The Effectiveness and Equity of Grammar Schools in England

BINWEI LU

The Effectiveness and Equity of Grammar Schools in England

Binwei Lu

PhD Thesis

School of Education

Durham University

2019

Table of Contents

Acknowledgement

I would like to thank my supervisors, Professor Stephen Gorard and Dr Beng Huat See, for always guiding me and encouraging me during this journey. Thanks also go to Dr Nadia Siddiqui for her ongoing help. Their support means a lot to me.

I would like to thank Professor Christine Merrell for her help and her valuable comments on this study at earlier stages.

I would like to thank Mrs Joanne Bartley and Dr Alan Bainbridge. They helped me patiently when I was struggling to collect some important data of this study.

I would like to thank Professor Prue Holmes, who has supported me since my master's programme.

Finally, I am grateful to my family, my pet, and my friends. They are the source of my strength.

List of Abbreviations

EAL	English as an Additional Language
FSM	Free School Meals
HE	Higher Education
HESA	Higher Education Statistics Agency
IDACI	Income Deprivation Affecting Children Index
KS	Key Stage
LA	Local Authority
ML	Multilevel
NPD	National Pupil Database
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
PISA	Programme of International Student Assessment
RCT	Randomised Controlled Trial
RDD	Regression Discontinuity Design
SES	Socioeconomic Status
SEN	Special Educational Needs
SEN-PS	Special Educational Needs - School Action Plus or Statement
TIMSS	Third International Mathematics and Science Study
VA	Value-added

Abstract

In England, secondary education is mainly comprehensive in nature. However, a small group of grammar schools retain ability-based selection at the age of 11. Students in these 163 schools obtain high grades at Key Stage 4 (KS4), aged 16. Some commentators and policy-makers believe that these high grades are an indication that grammar schools are therefore good schools, and suggest increasing the number of grammar school places. They also claim that grammar schools provide opportunities for economically poorer but able students. At the same time, other commentators believe that grammar schools are elitist and socially divisive, actually providing further opportunities mostly for the already privileged. They urge successive governments to abolish the remaining grammar schools. The issue has become a particular 'hot topic' in the last two elections in England.

This study looks at the impact of the presence of grammar schools in some local authorities (LAs), and considers whether the expansion or removal of grammar schools would lead to a more effective and equitable education system. This new study uses the National Pupil Database (NPD), the Higher Education Statistics Agency (HESA) data on university attendance, and a local dataset providing the results of the grammar schools' selection test in one LA for one year. The study looks at access to grammar schools, and its relationship with pupil's prior attainment, geographical location, and family background. It considers whether the opportunity to attend grammar schools, and outcomes of attending grammar schools, are distributed equally among different social groups. The effectiveness of grammar schools is evaluated based on pupil's KS4 attainment and their Higher Education (HE) participation patterns. The analysis is based on standard regression models, controlling for pre-existing differences between pupil groups, and uses the innovative design of regression discontinuity (RDD) which is more robust in making causal inferences. The study also looks at the influence of the presence of grammar schools on the overall performance level in the local area, looking at the effectiveness of selective LAs and non-selective LAs in a holistic way. Lastly, the link between family background and post-18 destination in both types of LA are compared. Based on access to and outcome of grammar schools, the analysis demonstrates the possible impacts of the selective system on the redistribution of educational opportunities. This study is in a position to help decide whether the policy of grammar school expansion would work as claimed by the government, or whether the removal of grammar schools would make any difference.

Overall, once prior attainment is accounted for, the study found a small advantage in KS4 attainment associated with grammar school attendance, only for borderline pupils, rather than the highest performers. The study found no advantage from grammar school attendance for HE participation. Selective LAs have no better results than non-selective LAs. And pupils in selective LAs who failed to attend grammar schools gained worse results than equivalent ones in comprehensive LAs. Therefore, the claim that increasing grammar school places will help raise national academic standard is unrealistic, and not founded in existing grammar school outcomes.

In addition, attainment at an early age differs considerably between social groups. Selection at an early age means that the opportunity to attend grammar school is heavily unbalanced between these groups. The layered early attainment compounded by the imperfect selection process means that the small advantage that may be associated with attending grammar schools for borderline entrants and the cost of not attending grammar schools is not equally spread between social groups. The mechanism of selection and differentiation thus amplifies the attainment gap between pupils from high and low socioeconomic groups. Therefore, the selective system fails to comply with the demands of equity, whether merit-based or needs-based, and whether taking pupil's prior attainment into account or not. Moreover, separating pupils into different schools at an early-age creates segregated school compositions, which ultimately endangers the integration and long-term development of the society. To conclude, this study does not find any substantial benefit associated with early-age academic selection in England, and the expansion of grammar schools is unlikely to raise national performance standard, or to provide more opportunities for the poor.

The study also includes the first attempt at using the RDD to evaluate the current effectiveness of grammar schools in England. However, the incomplete data available on the results of the selection test means the application of the RDD is more a feasibility trial than a definitive test intended to settle the debate on the effectiveness of grammar schools. Conducting this design with national data on grammar school selection would create the most powerful evidence so far. The study thus advocates that the government should make the responsible decision to disclose grammar school selection data for the purposes of research. Promoting an effective and equitable education system is important, and the public deserve to know more. If the results are confirmed with full data on the results of the selection test, then perhaps the next step would be to phases the 163 grammar schools out altogether.

1 Introduction

This chapter starts with a background discussion of the general practice of academic selection and the specific situation of academically selective secondary schools in England. It then discusses the purpose and significance of this study, explaining why this study is important not only to England, but also to most countries around the world. Lastly, it presents the study's research questions and the structure of the following chapters.

1.1 Background

The action of selecting pupils based on early-age academic ability, and then providing divergent educational routes depending on the selection results, is usually referred to as academic selection. Based on the perception that academic selection raises overall educational standards while also providing educational opportunities for the economically poorer but able, selective placement has become a long-established practice worldwide. However, detractors argue that the selective system is no more effective than a comprehensive system, and it may reinforce the link between family origins and future destinations. Considering the potential harm of early-age academic selection, the secondary education in most Anglophone countries has been converted into a more comprehensive system. However, between-school separation remains a widespread practice around the world. Among the 72 participants in The Programme of International Student Assessment (PISA), 15 countries (e.g. Germany and Singapore) still select their pupils before the age of 15 (OECD, 2015). The long-lasting debate on the role of early-age academic selection in social effectiveness and equity implies the importance of an accurate evaluation of its real influence.

In England, secondary education also has a long history of dividing pupils into pathways according to their academic abilities. Dating back to the 1940s, the general principle underlying the education system was that secondary education should be selective (Kerckhoff et al., 1998). Thus, the 1944 Education Act proposed the Tripartite system, which divided pupils into three types of secondary schools based on their abilities. Under the fully-selective system, grammar schools were intended for academically-oriented pupils. They would provide the primary route to Higher Education (HE) institutions. In contrast, secondary modern schools usually taught practical subjects, equipping pupils with the basic literacy and numeracy necessary for manual work. In addition to these two types of schools, the 1994 Education Act also envisaged technical schools intended to teach specialised technical subjects in addition to general

education, preparing their pupils for post-war industries (Morris & Perry, 2017). However, this strand of technical education failed, as few technical schools were established, and only 2% to 5% of English pupils attended technical schools after the Act (Foreman-Peck, 2004). Thus, the Tripartite System became a Bipartite system, distinguishing high-performers from the other. Along with the global trend of comprehensivisation, the division between grammar schools and secondary modern schools has also been largely effaced since the 1960s, and most schools were transitioned to mixed-ability schools. However, 163 grammar schools remain in England, and they educate about 5% of the pupils in the English state system today (Bolton, 2017). Despite their small number, grammar schools receive considerable political and public attention because of their link to effectiveness and equity, which are not only political buzzwords in England, but significant global issues.

Effectiveness is one of the most important aspects of education systems. Since grammar schools usually have high rankings in school league tables, there are beliefs that they may offer a model of effective schooling. Regarded by some in government as better at raising academic standards than other state-funded schools, grammar schools in England have prompted political attempts at expansion, and this tendency has intensified in recent years (DfE, 2016; The Conservative Party, 2017). However, the expansion of grammar schools has received resistance as well, most noticeably from campaign groups (e.g. Comprehensive Future and the Kent Education Network) and academics, due to concerns such as low academic benefits, segregated school composition, and impediments to disadvantaged pupils (Allen, Bartley & Nye, 2017; Andrews, Hutchinson & Johnes, 2016; Cribb et al., 2013). Based on these concerns, it is believed that existing grammar schools should not be expanded, and that they may even need to be closed. Despite a plethora of studies from both the government and researchers, evidence of grammar school effectiveness is mixed. Since children cannot be randomly allocated to schools in order to test the difference each route makes to their life chances, most existing research uses statistical models to control for pre-existing differences between grammar school pupils and those in non-selective schools. However, the reliance on passive designs not conducive to causal inference means that the results only reveal the correlation between grammar school attendance and academic performance. These estimations become biased whenever influential baseline variables between pupil groups are either neglected, unavailable, or unmeasurable. Thus, when differences in later attainment emerge, it is unclear whether they are due to the school attended, or imperfections in the modelling process. This casts doubt on the estimated grammar school effect.

Based on the limitations of previous research controlling for pre-existing differences between pupil groups, and the infeasibility of a randomised controlled trial (RCT), the best solution would be to adopt a regression discontinuity design (RDD) (Imbens & Lemieux, 2008). As robust as an experimental design, an RDD compares cases in the neighbourhood of a cut-off point. Since cases in the control group who have just missed the cut-off point, and those in the treatment group who have just made the threshold, are similar, this design reduces the problem of pre-existing differences between the two groups (Imbens & Lemieux, 2008; Lee & Lemieux, 2009). While there have been fruitful applications of RDDs for school effectiveness, so far only one study has used this design to evaluate grammar schools in England (Clark, 2010). However, since that study used data from the late 1960s and 1970s, the situation may be different in today's grammar schools. Due to the mixed results from research controlling for pre-existing differences between pupil groups, and the inadequate evidence from research applying a strong research design to make robust causal statements, grammar school effectiveness is still in need of evaluation.

While performance standards are a major concern for governments, outcome distribution among different social groups is also a critical issue. Educational equity is emphasised internationally, and it has become a core issue in the international development agenda proposed by the United Nations (2015). The critical role of education in creating economic and social benefits means that access to education may influence pupils' life trajectories. Therefore, an equitable distribution of educational opportunities and outcomes should not be dependent on characteristics beyond pupils' control, such as parental income, ethnicity, religion, and gender (Levin, 1990; Rawls, 1971; Roemer, 1998).

With their assumed roles in helping less advantaged pupils fulfil their potential constantly emphasised, grammar schools are believed to present a tight connection with educational equity. This is because grammar schools select pupils by ability alone, rather than by other family background characteristics such as social status and income. However, the underrepresentation of disadvantaged pupils in grammar schools suggests that grammar schools might not only be academically selective, but socially selective as well. Therefore, the assumption behind grammar schools' role in helping the poor also needs further assessment.

Most political attempts to expand grammar schools are based on two claims: that grammar schools raise national academic standards, and that they provide the economically poorer but able with more opportunities. This implies that grammar schools are both effective and equitable. However, as mentioned above, despite advocates' confidence, there are also opposing calls from researchers and parental groups, who emphasise the negative influence of grammar schools on overall academic standards and the distribution of academic opportunities, advocating for their closure. As the evidence of grammar schools' effectiveness and equity has been mixed and inadequate in previous research, this study provides a thorough evaluation of these issues.

1.2 Purpose and significance of the study

Based on the special roles of grammar schools in educational effectiveness and equity, as well as the mixed evidence in previous research, this study thus focuses on potential impacts of grammar schools and evaluates whether expanding or closing them would lead to a more effective and equitable education system.

This study examines whether the opportunity to attend grammar schools is spread equally across social groups, and whether attending grammar schools is associated with outcomes favourable to those of other state-funded schools. It also addresses how the potential influence of attending grammar schools is distributed across social groups. The study considers possible consequences of the expansion of grammar schools, and evaluates whether such an expansion policy would work as it has been claimed. The findings are especially crucial at the time of writing, as the expansion of grammar schools is being promoted by multiple groups, and actions are being taken towards its implementation. Based on the high costs of new grammar school places, the small number of potential grammar school participants, and the concurrent need to invest in basic educational areas in England, governments and practitioners need better evidence on these issues. This study provides implications for future development of secondary education in England which could promote an effective and equitable education system for generations.

This study's findings are not merely important to England. While the analysis only focuses on grammar schools in England, this study is relevant to the practices of selection and differentiation around the world. The findings provide implications for other areas with early-

age academic selection, such as between-school selection and tracked curricula. Furthermore, it also has relevance for the possible consequences of other forms of early-age selection which separate pupils into schools based on characteristics such as gender or religion. While the selection may adopt distinct forms under different specific circumstances, the underlying principle of selection is the same, which separates pupils according to certain characteristics.

1.3 Research questions and thesis outline

1. Does the opportunity to attend grammar schools vary systematically between pupil groups? How does it correlate with pupil's prior attainment, geographical location and family background?
2. Are grammar schools more effective than non-selective mainstream state-funded schools in improving pupils' academic performance?
3. Are selective local authorities (LAs) more effective than non-selective LAs in improving pupils' academic performance?
4. What is the relationship between grammar school attendance and participation in HE?
5. Is the link between pupil's family background and post-18 destination stronger in selective LAs than in non-selective LAs?

The remainder of this thesis is structured into four parts: a summary of relevant theories, policies, and existing evidence on selective schooling (Chapter 2-8); the methods used in the present study (Chapter 9); the findings relevant to each research question (Chapter 10-15); and the conclusions and study implications (Chapter 16-20).

The literature part starts with the nature and definition of selective education in general, and the reason why selection at the compulsory education stage is a sensitive issue. This is followed by a discussion of contemporary secondary education systems around the world, the possible increase in selection and differentiation in the purported quasi-market for schools, and the assumed benefits and possible consequences of the quasi-market. After demonstrating the emphasis on selection in the quasi-market, the discussion turns to examples of schools which select on the basis of different pupil characteristics. After discussing schools which select by pupils' academic ability, issues that arise from attempts to theoretically define and practically measure 'ability' are then revealed.

After mapping the broader field, the next part of literature review focuses on two critical terms in this study— 'effectiveness' and 'equity'. First, theories of and previous research on school effectiveness are discussed. This includes the definition of school effectiveness, the influence of schools, the role of peer groups, and the relationship between academic selection and school effectiveness. The literature of school effectiveness is followed by a discussion of equity, which presents difficulties in defining equity, the importance of equity, and why equity is relevant to academic selection. In addition to independent discussion of effectiveness and equity, whether there is a systematic trade-off between these two elements is also addressed. This part also discusses how academic selection is connected to the segregation of pupil characteristics between schools.

The last part of the literature focuses on the status of grammar schools in England. It first demonstrates their origin, historical changes since the last century, as well as current political attempts to both expand and to close them. The discussion then turns to existing evidence on grammar schools' effectiveness. Based on inconsistent conclusions in previous research controlling for baseline characteristics, it discusses possible reasons for why the evidence differs, as well as the limitations of these estimation results. Evidence in the only previous study using the RDD approach for grammar school effectiveness is also presented. This is followed by previous evidence of grammar schools' role in social equity. This section demonstrates how access to grammar schools differs between pupils, whether it has a systematic correlation with family background, and the segregated pupil compositions of grammar schools.

After discussing theories and previous empirical evidence, the methods chapter details the data applied and the choice of statistical approaches at each step. This is followed by the findings chapters which demonstrate the study's analytical results. First, the opportunity to attend grammar schools for different pupil groups is analysed. Its relationship with pupil's prior attainment, geographical location, and pupil's family background is systematically evaluated. The analysis of this part is conducted on the basis of an earlier version (Lu, 2018). After revealing the low rates of attending grammar schools for certain pupil groups, both before and after accounting for prior attainment, the study explores possible reasons for this pattern. The analysis in this part thus addresses the first research question. The second part of the findings chapters presents grammar schools' effectiveness in terms of improving pupil's Key Stage 4 (KS4) GCSE performance. The analysis is first conducted through traditional approaches,

applying Ordinary Least Squares (OLS) models and logistic regression models. The evaluation is then conducted using the RDD approach. A simplified version of the RDD analysis on the effectiveness of grammar schools is available in Lu's previous paper (2019). In this part, grammar schools and their non-selective counterparts in selective LAs are compared. For simplicity, LAs with grammar schools within the local area are referred to as 'selective LAs' in this study. This is consistent with the definition used in several previous studies (e.g. Gorard & Siddiqui, 2018). In addition to the national picture, the analysis is also complemented by the pattern of individual selective LAs. This considers the broader geographical and social context of each LA. The analysis also pays special attention to pupils eligible for Free School Meals (FSM), exploring the differential effectiveness of grammar schools and their role in narrowing the attainment gap between high and low socioeconomic status (SES) groups. Based on these findings, the analysis answers the second research question. Following the analysis of grammar schools' effectiveness, the study then discusses the influence of the presence of grammar schools on the overall performance standard of the local area. This compares the effectiveness of selective LAs with non-selective areas, thus answering the third research question. The next part focuses on the pattern of HE participation associated with grammar school attendance. The rate of attending HE institution in general and the rate of attending the Russell Group universities are compared between grammar school pupils and equivalent pupils in non-selective schools. The evidence in this part answers the fourth research question. Finally, combining the access to and the outcomes of different pathways within the selective system, the study discusses the relationship between the selective system and the distribution of later academic opportunities. This presents whether the link between pupil's family background and post-18 destination is stronger in selective LAs than in comprehensive LAs, answering the final research question.

After presenting detailed statistical analysis based on different models, the last chapter summarises possible impacts of the presence of grammar schools on effectiveness and equity. Implications for policy and future research, as well as limitations of the present study, are also discussed in the conclusion.

2 Using academic selection to control educational opportunities

While this study evaluates the effectiveness and equity of one specific type of academically selective school in England, this section first presents the background of the general practice of academic selection. It discusses academic selection at different education stages, and explains why academic selection at the compulsory education stage is a sensitive issue.

2.1 The widespread practice of academic selection

The nature of selective education is that access to some educational institutions is neither random nor universal, but is based on specific selection criteria. For example, in the long-established practice of academic selection, young persons are selected based on their academic ability, which then decides the provision of different routes at later stages. One of the most common practices of academic selection is apparent in HE. In most countries, access to HE is not universal. HE institutions select their prospective students based on strict requirements which they believe to be essential to future learning. The selection criteria may include academic qualifications indicating satisfactory performance at lower educational levels, as well as non-academic skills. International students are also required to demonstrate their language proficiency, typically evaluated through formal language tests such as the IELTS in the UK.

Unlike the selective nature of HE, academic selection is less common at the compulsory education stage. In some countries/areas where the secondary education system is selective, pupils are allocated to different pathways at young ages, some as low as 11. Under this system, pupils who perform well at the end of primary school are usually eligible to attend secondary schools with superior status. Those perceived as less-academically oriented usually attend general or vocational schools (Morris & Perry, 2017). The hierarchical status between different secondary education pathways qualifies their pupils for different destinations in the ensuing stage. Therefore, the allocation of secondary school places is linked to post-compulsory destinations, and is the key to later job opportunities (Cullinane et al., 2017).

2.2 The sensitive issue of selection at the compulsory education stage

While it is widely accepted that HE institutions can select their students, the practice of setting a selective admission criterion during the compulsory stage is controversial. This is primarily due to the different missions of HE and compulsory education. As Walzer (1984) mentioned, HE is geared towards developing professional capabilities in particular fields and preparing

young people for specific occupations. The nature of HE is focused on differentiating pupil groups. Compulsory education, on the contrary, is geared toward the basic knowledge and skills necessary for participation in society. Thus, the need for high-quality compulsory education is universal. This reveals the importance of carefully evaluating academic section in secondary education, as it may contradict the original purposes of compulsory education. Therefore, the main focus of this study is the consequence of academic selection at the compulsory education stage, evaluating the specific situation of grammar schools in England.

3 Secondary education in today's society

After demonstrating the sensitive role of academic selection, this chapter discusses the current status of secondary education and the global trend of establishing education 'quasi-markets'. The following sections present the assumed benefit of this quasi-market system, concerns over its real impacts, and its relationship with academic selection.

3.1 The quasi-market in education and its assumed benefits

Since World War II, secondary education in many countries has transitioned into a differentiated and segmented system. In England, the tendency can be witnessed after the Education Reform Act of 1988 and the 1993 Education Act. The reform in England has emphasised open enrolment and school autonomy, which also included a new school funding formula based on the size of each school's student body (Power et al., 2003). According to the new funding formula, schools struggling to fill their places lose some of their state funding. Meanwhile, schools which are popular among parents receive extra funding. The market principles in the education system thus encourage schools to be independent, autonomous, and responsible for their provided services, behaving like business groups. Meanwhile, parents are also asked to act like 'consumers' in choosing the right school for their children (Power & Frandji, 2010). These changes are usually summarised as a politically-regulated quasi-market in education (Power, Halpin & Whitty, 1997; Le Grand, 1991).

The idea of a quasi-market in education is based on the assumption that the welfare state is neither effective nor equitable. It is believed that bureaucratic redundancy wastes resources, increases maintenance costs, and ignores the needs of the public, especially those of disadvantaged groups (Le Grand & Bartlett, 1993). The quasi-market is expected to solve these problems. According to the ideal quasi-market, different types of schools suit the unique academic, religious, physical, or other personal needs of each group, which can enhance overall education standards (Kitchener, 2013). As the market system also intends to offer parents the freedom to make responsible and rational decisions for their children when choosing schools, pupils caught up in poor schools due to their catchment areas would have the right to attend schools elsewhere. It is argued that this system breaks the cycle of residential and school segregation, which in turn promotes social equity (Parsons, Chalkley & Jones, 2000). Under this system, both parents and schools are supposed to make decisions which maximise their

own advantages, and thus the process promotes equitable state development (Robertson & Lauder, 2001).

3.2 The quasi-market in education and concerns over its real impacts

3.2.1 Quasi-market effectiveness

In reality, the quasi-market system in education may not function as well as its idealised image. This is primarily due to the divergent natures of educational and conventional markets. From the provider side, the aim of schools is not to maximise their 'profits', which may make the rewards and penalties ineffective (Levin, 1992). Unlike conventional products, school expansion is subject to inevitable physical and practical constraints, despite the amount of funding. Similarly, schools rarely close, even if they are providing poor education. This is especially obvious if the number of pupils exceeds local provisions. Additionally, contrary to the assumption that the quasi-market can respond to the specific needs of different groups, changes in the education system are usually driven by politics rather than the needs of 'customers'. In other words, the system is top-down instead of bottom-up. For example, some new school types have been established in England, such as Academies, Free Schools, and Specialist Schools. However, they have been primarily the result of government strategies, rather than being reflections of parental demands. For consumers, the quasi-market in education is also different from a real market. While parents can name their preferred schools, parental preference is limited by provision and availability (Whitty, Halpin & Power, 1998). This is especially obvious in popular oversubscribed schools, where schools, rather than parents, are making the choices, adopting overt and covert forms of selection. Therefore, the quasi-market system may not necessarily be better than a state-controlled system at responding to pupils' needs.

