ic,1,st} = \alpha_0^2 + \alpha_1^2 CompLow_{1,st}^{pred} + \gamma X_{i,1,st} + \delta_s + \tau_t + \varepsilon_{ic,1,st} \tag{2.4}$$

Instrumental variable regressions, while consistent, always yield biased estimates, even if all identifying assumptions are met. Bound et al. (1995) show that weak instruments may massively exacerbate this finite sample bias that is inherent to Two-Stage-Least-Squares (2SLS) instrumental variable estimation. A common indicator of instrument strength is the first-stage F-statistic which is typically assessed against a cut-off (Stock and Yogo, 2002). Recent work by Lee et al. (2020) suggests that in order to achieve valid estimation parameters, an F-statistic of larger than 104 is required. Our estimations of Equation 2.3 and Equation 2.4 are shown in columns (1) and (2) of Table 2.4. In column (1) we regress multi-grade status on an indicator of whether the class planner predicts P1 to be the bottom grade, conditional on covariates, school and year fixed effects. Column (2) repeats the same exercise, but uses the number of older (P2) peers as a dependent variable. We also estimate variations of Equation 2.2, separately for grade 4 and 7, i.e. P4 and P7. We employ two versions of our instrument $Comp_{gst}^{pred}$, namely $CompLow_{gst}^{pred}$ and $CompUp_{gst}^{pred}$, based on whether the class planner predicts a specific grade to contribute toward the lower or upper part of the age distribution of a multi-grade class, respectively. Naturally, grades 1 and 7 can uniquely contribute toward the bottom and top grade in a multi-grade class respectively, whereas P4 pupils can either share the class with older or younger pupils. For instance, in columns (3) and (4), we use the class-planner to predict the probability that a fourth-grader becomes part of the bottom or top part of the age distribution of a multi-grade class respectively. Similarly, in columns (5) and (6) we look at the predictions on the the number of younger and older peers a fourth-graders can share the classroom with.

Table 2.4 indicates a strong first stage for our sample of first-graders with F-statistics of 368 and 556, respectively. These are also displayed at the bottom of our second stage results Table 2.5 in Section 2.4. All F-statistics are heteroscedasticity and autocorrelation consistent (HAC) and were obtained using the method developed by Kleibergen and Paap (2006). Note that by contrast, our first-stage results for P4 and P7 (the only other two stages with outcome data) are well below any $F > 104$ threshold. We therefore focus our analysis on first-graders but report our results for fourth and seventh-graders for completeness in Table B.2.

There are several reasons why we might expect that our instrument would be stronger for lower grades compared to later grades. The main driver is the way the class planner is set up. The most cost-efficient pupil allocation provided by the algorithm is not always a unique solution. The class planner

Table 2.4: First Stage Results

	First Graders (P1)		Fourth Graders (P4)				Seventh Graders (P7)	
	(1) Bottom Comp.	(2) Older Peers	(3) Bottom Comp.	(4) Top Comp.	(5) Younger Peers	(6) Older Peers	(7) Top Comp.	(8) Younger Peers
$CompLow^{pred}_{gst}$	0.087***	1.038***	0.011***	−0.006	0.006	0.231***		
	(0.005)	(0.044)	(0.004)	(0.005)	(0.039)	(0.040)		
$CompUp^{pred}_{gst}$			0.005	0.021***	0.265***	−0.014	0.018***	0.212***
			(0.005)	(0.005)	(0.041)	(0.041)	(0.005)	(0.050)
Observations	190,704	190,704	194,804	194,804	194,804	194,804	186,082	186,082
R-squared	0.191	0.151	0.243	0.246	0.139	0.136	0.290	0.217
No. of Schools	1437	1437	1428	1428	1428	1428	1435	1435
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Stat	368.2	556.5	4.838	4.918	12.61	12.82	12.91	17.99

Notes: ∗∗∗/∗∗/∗ indicate significance at the 1%/5%/10%-level. Heteroscedasticity-robust standard errors adjusted for clustering at the school and year level are reported in parentheses.
This table shows the results for our estimation of a first stage equation (2.2) in which we regress our endogeneous measures of class composition on our instruments which indicate whether a grade should contribute to a composite class.
Covariates include pupil age, sex, and ethnicity an indicator for whether pupil is from a neighborhood in bottom 20% of deprivation (SIMD), classize and grade enrolment counts (and its square), the size of the school, and the percentage of pupils in a school that are female, white British, native English speakers, and in the bottom 20% of deprivation respectively. All specifications contain a set of school and school-year fixed effects.
The reported F-statistic is heteroscedasticity and autocorrelation consistent (HAC) and was calculated using the method developed by Kleibergen and Paap (2006).

is coded to work sequentially through enrolment counts in each grade from P1 to P7 in calculating class allocations. It is thus more likely to suggest composite classes in earlier grades. This is also consistent with head teacher preference who may find pooling 5 and 6 year old pupils into a single classroom more appealing than pooling 11 and 12 year olds. After all, the former is just a continuation of the nursery/kindergarten setup, whereas the latter is a more discrete classroom composition break in pupils' primary school trajectory. In addition, head teachers may have concerns about the suitability of particular groups of pupils to learn effectively in mixed grade classrooms, and thus decide to not stick with a class planner suggestion. These suitability assessments are, however, much harder to make for school starters whose abilities schools have little information on.

Our identification strategy also requires our instrument to be exogenous and the exclusion restriction to hold.[6] Class planner predictions are ultimately generated by random population variation, making the exogeneity assumption credible. The exclusion restriction requires planner predictions to only

[6] It is unlikely that our research design features "defiers" which would violate the monotonicity assumption.

affect learning outcomes through class-type. While this assumption is not formally testable, planner predictions are in practice indeed only used to determine the number and types of classes. Moreover, random fluctuations in the enrolment counts for *any* grade may change planner predictions across all grades. It is, thus, not conceivable that head teachers or parents can manipulate enrolment counts in order to consciously trigger or prevent multi-grade classrooms in a specific grade. In addition, as composite class formation is impossible for parents to foresee, we confidently rule out sorting into schools on the basis of an anticipated placement in a composite class. Our instrument is, therefore, unlikely to be correlated with parent or school characteristics that have an independent effect on our outcome of interest.

Furthermore, we want to rule out an alternative channel through which class planner predictions might affect our outcome: class size. If multi-grade classes were deployed to save resources, one might worry that class sizes may be larger, on average, whenever there is an incentive to create a multi-grade class. While this might be plausible at the school level, it is not necessarily the case at the grade level. As we discuss in Subsection 2.4.2, class size in first grade is the same as for composite classes as it is for single-year classes. Nevertheless, we will be controlling for class size in all specifications.

Hence, β_1 of Equation 2.1 will yield a local average treatment effect (LATE). That is different from a population average treatment effect (ATE) for two reasons. First, head teachers may not always follow the suggestions of the class planner. Even though head teachers who do not stick with algorithmic suggestions face clear budgetary issues, we have outlined above that compliance, while strong, is not perfect. Second, while it is as good as randomly determined whether a *grade* contributes to a multi-grade class, the specific subset of *pupils* who, in turn, are assigned to such a multi-grade class is not a randomly selected sample.

The interpretation of our LATE hinges on who these "compliers" are. Table 2.3, for instance, shows that age is a strong positive predictor of attending a multi-grade class. The oldest pupils of a cohort are more likely to become the lower-grade component of a multi-grade class whereas the youngest members of a cohort are more likely to become the higher-grade part of a multi-grade class. The coefficients in Table 2.3 lack causal interpretation, but this pattern is consistent with insights from school officials and teachers who we consulted as part of our research. Other socio-economic characteristics are only weak predictors. For instance, girls are a mere 0.4 pp more likely to attend a P1/P2 than boys. Hence, the compliers in our study tend to be comparatively mature school-starters, but do not otherwise differ substantially from fellow school-starters in terms of observable background characteristics.

While age has an independent effect on attainment (Black et al., 2011), it is important to note that non-random selection of pupils who are taught in multi-grade classes does not induce bias into our estimated LATEs. It is what makes these effects "local". Indeed, our instrumental variable technique addresses exactly this selection issue. Intuitively, our identification strategy compares pupils who - by virtue of random variations in enrolment counts - end up in a multi-grade class with older peers, against pupils who would have ended up in a multi-grade class, had the enrolment count in their school-year just marginally differed from their actual enrolment count. While our LATE might thus not yield a universal average treatment peer effect, it is arguably more policy-relevant than the ATE. After all, we identify peer effects for those school starters who are, in practice, most likely to be exposed to second-graders by way of multi-grade classes.

A residual concern could be that, alongside the number of older peers, multi-grading also affects one's own relative age in classroom.[7] First, by comparing March-born first graders (older in own cohort) who end up in multi-grade, with March-born first-graders in single-year classes we are not holding within-class' age distribution constant. In fact, the latter are relatively older in their own classrooms, whilst the former are grouped with second graders. Hence, first graders 'lose' the relative age advantage when put in multi-grade classes. Therefore, our estimates might suffer from attenuation bias. However, we show in Section 2.4 that the effect we find is already large and in line with previous research, meaning it is unlikely that such a mechanism is masking larger gains in multi-grade classes. Second, there might be general equilibrium effects in schools with more than one class per grade, where multi-grading would also alter the age distribution in adjacent single-year classes, whose pupils will be more similar to each other in terms of age. If this were the case, one could worry that the control group would be indirectly affected by the treatment. One potential solution would be to compare the same

[7] A similar concern is raised by Bertoni and Nisticò (2023) in the context of ability tracking policies and in general there is a wealth of studies showing the importance of relative age and rank effect in school (see e.g Elder and Lubotsky, 2009; Black et al., 2011; Cascio and Schanzenbach, 2016; Peña, 2017; Ballatore et al., 2020; Fumarco and Schultze, 2020).

cohort pupils in single-year classes across schools with and without composite classes or limiting the control group to those pupils in single-year classes in schools that do not adopt multi-grading (within the same cohort). However, this approach might not be feasible for two reasons. First, it does not constitute a comparison among compliers necessary in an IV setup. Second, it would come at the cost of considerable sample size reduction. Therefore, we are not able to address this concern directly. Finally, in previous specifications (not presented here) we control for within-class and within-cohort age rank separately, and find that our estimates are virtually unchanged. This suggests that the extent to which multi-grading affects relative age in classroom and grade might be negligible.

2.4 Results

In this section we present our estimates for the effect of exposure to older, more (school-) experienced peers by way of multi-grade classes. For comparison, we report OLS estimates alongside 2SLS coefficients corresponding to Equation 2.1. All specifications control for individual pupil characteristics, time-variant school characteristics, school fixed effects, and school-year fixed effects. Standard errors are adjusted for clustering at the school and year level throughout.

2.4.1 Second Stage Results

Columns (2) and (3) of Panel A in Table 2.5 show that for first-graders, exposure to an additional older peer raises the probability of performing at level or better in numeracy by 0.8 to 1.1 percentage points. On average, P1/P2 classes contain about 10 P2 pupils, so this translates into an average increase of 9-11 percentage points for pupils attending a typical composite class (see columns (5) and (6)). These sizable effects stand in contrast to naïve OLS estimates in column (1) which indicate a precisely estimated zero effect. Panel B shows that our effects are slightly larger for literacy. Each P2 peer increases performance by 1.3 to 1.5 percentage points. The coefficients in both columns (2) and (3) are statistically significant at the 5% level. This translates into a 15-16 percentage point increase in the probability of performing at least at the expected level in literacy for pupils in a multi-grade class.[8]

Table 2.5: Second Stage Results - First Graders (P1)

Panel A: Numeracy - Performing at Least at Level	(1) OLS	(2) 2SLS	(3) 2SLS	(4) OLS	(5) 2SLS	(6) 2SLS
P2 Peers	0.001^{*}	0.008^{**}	0.011^{**}			
	(0.000)	(0.003)	(0.005)			
Composite				-0.002	0.091^{**}	0.108^{**}
				(0.004)	(0.037)	(0.054)
Class Size	0.002^{***}	0.001	0.006^{*}	0.002^{***}	0.001^{**}	0.005^{*}
	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.003)
Observations	190,704	190,704	190,704	190,704	190,704	190,704
No. of Schools	1,437	1,437	1,437	1,437	1,437	1,437
Class-Size Instrumented	No	No	Yes	No	No	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
F-Stat		556.5	212.6		368.2	190.3
Panel B: Literacy - Performing at Least at Level	(7) OLS	(8) 2SLS	(9) 2SLS	(10) OLS	(11) 2SLS	(12) 2SLS
P2 Peers	0.001^{**}	0.013^{***}	0.015^{**}			
	(0.000)	(0.004)	(0.007)			
Composite				0.003	0.159^{***}	0.153^{**}
				(0.004)	(0.046)	(0.067)
Class Size	0.002^{***}	0.001	0.004	0.002^{***}	0.002^{**}	0.002
	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.004)
Observations	190,704	190,704	190,704	190,704	190,704	190,704
No. of Schools	1,437	1,437	1,437	1,437	1,437	1,437
Class-Size Instrumented	No	No	Yes	No	No	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
F-Stat		556.5	212.6		368.2	190.3

Notes: $***/**/*$ indicate significance at the 1%/5%/10%-level. Heteroscedasticity-robust standard errors adjusted for clustering at the school and year level are reported in parentheses.
This table shows the results for our estimation of equation (2.1) by Ordinary Least Squares (OLS) and 2-Stage-Least-Squares (2SLS) regression. Our outcomes of interest are dummy indicators for whether a pupil performs at least at the expected level in numeracy or literacy, respectively. All results refer to our sample of first graders (P1).
Covariates include pupil age, sex, and ethnicity, an indicator for whether pupil is from a neighborhood in bottom 20% of deprivation (SIMD), grade enrolment counts and its square, the size of the school, and the percentage of pupils in a school that are female, white British, native English speakers, and in the bottom 20% of deprivation, respectively. All specifications contain a set of school and school-year fixed effects.
The reported first-stage F-statistic is heteroscedasticity and autocorrelation consistent (HAC) and was calculated using the method developed by Kleibergen and Paap (2006).

While our 2SLS estimates are large, they are in line with the previous literature. For instance, Leuven and Rønning (2014) find that multi-grade classes in Norway increase younger pupils' performance by 0.4 standard deviations, and Barbetta et al. (2019) document educational gains for sharing a class with

[8]Columns (1) and (2) of Table B.1 report the reduced form estimates. The ratio of our reduced-form and first stage effects is approximately equal to our 2SLS coefficients.

older peers of as much as 0.33 standard deviations. Our point estimates suggest improvements of 0.28 standard deviations for numeracy and 0.35 standard deviations for literacy. By way of comparison, these gains are large enough to close the attainment gap between the average pupil and a pupil in one of the 20% most deprived data zones in Scotland.

We also find no evidence that gains for first graders come at the expense of lower attainment among their second grade peers. The second stage results in columns (2) and (4) of Panel A in Table 2.6 indicates a small negative effect on maths assessments of second graders who shared a multi-grade classrooms with first-graders. However, the point estimate is not statistically significant at any reasonable level of significance. The standard errors are also small enough to rule out effects that are large enough to offset the gains to first graders. In the same vein, we find small statistically insignificant negative effects on second graders' literacy (see columns (6) and (8)). This is in contrast to OLS estimates (columns (5) and (7)) which suggest statistically significant detrimental effects, but which are biased due to negative selection of P2 pupils into P1/P2 multi-grade classes. One caveat here is that second-graders are not assessed in the same year that they share a classroom with first graders, but only once they get to fourth grade. Hence, the main takeaway from Panel A of Table 2.6 is that there is no evidence for medium-term adverse effects of P1/P2 multi-grade classes on those second graders who are placed in these classes.

Panel B of Table 2.6 shows our results for first-graders' performance once they have progressed to fourth grade (P4). Of course, pupils are subject to a variety of other influences as they progress from P1 to P4, all of which may amplify or mitigate the effects of starting school in a multi-grade class. This makes it challenging to identify a clean and precisely estimated effect of P1/P2 attendance on attainment in P4. The OLS estimates for multi-grade status in first grade are all positive, reflecting the positive selection of P1s into P1/P2 composite classes. However, our 2SLS estimates document that once they have progressed to fourth grade, there is no statistically significant difference in attainment between pupils who shared a classroom with second graders when they were in first grade and those who were in single-year groupings. In other words, the attainment gains shown in Table 2.5 appear to fade out over time. This pattern can partly be explained by the transitory nature of composite classes. Only about 22% of pupils who were in a multi-grade class in the previous school year remain in a multi-

Table 2.6: Second Stage Results - Performance in Fourth Grade (P4)

Panel A: Performance of Second Graders (P2) in Fourth Grade (P4)	Numeracy				Literacy			
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS
P1 Peers	-0.006^{***}	-0.002			-0.006^{***}	-0.003		
	(0.001)	(0.004)			(0.001)	(0.004)		
Composite			-0.054^{***}	-0.020			-0.052^{***}	-0.024
			(0.006)	(0.039)			(0.006)	(0.042)
Class Size	0.002^{***}	0.003^{**}	0.002^{***}	0.003^{**}	0.002^{***}	0.003^{**}	0.002^{***}	0.003^{**}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	194,666	194,666	194,666	194,666	194,666	194,666	194,666	194,666
No. of Schools	1449	1449	1449	1449	1449	1449	1449	1449
F-Stat		346.7		282.5		346.7		282.5

Panel B: Performance of First Graders (P1) in Fourth Grade (P4)	Numeracy				Literacy			
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS
P2 Peers	0.004^{***}	-0.005			0.004^{***}	-0.000		
	(0.000)	(0.004)			(0.000)	(0.004)		
Composite			0.032^{***}	-0.053			0.036^{***}	-0.005
			(0.004)	(0.046)			(0.004)	(0.049)
Class Size	0.001	0.001^{**}	0.001^{*}	0.001^{**}	0.001^{**}	0.001^{**}	0.001^{***}	0.001^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	192,428	192,427	192,428	192,427	192,428	192,427	192,428	192,427
No. of Schools	1443	1443	1443	1443	1443	1443	1443	1443
F-Stat		442.7		305.5		442.7		305.5

Notes: $***/**/*$ indicate significance at the 1%/5%/10%-level. Heteroscedasticity-robust standard errors adjusted for clustering at the school and year level are reported in parentheses.

This table shows the results for our estimation by Ordinary Least Squares (OLS) and 2-Stage-Least-Squares (2SLS) regression. In Panel A, our outcomes of interest are dummy (0/1) indicators for whether a second grader (P2) performs at least at the expected level in numeracy or literacy two years later in fourth grade (P4). In Panel B it is the same measure but for first-graders (P1) when assessed in P4. The explanatory variable measures the number of younger P1 peers or older P2 peers a pupil was exposed to by way of a P1/P2 composite class.

All specifications include covariates for pupil age, sex, and ethnicity, an indicator for whether pupil is from a neighborhood in bottom 20% of deprivation (SIMD), grade enrolment counts and their squared values, the size of the school, and the percentage of pupils in a school that are female, white British, native English speakers, and in the bottom 20% of deprivation, respectively. All specifications also contain a set of school and school-year fixed effects.

The reported first-stage F-statistic is heteroscedasticity and autocorrelation consistent (HAC) and was calculated using the method developed by Kleibergen and Paap (2006).

grade class the year after. This is because the class planning algorithm is sensitive to small changes in enrolment which can trigger a reshuffling of pupils into single-year and composite classes every year. After first-grade, pupils may be grouped with either older or younger peers, which makes these dynamics hard to model, but Panel B of Table 2.6 suggests that this lack of persistence in peer effects could drive a medium-run regression to the mean.

2.4.2 Mechanisms and Heterogeneity

So far, we have said little about the mechanisms that might underpin the large, statistically significant short-run effects of multi-grade classes that we set out in the previous section. Here we explore six potential explanations and describe what our analysis tells us about each: the role of class size, breaks in peer groups and social stigma, whether the type of activity assessed reveals anything about the mechanism, potential socio-economic channels, whether there might be additional staffing support and resources, and gender composition effects.

Throughout our analysis, we control for class size and report the corresponding regression output. In all tables it has been noticeable that the effect of class size tends to be both statistically and economically insignificant in virtually all specifications. Similar to Leuven et al.'s (2008) study of Norwegian middle schools, the class size coefficients in Table 2.5 are very small and positive, range from 0.001 to 0.006 depending on the specification, and are often not statistically significant at the 5% level. This is not surprising as both single-year P1 and multi-grade P1/P2 classes are capped at 25 pupils and consequently have virtually identical average class sizes (21.8 for single-year, 22.0 for composite classes). It is thus unlikely that class size is driving these positive effects.[9]

Note that our analysis focuses on school starters. That makes it unlikely that *breaks* in peer groups are driving our results. Scottish primary schools do not typically have kindergarten grades but take in first-graders from a variety of smaller day-cares. While school and social networks are clearly important (see Bramoullé et al. (2009), Crosnoe et al. (2003), De Giorgi and Pellizzari (2014), De Giorgi et al. (2010), Lavy and Sand (2019), Patacchini et al. (2017)), they are only beginning to form in first grade. In the same vein, it is unlikely that stigma or feelings of inferiority (or superiority) are driving our results. Five-year old school starters will have no reference point for their experienced class structure.[10]

Our finding that there are larger and more pronounced gains for literacy compared to numeracy suggests that the type of activity being assessed might shed some light on potential mechanisms. Ta-

[9]We also deployed an instrumental variable strategy in which class size is instrumented by class size predictions that are obtained exploiting maximum class size cutoffs (see columns (3) and (6) of Table 2.5. This identification is in the mould of Angrist and Lavy's (1999) seminal work and their recent follow-up study (Angrist et al., 2019). Appendix B elaborates on this approach.

[10]Stigma and loss of social networks may, of course, play a more important role in the case of second-graders who are chosen to attend P1/P2 classes.

Table 2.7: Second Stage Results (P1) for Literacy Subcategories

	Reading		Writing		Listening & Talking	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS	OLS	2SLS	OLS	2SLS
P2 Peers	0.001^{*}	0.008^{**}	0.001	0.012^{***}	0.000	0.004
	(0.000)	(0.003)	(0.000)	(0.004)	(0.000)	(0.003)
ClassSize	0.002^{***}	0.001	0.002^{***}	0.001	0.002^{***}	0.002^{**}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	190,704	190,704	190,704	190,704	190,704	190,704
No. of Schools	1,437	1,437	1,437	1,437	1,437	1,437
School FE	Yes	Yes	Yes	Yes	Yes	Yes
F-Stat		556.5		556.5		556.5

Notes: $***/**/*$ indicate significance at the 1%/5%/10%-level. Heteroscedasticity-robust standard errors adjusted for clustering at the school and year level are reported in parentheses.
This table shows the results for our estimation of equation 2.1 by Ordinary Least Squares (OLS) and 2-Stage-Least-Squares (2SLS) regression. Our outcomes of interest are dummy indicators for whether a pupil performs at least at the expected level in three subcategories of literacy. All results refer to our sample of first graders (P1)
Covariates include pupil age, sex, and ethnicity an indicator for whether pupil is from a neighborhood in bottom 20% of deprivation (SIMD), grade enrolment counts and its square, the size of the school, and the percentage of pupils in a school that are female, white British, native English speakers, and in the bottom 20% of deprivation respectively. All specifications contain a set of school and year fixed effects.
The reported first-stage F-statistic is heteroscedasticity and autocorrelation consistent (HAC) and was calculated using the method developed by Kleibergen and Paap (2006).

ble 2.7 breaks down our literacy assessment into its three components: reading, writing, and listening & talking. These subcategories may offer pointers on the channel through which exposure to more mature peers improves literacy. While listening and talking are – by definition – interactive activities, reading and writing can be improved by working on one's own. Columns (2) and (4) show that the gains appear to be concentrated in improvements in reading and writing ability respectively, whereas the effect for listening and talking (column (6)) are smaller and not statistically significant at the 5% level. While this breakdown does not allow us to fully disentangle these mechanisms, it suggests that it is not the direct interaction with older peers that is driving these improvements. Instead younger pupils may be motivated and spurred on by observing peers who are have already acquired reading and writing proficiency.

Table 2.8: Second Stage Results (P1): Effect Heterogeneity

	Numeracy				Literacy			
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS
Panel A: Heterogeneous Effects by Level of Deprivation								
	Top 60% SIMD		Bottom 40% SIMD		Top 60% SIMD		Bottom 40% SIMD	
P2 Peers	0.001^{**}	0.006^{*}	0.000	0.009^{*}	0.001^{**}	0.011^{**}	0.001	0.016^{***}
	(0.000)	(0.003)	(0.000)	(0.005)	(0.000)	(0.004)	(0.001)	(0.006)
Observations	106,653	106,653	84,051	84,051	106,653	106,653	84,051	84,051
No. of Schools	1411	1411	1269	1269	1411	1411	1269	1269
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Stat		376.9		338.5		376.9		338.5
Panel B: Heterogeneous Effects by School Location								
	Urban		Rural		Urban		Rural	
P2 Peers	0.001^{**}	0.008^{**}	-0.000	0.005	0.001^{**}	0.012^{***}	0.001	0.017^{**}
	(0.000)	(0.004)	(0.001)	(0.006)	(0.000)	(0.005)	(0.001)	(0.008)
Observations	143,834	143,834	46,870	46,870	143,834	143,834	46,870	46,870
No. of Schools	972	972	486	486	972	972	486	486
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Stat		424.8		131.2		424.8		131.2
Panel C: Heterogeneous Effects by Pupil Sex								
	Boys		Girls		Boys		Girls	
P2 Peers	0.000	0.009^{**}	0.001^{**}	0.006	0.001	0.016^{***}	0.001^{***}	0.011^{**}
	(0.000)	(0.004)	(0.000)	(0.004)	(0.001)	(0.005)	(0.000)	(0.004)
Observations	97,125	97,125	93,579	93,579	97,125	97,125	93,579	93,575
No. of Schools	1435	1435	1435	1435	1435	1435	1435	1435
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Stat		479.7		489		479.7		489

Notes: $***/**/*$ indicate significance at the 1%/5%/10%-level. Heteroscedasticity-robust standard errors adjusted for clustering at the school and year level are reported in parentheses.

This table shows the results for our estimation of equation (2.1) by Ordinary Least Squares (OLS) and 2-Stage-Least-Squares (2SLS) regression. Our outcomes of interest are dummy indicators for whether a pupil performs at least at the expected level in numeracy or literacy, respectively. All results refer to our sample of first graders (P1).

Unless they are the category of interest, covariates include pupil age, sex, and ethnicity, an indicator for whether pupil is from a neighborhood in bottom 20% of deprivation (SIMD), grade enrolment counts and its square, the size of the school, and the percentage of pupils in a school that are female, white British, native English speakers, and in the bottom 20% of deprivation respectively. All specifications contain a set of school and school-year fixed effects.

The reported first-stage F-statistic is heteroscedasticity and autocorrelation consistent (HAC) and was calculated using the method developed by Kleibergen and Paap (2006).

Another mechanism that might help to explain our results is if parents invest more effort into supporting their children if they end up in a multi-grade class. More generally, there is growing evidence suggesting that parents from lower socio-economic strata may provide less educational input to their offspring (Francesconi and Heckman 2016; Fredriksson et al. 2016). Multi-grade classes may exacerbate these inequalities if gains for first graders are driven by greater investment by affluent parents. While we cannot directly measure parental effort, we can explore whether there are differences in our results across socio-economic status. Panel A of Table 2.8 indicates few such differences. In fact, our point estimates suggest that pupils from postcodes which are ranked in the two bottom quintiles in terms of deprivation tend to benefit slightly more from exposure to more experienced peers than pupils in the top three quintiles. However, these differences are not significant at any reasonable level of statistical significance.

In discussions with educational decision makers in Scotland it became clear that there is neither special training, nor additional support for teachers who are in charge of multi-grade classes. That is because teaching approaches in primary school do not differ between single-year and multi-grade classes. In both setups, group-based teaching is the norm. At the beginning of the school year, teachers evaluate students' cognitive abilities and split them into groups accordingly, often seated separately within the same classroom. Throughout the year, teachers would teach each group separately and then assign different tasks to each of them. Teachers periodically evaluate students' learning progress and re-arrange groups accordingly. In other words, teaching approach is the same across multi-grade and single-year classes –which is also why multi-grading is such a dominant feature of this system. Put differently, all teachers in Scotland teach some sort of 'composite' class. The key difference is that in actual multi-grade classes P1 pupils may share a table with P2 pupils. First-graders in regular and multi-year classes are also taught the same curriculum. This not only means that P1 pupils are not exposed to more advanced material in multi-grade classes than they would be in a single-year class. It also means that their older class mates should not experience any deviations from the curriculum. This is a potential concern raised by Checchi and De Paola (2018) in motivating the negative effects they find for pupils in the final stages of primary school.

Nevertheless, it might be the case that more teaching resources are provided to the teaching of multi-grade classes and that this helps explain our findings. We explore this dimension as far as we can given the data available. While individual teachers cannot be identified in our data, Table B.3 shows that there are no differences in terms of staffing (e.g. presence of teaching assistants or additional teachers). Furthermore, urban schools tend to find teacher recruitment easier and are on average larger which may make them more likely to develop teachers who specialise in the instruction of multi-grade classes. But again, Panel B of Table 2.8 reveals little in the way of effect differences between urban and rural schools.

Finally, we stratify our sample by pupil gender. There is an extensive literature that shows gender differences (see e.g. Lavy et al., 2012b). Panel C of Table 2.8 shows that we cannot reject the null hypothesis of no gender differences, even though our point estimates for boys tend to be larger than those for girls in both literacy and numeracy.

2.5 Conclusion

This study explores the impact of sharing a multi-grade classroom with more experienced peers in early primary school. We combine population-level pupil data with an instrumental variables estimation strategy that exploits exogenous variation in the creation of multi-grade classes generated by a class planning algorithm. We find that the presence of second graders improves first-graders' reading, writing, and maths performance, as measured by teacher assessments that are informed by standardized test scores. It is important to note that we estimate a local average treatment effect (LATE). That is, these benefits may not accrue to the average school-starter but only to the oldest cohort members who - if assigned to multi-grade classes - are typically exposed to second-graders by way of a multi-grade classes. While these effects wash out over time, we also find no evidence of a detrimental impact of the classroom presence of younger first-graders on those second-graders who make up the older component of multi-grade classes.

Our paper adds to two strands of literature. First, our findings are consistent with, and generalize beyond, recent research on multi-grade classes that exploited that small population variations in sparsely populated areas of Norway (Leuven and Rønning 2014) and Italy (Checchi and De Paola 2018; Barbetta et al. 2019) lead to the lumping together of grades in rural middle and elementary school respectively. We show that the benefits of exposure to older pupils by way of a multi-grade class, also accrue in urban settings where multi-grade classes are created by design and where school-starters are placed in multi-grade classes often for only one year at a time. While further research in this area is certainly warranted, the overall body of evidence suggests that multi-grade classes, especially in the early years of primary education, have the potential to be a useful tool to stimulate the learning of academically strong and relatively mature pupils by exposing them to older, more experienced peers.

Second, we contribute to an important literature on peer effects. We demonstrate that first graders benefit from exposure to more mature peers with an additional year of primary schooling under their belt. Our research thus re-enforces the common finding that externalities from peers are important determinants of pupil attainment. In fact, our study suggests that these spillovers are more important than conventional education production inputs, such as class size. As such, our findings also have important implications for policymakers and education practitioners. Our study suggests that multi-grade classes deliver better learning outcomes for first-graders while simultaneously acting as a way for policymakers to allocate resources more efficiently.

Chapter 3

DOES THE PROVISION OF UNIVERSAL FREE SCHOOL MEALS IMPROVE SCHOOL ATTENDANCE AND BEHAVIOUR?

3.1 Introduction

Free school meals (henceforth, FSM) have recently been at the centre of a heated debate following disruption to their provision caused by Covid-19 school closures. This affected approximately 300 million children around the world and 1.6 million in the UK. For most of these children, school meals represent a crucial dietary component.[1] In recent years, many high-income countries have moved from means-tested programmes, in which eligibility is contingent on households' financial circumstances, to school-wide or even school system-wide universal provision. For instance, in 2014 the US launched the Community Eligibility Provision (CEP), which subsidises schools with at least 40% of pupils eligible for FSM to extend the provision to all their students. Around the same time, nations of the United Kingdom implemented similar policies, following a series of pilot schemes. England launched the Universal Infant Free School Meals (UIFSM) programme for all primary school children in year 1 and 2 in September 2014, followed by Scotland in January 2015, where eligibility was extended to pupils in the first three years of primary school.

The motivation for these policies are deep-rooted. In Sorhaindo and Feinstein (2006)'s review the authors identify four channels explaining the link between a healthy diet and school outcomes. First, deficiency of nutrients, such as iron, could affect cognitive development in children.[2] Second, a healthy diet can reduce the risk of illnesses and thus school absences. The third and fourth channels are behaviour and school life. Evidence from the medical literature suggest that a healthy diet should include nutrients which can reduce the risk of violent and anti-social behaviour as well as hyperactivity (Benton, 2007). In addition, FSM uptake can be associated with stigma and consequently victimization, both of which may be mitigated by making provision universal.

A broad literature documents the beneficial effects of school meal programmes (whether lunch or breakfast) on a variety of outcomes, and finds that these are indeed linked to gains in academic performance (Belot and James, 2011; Chakraborty and Jayaraman, 2019; Schwartz and Rothbart, 2020; Gordanier et al., 2020; Ruffini, 2021), reduction in body-mass index (Holford and Rabe, 2022), improve-

[1] See the contribution to the Economics Observatory, Link: https://www.economicsobservatory.com/how-coronavirus-affecting-provision-free-school-meals

[2] There is extensive evidence in the economic literature on the positive effect of nutrition on academic attainment (see e.g. Glewwe et al., 2001; Winicki and Jemison, 2003; Alderman et al., 2006; Victora et al., 2008).

ments in labour market outcomes (Bütikofer et al., 2018; Lundborg et al., 2021), but also households' finances and nutrition (Bhattacharya et al., 2006; Handbury and Moshary, 2021; Marcus and Yewell, 2021; Ozturk et al., 2021).[3]

Nonetheless, FSM implementation continues to be a debated topic, both among policy-makers and academics. One counter-argument is that FSM programmes do not improve children's health (Corcoran et al., 2016) and can even raise the risk of unhealthy body weight (Schanzenbach, 2009; Polonsky et al., 2019; Abouk and Adams, 2022). Clearly, this depends on the quality of the foods served, as well as the degree of substitutability/complementarity between school and home food. For instance, Belot and James (2011) and Bütikofer et al. (2018) argue that gains within the education space occur when school meal policies specifically target the nutritional content of the meals. In addition, Scholder (2013) finds that two reforms meant to restrict access to FSM in the UK resulted in a drop of school meals consumption by nearly 40%, suggesting a considerable substitution between school meals and home-packed lunch. Hence, one would expect further benefits from school meals when these complement or substitute relatively lower quality home food (Holford and Rabe, 2022).