These concerns are underscored by previous research which has failed to find robust evidence that the quasi-market system in education raises performance standards in either European countries or elsewhere (Arnott, Bullock & Thomas, 1992; Bullock & Thomas, 1994; Power, Fitz & Halpin, 1994; Whitty, Halpin & Power, 1998). On the contrary, it has been noticed that unnecessary competition between schools discourages collaboration, and the importance of attracting more parents misleads schools into emphasis on physical facilities and public image, at the expense of core issues such as pedagogical improvement (Power, Halpin & Whitty, 1997;

Wylie, 1994, 1995). Therefore, the perceived efficiency gains of the quasi-market may be unrealistic.

3.2.2 Quasi-market equity

3.2.2.1 Concerns over quasi-market inequality in education

In addition to concerns over the effectiveness of the educational market, many researchers also believe that the quasi-market system in education skews distribution between the rich and the poor. While the process of school choice may help some parents find more suitable schools, it likely discriminates against pupils from disadvantaged families (Gewirtz, Ball & Bowe, 1995; Wylie, 1994). Thus, the system may increase stratification and reinforce social inequality (Le Grand, 1991).

3.2.2.1.1 Parents' differentiated school choice patterns

Due to differentiation in education provision under the quasi-market, parental choice is an important factor in determining pupils' educational trajectories. This may strengthen the link between family background and later outcomes (Marks, 2005; Horn, 2009). Therefore, the assumption that the quasi-market allows freer school choice may be over-optimistic.

School choice in the educational market involves a complex process. A wise school choice can only be made on the basis of the willingness, adequate resources and skills to access information. However, the distribution of these factors is uneven across social groups. Since not all parents are willing or able to choose, the choice per se can become a process of separation (Echols & Willms, 1995; Siddiqui, 2017). Evidence from different countries adopting the market system reveals a similar trend that parents from higher social classes are more likely to execute their rights to choose than poorer and less-educated families are (Moore & Davenport, 1990). Additionally, although some parents from disadvantaged families do recognise the importance of high-quality education for their children, they are constrained by available resources (Tomlinson, 1997, 2005). For example, a crucial factor is school distance. The understanding of an acceptable home-school distance is usually influenced by family organisation, social networks, and patterns of family activities, each of which are conflated with class (Harvey, 2018). The financial and time costs of attending distant schools can be heavy for working-class families. The greater parental preference for local schools among certain social groups means that housing segregation may be transferred into school

composition segregation, which is similar to the consequence of catchment areas (Jacobs, 2013).

Parental free choice is based on some strict premises such as equal decision-making resources, the affordability of attending schools at a distance, and equal accessibility to favourable school places across social groups (Windle, 2009). However, around the world, it is the rare case in which each of these assumptions is applicable. Therefore, the quasi-market system may still enable middle-class families to retain their social advantages, thus fostering 'reproduction' (Bourdieu & Boltanski, 1979, p. 198). Since how parents make choices is related to their social backgrounds, the system may be no more equitable than its predecessor.

3.2.2.1.2 The quasi-market and school stratification

In addition to the differentiated school choice patterns among parents, researchers have also noted that the quasi-market system, which encourages competition and accountability, also promotes stratification between schools (Gewirtz, Ball & Bowe, 1995; Wylie, 1994). Schools which attract more middle-class pupils and retain their elite images are usually more popular among parents. They are also more likely to be oversubscribed. Since the funding formula depends on the number of pupils at each school, these schools benefit financially, and are thus better positioned to recruit high-qualified staff and update their facilities. However, some schools are hurt by the circle of parental perception, school image and funding. This has been descried as the 'spiral of decline' (Croxford & Raffe, 2007). Due to the correlation between pupils' backgrounds and later performance, disadvantaged schools tend to have inferior results in league tables, and thus become less attractive to parents. Struggling to fill their places, these schools are likely to have insufficient budgets, which inhibits improvement. Thus, some researchers believe that the quasi-market system in education deepens school stratification (Whitty, Halpin & Power, 1998).

3.2.2.2 The quasi-market's limited role in equity

Although concerns over the quasi-market's inequitable consequences are substantial, other studies have noted that these negative impacts have been exaggerated, and that the actual role of the educational market in social equity is similar to that of other state-controlled systems (Gorard, 1997; Herbert, 2000; Levin & Riffel, 1997). Focusing on pupils in England and Wales, Glatter et al. (1997) found no increase in stratification, and there might even be a small positive

pattern associated with the marketisation reform. Applying the English national data from 1989 to 1999, a later study by Gorard and his colleagues confirmed that there is little evidence of intensified stratification between schools. Furthermore, the existence of failing schools cannot be causally attributed to the market system (Gorard, Taylor & Fitz, 2002). Unlike opinions either for or against the role of the market in stratification, they also concluded that the quasi-market system has changed the education system little in England. This is consistent with findings from the US (Gorard, Taylor & Fitz, 2002; McGuinn & Hess, 2000). The slight increase or decrease in segregation rates after the implementation of the quasi-market system may only be a starting-gun effect—over time, the situation could revert to that of years past (Cookson, 1994; Gorard & Fitz, 1998).

To conclude, while there is considerable support for the superiority of a quasi-market, both in terms of effectiveness and equity, there are also substantial concerns that this fresh idea in education may be no better than the previous state-controlled system. Despite debates over the real impacts of the quasi-market, education systems in many countries have already become more autonomous and independent, following market principles. These changes influence educational selection substantially, which are discussed below.

3.3 The quasi-market, deepened awareness of school performance, and intensified selection

Although the quasi-market's real impact is debatable, its underlying principles, such as choice and accountability, have been more widely accepted. These principles have drawn public attention to school performance, and the publication of league tables is carrying more weight than ever before. As the primary (and sometimes the only) indicator of school quality, test scores have also been attached with more importance (although performance figures themselves are problematic, as will be discussed in Chapter 5) (Whitty, Halpin & Power, 1998). Within the market system, parents, teachers and the government have each become sensitive to school performance and school rankings, and awareness of the differences in school effectiveness has intensified (Power & Frandji, 2010).

The uni-dimensional indicator of school effectiveness and the emphasis on test performance encourage schools to overtly and covertly select their pupils. In order to rank higher in league tables, schools target pupils with adequate family support, and avoid those with greater needs.

This is due to the logic of the market which implies that schools are rewarded for their high performance, not their efforts to help the disadvantaged (Gewirtz, Ball & Bowe, 1995; Moore & Davenport, 1990). Some schools have even resorted to exclusionary practices such as complicating their enrolment processes and adding socially selective specialist curricula to select out those with limited family resources (Walford, 2001). In addition to schools intentionally excluding certain pupil groups, overt and covert selection could engender self-selection and self-elimination among the poor, who may regard popular and high-ranking schools as the choices for more advantaged groups (Ball, 2002).

Overall, the trend of marketisation has brought substantial changes to the education system around the world. While the quasi-market's real impact remains to be seen, its underlying principles have become more popular in many countries. Following its emphasis on choice, competition, autonomy and accountability, this new system may have the potential to strengthen public awareness of school effectiveness, and encourage different forms of educational selection.

4 Different forms of selective schools

After reviewing current changes to secondary education systems around the world and their relationship with educational selection, this chapter presents examples of different types of selective schools. It starts with the situation of single-sex schools and religious schools, which select by pupils' gender and religious beliefs respectively. After discussing the traits of these two school types, this chapter turns to schools selecting by aptitude and ability. It also presents the underlying reasons behind establishing these schools and concerns over their real impacts.

4.1 Schools selecting by gender and religious belief

Single-sex schools are common in many societies' education systems. A popular claim in support of single-sex schools is that based on the differentiated learning and behaving patterns of boys and girls, teachers can better accommodate the needs of each gender if boys and girls are separate. Along with this perception, there is evidence that girls in single-sex schools are more confident, and more willing to participate in classroom discussion (Jackson, 2002). It has been reported that boys in single-sex schools benefit as well, because teachers are able to reshape curricula to suit their learning pace (Arnot & Gubb, 2001). Furthermore, single-sex schools are also regarded as effective in countering traditional gender stereotypes in subject choice and future aspirations. However, opposing voices point out the negative influence of single-sex schools. For example, separated into different schools at an early age, young people lose the opportunity to interact with the opposite sex. Meanwhile, there are also concerns over the clustering of advantaged pupils in single-sex schools. Single-sex schools (especially girls' schools) are often over-subscribed because of their high test scores. This renders entry application to these schools more difficult, which may exclude certain social groups who are less likely to devote large amounts of time to school choice. The limited availability of single-sex schools within local areas also creates problems such as increasing home-school distance's salience in stratifying social groups (Spielhofer, Benton & Schagen, 2004).

Similar to the assumption that single-sex schools respond to the different needs of boys and girls, religious schools are also assumed to address the academic and social needs of pupils with certain religious beliefs. In England and Wales, over one third of primary schools and about a quarter of secondary schools are defined by the government as religious schools (DfEE, 2000). Even in countries with high degrees of secularisation, such as the Netherlands, a substantial proportion of pupils attend religious schools (Hofman & Hofman, 2001). Beyond

their distinct religious ethos, religious schools have been known to outperform secular schools academically in some countries. For example, Altonji, Elder and Taber (2005) noted that attending religious schools is positively correlated with a higher rate of high school completion and HE participation in the US. Higher exam scores at certain religious schools have also been noticed in European countries, which makes these schools attractive to pupils from secularised families as well (Hofman & Hofman, 2001; Morris, 2001). However, it has also been revealed that the segregation index tends to be higher in religious schools, partly due to links between faith and ethnicity. Apart from ethnicity, a growing number of religious schools are becoming more segregated in other aspects as well. While many religious schools are following their original purposes to help the poor, some other have become more selective. For example, Allen and West (2009) mentioned that some religious schools use their faith-based admission criteria to select more advantaged families and select out others, such as setting strict, and sometimes unrealistic, requirements on church attendance.

Claims in support of these two types of schools are based on the perception that they can accommodate certain needs of specific pupil groups better than co-educational schools. However, whether the real situations of these schools correspond to their ideals remains contentious.

4.2 School selecting by aptitude and ability

Apart from schools selecting by pupils' physical and social characteristics, some schools place more emphasis on pupil performance in their selection criteria. This is based on the same assumption that pupils benefit from the experience of being educated along with those who have similar academic needs (Reichelt, Collischon & Eberl, 2019). Thus, this section turns attention to schools which select pupils based on aptitude or ability.

The specialist schools programme in England, which has been in place since 1994, is an example of promoting school selection based on aptitude (Coldron, Willis & Wolstenholme, 2009). Under this act, comprehensive schools have been encouraged to specialise in specific areas such as sports, modern foreign languages, art or technology, and select up to 10% of their intakes based on these subjects (Penney, 2004). The programme was seen by policy makers as a means to raising academic standards, improving diversity and increasing school options for parents (Gorard & Taylor, 2001). However, while there is some evidence of better academic

results in specialist schools, it is difficult to distinguish the real school effects from the influence of government funding allocation (Coldron, Willis & Wolstenholme, 2009). At the same time, the issue of creaming advantaged pupils is also salient in schools that select by aptitude in specific subjects, as aptitude in sports or art is connected with SES factors such as family resources and parental involvement. Furthermore, due to its similarity to ability selection, the selection by aptitude in certain subjects is sometime misused as selection by general academic ability, which is more closely interrelated with pupils' backgrounds.

Schools selecting directly on general academic ability are believed to offer an even stronger form of selection (Stringer, 2008). Access to these academically selective schools is normally based on ability tests. In areas with a state-wide academic selection process after primary school, such as Northern Ireland (except for one area), pupils need to pass a selection test to attend more academically-oriented schools, which are the primary path to post-compulsory education and HE institutions. Those who did poorly on the test are allocated to the lower track, which usually cannot offer upper secondary education after the school-leaving age (Finch, McCreight & McAleavy, 2010). Although the example of state-wide early-age academic selection in Northern Ireland is not a common practice in European countries, the practice of ability selection is still prevalent. For example, there is a similar system of early-age academic selection in countries such as Hungary and Austria (Horn, 2009). In England, while the Bipartite System of dividing pupils into higher and lower tracks at 11 years-old has been reformed since the 1960s, there are still 163 grammar schools whose intakes are based solely on academic selection (Jesson, 2013; Morris & Perry, 2017). In Germany, pupils are divided into four tracks at the age of 10, with the stark divisions between each educational pathway (Hanushek & Wößmann, 2006). Even though the ability test has been replaced by a choice system, the underlying principle of academic selection remains unchanged. This is because parental choice is still dominated by head teacher evaluations, which act as an indirect form of academic assessment (Dustmann, 2004). The system of early-age academic selection is also a common practice in some Asian countries, where public education resources are more limited and unbalanced. In China, state-wide academic selection is held at the transitional stage between junior and senior high school for 15-year-old pupils. Although pupils (and parents) can name their preferred schools before or after the test, they are only eligible to attend their chosen school if they pass its selection threshold (OCED, 2016). Besides the overt between-school academic selection, it is also believed that many comprehensive schools are also academically selective, adopting covert academic selection within schools, such as the tracked

curricula (Coe et al., 2008). To conclude, despite variation in selection criteria from country to country, academic selection remains a common practice around the world.

4.3 Problems with selection by ability

4.3.1 Difficulty in defining ability

Unlike selection criteria such as religion and gender, a substantial problem of academic selection is how to define 'ability'. The UK government provides some official definitions of 'ability' in an attempt to distinguish it from 'aptitude'. They believe that the former is a presentation of a pupil's current capacity, and the latter evaluates an individual's future potential (House of Common, 2004). However, many researchers still find it hard to differentiate these two terms, as 'ability' is not only used to refer to the current capacity, but also the capacity to learn 'in the future' under certain circumstances (Kamin, 1981, p. 94). This has been thoroughly discussed by Barber and McCallum (1996), yet in reality, these two terms are still used and implemented with wide variation (West & Hind, 2003).

4.3.2 Difficulty in measuring ability

In addition to theoretical vagueness, there are also practical challenges to the measurement of ability. While a pupil's test score is usually used as the indicator of academic ability, this perception is resisted by some test designers who argue that general academic examinations, such as GCSE, only evaluate 'attainment', rather than the controversial 'ability' (Bourne & Moon, 1995; Gillborn & Youdell, 2000). A similar understanding of the nature of academic exams can be found among critics of academic selection, some of whom argue that academic tests only focus on whether pupils are able to answer particular types of questions based on memorised knowledge, yet pay little attention to creativity, problem solving, critical thinking, and other cognitive skills. In other words, these tests merely reflect how familiar pupils are with the test format and content; they are not relevant to 'ability' under most circumstances (ILEA, 1995). Furthermore, in addition to being poor indicators of 'ability', academic selection tests have also been criticised as presenting an inaccurate level of 'attainment'. This is because a short one-off exam is too simple to reflect pupils' learning experience over several years, especially when the format of the test is limited to written tasks (Brown, 1995). Furthermore, test result accuracy also depends on whether the test is designed for high validity and reliability. Evidence has shown that even in high-stakes tests, misallocation rates may be high. For example, in Northern Ireland, it is believed that every year a group of pupils is misallocated

into their current tracks (Gardner & Cowan, 2000). Similarly, the selection test for grammar schools in England shows a weak correlation with pupils' KS2 attainment—a national academic test taken several months later—and the average correlation figure is only around 0.6 (Allen, Bartley& Nye, 2017). Therefore, the issues of the validity and reliability of selection tests might be prevalent in academic selection. Additional concerns over high-stakes selection tests also derive from the fact that these tests may be culturally-biased. This would favour dominant groups, such as white pupils from affluent families (Coldron, Willis & Wolstenholme, 2009). In this way, the selection results for pupils from disadvantaged backgrounds might be underestimations of their real performance (Mortimore, 1997).

Another practical problem in assessing ability is that individual ability changes over time. Legitimising the value of any ability test by regarding ability as a fixed and single entity is problematic, as it implies that educational pathways should be decided based on generic standards. In fact, empirical evidence has demonstrated that the degree of consistency between pupils' entrance tests and their later scores in a given subject is low (Coffey & Whetton, 1996). This implies that any ability test which determines the eligibility of later educational pathways may become invalid after pupils actually enrol in these schools.

Even if the assumption that academic ability can be predicted with precision is met, plenty of unanswered questions remain. Standardised tests focusing on academic ability neglect social, psychological and other non-academic abilities, such as teamwork, communication skills, resilience, self-disciplinary, and a sense of responsibility. These are all vital to young people's future success and well-being. It has also been noted that academic and non-academic abilities are loosely correlated, revealing the possibility that not all people with high academic performance have adequate non-academic abilities (Sternberg et al., 1995). Therefore, overly-simplistic selection by academic ability may be inconsistent with the needs of society (Carneiro & Heckman, 2003). Based on the limitations of academic tests, researchers have suggested that assessment at an early age should be used in a formative rather than summative way. The assessment is more fruitful when used to gather information about young pupils' learning processes, rather than as a prediction of their future achievement (Brown, 1995). Thus, the entire exam system needs to move away from 'surveillance', which tracks the success of a minority, while restricting the rest.

4.3.3 Problems with early-age academic selection

Besides general concerns on academic selection, there are also specific problems with early-age academic selection. The application of academic tests at early ages is believed to yield even less reliable results than at later ages. Judson (1998) mentioned that the selection noise is higher among younger people who have received fewer years of education, as it is hard to differentiate between them. The longer pupils have learnt, the less noise remains in their test results. This means an exam conducted after a pupil finishes secondary education would reduce the error in the test score, thus yielding a more accurate evaluation of the learning outcome than an exam during primary school. Moreover, assessing the ability of primary school age pupils presents another issue—it is almost impossible to control for prior learning outcomes. This renders the calculation of progress within any given time period impossible (Brunello, Giannini & Ariga, 2007; Coldron, Willis & Wolstenholme, 2009).

In addition to the selection test noise, there is also evidence of a stronger connection between the result of early-age academic selection and family background. Empirical evidence has shown that parental influence is of varying importance at different educational stages—typically stronger earlier on (Feinstein, 2003). It has been noted that parental engagement is crucial for primary school pupils, both in terms of their academic performance, and their psychosocial well-being (Dustmann, 2004). Although most parents do not consciously reduce their level of involvement as children grow up, the significance of the family drops while the impact of the wider community increases. For 16-year-old pupils, the influence of social contexts, such as the SES of school composition, surpasses the impact of parental involvement (Feinstein, 2003). Therefore, although low social status and poverty continue to impose long-term negative effects on children's achievement, early-age selection is likely to exacerbate the influence of family background on educational trajectories. Since early-age choices qualify pupils for different educational pathways and occupations in the future, age of selection is critical to the life chances of different social groups, and thus influential to the landscape of society as a whole (Sacker, Schoon & Bartley, 2002).

In sum, the vagueness of the definition of selection by ability, and the practical difficulty in assessing ability, raise concerns over the fairness and accuracy of the selection process for academically selective schools. Apart from the controversial selection criteria, disputes over academically selective schools also result from their link with effectiveness and equity. Thus, the following chapters address these two core concepts.

5 Effectiveness and equity in education

This chapter discusses two core concepts of this study, which are effectiveness and equity. It starts with an introduction to educational effectiveness research and the reasons why evaluating academic outcome of schools has become a focus of this field. It also discusses the role of schools, peer groups, and academic selection in pupils' academic performance. The discussion then reveals the difficulty of assessing school effectiveness, as well as attempts to overcome this in previous research. After introducing school effectiveness, this chapter also presents the discussion of equity, including the complicated definitions of equity in education, and the connection between equity and academic selection. The last part of this chapter then turns to the relationship between effectiveness and equity.

5.1 School effectiveness: Are some schools more effective than others? Why?

5.1.1 Educational effectiveness research

The role of schooling in a modern society has long been emphasised. It is believed that education may be the answer to many societal challenges, such as economic productivity, democracy, social participation and cohesion (Dewey, 1916; Haveman & Wolfe, 1984). Meanwhile, education also benefits the educated individual. There are academic outcomes to schooling such as knowledge and skills, which ultimately lead to better occupational opportunities, higher incomes and privileged social status. Schooling also has psychological and social aims, which are not directly reflected in test scores and qualifications, such as passion for new things, receptive attitudes towards others, emotional adjustment, and positive social links (Finch, McCreight & McAleavy, 2010).

Despite the range of possible educational outcomes, academic achievement remains the predominant criterion in educational effectiveness research (such as in this study). This is not surprising based on the necessity of learning, both for individual's societal participation, and for the development of society as a whole. However, in the current system, the tendency is also tied to over-emphasised high-stakes standardised exams. The conflation of qualification and the narrowly-defined 'ability' means that individuals who fail to attain satisfactory scores on public exams are hindered in their future chances. Test scores are not only critical to individual pupils, but also to schools. Performance on standardised tests has become the most powerful indicator of school quality for all kinds of league tables. The performance figures are then used by parents as the basis of making school choices for their children, determining the position of

each school in the market (Au, 2009). Therefore, the importance of academic results, for both individual pupils and schools, largely explains the predominance of academic attainment in school effectiveness research (Gillborn & Youdell, 2000).

5.1.2 The role of school in pupils' academic performance

Following the importance of school effectiveness, researchers have made several conclusions on the role of school in pupils' academic progress. The first wave of school effectiveness studies began with influential reports by Coleman et al. (1966) and Jencks et al. (1972). Contrary to traditional perceptions of the role of schools in compensating for family disadvantage, both reports pointed out the weak connection between schools and pupils' life destinations, suggesting the unimportance of schools. Questioning the role of schools in improving academic performance, they believed that many other social factors, such as family background, are more powerful than schools (Coleman et al., 1966; Jencks et al., 1972). Since the potential effect of schools is limited, any intervention at school is less significant than the advantages of being brought up in an affluent and supportive family (Jencks et al., 1972; Mortimer & Whitty, 2000). The situation is similar in most industrialised states, and has not changed, even under educational system expansion. Thus, Jencks et al. (1972) criticised the reliance on education for social reform as utopian. Additionally, they predicted that if all schools were equally effective, the variation in pupil's attainment would be erased by no more than 1% (Rutter, 1979). Thus, it was concluded that the school does not hold a critical position, either in ameliorating or in intensifying, inequality (Anderson, 1961).

This pessimistic perception of school effectiveness has been resisted by other researchers such as Rutter et al. (1979), Smith and Tomlinson (1989), and Mortimore et al. (1988), who have found consistent school effects and believe that some schools foster academic achievement better than others. However, a typical finding is that schools only account for a small proportion of between-pupil difference in academic performance, normally between 10%-20% (Thomas & Mortimore, 1995). The proportion is higher for subjects with less exposure in the family, such as mathematics, but lower for subjects which are more frequently encountered outside school such as languages (Teddlie & Reynolds, 2000). Furthermore, the variance accounted for by school may drop to around 8% if any progressive scores such as the value-added (VA) are calculated (Gorard & Smith, 2010).

Despite the low variation in attainment attributable to school, many researchers believe that even a small school effect makes real differences, especially for disadvantaged pupils. For example, Mortimore (1997) explained how a 10% variation in academic performance can be translated into the distance between seven GCSE subjects at grades C and E. While the former represents a good pass in each of the tested subjects, thus revealing the suitability of pursuing the post compulsory education, the latter is a signal of inadequate academic performance. Therefore, the low difference in attainment attributed to schools can still lead to a substantial difference in pupils' later education pathways. This is especially meaningful in the current educational system, as test scores and qualifications are critical to future opportunities. Another meaningful aspect of schooling is the long time period it covers and the large number of pupils it involves (Reynolds et al., 2014). This means that any school effect will eventually add up to a cumulative result for the entire population.

5.1.3 The role of peers: The school compositional effect

Apart from the discussion of whether school matters, researchers have also paid attention to what makes some schools better than others, trying to elucidate efficient inputs of school quality (Vigdor & Nechyba, 2007). Some research has attributed school effectiveness to the composition effect, which is the aggregated influence of pupil's social background within each school (Hattie, 2009). The existence of such an effect has been systematically presented since Coleman's (1966) report, who concluded that the quality of the peer group is one of the most powerful predictors of pupil's later performance. Jackson (2013) also noticed that the size of the peer effect accounts for 7% to 14% of the school effect, and it is more obvious among schools with high selectivity (Jackson, 2013). More detailed research focusing on each background characteristic criterion has also been conducted, with evidence showing the positive composition effect of having a better average attainment, more affluent families, and a higher proportion of girls (e.g. Caldas & Bankston, 1999; Kang, Park & Lee, 2007; Van Houtte, 2004). However, most research claiming to find a causal relationship between peer group composition and academic achievement has been criticised as inconclusive, due to methodological limitations and data quality, most of which have been cross-sectional in nature with low specification at the individual level (Gorard, 2006; Nash, 2003). Vigdor and Nechyba (2007) noticed that the potential impact of the peer group exists prior to a pupil's actual contact with their peers. Considering the possible influence of within-school sorting, they mentioned that the perceived peer group effect may be the result of parental intervention and access to

more qualified teachers (Vigdor & Nechyba, 2007). Once the teacher-level effect was controlled for, the peer effect turned negative. Overall, as Nash (2003) and Reynolds et al. (2014) have concluded, despite plausible reasons to believe in the existence of the composition effect, it is not supported by robust empirical evidence.

In addition to focusing on the impact of the peer group on academic performance, researchers have also explored other potential influences it may have. For example, it has been mentioned that there is a positive connection between an advantaged peer group and pupil's higher aspirations (Gorard & Smith, 2010). Pupils in schools that serve those with better academic abilities and those from middle-class families are more likely to stay in school after the compulsory stage. Poor pupils also tend to have higher occupational aspirations, and this transcends the influence of their own family backgrounds (Gorard & Smith, 2010). Along with the positive effects of a more advantaged peer group, there are potential downsides as well. Marsh and Hau (2003) noticed how a high-performing peer group has negative influences on pupils' academic self-concept. Based on his theory of the Big-Fish-Little-Pond effect, Marsh (1984) believed that pupils evaluate their own achievement in comparison with the academic performance of their peer group. Pupil's academic self-concept tends to be lower in highly selective schools than in mixed-ability schools with lower average attainment (Marsh & Hau, 2003). Thus, while the importance of peer groups on academic performance remains controversial, the composition of the peer group might have other long-term effects.

5.1.4 The role of academic selection in pupils' academic performance

Apart from focusing on the composition of peer groups, which is usually beyond schools' control, researchers have also searched for more direct factors which may improve school effectiveness, such as input resources (e.g. funding and teacher salaries), school organisation, and teaching strategies (David, Teddlie & Reynolds, 2000; Rutter, 1983). However, the conclusions on the real impact of these interventions have also been met with controversy. In addition to the above-mentioned elements, some also believe that selective schools are better than other types of schools at fostering academic performance, revealing the possibility of a 'selective school effect'. Supporters of selective education believe that differentiation and competition raise overall academic standards (Mickelson, Nkomo & Wimberly, 2012). Using a suitable curriculum and teaching style, pupils progressing at different speeds can learn more effectively in a homogeneous environment where the specific needs of each group are better

accommodated than in a mixed-ability class (Van de Werfhorst & Mijs, 2010). It is believed that in a selective system, there is no need to worry about a ceiling on the most able, or about losing the low-achieving group. This maximises the benefits of both groups (Hanushek & Wößmann, 2006). There are also arguments showing the positive connection between a stratified system and smoother transition from school to work, revealing how selective schools may better prepare pupils for future destinations (Breen, 2005). In addition to tailoring teaching, the possible benefits of the selective system may also result from the emphasis on competition. The chance to attend more favourable schools can be an attractive and useful reward, which encourages pupils' academic performance (Van de Werfhorst & Mijs, 2010). Evaluation of the education systems in Asian countries has revealed the importance of high-stakes exit exams, which are not only reflective of pupils' previous learning experience, but also influence access to ensuing educational levels. Being competitive, these tests are believed to create higher academic aspirations and improve academic performance (Peterson & Woessmann, 2007).

Apart from the possible benefits of academic selection, there is also evidence of its negative influences. For example, the emphasis on competition in an academically selective system leads to subsequent problems such as teaching to the test, narrow curricula and stress-related issues. Some secondary school teachers in Northern Ireland have complained about the inadequate academic abilities of their pupils, as they had been taught to pass the selection test without enough attention to essential learning skills (Department of Education, 2000). Meanwhile, it has also been demonstrated how assessment at school transitional stages is related to stress. Stories in some selective single-sex girls' schools have uncovered the stressful status of young girls who have suffered from the intensive preparation for high-stakes tests and the underlying ideology of competition rooted in school culture (Allan, 2010; Howard, 2013). A similar pattern has been noted by Kiselica et al. (1994), revealing the correlation between the long-term exposure to challenging academic tasks and problems in well-being. Meanwhile, although the overall impact of the selective and differentiated system on pupil's performance is less clear, the evidence of their negative impact on disadvantaged pupils is stronger. Unlikely to earn school places in higher tracks, these pupils are more likely to be allocated to schools which have less experienced teachers, limited teaching resources, less financial support to maintain facilities, and thus less favourable learning environments (Harris & Williams, 2012; Kalogrides & Loeb, 2013; Massey & Fischer, 2006). A high concentration of disadvantaged pupils in certain types of schools also makes a school's operations more challenging. Moreover, any positive influence of having more advantaged peer groups may further disadvantage pupils

in lower tracks. Empirical evidence has shown that pupils in lower tracks in the selective system are more vulnerable to low aspirations and high drop-out rates, even if they outperform equivalent pupils in comprehensive settings (Gorard & Smith, 2010).

While there are several possible reasons why an academically selective system may be more effective in raising pupil performance, previous research has found only limited evidence supporting this statement. The strongest conclusion in favour of the effectiveness of selective systems may be that of Duflo, Dupas and Kremer (2011). They revealed more progress among both high-performing and low-performing pupils in a tracked system in Kenya. Most other studies have found either no selective effect, or even a negative effect. For example, an international study applying PISA data found no correlation between a selective system and overall academic performance (Van de Werfhorst & Mijs, 2010). Furthermore, a negative consequence of the tracked system on overall performance standards was also noted—pupils at both ends of the performance distribution suffer from tracking (Condron, 2013; Hanushek & Wößmann, 2006). To conclude, despite plausible claims of the benefits of academic selection in raising academic standards, there is no adequate empirical evidence supporting these claims.

5.1.5 Evaluating school effectiveness in academic achievement

5.1.5.1 The value-added approach: Controlling for pre-existing differences between pupils

Assessing school effectiveness in terms of academic performance is not easy. Schools have different intakes, and pupil characterises influence later learning outcomes. Although some schools usually have better test scores than others, this may be primarily a result of their advantaged intakes (Gorard & See, 2013). The unfairness of using raw test scores as the indicator of school effectiveness has become a consensus among researchers. In contrast, the growth in performance is believed as a more accurate indicator of the impact of school, which is gradually accepted as the essential criterion of school effectiveness (Teddlie & Reynolds, 2000). Following this perception, the most influential innovation has been the VA approach (Goldstein, 2001). Controlling for prior attainment, VA calculates pupils' relative progress within a fixed duration and compares like with like (Leckie & Goldstein, 2017). Following this principle came a more complex approach which takes other pupil-level characteristics into consideration. Even for pupils with equivalent prior attainment, pupils from more supportive families are expected to progress more than the disadvantaged. VA models which also control

for pupil's contextual backgrounds are thus believed to capture the influence of factors such as SES, gender and ethnicity, as these factors are also influential to pupils' later performance but are beyond schools' control (Perry, 2016). This is regarded as isolating the net effect of schools more thoroughly than models only accounting for prior attainment, alleviating the bias of unfairly assessing schools that serve disadvantaged pupils (Burgess & Thomson, 2013).

In addition to decisions in choosing pupil-level baseline variables, the choice of whether controlling for school-level variables or not raises more complicated discussion. Previous researchers have tried to distinguish two types of school effects, and each of them requires for a different set of baseline variables (Raudenbush & Willms, 1995). The Type A school effect is the estimate of school effectiveness which only controlled for pupil-level background variables. The estimated school effect is thus a combination of school's contextual effect as well as any positive policies and practices in schools (e.g. high teaching quality). Therefore, the estimation is usually relevant to parents, as schools with larger Type A effects will help their children to perform better. However, the Type A effect is unhelpful in identifying beneficial school practices and improving school effectiveness, as any benefit associated with advantageous school composition cannot be replicated in less advantaged schools. Furthermore, schools being rewarded for something they are not responsible for is also unfair. This promotes the idea of the Type B school effect. Further controlling for school compositional variables is believed to accounts for the compositional effect of schools. The estimation thus reveals the effect of school policies and practices but excludes compositional factors. The Type B school effect thus compares the performance of equivalent pupils in schools with similar compositions. The result is of the most interest to policy makers, as it reflects school actions which are genuinely beneficial to academic outcomes. However, the differences between these two types of school effects are not merely about including or excluding the controversial 'compositional effect'; both approaches also face validity issues. It has been pointed out that models that only control for pupil-level variables suffer from measurement errors on test scores. This may upwardly bias the effectiveness of more advantaged schools (Perry, 2019). Meanwhile, there may be pre-existing differences between pupil groups that are unaccounted for by surface pupil-level variables. This may be alleviated when school-level baseline variables are also controlled for (Coe et al., 2008). Despite claims in support of controlling for school-level baseline variables, concerns remain. It is possible that accounting for school-level prior attainment mitigates differences between schools, especially if beneficial school actions correlate with more advantaged intakes (Raudenbush & Willms, 1995). In this case, the

estimated effectiveness of more advantaged schools may be lower than their actual size, when both pupil-level and school-level baseline variables are controlled for.

The choice of baseline variables in the VA approach is difficult, and there is no perfect solution (Visscher, 2001). In addition to the difficulty in selecting baseline variables in the VA approach, there are further challenges to its validity despite its robust underlying logic (Harker & Tymms, 2004). First, the calculation is largely limited by the availability of background variables and the proportion of missing data. Even in high-quality databases such as the National Pupil Database (NPD), only around 70%-85% of pupils have complete data on attainment and contextual factors (Gorard & See, 2013). Secondly, the estimated VA scores are very unstable across time, and the correlation after 4 years may drop below 0.5 (Gorard, 2010; Leckie & Goldstein, 2009). Thirdly, even a very moderate rate of measurement error (e.g. 10%) in test scores would be accumulated to large errors, which could be 40 times larger than the estimation results (Gorard, 2010). Overall, the threats of the bias of unmeasured pre-existing differences between pupils, as well as the measured errors in baseline variables, are considerable (Perry, 2016). Regardless of threats to the validity of VA, this approach has been widely used (Leckie & Goldstein, 2017), primarily due to the lack of suitable alternatives.

5.1.5.2 The RDD approach: A strong design to make causal inference

As mentioned above, after conditioning on surface variables, pupils in different schools might still have distinct characteristics in unmeasurable aspects, which threatens the validity of any conclusion made. The most versatile solution to exclude pre-existing difference and evaluate the treatment effect is to conduct an RCT (Shadish, Cook & Campbell, 2002). However, in most school effectiveness research, pupils cannot be allocated randomly to different pathways to test the influence of school. Even when an RCT can be applied in certain educational settings, there are some difficult practical issues (Siddiqui, Gorard & See, 2018). The strongest alternative to the RCT is an RDD (Imbens & Lemieux, 2008; Lee & Lemieux, 2009). In an RDD, participants are allocated to either the treatment or the control group according to the cut-off point of a continuous assignment variable. Only those who reached the cut-off point are given the treatment, and those who missed the threshold are not. If participants' assignment variables could not be manipulated precisely, their chances of just making the threshold or just missing it can be regarded as locally random (Lee & Lemieux, 2009). Comparing participants in the neighbourhood of a cut-off point, the approach can attribute any discontinuity at the cut-

off point to the treatment effect. This process solves the problem of pre-existing differences between the treatment and control group, and provides perfect counterfactual results (Imbens & Lemieux, 2008). The advantage of the RDD makes it a feasible design not only to assess the absolute effect of attending schools but also the differences in effectiveness between schools (Gibbons, Machin, & Silva, 2013; Luyten, 2006; Luyten, Tymms, & Jones, 2009). However, as the RDD design requires a clear knowledge of the value in the assignment variable which can be hard to collect under certain circumstance (such as in this study), it has some limitations in terms of applicability.

5.1.5.3 The generalisability of the RDD

The generalisability of the RDD approach is a sensitive issue due to its special process of making the estimation. While the RDD approach is considered as a strong alternative to an RCT, the localised nature of this design makes its generalisability from the cut-off point to the whole data range a concern (Shadish, Cook, & Campbell, 2002). This is especially relevant to school effectiveness as the treatment effect may be heterogeneous at different data points. Previous research has revealed the differential school effectiveness for pupils with high and low prior attainment (e.g. Strand, 2010; Thomas et al., 1997). Similar results have been found in studies of grammar schools showing that grammar school attendance is more beneficial for pupils with lower attainment than for high-performing ones (Atkinson, Gregg, & McConnell, 2006; Levačić & Marsh, 2007). This means the treatment effect of grammar schools may be inconsistent across performance levels. Thus, we might expect the result at the cut-off point to be larger than at higher points.

Despite the potential limits of the RDD approach, researchers have also found evidence contrary to the pessimistic perception of the low generalisability of RDD (Bloom & Porter, 2012; Lee & Lemieux, 2009). According to Lee and Lemieux (2009), the discontinuity at the cut-off is a weighted average effect across all observations, and the weight calculates an individual's probability of being located near the cut-off point. Therefore, the estimate is relevant to all the observations, and the strength of relevance is largely determined by the rate of noise in the assignment variable. Larger errors in the assignment variable create a more heterogeneous pupil group near the cut-off point, increasing generalisability (Jacob et al., 2012). In extreme cases, if an assignment variable only contains random errors, then pupils would be allocated randomly, as in an RCT (Bloom & Porter, 2012). While the potential issues of low

reliability threaten the validity of the selection, the larger random error contained in the test score also means that pupils with different ability levels may score equally-well in the selection, and thus have a similar probability of being located near the cut-off point. In this case, the results at the cut-off point would be closer to the overall average treatment effect and more relevant to pupils at higher points.

To conclude, while previous research has reached some consensus on the possible size of school effectiveness, an accurate evaluation faces some difficulties. Therefore, the real impact of many elements which are believed to be influential to school effectiveness, such as academic selection, remains unsettled, revealing the need of further research. Apart from discussing school effectiveness and the role of academic selection, the following section turns to another major issue, which is social equity and its close connection with academic selection.

5.2 Education and social equity

5.2.1 Defining equity

Unlike the simpler definition of school effectiveness in raising pupils' academic performance, the definition of equity in education is more complicated. This section first presents the most commonly applied definition of equity in education, and then turns to two major categories of equity under different situations.

5.2.1.1 A general definition of equity in education

Equity usually refers to the distribution of important goods or conditions which are necessary to each individual (Deutsch, 1975, p. 137). According to UNESCO's definition, equity in education 'considers the social justice ramifications of education in relation to the fairness, justness and impartiality of its distribution at all levels or educational sub-sectors' (2018, p. 17). It is usually used as a synonym for 'fairness' or 'justice', and is the process of deciding whether and why something is (un)fair (Goarad & Smith, 2010, p. 65). Exploring what inequalities are justifiable, John Rawls' seminal work, *A Theory of Justice*, claims that a fair society should protect the equal liberty of each individual. Furthermore, inequalities can only be justified if they emerged under conditions of fair equality of opportunity, and if they somehow benefit the least-advantaged group in a society (Rawls, 1971).

Related to the perception of human capital, equity emphasises the utilitarian facet (Bentham, 1948; Rawls, 1971; Strike, 1979). Unlike 'equality', which usually requires the same treatment for all (Van de Werfhorst & Mijs, 2010), equity considers the practical circumstances of the individual in allocating educational resources. It attempts to overcome undesirable personal and home circumstances based on the perception that educational failure is not inevitable. Meanwhile, equity also encourages fair competition while accepting unequal results (Hick & Wrigley, 2009). Therefore, emphasising equity does not necessarily lead to more equality. On the contrary, to achieve equity, public policy may need to allocate educational resources unequally (Rawls, 1971).

5.2.1.2 Two important principles of equity: Merit-based and needs-based principles

While equity may mean equal shares under some circumstances, it can also require shares based on factors such as merit, effort, need, resources and opportunities (Larkin & Staton, 2001). Among perceptions of the principle of equity, merit-based equity and needs-based equity are two major categories. Merit-based equity means that merit should be the basis to evaluate whether the distribution of public resources is fair (UNESCO, 2018). Encouraging people to be productive participants in social life, merit-based equity protects the functions of a market economy. In the educational arena, pupils who perform better in high-stakes examinations are perceived to be those who have benefited the most from schools and would be more likely to contribute to society than their peers (Feinberg, 1970). Therefore, the system is fair if high-performers have access to educational pathways which are often unavailable to the rest—pathways that form the basis of academic selection. In contrast, needs-based equity pays more attention to individuals' disadvantages rather than their advantages (Feinberg, 1970). Following needs-based equity, society should compensate for the disadvantaged status of some social groups, and those with greater needs should receive more public resources, as they have inadequate family support. According to Levin (1990), a needs-based system should favour people under disadvantaged conditions such as those located at the bottom of the income distribution. While it is unrealistic to achieve a perfect needs-based system with equal access to the benefits of education for different social groups, public policies focusing on redistributing and allocating educational resources according to pupils' educational needs are consistent with this principle. Despite the contrast between these two principles of equity, the reality is not always dichotomous; the principles can coexist within the same system.

An underlying difference between merit-based and needs-based equity is whether individuals should be responsible for their own advantages/disadvantages. Emphasising agency, merit theorists believe that high-performing pupils are responsible for their advantages in attainment. As purposeful beings, pupils cannot achieve good results without their own effort, regardless of the degree of family support (Cooper, 2010). It is thus equitable if pupils with better performance receive better education. Since the distribution of educational resources is dependent on academic performance alone, it appears uncorrelated to a pupil's other characteristics. However, the challenge arises when there is a correlation between academic performance and social backgrounds, resulting in de facto privilege to pupils from advantaged families (UNESCO, 2018). Therefore, the merit-based principle of distributing educational resources might be unfair to pupils in disadvantaged families who lag behind before they even enter the education system.

Due to the importance of social context in pupils' learning process, merit-based equity is believed to over-emphasise factors over which individuals have little control, and award pupils for achievements for which they are not responsible. Unlike merit-based equity, needs-based equity attempts to distinguish circumstances from effort. Since people cannot determine their own circumstances, the advantages or disadvantages resulting from circumstances should not form the basis of distribution (UNESCO, 2018). Therefore, pupils who have achieved high academic performance as a result of favourable family backgrounds should not be rewarded for these achievements. Similarly, pupils born to families with limited resources, who have suffered from their experience of growing up in challenging conditions, should also not be punished for their low academic achievement (Roemer, 1998).

While it is clear in the needs-based equity principle that the (dis)advantages resulting from circumstances for which people are not accountable is not a fair basis of distribution, a more complicated question is whether the differences resulting from effort are justifiable. As Roemer and Trannoy (2016) mentioned, inequalities are never just if they result from factors other than effort, but not all inequalities due to effort are fair. For example, it is possible that the effort devoted to learning may be influenced by factors related to family background, such as motivation (Gorard & Smith, 2010). A detailed statement of why measuring effort is also problematic was given by Rawls:

> *That we deserve the superior character than enables us to make the effort to cultivate our abilities is also problematic; for such character depends in good part upon fortunate family and social circumstances in early life for which we can claim no credit* (1971, p. 89).

Based on the potential relationship between effort and circumstances, a more practical conceptualisation of effort is proposed. Roemer (1998) has advocated for considering individuals' background traits which are influential to the degree of effort. In this way, people with similar circumstances can be categorised into the same 'type' (e.g. gender, ethnicity, income). Following this principle, people with low effort because he/she belongs to a type which has a low mean of effort, should not be held responsible for their low achievement. This means that effort should not be evaluated in an absolute way. Rather, it needs to be presented in comparison with the degree of effort of other members of the same type. According to Roemer, calculating this new definition of effort should be based on the 'quantile of the effort distribution for his type at which an individual sits' (Roemer, 2002, p. 458). However, this complicated calculation is also subject to challenges such as practical difficulty and its underlying assumption that the differences between types are morally arbitrary (UNESCO, 2018).

The principles of equity stated above imply the difficulty of determining a universal definition of equity (Field, Kuczera & Pont, 2007). However, there is some evidence that when given a specific scenario, pupils show some agreement on what is fair or unfair, applying slightly different standards (Gorard & Smith, 2010). Proposing several principles of justice, such as recognising merit, equal opportunity, equal outcome, respecting individual, fair procedures, and appropriate treatment, Gorard and Smith (2010) believed that any single criterion of justice is inadequate, as it may lead to injustice under other circumstances. A similar statement was made by Walzer (1984) in *Spheres of Justice* proposing that no single standard of distributive justice is applicable to all goods. Rather, specific principles should be applied based on the meaning and the role of these goods, the domain of life, as well as the societal context (Miller, 1999; Walzer, 1984). This is especially true in education, as the aims of different levels of institutions are not the same. The equity principle in regard to access to HE may not be applied the same toward the opportunity to achieve basic reading and arithmetic skills during compulsory education. While the former may allow for a high degree of differentiation to accommodate individuals' future career needs, the latter should pay more attention to the needs

of disadvantaged groups, as it emphasises a universal provision of basic knowledge necessary for each citizen (UNESCO, 2018).

5.2.2 Equity in academic selection

The nature of selective secondary education is to create differentiated provisions for pupils with dissimilar characteristics after they finish primary school. Some selection criteria may seem 'neutral' in the sense that those who meet the requirements would not receive extra benefits, as in single-sex schools. However, this is not the case for schools selecting by academic ability. The underlying principle of academic selection is that academic achievement deserves to be rewarded. Therefore, pupils who perform better at the end of primary school should be given the chance to enrol in schools which are believed to be superior to other alternatives within the state system. Since the assumed benefit associated with academic selection is enjoyed by a small group of pupils, it is vital to evaluate whether access to these 'superior' schools is distributed equally across pupil groups. Therefore, academic selection is intertwined with the issue of equity.

5.2.3 Access to academically selective schools and family background

After discussing the complicated definitions of equity in education, and the reasons why equity is highly relevant to academic selection, this section begins to present the issue of equity in academic selection. It starts with the unbalanced opportunities to take the selection test across social groups, and then focuses on the reasons why certain groups tend to have better opportunities to pass selection tests. After revealing that access to academically selective schools might be unbalanced across social groups, this section also addresses the stratified future destinations associated with different educational pathways. This demonstrates how academic selection may influence the distribution of educational resources and outcomes.

5.2.3.1 The choice of whether to take the selection test

Given the wide variation in parents' school choice decisions, deciding whether or not to take the selection test is the first step in differentiating candidates for academically selective schools from their peers. In addition to the general process by which different social groups make their school choices, there are several unique elements involved in the choice to attend academically selective schools (Coldron, Willis & Wolstenholme, 2009; Parsons & Welsh, 2006).