However, the core of the debate relates to the trade-off between the cost of their universal implementation, and their effectiveness when they are means-tested. In fact, stigma attached to FSM status might be an explanation for the imperfect uptake rates observed even for means-tested provision. In such a case, FSM expansion can raise take-up not just among previously ineligible students but also among previously eligible students (Leos-Urbel et al., 2013; Holford, 2015; Corcoran et al., 2016; Ruffini, 2021).

In spite of the emphasis that has been put on the link between FSM and academic achievement, the evidence remains rather mixed and mostly confined to short-term outcomes. Existing studies largely overlooked the effect that school meals programmes have on students' outcomes which are more likely to shape their cognitive and non-cognitive abilities in the long-run. To fill this gap in the literature we look at the effect of Scotland's Universal Free School Meals (UFSM) programme on an array of student behaviours, such as attendance, detentions as well as health-related absences. The intervention extended the eligibility for free lunches to *all* pupils in the first three (P1, P2 and P3) of the seven grades

[3] In addition, see Cohen et al. (2021) for a systematic review of the literature on the link between universal free school meals and student performance, attendance, diet quality and body mass index.

of primary school, regardless of their families' financial circumstances.

For our analysis we use a panel of primary schools over more than ten years. We observe the percentage of pupils registered for free meals, alongside take-up rates, both among previously and newly eligible pupils. In our setup all primary schools in the nation are 'treated' at once, thus we implement a difference-in-differences estimation strategy with variation in the intensity of treatment in the same fashion as, for example, Card (1992) and Clemens et al. (2018). In particular, we leverage different levels of exposure to the policy, as determined by variation in the fraction of the school population taking FSM prior to the change in policy. We find that schools with a higher fraction of non-FSM takers (high exposure to the policy) experienced an increase in uptake of about 10 percentage points more than their lower exposure counterparts. We match these schools to their attendance and absence records, alongside a rich set of school characteristics.

Our study provides two main contribution to the literature and in turn the policy debate. First, Scotland is a nation with a high prevalence of childhood obesity. In 2016, approximately 29% of children in Scotland were at risk of becoming overweight, and about 14% were at risk of obesity (Public Health Scotland, 2021).[4] These figures are in line with the rest of the UK, and with countries like Cyprus, Greece, Italy and Spain, where between 18% and 21% of boys are obese, but far above Denmark, France, Ireland, Latvia and Norway where these figures amount to up to 9% (World Health Organization and others, 2018).[5] In fact, this policy is the culmination of a wider range of interventions and reforms aimed at regulating food standards in schools as well as investing schools with the role of health educators, on the same basis as mathematics and languages. Parnham et al. (2022) find that UFSM led to some improvement in pupils' diets.

Second, the outcomes we are interested in are strong correlates of the 'Big Five' personality traits, and thus of non-cognitive skills, which have been proven to be powerful predictors of adult life out-

[4]The likelihood of being at risk of obesity is based on the classification of obesity risk categories for children by Public Health Scotland. This classification is determined by children's BMI values as compared to reference values based on the expected 'healthy' BMI for each age and gender group. Each children's health status is then based on the probability of having a BMI lower (or higher) than the reference value. Since this comparison includes a probabilistic component, the health category is framed as a risk ('at risk of obesity' for children), as opposed to the discrete category of 'obese' for adults. More detail on these classifications can be found in Public Health Scotland (2021).

[5]Griffith et al. (2016) discuss the overall increase in obesity and document that, in England, between 1980 and 2013 the aggregate consumption of calories has declined, while physical activity has dropped by an even further extent. The authors argue that changes in life and work styles explain a large part of this phenomenon.

comes (see Heckman and Rubinstein, 2001; Chetty et al., 2011; Lindqvist and Vestman, 2011) even if not captured by test-scores (Jackson, 2018).[6] For instance, absenteeism is linked to worse academic attainment (Gottfried, 2009, 2011; Aucejo and Romano, 2016; Gershenson et al., 2017; Liu et al., 2021; Klein et al., 2022).

Alongside better nutrition, UFSM could reduce absenteeism (Belot and James, 2011; Leos-Urbel et al., 2013; Corcoran et al., 2016; Gordanier et al., 2020; Cuadros-Meñaca et al., 2022; Holford and Rabe, 2020). First, parents are incentivised to send their kids to school more often than they would absent the policy. Second, as the risk of stigma associated with FSM uptake reduces, school becomes a more attractive prospect for pupils. Furthermore, a better diet reduces the risk of illnesses and therefore school absences.

Finally, our study relates to school misbehaviour, the importance of which has been studied extensively in relation to mental health outcomes (see e.g. Bowes et al., 2015; Kim and Leventhal, 2008; Hertz et al., 2013), lower educational gains and peers' misbehaviour (Figlio, 2007b; Carrell and Hoekstra, 2010; Ammermueller, 2012; Ponzo, 2013; Eriksen et al., 2014; Ahn and Trogdon, 2017) in addition to lower schooling and future earnings (Brown and Taylor, 2008; Ammermueller, 2012; Carrell et al., 2018b). However, there is still little evidence on how FSM policies might mitigate these issues (Altindag et al., 2020; Gordon and Ruffini, 2021; Cuadros-Meñaca et al., 2021).

We present three main results. First, UFSM had only minimal impact on attendance and health-related absences. In particular, a 10 percentage points increase in FSM uptake translates into less than one school day gained. These results are precisely estimated. Second, we find no evidence that the policy improved student behavior. Third, the effect for attendance and health-related absences is slightly stronger within smaller schools and those with more resources per pupil, suggesting that uptake is strongly influenced by how efficiently the policy was rolled out.

Our analysis contributes to a wealth of studies, by bringing evidence from a new country. In particular, our work is closely related to Gordon and Ruffini (2021) and Altindag et al. (2020) but it departs

[6] Agreeableness, conscientiousness, neuroticism are associated with absences, tardiness and anti-social behaviour (see for example, John et al., 1994; Barbaranelli et al., 2003; Carneiro et al., 2007; Duckworth et al., 2007; Lleras, 2008; Jackson, 2018).

from these in two significant ways. First, relative to Gordon and Ruffini (2021) we are able to observe actual take-up rates. This is a substantial addition with respect to most of the literature which instead measures an 'intention-to-treat' (ITT) effect, where the predictor of interest is eligibility for a programme.[7] Similar to us, Altindag et al. (2020) use variation in take-up rates as surveyed on one day of the school year. By focusing on a variety of outcomes beyond infractions we are able to provide a more comprehensive investigation of the potential effects of FSM policies on non-cognitive skills. Our context, however, leads to different results. Gordon and Ruffini (2021) use the CEP, which as we outlined above, was meant to provide FSM to all pupils in schools above a certain threshold of eligibility. Therefore, they focus on contexts of deprivation, where arguably the effect of universal provision might be different. Similar to Altindag et al. (2020) we look at the effect of universal implementation. Like us, Altindag et al. (2020) leverage actual take-up rates, which might explain the larger effects they find relative to Gordon and Ruffini (2021) –35% versus up to 25%–, whom only estimate an ITT. The null result we find can be explained in two ways: first, unlike Altindag et al. (2020), UFSM only targeted three out of seven grades in primary school; second, the authors also focus on secondary schools, which are characterised by larger cross-sectional variation in school attendance and misbehaviour. In addition, our work contributes to an emergent literature on the effect of universal school meals on attendance (Leos-Urbel et al., 2013; Corcoran et al., 2016; Gordanier et al., 2020; Cuadros-Meñaca et al., 2022; Holford and Rabe, 2020) which find zero or small effects. On the other hand, Belot and James (2011) find a 14% decrease in authorised absences, which are mostly driven by illnesses, in response to an increase in school meals quality. Therefore, unlike most of the previous studies which cannot distinguish absences by type, we are able to observe the reason of each absence, and focus on health-related absenteeism, and unlike Belot and James (2011), we do this in an universal FSM scenario. Similarly to the other studies, the aim of our policy was to purely extend provision without altering the nutritional content of the meals. However, the positive, yet small, effect on attendance seems driven by a reduction in health-related absences, suggesting the policy might have improved dietary habits for some students. Unfortunately, unlike previous studies we are not able to run an heterogeneity analysis to check which type of students was more likely to benefit from the policy.

[7] Other work using this identification are (Belot and James, 2011; Chakraborty and Jayaraman, 2019; Gordanier et al., 2020; Gordon and Ruffini, 2021)

The remainder of the paper is structured as follows. Section 3.2 discusses the institutional background and UFSM implementation; Section 3.3 describes the data; Section 3.4 outlines the estimation strategy; Section 3.5 presents the results. Finally, Section 3.6 contains our concluding remarks.

3.2 Institutional Background

Over the last couple of decades, a series of reforms aimed at encouraging healthy eating habits in schools have taken place in Scotland. In particular, the launch of *Hungry for Success: A Whole School Approach to School Meals in Scotland (2003)* set national nutritional standards for school lunches. These endeavours have been subsequently formalised within the Schools (Health Promotion and Nutrition) Scotland Act (2007), which set out the responsibilities and duties of schools and Local Authorities in terms of health education and the administration of schools meals, and the Nutritional Requirements for Food and Drink in Schools (Scotland) Regulations (2008), which aligned the nutritional standards of all food and drink in schools with the Scottish Government's dietary goals for the population.

The first experience of FSM universality dates back to the school year 2007/08, when UFSM was piloted among P1-P3 pupils in five Local Authorities, namely East Ayrshire, Fife, Glasgow, Scottish Borders and West Dunbartonshire. These regions were chosen to cover a wide portion of the nation and also based on deprivation (MacLardie et al., 2008). As Holford (2015) documents, the trial was announced in the Summer 2007 to be launched in the following October, setting March 2008 as the initial deadline. It was subsequently extended until June, meaning the trial would ultimately cover the entire academic year. Prior to the launch of the trial, FSM entitlement was means-tested.[8] Thus, pupils eligible under the national criteria were able to take free school meals only upon registration, which is completed by parents by providing proof of eligibility to their Local Authority (Holford, 2015). During the pilot trial, on the other hand, all targeted pupils were eligible and automatically registered for free school meals regardless of their household's financial circumstances.

Starting in August 2010, Local Authorities launched a series of local initiatives aimed at increasing

[8]See https://www.gov.scot/publications/school-healthy-living-survey-statistics-2020/pages/4/, alongside Appendix C for more details.

eligibility among P1-P3 pupils. The goal was to promote healthy eating by stimulating take-up among pupils who would not otherwise be entitled. A 2011 report by the Scottish Government shows that free school meals registration increased by 10.3%, whereas the overall take-up (free or paid-for meals) increased only slightly.[9]

There are some important distinctions to be made at this point. Prior to UFSM implementation, eligibility alone, whether due to households financial circumstances or local initiatives, did not automatically entitle students to claim a free school meal from the school canteen. This was still contingent on registration. Similarly, registration did not guarantee that students were indeed taking free school meals every day.

The launch of Better Eating, Better Learning (2014) set out the government intention to make healthy eating habits a pillar of education in Scotland. It was paired with the introduction of the new national 'Curriculum for Excellence', which includes Health and Wellbeing as one of the eight curricular areas, alongside, for instance, mathematics, languages and sciences. Starting from January 2015, all P1-P3 pupils in Scotland would become eligible and were automatically registered for free school meals. The policy, which entailed £70.5 million funding from the Scottish Government to Local Authorities over the following two years, was estimated to provide households with financial savings of approximately £380 per child per year while also providing nutritional benefit to children (Beaton et al. (2014) and McAdams (2016)). A clearly declared goal of the policy was to reduce health inequalities.

According to official statistics, the number of FSM registrations in primary schools increased by 135,408 compared to the previous year, to a total of 213,199 pupils. This corresponds to 55.3% of the primary school population, compared to 20.6% in 2014, and is nearly entirely attributable to the change in policy. In fact, this roughly corresponds to the number of FSM-unregistered P1-P3 pupils the year before the policy change.[10] In terms of uptake, the fraction of pupils taking a meal (free or paid for)

[9]For more details, see `https://www.webarchive.org.uk/wayback/archive/20180518073257mp_/http://www.gov.scot/Resource/Doc/920/0119410.pdf`

[10]National Statistics (2015). If we assume that the fraction of FSM-registered pupils is roughly unchanged between P1-P3 and P4-P7, a back of the envelope calculation based on the formula $P(FSM) = P(FSM|P1-P3) \times P(P1-P3) + P(FSM) = P(FSM|P4-P7) \times P(P4-P7)$, where in school year 2013/2014 about 45% of the primary school population was in P1-P3 and 20.6% of the school population was FSM-registered, suggests that nearly 21% of P1-P3 pupils were FSM-registered in 2014. This means $169,485 \times (1 - .21) \approx 133,893$ is the number of P1-P3 pupils who were not FSM-registered one year before the policy.

increased from 53.2% to 64.6% in 2015, and to 78.9% among P1-P3 pupils. In addition, in 2016 approximately 66% of primary pupils were taking school meals, with the P1-P3 fraction increasing to 81.7%. On the other hand, the share of P4-P7 taking school meals remained fairly stable around 53% (McAdams, 2016). The policy seemed to have achieved at least some of the initial goals. For instance, a qualitative evaluation (Ford et al., 2015) found that parents identified three main benefits: financial savings, time savings from not having to pack lunches, and school meals being healthier.

3.3 Data

For most of the analysis, we use a balanced panel of 1,630 primary schools observed for eleven school years within the time period 2003/2004 – 2016/2017. We merge four main data sources: *i)* Healthy Living Survey; *ii)* Scottish Pupil Census; *iii)* School Contact Details; *iv)* Attendance and Absence Survey.

3.3.1 Healthy Living Survey

The main data for this project come from the School Meal Survey, renamed Healthy Living Survey (HLS) in 2012. The survey takes place in February every year. For every school, this collects the following information: *i)* number of pupils on the school roll; *ii)* number of pupils present on the day of the survey; *iii)* number of pupils registered for FSM; *iv)* number of pupils present and registered for FSM; *v)* number of pupils present and who took a school meal, whether free or paid-for; *vi)* number of pupils present and who took a FSM. Additional information are collected for a subset of waves only. For example, until 2009 the survey would also report the number of pupils who are eligible to receive FSM under the national criteria, and until the following wave the survey would include information on whether the school: *i)* had an anonymised payment system for school meal collection; *ii)* provided pupils with fresh fruit and water; *iii)* had a breakfast club.[11]

Therefore, for each wave it is possible to calculate FSM, paid-for school meals (PSM) and overall school meals uptakes. Holford (2015) presents the overall uptake in the following way. Given school

[11]Not surprisingly, prior to 2015 the correlation between the fraction of pupils who are eligible and those registered is nearly 1.

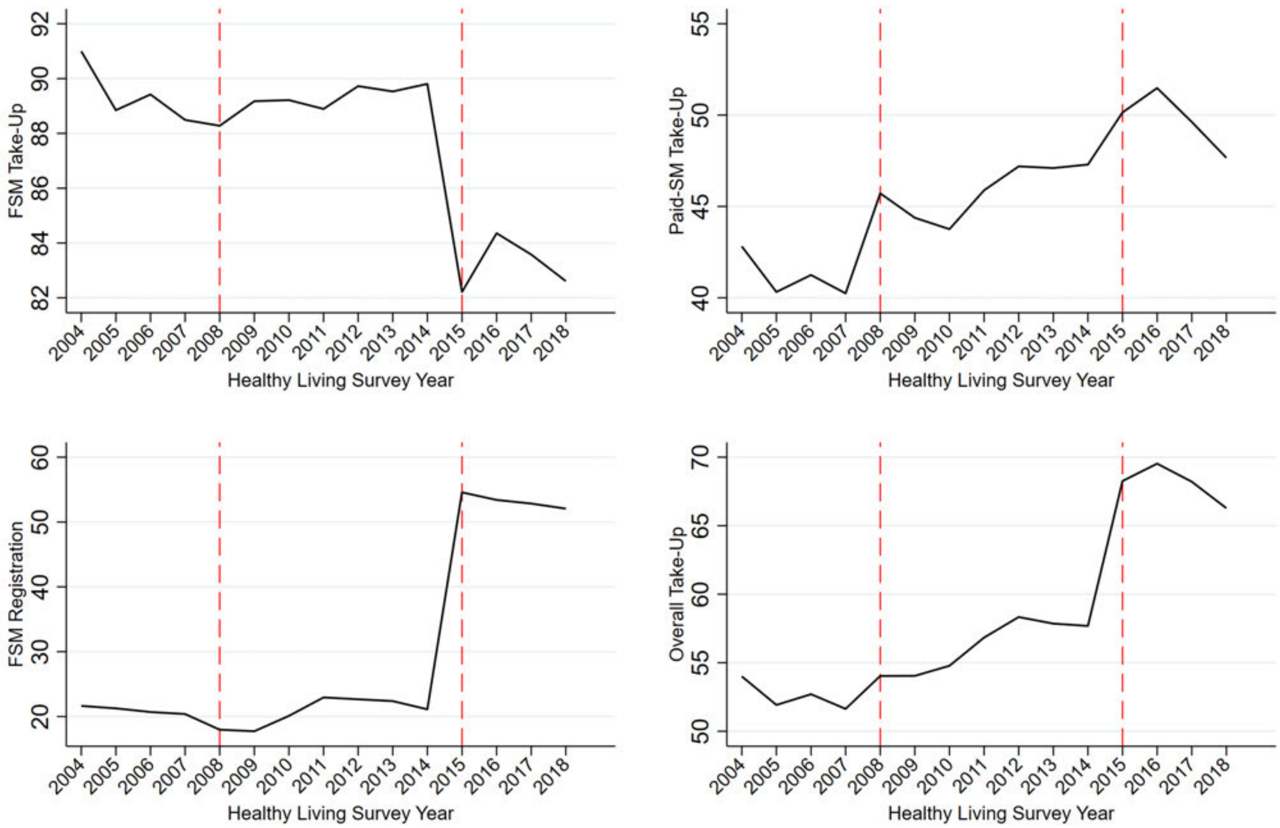
s, ρ_s represents the fraction of pupils registered for free school meals, while y_s^f and y_s^p are respectively FSM and PSM uptakes. Hence, the overall uptake can be outlined as the weighted sum of FSM and PSM uptakes, whose weights are the fractions of registered/unregistered pupils:

$$\begin{aligned} \underbrace{\frac{\#Meal-Takers}{School\ Population}}_{\text{Overall Uptake, } Y_s} = \overbrace{\frac{\#FSM-Registered}{School\ Population}}^{\rho_s} \times \overbrace{\frac{\#FSM-takers}{\#FSM-Registered}}^{y_s^f} + \\ + \underbrace{\frac{\#notFSM-Registered}{School\ Population}}_{1-\rho_s} \times \underbrace{\frac{\#PSM-takers}{\#notFSM-Registered}}_{y_s^p} \end{aligned} \tag{3.1}$$

We calculate all the elements of Equation 3.1 from the information provided in the HLS. For instance, of all pupils present on the survey day, y_s^f is calculated as the fraction of FSM-registered pupils who took a (free) school meal. In addition, we calculate the uptake of paid-for meals as a ratio whose numerator is the difference between the number of (present) pupils taking any meal minus the number of pupils specifically taking FSM, while its denominator is the number of pupils non-registered for FSM, which in our data is just the number of pupils present minus those present and registered. From here, the overall uptake can be calculated in accordance with Equation 3.1 or simply as the ratio between pupils taking any meal and the number of pupils present.

Figure 3.1 provides trends in each of the terms outlined in Equation 3.1. We can see how registrations start increasing after 2009, following a series of local initiatives to expand eligibility. After plateauing from 2011, we then observe a sharp hike by about 35 percentage points in 2015. Take-up of paid meals only increased by a couple of percentage points in 2015, following an upward trend started prior to 2010. It should be noted that from 2015 on, it was exclusively P4 to P7 students who were paying for school meals, as pupils in those grade were not covered by the UFSM policy. Uptakes of free school meals dropped remarkably following the introduction of their universality in 2015. This happened for two reasons. First, all P1-P3 students were automatically registered from 2015, thus expanding y_s^f's denominator. Second, the newly eligible group had a lower take-up rate than the existing eligible group.

Figure 3.1: Trends



Note: The above charts are calculated as raw yearly-averages of the following ratios: FSM-Take-up: $\frac{\#FSM-Takers}{\#FSM-Registered}$; Paid-SM Take-Up: $\frac{\#PSM-Takers}{\#non-FSM-Registered}$; FSM Registration: $\frac{\#FSM-Registered}{School\ Population}$; Overall Take-Up: $\frac{\#Meal-Takers}{School\ Population}$. Because the survey is run in one day, the raw counts refer to pupils present on the day of the survey.

This is also an important departure from Holford (2015)'s evaluation of the pilot. His identifying assumption relied on the fact that the policy would only affect Y_s through y_s^p due to the reduction in price, which would in turn leave y_s^f unaffected. But importantly, due to the way the pilot was implemented, the share of FSM-registered pupils was not impacted. This obviously does not apply to the UFSM set up, which automatically registered all P1-P3 pupils. This explains the apparent drop in FSM uptake. As the goal of this analysis is to examine how policy-induced variation in the fraction of school population taking FSM affected school-level behaviour, we therefore consider a slightly different measure of FSM uptakes. This can be seen after re-arranging Equation 3.1:

$$\underbrace{\frac{\#Meal-Takers}{School\ Population}}_{\text{Overall Uptake, } U_s} = \overbrace{\frac{\#FSM-takers}{School\ Population}}^{u_s^f} + \underbrace{\frac{\#PSM-takers}{School\ Population}}_{u_s^p} \tag{3.2}$$

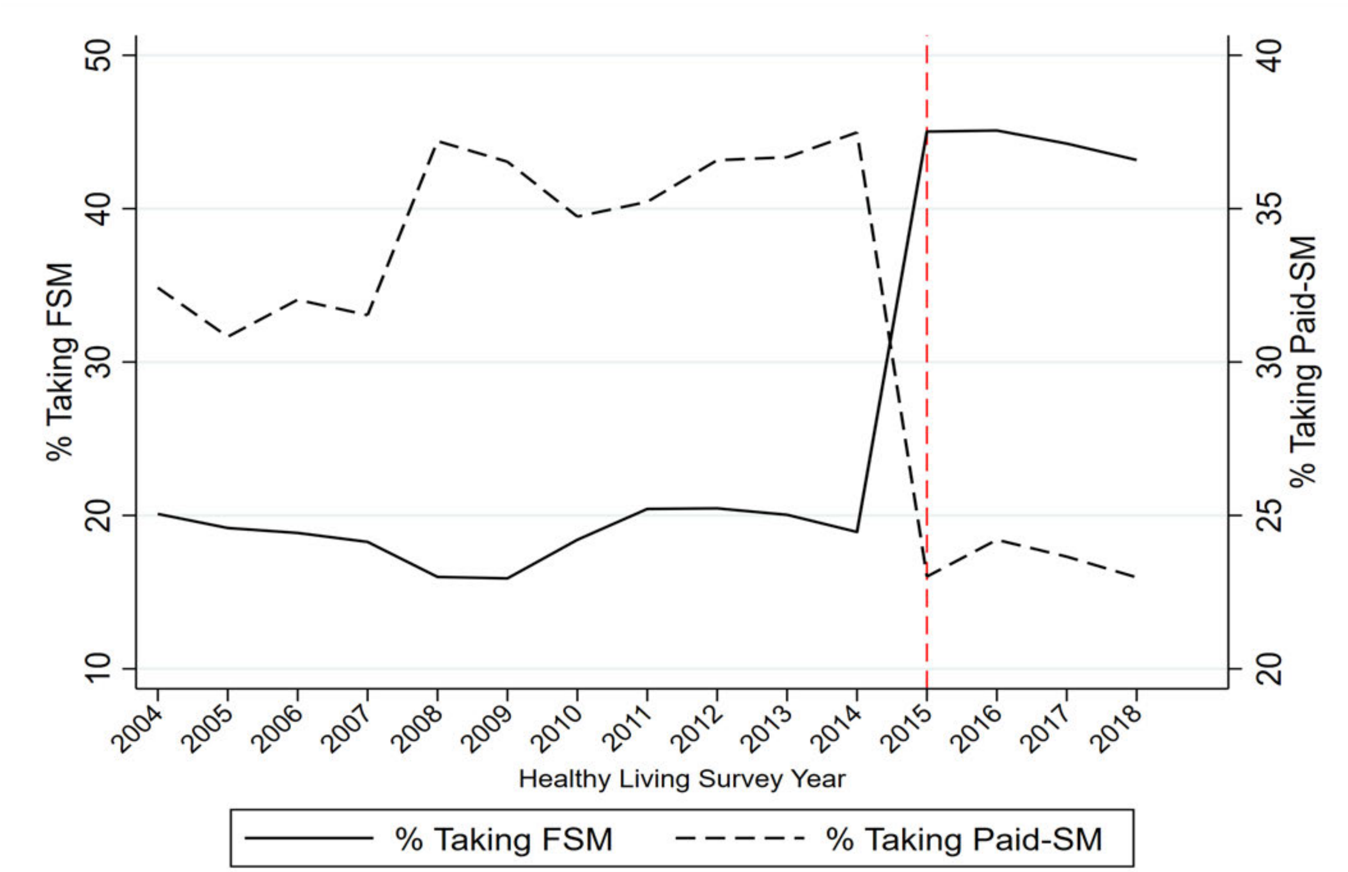
Whereby the overall uptake is the sum of FSM and PSM takers divided by the number of pupils in school.[12] Equation 3.2 is equivalent to Equation 3.1 after cross-deleting registration rates. Figure 3.2 shows the re-defined measures of FSM take-up rate, defined as the % of pupils present on the survey day who took a free meal (solid line) and who took a paid-for meal (dashed line). Once we do not account for FSM registration rates, we observe the expected pattern. The share of pupils taking FSM - which is the focus of this work - increased by approximately 25 percentage points following the implementation of UFSM. Conversely, the fraction of those paying for their meals went down by about 14 percentage points.

It is worth reiterating that the policy targets pupils in the first three stages of primary school (P1-P3). While we cannot observe a breakdown of registrations and uptakes by stages throughout our sample period, official statistics from 2015 report this information at the national level. These are plotted in Figure 3.3a. The dotted black line represents the entire primary school-level FSM registration trend as in Figure 3.1. The solid red line disaggregates the trend for the P1-P3 group (targeted by the policy) whereas the solid black line shows the P4-P7 trend. Starting in 2015, all P1-P3 pupils become automatically FSM-registered (horizontal line) whereas P4-P7-level registrations roughly maintain the overall pre-policy trend, yet with a slight downturn. Figure 3.3b follows the same approach, but it plots the percentage of primary school population taking free school meals, split by stages.[13] Under the plausible assumption that uptakes and registrations did not differ substantially across grades, i.e. the fractions of FSM-registered (and/or FSM-takers) are roughly the same within P1-P3 and P4-P7 cohorts, we can see that following the change in policy, registrations and uptakes only changed significantly within the

[12] Due to the survey taking place on a single day, this is not the actual school roll but the number of pupils present on the survey day. However, the latter is a good proxy for the former, with a correlation coefficient above 0.9.

[13] The official statistics only report the percentage of FSM-registered pupils who took a meal on survey day, i.e. y_s^f from Equation 3.1. Therefore, by multiplying this by the P1-P3 shares of FSM-registered pupils we obtain our P1-P3 uptake measure as in Figure 3.2. The same approach is applied to obtain P4-P7 uptakes.

Figure 3.2: FSM vs PSM Uptakes



Note: The above trends are calculated as raw yearly-averages of the following ratios: % Taking FSM: $\frac{\#FSM-Takers}{School\ Population}$; % Taking Paid-SM: $\frac{\#PSM-Takers}{School\ Population}$. Because the survey is run in one day, the raw counts refer to pupils present on the day of the survey.

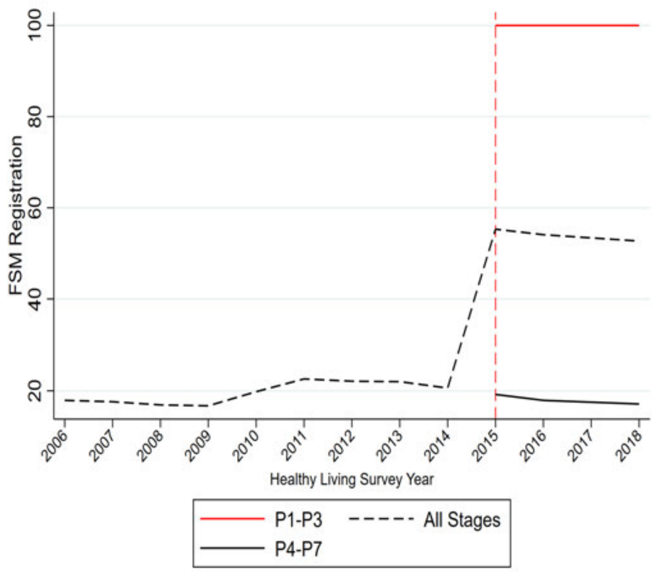
P1-P3 group.[14]

3.3.2 School Information

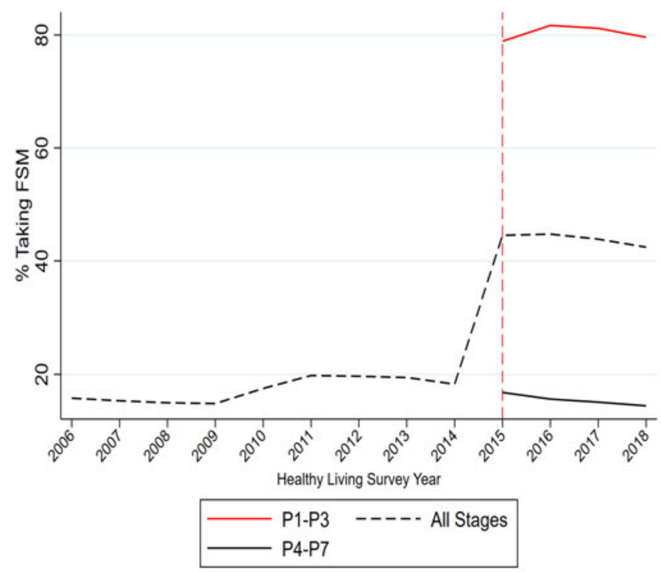
Data on school characteristics are taken from the 'school contact details' database, alongside the Scottish Pupils Census (SPC), which runs shortly after the beginning of every school year. Hence, SPC and HLS which pertain to the same school year are identified from subsequent waves, as they take place in two different calendar years. From these data sources we are able to obtain information such as: *i)* school

[14]For example, the registration rate can be decomposed as $P(FSM) = P(FSM|P1-P3) \times P(P1-P3) + P(FSM) = P(FSM|P4-P7) \times P(P4-P7)$. As $P(P1-P3) \approx P(P4-P7)$, then $P(FSM|P1-P3)$ and $P(FSM|P4-P7)$ must not diverge excessively.

Figure 3.3: FSM Registrations and Uptakes - by Stages



(a) FSM Registration



(b) FSM Uptakes

Note: The above charts are calculated as raw yearly-averages of the following ratios: FSM Registration: $\frac{\#FSM-Registered}{School\ Population}$; % Taking FSM: $\frac{\#FSM-Takers}{School\ Population}$. Because the survey is run in one day, the raw counts refer to pupils present on the day of the survey. The dotted black line represents the entire primary school-level trend. The solid red line dis-aggregates the trend for the P1-P3 group (targeted by the policy) whereas the solid black line shows the P4-P7 trend. The breakdown of registration and take-up is only available, at the national level, starting from HLS wave 2015.

postcode, which can be linked to a variety of neighbourhood characteristics; *ii)* number of pupils in school; *iii)* FTE number of teachers; *iv)* fraction of pupils from ethnic minorities; *v)* breakdown of pupil numbers by stage; *iv)* average class size and fraction of composite classes. We use the Scottish Index of Multiple Deprivation (SIMD) of the school location as a proxy for school composition. SIMD is an index developed based on seven domains, i.e. income, employment, health, education, crime, housing and access to services. This information is collected at the data zone level. A data zone is a block containing between 500 to 1,000 residents and currently Scotland is divided into 6,976 data zones.

3.3.3 Attendance and Absence Survey

Data on attendance, absences and exclusions have been consistently collected every school year until 2010/11. Thereafter, the survey took place every second year with 2018/19 being the latest wave.

Appendix Table C.10 provides an overview of when and how the survey was carried out. It follows the same wave structure as the SPC. For instance, wave 2010 includes attendance and absences which occurred in school year 2010/2011. HLS waves instead refer to February of the same school year, hence the calendar year following the one when the school year begins.

For a given school in a given year, the survey collects the total number of episodes of attendance, authorised absence, unauthorised absence and exclusions, broken down by reason. These refer to all stages in school. For instance, days spent in school, whether arriving on time or slightly late, days of work experience and instances of educational provision during illness all count toward attendance. Table 3.1 reports summary statistics for the period 2003-2018.

In panels A to C we calculate the incidence of each episode within their own category. It can be seen that attendance is mainly composed of "in school" attendance (99.14%) with a residual part being mostly due to students being late and studying while sick.[15] From Panel B we see that about 67% of authorised absences are health-related, followed by nearly 27% of the episodes being attributed to "other" authorised absences. In addition, unauthorised absence are mostly due to holidays and 'only' about 30% are due to truancies.[16] Panel D shows that approximately 2.5% of all school sessions are missed due to illness and even lower shares are attributable to lateness and truancy. Moreover, exclusions are rare in primary schools (.02%). Unfortunately, we are not able to observe the type of misbehaviour which led to a disciplinary action. Rates in panels D and E are calculated in relation to the number of all possible attendances, which for each school corresponds to the roll times the total number of half-day openings.[17]

We face some missing data issues in the last wave, primarily related to statistical disclosure control (SDC) protocols which suppress small cells based on less than 5 instances. For instance, if in wave 2016, the cell for school A for "very late" is suppressed, the entire authorised absence rate cannot be calculated for that school in that wave. This issue affected primarily the authorised and unauthorised absences categories. For this reason we focus on three outcomes with relatively few instances of SDC

[15]This most likely entails chronic conditions forcing pupils staying away from school.