It may seem strange that some parents forego the selection tests and do not let their children at least have a try. The first reason for not taking the selection test is simple—a lack of the perceived academic performance required by selective schools. The association between family background and academic performance implies that pupils from disadvantaged backgrounds are less likely to perform as well as their advantaged peers (this is explained in detail in the next section). For many working-class families, the opportunity to attend the selection test may seem irrelevant and the test is just a waste of time. When an opt-out test system was adopted in Northern Ireland before 2008, a noticeable number of parents still withdraw from the selection test and gave up the possibility of attending higher tracks within the selective system. They did so because they thought that their children would not score high enough anyway (Department of Education, 2000). In many areas, the selection of secondary schools adopts an 'opt-in', rather than 'opt-out', system. This means that only pupils whose parents have registered will have the chance to sit the selection test. We might expect that foregoing the test among pupils with lower achievement would be more common in an opt-in system.

In addition to lower performance, another reason for less advantaged pupils not attending the selection test lies in the elite images many selective schools uphold. The division between academically selective schools and other state-funded schools is symbolic in many countries (Parsons & Welsh, 2006). Targeting the most able, selective schools usually have more pupils from advantaged families with middle-class parents. Offering a wider range of extracurricular activities, and usually requiring students to wear 'posh' uniforms, these schools strive to maintain their elite images. While these efforts attract pupils from advantaged families whose parents value these images, working-class parents may be hesitant to send their children to schools where the dominant culture is so different from their own, so as to avoid humiliation and potential failure (Windle, 2009). Academic performance is a key element in school choice, but it is not everything. Parents are sensitive to the complex of behaviour and activities in school, and they also have practical considerations over whether their children can fit in and get along well with their classmates (Bernstein, 1975). Evidence from England has demonstrated that among the 100 most socially selective comprehensive schools, some do not produce strong academic results, with only a minority of their pupils achieving five good GCSEs by the end of compulsory education (Smithers & Robinson, 2010). This reveals that school culture and atmosphere are sometimes regarded as even more crucial than school performance. Parents often hope to send their children to schools which have been chosen by

families similar to them. This evaluation is not limited to family income, but also includes elements such as ethnicity and religious background.

While some parents from less advantaged backgrounds may refuse to approach academically selective schools due to distinct school ethos, it is also possible that others have inadequate interest or ability to evaluate whether these selective schools are suitable for their children, as the selection processes are often complicated. This may also reduce the rate of attending the selection test among less advantaged pupils, based on the extra time and effort involved (Weldon, 2018).

5.2.3.2 The link between selection results and family background

By definition, access to academically selective schools is determined by academic tests. These tests are usually believed to evaluate pupil's ability—the basis for determining whether they are suitable to be educated in higher pathways (Coldron, Willis & Wolstenholme, 2009). The selection criteriaa appear fair because only high-performing pupils receive better opportunities. However, it is generally believed that there is an undeniable connection between early-age performance and family background (Ermisch & Francesconi, 2001). The negative influence of poverty on academic performance has been discussed by Coleman since the last century (1966). Analysis of OECD countries has also shown that SES is a powerful predictor of pupil's performance, as the SES of families and schools accounts for about 60% of the differences in test scores, on average (OECD, 2010). A similar conclusion was reached in a study of the Third International Mathematics and Science Study (TIMSS) data. It revealed that family status (e.g. parental occupation and the number of books in the household) are predictive of children's academic achievement (Peterson & Woessmann, 2007). There is also an increasing body of research uncovering possible reasons for the stubborn pattern of underachievement among the poor (Smith, 2005). Household poverty and its associated problems, such as poor health conditions, violence, insecurity, and the lack of positive role models in the community, are all detrimental to pupils' living conditions and academic aspirations. Meanwhile, pupils' academic progress is influenced by parental involvement. Parents who positively engage in their children's learning process and work in partnership with schools benefit both the cognitive and non-cognitive skills of their children (Lareau, 1997; Siddiqui & Ventista, 2018). However, the degree of parental engagement is also constrained by family background. Since

pupils from families with insufficient support are unlikely to get proper guidance at home, they may rely on school as the only source of learning (Tudge et al., 2000).

This means the academic competition between different social groups has always entailed a degree of unfairness. The gap in cognitive ability is significant from the starting gate, with evidence in the US revealing that children in the top SES group achieve 1.6 times higher than the lowest SES group, before kindergarten (Lee & Burkam, 2002). Similar results have been found in the UK, demonstrating ability differences among 22-month-old children (Feinstein, 2003). Children from poorer families are less prepared to start nursery than more advantaged peers, and they are more likely to be enrolled into primary schools of lower quality. This reinforces the inequality in attainment between social groups. Therefore, even if the opportunity to attend schools in higher tracks is decided by performance, we can expect it to be distributed unevenly across social groups.

There are additional reasons why an advantaged group may be more likely to pass selection tests. Apart from the unbalanced development of early-age cognitive ability across social groups, those more familiar with the test content, and with better preparation, are also more likely to pass selection tests. The different levels of preparedness may be the result of teaching quality in primary schools, but the pattern is also believed to be relevant to private coaching for the more advantaged. According to evidence from Northern Ireland, a majority of parents pay for private coaching to help their children pass the selection test (Department of Education, 2000). However, this is a less affordable option for disadvantaged pupils who rely on school to prepare them for the selection test. Lacking proper preparation, some children from poorer families fail to pass the selection test, despite their exceptional academic abilities (Galindo-Rueda & Vignoles, 2005). In contrast, there are other pupils who pass the selection test and go to selective schools, but have been found to perform far below the level predicted by their high selection scores. These pupils are believed to have been over-coached for the selection, with their test results overestimating their real ability (Department of Education, 2000). The tendency for the rich to manipulate examination systems is not limited to secondary school choice. It has also been found that private schools pupils tend to do worse and have lower academic performance in universities than pupils from state-funded schools with equivalent GCSE scores. This has also been considered a reflection of the wealthy's over-preparation for high-stakes tests (Smith & Naylor, 2001).

To conclude, the decision of whether or not to take the selection test may systematically differ between social groups, especially in areas with opt-in selection systems. This leads to disparities in access to selective schools, which is then compounded by the association between test scores in selection and pupil's family backgrounds. Therefore, the opportunities to attend these academically selective schools might be unequally distributed across social groups, and selective schools are more likely to be populated by advantaged pupils.

5.2.3.3 Different educational pathways and pupil's future destinations

The process of how parents decide whether or not their children will take the selection test, and the unbalanced rate of passing the test, imply that whatever the benefits (or drawbacks) of attending academically selective schools are, they will be unevenly distributed across social groups. Therefore, it is necessary to evaluate future opportunities accessible to pupils in different pathways.

The pathways of secondary education influence subsequent choices, as they usually qualify pupils for different post-compulsory destinations. Based on the definition of academic selection, pupils who have passed selection tests are provided with more challenging and tougher academic tasks which usually lead to promising future destinations. Although the real effectiveness of academically selective schools remains contestable, these schools are originally designed to promote academic performance superior to that of other schools. Enrolling high-performing pupils and producing high test results, these schools in favourable positions may be more likely to attract highly qualified teachers and managers (Allen, Mian & Sims, 2016). This means that whether or not the proposed effectiveness of selective schools is real, the system of academic selection strives to create favourable learning environments for those who have passed the selection.

Global evidence has revealed that pupils in academically selective schools usually produce better results at the end of secondary school, and have higher rates of post-16 educational participation. They are also more likely to perceive HE as an essential step in their lives. In a study in Northern Ireland, all of the grammar school pupils expressed their intention to stay for post-16 education, and most planned to apply to HE institutions. In contrast, many pupils in non-selective secondary schools were unsure whether to stay after the compulsory stage; few expressed their willingness to attend university (Department of Education, 2000).

Advantageous status for pupils in grammar schools has also been found in England, as they usually have better academic performance and are more likely to attend university than their peers in other state-funded schools (Atkinson, Gregg, & McConnell, 2006; Clark, 2010; Mansfield, 2019). A similar situation has been reported in Germany. Although the process of dividing pupils into different secondary schools is no longer based on ability tests, the disparity in the rate of HE participation between different educational pathways remains salient. For German pupils, there are strong connections between high school and academic HE degrees, intermediate school and white-collar schemes, and general school and blue-collar schemes (Dustmann, 2004). This means the route taken during the compulsory stage is predictive of pupil's later academic destinations. The differentiated patterns of post-18 participation between educational pathways also mediate job status at later stages.

The selective system not only imposes academic and occupational effects, but also influences identity formation and emotional adjustment. Claiming to provide an evaluation of ability, the selection results carry considerable weight (Coldron, Willis & Wolstenholme, 2009). As the academic selection system entails inevitable 'losers', it may create a sense of failure for those rejected at an early age, as it would be viewed as inappropriate for them to pursue destinations associated with higher tracks (Coldron, Willis & Wolstenholme, 2009). It has been reported that pupils are clearly aware of the school hierarchy and the social meaning attached to each educational pathway. Some teachers in non-selective secondary schools serving pupils who had failed the selection mentioned that these pupils arrived with frustration and a low sense of self-worth (Department of Education, 2000). As a result, apart from daily learning, an extra task for these secondary schools was to rebuild their pupils' self-confidence, boost their academic aspirations, and avoid the negative circle of self-verification. The story for those who passed the selection is the opposite. Pupils who passed the selection are positioned with status superior to that of their peers, as they are 'officially approved' as academic high-achievers. These pupils usually enjoy high aspirations toward their futures (Department of Education, 2000). Therefore, selection is not only the process of 'predicting' how pupils will perform in the future based on their previous learning experience; it may also mediate their performance.

The results of the segregated school choice pattern in a selective system, as well as the unbalanced success rate in the selection process, reveal how the opportunity to attend selective schools may overlap with family background. Therefore, any possible benefits of selective schools in creating access to more privileged destinations may be disproportionately enjoyed

by those from more affluent families, with their advantages exacerbated. This implies the potential role of the selective system in reinforcing social inequalities, revealing the strong connection between academic selection and the issue of equity (Reichelt, Collischon & Eberl, 2019).

5.3 The relationship between effectiveness and equity

While previous sections have discussed effectiveness and equity independently, this section focuses on the relationship between these concepts. The primary issue of interest is whether there is a trade-off between effectiveness and equity, and whether these two elements can be achieved concurrently.

Some researchers believe that the effort devoted to enhancing equity inevitably slows down effectiveness, as the former detracts resources from more productive groups in favour of those with greater needs (Field, Kuczera & Pont, 2007). However, evidence shows that effectiveness and equity are not contradictory, and might even be mutually beneficial in terms of economic development (World Bank, 2006). Similar conclusions have been reached in the educational arena, as no obvious trade-off between educational quality and equity has been found (Pfeffer, 2015). International research conducted among OECD countries has shown that there is no systematic correlation between PISA test scores and the segregation index (Andrews, Hutchinson & Johnes, 2016). Several countries, such as those in Scandinavia, outperformed others in terms of both effectiveness and equity. Others, like the Czech Republic, did poorly, both in terms of raising overall education standards and in reducing educational gaps. Meanwhile, countries such as the UK and France, did well in terms of education quality, but needed more effort to tackle the issue of equity (Gorard & Smith, 2010). In addition to the country-level evidence, the evaluation of school-level data in England also confirms that schools with better performance are not necessarily those with higher rates of segregation. Many high-achieving comprehensive schools enrol fair shares of disadvantaged pupils and have socially-balanced intakes. Some comprehensive schools even admit more FSM pupils than the overall proportion of their local areas (Cullinane et al., 2017). In contrast, the performance of the most socially selective comprehensive schools does not always match their prestigious status, with some even yielding unsatisfactory academic results (Smithers & Robinson, 2010). Therefore, there is no evidence that a more segregated and less equitable

system promotes effectiveness. On the contrary, effectiveness can be achieved in conjunction with equity (Peterson & Woessmann, 2007).

Overall, this chapter has focused the two core terms of this study—effectiveness and equity. In addition to discussing these two terms and their relationship with academic selection independently, this chapter has also revealed that the perception of a trade-off between effectiveness and equity might be illusionary. Previous research has found no robust evidence that effectiveness cannot be achieved in conjunction with equity. On the contrary, these two elements might be mutually beneficial, implying the possibility of education systems that are both effective and equitable.

6 Selection, segregation and integration

Apart from explaining why academic selection is intertwined with issues of effectiveness and equity, this chapter discusses the impact of selection on segregation and integration.

Integration is an important yardstick of a healthy society. However, coherence and integration are difficult tasks, and many practical issues impede their realisation. Logistically, pupils attend schools with peers of similar backgrounds because similar families tend to live near each other. Thus, the dilemma of residential segregation based on housing price and community culture influences schooling patterns (Gorard, 2006). The quasi-market system and the emphasis on educational agency through parental choice is an attempt to break the association between housing segregation and school segregation. However, as presented previously, the assumed benefit of the educational market may be unrealistic and the system would still segregate pupils, both through overt and covert selection criteria. Therefore, the issue of desegregation and integration is not only relevant to selective state schools, but also to comprehensive schools as well as independent schools.

According to previous research, education integration may be beneficial (Gewirtz, Ball & Bowe, 1995; Gorard & Smith, 2010). One of the most significant positive outcomes is improved educational achievement, especially for less advantaged pupils and minority groups. For example, the process of comprehensivisation in Scotland is believed to raise the performance level of the working-class who have benefited from the favourable integrated school context, even though comprehensive schools may also face with the challenge of integration due to residential segregation as well as within-school selective places (e.g. tracked curricula) (Coe et al., 2008; Echols & Willms, 1995). Evidence from the US has shown that schools with a balanced racial and ability composition improve the academic achievement of ethnic minorities and decrease their drop-out rates (Gewirtz, Ball & Bowe, 1995). Similarly, pupils with special educational needs (SEN) may achieve higher test results if they are placed in mainstream schools (Warnock, 1978). In addition to academic performance, school integration also improves occupational aspirations. The effect of school composition on pupils' aspirations discussed in previous sections means that disadvantaged pupils in a mixed school tend to aim higher, and schools dominated by working-class pupils are harmful to pupils' occupational aspirations (Gorard & Smith, 2010). Furthermore, since attending a socially mixed school is negatively correlated with having criminal record in adulthood, being educated

in a heterogeneous school context is also believed to reduce the risk of social problems (Lankford, Loeb & Wyckoff, 2002; Massey, 2006).

The possible benefits of integration are not merely relevant to disadvantaged pupils (Gorard & Smith, 2010). According to Piaget's theory of disequilibrium (1983), learning in a setting with diversity, contradictions and discrepancies challenges young people's taken-for-granted perspectives. This provides valuable practice for their cognitive capacity, and stimulates their intellectual potential (Cohen & Lotan, 1995; Gurin et al., 2003). As a result, attending racially and socioeconomically balanced schools has been found to be positively correlated with HE participation rates (Mickelson, Nkomo & Wimberly, 2012). Therefore, there is evidence that pupils from more advantaged groups also tend to progress better in diverse schools (Gurin et al., 2003).

In addition to academic achievement, an equally valuable outcome of education is preparing pupils to work and live with other people. It has been argued that one of the greatest detriments of a segregated peer group is the lack of experience in getting along with people from other backgrounds at an early age. International research has revealed that pupils with friends from other ethnic groups are more amenable to mixed environments, and those with positive diversity experiences also attach more importance to it (Gorard & Smith, 2010). Researchers have noticed that diverse schools foster pupil's awareness of others, their respect for cultural differences, and their ability to negotiate conflict (Mickelson, Nkomo & Wimberly, 2012). Meanwhile, the mixed school composition also helps young people develop positive attitudes towards democracy, civic responsibility and social engagement (Kurlaender & Yun, 2006). On the contrary, a segregated and differentiated school system may undermine the public goals of a society, as pupils in different schools have fewer shared objects and less similar learning experiences (Levin, 2018). Therefore, an integrated school composition is crucial to social cohesion, harmony, and long-term development.

It is also believed that children are highly sensitive to their environments, and early-age exposure to diversity fosters positive attitudes. Therefore, any efforts to create social coherence and integration are most effective among younger pupils (Allport, 1954). When the optimal time window is missed, it is difficult to compensate with later interventions. This means that who attends school with whom not only matters; it matters the most for young pupils in primary and secondary school when they are forming their cultural attitudes and value systems. In order

to foster integration, pupil's actual diversity experiences need to be guided through fair and just interactions, active conversations, as well as tolerant and inclusive campus events (Gurin, Nagda & Lopez, 2004; See, Gorard & Siddiqui, 2017). However, it is clear that a segregated school composition is unable to provide these opportunities.

Similar to the important role of academic selection in terms of effectiveness and equity, the influence of academic selection on social integration is another major area of discussion. Besides the general situation of academic selection, the next chapter starts to focus on the specific issue of grammar schools in England.

7 Introduction of academically selective schools in England: Grammar schools

After discussing academic selection, effectiveness, equity and social integration in general, this chapter finally turns to the specific situation of academically selective schools in England. It starts with the origins of grammar schools in England and their historical status. This is followed by current debates on the future development of today's grammar schools. While there are strong political attempts to expand grammar schools, there are also opposite voices suggesting their closure, making the issue of grammar schools a hot topic in England.

7.1 History of grammar schools

Secondary education in England has a long history of dividing pupils into different pathways according to their academic, religious, or other personal characteristics. According to the 1944 Education Act, pupils with different levels of academic attainment in England should be allocated to different pathways within the Tripartite System, which includes grammar schools for academically-oriented pupils, secondary modern schools for pupils aiming for practical works, and technical schools for pupils with interest in technical areas (Morris & Perry, 2017; Foreman-Peck). However, the Tripartite System is also referred to as the Bipartite System in reality, as there were very few technical schools post the Act. While technical schools were envisaged in the Act, they were only established in a few LAs, due to reasons such as the absence of local interest in technical education, and the lack of skilled teachers to meet the demand of technical education. Under the de facto Bipartite System, pupils who passed the ability test at the age of 11 could attend grammar schools, which provided access to HE and more prestigious occupations in the future. In contrast, those who did not pass the academic selection were allocated to secondary modern schools, which were not intended to prepare their pupils for post-compulsory education. The courses taught in secondary schools were more practically-focused, equipping their pupils with basic literacy and numeracy (Kerckhoff et al., 1998). The system thus creates a strong dichotomous status between grammar school pupils and pupils in secondary modern schools through the academic selection at the age of 11.

Based on the consensus from the 1940s to the mid-1960s that secondary education should be selective, the proportion of grammar school pupils continued growing and peaked at 38% in 1964 (Bolton, 2017). However, starting in the 1950s, the selective system has already received public attention due to its sensitive linkage with social equity and segregation. In short, grammar schools were dominated by pupils from professional or managerial families, and they

did not appear to lower the effects of family background on attainment (Power & Whitty, 2015). As a result, grammar schools were gradually replaced by comprehensive secondary schools, and the national secondary education system had been converted into a non-selective system as proposed by the Labour Government in 1964 (Jesson, 2013). Since the 1998 School Standards and Frameworks Act, the opening of new grammar schools has been prohibited, and the total number of grammar schools has settled at around 163. Currently, only 5% of English pupils attend grammar schools (Bolton, 2017).

7.2 Current debates on grammar schools

Today's grammar schools are probably the most controversial schools in England, which received public attention disproportionate to their small number. Uniquely retaining academic selection within the comprehensive system, grammar schools have sensitive links with effectiveness and equity, which are key issues and buzzwords for England.

Despite Labour's 1998 ban on establishing new grammar schools and their stable number for several decades, there have been constant voices in support of the reintroduction of grammar schools. The effort to revive grammar schools has continued, mostly from the Conservative Party (Morris & Perry, 2017). The past five years have witnessed the intensification of this tendency. In 2015, the expansion of the Weald of Kent was approved. Although grammar schools can expand within a legal frame, the approval aroused fierce public discussion. Instead of merely increasing school places within their original campus, the Weald of Kent was allowed to establish an annex 10 miles away. The 'expansion' has thus been regarded as the first 'new' grammar school in 50 years (Coughlan, 2016). As an encouraging sign of the looser restrictions on grammar schools, it was reported that at least 8 more regions were planning a similar expansion for their grammar schools (Espinoza, Finnigan & Gurney-Read, 2015). The debate was further intensified by the government led by Theresa May in 2016. As an important part of Theresa May's education reform, selective schools were encouraged to be more active in raising academic performance nationally. Furthermore, as the government claimed to make England a place that 'works for everyone, not just the privileged few' (DfE, 2016, p. 5), grammar schools' assumed role in reducing social inequality to fulfill the potential of pupils from less advantaged backgrounds has been more strongly emphasised. In a Green Paper, *Schools That Work for Everyone* (DfE, 2016), several ways of developing grammar schools at the national level were suggested. Existing selective schools would be allowed to expand and

new ones could be established when parental demand for selection was demonstrated. Well-performing non-selective schools could also be converted into selective schools if they meet the conditions for serving more disadvantaged pupils. Examples of these conditions would be accepting a proportion of underprivileged pupils, establishing a primary feeder in poorer areas, and providing more opportunities to join the school at different ages (DfE, 2016, p. 25). An additional statement supporting the proposal was published by the government in February 2017, in which this new education system with more selective schools was believed to 'increase parental choice, create more good schools and decrease the attainment gap between children from high and low socioeconomic groups' (House of Commons, 2017, p. 5).

Although the attempt to lift the ban on grammar schools was suspended after the loss of a majority in parliament during the election of 2017 and the aforementioned Green Paper was thus abandoned, the intention to expand grammar schools has not gone away (Harding, 2017). In May 2018, the Department for Education (DfE) announced a £50m expansion fund to create more grammar school places. Targeting only a handful of grammar schools, this is a simplified and compromised plan of its original intention to reintroduce the selective education system (George, 2018). In December 2018, the list of 16 grammar schools which would receive expansion funds was released, with 4,000 new school places planned (Sellgren, 2018). While the number of grammar schools supported by the fund at this stage is small, the message is profound. It emphasises the importance of grammar schools and the continued possibility of returning to the selective system. While the changes in political structure of the Conservative government in July 2019 imply that some previous policies may not be implemented, the whole back-and-forth process since the last century has confirmed the importance of grammar schools and their sensitive roles in England's education system. Moreover, although only around 5% of pupils in England attend grammar schools now, existing grammar schools can still increase their enrolments nationally under the current system.

The expansion of grammar schools requires financial support. More funds allocated to expand grammar schools means less support for other areas of education, within set government budgets. The need to legitimise public spending on increasing grammar school places is even more pressing at the present time when schools across England are facing tightening budgets. Released by the National Audit Office, per capita education funding in England has shrunk by 8% in real terms since 2010, with about three billion pounds of funding cut by 2019 (Sibieta, 2018). The annual expenditure released by DfE shows that one third of mainstream secondary

schools were experiencing budget crises in 2018, about three times as high as the figure from five years before (DfE, 2018). Due to the funding shrinkage, many secondary schools have to reduce the number of available GCSE subjects and increase class sizes (Adams, 2017). The funding cuts are more detrimental to schools with poorer intakes, where parents are unable to offer extra financial support. It has been warned that disadvantaged pupils, especially children with special educational needs and disabilities (SEND), are more vulnerable to the negative consequences of financial shortfalls (Adams, 2018). There are also complaints that the large-scale funding shrinkage will mitigate the government's plan to increase literacy and threaten the quality of basic education (Kentish, 2018). It is thus believed that investment in the expansion of grammar schools, which targets a small group of academically able pupils, comes at the cost of budget cuts to basic educational support for the vast majority. Therefore, unlike the strong political attempts advocating the expansion of grammar schools, there are also opposite voices resisting the expansion. The expansion of grammar schools has received strong resistance most noticeably from campaign groups (e.g. Comprehensive Future and the Kent Education Network) and academics, due to concerns such as low academic benefits, segregated school composition, and impediments for disadvantaged pupils (Allen, Bartley & Nye, 2017; Andrews, Hutchinson & Johnes, 2016; Cribb et al., 2013). Based on these concerns, it is believed that grammar schools should not be expanded. On the contrary, they may even need to be closed.