[16]The same information is collected for secondary schools, for which we observe different patterns. For instance, truancies tend to be the main reason for unauthorised absences and episodes of work experience features more often.

[17]These are about 380 per school year for the vast majority of schools and Local Authorities.

Table 3.1: Summary Statistics

Panel A: Attendance	Mean	St.Dev.
In School	99.14	1.34
Late	0.85	1.34
Work Experience	0.00	0.01
Sick With Educational Provision	0.02	0.09
Panel B: Authorised Absence		
Sick w/o Educational Provision	66.85	27.58
Very Late	0.21	1.07
Authorised Holiday	3.83	4.41
Exceptional Domestic Circumstances	2.22	3.78
Other Authorised Absence	26.88	28.01
Panel C: Unauthorised Absence		
Unauthorised Holiday	55.24	27.98
Truancy	30.13	28.61
Unauthorised Exceptional Domestic Circumstances	2.33	4.85
Other Unauthorised	12.30	17.01
Panel D: Sub-categories		
Sickness Rate (% Possible Attendance)	2.50	1.24
Lateness Rate (% Possible Attendance)	0.81	1.43
Truancy Rate (% Possible Attendance)	0.45	0.75
Panel E: Aggregates		
Exclusion Rate	0.02	0.04
Attendance Rate	95.02	1.63
Authorised Absence Rate	3.90	1.24
Unauthorised Absence Rate	1.06	0.87
No. of Schools	2,335	

Notes: Panels A, B and C report for each group the average proportion of each subcategory within attendance, authorised and unauthorised absence respectively. For example, on average 99.14% of overall attendance is "in school", and about 67% of authorised absences are due to pupils being ill. In Panel E, each aggregate is calculated as % of all possible attendance. These are 2003-2018 averages.

suppression, i.e. attendance rate, health-related absences (sickness rate from panel D) and exclusion rates.

3.3.4 Analytical Sample

Our final sample spans from school year 2003/2004 through to 2016/2017. Within this interval, we observe school meals uptake and registration for every year. However, due to the fact that the attendance survey was collected every second year from school year 2010/2011, we experience some gaps in the outcome variables, specifically for school years 2011/2012, 2013/2014 and 2015/2016. While wave 2018/2019 is in principle available to us, we decided not to include it in our analysis primarily as in August 2018 Glasgow City Council extended FSM provision to fourth graders, therefore we want to avoid an overlapping of policies. In summary, our analysis sample includes eleven time periods. Specifically, we have nine pre-treatment periods, i.e. 2003/2004, 2004/2005, 2005/2006, 2006/2007, 2007/2008, 2008/2009, 2009/2010, 2010/2011 and 2012/2013, and two post-treatment periods, namely 2014/2015 and 2016/2017. Appendix Table C.10 provides more details.[18]

In addition, the secondary data used for this project also contain a series of suppressed values due to the application of statistical disclosure control. In general, any percentage whose underlying sample size is between 1 and 4 is reported as missing. In Figure C.2 we compare the trends calculated with the secondary data on school meals to those obtained from the official, publicly available aggregates. The patterns are virtually unchanged, despite the fact that our data seem to overstate these measures. This is mechanical, since disclosure control essentially excludes instances with very low values. This issue pertains to our treatment as well as outcome variables. We show in a robustness check that our results are not driven by missing data. We randomly impute missing counts with values between 1 and 4, re-calculate the relevant fractions and re-run the analysis with the imputed outcomes and treatment variables. Finally, we retain only those schools that can be observed in all eleven of our time periods making our panel balanced.

3.4 Empirical Strategy

Our research question entails a number of methodological challenges, primarily in relation to the endogeneity of school-level uptakes. First, there can be omitted variable bias induced by unobserved school-specific and time-invariant characteristics which are correlated with uptake. For instance, selection into schools might be a driver of uptake. Second, shocks in behaviour in one year might drive FSM uptake in the future, leading to simultaneity between behaviour and uptake. Comparing be-

[18]Treatment occurs half-way through the 2014/2015 school year.

haviour across schools with different levels of uptake would certainly lead to spurious correlations. To overcome these issues, we employ a difference-in-differences (DiD) model with continuous treatment (Clemens et al., 2018; Card, 1992) which is estimated using the following two-way fixed effects (TWFE) regression equation:

$$y_{sct} = \gamma(I_{t\geq 2015} \times E_{sc,2014}) + \beta' X_{sct} + \alpha_s + \lambda_t + \varepsilon_{sct} \tag{3.3}$$

where y_{sct} is a set of behavioural outcomes averaged for school s, in Local Authority (council) c and school year t. E_{sc} is our measure of exposure to the UFSM intervention. $I_{t\geq 2015}$ is a dummy variable switching to one for every wave following the policy implementation.[19] Our coefficient of interest is γ, measuring difference in behaviours across different levels of exposure to the policy. X_{sct} is a set of pre-treatment covariates, which we interact with time dummies. These include school composition, school population and resources.[20] Since these could constitute potential outcomes of the policy we fix them to 2004, well before the change in policy. In addition, X_{sct} contains time-invariant characteristics such as religious status and whether the school is in an urban area, also interacted with time dummies. Finally, α_s and λ_t are school and year fixed effects and ε_{sct} is the idiosyncratic shock component of behaviour. Our source of identifying variation comes from the before-after comparison of treatment and control group, paired with within-school variation leveraged by α_s. Residual concerns entail school-specific time trends which might simultaneously drive uptakes and outcomes. For instance, periodic staff shortages could undermine the effectiveness of lunch deliveries but also impact students' engagement and behaviour. In this vein, in some specifications we control for school and Local Authority-specific time trends, i.e. $\alpha_s t$ and $\alpha_c t$. In addition, as UFSM was widely discussed in the media at least one year before the implementation, one could worry about anticipation effects. This would entail pupils switching schools on account of pre-treatment level of uptakes. We believe this is highly unlikely as the Scottish school attendance system is residence-based and implies households providing evidence of residence within the attendance area at least six months in advance of the school

[19] These refer to the calendar years before/after the policy was implemented.

[20] We mostly employ school-level average class size as this is a good proxy of school population and resources. Its correlation with $ln(school\ population)$ and pupil-teacher ratio is .84 and .77 respectively.

year start (see Rossi (2021) and Borbely et al. (2021)).

Our preferred measure of exposure is:

$$E_{sc,2014} = \frac{\#\ Non - FSM - Takers_{sc,2014}}{School\ Population_{sc,2014}}$$

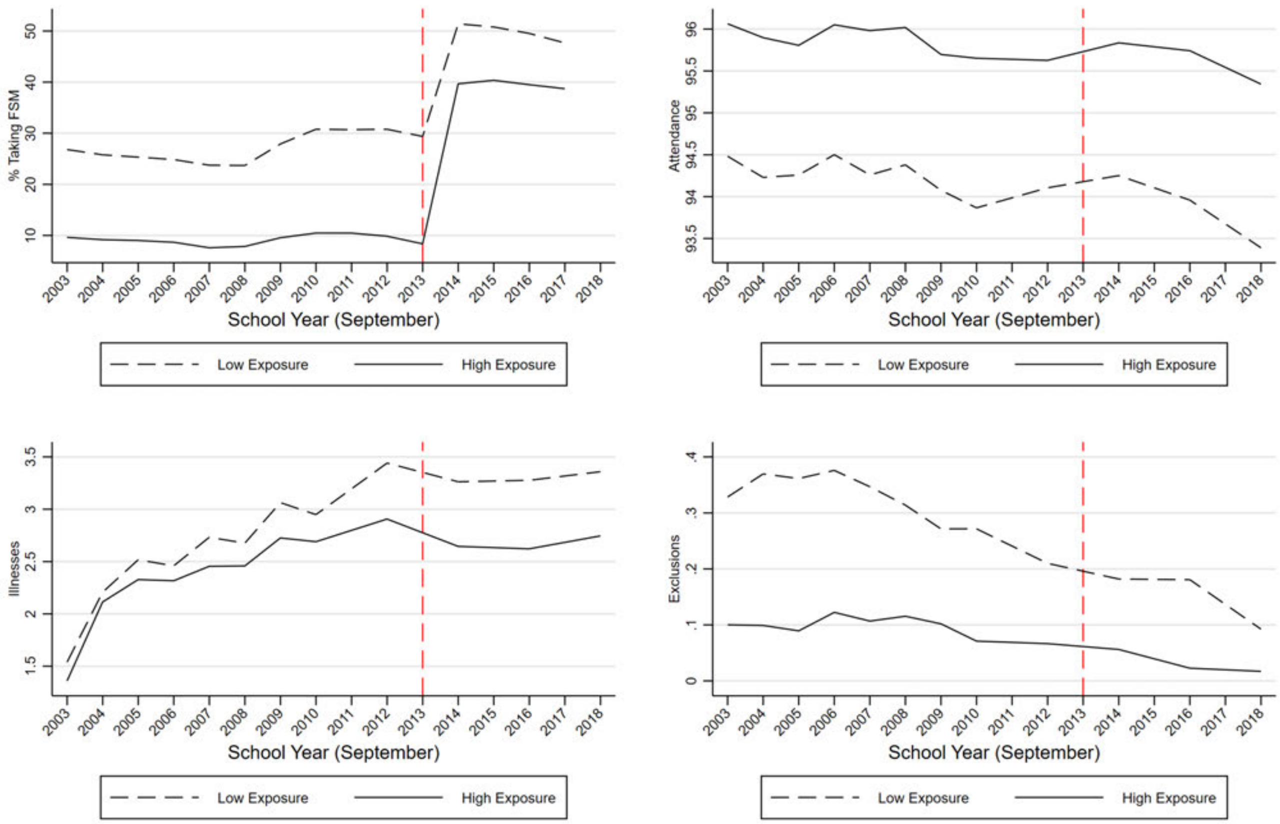
whose rationale we illustrate with an example. Consider two equally sized schools, A and B. In school A, 90% of the students take FSM before FSM become universal, whereas only 10% do in school B. Conditional on observables, the policy is thus likely to have a stronger effect in school B, where a larger share of the school population do not already take FSM. Put simply, our strategy compares schools where the policy induced a larger change in uptake with schools where the policy did not do so, on the account of uptake already being high prior to the policy.

Of course, we need to corroborate this with evidence that the policy led to such an increase in FSM uptake. This 'first stage' is presented in the top-left panel of Figure 3.4.

Here we plot trends in school-level percentage of pupils taking FSM by level of exposure to the policy. For ease of representation we distinguish between school with high versus low exposure to the policy, based on the school having a share of non-FSM-takers in school year 2013/2014 that is above or below the median. The group with high exposure to the policy is characterised by a stable trend of about 10% of FSM-takers prior the UFSM implementation, whereas the group with low exposure, where relatively more pupils were already taking FSM by 2015, had around 25 to 30% FSM-takers. Here we can see how FSM uptakes have sharply increased regardless of the pre-policy level of exposure, however, these increased by approximately 10 percentage points more within schools with high policy exposure.

Now that we have ascertained that the UFSM has indeed induced a stronger increase in uptakes for schools with fewer FSM takers before the treatment, we need to make sure these two groups are comparable over time. As stated before, a simple 'between-schools' comparison would result in inconsistent estimates of the effect of the policy. Figure 3.4 shows trends in outcomes for schools within the treatment and control group.

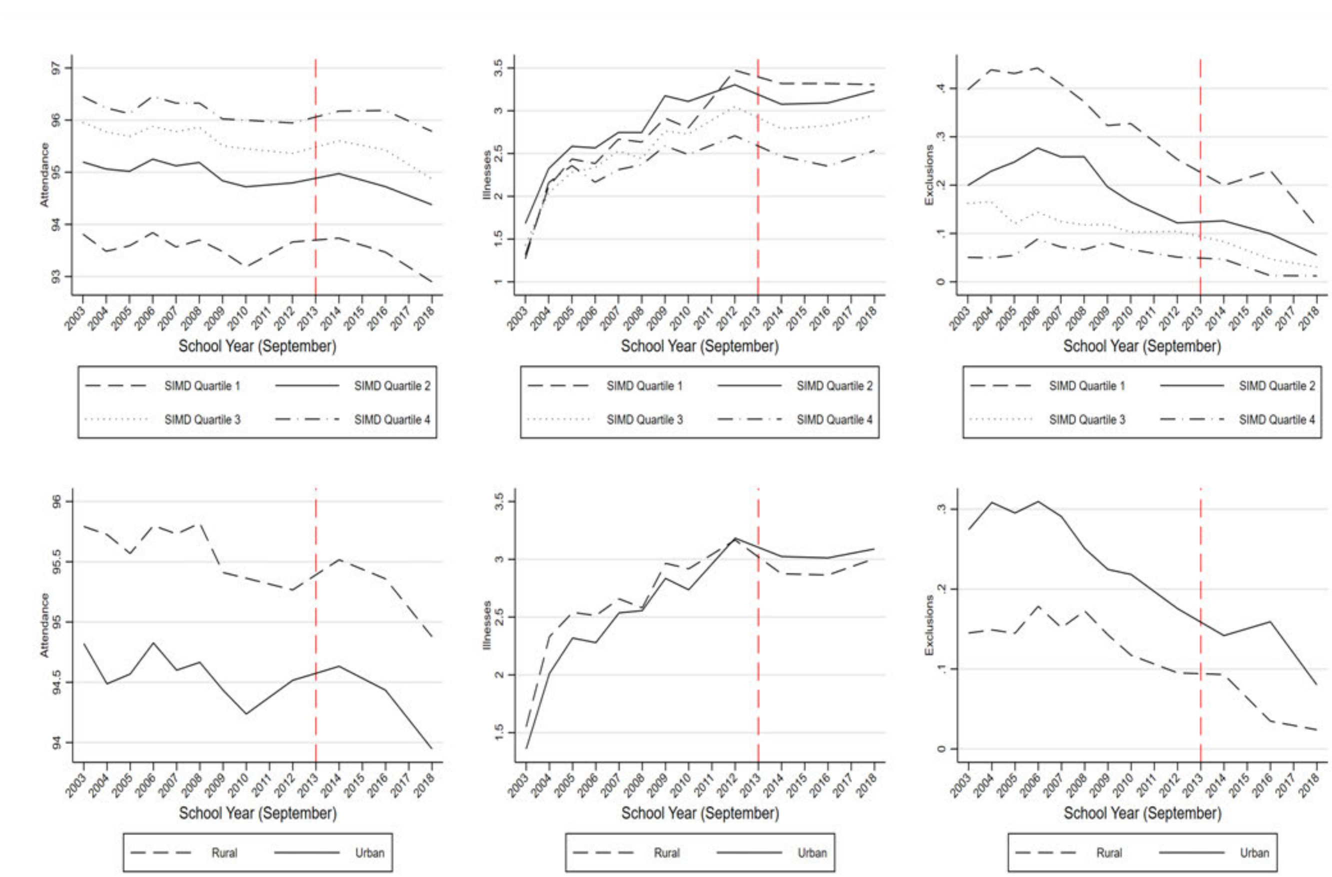
Figure 3.4: Parallel Trends



Note: The above charts are calculated as raw yearly-averages of the reported variables, across levels of exposure. In particular, exposure is defined as the % of school population not-taking FSM in 2013, one year before the policy change. A high exposure is denoted by a value above the median.

While for our design to be valid we need the timing of the policy not to be associated with *changes* in outcomes – i.e. the parallel trends assumption needs to hold – a recent, yet small, literature (Jaeger et al., 2020; Kahn-Lang and Lang, 2020) has stressed the importance for treatment and control group to be similar also in *levels*. If these diverge considerably, one might wonder whether the factors driving those discrepancies might also affect *changes* in trends. We can see from Figure 3.4 that while the outcomes seem to follow parallel trends up to the change in policy, different exposures are associated with different *levels* of the outcomes. For instance, high exposure schools are characterised by approximately 1.5 percentage points higher attendance rate on average, as well as lower illnesses and exclusions. This is not surprising considering that, prior to the policy change, FSM eligibility was contingent to disadvantaged socio-economic status, which can therefore be related to higher absenteeism (Sosu et al., 2021), worse health outcomes (Adams et al., 2003) and anti-social behaviour (Piotrowska et al., 2015).

Figure 3.5: Trends Within Sub-Groups



Note: The above charts are calculated as raw yearly-averages of the reported variables, across Scottish Index of Multiple Deprivation quartiles and whether the schools are in urban vs small town or rural area.

For this reason, similarly to Jaeger et al. (2020) we present in Figure 3.5 outcomes' trends by sub-groups, namely quartiles of the SIMD as well as for urban and rural schools. We can observe how there are differential trends across groups, especially for exclusions and illnesses. While this is not invalidating per se, we need to keep in mind that our measure of exposure is directly related to deprivation and this might cause the parallel trends assumption not to hold. For this reason, even our most basic specifications will contain interactions between year dummies and schools' postcode SIMD score in 2004 (well before the change in policy) which proxies school composition.[21] One concern is that after controlling for SIMD scores-specific time trends there might be little variation left in our outcomes and

[21]Primary schools' catchment areas are small enough that the school building neighbourhood is a good enough proxy of the entire catchment. In addition, the correlation between this 2004 index and the percentage of school population entitled for FSM under the national scheme in 2012 is .70.

treatment variables. We investigate this in Figure C.1. Histograms on the left-hand side refer to the raw variables, whereas those on the right-hand side are the residuals after regressing the variables on the Scottish Index of Multiple Deprivation Score in 2004. We can see that even after controlling for our proxy of school composition, there is still a decent amount of cross-sectional variation left in our measure of exposure. On the other hand, this seems to hardly matter for the outcomes, whose 'residualised' distributions look very much the same as the raw ones. This might be due to limited amount of variability in the first place, which is a feature that emerged previously in Table 3.1. We will address this in Subsection 3.5.1.

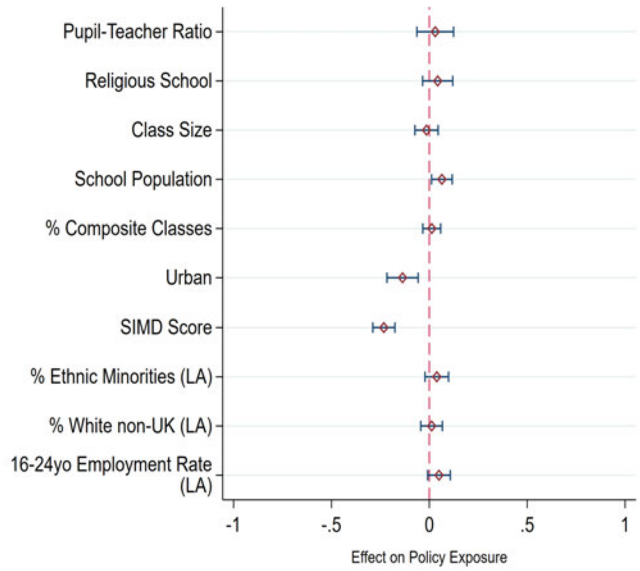
Furthermore, we want to rule out the possibility that divergence in outcome *levels* between high and low-exposure groups may mask differences in drivers of exposure to the policy. Figure 3.6 provides some evidence on this. Each point estimate is from a multiple regression model where the dependent variable is a binary indicator for whether the school was highly exposed to the policy in school year 2013/14 (above median, i.e. 83% of school population non-FSM takers) and the predictors are the variables listed on the left-hand side of the figure. The whiskers indicate 95% level confidence intervals. The analysis is conducted on a cross-section of 1,630 primary schools in school year 2013/14. While we mostly include school-level variables, we also use Local Authority-level percentages of primary school pupils from ethnic minorities and from a non-UK white background, as well as the employment rate among people presumably too young to be the parents of primary school pupils in school year 2013/14. Results are similar whether we leverage between (Figure 3.6a) or within-Local Authority variation (Figure 3.6b) and suggest that none of the available school (or Local Authority) characteristics are significant predictors of UFSM exposure. Urban and more deprived schools seem to be significantly less exposed, i.e. they have larger pre-policy uptakes. Whilst we already control for deprivation in all of our specifications, we also show that our results are unaffected when controlling for variation in school location (urban or rural).

Finally, our data structure allows us to formally test parallel trends by estimating the following event-study equation:

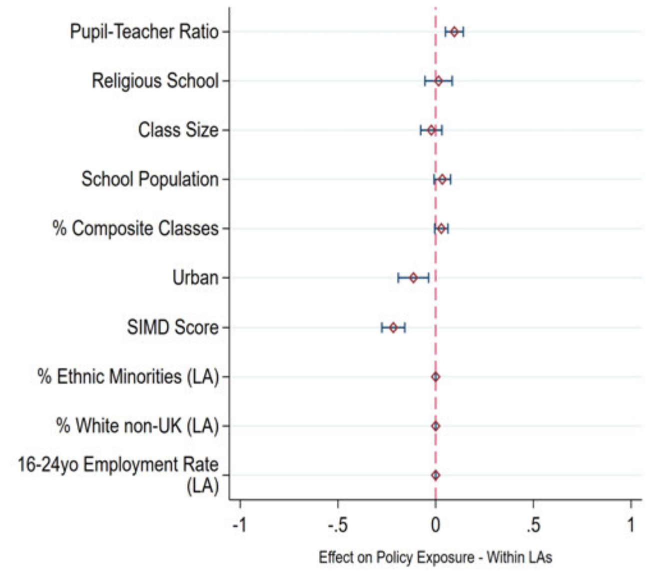
$$y_{sct} = \sum_{t=2003}^{2016} \gamma_t (I_t \times E_{sc}) + \beta' X_{sct} + \alpha_s + \lambda_t + \varepsilon_{sct} \tag{3.4}$$

Figure 3.6: Balancing Tests

(a) w/o LA Fixed Effects



(b) w LA Fixed Effects

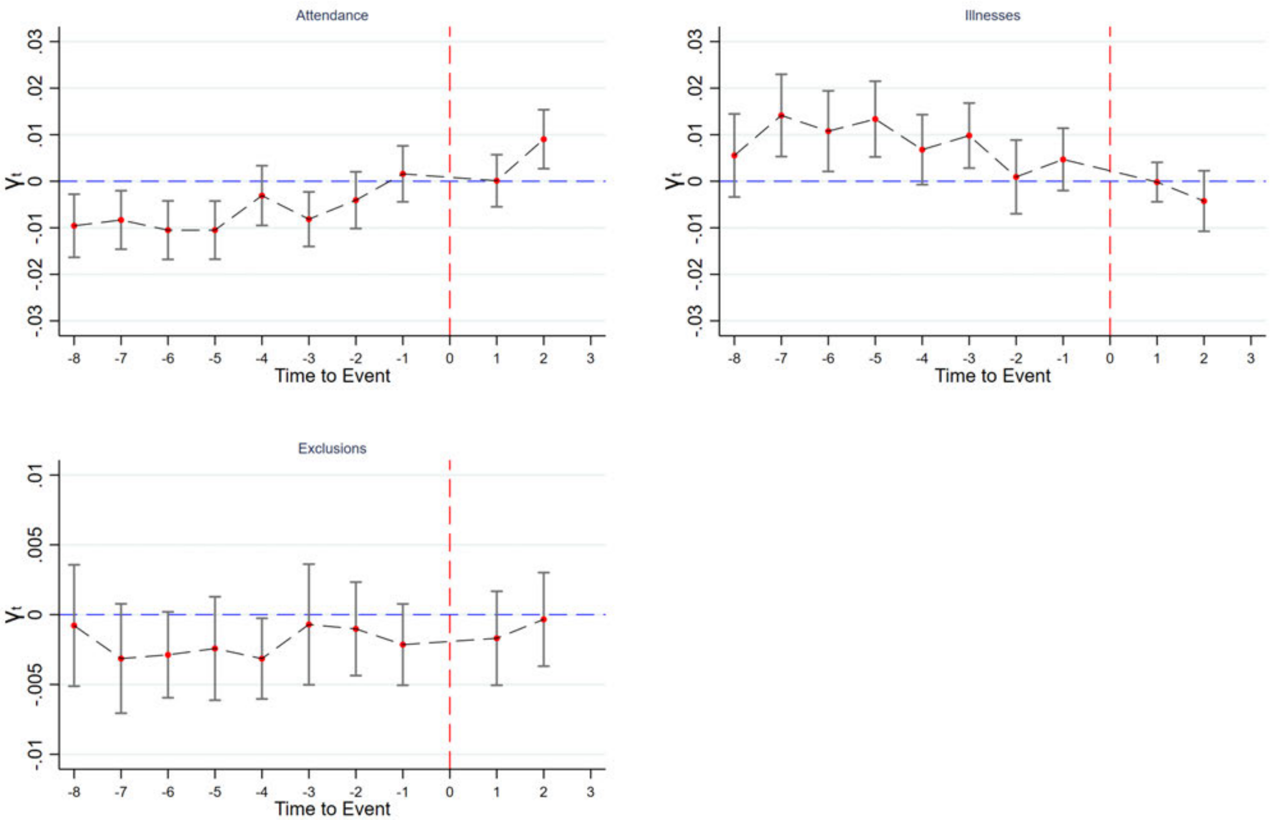


Note: Each coefficient results from a multiple regression of the sort $HighExposure_s = \beta_0 + \beta_1 x_{1,s} + \beta_2 x_{2,s} + ... + \beta_k x_{k,s} + \varepsilon_s$ whereby the dependent variable is a binary indicator for whether the school was highly exposed to the policy in school year 2013/14 (above median, i.e. 83% of school population non-FSM takers) and the predictors are the variables listed on the left-hand side of the figure. The whiskers indicate 95% level confidence intervals. The analysis is run on a cross-section of 1,630 school in school year 2013/14.

Our coefficients γ_t are plotted in Figure 3.7, alongside their 95% confidence intervals. The horizontal axis measures the number of periods to and from the reference period (t=0), which is school year 2012/2013, just before the change in policy. The reason for this choice of reference period is due to us not observing the outcomes in 2013/2014, as the survey was not carried out. For instance, t = – 8 is school year 2003/2004, eight periods prior to 2012/2013, once accounting for another gap in outcomes in 2011/2012. Period one is year 2014/2015, namely when UFSM is in place. Our models contain interactions between year dummies and SIMD. Standard errors are clustered at the school level.

These plots confirm the finding of the parallel trends charts, i.e. higher exposure is not associated with diverging trends in outcomes prior to the policy change. The coefficients are not statistically significant at the 5% level up to four years prior to UFSM implementation. While one could point out that some of the coefficients are individually significant at the 5% level, we argue that these point estimates are very small and not economically significant. For example, looking at the top-left panel, the

Figure 3.7: Event Study



Note: Coefficients are obtained by estimating γ_t from Equation 3.4. Each coefficient can be interpreted as the difference in outcome for that period relative to the reference period, which is school year 2012/2013, the earliest we observe the outcomes before the change in policy. Time is expressed in number of periods to and from the reference period, which is 0. Period one is year 2014/2015, namely when UFSM is in force. Exposure is defined as the % of school population not-taking FSM in 2013, one year before the policy change and outcomes are in levels. Data span from school year 2003/2004 through to 2016/2017 and outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. Whiskers represent coefficients' 95% confidence interval. Standard errors are clustered at the school level.

largest lead coefficient suggests a 0.1 percentage points reduction in attendance, which in turn suggests nearly a 0% reduction. We can also observe from these charts that following the implementation of the policy, while the outcomes changed in the expected direction, the size of the coefficient is also very small and suggestive of a null effect. We explore this further in the next section.

3.4.1 Limitation

Our main specification follows a difference-in-differences (DiD) with continuous treatment approach, where the increase in the share of FSM-takers is our dose d of treatment. In our example, $d = 10$ can be seen as a 10pp increase in FSM uptake. Recent work in the econometric literature (Callaway et al., 2021) shows that, under parallel trends assumption, TWFE returns a weighted average of a range of average treatment effects on treated (ATT) parameters, based on different levels of dosage d.[22] For instance, $ATT(d_j|d_j)$ is the average effect of dose d_j for those who effectively experienced dose d_j. In our case, the change in attendance rate a school would experience following a 5pp increase in FSM uptake. However, these weights can be negative, and TWFE can also put larger weights on doses closer to the average exposure, thus leading to misleading estimates. The authors bring what Angrist and Imbens (1995) coined as the *average causal response* to identify a change in outcome in response to a change in dose of the treatment. In the case at hand, this is called the average causal response on the treated (ACRT), and it identifies the change in ATT from changing dosage, namely:

$$ACRT(d_j|d_j) = E[Y_{it}(d_j) - Y_{it}(d_{j-1})|D_{it} = d_j]$$

which is the ACRT for dosage group d_j from taking dosage d_j, i.e. switching from dosage d_{j-1} to d_j. In our example, this would be the change in attendance rate a school would record, were a school with a 5 percentage points increase in FSM uptake to record additional 5 percentage points increase in uptakes. The identification of ACRT requires, however, one additional assumption relative to those needed for ATT. Calculating the ACRT entails comparing ATTs across different dosage groups, i.e. change in attendance rate following a 'dose' of 10 percentage points increase in uptake versus the one due to a 'dose' of 5 percentage points change. Mathematically, this is:

$$\underbrace{ATT(d_j|d_j) - ATT(d_{j-1}|d_{j-1})}_{\gamma} = \underbrace{(ATT(d_j|d_j) - ATT(d_{j-1}|d_j))}_{\gamma^{ACRT}} + \underbrace{(ATT(d_{j-1}|d_j) - ATT(d_{j-1}|d_{j-1}))}_{\text{selction bias}}$$

[22]Alongside parallel trends assumption, i.e. $E[Y_t(0) - Y_{t-1}(0)|D = d] = E[Y_t(0) - Y_{t-1}(0)|D = 0]$, identification of $ATT(d|d)$ also requires random sampling, support and no anticipation.

therefore, what we estimate in Equation 3.3 is γ, which captures γ^{ACRT}, i.e. what we hope to estimate plus a 'selection bias'. The latter term is the difference in ATT for the same dosage across different dosage groups, i.e. the difference in potential outcomes had different group experienced the same dosage. This bias reflects heterogeneity in treatment effects for the same dosage, and identification of ARCT requires homogeneity in treatment effect, on the top of canonical parallel trends assumption. This is what the authors call 'strong parallel trends'. In our example, the question is whether a school with a 5 percentage points increase in uptake and one with a 10 percentage points increase would experience the same change in attendance rate, had they both experienced the same dose. Whilst imposing treatment effect heterogeneity in such set up seems too restrictive, in their most recent version Callaway et al. (2021) do not yet provide a solution to this type of problem. For this reason, an admitted limitation of our study is that our identification assumes homogenous treatment effect across schools that 'select' into different doses of treatment.

3.5 Results

Table 3.2 presents estimates of γ from Equation 3.3. Just as in the event-study, we keep only those schools which we are able to observe every year from 2003 to 2016.[23] Every specification contains school and year fixed effects. Column (1) is our most basic specification where we additionally control for interactions between schools' SIMD score and year dummies. In column (2) alongside our proxy of school composition we control for average class size, religious status and indicator of whether the school is in a urban area, all appropriately interacted with time dummies. Finally, in columns (3) and (4) we control for school and Local Authority-specific linear time trends.

Our preferred specification is the one in column (2), which suggests that a one percentage point increase in exposure, i.e. pre-treatment share of pupils not taking FSM, increases attendance by .011 percentage points, reduces heath-related absences by .010, that is by only about 1/100 of a percentage point, and has no impact on days missed due to disciplinary action. The first two coefficients are statistically significant at any conventional significance level, while we fail to reject the null hypothesis

[23] Table C.1 and Figure C.5 show estimates when using the unbalanced panel. Results do not change.

of no effect on exclusions. These results, which are consistent across specifications, confirm what we found in the previous section, i.e. null effects of the policy on the variables of interest.

Table 3.2: Main Results

Panel A: Attendance				
	(1)	(2)	(3)	(4)
Post × Exposure	0.010***	0.011***	0.002	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)
Observations	17,766	17,766	17,766	17,766
No. of Schools	1630	1630	1630	1630
Mean Dep. Var.	95.05	95.05	95.05	95.05
SD Dep. Var.	1.57	1.57	1.57	1.57
R-squared	0.757	0.762	0.713	0.766
Panel B: Illnesses				
	(1)	(2)	(3)	(4)
Post × Exposure	-0.009***	-0.010***	-0.003	-0.009***
	(0.003)	(0.003)	(0.003)	(0.002)
Observations	17,781	17,781	17,781	17,781
No. of Schools	1630	1630	1630	1630
Mean Dep. Var.	2.57	2.57	2.57	2.57
SD Dep. Var.	1.54	1.54	1.54	1.54
R-squared	0.585	0.587	0.443	0.667
Panel C: Exclusions				
	(1)	(2)	(3)	(4)
Post × Exposure	0.001	0.001	-0.000	0.001
	(0.001)	(0.001)	(0.002)	(0.001)
Observations	17,308	17,308	17,308	17,308
No. of Schools	1630	1630	1630	1630
Mean Dep. Var.	0.19	0.19	0.19	0.19
SD Dep. Var.	0.48	0.48	0.48	0.48
R-squared	0.300	0.303	0.198	0.306
Baseline Controls		✓	✓	✓
School FE	✓	✓	✓	✓
School Trends			✓	
Local Authories Trends				✓

Notes: Coefficients are obtained by estimating γ from Equation 3.3. Outcomes are calculated in % of all possible half-day openings. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) score and school average class size in 2004, both interacted with year dummies, alongside indicators for urban and religious all appropriately interacted with year dummies. Data span from school year 2003/2004 through to 2016/2017. Outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. Standard errors (in parentheses) are clustered at the school level. *** p¡0.01, ** p¡0.05, * p¡0.1

In particular, even a full standard deviation increase in exposure translates into less than one session missed.[24]

3.5.1 Robustness

Our first set of results clearly suggests that, while the policy undoubtedly increased FSM uptakes, this did not translate into improved attendance, behaviour and short-term health conditions. In this section we address a series of concerns threatening our identification strategy. First, we want to address another point raised by Kahn-Lang and Lang (2020) in relation to parallel trends when treatment and control group differ in levels. In Figure 3.4 and Figure 3.7 we ascertain that high and low exposure schools, while on different levels of the outcomes, had been trending similarly with respect to the outcomes themselves. However, as Kahn-Lang and Lang (2020) point out, it is not clear what "similarly" means in circumstances like the one at hand. For instance, parallel trends might hold in levels, meaning that treatment and control groups experience the same absolute changes, but not in a logarithmic specification as their absolute changes lead to divergent percentage changes. Figure C.6 and Table C.2 reports the same event study and aggregate regression exercises in Figure 3.7 and Table 3.2 but with log-transformed outcomes.[25] We can immediately notice that not only does a different functional form not alter our conclusion about parallel trends, but the regression results also point towards a null effect.

Another concern, continuing the point raised in Section 3.4, relates to the possibility that variables explaining the difference in levels of the outcomes might have diverging trends. Table C.3 assess the robustness of our results to alternative specifications. While the majority of the analysis is conducted interacting covariates with year fixed effects - as in column (2) - we also run similar exercises using interactions with a linear time trend and the post treatment dummy variable. Results are largely unchanged.

In addition, as students in the first three grades (P1, P2 and P3) were targeted by the policy, perhaps

[24] The Scottish system features a minimum of 190 days (and 380 half-day sessions) per school year. The standard deviation of our exposure measure is about 13 percentage points, so one standard deviation increase in exposure translates to 13 $\times$ .011 = .14 percentage points increase in attendance. These are measured in half-day sessions, so 3.8 $\times$ .14 = .55 sessions, which is far below a whole school day.

[25] To accommodate this transformation in the presence of null values for health-related absences and exclusions, we took the natural log of 1 + the variables in levels, e.g. $ln(x) = ln(1 + x)$

higher importance should be put on those observations where a larger fraction of the school population is enrolled in those grades around the time the policy went into effect. We therefore estimate a weighted version of Equation 3.3. The results are reported in Table C.4. Regressions are weighted using the percentage of students in P1-P3 in 2013/2014 in each school as weights. Our results are not significantly affected by this new specification. Furthermore, we repeat the estimation exercise in Equation 3.3 and Equation 3.4 excluding from our sample: *i)* schools which took part in the FSM pilot in school year 2007/08; *ii)* schools which extended FSM eligibility among P1-P3 students based on local initiatives in 2010. The event study charts for these samples are reported in Figure C.7 and Figure C.8 whereas the aggregate regression results are in Table C.5 and Table C.6. In both cases, our conclusion does not change. Table C.7 and Table C.8 repeat once again the estimation exercise. In Table C.7 we impute observations which are missing due to statistical disclosure control. For both outcomes and treatment variables, we randomly imputed counts between 1 and 4 and recalculated the rates. In Table C.8 instead, we use outcomes in inverse hyperbolic sine (IHS) in order to prevent outliers to drive the results. In either case, the results do not change.[26]

A further potential concern relates to our chosen measure of exposure, which might not generate enough variation in uptake to trigger the expected benefits in terms of attendance and behaviour. For this reason, we expand our model in Equation 3.3 and estimate a triple difference-in-differences (DDD) similarly to Muralidharan and Prakash (2017). Our third term of comparison is secondary schools, which were not affected by the policy. Figure C.9 shows a sharper change in uptake (nearly 30 percentage points) when primary schools are compared against secondary schools. Therefore, we estimate the model[27]:

$$y_{sct} = \gamma(I_{t\geq 2015} \times E_{sc,2014} \times Primary_{sc}) + \beta' X_{sct} + \alpha_s + \lambda_t + \varepsilon_{sct} \tag{3.5}$$

Where $Primary_{sc}$ is a dummy variable taking value 1 if a school is in the primary sector. Results are presented in Table C.9. As the outcome variables follow different distributions across sectors, we estimate the models in a log-linear fashion. Similarly to the DiD results, we can see that highly exposed

[26]Unsurprisingly, IHS results are virtually identical to those from specifications using the natural logarithm of the outcomes.

[27]Triple DiD entails the inclusion of a minimum of seven terms on the right-hand side of the equation. For the sake of simplicity, we only display the main interaction.

(primary) schools experienced an increase in attendance by approximately 0.1% conditional on school composition and characteristics. The effect, however, fades out when we control for school and Local Authorities' trends. Health-related absences seemed to have increased in primary schools following the UFSM implementation by up to 1%. These results, however, do not hold up in all specifications. Finally, exclusions have experienced a reduction of about 1.9%. This new set of results suggest that even when comparing highly exposed primary schools before and after the change in policy with secondary schools, the results are quite modest. In addition, the event-study in Figure C.10 shows how primary schools were exhibiting positive pre-trends in attendance and health-related absence and negative ones in exclusions, relative to secondary schools. Therefore, the estimates from Table C.9 are at best an upper bound for the effect of FSM uptake on attendance and health-related absences and a lower bound for the effect on exclusions. This strengthens our main conclusion that there has been no effect of the policy on behavioural outcomes.

3.5.2 Mechanisms

One possible explanation of the null effects we have been observing so far may be related to the effectiveness with which the policy was implemented. In fact, we might expect the ease of implementation to differ across schools' characteristics. In this section we explore potential mechanisms, building on the findings of two evaluations of UFSM. McAdams (2016) found that the largest uptakes following UFSM were recorded in rural schools, as well as schools with the largest share of pupils from disadvantaged background. Conversely, urban schools experienced the lowest uptakes. Additionally, Chambers et al. (2016) pointed out that a lack of funding, staff, and school spaces worked as a barrier to policy implementation, thus affecting take-up. Based on these findings, we might expect larger effects in those contexts in which policy roll-out went more smoothly.

We investigate this by stratifying our sample and looking at the heterogeneous effects of the policy across sub-groups. The sub-groups are defined as follows. Urban schools are based on the six-fold classification of the Scottish Government and include schools in large and other urban areas. We proxy staff recruitment and funding by using information on class size, school roll and pupil-teacher ratio from the academic year commencing in 2005. This is well before the treatment period and is the earliest we can observe data on teachers' numbers. We generate dummy variables for small schools and classes by using the first quartile of their distributions in 2005. These correspond respectively to schools with less than 122 pupils and whose average class size is less than 20 pupils. Similarly, a low pupil-teacher ratio is identified by the bottom quartile of its distribution in 2005, i.e. an average of 15 pupils per teacher or less. School internal area is collected from the 2008 School Estates Survey and is measured in square-metres. Finally, we identify a school as *'most deprived'* if it is located in a data zone which is classified as within the 25% most deprived according to the 2004 SIMD.[28] We control for school and year fixed effects in every specification, alongside interaction between year dummies and SIMD score - except, of course, in column 6. While the point estimates appear to be rather small, they suggest interesting insights.

Table 3.3 shows that within the most exposed schools, urban schools seem to have experienced smaller improvements than schools in rural areas or small towns.

Table 3.3: Heterogeneity - Attendance

Dependent Variable: Attendance	(1)	(2)	(3)	(4)	(5)	(6)
Post × Exposure × Urban	-0.012***					
	(0.004)					
Post × Exposure × Small School		0.014***				
		(0.005)				
Post × Exposure × Small Classes			0.009*			
			(0.005)			
Post × Exposure × Low PT-ratio				0.004		
				(0.005)		
Post × Exposure × ln(Internal Area)					-0.003	
					(0.003)	
Post × Exposure × Most Deprived						-0.020***
						(0.005)
Observations	17,766	17,766	17,766	17,766	17,766	17,766
No. of Schools	1630	1630	1630	1630	1630	1630
Mean Dep. Var.	95.05	95.05	95.05	95.05	95.05	95.05
SD Dep. Var.	1.57	1.57	1.57	1.57	1.57	1.57
R-squared	0.761	0.761	0.761	0.761	0.759	0.758

Notes: Outcomes are calculated in % of all possible half-day openings. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) in 2004 score interacted with year dummies, alongside indicators for urban, religious, below-median school population indicators, all appropriately interacted with year dummies. Standard errors (in parentheses) are clustered at the school level. *** p¡0.01, ** p¡0.05, * p¡0.1

In addition, column (2) to (4) suggest that small schools and those with higher per-pupil expendi-

[28] Variation in policy exposure within schools from most deprived areas is not considerably different from the one within schools from less deprived areas. Their coefficients of variation ($CV = \frac{\sigma}{\mu}$) are .18 and .12 respectively.

tures benefited more in terms of reduced absenteeism. All coefficients are, however, again very small and not statistically significant at any conventional level, with the exception of school population. We conjecture that this is due to larger variation in number of pupils per school, relative to class size and pupil-teacher ratio.[29] However, given the large correlation shared by these three variables we can safely assume they all reflect the same dimension of school resources. Furthermore, we find no evidence that school space might have contributed to the effect of the policy, while highly exposed schools, and those with the highest levels of deprivation have seen an increase in absenteeism. Table 3.4 and Table 3.5 conduct the same exercise but for health-related absences and exclusions.

Table 3.4: Heterogeneity - Health-related Absences

Dependent Variable: Illnesses	(1)	(2)	(3)	(4)	(5)	(6)
Post × Exposure × Urban	-0.014***					
	(0.005)					
Post × Exposure × Small School		0.002				
		(0.005)				
Post × Exposure × Small Classes			0.011**			
			(0.005)			
Post × Exposure × Low PT-ratio				0.006		
				(0.005)		
Post × Exposure × ln(Internal Area)					0.003	
					(0.003)	
Post × Exposure × Most Deprived						-0.013**
						(0.006)
Observations	17,781	17,781	17,781	17,781	17,781	17,781
No. of Schools	1630	1630	1630	1630	1630	1630
Mean Dep. Var.	2.57	2.57	2.57	2.57	2.57	2.57
SD Dep. Var.	1.54	1.54	1.54	1.54	1.54	1.54
R-squared	0.586	0.585	0.585	0.585	0.585	0.582

Notes: Outcomes are calculated in % of all possible half-day openings. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) in 2004 score interacted with year dummies, alongside indicators for urban, religious, below-median school population indicators, all appropriately interacted with year dummies. Standard errors (in parentheses) are clustered at the school level. *** p¡0.01, ** p¡0.05, * p¡0.1

Results for health-related absences are similar in magnitude to those for attendance, and share the same sign. In other words, these findings suggest that where uptake of the policy was greater, i.e. rural schools and those with more resources, students' short-term health condition slightly worsened, although the effect sizes are still very small. In the most deprived schools the beneficial effects of FSM on illnesses seem to be slightly stronger. This may indicate an improvement in nutritional intakes for pupils from disadvantaged backgrounds. For exclusions, our analysis yields regression coefficients

[29] Average school roll in 2005 is 210 with a standard deviation of 124. Average class size has a mean of 22.8 and standard deviation 4.27. Finally, PT ratio's mean is 16.7 with only 3.5 of standard deviation

Table 3.5: Heterogeneity - Exclusions

Dependent Variable: Exclusions	(1)	(2)	(3)	(4)	(5)	(6)
Post × Exposure × Urban	0.005**					
	(0.002)					
Post × Exposure × Small School		-0.003				
		(0.003)				
Post × Exposure × Small Classes			-0.002			
			(0.002)			
Post × Exposure × Low PT-ratio				-0.001		
				(0.002)		
Post × Exposure × ln(Internal Area)					0.002*	
					(0.001)	
Post × Exposure × Most Deprived						0.005*
						(0.002)
Observations	17,308	17,308	17,308	17,308	17,119	17,308
No. of Schools	1630	1630	1630	1630	1630	1630
Mean Dep. Var.	0.19	0.19	0.19	0.19	0.19	0.19
SD Dep. Var.	0.48	0.48	0.48	0.48	0.48	0.48
R-squared	0.300	0.300	0.300	0.300	0.585	0.582

Notes: Outcomes are calculated in % of all possible half-day openings. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) in 2004 score interacted with year dummies, alongside indicators for urban, religious, below-median school population indicators, all appropriately interacted with year dummies. Standard errors (in parentheses) are clustered at the school level. *** p¡0.01, ** p¡0.05, * p¡0.1

that are in line with expected reductions in misbehaviour in schools where uptakes have been higher, although coefficients are only weakly statistically significant. Taken together, our subgroup analysis fails to reveal any substantial benefits for any group. This suggests that the aggregate null effect that we find, is not masking any important sub-group effect.

3.6 Conclusion

The provision of universal free school meals (UFSM) has become a commonly used form of welfare policy in recent years, yet its impact is still widely debated. In this paper, we evaluate UFSM implementation in relation to an overlooked set of outcomes: school attendance, short-term health conditions and misbehaviour. We do so by focusing on the case of Scotland, where in 2015 all pupils in the first three grades of primary school became eligible to receive FSM, regardless of their household's financial circumstances. We employ a difference-in-differences (DiD) design where treatment intensity is determined by pre-policy levels of FSM uptakes. That is, the introduction of UFSM had more "bite" in schools with few pre-implementation FSM takers than in schools where FSM enrollment was high at baseline.[30] We find precisely estimated null effects on attendance and health-related absences. A 10-percentage-points increase in school population taking free school meals leads to a gain of less than one school day. We also find no evidence that the policy helps to prevent misbehaviour. Nor do the null effects mask benefits for subgroups. Rural schools, small schools, and schools with more resources see only marginally larger and economically very small benefits.

Our study has some limitations. First, the policy targeted only the first three grades of primary schools (45% of school population on average), while the outcomes are aggregated across all seven grades. Second, our outcome data allow us to only track effects for up to 3 years after policy implementation. Third, attendance rates - our main outcome - are very high to begin with in Scottish primary schools (95% on average) which means that there is limited room for improvements. Exclusions refer to a disciplinary procedure being issued, rather than the actual misbehaviour occurring (Altindag et al., 2020). Either way, these are rare in primary school.

With this in mind, one should be careful about concluding that UFSM are in general ineffective at encouraging attendance or improving pupils' short-term health condition. Both dimensions should be carefully considered by policymakers. However, the small effects so far found in the literature (see Corcoran et al., 2016; Gordanier et al., 2020; Cuadros-Meñaca et al., 2022; Holford and Rabe, 2020) as well as our study should raise the question of how UFSM policies are rolled out, how nutritious the meals provided are, and whether implementing such programmes on a large scale can be done while maintaining a high food quality (Parnham et al., 2022). One assumption of our study was that the policy did not change the nutritional content of school meals. Notably the only study finding a sizeable reduction in absenteeism is Belot and James (2011) where the authors assess the effect of a change in nutritional content, rather than just an extension in the provision of school meals. This is a key area of future research. Similarly, it would be beneficial to utilise individual pupil level data to test the effect of the more direct exposure to the policy, i.e. being enrolled in first, second or third grade on pupil level outcomes. Additionally, future work could explore how this policy has increased uptake among previously eligible students, in line with Holford (2015). Finally, as one of the goals of the policy was to reduce food insecurity among children, scholars could explore the impact of this policy on households'

[30] Prior to the change in policy uptakes among eligible students were on average 90%.

finances and benefit claims.

Concluding Remarks

This thesis is a collection of three empirical essays on school choice and child development.

School quality is an important aspect of households' residential choice. In **Chapter 1** I use housing and school data from Scotland to estimate house price capitalisation of an array of school characteristics. To isolate the effect of school performance on house prices from the one generated by other amenities, I compare properties in close proximity to catchment area boundaries. I contribute to the existing literature in four main ways. First, I reconcile previous finding on the lack of interest in value-added by showing that, even in presence of detailed information on school effectiveness, parents largely overlook this dimension; second, this is the first study examining parental preference for school-level non-cognitive skills, and finds that these do not affect school choice; third, despite the large set of indicators being used, I find that school quality is bi-dimensional, however, households mostly focus on peers' performance, even if this is driven by school composition; finally, while most of the literature focuses on one metropolitan area at the time, this is one of few studies looking at this phenomenon on national scale. As parents infer school quality from the average performance of the pupils attending it, rather than the contribution to their learning, future research should focus on the cognitive and non-cognitive return to effective schools versus those to 'better' peers.

Peer effects are soundly proven to be an important input of the education production function. In **Chapter 2** we study the educational gains occurring from sharing the classroom with older, more experienced, peers in the early stages of primary schools. We do so by levering multi-grading, i.e. when schools join adjacent year groups within the same classroom. Contrary to the existing literature focusing on institutional contexts whereby multi-grading is typically a rural schools' phenomenon, an important novelty of our work lies in the widespread use of multi-grade classes within the Scottish educational system, whereby these are used in rural as well as in urban areas. We disentangle the endogeneity of multi-grading by leveraging random fluctuation in population, jointly with an algorithm employed by Scottish local governments to decide on the optimal number of classrooms within each school. Hence, we construct an instrumental variable predicting the formation of multi-grade classes. We find that sharing a classroom with older peers leads to improvement in both literacy and numeracy. In addition, we do not find evidence that older peers gains are hindered by such a set up. While our results align with a small and recent literature on the cognitive effects of multi-grading, future work should investigate whether multi-grading affects non-cognitive development.

A healthy diet is a crucial determinant of physical, as well as cognitive and non-cognitive development. For some children, school meals represents a crucial component of their diets. In **Chapter 3** we conduct an evaluation of the extension of the provision of free school meals occurred in Scotland in 2015, which targeted pupils in grade 1, 2 and 3 of primary school. Unlike most of the literature to date, which focuses on the impact on test-scores, we look at the effect of such a policy on school attendance, health-related absences and exclusions. These outcomes are highly correlated with pupils' non-cognitive development. As the policy was implemented at the national level, we exploit cross-sectional variation in policy exposure, proxied by the share of pupils taking FSM prior to the change in policy, and estimate a difference-in-differences model with variation in treatment intensity. We find that such a policy did not lead to any improvements in the outcomes being considered. Potential extension of this work should examine how this policy affected households' finances and whether it improved uptakes among previously eligible children by reducing stigma. Additionally, as new data on absenteeism and discipline will be collected, future work should examine the impact of the recent extension of the provision to all grades.

Bibliography

Abdulkadiroğlu, Atila and Tayfun Sönmez, "School choice: A mechanism design approach," *American economic review*, 2003, *93* (3), 729–747.

—, Joshua Angrist, and Parag Pathak, "The elite illusion: Achievement effects at Boston and New York exam schools," *Econometrica*, 2014, *82* (1), 137–196.

—, Nikhil Agarwal, and Parag A Pathak, "The welfare effects of coordinated assignment: Evidence from the New York City high school match," *American Economic Review*, 2017, *107* (12), 3635–89.

—, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters, "Do Parents Value School Effectiveness?," *American Economic Review*, 2020, *110* (5), 1502–1539.

Abdulkadroğlu, Atila, Joshua D Angrist, Yusuke Narita, Parag A Pathak, and Roman A Zarate, "Regression Discontinuity in Serial Dictatorship: Achievement Effects at Chicago's Exam Schools," *American Economic Review*, 2017, *107* (5), 240–45.

Abouk, Rahi and Scott Adams, "Breakfast After the Bell: The Effects of Expanding Access to School Breakfasts on the Weight and Achievement of Elementary School Children," *Economics of Education Review*, 2022, *87*, 102224.

Adams, Peter, Michael D Hurd, Daniel McFadden, Angela Merrill, and Tiago Ribeiro, "Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status," *Journal of Econometrics*, 2003, *112* (1), 3–56.

Agarwal, Nikhil and Paulo Somaini, "Demand analysis using strategic reports: An application to a school choice mechanism," *Econometrica*, 2018, *86* (2), 391–444.

Agarwal, Sumit, Satyanarain Rengarajan, Tien Foo Sing, and Yang Yang, "School allocation rules and housing prices: A quasi-experiment with school relocation events in Singapore," *Regional Science and Urban Economics*, 2016, *58*, 42–56.

Ahn, Tom and Justin G Trogdon, "Peer delinquency and student achievement in middle school," *Labour Economics*, 2017, *44*, 192–217.

Ainsworth, Robert, Rajeev Dehejia, Cristian Pop-Eleches, and Miguel Urquiola, "Information, Preferences, and Household Demand for School Value Added," *IZA Discussion Paper No. 13980*, 2020.

Aizer, Anna, "Peer effects and human capital accumulation: The externalities of ADD," *National Bureau of Economic Research (NBER) Working Paper no. 14354*, 2008.

Alderman, Harold, John Hoddinott, and Bill Kinsey, "Long term consequences of early childhood malnutrition," *Oxford Economic Papers*, 2006, *58* (3), 450–474.

Altindag, Duha T, Deokrye Baek, Hong Lee, and Jessica Merkle, "Free lunch for all? The impact of universal school lunch on student misbehavior," *Economics of Education Review*, 2020, *74*, 101945.

Ammermueller, Andreas, "Violence in European schools: A widespread phenomenon that matters for educational production," *Labour Economics*, 2012, *19* (6), 908–922.

Anelli, Massimo and Giovanni Peri, "The effects of high school peers' gender on college major, college performance and income," *The Economic Journal*, 2019, *129* (618), 553–602.

Angrist, Joshua D and Guido W Imbens, "Two-stage least squares estimation of average causal effects in models with variable treatment intensity," *Journal of the American statistical Association*, 1995, *90* (430), 431–442.

— **and Kevin Lang**, "Does school integration generate peer effects? Evidence from Boston's Metco Program," *American Economic Review*, 2004, *94* (5), 1613–1634.

— **and Victor Lavy**, "Using Maimonides' rule to estimate the effect of class size on scholastic achievement," *The Quarterly Journal of Economics*, 1999, *114* (2), 533–575.

— **, — , Jetson Leder-Luis, and Adi Shany**, "Maimonides' Rule Redux," *American Economic Review: Insights*, 2019, *1* (3), 309–24.

Aucejo, Esteban M and Teresa Foy Romano, "Assessing the effect of school days and absences on test score performance," *Economics of Education Review*, 2016, *55*, 70–87.

Ballatore, Rosario Maria, Marco Paccagnella, and Marco Tonello, "Bullied because younger than my mates? The effect of age rank on victimisation at school," *Labour Economics*, 2020, *62*, 101772.

— **, Margherita Fort, and Andrea Ichino**, "Tower of Babel in the classroom: immigrants and natives in Italian schools," *Journal of Labor Economics*, 2018, *36* (4), 885–921.

Barbaranelli, Claudio, Gian Vittorio Caprara, Annarita Rabasca, and Concetta Pastorelli, "A questionnaire for measuring the Big Five in late childhood," *Personality and Individual Differences*, 2003, *34* (4), 645–664.

Barbetta, Gian Paolo, Giuseppe Sorrenti, and Gilberto Turati, "Multigrading and child achievement," *Journal of Human Resources*, 2019, pp. 0118–9310R4.

Batson, Tammy, "The Impact of Controlled Choice School Assignment Policy on Housing Prices: Evidence from Rockford, Illinois," *Measuring Amenities and Disamenities in the Housing Market: Applications of the Hedonic Method*, 2021, pp. 121–140.

Battistin, Erich and Lorenzo Neri, "School performance, score inflation and economic geography," *CEPR Discussion Paper No. DP12445*, 2017.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan, "A unified framework for measuring preferences for schools and neighborhoods," *Journal of Political Economy*, 2007, *115* (4), 588–638.

Beaton, Maura, P Craig, SV Katikireddi, R Jepson, and A Williams, "Evaluability assessment of free school meals for all children in P1 to P3," *Proj. Rep. NHS Health Scotl. Edinb*, 2014.

Bedard, Kelly and Elizabeth Dhuey, "The persistence of early childhood maturity: International evidence of long-run age effects," *The Quarterly Journal of Economics*, 2006, *121* (4), 1437–1472.

Belot, Michèle and Jonathan James, "Healthy school meals and educational outcomes," *Journal of Health Economics*, 2011, *30* (3), 489–504.

Benton, David, "The impact of diet on anti-social, violent and criminal behaviour," *Neuroscience & Biobehavioral Reviews*, 2007, *31* (5), 752–774.

Bertoni, Marco and Roberto Nisticò, "Ordinal rank and the structure of ability peer effects," *Journal of Public Economics*, 2023, *217*, 104797.

Bertrand, Marianne and Jessica Pan, "The trouble with boys: Social influences and the gender gap in disruptive behavior," *American Economic Journal: Applied Economics*, 2013, *5* (1), 32–64.

Betts, Julian R, "The economics of tracking in education," in "Handbook of the Economics of Education," Vol. 3, Elsevier, 2011, pp. 341–381.

Beuermann, Diether, C Kirabo Jackson, Laia Navarro-Sola, and Francisco Pardo, "What is a Good School, and Can Parents Tell? Evidence on the Multidimensionality of School Output," *National Bureau of Economic Research, Working Paper No. 25342*, 2018.

Beuermann, Diether W and C Kirabo Jackson, "The Short-and Long-Run Effects of Attending the Schools that Parents Prefer," *Journal of Human Resources*, 2022, *57* (3), 725–746.

Bhattacharya, Aveek, "How much choice is enough? Parental satisfaction with secondary school choice in England and Scotland," *Journal of Social Policy*, 2021, pp. 1–21.

Bhattacharya, Jayanta, Janet Currie, and Steven J Haider, "Breakfast of champions? The School Breakfast Program and the nutrition of children and families," *Journal of Human Resources*, 2006, *41* (3), 445–466.

Bifulco, Robert, Jason M Fletcher, and Stephen L Ross, "The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health," *American Economic Journal: Economic Policy*, 2011, *3* (1), 25–53.

_ , _ , Sun Jung Oh, and Stephen L Ross, "Do high school peers have persistent effects on college attainment and other life outcomes?," *Labour Economics*, 2014, *29*, 83–90.

Black, Sandra E, "Do better schools matter? Parental valuation of elementary education," *The Quarterly Journal of Economics*, 1999, *114* (2), 577–599.

_ , Paul J Devereux, and Kjell G Salvanes, "Too young to leave the nest? The effects of school starting age," *The Review of Economics and Statistics*, 2011, *93* (2), 455–467.

_ , _ , and _ , "Under pressure? The effect of peers on outcomes of young adults," *Journal of Labor Economics*, 2013, *31* (1), 119–153.

Bogart, William T and Brian A Cromwell, "How much is a neighborhood school worth?," *Journal of Urban Economics*, 2000, *47* (2), 280–305.

Bonilla-Mejía, Leonardo, Esteban Lopez, and Daniel McMillen, "House prices and school choice: Evidence from Chicago's magnet schools' proximity lottery," *Journal of Regional Science*, 2020, *60* (1), 33–55.

Bono, Emilia Del, Marco Francesconi, Yvonne Kelly, and Amanda Sacker, "Early maternal time investment and early child outcomes," *The Economic Journal*, 2016, *126* (596), F96–F135.

Booij, Adam S, Edwin Leuven, and Hessel Oosterbeek, "Ability peer effects in university: Evidence from a randomized experiment," *The review of economic studies*, 2017, *84* (2), 547–578.

Borbely, Daniel, Markus Gehrsitz, Stuart McIntyre, Gennaro Rossi, and Graeme Roy, "Early-Years Multi-Grade Classes and Pupil Attainment," *IZA Discussion Paper No. 14678*, 2021.

Bound, John, David A Jaeger, and Regina M Baker, "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak," *Journal of the American Statistical Association*, 1995, *90* (430), 443–450.

Bowes, Lucy, Carol Joinson, Dieter Wolke, and Glyn Lewis, "Peer victimisation during adolescence and its impact on depression in early adulthood: prospective cohort study in the United Kingdom," *bmj*, 2015, *350.*

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin, "Identification of peer effects through social networks," *Journal of econometrics*, 2009, *150* (1), 41–55.

Brasington, David and Donald R Haurin, "Educational outcomes and house values: A test of the value added approach," *Journal of Regional Science*, 2006, *46* (2), 245–268.

Brown, Sarah and Karl Taylor, "Bullying, education and earnings: evidence from the National Child Development Study," *Economics of Education Review*, 2008, *27* (4), 387–401.

Burgess, Simon, Ellen Greaves, and Anna Vignoles, "School choice in England: evidence from national administrative data," *Oxford Review of Education*, 2019, *45* (5), 690–710.

_ , _ , _ , and Deborah Wilson, "What parents want: School preferences and school choice," *The Economic Journal*, 2015, *125* (587), 1262–1289.

Burke, Mary A and Tim R Sass, "Classroom peer effects and student achievement," *Journal of Labor Economics*, 2013, *31* (1), 51–82.

Bütikofer, Aline, Eirin Mølland, and Kjell G Salvanes, "Childhood nutrition and labor market outcomes: Evidence from a school breakfast program," *Journal of Public Economics*, 2018, *168*, 62–80.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro HC Sant'Anna, "Difference-in-differences with a continuous treatment," *arXiv preprint arXiv:2107.02637*, 2021.

Calsamiglia, Caterina and Maia Güell, "Priorities in school choice: The case of the Boston mechanism in Barcelona," *Journal of Public Economics*, 2018, *163*, 20–36.

_ , Chao Fu, and Maia Güell, "Structural estimation of a model of school choices: The boston mechanism versus its alternatives," *Journal of Political Economy*, 2020, *128* (2), 642–680.

Card, David, "Using regional variation in wages to measure the effects of the federal minimum wage," *Industrial and Labor Relations Review*, 1992, *46* (1), 22–37.

Carneiro, Pedro, Claire Crawford, and Alissa Goodman, "The impact of early cognitive and non-cognitive skills on later outcomes," *Centre for Economics of Education*, 2007.

Carrell, Scott E and Mark Hoekstra, "Family business or social problem? The cost of unreported domestic violence," *Journal of Policy Analysis and Management*, 2012, *31* (4), 861–875.

_ and Mark L Hoekstra, "Externalities in the classroom: How children exposed to domestic violence affect everyone's kids," *American Economic Journal: Applied Economics*, 2010, *2* (1), 211–28.

_ , Mark Hoekstra, and Elira Kuka, "The long-run effects of disruptive peers," *American Economic Review*, 2018, *108* (11), 3377–3415.

_ , _ , and _ , "The long-run effects of disruptive peers," *American Economic Review*, 2018, *108* (11), 3377–3415.

Cascio, Elizabeth U and Diane Whitmore Schanzenbach, "First in the class? Age and the education production function," *Education Finance and Policy*, 2016, *11* (3), 225–250.

Chakraborty, Tanika and Rajshri Jayaraman, "School feeding and learning achievement: evidence from India's midday meal program," *Journal of Development Economics*, 2019, *139*, 249–265.

Chambers, Stephanie, Allison Ford, Nicola Boydell, Laurence Moore, Martine Stead, and Douglas Eadie, "Universal free school meals in Scotland: A process evaluation of implementation and uptake: Stephanie Chambers," *The European Journal of Public Health*, 2016, *26* (suppl_1), ckw169–025.

Chan, Jimmy, Xian Fang, Zhi Wang, Xianhua Zai, and Qinghua Zhang, "Valuing primary schools in urban China," *Journal of Urban Economics*, 2020, *115*, 103183.

Checchi, Daniele and Maria De Paola, "The effect of multigrade classes on cognitive and non-cognitive skills. Causal evidence exploiting minimum class size rules in Italy," *Economics of Education Review*, 2018, *67*, 235–253.

Chetty, Raj, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan, "How does your kindergarten classroom affect your earnings? Evidence from Project STAR," *The Quarterly Journal of Economics*, 2011, *126* (4), 1593–1660.

Cheung, Ka Shing, Chung Yim Yiu, and Yuyu Zhang, "What matters more, school choices or neighbourhoods? Evidence from a socioeconomic based school zoning," *Cities*, 2022, *128*, 103772.

Chung, Il Hwan, "School choice, housing prices, and residential sorting: Empirical evidence from inter-and intra-district choice," *Regional science and urban economics*, 2015, *52*, 39–49.

Clapp, John M, Anupam Nanda, and Stephen L Ross, "Which school attributes matter? The influence of school district performance and demographic composition on property values," *Journal of Urban Economics*, 2008, *63* (2), 451–466.

Clark, Damon, "Selective schools and academic achievement," *The BE Journal of Economic Analysis & Policy*, 2010, *10* (1).

_ and Emilia Del Bono, "The long-run effects of attending an elite school: evidence from the United Kingdom," *American Economic Journal: Applied Economics*, 2016, *8* (1), 150–76.

Clemens, Michael A, Ethan G Lewis, and Hannah M Postel, "Immigration restrictions as active labor market policy: Evidence from the mexican bracero exclusion," *American Economic Review*, 2018, *108* (6), 1468–87.

Cohen, Juliana FW, Amelie A Hecht, Gabriella M McLoughlin, Lindsey Turner, and Marlene B Schwartz, "Universal school meals and associations with student participation, attendance, academic performance, diet quality, food security, and body mass index: A systematic review," *Nutrients*, 2021, *13* (3), 911.

Corcoran, Sean P, Brian Elbel, and Amy Ellen Schwartz, "The effect of breakfast in the classroom on obesity and academic performance: Evidence from New York City," *Journal of Policy Analysis and Management*, 2016, *35* (3), 509–532.

Crawford, Claire, Lorraine Dearden, and Ellen Greaves, "The drivers of month-of-birth differences in children's cognitive and non-cognitive skills," *Journal of the Royal Statistical Society. Series A,(Statistics in Society)*, 2014, *177* (4), 829.

Crosnoe, Robert, Shannon Cavanagh, and Glen H Elder Jr, "Adolescent friendships as academic resources: The intersection of friendship, race, and school disadvantage," *Sociological perspectives*, 2003, *46* (3), 331–352.

Cuadros-Meñaca, Andres, Michael R Thomsen, and Rodolfo M Nayga Jr, "The Effect of School Breakfast on Student Behavior: An Evaluation of Breakfast After the Bell," *Available at SSRN 3806620*, 2021.

_ , _ , and _ , "The effect of breakfast after the bell on student academic achievement," *Economics of Education Review*, 2022, *86*, 102223.

Cullen, Julie Berry, Brian A Jacob, and Steven Levitt, "The effect of school choice on participants: Evidence from randomized lotteries," *Econometrica*, 2006, *74* (5), 1191–1230.

Davidoff, Ian and Andrew Leigh, "How much do public schools really cost? Estimating the relationship between house prices and school quality," *Economic Record*, 2008, *84* (265), 193–206.

Dee, Thomas and Xiaohuan Lan, "The achievement and course-taking effects of magnet schools: Regression-discontinuity evidence from urban China," *Economics of Education Review*, 2015, *47*, 128–142.