In sum, the existence of grammar schools has become a sensitive issue in the educational system of England since the last century. The unsettled debates on their impact have aroused both political attempts at their expansion, and resisting voices advocating their closure. This reveals the importance of an accurate and unbiased evaluation of the effectiveness and equity of today's grammar schools. The following chapter thus focuses on previous research on grammar schools and evidence so far.

8 Previous research on grammar schools' effectiveness and equity

This chapter focuses on previous evidence of the effectiveness and equity of grammar schools in England. It starts with the possible impacts of grammar schools on pupil's later academic performance and post-18 destination, and then turns to the unbalanced access to grammar schools between social groups.

8.1 Grammar school attendance and academic performance

While the costs of expanding grammar schools are substantial, it remains unclear whether there is any benefit to the expansion, as well as the extent of such benefit if it exists. This section thus pays attention to previous evidence of the effectiveness of grammar schools.

Similar to supporting opinions of any forms of selection, proponents of academically selective systems argue that pupils with different academic abilities can learn more effectively if they study together with peers of similar ability. Officially, the government claimed that the grammar school expansion policy has two major benefits. First, grammar schools can enhance national academic standards. Second, they will promote social equity by providing better results regardless of pupil's family backgrounds (The Conservative Party, 2017; DfE, 2016). Since both reasons are based on the perception that grammar schools are more effective than other state-funded schools in raising academic performance, the real effectiveness of grammar schools is thus the most fundamental issue. Without knowing the answer, it is impossible to evaluate the influence of the expansion of grammar schools.

8.1.1 The application of the value-added approach in grammar schools' effectiveness

As mentioned in previous chapters, the VA approach is probably the most widely used approach in educational effectiveness research. Many researchers in England have also evaluated the effectiveness of grammar schools following the principle of VA. So far, the evidence is mixed, but some systematic patterns emerged. A common practice in each of the studies discussed below is to control for both pupil-level and school-level variables. The major differences have been found between those studies which account for school-level prior attainment, and those which do not. While the latter usually finds a positive grammar school effect, the former tends to report no difference between grammar schools and others (see Appendix 1 for details of these studies).

8.1.1.1 Studies without controlling for school-level prior attainment

For studies without school-level prior attainment, the general pattern of the estimated grammar school effect is positive. For example, applying ML models, Schagen and Schagen (2003) found larger progress in grammar schools from Key Stage 2 (KS2) to Key Stage 3 (KS3), and a small positive advantage from KS3 to KS4 for average pupils. But for high-performing pupils at KS3 (level 7 or above), there is no difference in progress associated with school type. Furthermore, they also concluded that selective LAs are slightly more effective from KS2 to KS3 than comprehensive areas. A later study by Atkinson, Gregg, and McConnell (2006) applied the OLS regression, concluding that grammar school pupils achieved four grades higher on capped GCSE than equivalent pupils in non-selective areas, especially for borderline grammar school students. Meanwhile, those who were not accepted into grammar schools did slightly worse than equivalent pupils in non-selective areas. This leads to a small net gain of the selective system. Similarly, Levačić and Marsh (2007) noticed a six-grade advantage on total GCSE/GNVQ associated with grammar school attendance, but again the effect drops among high-attaining groups. Paying attention to the negative consequence of the presence of grammar schools on surrounding schools in selective LAs, they also mentioned that pupils in secondary modern schools are doing one grade worse than equivalent ones in comprehensive schools. Apart from the national pattern, a study focusing on Buckinghamshire also found a positive grammar school effect (Harris & Rose, 2013). According to Harris and Rose (2013), grammar school pupils are 10% more likely to achieve 5 A*-C on GCSE/GNVQ than equivalent pupils in non-selective schools, but the negative impacts on pupils who did not get into grammar schools is also found. In a report of Education Policy Institute, Andrews and his colleagues (2016) found a relationship between the benefit of grammar schools and the overall degree of selectivity. While the grammar school effect is about 1/3 grade per GCSE subject, this effect drops when the proportion of grammar school places in the local area goes up. The effect decreases to 1/10 grade per subject when grammar schools offer more places than the number of high achievers in that area (Andrews, Hutchinson & Johnes, 2016). Meanwhile, the overall effectiveness of the selective system is the same as the non-selective system (Andrews, Hutchinson & Johnes, 2016). Overall, the differences in statistical models, outcome variables, and geographical areas between these studies complicate the comparison of results. However, in all of these studies, grammar school pupils made more progress than their counterparts in comprehensive schools and secondary modern schools.

One exception is Gorard and Siddiqui's (2018) study, which did not account for school-level prior attainment, but still found no positive grammar school effect. They noticed that adding 'whether a pupil went to a grammar school or not' did not improve the fitness of the model for predicting capped GCSE results. This implies that grammar schools are not more effective than other schools in raising pupils' academic performance. Unlike previous research using traditional binary indicators of FSM eligibility, they calculated the total years of being eligible for FSM. Since this new variable carries more information than traditional ones, it may be more powerful in removing unmeasured differences between grammar school pupils and others, thus reducing the estimated effect of grammar schools.

8.1.1.2 Studies controlling for school-level prior attainment

One of the most extensive attempts at evaluating grammar schools' effectiveness is that of Coe et al. (2008). Based on both a systematic review of previous research as well as their own analysis applying OLS and ML models, they concluded that pupils in grammar school might have an advantage of 0 to 3/4 of a GCSE grade per subject. The wide variance in the estimated effect is primarily the result of the choice of baseline variables. If regression models only control for pupil-level variables, the estimates are substantially positive (reaching 0.75 grade per subject at most). However, once school-level variables, which include prior attainment, proportion of FSM, the Income Deprivation Affecting Children Index (IDACI), and single-sex status, are also controlled for, the difference between grammar schools and others drops to around 0. Coe and his colleagues (2008) believed that the lower end of the grammar school effect is more reliable than those substantially positive results, since grammar school pupils have already progressed more than their peers, prior to attending secondary schools. The conclusion thus presents salient differences from most studies discussed in the previous section which also controlled for pupil-level and some school-level variables, but omitted school-level prior attainment.

The substantially decreased grammar school effect when school-level prior attainment is added might be related to the controversial 'compositional effect' (Harker & Tymms, 2004). It is also possible that adding this variable removes more unmeasured differences between pupil groups, differences which are not sufficiently accounted for by pupil-level surface variables or other school-level compositional variables (Coe et al., 2008). More importantly, adding school-level prior attainment is believed to correct measurement errors in baseline variables which

otherwise upwardly bias the effectiveness of more advantaged schools (Perry, 2019). Based on this finding, Perry (2019) concluded that studies without accounting for school-level prior attainment present a 'phantom' grammar school effect. Despite supporting claims of controlling for school-level prior attainment, concerns remain. It is possible that accounting for school-level prior attainment removes genuine difference between schools, especially if beneficial school actions correlate with more advantaged intakes (Raudenbush & Willms, 1995). In this case, the estimated grammar school effectiveness may be lower than its actual size.

In sum, the choice of baseline variables in the VA approach is difficult, and there is no perfect solution (Visscher, 2001). Previous studies do not reach a consensus on the effectiveness of grammar schools, and the mixed evidence suggests that grammar schools may perform slightly better than other state-funded schools, yet we are unsure to what extent.

8.1.2 The regression discontinuity design in grammar schools' effectiveness

The difficulty of choosing the appropriate baseline variables in the VA approach reveals the importance of applying stronger research designs to evaluate the effectiveness of grammar schools. As one of the most powerful alternatives to an RCT, the RDD approach overcomes the difficult issue of baseline variables, which is strong enough to make causal inferences. Based on the strength of the RDD approach, there have been fruitful applications for school effectiveness (e.g. Gibbons, Machin, & Silva, 2013; Luyten, Tymms, & Jones, 2009). However, there is only one study which has used this design to evaluate the effectiveness of grammar schools in England. Clark (2010) focused on four grammar schools in East Riding and detected a small grammar school effect in Year 9 test scores, which is 7% higher than pupils just below the cut-off point. However, apart from the small number of grammar schools, this study applied data in the late 1960s, when the transformation of comprehensive schooling was prevalent. Data collected decades ago raises doubts on the external validity of this research in the present (Coe et al., 2008).

To conclude, despite a plethora of evidence both from the government and researchers, grammar schools' effectiveness in raising academic performance remains unclear due to reliance on passive designs which are not conducive to causal inference, and the lack of evidence from stronger research designs such as the RDD. Overall, there is no consistent

conclusion on whether there is any benefit of grammar school attendance, as well as the extent of such benefit if it exists. Apart from the role of grammar schools in raising pupils' academic performance, previous research has also focused on the relationship between grammar school attendance and participation in HE. The detailed evidence is discussed in the following section.

8.1.3 Grammar school attendance and participation in HE

Previous research on the impact of grammar schools has focuses not only on their role in raising academic performance, but also on patterns of HE participation. Although less research has addressed the relationship between grammar school attendance and HE participation, the conclusions are still mixed.

Some researchers have found a positive grammar school effect on pupil's post-18 academic destinations. A 2019 study on the progression to HE revealed that while there is no obvious difference in the general HE participation rates of impoverished pupils from selective and non-selective areas, the former are more than twice as likely to go to Oxbridge (Mansfield, 2019). Similarly, Clark's (2010) study also provides evidence of the positive impact of grammar schools on HE participation, as their pupils are 6% more likely to attend HE institutions than those who had just been rejected by grammar schools (Clark, 2010). However, a contrary pattern has been found in other research which has shown that grammar schools do not influence pupils' academic trajectories. For example, according to Crawford (2014), grammar school pupils have an initial advantage in accessing HE institutions—they are 40% more likely to attend universities than other pupils within the state system. However, the advantage drops to 4% when initial differences between pupil groups are considered. Therefore, similar to the discussion of the relationship between grammar school attendance and later attainment, the higher rates of attending universities among grammar school pupils can also be mostly explained by their advantaged characteristics (Chowdry et al., 2013). The weak link between grammar school attendance and HE participation has also been noticed by Boliver and Swift (2011) who pointed out that grammar schools have little impact on pupils' future paths, even for pupils from lower origins. Similarly, Sullivan et al. (2016) found no positive link between attending grammar school and achieving HE degrees. Likewise, Iannelli (2013) noticed that school type does not influence children's future social class.

Similar to the inconclusive evidence for grammar school's effect on academic performance, conclusions on the relationship between grammar school attendance and HE participation in previous studies are also mixed. This reveals the need of further evaluations again. Apart from discussing existing evidence for whether attending grammar schools is associated with better outcomes, the following section also presents previous research assessing the roles of grammar schools in social equity.

8.2 Grammar schools and social equity

As mentioned previously, besides grammar schools' assumed superiority in effectiveness, another important reason behind the grammar school policy is their perceived role in promoting social equity. While the secondary education system in England is largely comprehensive, indirect forms of selection based on pupil's family backgrounds still exists. For example, parents need to pay £45,700 more than the average price to buy a house in the catchment area of a top comprehensive school (Cullinane, Hillary, Andrade & McNamara, 2017). As a result, good comprehensive schools are accepting many fewer FSM pupils than their fair share, and the rate is only 9.4% among the top 500 comprehensive schools. Unlike the admission principle of comprehensive schools which carries some forms of social selection, the admission of grammar schools is solely decided by pupils' academic ability. Based on the perception that grammar schools provide high-ability students in lower SES families with more opportunities and that grammar schools benefit those students eligible for FSM, the presence of grammar schools is assumed to make pupils' future success less determined by family backgrounds but more so on their own talents and efforts. Following this assumption, grammar schools are believed to provide social ladders for children from low SES families and have a positive impact on social equity (Randall, 2009).

However, grammar schools' perceived role in promoting social equity is based on the premise that a considerable group of disadvantaged pupils are actually enrolled into grammar schools. Contrary claims that grammar schools might deepen the gap of inequality is thus made by others due to the underrepresentation of low SES pupils in grammar schools, whether taking prior attainments into account or not (Harris & Rose, 2013). Similar to the general patterns of how parents are making different school choices for their children (Chapter 3 and Chapter 5), FSM pupils are substantially less likely to attend the selection test of grammar schools (Allen, Bartley, & Nye 2017). Even if they do so, the pass rate for FSM pupils is only 12%, but the

rate is 30% for non-FSM pupils (Allen, Bartley, & Nye 2017). As a result, the proportion of FSM pupils in grammar schools is less than 3%, which is only about 1/4 of the national rate. Additionally, only 7% of grammar school pupils have been eligible for FSM at any point over the previous 6 years, but the national rate is 31% (Andrews, Hutchinson & Johnes 2016; Nye, 2016; Sibieta, 2016). Moreover, evidence also shows that grammar school pupils are more likely from primary schools where the average proportions of FSM and SEN pupils are lower, and they are twice as likely to be educated in private schools during primary education as the national average (Cribb et al., 2013). The systematically advantageous status of grammar school pupils thus implies that grammar schools are socially selective as well (McCulloch, 2015; Rasbash et al., 2010).

Additionally, when prior attainment is taken into account, the gap in grammar school opportunities persists (Harris & Rose, 2013). For example, in 2011, only 40% of high-achieving FSM pupils entered grammar schools, but the proportion of non-FSM pupils was 60% in selective LAs (Cribb et al., 2013). In Kent, amongst FSM pupils who achieved Level 5 or above in English and maths in 2015, only 51.4% of them attended grammar schools. Meanwhile, the corresponding rate for non-FSM pupils was 72.7% (Andrews, Hutchinson, & Johnes 2016). This is also the case for SEN pupils, who are less likely to attend grammar schools even when attainment is taken into consideration (Cribb et al., 2013). A similar discrepancy can be found among different ethnic groups (Bolton, 2017). While Chinese and Indian pupils are overrepresented in grammar schools, black pupils are often underrepresented (Andrews, Hutchinson & Johnes 2016). The unbalanced grammar school opportunities for pupils with equivalent performance further implies that prior attainment cannot fully explain the underrepresentation of certain social groups in grammar schools, which questions the fairness of the selection process of grammar schools.

Overall, the access to grammar schools is a critical area of inquiry. It reveals how different pupil groups might be benefited or disadvantaged if the grammar school effect truly exists. Besides providing evidence to grammar schools' role in promoting social equity, the answer is also valuable for its own sake, as it demonstrates whether the existence of grammar schools affects social coherence and integration.

8.3 The structure of this study's analysis

Based on the mixed evidence for grammar school effectiveness and equity in previous research, this project evaluates the potential impact of grammar school in these two aspects.

This study's analysis starts with access to grammar schools for different pupil groups. It reveals how grammar school opportunities are related to attainment, as well as other factors which should be irrelevant to the selection criteria, such as geographic location and family background. Following the unbalanced patterns of participation, possible explanations for the underrepresentation of certain disadvantaged groups in grammar schools are analysed.

The second part of the evaluation focuses on grammar schools' effect on pupils' later academic performance, both through traditional regression models such as OLS, as well as the less-common RDD approach. Although the RDD approach can make a robust causal inference on the relationship between grammar school attendance and later attainment, it is restricted by the 11+ data in England. Therefore, despite the limitations of traditional regression models, these models are still applied in this project, as the rich NPD data allows them to present a nation-wide analysis of the entire population. Apart from comparing the effectiveness between grammar schools and non-selective schools, the impact of the presence of grammar schools on local areas' overall academic standards is also examined. The effectiveness of selective LAs is then compared with that of non-selective LAs.

After revealing the association between school type and pupils' academic performance, the HE participation pattern between grammar school pupils and pupils in non-selective schools is also compared. Both the general HE participation rate, and the opportunity of attending the Russell Group universities are compared between pupils in grammar schools and equivalent pupils in non-selective schools. The last part then combines the varied access to grammar school and the outcome of grammar school attendance, which reveals whether the link between family background and post-18 destination is stronger in selective LAs than in comprehensive areas in England.

Before presenting the statistical answers to these questions, the next chapter explains the detailed process of conducting the analysis in this study.

9 Research methods

After discussing previous research on the effectiveness and equity of grammar schools, this chapter explains how the analysis in this study was conducted, which includes the data applied and the choice of statistical approaches in each step.

9.1 The opportunity to attend grammar schools for different pupil groups

This section explains the process of evaluating the opportunity to attend grammar schools. In addition to the general patterns of access to grammar schools, the analysis also addresses the relationship between grammar school opportunities and pupils' prior attainment, geographical location, and family background.

9.1.1 Grammar school opportunities across LAs

Most of the data analysed in this study was acquired through the NPD, which collects annual performance and family background data for all pupils in England. To analyse access to grammar schools, the 2010/2011 cohort of KS2 pupils recorded in the NPD was selected (see Appendix 10 for a summary of cohort member and data resource of this study). Among the 612,027 pupil records for this year group, there are 186,461 pupils in 36 selective LAs. The analysis first shows how the difficulty of grammar school selection varies across LAs by comparing the KS2 performance of prospective grammar school pupils in each selective LA. This process includes 160,070 valid cases and excludes 26,391 (14%) cases with missing KS2 attainment data. Lacking national data for the 11+, the KS2 performance indicator used in this stage is pupils' English and maths results from the KS2 national test, with a total mark of 200 (100 for each subject). Pupils' KS2 science results are excluded not only because they are based on teacher assessment (which is less consistent across schools and LAs), but also because the 11+ usually includes English, verbal reasoning, numerical reasoning, and non-verbal reasoning, the contents of which have more direct links with English and maths than with science.

9.1.2 Grammar school opportunities for pupils moving across LAs

After revealing the difference in selection requirement across LAs, the analysis then pays attention to several subgroups. First, the analysis focuses on a small group of pupils whose home LAs are different from the LAs of their secondary schools. Their proportion in grammar schools and the probability of attending grammar schools are compared with their counterparts who stayed within their home LAs for secondary education. This is done both with and without

considering prior attainment. The process includes 169,691 valid cases and excludes 16,770 (9%) pupils with no valid LA information.

9.1.3 Grammar school opportunities for FSM, SEN and EAL pupils

The following step analyses the probability of attending grammar schools for three minority groups: pupils eligible for FSM, pupils with SEN school action plus or statement (SEN-PS) and pupils speaking English as an additional language (EAL). While the SEN School Action Plus code has been replaced by SEN Support, and the SEN Statement has been replaced by the Education, Health and Care Plans in September 2014 (Department for Education and Department of Health, 2015), the SEN code for this cohort still uses the older version. The analysis includes 168,023 valid cases with FSM data (18,438 missing), 186,461 with SEN data (no missing data), and 168,023 with EAL data (18,438 missing) in all selective LAs.

In the next part, KS2 attainment is also considered and only *high performers* in each LA are selected. Again, lacking national data for the 11+, the standard of *high performers* is set as pupils in each LA whose KS2 marks were higher than the lowest KS2 marks for grammar school pupils. This distinguishes potential grammar school candidates from the entire year group, as pupils who achieve this mark may attend grammar schools, while those who do not were given no opportunity in their LAs. Based on these standards, the analysis includes 103,558 valid cases with KS2 attainment data, and excludes 42 with missing FSM and EAL data, in all 36 LAs. Instead of comparing the probability of minority groups attending grammar schools to all of their peers in the same year group, this step only compares high performers in each selective LA.

9.1.4 The relationship between grammar school opportunities, pupil's prior attainment, geographical location and family background

In addition to snapshots of access to grammar schools, logistic regression models are used to evaluate whether grammar school opportunities can be explained by pupil's attainment, geographical location, and their family background. A logistic regression model predicts the probabilities of a binary outcome and provides the relative odds (e.g. probability of getting into a grammar school / probability of not getting into a grammar school). The most important outcome indicators in the model are 1) increase in the percentage correctness, which reveals how knowing certain sets of background variables increases the predictive ability of the model,

and 2) the Exp (B) of each baseline variable, which compares the odds of getting into grammar schools for one group of pupils with the odds for another pupil group, producing an odds ratio. For categorical independent variables, the Exp (B) compares the odds for each subgroup with the reference category. For numerical variables, it shows the changes in odds ratios with a one-unit increase in the independent variables. For the analysis in this step, only LAs with over 20% grammar school pupils are chosen, so as to make the base figure of predictive accuracy closer to 50%. Another reason for this choice is that in these LAs, attending grammar school is a common option for their pupils, rather than a rare route for a tiny minority. Therefore, the analysis is limited to 12 selective LAs and 55,831 pupils in total, which includes 45,048 valid cases with complete records for the baseline variables. The first logistic model considers pupils' personal backgrounds. The analysis includes dummy variables for girls contrasted with boys, pupils staying within their LAs during secondary education contrasted with those who move away; for FSM eligible pupils contrasted with non-FSM pupils; for SEN pupils contrasted with pupils with no SEN; and for each ethnic group contrasted with the majority white group. The recoded birth month, *Month Age*, converts children's birth months into ordinal numbers, with pupils born in August (the youngest) equalling 1 and those born in September equalling 12, thus accounting for the relative age within a year group. The second logistic regression model adds KS1 prior attainment (i.e., English and maths point scores). In the third model, KS1 attainment variables are replaced by KS2 attainment. The indicators for KS2 performance are the KS2 fine grades in English and maths, as they demonstrate more predictive power than the point scores (but the KS1 point score is unavailable). Lastly, the fourth model combines all of the aforementioned variables.

9.1.5 Possible explanations for the underrepresentation of certain pupil groups in grammar schools: The case of FSM pupils

Apart from focusing on the national pattern of the relationship between the opportunity to attend grammar schools and pupil's background characteristics, a local dataset of the 11+ is provided by a non-governmental group. This offers the opportunity to analyse possible explanations for the unbalanced grammar school chances across pupil groups. This 11+ data file has individual records for 2011/2012 KS2 local pupils who sat the 11+ in this selective LA in 2011. In addition to the 11+ results, the file also contains pupils' FSM status, IDACI, ethnicity, and KS2 level. Using this 11+ dataset, the proportion of pupils attending the selection, and the rate of passing the selection, are calculated for different pupil groups. Meanwhile, the

relationship between KS2 and 11+ results is also examined to assess whether the chances of passing the selection are fair after accounting for KS2 attainment level. The analysis in this part includes 7,917 valid cases out of the total 8,698 records.

After presenting the process of evaluating access to grammar schools, its relationship with pupil's background, and possible reasons why certain pupil groups are less likely to attend grammar schools, the next section presents the methods of grammar school effectiveness.

9.2 The effectiveness of grammar schools in improving pupils' academic performance

In addition to access to grammar schools, the academic outcome of attending grammar schools is also a major focus of this study. This section describes the process of evaluating grammar schools' effectiveness in improving pupil's KS4 academic performance.

The effectiveness of grammar schools in comparison with other mainstream state-funded schools in selective LAs is first analysed through traditional regression models controlling for pre-existing differences between pupil groups. As discussed in the literature review, one issue in regression models controlling for baseline differences is whether or not to include school composition variables. The consequences of including school compositional variables can be both positive and negative simultaneously. After evaluating the pros and cons, this study's conclusion is based on the results of models which include both pupil-level and school-level baseline variables, as the danger of omitting school-level variables may be greater. The estimation of grammar school effectiveness after accounting for pupil-level and school-level baseline variables is believed to provide a lower bound for its real effect. The following sections present the detailed information of applying linear regression models and logistic regression models to evaluate the effectiveness of grammar schools.

9.2.1 Linear regression models of the effectiveness of grammar schools

9.2.1.1 The national pattern and the patterns of individual LAs

The effectiveness of grammar schools in comparison with other mainstream state-funded schools in selective LAs is first conducted through OLS linear regression models. This analysis is conducted on the same 2010/2011 KS2 (2015/2016 KS4) cohort. Since the comparison is only relevant to mainstream state-funded schools, 20,344 pupils in independent schools and special schools in selective LAs are excluded from the model. Thus, the analysis includes

149,072 cases with valid records in selective LAs, out of a total of 186,461. Multi-stage linear regression models are applied to control for different sets of baseline variables at each stage. The first stage enters pupil background variables—gender, month age, IDACI, FSM eligibility, SEN-PS, EAL group, and ethnicity (converted to dummy variables in reference to white pupils). The second stage also includes pupils' KS2 total marks for English and maths. The next stage adds school-level variables, which are aggregated pupil-level variables from each secondary school. After testing all possible combinations of compositional variables, only two school-level variables influence the estimation results, which are average KS2 total mark and the proportion of FSM pupils in secondary school. Therefore, only these two school-level variables are added to the model. The final stage introduces school type into the model. This is a binary variable flagging grammar school attendance. The outcome of interest is pupil's 2015/2016 KS4 attainment. There are three widely used GCSE indicators—total GCSE, capped GCSE and average GCSE point scores (all include equivalents). For comparison, the analysis presents the results for all three GCSE outcome variables at the beginning. However, to avoid repetition, it only gives the results of capped GCSE at later parts. In addition to overall GCSE performance, the analysis also separates the two most fundamental GCSE subjects—English and maths. According to the Secondary Accountability Measures (DfE, 2018), the highest point score for a GCSE subject in 2016 was 8 (A*), and the interval between each grade is 1 point score. This means that the highest possible capped GCSE point score for each individual is 64. However, there were pupils whose GCSE results surpassed this. A similar situation has been encountered by Coe and his colleagues (2008) when using NPD for grammar school effectiveness evaluation. They have noted that comparisons between schools are not affected by calculation violations, since the point score scale is consistent for all (p. 200).

Apart from the national picture of the effectiveness of grammar schools, the analysis also presents the patterns of individual LAs. Furthermore, ML regression models are applied as the supplementary approach to test the stability of the estimation results. The analysis thus includes results from fixed slope ML models and random slope ML models (the slope of each school varies as a function of KS2 attainment). These models are conducted for the same group of pupils, using the same sets of baseline and outcome variables from the OLS models.

9.2.1.2 School effectiveness and the degree of selectivity

In addition to the general effectiveness of grammar schools, the relationship between their effectiveness and the degree of selectivity of each LA are evaluated using data from the same cohort (2010/20111 KS2). The indicators of the selectivity of each LA are 1) the lowest KS2 mark for grammar school pupils, and 2) the proportion of grammar school places. While the first indicator presents the selection difficulty for grammar schools in each LA, this indicator is less stable when the number of grammar school pupils is small. Therefore, the proportion of grammar school places is used as a complementary indicator, revealing the unbalanced opportunity of getting into grammar schools in each LA. The correlation figures between the grammar school coefficients and these two selectivity indicators are calculated.

Apart from the estimation results based on linear regression models, the following section discusses the use of logistic regression models to evaluate grammar school effectiveness. Evidence from both types of models is helpful in comparing whether the conclusions are consistent for different statistical approaches.

9.2.2 Logistic regression models of the effectiveness of grammar schools

This section details the process of using logistic regression models to compare the effectiveness of grammar schools with that of non-selective schools in their local areas, focusing on whether attending grammar schools is positively predictive of achieving certain GCSE levels at KS4. The analysis is still conducted on the 2010/2011 KS2 cohort in selective LAs. Four sets of explanatory variables—KS2 pupil background, KS2 pupil attainment, school compositional characteristics, and school type—are added to the multi-stage logistic regression models, which are the same as those in the multi-stage linear regression models. Unlike the OLS models, which can set capped GCSE as the outcome variable to control for the total number of exams entered, the outcome variables in logistic regression models cannot present this result. Therefore, the number of exams entered is also added into the logistic regression models at a later stage to assess how much it explains the difference between grammar schools and non-selective schools.

Two binary outcome variables are applied in the logistic regression models to flag whether pupils achieved good and high levels at KS4. The indicator of good KS4 results is achieving 5 or more GCSE and equivalent qualifications at grades A*-C (including English and maths).

The indicator of high KS4 results is achieving 5 or more GCSE and equivalent qualifications at A*-A grades. The logistic regression models for 5 A*-C levels include all 149,072 valid cases in selective LAs. This is the same number as in the OLS regression models. The models for achieving 5 A*-A grades only include pupils who scored higher than the KS2 median in selective LAs (134 total marks). In this way, the model's base figure is closer to 50%. As a result, only 70,683 valid cases are included in the models of 5 A*-A results. While the process excludes many valid cases, a parallel analysis conducted on the whole group produces similar results.

After introducing the process of applying traditional regression models which control for pre-existing differences between pupil groups to evaluate the effectiveness of grammar schools in improving pupils' academic performance at KS4, the next section details the use of RDD, which is a stronger research design to make causal inferences.

9.2.3 Regression discontinuity analysis of the effectiveness of grammar schools

9.2.3.1 The process of selecting pupils into grammar schools in the chosen LA

Before describing the detailed steps of conducting the RDD, this section first introduces the process of grammar school selection. As mentioned previously, the 11+ data applied in this study only covers one single LA due to the lack of national data. Therefore, the RDD is only conducted to this LA where the overall proportion of grammar school places is high.

In this participating LA, grammar schools apply an opt-in selection system. Only pupils whose parents have registered them for the test are allowed to take the selection test. Like all the other selective LAs, the eligibility for grammar school attendance in this LA is primarily decided by the selection test (the 11+), which is held in the last year of primary school. In 2011, the 11+ in this LA included three subjects. A full mark for each subject was 140, adding up to 420 in total. In order to be qualified to attend grammar schools, pupils in this LA not only need to cross the threshold in total score (360), but also pass the minimum requirement of each individual subject. Apart from the formal test, head teachers of primary schools can appeal through the Head Teacher Panel if they are not satisfied with their pupils' test results. In this case, extra supporting materials are evaluated.

The following sections start to discuss statistical steps of conducting the RDD to evaluate the effectiveness of grammar schools in this participating LA.

9.2.3.2 The theoretical framework of RDD

The underlying logic of making a causal inference is to provide counterfactual results of what would have happened to the same person without the treatment (Shadish, Cook, & Campbell, 2002). In an RCT, since participants are randomly assigned to the treatment and control group, baseline characteristics are also randomly spread across the two groups. Therefore, any difference in outcome is due to the treatment. The function of an RDD is similar. The basis of an RDD is that participants are allocated to either the treatment or the control group according to the cut-off point of a continuous assignment variable. Only those who reach the cut-off point are given the treatment. If participants' assignment variables could not be manipulated with precision, their chances of just making it or just missing it can be regarded as locally random (Lee & Lemieux, 2009). As the values of the assignment variable are similar among participants in the neighbourhood of the cut-off point, a comparison of the outcome variable between the treatment and control group can attribute any discontinuity at the cut-off point to the treatment (Lee & Lemieux, 2010). This process provides perfect counterfactual results and solves the problem of pre-existing differences between the treatment and control group.

In the ideal 'sharp' RDD, all individuals who have passed the cut-off point would get the treatment and those who have missed it would not. However, in reality, it is more common to encounter programmes with imperfect compliance and programmes in which the eligibility to get the treatment is not decided by one assignment factor alone. This means an individual who reaches the threshold may not get the treatment ('no shows'), while one who does not reach the threshold may in fact get it ('crossovers'). For example, there might be some pupils who did not achieve the passing score on the 11+, but still attended grammar schools. On the contrary, it is also reasonable that not all pupils who passed the selection attended grammar schools. This is similar to an RCT with imperfect compliance (Lee & Lemieux, 2009, p. 23). These situations are categorised as 'fuzzy' RDDs, in which the treatment is not perfectly decided by the assignment variable, and we will encounter some cases which violate the assignment rule, especially near the cut-off point (Trochim, 1984). The rate of incompliance thus needs to be considered when calculating the treatment effect.

The selection process for grammar schools in our sample LA is a typical 'fuzzy' RDD in which the assignment variable (total score on the 11+) is not the only factor deciding grammar schools' eligibility. Pupils' test scores on the three individual subjects of the 11+, as well as the results of the Head Teacher Panel, also influence eligibility. Therefore, a 'fuzzy' RDD is applied to estimate the treatment effect.

9.2.3.3 Empirical strategy of RDD

According to the definition of the treatment effect in a 'fuzzy' RDD (Jacob et al., 2012; Lee & Lemieux, 2009), the estimation can be written as:

$$Y_i = \alpha + \beta T_i + f(X_i) + u_i, \quad (1)$$

$$T_i = \gamma + \delta D_i + g(X_i) + v_i, \quad (2)$$

where Y_i is the outcome measure for each individual i; T_i is the treatment dummy; X_i is the assignment variable (X_i=0 is the cut-off point); D_i is the binary indicator of whether individual i reached the cut-off point (D_i=1 if $X_i \geq 0$); u_i and v_i are the random error for each individual. The effect of attending grammar schools which needs to be estimated equals β. To make it easier to understand, these two equations can be simplified as:

KS4 performance = grammar school effect * grammar school or not +
the effect of prior attainment,
Grammar school or not = compliance rate * passed threshold or not +
the effect of prior attainment

The treatment effect in the 'fuzzy' RDD revealed in equation (1) and (2) is consistent with a standard instrumental variable setting, and thus it can be estimated using a Two-Stage Least Squares (2SLS) model (Hahn, Todd, & Van der Klaauw, 2001). The parametric approach involves finding appropriate regression lines to fit data points. A correct estimation thus requires accurately modelling the relationship between KS4 performance and prior attainment ($f(X_i)$), and the relationship between 'grammar school or not' and prior attainment ($g(X_i)$). For example, if the relationship between prior and later attainment can be graphically presented as a straight line, then a linear functional form (i.e., Y=a+ bX) can be used as $f(X_i)$. However, although it is a widespread practice to use a linear function to depict the relationship between prior and later attainment, the actual relationship between these two variables may be a curve line, as it could be harder to make equivalent progress at a high level than at lower ones. Therefore, quadratic function forms (i.e., $Y=a+bX+cX^2$) are also fitted in this study to avoid

misspecification. Meanwhile, the slopes of the regression lines are also allowed to vary on two sides of the cut-off point. To comply with the calculation rule of the 2SLS analysis, the same type of regression line (functional form) is used for both equations (Jacob et al., 2012; Lee & Lemieux, 2009). More details of identifying the treatment effect are attached in Appendix 2.

Apart from the parametric approach, which finds regression lines to fit data, the estimation of the treatment effect can also be realised through the non-parametric approach, which selects data to fit regression lines. The non-parametric approach applies local linear regression to depict the relationship between explanatory and outcome variables. While the overall pattern between these variables may not be linear, if we only select data points within a small range, it is likely to see a linear relationship (Hahn, Todd, & van der Klaauw, 2001). Therefore, unlike the parametric approach which makes an estimation based on all the data, the non-parametric approach only uses data within a limited range. Since the estimation of interest in this study is at the cut-off point, data should also be selected on both sides of the cut-off point. The range of selected data on each side of the cut-off point is also referred to as a 'bandwidth', and an accurate estimation heavily depends on choosing a right bandwidth. Instead of using visual inspection, the optimal data bandwidth is calculated according to the data-driven algorithm proposed by Imbens and Kalyanaraman (2012). To avoid redundancy, the same bandwidth is used on both sides of the cut-off point, and in equations (1) and (2) (Imbens & Lemieux, 2008). Despite the different calculation processes, the non-parametric estimate should be similar to the estimate in the parametric approach. It is thus used as the complimentary approach to the parametric estimation.

In this study, the treatment effect of interest is the effectiveness of grammar schools compared with other non-selective mainstream state secondary schools. The outcome variable used as the indicator of school effectiveness is the capped GCSE point score, which was given five years after the grammar school selection test. Capped GCSE point score instead of the total GCSE point score is used due to the larger number of tests taken by grammar school pupils than their counterparts in non-selective schools. Using the total GCSE point score thus may upwardly bias the performance of grammar school pupils. According to the Secondary Accountability Measures (DfE, 2018), the highest point score for a GCSE subject in 2017 was 8.5 (A*). The interval between each grade is 1.5 point score for A*-C grades and 1 point score for C-E grades. This means the highest possible capped GCSE point score for each individual is 68. Again, there are pupils in the sample who achieved GCSE results higher than this, and the maximum

had point scores of 71.25. But the comparisons between schools would not be affected based on the previously-stated reason. As the total test score on the 11+ is the major factor deciding pupils' eligibility to attend grammar schools, it is centered at the lowest passing score and set as the assignment variable (point 0 is the cut-off point). The value of the assignment variable ranges from -140 to 60, but there are only about 10% of pupils scored lower than -60. An important premise of a valid RD design is participants' inability to precisely control the assignment variable (Lee & Lemieux, 2009). This condition can be easily met in the 11+. As the passing score of the 11+ may change each year, pupils do not know the exact cut-off point when attending the test. While pupils have some influences on the test score, they are unable to accurately manipulate it.

Based on the principle of the RDD, baseline covariates are believed to be randomly distributed in the treatment and control group near the cut-off point. Thus, there is no need to control for these variables. However, the regression estimates between models with and without baseline variables are still compared to evaluate the internal validity of the design, as theoretically both types of models should yield similar results. A robustness check is also conducted by trimming the 10% outermost observations at both ends of the assignment variable. The process thus excludes data points above 57 or below -77 in the assignment variable.

9.2.3.4 The data set of RDD

After introducing the detailed process of empirically identifying the treatment effect in the RDD, this section discusses how the 11+ data is dealt with.

The underlying assumption of the RDD requires the data of the assignment variable. However, absent from the NPD and all the other major databases in England, the national result of the 11+ is not publicly available, as mentioned before. The analysis in this section thus still applies the 11+ file provided by a local group. This 11+ file has 7,917 valid cases of local pupils who sat the 11+ in this LA in 2011 (2011/2012 KS2 cohort). It contains the 11+ test data which include test score for each subject, whether a pupil has been entered in the Head Teacher Panel, and the result of the selection. It also keeps a record of pupils' background variables, including FSM status, ethnicity, IDACI, and KS2 level. While the 11+ file tells whether a pupil passed the selection, it provides no information on actual attendance. A comparison between the NPD and the 11+ file demonstrates that the total number of local pupils in grammar schools in the

NPD is close to the number of local pupils in this LA who passed the selection as recorded in the 11+ file, with an attrition rate below 3%. This is a small number compared to the overall effect size as shown in the findings chapters. Therefore, the selection result in the 11+ file is used as the indicator of actual participation in grammar schools.

Lacking any record of academic performance at later stages, the 11+ file is linked to the NPD data of the same 2011/2012 KS2 cohort for the 2016/2017 GCSE results. However, since the 11+ data is anonymous (without any form of identifier), the 11+ file and the NPD data extract are matched through family backgrounds and KS2 attainment. While FSM status, ethnicity, KS2 point score and school types can be exactly matched between the two files, IDACI scores are slightly different in the 11+ file and the NPD, which is thus matched with a 0.01 tolerance rate (IDACI scores can be matched within the interval of 0.01). Another problem which occurred in the matching process is duplicate cases of pupils who share the same combinations of all the available demographic and attainment variables. In order to make one-to-one unique matches between the two files, these duplicate cases are deleted. This process excludes 52% (4,119) of the total samples in the 11+ file. While this process might threaten the representativeness of the sample, it is the best available option due to the limited information in the 11+ file (alternative sampling strategies in Appendix 3). After data clearing, 2,628 valid cases in the 11+ file are matched to their NPD records, and 2,541 cases in the mainstream state-funded schools are kept for the RDD analysis. Unlike the national data, the small pupil number of this sample group means that when figures are draw, there would be some data points representing fewer than 5 cases. For privacy reasons, these data points are not presented in all of the figures in the RDD analysis part.

Overall, after carefully dealing with the imperfect 11+ data, the analysis is finally able to present the estimated treatment effect in the RDD. It should be noted that typical RDDs do not involve matching. The complicated process of matching pupils' prior attainment with later performance in this study is a result of the limited 11+ data in England. Cases omitted during the matching process imply that the estimation is not definitive, and the results are more about the feasibility of the RDD approach in causally evaluating the effectiveness of grammar schools. While all the previous sections discussed so far pay attention to the general effect of grammar schools, the next section starts to describe the process of evaluating whether grammar schools are especially effective for disadvantaged pupils.

9.3 Differential effectiveness of grammar schools for FSM pupils

Unlike previous sections which have evaluated the effectiveness of grammar schools for all pupils, this section focuses on a sub-group—pupils eligible for FSM at KS2. The analysis of the differential effectiveness of grammar schools for FSM pupils is presented first through the comparison of raw performance. Then, separate OLS models analysing FSM and non-FSM pupils are presented. The results are supplemented by OLS models using the interaction term, FSM*Grammar School, which flags the difference between FSM and non-FSM pupils in the two types of schools. In addition to the OLS models, logistic regression models are also applied. All models in this section use the same data from the NPD 2010/2011 KS2 cohort, as was used in previous OLS and logistic regression models of grammar school effectiveness. The baseline and outcome variables are also the same as in the general models of grammar school effectiveness.

So far, all the steps of comparing the effectiveness of grammar schools with non-selective schools in terms of improving pupils' academic performance have been presented. After using different approaches to assess the effectiveness of grammar schools and non-selective state-funded schools, the next section presents the process of evaluating the effectiveness of selective and non-selective LAs, focusing on the influence of the presence of grammar schools on the overall academic performance of local areas.

9.4 The effectiveness of selective LAs in improving pupils' academic performance

The existence of academically selective schools not only matters to their own pupils, but also to surrounding schools. Therefore, the evaluation of grammar school effectiveness in comparison with non-selective schools is only a partial answer. In order to assess the influence of the presence of grammar schools on the overall academic standards of broad areas, the effectiveness of selective LAs is compared with that of non-selective LAs.

According to the original design of the selective system in the 1940s, pupils who did not pass the selection test of grammar schools were allocated to secondary modern schools which offered training in basic subjects. In order to assess the impact of the selective system, pupils in grammar schools and secondary modern schools needed to be compared with those in comprehensive schools. In this way, the difference between the two systems could be presented. However, after the comprehensivisation reform, the difference between secondary modern

schools and comprehensive schools has become unclear. Courses taught in contemporary secondary modern schools are the same as those of comprehensive schools. The intakes in both types of schools in selective LAs are also converging, both admitting pupils who did not pass the 11+ (Allen, 2016). Therefore, school name is no longer an accurate reflection of the difference between secondary modern schools and comprehensive schools (Bolton, 2017). However, as the school identifier provided in the NPD is anonymised, it is impossible to know the real status of each school. Therefore, instead of distinguishing the difference at the school level, the overall effectiveness of selective LAs is compared with that of non-selective LAs. While it is possible that pupils in non-selective LAs may travel to selective LAs to attend grammar schools, this comparison would still be relevant to the vast majority.

The effectiveness of selective LAs in contrast with non-selective LAs is evaluated through OLS linear and logistic regression models. The analysis is still conducted on the 2010/2011 KS2 cohort. Instead of focusing on selective LAs, the analysis now includes all pupils in England of the same year group. Out of the total 612,027 cases within the whole cohort, there are 539,610 cases in mainstream state-funded schools (excluding special schools and independent schools). Among these pupils, 481,681 cases have complete records for all of the variables, and are thus included for analysis. While in the logistic regression model of 5 A*-C results, all valid cases are entered, the logistic regression model for 5 A*-A is still conducted on pupils with high KS2 marks (over 134) to make the base figure of predictive accuracy closer to 50%. This includes 216,099 valid cases. The indicator for the type of LA is a binary variable distinguishing 36 selective LAs from the rest. A categorical school type variable is also created, which includes grammar schools, non-selective schools in selective LAs, and schools in comprehensive LAs. This variable presents the effectiveness of grammar schools and non-selective schools in selective LAs against schools in comprehensive LAs. All of the other variables in the OLS and logistic models are the same as those in the previous models of grammar schools' effectiveness.

Following the detailed steps as described in this section, the estimation results would reveal the differences between selective and non-selective LAs. In addition to focusing on academic outcomes associated with the existence of grammar schools, the next section addresses the process of assessing whether grammar schools are more successful in improving their pupils' post-18 opportunities.

9.5 Grammar school attendance and participation in HE

This section discusses the methods of evaluating how grammar school attendance predicts the opportunity of HE participation. The analysis of HE participation applies the 2007/2008 KS2 NPD cohort, who finished KS4 in 2012/2013 and Key Stage 5 (KS5) in 2014/2015. The NPD cohort is then linked to the 2015/2016 Higher Education Statistics Agency (HESA) data. As this NPD data extract only includes pupils with valid KS5 records, it contains approximately 438,000 pupils, accounting for 70% of the total 633,000 cases in the same year. This means pupils who left schools immediately after the compulsory stage are not recorded in this data extract. While this excludes a substantial number of pupils, a parallel analysis was conducted (by my supervision team) using the whole KS2 cohort, and the results are consistent with the conclusions of this data extract. Furthermore, it needs to be noted that the NPD data for this cohort is only linked to the 2015/2016 HESA data. This means the analysis excludes those who did not start post-18 education immediately after finishing lower levels, but did choose to attend HE institutions in subsequent years. These are usually pupils from the most advantaged families (Hammer, 2003). However, as the proportion of returning pupils is low, which is only around 4% according to Raffe et al. (2001), the omission would not exert substantial influences on the general pattern. While the decision to omit returning pupils was made based on their small number, it is also due to data application time limits. The NPD-HESA linked data for the 2016/2017 academic year was not available online until April 2018, and the application for the linked data is long and complicated (taking over a year for the 2015/2016 one).

After introducing the dataset, the evaluation process for HE participation between grammar schools and other mainstream state-funded schools in selective LAs is described below. The analysis is conducted on all pupils who finished KS5 in 2014/2015 in 36 selective LAs. After deleting 15,327 pupils who were in special schools or independent schools during KS4 or KS5, 117,506 valid cases are included for analysis. Unlike in the logistic regression models predicting 5 A*-A GCSE grades, which only include pupils whose KS2 marks are higher than 134 to make the base figure closer to 50%, all cases are included in the models for HE participation and Russell group participation. This is done to make the comparison of the Exp (B) between the two model sets easier.

9.5.1 Explanatory variables

While some pupil characteristics are unlikely to change in different school years, such as gender, ethnicity, EAL and birth month, others may fluctuate. Therefore, in order to exclude the influence of changes in family backgrounds after KS2, the model may need to include several pupil-level background variables at KS4.

In order to assess which background variables are subject to change, and thus should be entered into the models repeatedly, the stability of the pupil characteristics is evaluated. For all pupils eligible for FSM in KS2, only 56.5% are still FSM eligible at KS4. This reveals that as an indicator of poverty, FSM eligibility is unstable. Unlike the FSM status, pupils' IDACI scores at different stages are similar, with a correlation of 0.82 between KS2 and KS4. Therefore, it may be unnecessary to include KS4 IDACI score once KS2 score is controlled for. Meanwhile, most pupils' SEN categories also remain unchanged. For all the pupils who have SEN statement at the end of KS2, over 90% still have the statement five years later at the end of KS4. Therefore, both KS2 and KS4 variables for FSM eligibility are entered into the model, and the other characteristics use the KS2 variable only, which are gender, ethnicity, EAL, birth month, IDACI and SEN provision. All background variables are coded the same way, as in previous GCSE performance logistic regression models.