Deming, David J, "Better schools, less crime?," *The Quarterly Journal of Economics*, 2011, *126* (4), 2063–2115.

_ , Justine S Hastings, Thomas J Kane, and Douglas O Staiger, "School choice, school quality, and postsecondary attainment," *American Economic Review*, 2014, *104* (3), 991–1013.

Dickerson, Andy, Konstantina Maragkou, and Steve McIntosh, "The causal effect of secondary school peers on educational aspirations," *Centre for Vocational Education Research Discussion Paper*, 2018, *17.*

Ding, Weili and Steven F Lehrer, "Do peers affect student achievement in China's secondary schools?," *The Review of Economics and Statistics*, 2007, *89* (2), 300–312.

Doris, Aedín, Donal O'Neill, and Olive Sweetman, "Good schools or good students? The importance of selectivity for school rankings," *Oxford Review of Education*, 2022, pp. 1–23.

Duckworth, Angela L, Christopher Peterson, Michael D Matthews, and Dennis R Kelly, "Grit: perseverance and passion for long-term goals.," *Journal of Personality and Social Psychology*, 2007, *92* (6), 1087.

Duflo, Esther, Pascaline Dupas, and Michael Kremer, "Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya," *American Economic Review*, 2011, *101* (5), 1739–74.

Dustan, Andrew, Alain De Janvry, and Elisabeth Sadoulet, "Flourish or fail? The risky reward of elite high school admission in Mexico City," *Journal of Human Resources*, 2017, *52* (3), 756–799.

Elder, Todd E and Darren H Lubotsky, "Kindergarten entrance age and children's achievement impacts of state policies, family background, and peers," *Journal of human Resources*, 2009, *44* (3), 641–683.

Elsner, Benjamin and Ingo E Isphording, "A big fish in a small pond: Ability rank and human capital investment," *Journal of Labor Economics*, 2017, *35* (3), 787–828.

Epple, Dennis N and Richard Romano, "Neighborhood schools, choice, and the distribution of educational benefits," in "The Economics of School Choice," University of Chicago Press, 2003, pp. 227–286.

Eriksen, Tine Louise Mundbjerg, Helena Skyt Nielsen, and Marianne Simonsen, "Bullying in elementary school," *Journal of Human Resources*, 2014, *49* (4), 839–871.

Fack, Gabrielle and Julien Grenet, "When do better schools raise housing prices? Evidence from Paris public and private schools," *Journal of Public Economics*, 2010, *94* (1-2), 59–77.

_, _, and Yinghua He, "Beyond truth-telling: Preference estimation with centralized school choice and college admissions," *American Economic Review*, 2019, *109* (4), 1486–1529.

Figlio, David N, "Boys named Sue: Disruptive children and their peers," *Education finance and policy*, 2007, *2* (4), 376–394.

_, "Boys named Sue: Disruptive children and their peers," *Education Finance and Policy*, 2007, *2* (4), 376–394.

_ **and Maurice E Lucas**, "What's in a grade? School report cards and the housing market," *American Economic Review*, 2004, *94* (3), 591–604.

Fiva, Jon H and Lars J Kirkebøen, "Information shocks and the dynamics of the housing market," *The Scandinavian Journal of Economics*, 2011, *113* (3), 525–552.

Flowerdew, R, E Graham, and Z Feng, "The production of an updated set of data zones to incorporate 2001 census geography and data," *Report to the Scottish Executive School of Geography and Geosciences. Scotland: University of St Andrews*, 2004.

Ford, Allison, Douglas Eadie, and Martine Stead, "Process evaluation of the implementation of Universal Free School Meals: Research with parents," *NHS Health Scotland*, 2015.

Fort, Margherita, Andrea Ichino, and Giulio Zanella, "Cognitive and noncognitive costs of day care at age 0–2 for children in advantaged families," *Journal of Political Economy*, 2020, *128* (1), 158–205.

Francesconi, Marco and James J Heckman, "Child development and parental investment: Introduction," *The Economic Journal*, 2016, *126* (596), F1–F27.

Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek, "Parental responses to public investments in children: Evidence from a maximum class size rule," *Journal of Human Resources*, 2016, *51* (4), 832–868.

Fruehwirth, Jane Cooley, "Identifying peer achievement spillovers: Implications for desegregation and the achievement gap," *Quantitative Economics*, 2013, *4* (1), 85–124.

Fumarco, Luca and Gabriel Schultze, "Does relative age make Jack a dull student? Evidence from students' schoolwork and playtime," *Education Economics*, 2020, *28* (6), 647–670.

Gershenson, Seth, Alison Jacknowitz, and Andrew Brannegan, "Are student absences worth the worry in US primary schools?," *Education Finance and Policy*, 2017, *12* (2), 137–165.

Gibbons, Stephen and Stephen Machin, "Paying for primary schools: admission constraints, school popularity or congestion?," *The Economic Journal*, 2006, *116* (510), C77–C92.

_, _, and Olmo Silva, "Valuing school quality using boundary discontinuities," *Journal of Urban Economics*, 2013, *75*, 15–28.

Gibbons, Steve and Stephen Machin, "Valuing English primary schools," *Journal of Urban Economics*, 2003, *53* (2), 197–219.

Giorgi, Giacomo De and Michele Pellizzari, "Understanding social interactions: Evidence from the classroom," *The Economic Journal*, 2014, *124* (579), 917–953.

_, _, and Silvia Redaelli, "Identification of social interactions through partially overlapping peer groups," *American Economic Journal: Applied Economics*, 2010, *2* (2), 241–75.

Glazerman, Steven and Dallas Dotter, "Market signals: Evidence on the determinants and consequences of school choice from a citywide lottery," *Educational Evaluation and Policy Analysis*, 2017, *39* (4), 593–619.

Glewwe, Paul, Hanan G Jacoby, and Elizabeth M King, "Early childhood nutrition and academic achievement: a longitudinal analysis," *Journal of Public Economics*, 2001, *81* (3), 345–368.

Gordanier, John, Orgul Ozturk, Breyon Williams, and Crystal Zhan, "Free lunch for all! the effect of the community eligibility provision on academic outcomes," *Economics of Education Review*, 2020, *77*, 101999.

Gordon, Nora and Krista Ruffini, "Schoolwide free meals and student discipline: Effects of the community eligibility provision," *Education Finance and Policy*, 2021, *16* (3), 418–442.

Gorman, Emma and Ian Walker, "Heterogeneous effects of missing out on a place at a preferred secondary school in England," *Economics of Education Review*, 2021, *81*, 102082.

Gottfried, Michael A, "Excused versus unexcused: How student absences in elementary school affect academic achievement," *Educational Evaluation and Policy Analysis*, 2009, *31* (4), 392–415.

_ , "The detrimental effects of missing school: Evidence from urban siblings," *American Journal of Education*, 2011, *117* (2), 147–182.

Gould, Eric D, Victor Lavy, and M Daniele Paserman, "Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence," *The Economic Journal*, 2009, *119* (540), 1243–1269.

Greaves, Ellen, "How do schools shape neighbourhoods? Endogenous residential location in response to local school quality," *Working Paper*, 2022.

_ **and Hélène Turon**, "School Choice and Neighbourhood Sorting," *Working Paper*, 2021.

_ **, Iftikhar Hussain, Birgitta Rabe, and Imran Rasul**, "Parental Responses to Information About School Quality: Evidence from Linked Survey and Administrative Data," *CEPR Discussion Paper No. DP16406*, 2021.

Griffith, Rachel, Rodrigo Lluberas, and Melanie Lührmann, "Gluttony and sloth? Calories, labor market activity and the rise of obesity," *Journal of the European Economic Association*, 2016, *14* (6), 1253–1286.

Han, Jun, Liyuan Cui, and Huayi Yu, "Pricing the value of the chance to gain admission to an elite senior high school in Beijing: The effect of the LDHSE policy on resale housing prices," *Cities*, 2021, *115*, 103238.

Handbury, Jessie and Sarah Moshary, "School food policy affects everyone: Retail responses to the national school lunch program," *NBER Wroking Paper No. 29384*, 2021.

Hanushek, Eric A and Steven G Rivkin, "Harming the best: How schools affect the black-white achievement gap," *Journal of Policy Analysis and Management*, 2009, *28* (3), 366–393.

_ **, John F Kain, and Steven G Rivkin**, "New evidence about Brown v. Board of Education: The complex effects of school racial composition on achievement," *Journal of Labor Economics*, 2009, *27* (3), 349–383.

_ **,** _ **, Jacob M Markman, and Steven G Rivkin**, "Does peer ability affect student achievement?," *Journal of Applied Econometrics*, 2003, *18* (5), 527–544.

Harjunen, Oskari, Mika Kortelainen, and Tuukka Saarimaa, "Best education money can buy? Capitalization of school quality in Finland," *CESifo Economic Studies*, 2018, *64* (2), 150–175.

Harris, Douglas N and Matthew Larsen, "What schools do families want (and why)," *Policy Brief (New Orleans, LA: Education Research Alliance for New Orleans*, 2015.

Hassani-Nezhad, Lena, Dan Anderberg, Arnaud Chevalier, Melanie Lührmann, and Ronni Pavan, "Higher education financing and the educational aspirations of teenagers and their parents," *Economics of Education Review*, 2021, *85*, 102175.

Hastings, Justine S and Jeffrey M Weinstein, "Information, school choice, and academic achievement: Evidence from two experiments," *The Quarterly Journal of Economics*, 2008, *123* (4), 1373–1414.

Hastings, Justine, Thomas J Kane, and Douglas O Staiger, "Heterogeneous preferences and the efficacy of public school choice," *NBER Working Paper*, 2009, *2145*, 1–46.

Heckman, James J and Stefano Mosso, "The economics of human development and social mobility," *Annual Review of Economics*, 2014, *6* (1), 689–733.

— **and Tim Kautz**, "Hard evidence on soft skills," *Labour Economics*, 2012, *19* (4), 451–464.

— **and Yona Rubinstein**, "The importance of noncognitive skills: Lessons from the GED testing program," *American Economic Review*, 2001, *91* (2), 145–149.

— **, John Eric Humphries, and Gregory Veramendi**, "Dynamic treatment effects," *Journal of Econometrics*, 2016, *191* (2), 276–292.

Hertz, Marci Feldman, Ingrid Donato, and James Wright, "Bullying and suicide: a public health approach," *Journal of Adolescent Health*, 2013, *53* (1), S1–S3.

Hoekstra, Mark, Pierre Mouganie, and Yaojing Wang, "Peer quality and the academic benefits to attending better schools," *Journal of Labor Economics*, 2018, *36* (4), 841–884.

Holford, Angus, "Take-up of free school meals: price effects and peer effects," *Economica*, 2015, *82* (328), 976–993.

— **and Birgitta Rabe**, "Impact of the Universal Infant Free School Meal Policy," *Institute for Social and Economic Research*, 2020.

— **and** — , "Going universal. The impact of free school lunches on child body weight outcomes," *Journal of Public Economics Plus*, 2022, p. 100016.

Hoxby, Caroline, "Peer effects in the classroom: Learning from gender and race variation," *National Bureau of Economic Research (NBER) Working Paper no. 7867*, 2000.

Hoxby, Caroline M, "Does competition among public schools benefit students and taxpayers?," *American Economic Review*, 2000, *90* (5), 1209–1238.

— **and Gretchen Weingarth**, "Taking race out of the equation: School reassignment and the structure of peer effects," *Unpublished Manuscript*, 2005.

Hoxby, Caroline Minter, "School choice and school productivity. Could school choice be a tide that lifts all boats?," in "The economics of school choice," University of Chicago Press, 2003, pp. 287–342.

Imberman, Scott A and Michael F Lovenheim, "Does the market value value-added? Evidence from housing prices after a public release of school and teacher value-added," *Journal of Urban Economics*, 2016, *91*, 104–121.

Ismail, Mohammad, Abukar Warsame, and Mats Wilhelmsson, "Schools' capitalization into housing values in a context of free school choices," *Working Paper*, 2022.

Jackson, C Kirabo, "Do students benefit from attending better schools? Evidence from rule-based student assignments in Trinidad and Tobago," *The Economic Journal*, 2010, *120* (549), 1399–1429.

Jackson, C Kirabo, "What do test scores miss? The importance of teacher effects on non–test score outcomes," *Journal of Political Economy*, 2018, *126* (5), 2072–2107.

Jaeger, David A, Theodore J Joyce, and Robert Kaestner, "A cautionary tale of evaluating identifying assumptions: did reality TV really cause a decline in teenage childbearing?," *Journal of Business & Economic Statistics*, 2020, *38* (2), 317–326.

John, Oliver P, Avshalom Caspi, Richard W Robins, Terrie E Moffitt, and Magda Stouthamer-Loeber, "The "little five": Exploring the nomological network of the five-factor model of personality in adolescent boys," *Child Development*, 1994, *65* (1), 160–178.

Kahn-Lang, Ariella and Kevin Lang, "The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications," *Journal of Business & Economic Statistics*, 2020, *38* (3), 613–620.

Kane, Thomas J, Douglas O Staiger, Gavin Samms, Edward W Hill, and David L Weimer, "School accountability ratings and housing values [with comments]," *Brookings-Wharton Papers on Urban Affairs*, 2003, pp. 83–137.

Kapor, Adam J, Christopher A Neilson, and Seth D Zimmerman, "Heterogeneous beliefs and school choice mechanisms," *American Economic Review*, 2020, *110* (5), 1274–1315.

Kautz, Tim, James J Heckman, Ron Diris, Bas Ter Weel, and Lex Borghans, "Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success," *National Bureau of Economic Research, Working Paper No. 20749*, 2014.

Kim, Young Shin and Bennett Leventhal, "Bullying and suicide. A review," *International Journal of Adolescent Medicine and Health*, 2008, *20* (2), 133–154.

Kinsler, Josh and Ronni Pavan, "Local distortions in parental beliefs over child skill," *Journal of Political Economy*, 2021, *129* (1), 81–100.

Kleibergen, Frank and Richard Paap, "Generalized reduced rank tests using the singular value decomposition," *Journal of Econometrics*, 2006, *133* (1), 97–126.

Klein, Markus, Edward M Sosu, and Shadrach Dare, "School absenteeism and academic achievement: does the reason for absence matter?," *AERA Open*, 2022, *8*, 23328584211071115.

Lavy, Victor, "Effects of free choice among public schools," *The Review of Economic Studies*, 2010, *77* (3), 1164–1191.

_ **and Analia Schlosser**, "Mechanisms and impacts of gender peer effects at school," *American Economic Journal: Applied Economics*, 2011, *3* (2), 1–33.

_ **and Edith Sand**, "The effect of social networks on students' academic and non-cognitive behavioural outcomes: evidence from conditional random assignment of friends in school," *The Economic Journal*, 2019, *129* (617), 439–480.

_ **, M Daniele Paserman, and Analia Schlosser**, "Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom," *The Economic Journal*, 2012, *122* (559), 208–237.

_ **, Olmo Silva, and Felix Weinhardt**, "The good, the bad, and the average: Evidence on ability peer effects in schools," *Journal of Labor Economics*, 2012, *30* (2), 367–414.

Learning Directorate, "Choosing a school: a guide for parents, revised November 2016," https://www.gov.scot/publications/choosing-school-guide-parents-nov-16/ 2016.

Lee, David L, Justin McCrary, Marcelo J Moreira, and Jack Porter, "Valid t-ratio Inference for IV," *arXiv preprint arXiv:2010.05058*, 2020.

Lee, David S and Thomas Lemieux, "Regression discontinuity designs in economics," *Journal of Economic Literature*, 2010, *48* (2), 281–355.

Lee, Yong Suk, "School districting and the origins of residential land price inequality," *Journal of Housing Economics*, 2015, *28*, 1–17.

Lefgren, Lars, "Educational peer effects and the Chicago public schools," *Journal of Urban Economics*, 2004, *56* (2), 169–191.

Leos-Urbel, Jacob, Amy Ellen Schwartz, Meryle Weinstein, and Sean Corcoran, "Not just for poor kids: The impact of universal free school breakfast on meal participation and student outcomes," *Economics of Education Review*, 2013, *36*, 88–107.

Leuven, Edwin and Marte Rønning, "Classroom grade composition and pupil achievement," *The Economic Journal*, 2014, *126* (593), 1164–1192.

_ **, Hessel Oosterbeek, and Marte Rønning**, "Quasi-experimental estimates of the effect of class size on achievement in Norway," *Scandinavian Journal of Economics*, 2008, *110* (4), 663–693.

Lindqvist, Erik and Roine Vestman, "The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment," *American Economic Journal: Applied Economics*, 2011, *3* (1), 101–28.

Liu, Jing, Monica Lee, and Seth Gershenson, "The short-and long-run impacts of secondary school absences," *Journal of Public Economics*, 2021, *199*, 104441.

Lleras, Christy, "Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings," *Social Science Research*, 2008, *37* (3), 888–902.

Lucas, Adrienne M and Isaac M Mbiti, "Effects of school quality on student achievement: Discontinuity evidence from kenya," *American Economic Journal: Applied Economics*, 2014, *6* (3), 234–63.

Lundborg, Petter, Dan-Olof Rooth, and Jesper Alex-Petersen, "Long-Term Effects of Childhood Nutrition: Evidence from a School Lunch Reform," *The Review of Economic Studies*, 2021.

Machin, Stephen and Kjell G Salvanes, "Valuing school quality via a school choice reform," *The Scandinavian Journal of Economics*, 2016, *118* (1), 3–24.

MacLardie, Jane, Chris Martin, Lorraine Murray, and Kate Sewel, "Evaluation of the free school meals trial for P1 to P3 pupils.," 2008, *Technical Report*, Scottish Government Social Research.

MacLeod, W Bentley and Miguel Urquiola, "Reputation and school competition," *American Economic Review*, 2015, *105* (11), 3471–88.

Marcus, Michelle M and Katherine G Yewell, "The Effect of Free School Meals on Household Food Purchases: Evidence from the Community Eligibility Provision," *NBER Working Paper No. 29395*, 2021.

McAdams, Rachel, "Evaluating universal infant free school meals, Scotland: second monitoring report of school-meal uptake: October 2016," 2016.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

Muralidharan, Karthik and Nishith Prakash, "Cycling to school: Increasing secondary school enrollment for girls in India," *American Economic Journal: Applied Economics*, 2017, *9* (3), 321–50.

National Statistics, "Summary statistics for attainment, leaver destination and healthy living," 2015.

Neidell, Matthew and Jane Waldfogel, "Cognitive and noncognitive peer effects in early education," *The Review of Economics and Statistics*, 2010, *92* (3), 562–576.

Neri, Lorenzo et al., "Moving Opportunities: The Impact of Public Housing Re-generations on Student Achievement," *School of Economics and Finance Working Paper No. 907, Queen Mary University of London*, 2020.

Norris, Jonathan, "Peers, parents, and attitudes about school," *Journal of Human Capital*, 2020, *14* (2), 290–342.

Orford, Scott, "The capitalisation of school choice into property prices: a case study of grammar and all ability state schools in Buckinghamshire, UK," *Geoforum*, 2018, *97*, 231–241.

Ovidi, Marco, "Parents know better: Primary school choice and student achievement in London," *Working Paper*, 2021.

Ozturk, Orgul, Pelin Pekgun, and Krista Ruffini, "Free School Meals and Demand for Community Food Resources," *Available at SSRN 3899868*, 2021.

Park, Hyunjun, Alan Tidwell, Sungho Yun, and Changha Jin, "Does school choice program affect local housing prices?: Inter-vs. intra-district choice program," *Cities*, 2021, *115*, 103237.

Parnham, Jennie C, Kiara Chang, Christopher Millett, Anthony A Laverty, Stephanie von Hinke, Jonathan Pearson-Stuttard, Frank de Vocht, Martin White, and Eszter P Vamos, "The impact of the Universal Infant Free School Meal policy on dietary quality in English and Scottish primary school children: evaluation of a natural experiment," *Nutrients*, 2022, *14* (8), 1602.

Patacchini, Eleonora, Edoardo Rainone, and Yves Zenou, "Heterogeneous peer effects in education," *Journal of Economic Behavior & Organization*, 2017, *134*, 190–227.

Peña, Pablo A, "Creating winners and losers: Date of birth, relative age in school, and outcomes in childhood and adulthood," *Economics of Education Review*, 2017, *56*, 152–176.

Piotrowska, Patrycja J, Christopher B Stride, Simone E Croft, and Richard Rowe, "Socioeconomic status and antisocial behaviour among children and adolescents: A systematic review and meta-analysis," *Clinical Psychology Review*, 2015, *35*, 47–55.

Polonsky, Heather M, Katherine W Bauer, Jennifer O Fisher, Adam Davey, Sandra Sherman, Michelle L Abel, Alexandra Hanlon, Karen J Ruth, Lauren C Dale, and Gary D Foster, "Effect of a breakfast in the classroom initiative on obesity in urban school-aged children: a cluster randomized clinical trial," *JAMA Pediatrics*, 2019, *173* (4), 326–333.

Ponzo, Michela, "Does bullying reduce educational achievement? An evaluation using matching estimators," *Journal of Policy Modeling*, 2013, *35* (6), 1057–1078.

Pop-Eleches, Cristian and Miguel Urquiola, "Going to a better school: Effects and behavioral responses," *American Economic Review*, 2013, *103* (4), 1289–1324.

Public Health Scotland, "Body Mass Index of Primary 1 Children in Scotland," 2021.

Reback, Randall, "House prices and the provision of local public services: capitalization under school choice programs," *Journal of Urban Economics*, 2005, *57* (2), 275–301.

Ries, John and Tsur Somerville, "School quality and residential property values: evidence from Vancouver rezoning," *The Review of Economics and Statistics*, 2010, *92* (4), 928–944.

Rossi, Gennaro, "School Performance, non-cognitive skills and house prices," *Strathclyde Discussion Papers in Economics no. 21 - 2*, 2021.

Ruffini, Krista, "Universal access to free school meals and student achievement: Evidence from the Community Eligibility Provision," *Journal of Human Resources*, 2021, pp. 0518–9509R3.

Sacerdote, Bruce, "Peer effects in education: How might they work, how big are they and how much do we know thus far?," in "Handbook of the Economics of Education," Vol. 3, Elsevier, 2011, pp. 249–277.

Schanzenbach, Diane Whitmore, "Do school lunches contribute to childhood obesity?," *Journal of Human Resources*, 2009, *44* (3), 684–709.

Schwartz, Amy Ellen and Michah W Rothbart, "Let them eat lunch: The impact of universal free meals on student performance," *Journal of Policy Analysis and Management*, 2020, *39* (2), 376–410.

Scottish, Government, "Hungry for Success: A Whole School Approach to School Meals in Scotland," *Scottish Government*, 2003.

__ , "Better Eating, Better Learning: A new context for school food," *Scottish Government*, 2014.

Scottish, Parliament, "Schools (Health Promotion and Nutrition) Scotland Act 2007," *Scottish Parliament*, 2007.

__ , "Nutritional Requirements for Food and Drink in Schools (Scotland) Regulations 2008," *Scottish Parliament*, 2008.

SG Education Analytical Services: Learning Analysis, "School Information Dashboard - Secondary," https://public.tableau.com/app/profile/sg.eas.learninganalysis/viz/SchoolInformationDashboard-Secondary/Introduction 2018.

Sheppard, Stephen, "Hedonic analysis of housing markets," *Handbook of Regional and Urban Economics*, 1999, *3*, 1595–1635.

Shure, Nikki, "Non-cognitive peer effects in secondary education," *Labour Economics*, 2021, *73* (102074), 1–26.

Sims, David, "A strategic response to class size reduction: Combination classes and student achievement in California," *Journal of Policy Analysis and Management*, 2008, pp. 457–478.

Slavin, Robert E, "Ability grouping and student achievement in elementary schools: A best-evidence synthesis," *Review of Educational Research*, 1987, *57* (3), 293–336.

Sorhaindo, Annik and Leon Feinstein, "What is the relationship between child nutrition and school outcomes?," *Journal of the Home Economics Institute of Australia*, 2006, *13* (3), 21–23.

Sosu, Edward M, Shadrach Dare, Claire Goodfellow, and Markus Klein, "Socioeconomic status and school absenteeism: A systematic review and narrative synthesis," *Review of Education*, 2021, *9* (3), e3291.

Stock, James H and Motohiro Yogo, "Testing for weak instruments in linear IV regression," *National Bureau of Economic Research (NBER) Working Paper no. 0284*, 2002.

The Edinburgh Council, "Find Your Catchment School," https://www.edinburgh.gov.uk/catchmentmap 2019.

Valuation Office Agency, "How domestic properties are assessed for Council Tax bands," https://www.gov.uk/guidance/understand-how-council-tax-bands-are-assessed 2022.

Victora, Cesar G, Linda Adair, Caroline Fall, Pedro C Hallal, Reynaldo Martorell, Linda Richter, Harshpal Singh Sachdev, Maternal, Child Undernutrition Study Group et al., "Maternal and child undernutrition: consequences for adult health and human capital," *The Lancet*, 2008, *371* (9609), 340–357.

von Hinke Kessler Scholder, Stephanie, "School meal crowd out in the 1980s," *Journal of Health Economics*, 2013, *32* (3), 538–545.

Winicki, Joshua and Kyle Jemison, "Food insecurity and hunger in the kindergarten classroom: its effect on learning and growth," *Contemporary Economic Policy*, 2003, *21* (2), 145–157.

Wong, Siu Kei and Kuang Kuang Deng, "School Catchment Zone Mergers and Housing Wealth Redistribution," *Journal of Planning Education and Research*, 2021, p. 0739456X211006760.

World Health Organization and others, "Childhood obesity surveillance initiative highlights 2015-17 preliminary data," 2018.

Zheng, Angela, "The Valuation of Local School Quality under School Choice," *American Economic Journal: Economic Policy*, 2022, *14* (2), 509–37.

Zou, Yonghua, "The Effects of Double-Attendance Zones and School Rankings on Housing Prices: Case of Nanjing, China," *Journal of Urban Planning and Development*, 2022, *148* (1), 05021064.

Appendix A

Appendix - Chapter 1

Table A.1: Summary Statistics

Panel A: House Prices and School-Level Variables	Full Sample		350 metres		300 metres		250 metres	
	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.
ln(House Price)	11.81	0.68	11.84	0.69	11.84	0.69	11.85	0.69
House Price (Thousands £)	166.56	115.96	173.95	127.73	174.64	128.53	175.70	129.82
% Achieving SCQF level 6	39.05	12.87	37.91	13.77	37.76	14.00	37.02	13.73
% Achieving SCQF level 5	66.30	11.50	64.40	12.10	64.27	12.33	63.80	12.30
% Achieving Literacy & Numeracy SCQF level 5	58.05	11.11	56.44	11.97	56.40	12.15	56.15	11.96
% of S6 Leavers	64.27	11.28	63.61	11.32	63.46	11.53	63.20	11.52
% of Leavers in Higher Education	36.38	11.28	35.97	12.23	35.76	12.47	35.16	12.29
% of Leavers in Further Education	24.30	6.62	24.51	6.65	24.52	6.58	24.76	6.46
% of Leavers Working	29.36	7.00	27.85	6.21	27.86	6.26	27.89	6.36
% of Leavers in Positive Destination	91.40	3.86	90.72	4.02	90.62	4.07	90.38	4.05
% of Pupils not on FSM	85.53	8.69	83.53	9.33	83.27	9.45	82.71	9.38
SCQF level 6 Value-Added	−0.68	5.25	−0.70	5.30	−0.68	5.31	−0.97	5.28
SCQF level 5 Value-Added	−0.23	5.03	−0.72	4.99	−0.63	5.05	−0.69	5.15
Literacy & Numeracy - Value-Added	−0.56	6.43	−0.71	6.41	−0.62	6.47	−0.65	6.36
Attendance Rate (In School)	88.70	2.70	88.09	2.84	87.98	2.87	87.82	2.85
Lateness Rate	2.70	1.41	3.01	1.52	3.09	1.52	3.17	1.51
Truancy Rate	2.02	1.52	2.20	1.63	2.23	1.65	2.27	1.66
Exclusion Rate	0.08	0.07	0.09	0.07	0.09	0.07	0.09	0.07
School Size	814.39	322.22	882.57	295.92	883.17	299.07	884.72	305.62
Panel B: Neighbourhood-Level Variables	Full Sample		350 metres		300 metres		250 metres	
	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.
Average No. of Rooms	5.02	0.88	4.79	0.89	4.80	0.90	4.79	0.90
Population Density	41.03	36.83	50.45	42.27	50.06	42.24	49.66	41.45
% on Social Renting	22.75	20.41	24.02	22.38	23.85	22.36	23.89	22.39
% Renting	11.83	10.75	13.78	12.77	13.82	12.81	13.85	12.80
% other than white	3.87	5.78	5.54	6.12	5.54	6.06	5.50	5.82
% no Qualification	26.82	11.76	25.93	12.74	25.86	12.80	25.89	12.82
% Higher Qualification	25.93	13.73	27.91	15.92	28.08	15.96	28.04	16.01
Median Age	41.33	6.75	39.74	7.05	39.76	7.09	39.72	7.09
% Female Alone	6.53	4.43	6.82	4.80	6.82	4.84	6.87	4.88
Crime Rate	310.81	446.26	410.12	646.06	416.79	667.52	422.76	687.96
Income Deprivation Rate	12.37	9.63	13.40	10.54	13.37	10.56	13.43	10.59
Overcrowding Rate	10.95	7.90	13.06	8.70	13.06	8.70	13.02	8.60
No. of DZs	6,623		2,332		2,131		1,950	
No. of Schools	275		188		179		170	
No. of Sales	221,073		52,471		45,417		38,090	

Note: House prices refer to transactions occurred between 2015 and 2018, whereas school-level information referes to academic years 2013/14, 2014/15 and 2015/16. Attainment data are elaborated by the Learning Directorate at the Scottish Government (SG) Analytical Services and available on the SG website.

Table A.2: Secondary School Qualifications

Grade	S4	S5	S6
SQA	National 5	Higher	Higher/Advanced Higher
SVQs	Modern Apprenticeship SVQ	Modern Apprenticeship SVQ	Modern Apprenticeship SVQ
SCQF	Level 5	Level 6	Level 7

Notes: Scottish Qualification Authority qualifications, Scottish Vocational Qualifications (SVQs) and their equivalent in terms of Scottish Credit and Qualifications Framework levels, by "Senior Phase" stage.

Table A.3: Robustness Checks - Composition

Dependent Variable: ln(House Prices)	(1) Baseline	(2) Extra Controls	(3) Data Zones × Years	(4) Data Zones × Boundary Distance	(5) RC Sample	(6) RC Sample
Socio-Economic Composition	0.038**	0.042**	0.020	0.039**	0.039*	0.038*
	(0.019)	(0.021)	(0.020)	(0.018)	(0.020)	(0.020)
PT ratio		0.014				
		(0.014)				
(Roll/Capacity)×100		−0.003				
		(0.002)				
Private School Index		0.044				
		(0.033)				
RC Higher						0.028
						(0.023)
Observations	13,368	13,368	13,368	13,368	12,390	12,390
No. of Schools	155	155	155	155	124	124
Mean	11.89	11.89	11.89	11.89	11.89	11.89
SD	0.71	0.71	0.71	0.71	0.72	0.72
Datazone FE	✓	✓	✓	✓	✓	✓
Adjusted R-squared	0.594	0.594	0.607	0.637	0.599	0.599

Notes: All of the above specifications refer to the sample of houses within those data zones which stretch across two catchment areas. Every specification includes data zone fixed effects, alongside indicators for year of transaction, urban area and school size. Data zone fixed effects make neighbourhood covariates redundant, hence these are not included. In columns (3) and (4) data zone dummies are interacted with year indicators and distance from boundary respectively. Columns (5) and (6) focus on sales which are assigned to a Roman Catholic school. In column (6) I control for the local Roman Catholic school's level of 'Higher'. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.4: Main Results - Academic Components

Dependent Variable: ln(House Prices)	(1) Full Sample	(2) 350 Metres	(3) Same DZ	(4) Same DZ	(5) 300 Metres	(6) 250 Metres
% SCQF level 6	0.034***	0.032***	0.043***	0.042***	0.040***	0.039***
	(0.008)	(0.007)	(0.014)	(0.014)	(0.014)	(0.014)
% SCQF level 5	0.024***	0.020***	0.024*	0.024*	0.026**	0.028**
	(0.007)	(0.006)	(0.014)	(0.014)	(0.013)	(0.013)
% Lit and Num	0.021***	0.023***	0.034**	0.034**	0.035**	0.036***
	(0.008)	(0.006)	(0.014)	(0.014)	(0.014)	(0.013)
% Higher Education	0.029***	0.025***	0.035**	0.036**	0.035**	0.034**
	(0.008)	(0.006)	(0.014)	(0.014)	(0.014)	(0.014)
% S6 Leavers	0.029***	0.019***	0.022*	0.022*	0.023*	0.026*
	(0.008)	(0.007)	(0.013)	(0.013)	(0.013)	(0.013)
Observations	221,073	52,471	13,368	13,368	12,659	11,668
No. of Schools	275	188	155	155	149	146
Mean	11.81	11.84	11.89	11.89	11.89	11.88
SD	0.68	0.69	0.71	0.71	0.71	0.72
Neighbourhood Controls	✓	✓				
Local Authority FE	✓					
Boundary FE		✓				
Data Zone FE			✓	✓	✓	✓
School Distance				✓	✓	✓
Boundary Distance Squared				✓	✓	✓

Notes: Each Column represents a specification in which $ln(HousePrices)$ is regressed on one indicator at the time. Indicators are standardised (mean = 0 and variance = 1). Each model contains a set of year of transaction dummies, school size and an indicator of geographic location (urban). Neighbourhood controls include the covariates from Figure 1.5. Data zone fixed effects make neighbourhood covariates redundant, hence these are not included from column (3) onward. Column (4) includes distance to catchment school (in 100s metres) and distance-from-boundary polynomials. Columns (5) and (6) pertain to sample of houses within 300 and 250 metres from catchent borders and within the same data zones. Adjusted R^2 range between .56 and .59. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.5: Main Results - Value-Added

Dependent Variable: ln(House Prices)	(1) Full Sample	(2) 350 Metres	(3) Same DZ	(4) Same DZ	(5) 300 Metres	(6) 250 Metres
% SCQF level 6 - Value-Added	0.055***	0.036***	0.021*	0.021*	0.018	0.016
	(0.016)	(0.009)	(0.011)	(0.011)	(0.011)	(0.011)
% SCQF level 5 - Value-Added	0.009	0.013	−0.013	−0.013	−0.012	−0.012
	(0.019)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)
% Lit and Num - Value-Added	−0.012	0.006	0.008	0.008	0.008	0.009
	(0.016)	(0.008)	(0.011)	(0.011)	(0.011)	(0.011)
Observations	221,073	52,471	13,368	13,368	12,659	11,668
No. of Schools	275	188	155	155	149	146
Mean	11.81	11.84	11.89	11.89	11.89	11.88
SD	0.68	0.69	0.71	0.71	0.71	0.72
Neighbourhood Controls	✓	✓				
Local Authority FE	✓					
Boundary FE		✓				
Data Zone FE			✓	✓	✓	✓
School Distance				✓	✓	✓
Boundary Distance Squared				✓	✓	✓

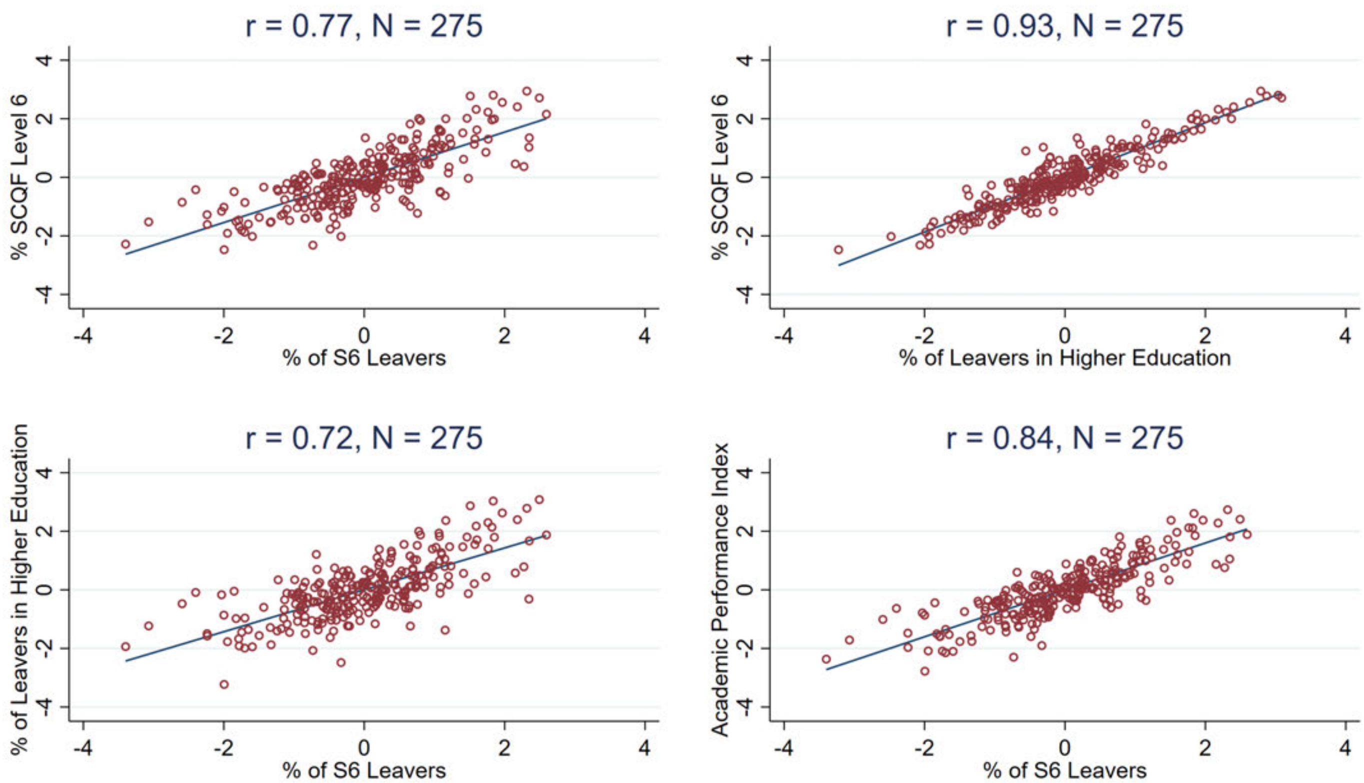
Notes: Each Column represents a specification in which $ln(HousePrices)$ is regressed on one indicator at the time. Indicators are standardised (mean = 0 and variance = 1). Each model contains a set of year of transaction dummies, school size and an indicator of geographic location (urban). Neighbourhood controls include the covariates from Figure 1.5. Data zone fixed effects make neighbourhood covariates redundant, hence these are not included from column (3) onward. Column (4) includes distance to catchment school (in 100s metres) and distance-from-boundary polynomials. Columns (5) and (6) pertain to sample of houses within 300 and 250 metres from catchent borders and within the same data zones. Adjusted R^2 range between .56 and .59. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.6: Main Results - Non-Cognitive Skills

Dependent Variable: ln(House Prices)	(1) Full Sample	(2) 350 Metres	(3) Same DZ	(4) Same DZ	(5) 300 Metres	(6) 250 Metres
Attendance	0.141***	0.070***	0.027**	0.027**	0.029**	0.029**
	(0.017)	(0.012)	(0.013)	(0.012)	(0.012)	(0.013)
Lateness	−0.054***	−0.042***	−0.013	−0.014	−0.015	−0.012
	(0.016)	(0.011)	(0.015)	(0.015)	(0.014)	(0.014)
Truancy	−0.067***	−0.012	−0.016	−0.017	−0.019	−0.018
	(0.016)	(0.012)	(0.013)	(0.013)	(0.013)	(0.014)
Exclusions	−0.071***	−0.017*	−0.004	−0.004	−0.005	−0.007
	(0.019)	(0.010)	(0.012)	(0.012)	(0.013)	(0.012)
Observations	221,073	52,471	13,368	13,368	12,659	11,668
No. of Schools	275	188	155	155	149	146
Mean	11.81	11.84	11.89	11.89	11.89	11.88
SD	0.68	0.69	0.71	0.71	0.71	0.72
Neighbourhood Controls	✓	✓				
Local Authority FE	✓					
Boundary FE		✓				
Data Zone FE			✓	✓	✓	✓
School Distance				✓	✓	✓
Boundary Distance Squared				✓	✓	✓

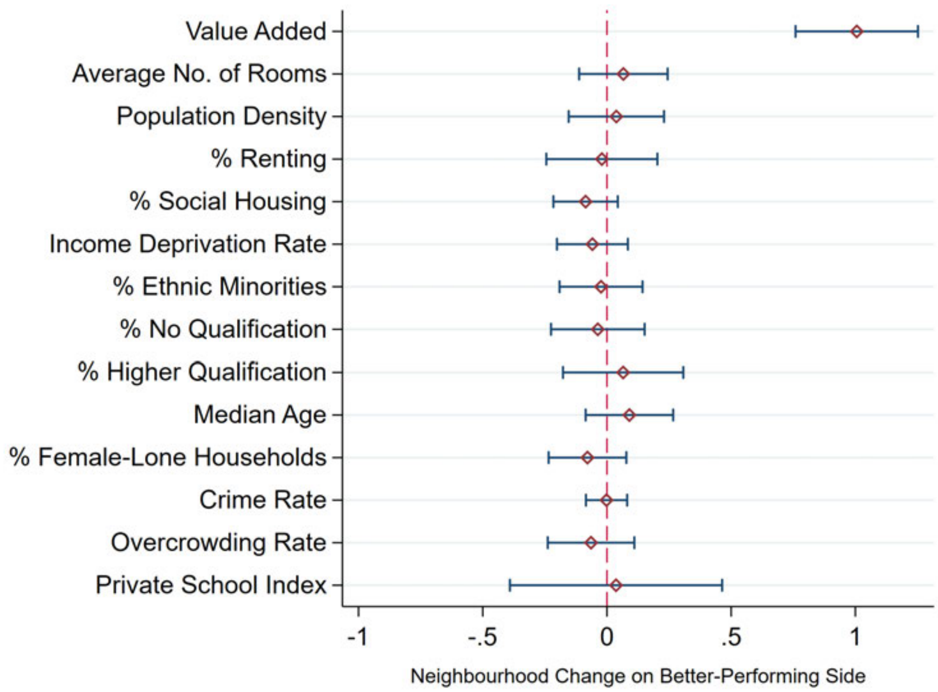
Notes: Each Column represents a specification in which $ln(HousePrices)$ is regressed on one indicator at the time. Indicators are standardised (mean = 0 and variance = 1). Each model contains a set of year of transaction dummies, school size and an indicator of geographic location (urban). Neighbourhood controls include the covariates from Figure 1.5. Data zone fixed effects make neighbourhood covariates redundant, hence these are not included from column (3) onward. Column (4) includes distance to catchment school (in 100s metres) and distance-from-boundary polynomials. Columns (5) and (6) pertain to sample of houses within 300 and 250 metres from catchent borders and within the same data zones. Adjusted R^2 range between .56 and .59. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

Figure A.1: SCQF Level 6 and other Academic Performance Indicators.

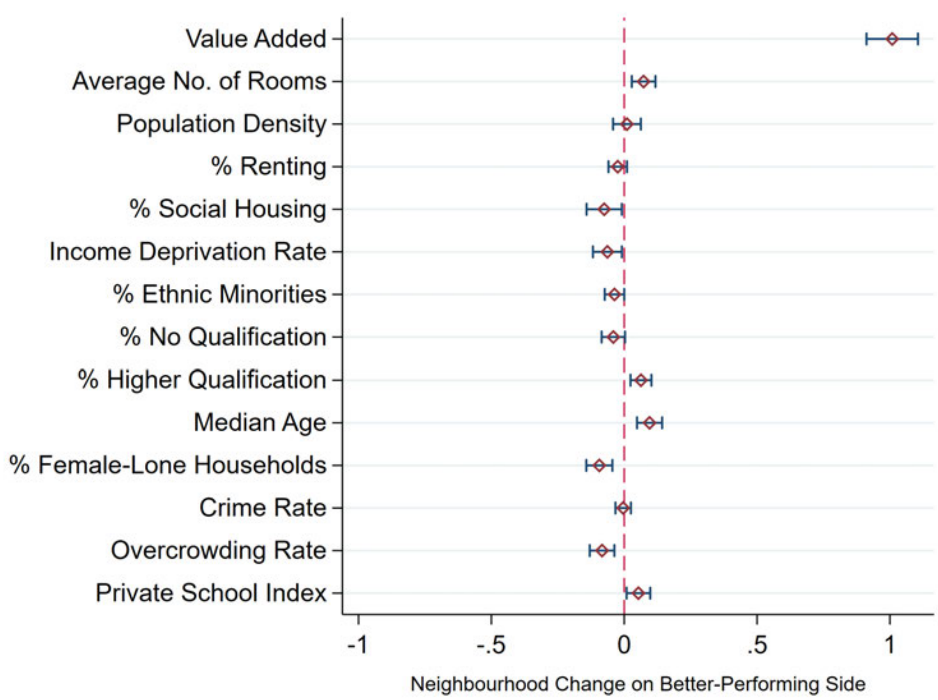


Note: In this set of scatter plots is reported the correlation between the % of Pupils achieving 4 or more awards at SCQF level 6 (or better) and other, more obvious, indicators of academic performance at the school level. In particular, each dot corresponds to a school whose value has been standardised to have mean zero and unit-variance and averaged across all school years (2013/14, 2014/15 and 2015/16).

Figure A.2: Balancing Test - Value-Added



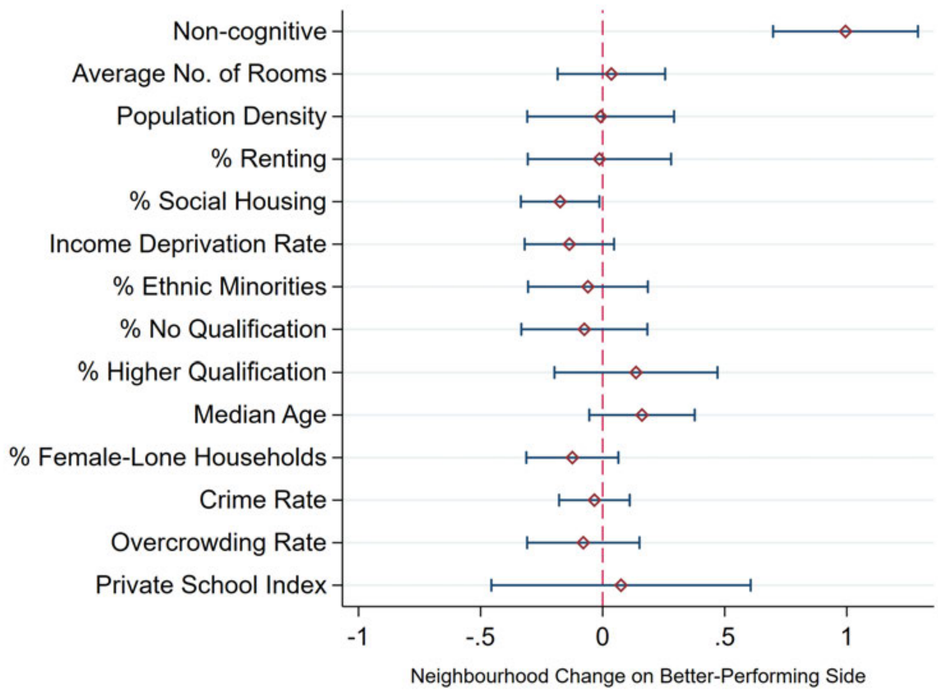
(a) w/o Boundary Fixed Effects



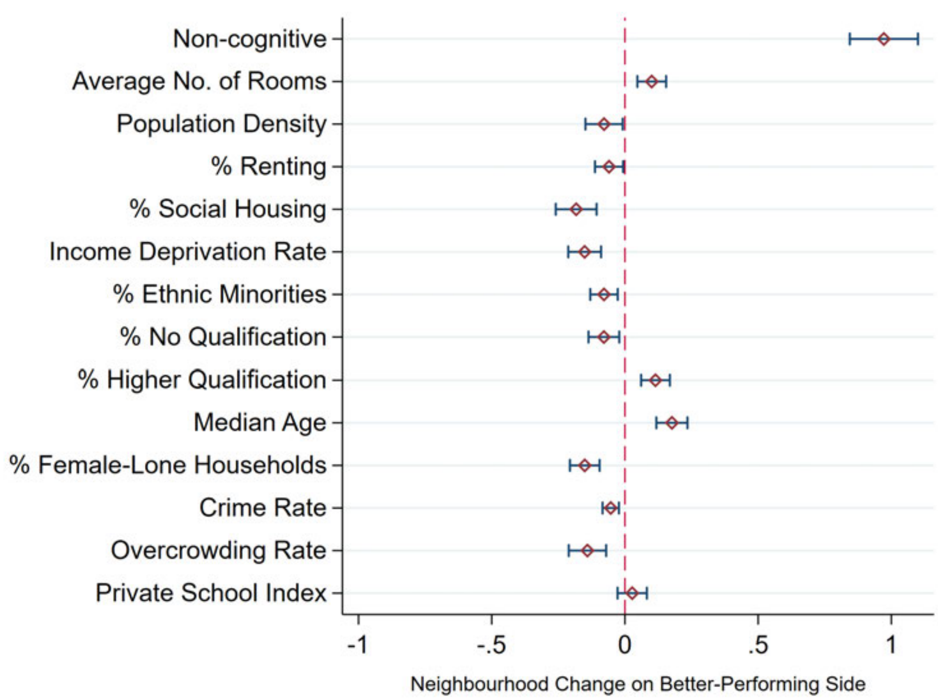
(b) w Boundary Fixed Effects

Note: Each coefficient results from a regression of the variable at hand on a binary indicator of whether the property is located on the better-performing side of the catchment area boundary, based on the *Value-Added* performance indicator. The whiskers indicate 95% level confidence intervals.

Figure A.3: Balancing Test - Non-cognitive Skills



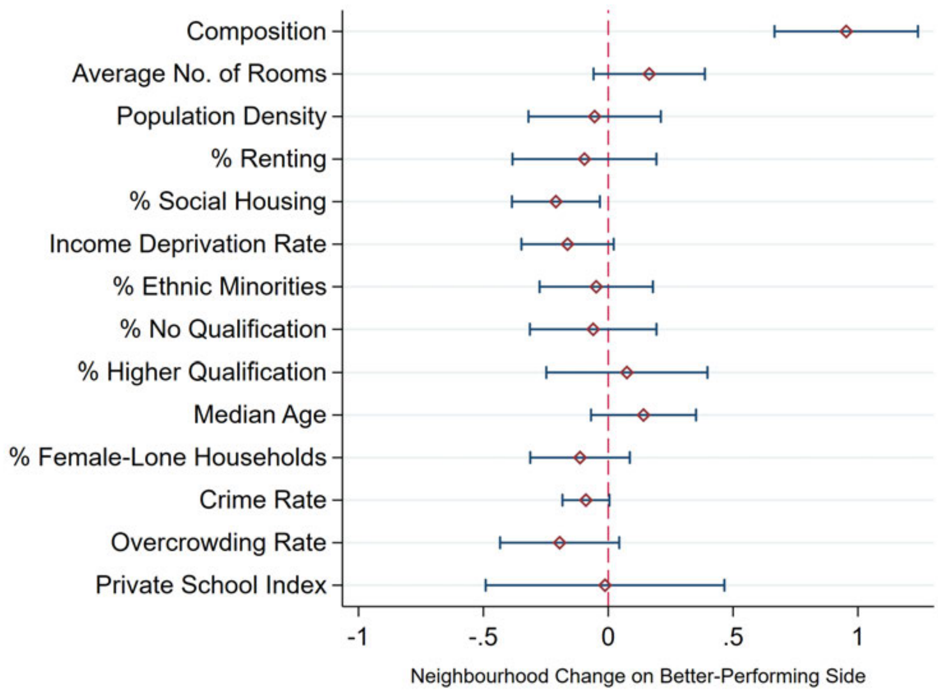
(a) w/o Boundary Fixed Effects



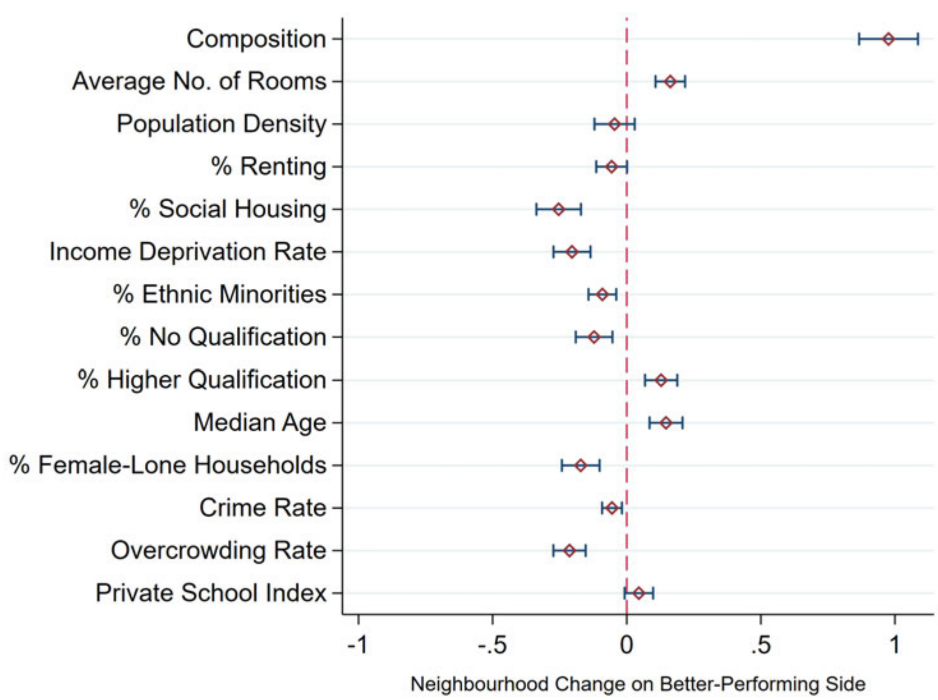
(b) w Boundary Fixed Effects

Note: Each coefficient results from a regression of the variable at hand on a binary indicator of whether the property is located on the better-performing side of the catchment area boundary, based on the *Non-cognitive Skills* performance indicator. The whiskers indicate 95% level confidence intervals.

Figure A.4: Balancing Test - Socio-Economic Composition



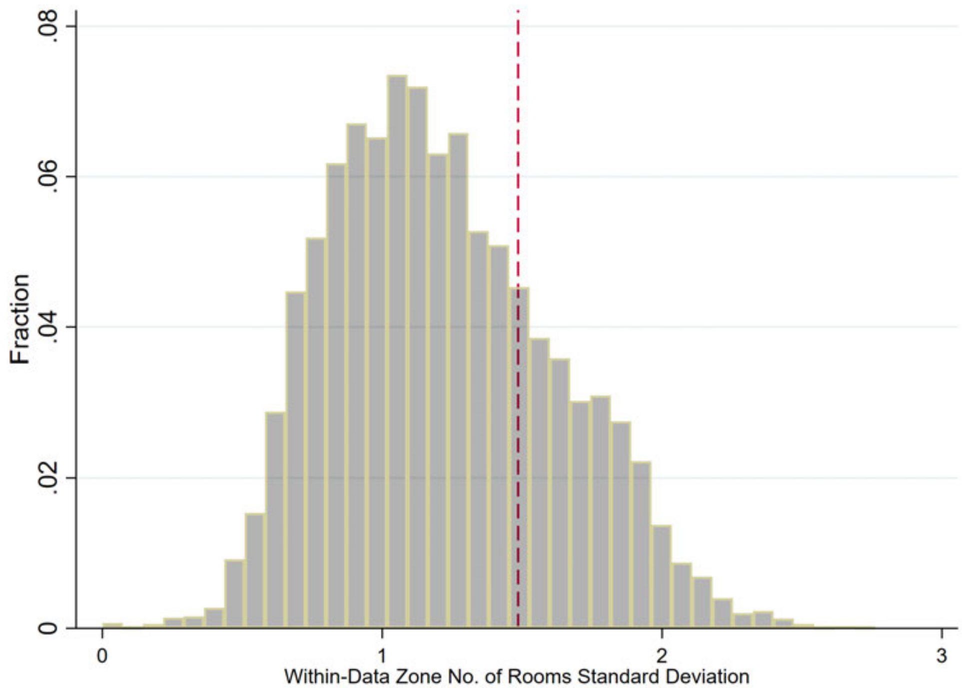
(a) w/o Boundary Fixed Effects



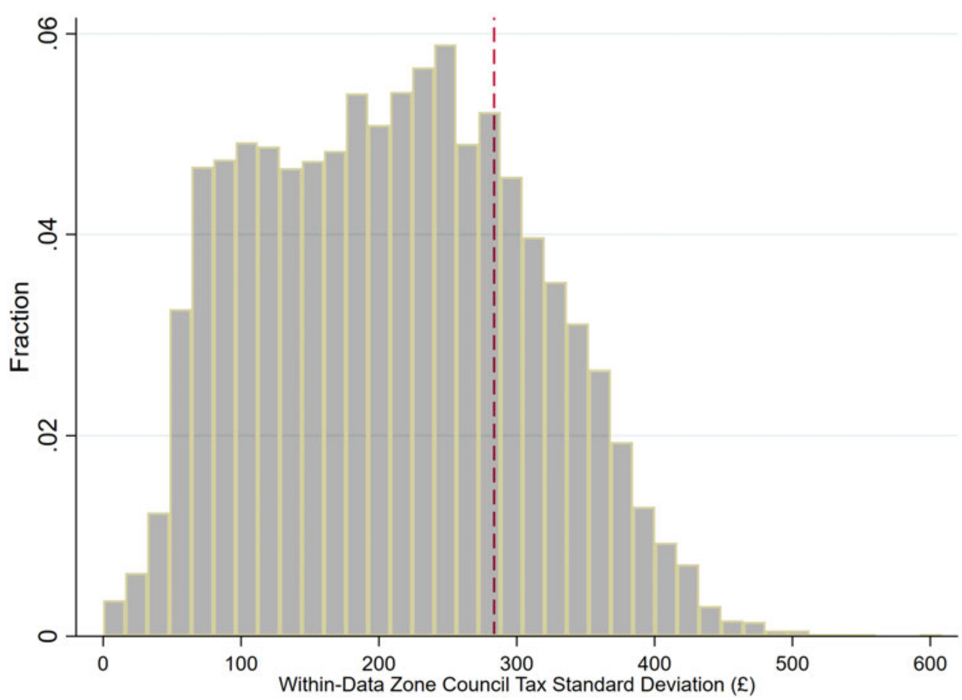
(b) w Boundary Fixed Effects

Note: Each coefficient results from a regression of the variable at hand on a binary indicator of whether the property is located on the better-performing side of the catchment area boundary, based on the *Composition* performance indicator. The whiskers indicate 95% level confidence intervals.

Figure A.5: Data Zones Homogeneity



(a) Number of Rooms



(b) Council Tax Value

Note: This figure shows the distribution of the within-data zone standard deviation in: *i)* number of rooms; *ii)* council tax value.

Table A.7: Variables Description

	Time	Description
% SCQF level 6	2014, 2015, 2016	% leavers with four or more awards at SCQF level 6
% SCQF level 5	2014, 2015, 2016	% leavers with four or more awards at SCQF level 5 (or above)
% Achieving Literacy & Numeracy SCQF level 5	2014, 2015, 2016	Level 5 or above
% of S6 leavers	2014, 2015, 2016	see variable name
% of Leavers in Higher Education	2014, 2015, 2016	see variable name
% of Leavers in Further Education	2014, 2015, 2016	see variable name
% of Leavers Working	2014, 2015, 2016	see variable name
% of Leavers in Positive Destination	2014, 2015, 2016	see variable name
% of Pupils not on FSM	2014, 2015, 2016	see variable name
SCQF level 6 - Value-Added	2014, 2015, 2016	SCQF level 6 - SCQF level 6 Virtual Comparator
SCQF level 5 - Value-Added	2014, 2015, 2016	SCQF level 5 - SCQF level 5 Virtual Comparator
Literacy & Numeracy - Value-Added	2014, 2015, 2016	Lit&Num - Lit&Num Virtual Comparator
Attendance Rate (In School)	2013, 2015	$(\frac{Half-days\ In\ School\ Attendance}{No.Half-day\ Openings})\times 100$
Lateness Rate	2013, 2015	$(\frac{Half-days\ Lateness\ Episodes}{No.Half-day\ Openings})\times 100$
Truancy Rate	2013, 2015	$(\frac{Half-days\ Truancy\ Episodes}{No.\ Half-day\ Openings})\times 100$
Exclusion Rate	2013, 2015	$(\frac{Half-days\ missed\ because\ of\ exclusion}{No.\ Half-day\ Openings})\times 100$
School Size	2014, 2015, 2016	No. of pupils in school

Notes: Data on Attendance, Absence and Exclusions are collected every two years and as such they are the average of the 2013 and 2015 values. All the other variables are used individually for school years 2013/14, 2014/15 and 2015/16.

Appendix B

Appendix - Chapter 2

Table B.1: Reduced Form Results

	Numeracy			Literacy		
	(1) P1	(2) P4	(3) P7	(4) P1	(5) P4	(6) P7
$CompLow^{pred}_{gst}$	0.008**	0.012***		0.014***	0.008*	
	(0.003)	(0.004)		(0.004)	(0.004)	
$CompHigh^{pred}_{gst}$		−0.004	−0.008**		−0.003	−0.001
		(0.004)	(0.004)		(0.004)	(0.004)
Observations	190,704	194,804	186,082	190,704	194,804	186,082
No. of Schools	1,437	1,428	1,435	1,437	1,428	1,435
School FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: $***/**/*$ indicate significance at the 1%/5%/10%-level. Heteroscedasticity-robust standard errors adjusted for clustering at the school and year level are reported in parentheses.

This table shows the reduced form results, i.e. the results of an Ordinary Least Squares (OLS) regression in which we regress our outcomes of interest - which are proficiency in numeracy and literacy, respectively - on our instruments - which are class planner predictions of whether a pupil's grade should contribute to a composite class.

Covariates include pupil age, sex, and ethnicity, an indicator for whether pupil is from a neighborhood in bottom 20% of deprivation (SIMD), class size, the size of the school, and the percentage of pupils in a school that are female, white British, native English speakers, and in the bottom 20% of deprivation respectively. All specifications contain a set of school and school-year fixed effects.

Table B.2: Second Stage Results - Fourth (P4) and Seventh (P7) Graders

Panel A: Second Stage Results for P4

	Numeracy				Literacy			
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS
Older Peers	0.002^{***}	0.044^{***}			0.002^{***}	0.024		
	(0.001)	(0.017)			(0.001)	(0.017)		
Younger Peers	-0.004^{***}	-0.009			-0.004^{***}	-0.008		
	(0.001)	(0.015)			(0.001)	(0.015)		
Bottom Comp.			0.033^{***}	0.509^{**}			0.027^{***}	0.291
			(0.006)	(0.215)			(0.007)	(0.206)
Top Composite			-0.032^{***}	-0.250			-0.037^{***}	-0.180
			(0.007)	(0.257)			(0.007)	(0.247)
Class Size	0.004^{***}	0.004	0.004^{***}	0.001	0.004^{***}	0.002	0.004^{***}	0.001
	(0.001)	(0.003)	(0.001)	(0.004)	(0.001)	(0.003)	(0.001)	(0.004)
Observations	194,804	194,803	194,804	194,803	194,804	194,803	194,804	194,803
No. of Schools	1428	1,428	1,428	1,428	1,428	1,428	1,428	1,428
Class-Size Instr.	No	Yes	No	Yes	No	Yes	No	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Stat		12.82		4.918		12.82		4.918

Panel B: Second Stage Results for P7

	Numeracy				Literacy			
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS
Younger Peers	-0.005^{***}	-0.051			-0.005^{***}	0.008		
	(0.001)	(0.035)			(0.001)	(0.031)		
Top Composite			-0.054^{***}	-0.704			-0.051^{***}	0.114
			(0.007)	(0.527)			(0.007)	(0.425)
Class Size	0.004^{***}	-0.017	0.004^{***}	-0.021	0.004^{***}	0.005	0.004^{***}	0.005
	(0.001)	(0.012)	(0.001)	(0.017)	(0.001)	(0.011)	(0.001)	(0.014)
Observations	186,082	186,078	186,082	186,078	186,082	186,078	186,082	186,078
No. of Schools	1,435	1,435	1,435	1,435	1,435	1,435	1,435	1,435
Class-Size Instr.	No	Yes	No	Yes	No	Yes	No	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Stat		17.99		12.91		17.99		12.91

Notes: $***/**/*$ indicate significance at the 1%/5%/10%-level. Heteroscedasticity-robust standard errors adjusted for clustering at the school and year level are reported in parentheses.

This table shows the results for our estimation of equation (2.1) by Ordinary Least Squares (OLS) and 2-Stage-Least-Squares (2SLS) regression. Our outcomes of interest are dummy indicators for whether a pupil performs at least at the expected level in numeracy or literacy, respectively. Results in Panel A refer to our sample of fourth graders (P4), results in Panel B refer to seventh graders (P7).

Covariates include pupil age, sex, and ethnicity, an indicator for whether pupil is from a neighborhood in bottom 20% of deprivation (SIMD), grade enrolment counts and its square, the size of the school, and the percentage of pupils in a school that are female, white British, native English speakers, and in the bottom 20% of deprivation respectively. All specifications contain a set of school and school-year fixed effects.

The reported first-stage F-statistic is heteroscedasticity and autocorrelation consistent (HAC) and was calculated using the method developed by Kleibergen and Paap (2006).

Table B.3: Composites and Number of Teachers per Classroom

	First Graders (P1)		Fourth Graders (P4)		Seventh Graders (P7)	
	(1)	(2)	(3)	(4)	(5)	(6)
	>1 Teacher Present		>1 Teacher Present		>1 Teacher Present	
Composite Class (Binary)	-0.008^{**}	-0.008^{**}	0.004^{**}	0.004^{**}	0.008^{***}	0.008^{***}
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
Female		0.000		-0.001^{**}		0.000
		(0.001)		(0.000)		(0.000)
White		0.001		0.001		0.000
		(0.001)		(0.001)		(0.001)
Native English Speaker		0.000		−0.001		−0.001
		(0.001)		(0.001)		(0.001)
Bottom 20% SIMD		−0.001		0.000		−0.000
		(0.001)		(0.000)		(0.000)
Age (in Years)		0.000		0.000		0.001
		(0.002)		(0.001)		(0.000)
% Female in School		0.096		−0.005		0.043
		(0.109)		(0.047)		(0.039)
% White British		0.163		-0.074^{*}		−0.080
		(0.115)		(0.044)		(0.054)
% Native English Speakers		−0.073		0.098^{**}		−0.026
		(0.138)		(0.047)		(0.057)
% in Bottom 20% SIMD		−0.146		-0.081^{**}		0.073
		(0.108)		(0.040)		(0.056)
Number of Students in School		0.000^{*}		0.000^{***}		0.000
		(0.000)		(0.000)		(0.000)
Observations	190,040	190,040	193,627	193,627	185,416	185,416
R-squared	0.424	0.424	0.274	0.276	0.228	0.229
School FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: $***/**/*$ indicate significance at the 1%/5%/10%-level. Heteroscedasticity-robust standard errors adjusted for clustering at the school and year level are reported in parentheses.
This table shows the results of an OLS regression where the dependent variable a dummy indicator for whether there is more than 1 teacher present in the classroom.

Estimating the Effect of Class Size

Throughout this paper, we control for the effect of class size. By virtue of a lower cap, multi-grade classes tend to be smaller than single-year classes (see Section 2.2 for maximum class size rules). As a result, they may affect achievement not just through peer effects but also due to a lower pupil to teacher ratio. In order to disentangle these two competing mechanisms, we have included a control variable for class size in all specifications.