In addition to the influence of pupil backgrounds, whether or not to include attainment after KS2 must also be decided. If only KS2 attainment is controlled for, the results would reveal whether attending grammar schools correlates with higher rates of HE participation for pupils with equivalent attainment at the end of primary school. When both KS2 and KS4 attainments are controlled for, the results present whether the HE participation pattern differs between the two school types, even accounting for KS4 attainment. Since both answers are meaningful, the analysis includes both results for comparison. The indicator of KS2 attainment is the total mark of English and maths, and the indicator of KS4 attainment is the capped GCSE (and equivalent) point scores. While the calculation rules for GCSE grades and point scores vary slightly between the 2016 and 2017 KS4 cohort, as described in previous sections, it needs to be noted that the calculation rule changed more dramatically in 2014. Therefore, the GCSE point score for this 2013 KS4 cohort in the analysis of HE participation (Chapter 14) is very different from that of the two cohorts in the analysis of grammar schools' effectiveness at KS4 attainment (Chapter 11-13). For pre-2014 GCSE results, grade G equals 16 point scores, and the interval between each grade is 6 point scores. This is much higher than the scales for 2016 and 2017.

After adding KS2 and K4 pupil-level variables, KS5 attainment variables are not added into the model because including A level results may be over-controlling. As the major deciding factor of the opportunity of HE participation, A level is too closely correlated with HE participation. This decision is also made due to the fact that pupils in grammar schools at the compulsory stage are not the same group as those in grammar schools at KS5. Among the pupils in grammar schools at KS4 (20,623), 89.6% of them (18,485) are still in grammar schools at KS5. Meanwhile, among all KS5 grammar school pupils (22,837), only 80.8% were also in grammar schools at KS4. Therefore, there are 6,526 pupils who had moved out or moved in during KS4 and KS5, which accounts for 35% of pupils who stayed in grammar schools during both stages. This shows the changes in composition in grammar schools between KS4 and KS5. As a result, controlling for KS5 attainment may create confusing results associated with grammar school attendance. Despite these reasons for excluding KS5 attainment, models with all variables from KS2 to KS5 are attached in Appendix 7 for comparison.

Similar to previous OLS and logistic models of school effectiveness, two school-level aggregated variables are also controlled for, in addition to the pupil-level baseline variables. These two school-level variables are still the mean KS2 total mark and the average proportion of FSM pupils at KS2 in each secondary school. Lastly and most importantly, school type is added into the model to present differences between the two types of schools. As the debate on grammar schools emphasises their lower secondary education instead of the sixth form, which leads to the discussion of the legitimation of early-age selection during the compulsory education stage, the indicator for grammar school attendance adopts the school type at KS4 instead of KS5. All the explanatory variables are added into the multi-stage logistic regression models chronologically. The first stage includes KS2 pupil-level variables. The second stage adds school-level variables. The third stage also controls for school type, and the last stage introduces KS4 pupil variables in addition to the above-mentioned baseline variables.

9.5.2 Outcome variables

The two binary outcome variables applied in the logistic regression models are 1) whether pupils have HE participation records, and 2) whether pupils attend the Russell Group universities. The information on whether pupils went to universities and the type of institutions they have attended is provided in the 2015/2016 HESA data extract. After matching the NPD

data with the HESA data, pupils with valid records in the HESA data are flagged as HE participants. Meanwhile, the institution names in the HESA dataset are used to identify pupils in the Russell Group universities.

Following the methods described in these sections, the relationship between grammar school attendance and participation in HE is revealed. The next section turns to the link between pupil's background and their post-18 destination.

9.6 The link between pupil's family background and post-18 destination

In addition to the connection between family backgrounds and access to grammar schools, and the relationship between grammar school attendance and later outcomes, this section focuses on the relationship between family background and pupil's post-18 destination in selective and non-selective LAs. The analysis is still conducted on the 2007/2008 KS2 (2014/2015 KS5) cohort in the NPD, which is linked to the 2015/2016 HESA data, just like the analysis of HE participation. The findings first present the correlation between pupils' attainment at KS2, KS4 and KS5, revealing how their early-age attainment connects with their performance five and seven years later. The analysis separates the patterns into selective and non-selective LAs. Comparing the two types of LAs reveals the difference between the selective system and the comprehensive system.

In addition to the correlation figures, the systematic connection between background and post-18 destination is also examined through multi-stage logistic regression models. After adding pupils' KS2, KS4, and KS5 background variables at each stage, the growth in the predictive accuracy of the model presents how family background and attainment at each stage explain pupils' later chances of attending HE institutions and the Russell Group universities. Thus, the comparison between selective and non-selective LAs provides evidence on whether the academic selection system during the compulsory stage is associated with a stronger connection between family background and pupil's future opportunity.

Overall, the analysis in this study is mainly based on data from the NPD and HESA, which includes valid cases from the entire population. The RDD approach also applies all eligible pupil records within a local area, without randomly selecting samples from the population. Therefore, discussion of issues such as significance test is not relevant to this study's estimation

results. After explaining the steps entailed in answering this study's research questions, the next chapter start to discuss the statistical results of these questions, following each step as described in this chapter.

10 Findings about the opportunity to attend grammar schools

This chapter presents the results of pupil's opportunities to attend grammar schools, which answer the first research question. It starts with the introduction of the selection process of grammar schools in England, the descriptive results of selective LAs, and demographic characteristics of pupils in selective LAs. It then reveals the unbalanced opportunity to attend grammar schools for different pupil groups and possible reasons under this pattern. Lastly, this chapter presents the systematic relationship between the opportunity to attend grammar schools and pupil's prior attainment, geographical location and family background, applying multi-stage logistic regression models.

10.1 The process of selecting pupils into grammar schools

In most selective LAs in England, the selection test of grammar schools (the 11+) is held towards the start of Year Six, which is last year of KS2. The time of the selection test is very close to the national assessment of KS2, which is usually held towards the end of Year Six in May. There is no national system of selecting pupils for grammar schools, and each selective LA (and sometimes individual school) has its own procedure. First, the test registration system is different geographically. Some selective LAs adopt an opt-out system, in which pupils are automatically entered for the 11+ unless their parents intentionally choose to withdraw. Other areas apply opt-in systems, in which only pupils whose parents have registered them for the test can attend the selection. Second, the format and content of the 11+ also varies from LA to LA, and the test is designed by several independent organisations (most notably CEM and GL Assessment) to suit the distinct need of each local area. In general, most of the 11+ tests focus on English, maths, verbal reasoning and non-verbal reasoning. The first two subjects are perceived as 'curriculum-aligned' due to their closer connection with the National Curriculum (Allen, Bartley & Nye, 2017). In contrast, pupils have limited exposure to relevant materials of reasoning tests in classrooms. These subjects are thus believed to be more closely related to the experience of private coaching (Allen, Bartley & Nye, 2017). The criticisms of coaching in rich families have forced test designers to increase the weighting of 'curriculum-aligned' subjects and lower the emphasis on reasoning subjects, which is one of the efforts to make assessments 'tutor-proof' (Allen, Bartley & Nye, 2017). Apart from formal written tests, some LAs also give head teachers in primary schools the right to appeal for the selection result and ask for a re-evaluation. This means in these LAs, pupils who did not reach the selection requirement may still have the chance to attend grammar schools. Overall, the process of

selecting pupils into grammar schools has substantial differences across LAs, and there are no universal state-wide practices in England.

10.2 The characteristics of selective LAs

In 2011, there were 163 grammar schools in England. They were located in 36 of 152 LAs, and educated about 5% of the pupils in England. Grammar schools are mostly concentrated in South East England, where the number of grammar schools accounts for about 1/3 of their total number in England. The largest number of grammar schools within a single LA can be found in Kent, with 32 grammar schools. Contrastingly, the smallest number of grammar school within a selective LA is only 1, which occurs in Cumbria, Liverpool, Kirklees, Stoke-on-Trent, Wolverhampton, Devon, and Enfield.

According to the classification standard by the UK government (Department for Environment, Food & Rural Affairs, 2014), rural authorities are areas with a more than 50% rural population. Based on this standard, among all the 36 selective LAs, only 5 of them are rural authorities. These 5 rural authorities are Devon, Cumbria, Wiltshire, North Yorkshire, and Lincolnshire. In the remaining LAs, the proportion of urban population exceeds rural population. There are 4 urban areas with a significant number of rural population (from 26%-49%), and all of the other selective LAs are dominated by an urban population with fewer than 26% people living in rural areas. Overall, the number of urban LAs accounts for 86% of the total number of selective LAs.

In terms of political preference, most of the selective LAs are now controlled by the Conservative Party, which is consistent with the historical reform of the comprehensive system since the 1960s. There are 117 grammar schools concentrated in 19 Conservative LAs, accounting for 72% of the total number of grammar schools. Only 9 selective LAs are governed by the Labour Party, where the total number of grammar schools is 27. In addition to LAs led by these two major parties, there are 7 selective LAs with no overall political control, with 14 grammar schools. There is also one selective LA led by the Liberal Democrats Party, and there are 5 grammar schools in the local area. The number of grammar school pupils is also the highest in the Conservative LAs. According to the 2011 NPD data extract, there were 16,408 grammar school pupils in the Conservative LAs, 3,541 in the Labour LAs, 1,907 in LAs with no overall control and 763 in the Liberal Democrats LA. Therefore, among all the 22,619

grammar school pupils, about 73% of them are educated in the Conservative LAs. Besides the large total number of grammar schools in the Conservative LAs, the average school number within each LA is also the highest in these LA. Each Conservative LA on average has 6 grammar schools, and the numbers in the Labour LAs and LAs with no overall control are only 3 and 2 respectively.

Overall, selective LAs in England have some distinct features in terms of geographical location, urban-rural categorisation, and political preference. According to the above description, grammar schools are more likely to be found in LAs in South East England, dominated by an urban population, and controlled by the Conservative Party.

10.3 Pupil characteristics in selective LAs

Besides presenting the features of selective LAs in England, pupil characteristics in these 36 LAs are also evaluated. As can be seen in Table 10.1, pupils in selective LAs have some differences to the national picture. First, pupils in selective LAs on average live in richer areas according to the IDACI, which is 0.22 compared with 0.24 in non-selective LAs. Consistently, the proportion of disadvantaged pupils is lower in selective LAs, such as pupils eligible for FSM, pupils with SEN-PS, and non-native English pupils. Furthermore, selective LAs have fewer pupils with missing data in the aforementioned aspects, who are believed to be more difficult to reach and more disadvantaged if they are educated within the state system (Gorard & See, 2013). This may result from that selective LAs have a lower proportion of independent school pupils (6.7%) than non-selective LAs (8.8%) do, and that selective LAs also have proportionally more trackable mainstream pupils.

Table 10.1: Characteristics of selective and non-selective LAs

	IDACI mean	**FSM percentage**	**SEN-PS percentage**	**EAL percentage**
Selective LAs	0.22	15.0	9.7	12.9
Non-selective LAs	0.24	16.5	10.0	13.4

The proportions of grammar school pupils in these 36 LAs differ, ranging from 1.4% to 37.4%. There are 18 LAs in which fewer than 10% of pupils attended grammar schools in 2011, including 10 LAs where the attendance rate was lower than 5%. There are also 2 LAs that selected more than 30% of their pupils into grammar schools in 2011. This shows that pupils

in different LAs have unbalanced opportunities to attend grammar school, even despite other factors. As presented in Table 10.2, LAs with different proportions of grammar school pupils are dissimilar in terms of local family income, and proportions of FSM, SEN-PS and EAL pupils.

Table 10.2: Proportion of grammar school places and characteristics of selective LAs

Percentage of grammar school places within the LA	**Number of LAs**	**IDACI mean**	**FSM percentage**	**SEN-PS percentage**	**EAL percentage**
Lower than 10	18	0.24	16.8	9.7	15.1
Between 10 to 20	6	0.19	12.3	7.9	9.2
Between 20 to 30	10	0.20	13.5	10.8	9.1
Between 30 to 40	2	0.15	7.3	8.1	13.2

In addition to the difference between selective LAs and non-selective LAs, the characteristics of pupils in grammar schools are also distinct to those in non-selective schools in selective LAs (Table 10.3). Grammar school pupils not only have higher KS2 total marks, but also have more advantaged family backgrounds as presented below. They usually come from richer areas, and they are very unlikely to be eligible for FSM. The proportion of SEN-PS pupils in grammar schools is also considerably lower than non-selective schools within the area. Since the selection of grammar schools is based on attainment, it is not surprising that pupil groups with lower performance have fewer grammar school opportunities. A later section will examine whether the underrepresentation of certain pupil groups can be fully explained by their prior attainment.

Table 10.3: Characteristics of grammar schools and non-selective schools in selective LAs

	IDACI mean	**FSM percentage**	**SEN-PS percentage**	**EAL percentage**	**KS2 total mark**
Grammar schools	0.13	3	1.3	17.6	165
Non-selective schools in selective LAs	0.20	13	8.2	13.8	122

10.4 Grammar school opportunities across LAs

As mentioned in the previous section, the proportion of grammar school places varies across LAs. This means the chance of attending grammar schools may also be very different from LA

to LA. This section thus reveals how the difficulty of the selection test of grammar schools in each LA varies, before considering all the other relevant factors.

In order to elucidate the difficulty of being accepted into grammar schools in each LA, the lowest KS2 marks for grammar school pupils are compared as presented in Figure 10.1. Selective LAs are sorted from left to right in ascending order of the proportion of grammar school places. Not surprisingly, the selection difficulty at each LA is very different, with the lowest mark ranging from 67 to 145. Based on the large variation, pupils in certain LAs may need to achieve more than twice the KS2 marks of those in others LAs to have any possibility of being admitted into grammar schools. In 2011, the national average total KS2 mark for English and maths was 126, with a bottom quartile mark of 103, and the highest quartile was 153. Contrasted with this national performance level, there are 12 selective LAs that admitted pupils in the bottom quartile of KS2 national performance into grammar schools. Meanwhile, 10 selective LAs did not admit any pupil with below-average KS2 results. Furthermore, in LAs with small proportions of grammar school places, high performance does not always guarantee a grammar school offer. Even the top-performing pupil may not have the opportunity to attend grammar school due to limits imposed by the overall proportion of available places.

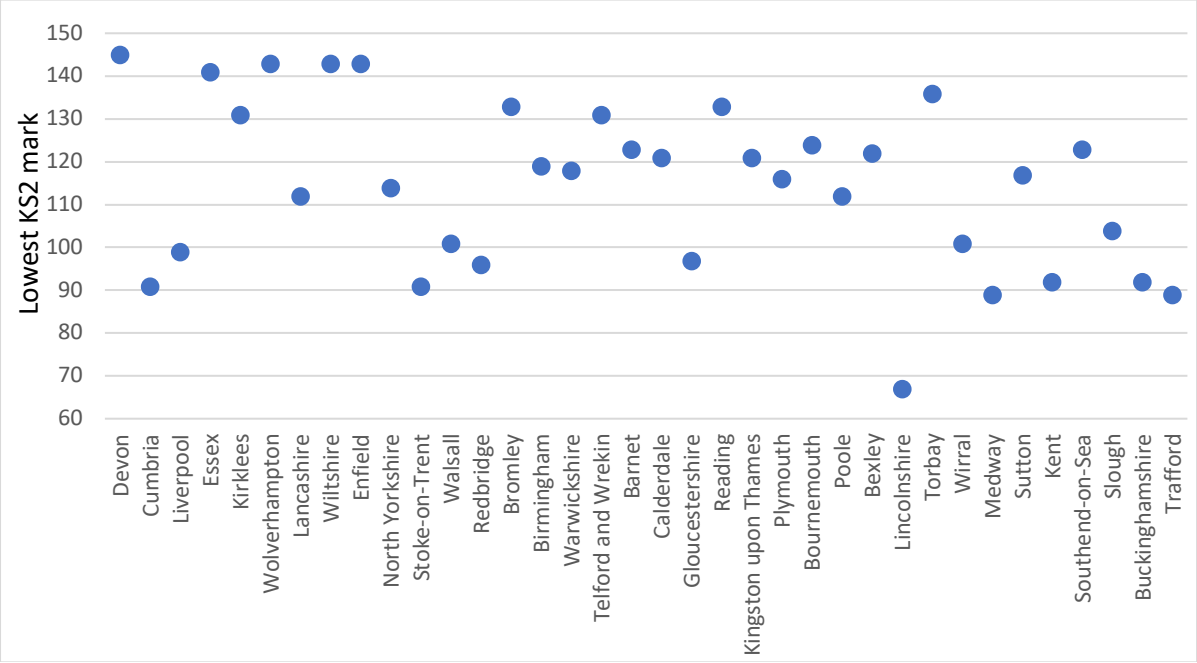

Figure 10.1: the lowest KS2 mark of grammar school pupils in each selective LA

10.5 Grammar school opportunities for pupils moving across LAs

As revealed by the KS2 marks of grammar school pupils, the difficulty of being accepted into grammar schools varies across LAs, depending on the provision of available school places, rather than a certain threshold of academic performance. The chance of attending grammar schools diverges when looking solely at where children live and where they apply for grammar schools. As changing the location of school application might influence grammar school opportunities, this section thus focuses on a small group of pupils who have moved outsides their home LAs for secondary education and examines the relationship between grammar school opportunities and relocation.

For all the secondary school pupils in selective LAs, only 9% of them moved outsides their home LAs for secondary education. However, the corresponding proportion in grammar schools is considerably higher, which reached 25.3%. Consistent with the high proportion of relocated pupils in grammar schools, pupils moving across LAs are also three times as likely to attend grammar schools as those who stay within the home LA. In order to see whether the difference is purely due to performance, the two groups' KS2 marks are compared. Based on the differentiated pattern of participation, it is not surprising that pupils who move outside the home LA have a higher average KS2 mark than those who stay within (142 vs. 126). However, when the probability of attending grammar schools for each KS2 mark is compared between the two groups, the higher average performance of pupils moving outside the LA can no longer explain their higher grammar school opportunities. As can be seen from Figure 10.2, pupils with low attainment (usually below 120) are not considered potential candidates for grammar schools in either case. The pattern of pupils at the right end is not stable, with large fluctuations, because there are very few cases in each KS2 mark above the point of 190. In all the other performance levels, pupils moved outside the home LA enjoy higher grammar school opportunities. The cleavage between the two groups is substantial, especially for pupils scoring between 150 and 190. Within this range, pupils who have moved across LA have in excess of 20%-30% higher probability of attending grammar schools than their counterparts do, and this is the performance range into which more than 80% of grammar school pupils fall.

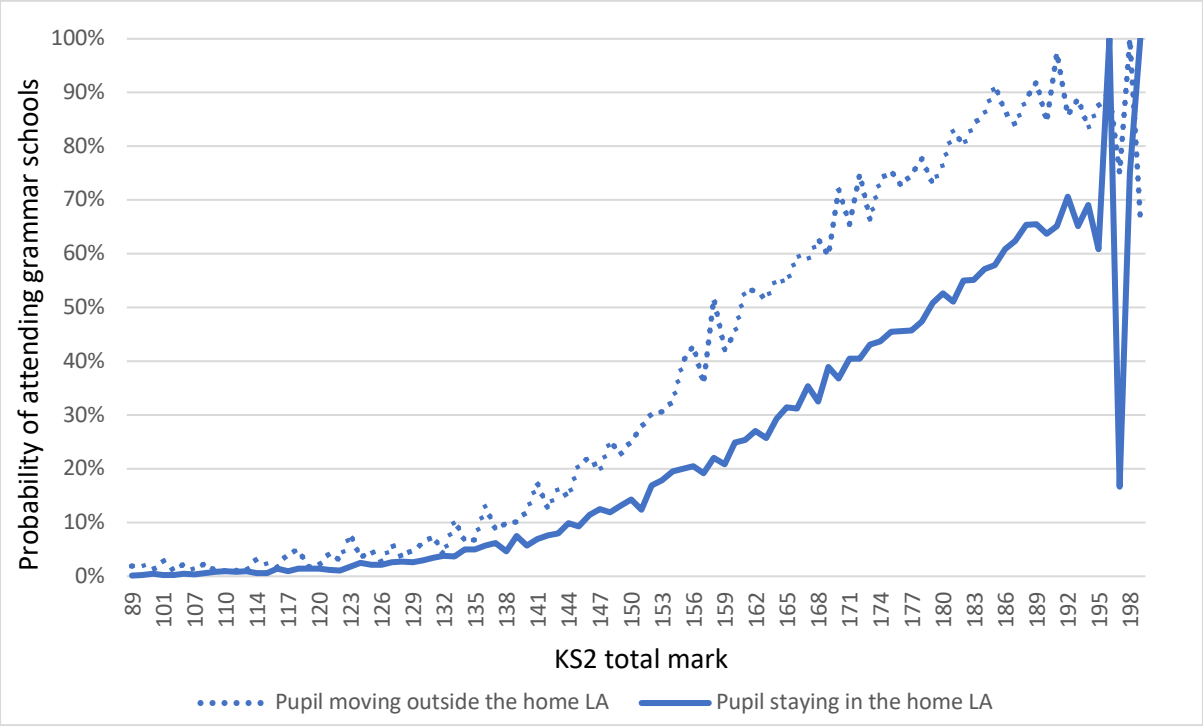

Figure 10.2: The opportunity to attend grammar schools for pupils moving across LAs

Since the clustered pattern of pupils crossing the LA boundary in grammar schools and their noticeably higher probability of attending grammar schools can hardly be explained by prior performance, the characteristics of this group are examined. After making the comparison, it is clear that pupils moving outside the home LAs for secondary education demonstrate systematic differences from those staying within (Panel A in Table 10.4). The former usually comes from slightly richer areas, as is revealed by the average IDACI (0.217 vs 0.221), and have many fewer FSM eligible and SEN-PS pupils. In terms of ethnicity, the relocated group has fewer white pupils but more Asian and black pupils proportionately.

Table 10.4: Characteristics of pupils moving across LAs and staying within home LAs

	KS2 mark	IDACI mean	FSM (%)	SEN-PS (%)	White pupils (%)	Asia pupils (%)	Black pupils (%)
Panel A: Selective LAs							
Moving across LAs	142	0.217	11.9	7.7	63.4	11.9	7.2
Staying within home LAs	126	0.221	17.1	10.8	76.4	8.8	3.2
Panel B: Non-selective LAs							
Moving across LAs	128	0.200	17.0	10.0	67.9	8.9	10.9
Staying within home LAs	124	0.200	17.8	10.7	79.1	9.2	5.1

In order to check whether it is a national pattern that pupils moving outside the home are more advantaged, the situation in non-selective LAs is examined. The overall proportion of relocated pupils in non-selective LAs is smaller, which is 8.4%. When the characteristics of relocated and non-relocated groups in comprehensive LAs are compared, the result surprisingly shows small differences between them (Panel B in Table 10.4). The proportions of FSM pupils and SEN-PS pupils were similar—the difference was less than 1%. The IDACI scores for the two groups were also the same. However, there were more ethnic minorities in the relocated group proportionately, similar to the pattern in selective LAs. Black pupils were still about twice as clustered in the relocated group, but Asian pupils were no longer overrepresented in the relocated group. Although there was still a 4-mark KS2 performance advantage in the relocated group, this comparison indicates the compatible SES and academic performance of the two groups in comprehensive LAs. Therefore, the advantaged background of pupils who move across LAs for secondary school is not a national pattern. It only exists in selective LAs.