While not the primary focus of this paper, the effect of class size is also interesting in itself. However, class size may well be endogenously determined even after controlling for school fixed effects. We therefore also construct an instrument for class size based on enrolment counts in the mould of Angrist and Lavy's (1999) seminal study. Our approach exploits that at an enrolment level just above a maximum class size cut-off, a new class needs to be created.

The key condition in such a regression discontinuity design is that enrolment counts are as good as randomly determined, for instance, by natural fluctuations in birth rates within catchment areas. Bunching in enrolment count on the other hand indicates a violation of the main identifying assumption. Figure B.1a shows the enrolment counts for P1 for the school years 2011/12 to 2018/19 when the maximum class size for this grade was 25. We can see obvious sorting at multiples of 25. For instance, there are almost twice as many schools with enrolment counts of exactly 50 than with 51. We can also see bunching in P4 (Figure B.1c) and P7 (Figure B.1e) at multiples of 33.

Angrist et al. (2019) find similar patterns for their data from Israeli schools. In their case, financial incentives lead to enrolment count manipulation. Israeli schools receive further funding for every additional class that needs to be created. School head teachers selectively use deferment and retention or class skipping, to create enrolment counts that are just large enough to trigger additional classes. In Scotland, the incentives line up exactly in reverse. Scottish head teachers have virtually no discretion over their enrolment counts. Grade retention is also almost unheard of. The sorting that is apparent in Figures B.1a, B.1c, and B.1e is instead driven by strategic acceptance of placing requests by Councils. As mentioned in the main text (see Section 2.2), parents can request their children go to schools outside their catchment area, but councils will only grant such placing requests if the requested school has space available. In practice, councils will often accept placing requests for oversubscribed schools up to the point at which the enrolment count is equal to a multiple of the class size limit. Because funding to schools is on a per-class basis (rather than a per-pupil basis), this reshuffling and "filling-up" approach helps councils to cut costs.

Angrist et al. (2019) remedy the bunching issue by calculating an imputed enrolment count that assigns each pupil to the grade in which they should be in, had birthday cut-offs been strictly adhered to. We follow this approach in spirit and assign each pupil to the school they should attend based on the catchment area they reside in. In order to do so, we exploit information on each pupil's postcode area.In other words, we calculate each school's (imputed) enrolment count as if placing requests were not an option. This creates two issues. First, a small set of pupils who are in a school might be there by virtue of a placing request, but we cannot identify them as such because only part of their postcode area overlaps with the catchment area. Second, if we identify a pupil who is attending a school by virtue of a placing request but whose postcode area stretches over multiple catchment areas, it is not obvious against which school's imputed enrolment count such a student should count.

We address both issues by calculating postcode area frequency distributions for all schools in all years, as well as school frequency distributions for each postcode area in all years. If a pupil's postcode area makes up less than 5 percent of her school's pupil population, we re-assign the pupil to her catchment area school[1]. In other words, we assume that students from infrequent postcode areas are in a school due to placing requests. If that same pupil's postcode feeds into two schools, we assign her to the first school with a probability equal to the percentage of pupils from the same postcode area that attend this first school; and to the second school with a probability equal to the percentage of pupils from the same postcode area that attend this second school.[2]

Figures B.1b, B.1d, and B.1f show our imputed enrolment counts for P1, P4, and P7 respectively. We can see that our imputed enrolment counts no longer suffer from bunching. All three distributions are smooth and the heaping at multiples of maximum class sizes has disappeared. Appendix Figures B.2a to B.2f show the corresponding density plots that accompany McCrary's (2008) formal test for sorting. We firmly reject the null hypothesis of no discontinuity for the original enrolment counts but fail to do so for imputed enrolment counts.

We therefore use the imputed grade enrolment counts rather than the actual enrolment counts to predict the class sizes for class c in grade g in school s and year t as:

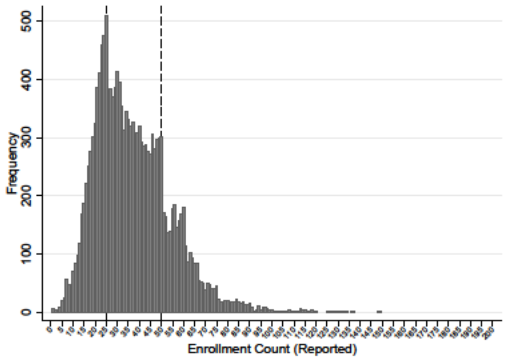
$$f_{cgst} = \frac{r_{gst}^{imp}}{int(\frac{r_{gst}^{imp}-1}{cutoff_{gt}}) + 1}$$

Where r_{gst}^{imp} is the imputed enrolment in school s' g^{th} grade as of September in year t. In both first-stage and second-stage regressions we also flexibly control for r_{gst}^{imp}. Lastly, cutoff$_{gt}$ represents the class size limit, which varies by grade g and school-year t Ultimately, we use the predicted class size f_{cgst} as an instrument for actual class size, $\widehat{CS_{cgst}}$.

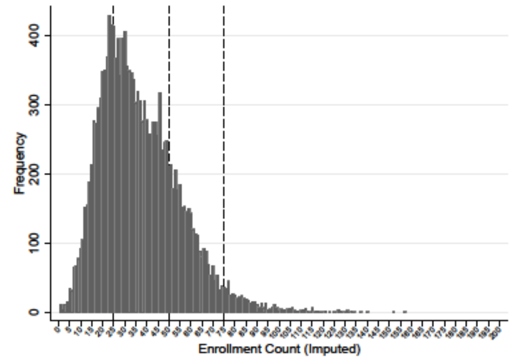
[1]We also experimented with slightly lower and higher thresholds, the results are qualitatively identical.

[2]Pupils on placing requests are excluded from these probability calculations such that the probabilities add up to 1.

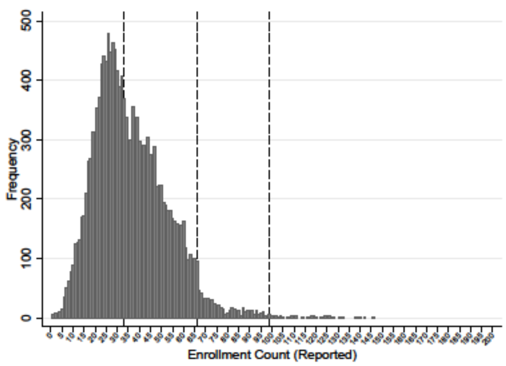
Figure B.1: Enrollment Distributions 2007-2018



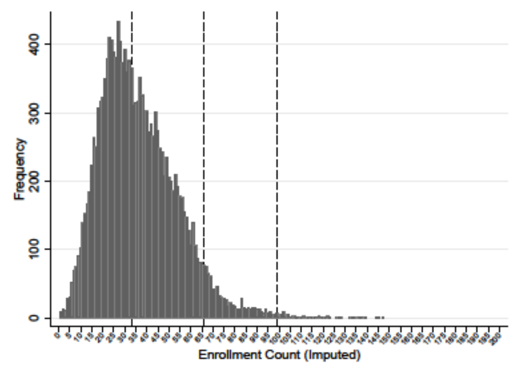
(a) P1 - Reported



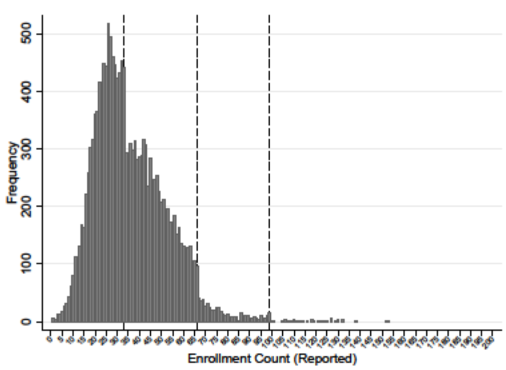
(b) P1 - Imputed



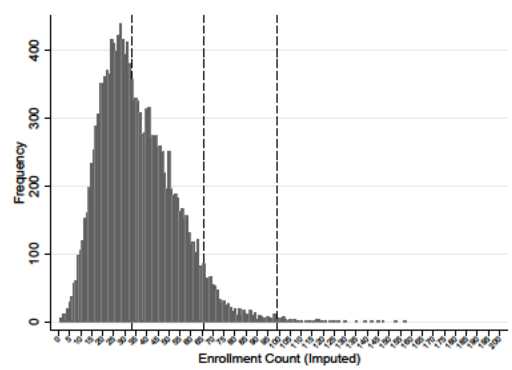
(c) P4 - Reported



(d) P4 - Imputed



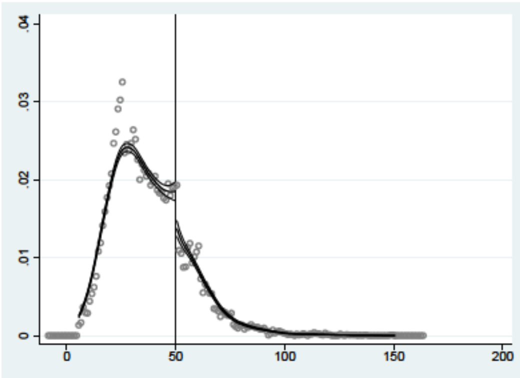
(e) P7 - Reported



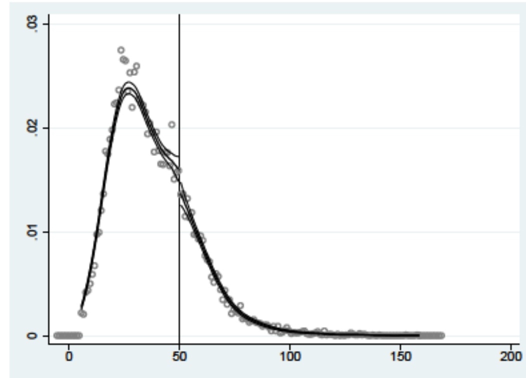
(f) P7 - Imputed

Notes: These figures show the enrolment counts for all schools in our data from 2007/08 to 2018/19, separately for first grade (P1), fourth grade (P4), and seventh grade (P7). On the left are the original enrolment counts which show bunching at multiples of the corresponding maximum class size. On the right are the corresponding imputed enrolment counts in which pupils who we believe are in a school due to placing requests were re-allocated to their catchment area school.

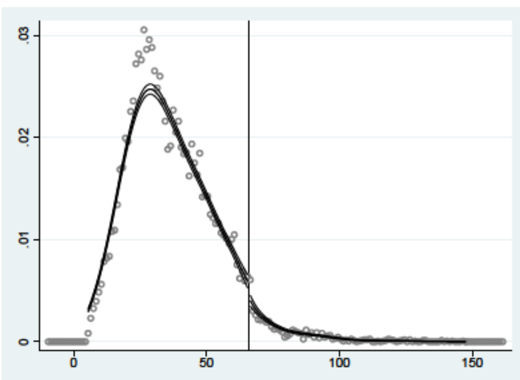
Figure B.2: Density Tests - Illustrations



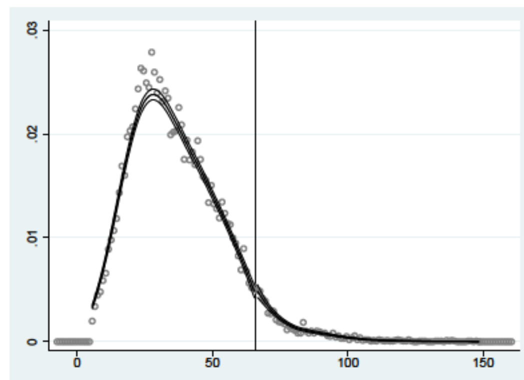
(a) P1 - Reported



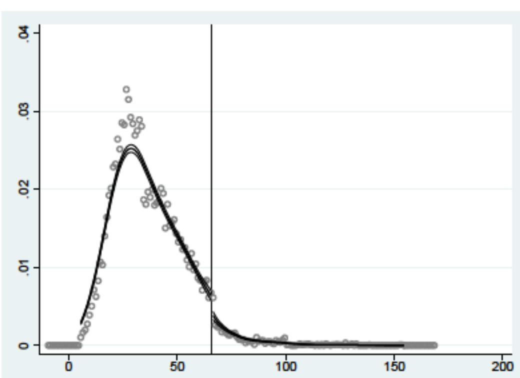
(b) P1 - Imputed



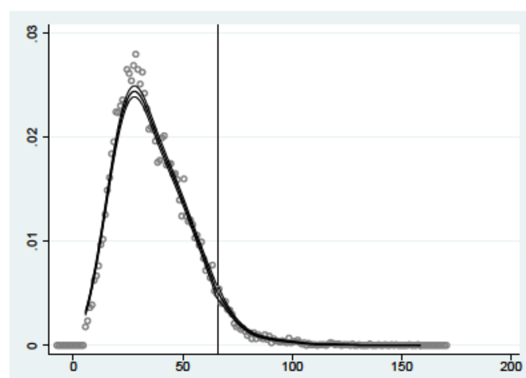
(c) P4 - Reported



(d) P4 - Imputed



(e) P7 - Reported



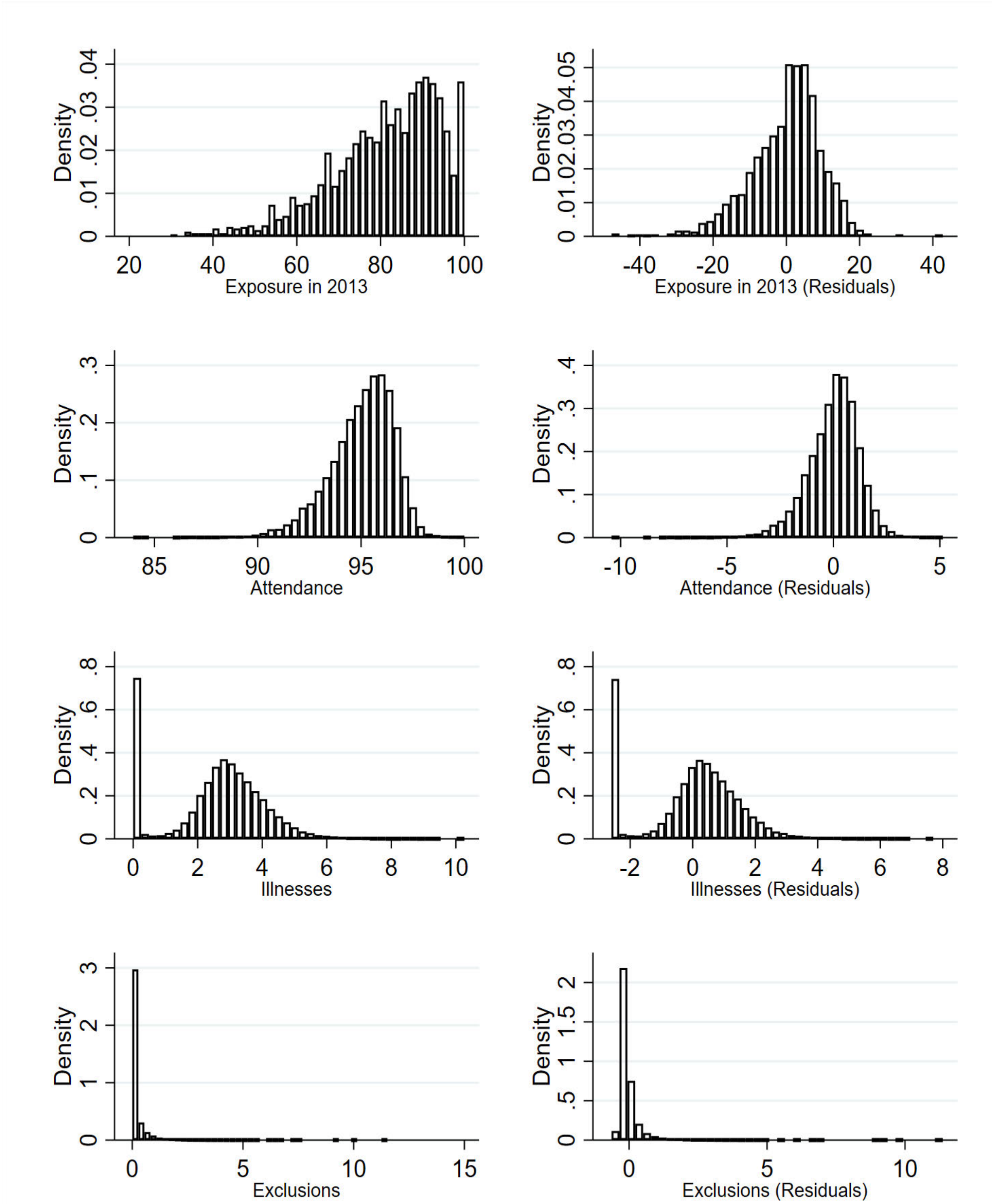
(f) P7 - Imputed

Notes: These figures show the density in cohort size distributions. Vertical lines are placed at the second multiple of the respective maximum class size thresholds.

Appendix C

Appendix - Chapter 3

Figure C.1: Treatment and Outcomes Variation



Notes: Histograms on the left-hand side refer to the raw variables, whereas those on the right-hand side are the residuals after regressing the variables on the Scottish Index of Multiple Deprivation Score in 2004.

Table C.1: Main Results - Unbalanced Panel

Panel A: Attendance	(1)	(2)	(3)	(4)
Post × Exposure	0.009***	0.010***	0.001	0.009***
	(0.003)	(0.003)	(0.003)	(0.003)
Observations	18,346	17,995	17,995	17,995
No. of Schools	1714	1714	1714	1714
Mean Dep. Var.	95.03	95.03	95.03	95.03
SD Dep. Var.	1.58	1.58	1.58	1.58
R-squared	0.764	0.76	0.712	0.765
Panel B: Illnesses	(1)	(2)	(3)	(4)
Post × Exposure	-0.009***	-0.011***	-0.003	-0.009***
	(0.003)	(0.003)	(0.003)	(0.002)
Observations	18,365	18,010	18,010	18,010
No. of Schools	1714	1714	1714	1714
Mean Dep. Var.	2.58	2.58	2.58	2.58
SD Dep. Var.	1.54	1.54	1.54	1.54
R-squared	0.584	0.585	0.44	0.666
Panel C: Exclusions	(1)	(2)	(3)	(4)
Post × Exposure	0.001	0.000	0.000	0.001
	(0.001)	(0.001)	(0.002)	(0.001)
Observations	17,872	17,537	17,537	17,537
No. of Schools	1714	1714	1714	1714
Mean Dep. Var.	0.19	0.19	0.19	0.19
SD Dep. Var.	0.49	0.49	0.49	0.49
R-squared	0.323	0.304	0.200	0.308
Baseline Controls		✓	✓	✓
School FE	✓	✓	✓	✓
School Trends			✓	
Local Authories Trends				✓

Notes: Coefficients are obtained by estimating γ from Equation 3.3. Outcomes are calculated in % of all possible half-day openings. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) score and school average class size in 2004, both interacted with year dummies, alongside indicators for urban and religious all appropriately interacted with year dummies. Data span from school year 2003/2004 through to 2016/2017. Outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. This sample does not include schools which extended eligibility following local initiatives from 2010. This sample also includes schools which are not observed every year. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table C.2: Main Results (log)

	(1)	(2)	(3)	(4)
Panel A: ln(Attendance)				
Post × Exposure	0.000***	0.000***	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	17,766	17,766	17,766	17,766
No. of Schools	1630	1630	1630	1630
Mean Dep. Var.	4.55	4.55	4.55	4.55
SD Dep. Var.	0.02	0.02	0.02	0.02
R-squared	0.761	0.762	0.714	0.766
Panel B: ln(1 + Illnesses)	(1)	(2)	(3)	(4)
Post × Exposure	-0.003**	-0.003***	-0.002*	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	17,781	17,781	17,781	17,781
No. of Schools	1630	1630	1630	1630
Mean Dep. Var.	1.14	1.14	1.14	1.14
SD Dep. Var.	0.58	0.58	0.58	0.58
R-squared	0.597	0.600	0.421	0.701
Panel C: ln(1 + Exclusions)	(1)	(2)	(3)	(4)
Post × Exposure	0.001	0.000	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	17,308	17,308	17,308	17,308
No. of Schools	1630	1630	1630	1630
Mean Dep. Var.	0.13	0.13	0.13	0.13
SD Dep. Var.	0.25	0.25	0.25	0.25
R-squared	0.386	0.387	0.289	0.389
Baseline Controls		✓	✓	✓
School FE	✓	✓	✓	✓
School Trends			✓	
Local Authories Trends				✓

Notes: Coefficients are obtained by estimating γ from Equation 3.3. Outcomes are calculated in % of all possible half-day openings. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) score and school average class size in 2004, both interacted with year dummies, alongside indicators for urban and religious all appropriately interacted with year dummies. Data span from school year 2003/2004 through to 2016/2017. Outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table C.3: Different Trends Specifications

	(1) Baseline Model	(2) Year FE ×	(3) Linear Trend ×	(4) Post ×
Panel A: Attendance				
Post × Exposure	0.010***	0.011***	0.006***	0.011***
	(0.003)	(0.003)	(0.002)	(0.003)
Observations	17,766	17,766	17,766	17,766
No. of Schools	1630	1630	1630	1630
Mean Dep. Var.	95.05	95.05	95.05	95.05
SD Dep. Var.	1.57	1.57	1.57	1.57
R-squared	0.757	0.762	0.759	0.758
Panel B: Illnesses	(1) Baseline Model	(2) Year FE ×	(3) Linear Trend ×	(4) Post ×
Post × Exposure	-0.009***	-0.010***	-0.013***	-0.010***
	(0.003)	(0.003)	(0.002)	(0.003)
Observations	17,781	17,781	17,781	17,781
No. of Schools	1630	1630	1630	1630
Mean Dep. Var.	2.57	2.57	2.57	2.57
SD Dep. Var.	1.54	1.54	1.54	1.54
R-squared	0.585	0.587	0.586	0.584
Panel C: Exclusions	(1) Baseline Model	(2) Year FE ×	(3) Linear Trend ×	(4) Post ×
Post × Exposure	0.001	0.001	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	17,308	17,308	17,308	17,308
No. of Schools	1630	1630	1630	1630
Mean Dep. Var.	0.19	0.19	0.19	0.19
SD Dep. Var.	0.48	0.48	0.48	0.48
R-squared	0.300	0.303	0.300	0.297
Baseline Controls		✓	✓	✓
School FE	✓	✓	✓	✓
School Trends				
Local Authories Trends				

Notes: Coefficients are obtained by estimating γ from Equation 3.3. Outcomes are calculated in % of all possible half-day openings. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) score and school average class size in 2004, both interacted with year dummies, alongside indicators for urban and religious all appropriately interacted with year dummies. Data span from school year 2003/2004 through to 2016/2017. Outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table C.4: Main Results P1-P3 Weighted

Panel A: Attendance	(1)	(2)	(3)	(4)
Post × Exposure	0.008***	0.009***	0.002	0.008***
	(0.003)	(0.003)	(0.003)	(0.003)
Observations	15,662	15,662	15,662	15,662
No. of Schools	1630	1630	1630	1630
Mean Dep. Var.	95.05	95.05	95.05	95.05
SD Dep. Var.	1.57	1.57	1.57	1.57
R-squared	0.796	0.797	0.752	0.801
Panel B: Illnesses	(1)	(2)	(3)	(4)
Post × Exposure	-0.011***	-0.013***	-0.005	-0.009***
	(0.004)	(0.004)	(0.004)	(0.003)
Observations	15,681	15,681	15,681	15,681
No. of Schools	1630	1630	1630	1630
Mean Dep. Var.	2.57	2.57	2.57	2.57
SD Dep. Var.	1.54	1.54	1.54	1.54
R-squared	0.604	0.606	0.456	0.687
Panel C: Exclusions	(1)	(2)	(3)	(4)
Post × Exposure	0.003**	0.002*	0.001	0.003*
	(0.001)	(0.001)	(0.002)	(0.001)
Observations	15,177	15,177	15,177	15,177
No. of Schools	1630	1630	1630	1630
Mean Dep. Var.	0.19	0.19	0.19	0.19
SD Dep. Var.	0.48	0.48	0.48	0.48
R-squared	0.342	0.346	0.238	0.350
Baseline Controls		✓	✓	✓
School FE	✓	✓	✓	✓
School Trends			✓	
Local Authories Trends				✓

Notes: Coefficients are obtained by estimating γ from Equation 3.3. Outcomes are calculated in % of all possible half-day openings. Regressions are weighted using percentage of P1-P3 students in 2013/2014 as weights. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) score and school average class size in 2004, both interacted with year dummies, alongside indicators for urban and religious all appropriately interacted with year dummies. Data span from school year 2003/2004 through to 2016/2017. Outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table C.5: Main Results - No Local Initiatives

	(1)	(2)	(3)	(4)
Panel A: Attendance				
Post × Exposure	0.011***	0.012***	0.004	0.012***
	(0.003)	(0.003)	(0.004)	(0.003)
Observations	16,097	16,097	16,097	16,097
No. of Schools	1477	1477	1477	1477
Mean Dep. Var.	95.01	95.01	95.01	95.01
SD Dep. Var.	1.58	1.58	1.58	1.58
R-squared	0.764	0.765	0.726	0.769
Panel B: Illnesses	(1)	(2)	(3)	(4)
Post × Exposure	-0.012***	-0.012***	-0.002	-0.008***
	(0.003)	(0.003)	(0.004)	(0.003)
Observations	16,120	16,120	16,120	16,120
No. of Schools	1477	1477	1477	1477
Mean Dep. Var.	2.58	2.58	2.58	2.58
SD Dep. Var.	1.56	1.56	1.56	1.56
R-squared	0.582	0.584	0.437	0.665
Panel C: Exclusions	(1)	(2)	(3)	(4)
Post × Exposure	0.001	0.001	-0.000	0.001
	(0.001)	(0.002)	(0.002)	(0.002)
Observations	15,687	15,687	15,687	15,687
No. of Schools	1477	1477	1477	1477
Mean Dep. Var.	0.19	0.19	0.19	0.19
SD Dep. Var.	0.49	0.49	0.49	0.49
R-squared	0.309	0.312	0.206	0.316
Baseline Controls		✓	✓	✓
School FE	✓	✓	✓	✓
School Trends			✓	
Local Authories Trends				✓

Notes: Coefficients are obtained by estimating γ from Equation 3.3. Outcomes are calculated in % of all possible half-day openings. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) score and school average class size in 2004, both interacted with year dummies, alongside indicators for urban and religious all appropriately interacted with year dummies. Data span from school year 2003/2004 through to 2016/2017. Outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. This sample does not include schools which extended eligibility following local initiatives from 2010. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table C.6: Main Results - No Pilot

Panel A: Attendance	(1)	(2)	(3)	(4)
Post × Exposure	0.013***	0.013***	0.004	0.012***
	(0.003)	(0.003)	(0.004)	(0.003)
Observations	13,756	13,756	13,756	13,756
No. of Schools	1263	1263	1263	1263
Mean Dep. Var.	95.24	95.24	95.24	95.24
SD Dep. Var.	1.44	1.44	1.44	1.44
R-squared	0.727	0.728	0.668	0.731
Panel B: Illnesses	(1)	(2)	(3)	(4)
Post × Exposure	-0.007**	-0.008***	0.001	-0.009***
	(0.003)	(0.003)	(0.004)	(0.003)
Observations	13,766	13,766	13,766	13,766
No. of Schools	1263	1263	1263	1263
Mean Dep. Var.	2.73	2.73	2.73	2.73
SD Dep. Var.	1.44	1.44	1.44	1.44
R-squared	0.571	0.575	0.473	0.662
Panel C: Exclusions	(1)	(2)	(3)	(4)
Post × Exposure	-0.000	-0.001	-0.001	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)
Observations	13,424	13,424	13,424	13,424
No. of Schools	1263	1263	1263	1263
Mean Dep. Var.	0.18	0.18	0.18	0.18
SD Dep. Var.	0.47	0.47	0.47	0.47
R-squared	0.267	0.271	0.173	0.275
Baseline Controls		✓	✓	✓
School FE	✓	✓	✓	✓
School Trends			✓	
Local Authories Trends				✓

Notes: Coefficients are obtained by estimating γ from Equation 3.3. Outcomes are calculated in % of all possible half-day openings. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) score and school average class size in 2004, both interacted with year dummies, alongside indicators for urban and religious all appropriately interacted with year dummies. Data span from school year 2003/2004 through to 2016/2017. Outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. This sample does not include schools which took part in the FSM pilot in school year 2007/08. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table C.7: Main Results - Imputed Missing

Panel A: Attendance	(1)	(2)	(3)	(4)
Post × Exposure	0.009***	0.009***	0.001	0.008***
	(0.003)	(0.003)	(0.003)	(0.003)
Observations	21,082	21,082	21,082	21,082
No. of Schools	1920	1920	1920	1920
Mean Dep. Var.	95.19	95.19	95.19	95.19
SD Dep. Var.	1.59	1.59	1.59	1.59
R-squared	0.723	0.725	0.682	0.729
Panel B: Illnesses	(1)	(2)	(3)	(4)
Post × Exposure	-0.009***	-0.008***	0.000	-0.007***
	(0.003)	(0.003)	(0.003)	(0.002)
Observations	21,082	21,082	21,082	21,082
No. of Schools	1920	1920	1920	1920
Mean Dep. Var.	2.50	2.50	2.50	2.50
SD Dep. Var.	1.52	1.52	1.52	1.52
R-squared	0.552	0.556	0.419	0.632
Panel C: Exclusions	(1)	(2)	(3)	(4)
Post × Exposure	0.002	0.001	0.000	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	21,082	21,082	21,082	21,082
No. of Schools	1920	1920	1920	1920
Mean Dep. Var.	0.16	0.16	0.16	0.16
SD Dep. Var.	0.45	0.45	0.45	0.45
R-squared	0.282	0.284	0.174	0.288
Baseline Controls		✓	✓	✓
School FE	✓	✓	✓	✓
School Trends			✓	
Local Authories Trends				✓

Notes: Coefficients are obtained by estimating γ from Equation 3.3. Missing values in outcomes and exposure due to statistical disclosure control have been randomly imputed with valued with values [1,4]. Outcomes are calculated in % of all possible half-day openings. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) score and school average class size in 2004, both interacted with year dummies, alongside indicators for urban and religious all appropriately interacted with year dummies. Data span from school year 2003/2004 through to 2016/2017. Outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table C.8: Main Results - Inverse Hyperbolic Sine

Panel A: Attendance				
	(1)	(2)	(3)	(4)
Post × Exposure	0.000***	0.000***	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	21,082	21,082	21,082	21,082
No. of Schools	1920	1920	1920	1920
Mean Dep. Var. (levels)	95.19	95.19	95.19	95.19
SD Dep. Var. (levels)	1.59	1.59	1.59	1.59
R-squared	0.723	0.725	0.682	0.729
Panel B: Illnesses				
	(1)	(2)	(3)	(4)
Post × Exposure	-0.004***	-0.003**	-0.001	-0.003**
	(0.002)	(0.002)	(0.001)	(0.001)
Observations	21,082	21,082	21,082	21,082
No. of Schools	1920	1920	1920	1920
Mean Dep. Var. (levels)	2.50	2.50	2.50	2.50
SD Dep. Var. (levels)	1.52	1.52	1.52	1.52
R-squared	0.560	0.566	0.399	0.664
Panel C: Exclusions				
	(1)	(2)	(3)	(4)
Post × Exposure	0.001**	0.001**	0.001	0.001**
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	21,082	21,082	21,082	21,082
No. of Schools	1920	1920	1920	1920
Mean Dep. Var. (levels)	0.16	0.16	0.16	0.16
SD Dep. Var. (levels)	0.45	0.45	0.45	0.45
R-squared	0.359	0.360	0.253	0.363
Baseline Controls		✓	✓	✓
School FE	✓	✓	✓	✓
School Trends			✓	
Local Authories Trends				✓

Notes: Coefficients are obtained by estimating γ from Equation 3.3. Missing values in outcomes and exposure due to statistical disclosure control have been randomly imputed with valued with values [1,4]. Outcomes are calculated in % of all possible half-day openings. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) score and school average class size in 2004, both interacted with year dummies, alongside indicators for urban and religious all appropriately interacted with year dummies. Data span from school year 2003/2004 through to 2016/2017. Outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table C.9: Secondary Schools Triple DiD

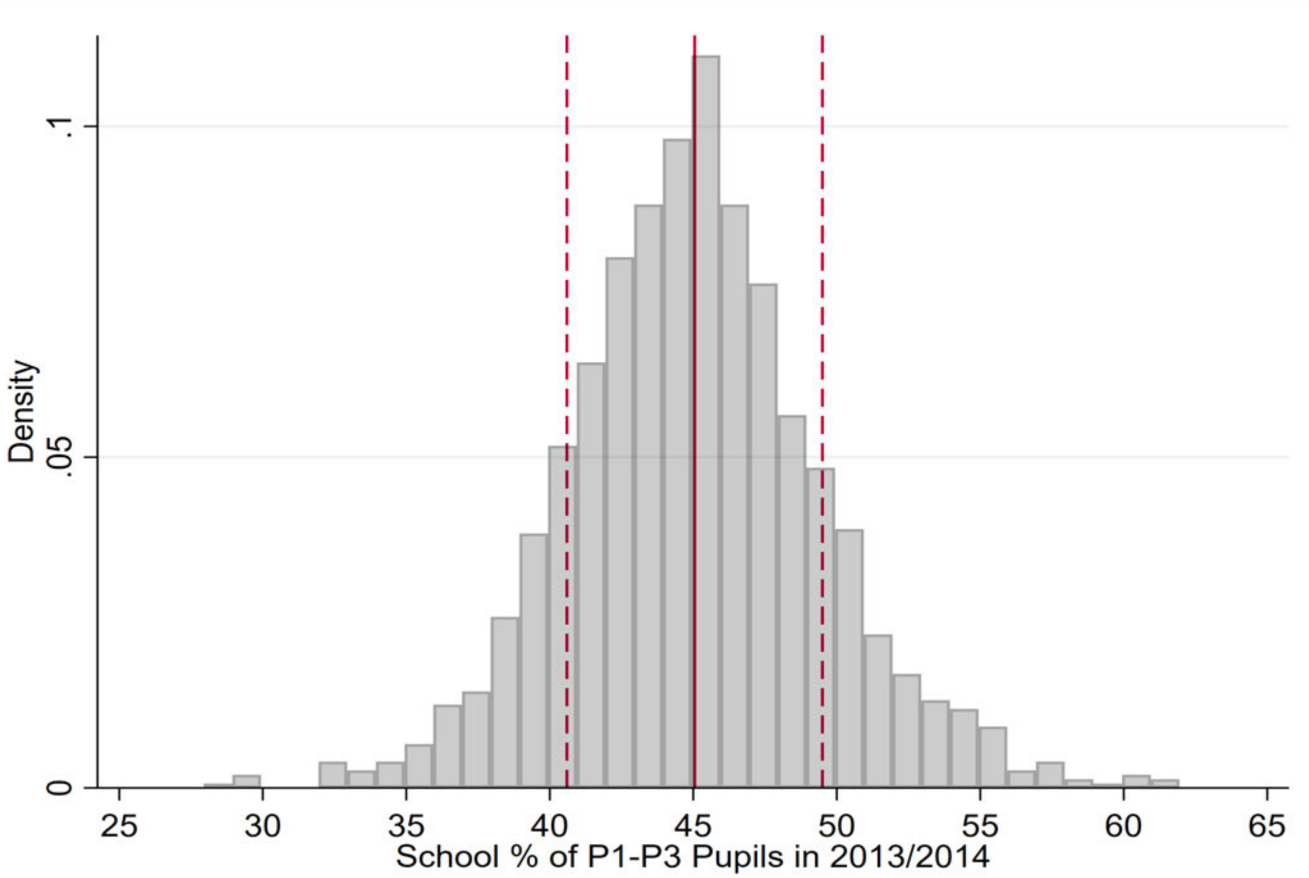
Panel A: ln(Attendance)	(1)	(2)	(3)	(4)
Post × Exposure × Primary	0.001***	0.001***	-0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	21,209	21,209	21,209	20,547
No. of Schools	1949	1949	1949	1949
Mean Dep. Var.	4.55	4.55	4.55	4.55
SD Dep. Var.	0.03	0.03	0.03	0.03
R-squared	0.844	0.845	0.806	0.841
Panel B: ln(1 + Illnesses)	(1)	(2)	(3)	(4)
Post × Exposure × Primary	0.010**	0.009**	0.001	0.003
	(0.005)	(0.004)	(0.005)	(0.003)
Observations	21,346	21,346	21,346	20,683
No. of Schools	1949	1949	1949	1949
Mean Dep. Var.	1.19	1.19	1.19	1.19
SD Dep. Var.	0.61	0.61	0.61	0.61
R-squared	0.614	0.616	0.436	0.725
Panel C: ln(1 + Exclusions)	(1)	(2)	(3)	(4)
Post × Exposure × Primary	-0.019***	-0.019***	-0.002	-0.016***
	(0.003)	(0.003)	(0.004)	(0.003)
Observations	20,850	20,850	20,850	20,187
No. of Schools	1949	1949	1949	1949
Mean Dep. Var.	0.25	0.25	0.25	0.25
SD Dep. Var.	0.40	0.40	0.40	0.40
R-squared	0.725	0.713	0.664	0.677
Baseline Controls		✓	✓	✓
School FE	✓	✓	✓	✓
School Trends			✓	
Local Authories Trends				✓

Notes: Coefficients are obtained by estimating γ from Equation 3.5. Outcomes are calculated in % of all possible half-day openings and are in logarithmic form. Exclusion rate is reported per 1,000 students and refer to the number of half-days missed on the account of temporary exclusion. Baseline controls include Scottish Index of Multiple Deprivation (SIMD) score and school average class size in 2004, both interacted with year dummies, alongside indicators for urban and religious all appropriately interacted with year dummies. Data span from school year 2003/2004 through to 2016/2017. Outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Figure C.2: Trends

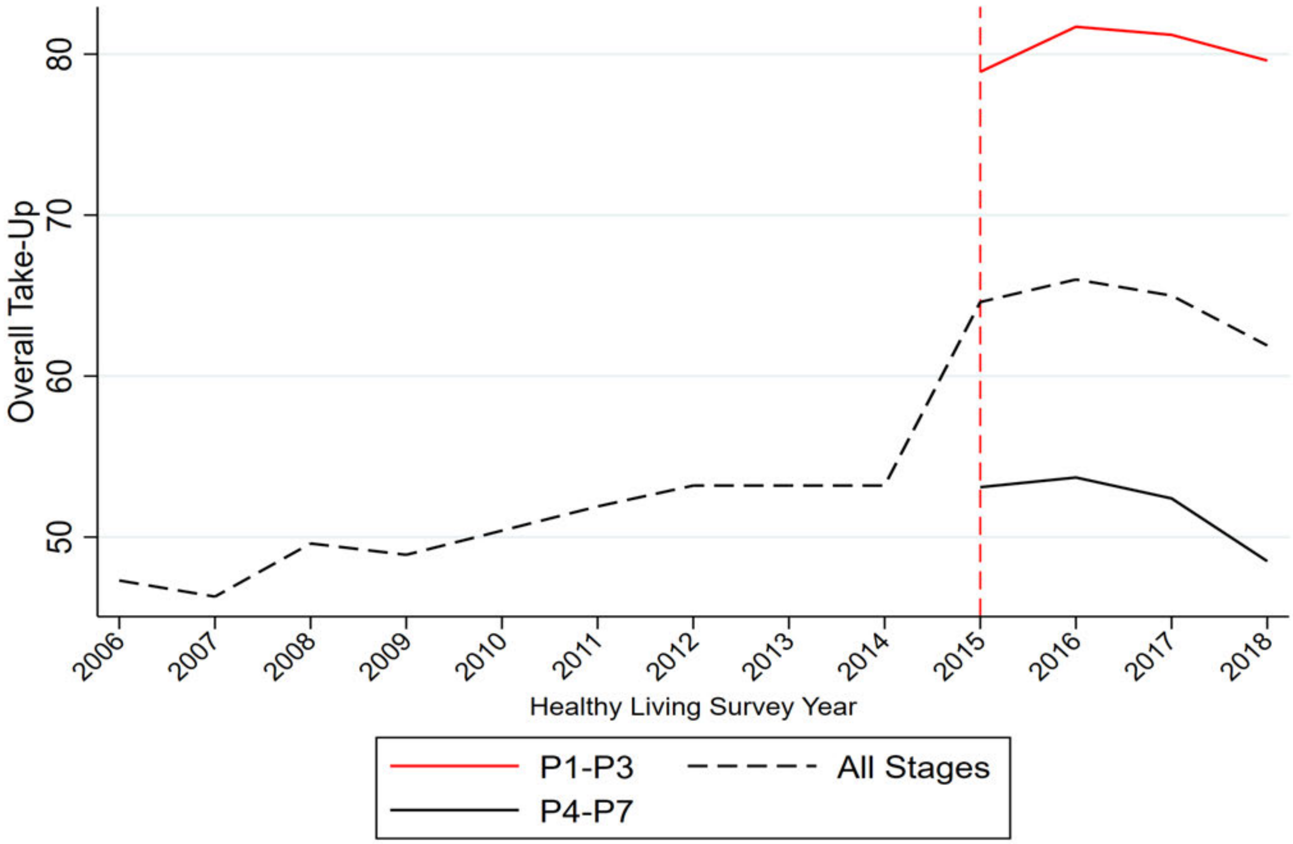
Notes: Our sample seems to follow the same trends (except for FSM uptake around 2008) of the official statistics released within SG reports. However, there it seems to be a systematic overstatement of all the variables. We suspect this is due to statistical disclosure control which suppresses values between 0 and 5. This way, the variable in question will necessarily have a large mean as the smallest values are removed from the sample.

Figure C.3: P1-P3 distribution in School Year 2013/2014



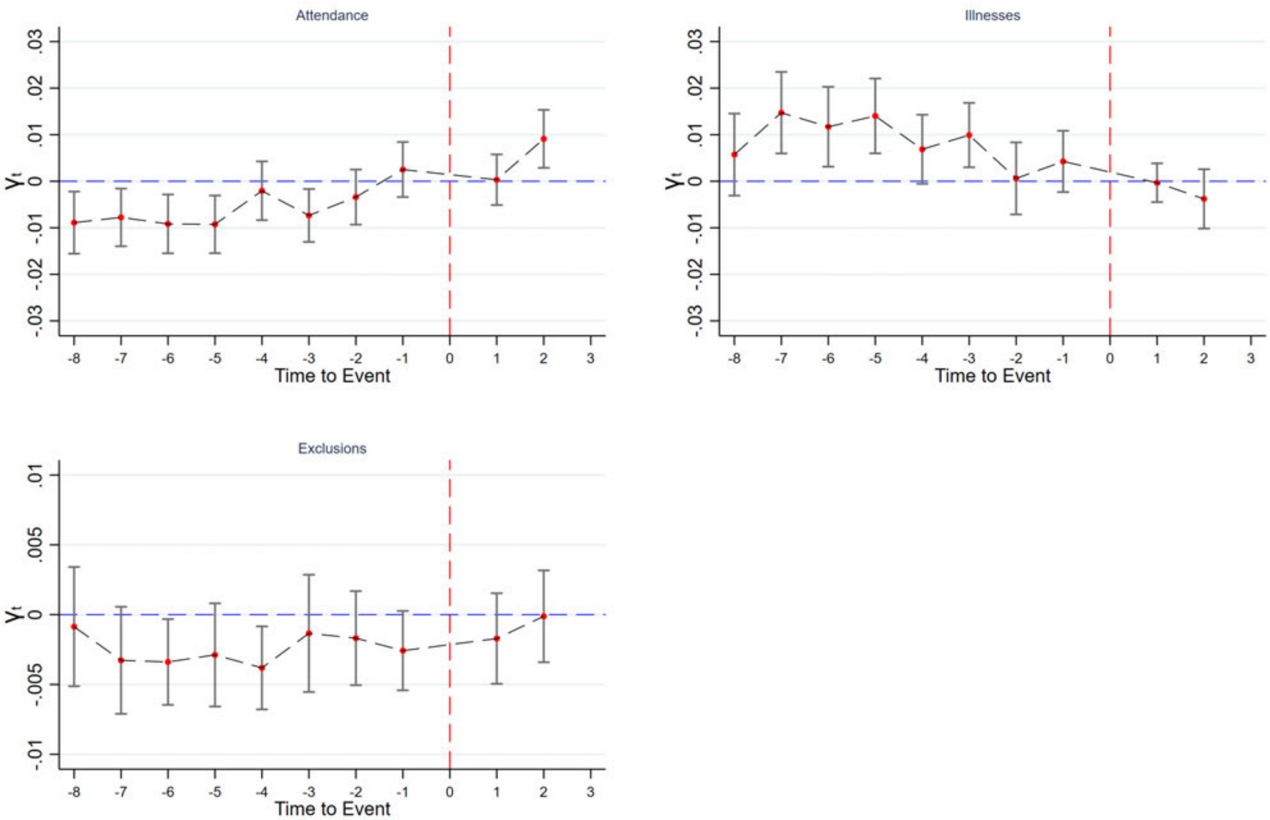
Notes: This is the distribution of P1-P3 pupils in school year 2013/2014, the one immediately before UFSM implementation. The solid red line is the average, whereas the dashed lines identify the mean +/- one standard deviation. Each bin correspond to one percentage point.

Figure C.4: FSM vs PSM Uptakes



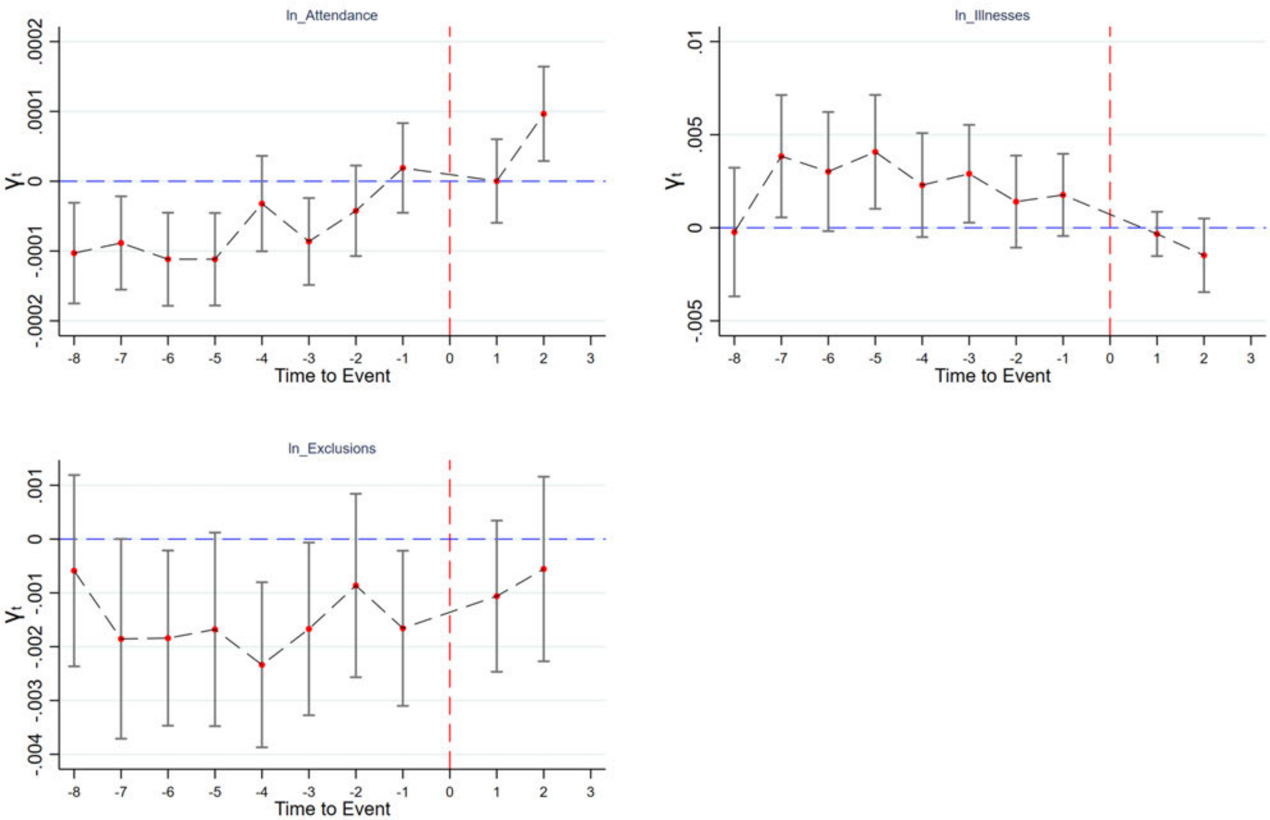
Note: The above chart is calculated as raw yearly-averages of the overall school meal take-up, i.e. $\frac{\#Meal-Takers}{School\ Population}$. Because the survey is run in one day, the raw counts refer to pupils present on the day of the survey. The dotted black line represents the entire primary school-level trend. The solid red line dis-aggregates the trend for the P1-P3 group (targeted by the policy) whereas the solid black line shows the P4-P7 trend. The breakdown of registration and take-up is only available, at the national level, starting from HLS wave 2015.

Figure C.5: Event Study - Unbalanced Panel



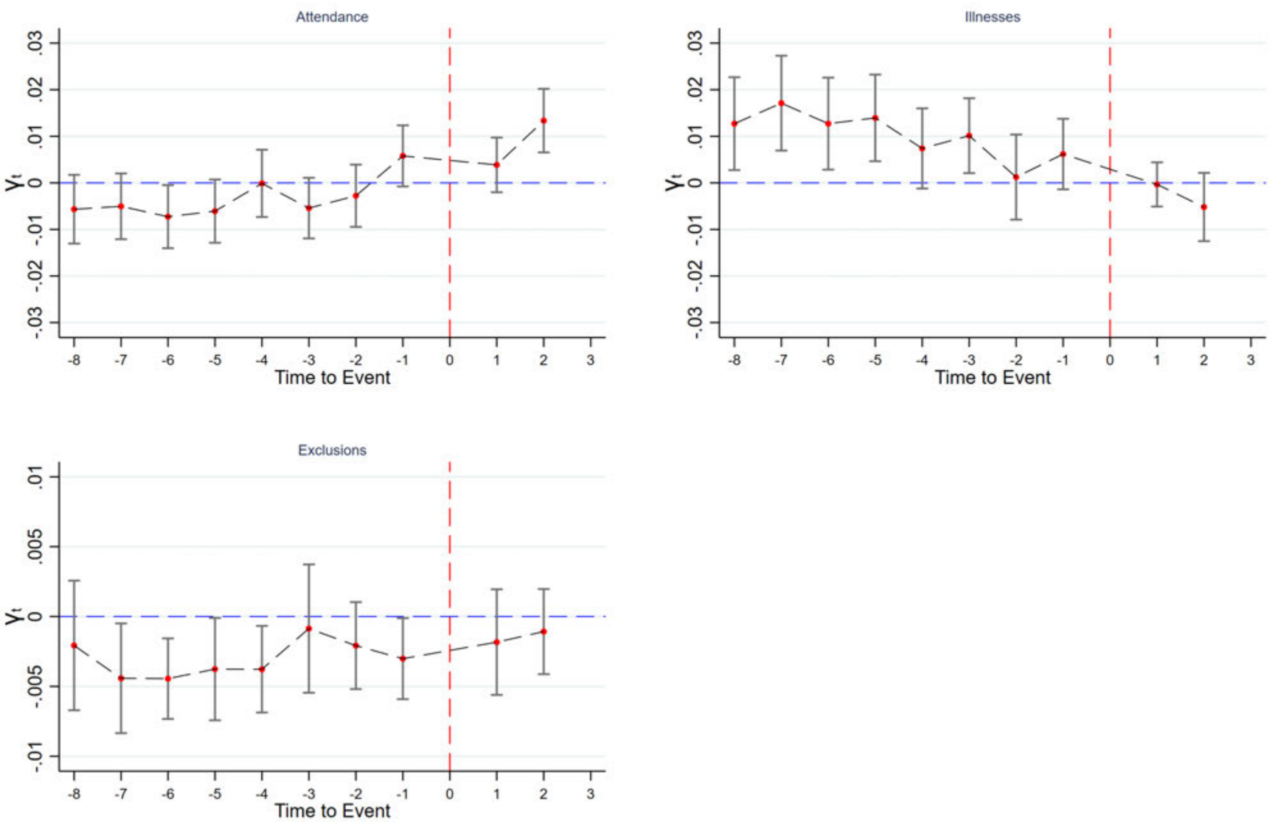
Notes: Coefficients are obtained by estimating γ_t from Equation 3.4. Each coefficient can be interpreted as the difference in outcome for that period relative to the reference period, which is school year 2012/2013, the earliest we observe the outcomes before the change in policy. Time is expressed in number of periods to and from the reference period, which is 0. Period one is year 2014/2015, namely when UFSM is in force. Exposure is defined as the % of school population not-taking FSM in 2013, one year before the policy change and outcomes are in levels. Data span from school year 2003/2004 through to 2016/2017 and outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. This sample also includes schools which are not observed every year. Whiskers represent coefficients' 95% confidence interval. Standard errors are clustered at the school level.

Figure C.6: Event Study in Logarithmic Form



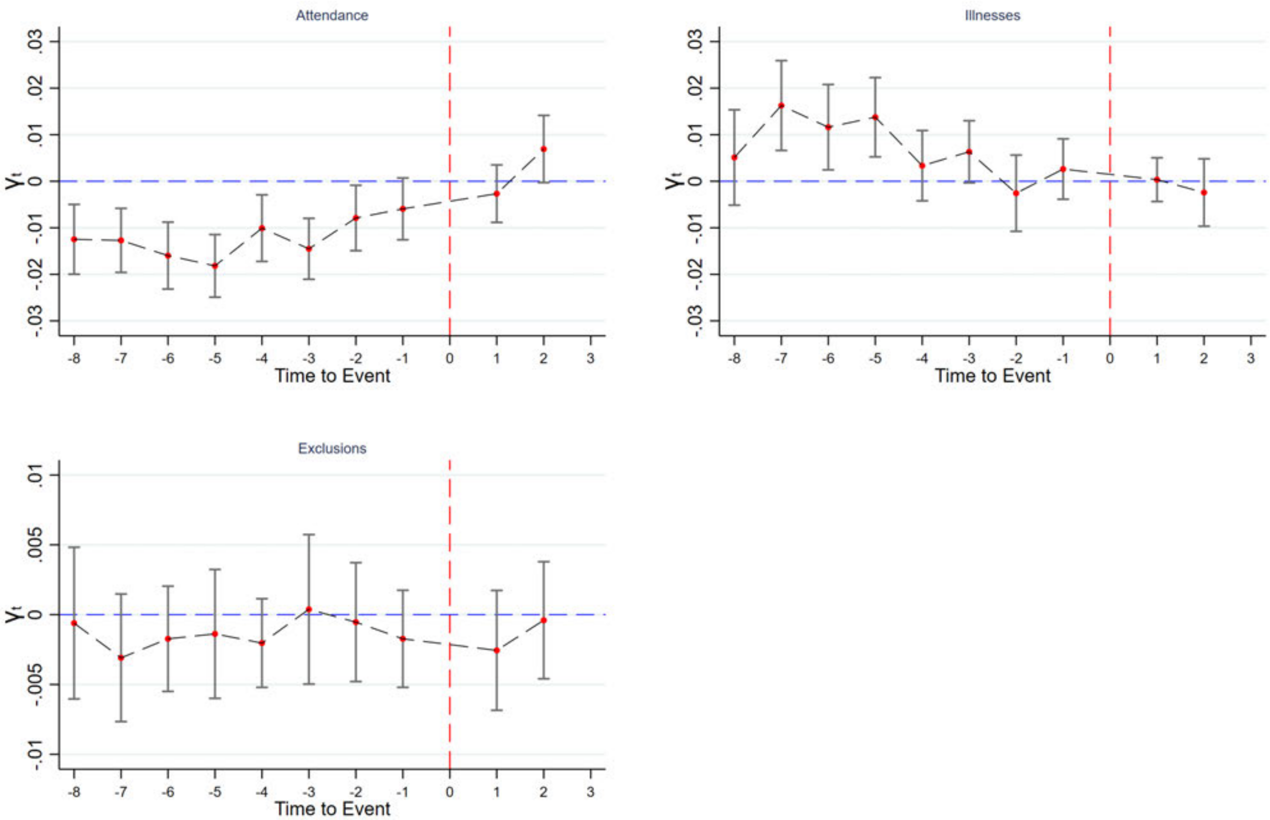
Notes: Coefficients are obtained by estimating γ_t from Equation 3.4. Each coefficient can be interpreted as the difference in outcome for that period relative to the reference period, which is school year 2012/2013, the earliest we observe the outcomes before the change in policy. Time is expressed in number of periods to and from the reference period, which is 0. Period one is year 2014/2015, namely when UFSM is in force. Exposure is defined as the % of school population not-taking FSM in 2013, one year before the policy change and outcomes are in log. Data span from school year 2003/2004 through to 2016/2017 and outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. Whiskers represent coefficients' 95% confidence interval. Standard errors are clustered at the school level.

Figure C.7: Event Study - No Local Initiatives



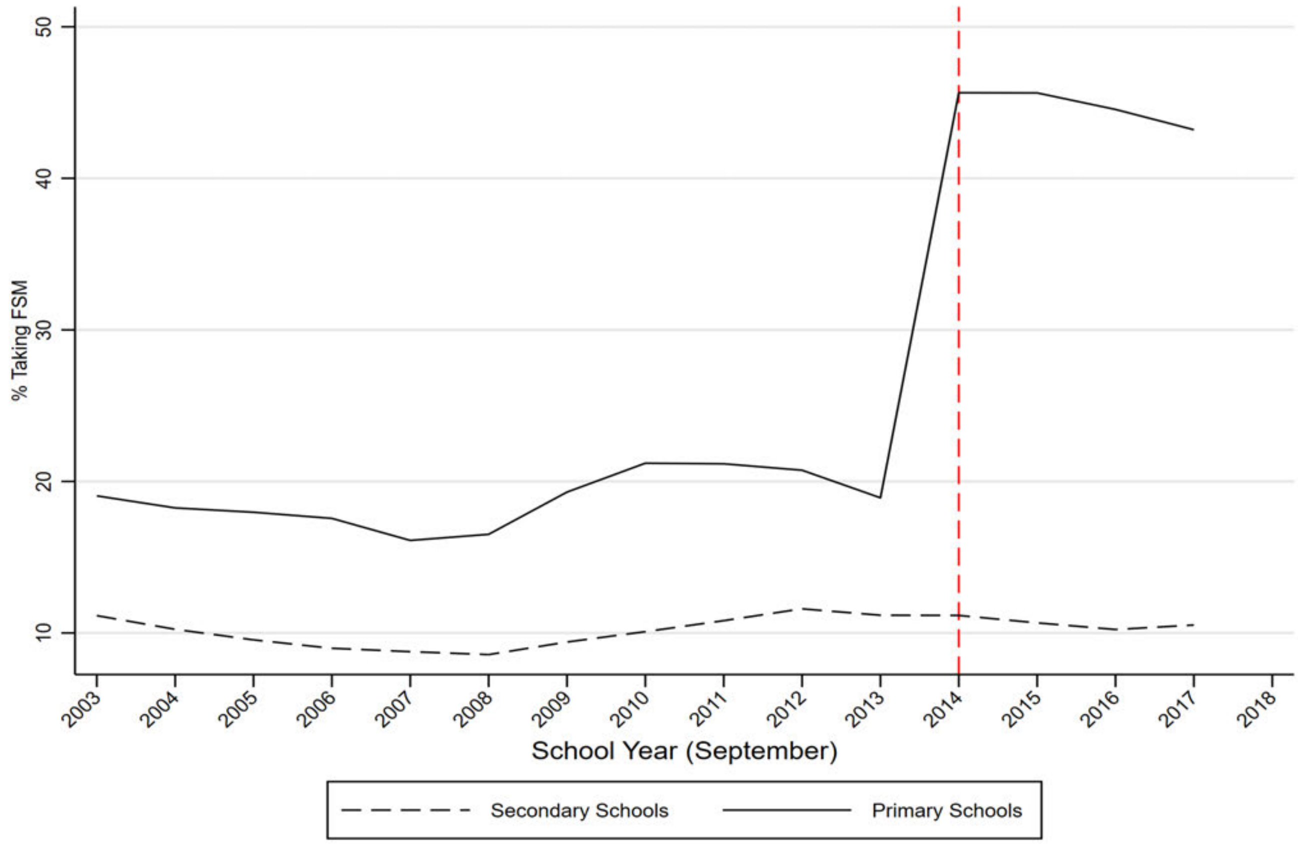
Notes: Coefficients are obtained by estimating γ_t from Equation 3.4. Each coefficient can be interpreted as the difference in outcome for that period relative to the reference period, which is school year 2012/2013, the earliest we observe the outcomes before the change in policy. Time is expressed in number of periods to and from the reference period, which is 0. Period one is year 2014/2015, namely when UFSM is in force. Exposure is defined as the % of school population not-taking FSM in 2013, one year before the policy change and outcomes are in levels. Data span from school year 2003/2004 through to 2016/2017 and outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. This sample does not include schools which extended eligibility following local initiatives from 2010. Whiskers represent coefficients' 95% confidence interval. Standard errors are clustered at the school level.

Figure C.8: Event Study - No Pilot



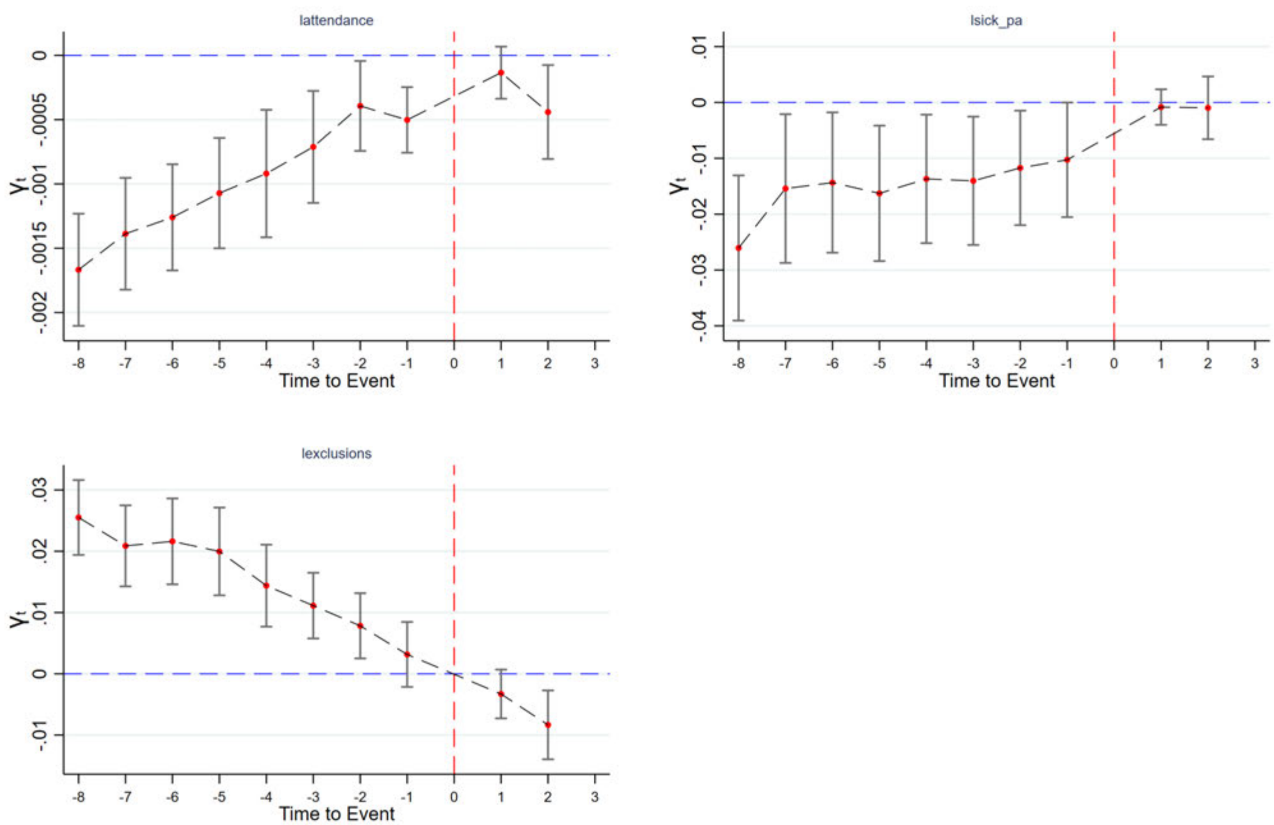
Notes: Coefficients are obtained by estimating γ_t from Equation 3.4. Each coefficient can be interpreted as the difference in outcome for that period relative to the reference period, which is school year 2012/2013, the earliest we observe the outcomes before the change in policy. Time is expressed in number of periods to and from the reference period, which is 0. Period one is year 2014/2015, namely when UFSM is in force. Exposure is defined as the % of school population not-taking FSM in 2013, one year before the policy change and outcomes are in levels. Data span from school year 2003/2004 through to 2016/2017 and outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. This sample does not include schools which took part in the FSM pilot in school year 2007/08. Whiskers represent coefficients' 95% confidence interval. Standard errors are clustered at the school level.

Figure C.9: Uptakes Trends Primary vs Secondary Schools



Notes: The above trends are calculated as raw yearly-averages of the following ratio: % Taking FSM: $\frac{\#FSM-Takers}{School\ Population}$ separately for primary and secondary schools, which were not targeted by the policy.

Figure C.10: Secondary Schools Triple DiD Event Study



Notes: Coefficients are obtained by estimating γ_t from Equation 3.4. Each coefficient can be interpreted as the difference in outcome for that period relative to the reference period, which is school year 2012/2013, the earliest we observe the outcomes before the change in policy. Time is expressed in number of periods to and from the reference period, which is 0. Period one is year 2014/2015, namely when UFSM is in force. Exposure is defined as the % of school population not-taking FSM in 2013, one year before the policy change and outcomes are in log. Data span from school year 2003/2004 through to 2016/2017 and outcomes are not available for years 2011/2012, 2013/2014 and 2015/2016. Whiskers represent coefficients' 95% confidence interval. Standard errors are clustered at the school level.

National Eligibility Criteria

The national criteria for eligibility to free school meals includes the following:[1]

- Pupils within families who receive Income Support, Income-based Job Seekers Allowance or any income related element of Employment and Support Allowance.
- Pupils within families who receive support under Part VI of the Immigration and Asylum Act 1999.
- Pupils whose parents or carers receive Child Tax Credit, do not receive Working Tax Credit and had an annual income (as assessed by the Inland Revenue) of below £16,105 (from April 2013).
- Pupils whose parents or carers are in receipt of both Child Tax Credit and Working Tax Credit and their income is up to £6,900 were also entitled (from August 2009).
- Pupils whose parents or carers are in receipt of Universal Credit and their monthly earned income does not exceed £610 were also entitled (from August 2017).
- Pupils in school education who receive any of these benefits in their own right are also entitled to receive free school meals.

[1]Source: School Healthy Living Survey Statistics: 2020. Scottish Government.

Table C.10: Sample Description

School Year	SPC Wave	HLS Wave	School Meals Data	Attendance Data	Final Sample
2003/2004	2003	2004	✓	✓	✓
2004/2005	2004	2005	✓	✓	✓
2005/2006	2005	2006	✓	✓	✓
2006/2007	2006	2007	✓	✓	✓
2007/2008	2007	2008	✓	✓	✓
2008/2009	2008	2009	✓	✓	✓
2009/2010	2009	2010	✓	✓	✓
2010/2011	2010	2011	✓	✓	✓
2011/2012	2011	2012	✓		
2012/2013	2012	2013	✓	✓	✓
2013/2014	2013	2014	✓		
2014/2015	2014	2015	✓	✓	✓
2015/2016	2015	2016	✓		
2016/2017	2016	2017	✓	✓	✓
2017/2018	2017	2018	✓		
2018/2019	2018	2019	FSM-Registration Only	✓	

Notes: Scottish Pupils Census (SPC) is run every school year in September. School Meals Survey (subsequently renamed Healthy Living Survey, HLS) is run in February of the school year beginning the previous August. After 2010/2011 school year attendance and absence survey were administered every second year. This explains the gaps.