Overall, pupils who moved outside home LAs for secondary education in selective LAs usually come from more advantaged families. They enjoy higher grammar school opportunities, which is beyond the explanation of their prior attainment. The unique feature of pupils crossing LA boundaries demonstrate that the unbalanced grammar school opportunities among LAs, combined with the freedom to move across LAs for grammar school places, has resulted in different access levels for pupils from different backgrounds. This has systematically benefited a group of more advantaged pupils.

10.6 Grammar school opportunities for FSM, SEN-PS and EAL pupils

Besides evaluating grammar school opportunities for pupils who have moved across LAs for secondary schools, in this section, three minority groups are examined: pupils known to be eligible for FSM, pupils who have SEN-PS, and EAL pupils. Since these groups have some distinct features, their patterns of attending grammar schools may also be very different from the majority.

The chance to attend grammar schools is relatively low for pupils eligible for FSM and those with SEN-PS. While the average probability of attending grammar schools in selective areas is 12.1%, it is only 2.4% for the FSM group and 1.5% for the SEN-PS group, which are considerably lower. Unlike these two groups, EAL pupils are overrepresented in grammar

schools despite their disadvantage in language. The probability of this group attending grammar school is also higher than that of their peers within the same year group, which reached 14.9%. However, based on the varied pupil characteristics within the EAL group, this overall pattern does not necessarily mean that each minority ethnicity all has an above-average opportunity to attend grammar schools. A more detailed in-group analysis will be conducted in a later section distinguishing pupils from different ethnic groups.

The overall trend is then confirmed by the following table which elucidates the systematic difference in the probability of attending grammar schools between three minority groups and all pupils in each LA. In Table 10.5, all selective LAs are sorted in ascending order according to their proportions of grammar school places, with the first LA containing the smallest proportion. The ratio difference of the probability between minority groups and all pupils in each LA is calculated (Ratio difference = Probability for each subgroup / Probability for all pupils).

For FSM and SEN-PS pupils, there is no exception to them having a lower probability of attending grammar schools than their peers in each LA. As the total proportion of grammar school pupils grows, the proportion of FSM and SEN-PS pupils also increases within the LA, yet the gaps remain obvious. Overall, the ratio differences between FSM and all pupils in all the selective LAs are smaller than 0.4, revealing that FSM pupils are not even half as likely to attend grammar schools as the average rate. The pattern for SEN-PS pupils is even worse, with their chances being further smaller than the FSM group. In addition to their low ratios in most areas, in 6 selective LAs, there are no SEN-PS pupils in grammar schools at all. Overall, the pattern in each LA reveals consistent disadvantages of FSM and SEN-PS pupils, regardless of geographical differences. In contrast, the EAL group demonstrates a more complicated pattern. In most selective LAs, EAL pupils have a higher probability of attending grammar schools than average pupils do. Exceptions exist in 8 LAs, where the EAL group has a lower probability, but only slightly. Therefore, despite the diverging patterns among LAs, the EAL group has an overall advantage in the probability of attending grammar schools. Unlike the dramatic underrepresentation of the FSM and the SEN-PS group, the advantage of the EAL group is relatively mild.

Table 10.5: Ratio difference of probabilities of attending grammar schools (GS) between FSM, SEN-PS, EAL pupils and all pupils in each LA

Local authority (Proportion of GS places)	**Probability for FSM / Probability for all pupils**	**Probability for SEN-PS / Probability for all pupils**	**Probability for EAL / Probability for all pupils**
Devon (1.4)	0.07	No SEN-PS in GS	2.57
Cumbria (2.2)	0.09	0.09	1.59
Liverpool (2.4)	0.08	No SEN-PS in GS	2.33
Essex (2.8)	0.14	0.04	3.71
Kirklees (3.3)	0.18	0.15	1.48
Wolverhampton (3.9)	0.13	0.10	1.41
Lancashire (4.0)	0.10	0.15	0.85
Wiltshire (4.1)	0.05	No SEN-PS in GS	0.63
Enfield (4.8)	0.21	0.15	1.00
North Yorkshire (4.9)	0.20	0.31	2.86
Stoke-on-Trent (5.0)	0.10	0.20	0.60
Walsall (5.3)	0.21	0.15	1.94
Redbridge (6.4)	0.31	No SEN-PS in GS	1.44
Bromley (6.7)	0.10	0.13	2.03
Birmingham (7.3)	0.23	0.07	0.82
Warwickshire (7.3)	0.08	0.07	1.19
Telford and Wrekin (7.8)	0.09	No SEN-PS in GS	1.28
Barnet (8.4)	0.13	0.19	1.45
Calderdale (11.4)	0.25	0.16	0.80
Gloucestershire (11.4)	0.20	0.05	2.24
Reading (14)	0.10	0.19	1.59
Kingston upon Thames (14.2)	0.20	0.04	1.92
Plymouth (14.4)	0.29	0.04	1.61
Bournemouth (17.1)	0.12	No SEN-PS in GS	0.85
Poole (20.8)	0.10	0.10	0.82
Bexley (22.7)	0.37	0.21	1.39
Lincolnshire (22.7)	0.26	0.15	0.84
Torbay (24.9)	0.24	0.11	1.10
Wirral (25.8)	0.20	0.04	1.28
Medway (25.9)	0.33	0.22	1.33

Sutton (26.3)	0.35	0.08	1.96
Kent (26.3)	0.24	0.16	1.51
Southend-on-Sea (26.9)	0.15	0.07	1.50
Slough (29.9)	0.30	0.10	1.13
Buckinghamshire (34.9)	0.21	0.11	1.14
Trafford (37.4)	0.20	0.05	1.25

10.7 Grammar school opportunities for high performing FSM, SEN-PS and EAL pupils

As grammar schools select their pupils based on attainment, which is correlated with pupil backgrounds, it might not be surprising that there are fewer disadvantaged pupils and more advantaged ones in grammar schools. However, if attainment is taken into consideration, are these three groups still disproportionately enrolled in grammar schools?

To evaluate whether performance can explain the unbalanced patterns of participation of these three sub-groups, KS2 attainment is also taken into consideration. Instead of including all the pupils for analysis, this section thus only includes high-performing pupils in each LA. The difference in the probability between high performing minority groups and all high performing pupils attending grammar schools in each LA is calculated (Ratio difference = Probability for high performers in each subgroup / Probability for all high performers).

Overall, as can be seen in Table 10.6, FSM and SEN-PS pupils still are less likely to attend grammar schools, but the gap is lower than that of Table 10.5. The rate for FSM high performers is less than half the rate for all high performers in 33 LAs, with the smallest rate being 0.09. For SEN-PS high performers, although they were also underrepresented, the gap is smaller than that of the FSM group (except 6 LAs where no SEN-PS pupils attended grammar school in 2011). Alongside these 36 LAs, there are 18 LAs where SEN-PS high performers still have lower probabilities of attending grammar schools as all high performers do. But exceptions can be found in Wolverhampton and North Yorkshire, where the SEN-PS group has slightly higher grammar school opportunities than the average rate within the area. Therefore, although in the previous section, SEN-PS are less likely to attend grammar schools than FSM pupils, the patterns are reversed in most LAs when attainment is taken into account. However, due to the contradiction between the definition of 'SEN' and the actual high-performance of pupils in this group, high-performing SEN-PS pupils may be systematically different from others.

Table 10.6: Ratio difference of probabilities of attending grammar schools (GS) between FSM, SEN-PS, EAL high performers and all high performers in each LA

Local authority	**Probability for FSM / Probability for all (high performers)**	**Probability for SEN-PS / Probability for all (high performers)**	**Probability for EAL / Probability for all (high performers)**
Devon	0.18	No SEN-PS in GS	3.25
Cumbria	0.16	0.19	1.88
Liverpool	0.09	No SEN-PS in GS	2.41
Essex	0.34	0.29	3.92
Kirklees	0.32	0.96	1.64
Wolverhampton	0.25	1.25	1.40
Lancashire	0.16	0.51	1.05
Wiltshire	0.18	No SEN-PS in GS	0.77
Enfield	0.35	0.55	1.10
North Yorkshire	0.29	1.16	2.79
Stoke-on-Trent	0.13	0.61	0.66
Walsall	0.28	0.54	2.16
Redbridge	0.36	No SEN-PS in GS	1.34
Bromley	0.19	0.65	2.06
Birmingham	0.31	0.37	0.85
Warwickshire	0.13	0.28	1.27
Telford and Wrekin	0.17	No SEN-PS in GS	2.41
Barnet	0.17	0.53	1.37
Calderdale	0.41	0.49	1.12
Gloucestershire	0.27	0.11	2.18
Reading	0.23	0.79	1.61
Kingston upon Thames	0.34	0.23	1.79
Plymouth	0.43	0.13	1.61
Bournemouth	0.22	No SEN-PS in GS	1.11
Poole	0.15	0.37	0.74
Bexley	0.56	0.66	1.31
Lincolnshire	0.30	0.27	0.98
Torbay	0.50	0.52	1.51
Wirral	0.26	0.15	1.51
Medway	0.42	0.43	1.30
Sutton	0.51	0.31	1.77
Kent	0.32	0.37	1.50
Southend-on-Sea	0.28	0.49	1.51

Slough	0.36	0.27	1.08
Buckinghamshire	0.29	0.30	1.24
Trafford	0.26	0.13	1.29

For EAL high performers, their probability of attending grammar school is still greater than the average rate in most LAs, and the gap is slightly larger than that of when attainment is not considered (Table 10.5). Furthermore, the number of LAs presenting a contradictory pattern also decreases from 8 to 5, with 31 LAs each presenting the same trend that the high-performing EAL pupils have a higher probability than the average rate of all high performers. Therefore, it seems that speaking a first language other than English is not a barrier for pupils in terms of academic performance during early-year education. On the contrary, EAL pupils enjoy more opportunities for grammar school education than native pupils do, regardless of whether attainment is taken into consideration.

10.8 Possible reasons for the underrepresentation of certain pupil groups in grammar schools: The case of FSM pupils

Previous sections have demonstrated that the probability of attending grammar school is low for disadvantaged pupils, even after taking attainment into account. Therefore, this section explores possible reasons why certain pupils are less likely to attend grammar schools, using FSM pupils as an example. As mentioned in the methods chapter, this section applies the 11+ data provided by a local group, and the following analysis only focuses on this LA.

10.8.1 The rate to take the grammar school selection test

FSM and non-FSM pupils' rates of taking the selection test are first analysed. While it is unsurprising that FSM pupils may have lower rates of taking the selection test due to their lower average performance, this section explores whether their underperformance explains the low attendance rates for the selection test.

According to 2011/2012 KS2 data for the LA chosen, among pupils who had taken the grammar school selection test, only 6.8% were eligible for FSM. This is less than half of the local proportion of FSM pupils. Unlike the majority non-FSM participants, this small group of FSM candidates has slightly worse KS2 performance. Meanwhile, the average total score on the 11+ is also lower among the FSM group than among non-FSM pupils—328 and 358, respectively. At first glance, the lower rate of FSM pupils taking the selection test is the result

of their inadequate academic performance. However, if the proportion of pupils taking the selection test is calculated by KS2 average point score, the participation rate for FSM pupils is still well below that of non-FSM pupils with equivalent KS2 performance.

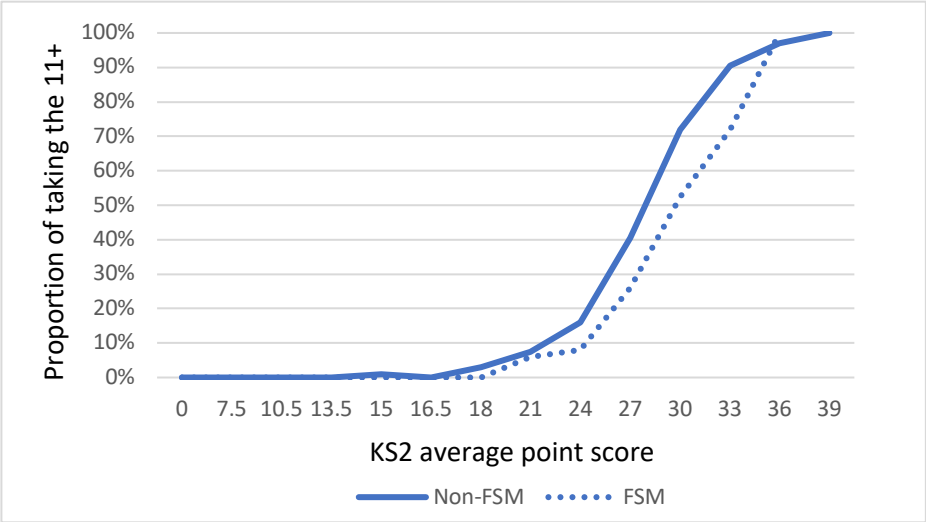

Figure 10.3: Proportion taking the 11+ for FSM and non-FSM pupils

As presented in Figure 1.3, for all the pupils with lower than 18 KS2 average point scores, participation rates are low, regardless of whether they are eligible for FSM. The rate grows for both groups at point 21, where the proportion of taking the 11+ reaches 7% for non-FSM pupils and 6% for FSM pupils. While the difference between the two groups is not obvious for pupils with low KS2 performance (who are usually not targeted by grammar schools), the FSM group's lower participation rate becomes more salient as the performance level rises. For non-FSM pupils with average point scores of 24 on the KS2 assessment, the proportion of taking the selection test is 15.5%, but the rate for FSM pupils with the same KS2 result is only 9%. Furthermore, although the participation rate for non-FSM pupils spiked at point 27, where 41% took the grammar school selection test, the proportion among the FSM group remains low, which is only 26%. The largest gap between the two groups is at point 30, where only half of the FSM pupils took the selection test, but 72% of non-FSM pupils did. At KS2 point score 33 (which is the performance level of half of the pupils in grammar schools), the gap between the two groups still remains salient, and the proportion for the FSM group is 17% lower than the non-FSM group. The only exception is at point 36, where the rate for the FSM group surpassed that of its counterparts. Since all FSM pupils with this high KS2 level sat the selection test, the participation rate is slightly higher than the one for non-FSM pupils (97%). This shows that FSM pupils with exceptional performance eventually enjoyed equal chances as their non-FSM

peers of taking the grammar school selection test. However, compared with the low rate of taking the selection test for FSM pupils at other KS2 levels, their advantage at this KS2 level is slight. Meanwhile, the proportion of pupils at this performance level is also low, which is 3.5% within the whole year group, and 15% for grammar schools. Due to the small number of pupils at this performance level, and the even smaller number of FSM pupils at this point, the FSM group's slightly higher rate at this point does not reverse their overall low rates of taking the selection test.

These results show that with the exception of a handful of pupils ranked at the top on KS2 performance, the rate of taking the selection test for grammar schools was higher for non-FSM pupils than for FSM peers, at most performance levels. The gap between the two groups is small among pupils with low attainment, but widens among pupils who have reached the potential selection threshold and are academically prepared for grammar school. The most pronounced disadvantage is for FSM pupils with 30 and 33 KS2 average point scores, which are the performance levels of more than 80% of grammar school pupils. The evaluation thus reveals that even for FSM pupils who have reached adequate performance levels, they are still less likely to take the selection test of grammar schools. This may be relevant to the opt-in test system adopted in this LA.

10.8.2 The rate to pass the grammar school selection test

Following FSM pupils' lower participation rates in the selection test, this section explains another possible reason for FSM pupils' underrepresentation in grammar schools, which evaluates whether the opportunity of passing the test differ between FSM and non-FSM pupils.

In order to elucidate pupils' chances of passing grammar school selection, the average success rate for pupils who attended the 11+ is calculated. Based on the higher average performance of non-FSM candidates, it is within expectations that they would have a higher rate of passing the selection than FSM pupils would. While FSM candidates' rate of passing the selection is only 17.4%, the rate for their non-FSM counterparts is 45.9%. When the passing rate at each KS2 level is calculated, the rate for FSM pupils is systematically lower, as shown in Figure 1.4. To present a clearer pattern, the figure only includes pupils who scored at least 27 point scores on the KS2, as it was impossible for those who had failed to reach this KS2 level to pass the selection test in this LA, regardless of their FSM status. At all the KS2 levels presented in

Figure 1.4, FSM pupils have lower passing rates than non-FSM pupils, and the gap is more obvious for pupils with high KS2 attainment. While the gap is only 16% for pupils who scored 30 on the KS2 test, it is 33% for those who scored 33. The passing rate for the non-FSM group is nearly 100% for those scoring 36 on the KS2, but the rate is still low for FSM pupils, only around 80%.

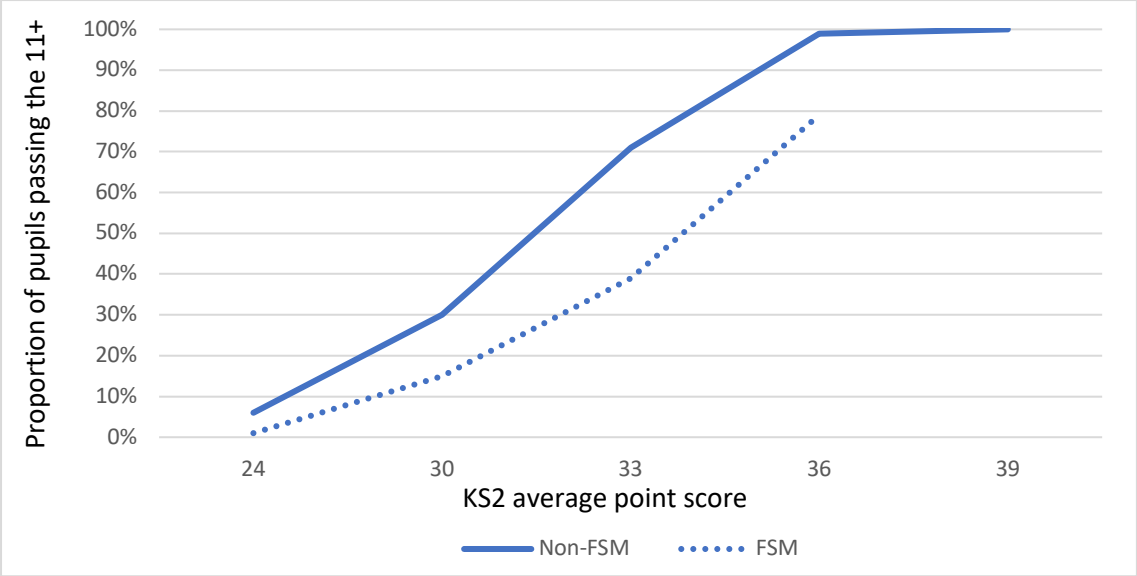

Figure 10.4: Proportion of pupils passing the 11+

The differentiated passing rate for FSM and non-FSM pupils with equivalent KS2 performance is largely a result of the fact that the KS2 national assessment and the 11+ do not correspond in the same way for FSM and non-FSM pupils. A detailed comparison between these two sets of tests reveals that at all KS2 performance levels, the 11+ scores are lower for FSM pupils (Figure 10.5). The gap is mild for pupils with low KS2 attainment, but is widened as the performance level rises. For pupils with KS2 average point scores from 24-27, the gap on the 11+ between FSM and non-FSM pupils is less than 15 (279 vs. 293, 308 vs. 321). However, for pupils with 33 and 36 average point scores on the KS2, FSM pupils have 23 scores lower on the 11+ (356 vs. 379, 386 vs. 409). The gap in the 11+ between FSM and non-FSM pupils after accounting for KS2 point scores might be a consequence of the extra coaching in more affluent families, as paying for private coaching is not equally affordable to FSM pupils. This is also likely a result of the different content between the 11+ and the KS2 assessment, which are not designed to measure the same aspect of 'ability'. It may also be attributable to bias on the test, which systematically favours pupils with certain background characteristics. Since there is no public data on the validity and reliability of the 11+ in England, it is impossible to

determine which test is a fairer measurement of pupil's early-age attainment at the end of primary school. However, the pattern in this section at least reveals that the 11+ assesses FSM pupils as less able than the KS2 national test does. For pupils with equivalent KS2 performance levels, FSM pupils are systematically disadvantaged in the grammar school selection test in this chosen LA.

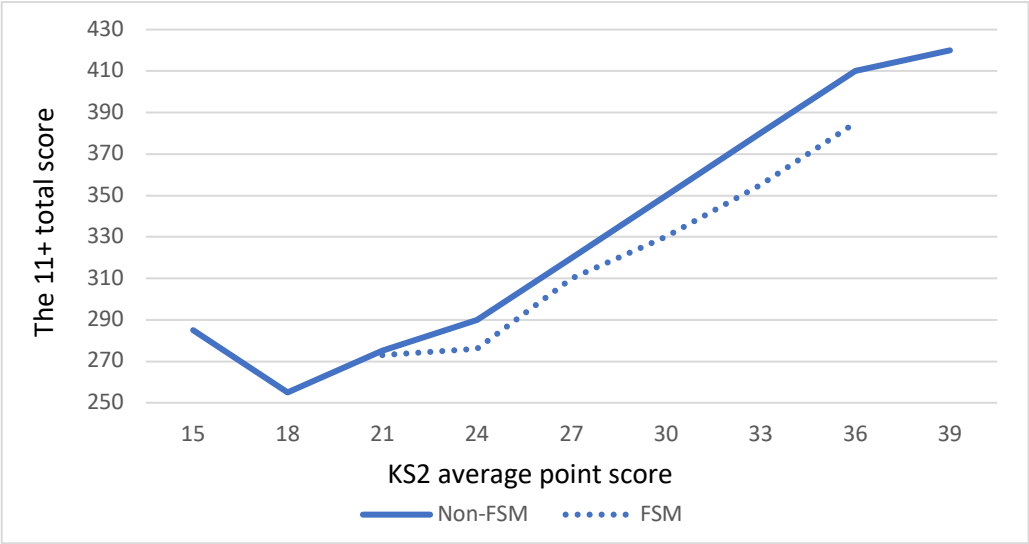

Figure 10.5: Relationship between KS2 average point score and the 11+ total score

To conclude, based on the analysis of 2011/2012 KS2 pupils in one selective LA, the lower proportion of FSM pupils taking the grammar school selection test, compounded by their lower rate of passing the test, has led to a disproportionately low probability of securing grammar school places for FSM pupils, even when KS2 performance is accounted for. The difference between the two groups is small for pupils with low KS2 attainment, as they are usually not academically prepared for grammar school selection, regardless of their FSM status. The gap between FSM and non-FSM pupils, both in terms of the rate of taking the selection test, and the rate of passing the test, affect FSM pupils who have reached the expected performance level of grammar schools. Thus, the analysis in these two sections provides possible explanations for the underrepresentation of FSM pupils in grammar schools, even after accounting for prior attainment.

10.9 The relationship between the opportunity to attend grammar schools and pupil's prior attainment, geographical location and family background

While section 10.8 pays most attention to the underrepresentation of FSM pupils in grammar schools, this section is not limited to this subgroup and returns to the general patterns of all the

pupils in selective LAs. The following analysis elucidates the relationship between the opportunity to attend grammar schools and pupil's attainment, geographical location and family background in a systematic way. All background variables which are believed to be related to grammar school opportunities are evaluated together using logistic regression models.

Table 10.7 presents the results of the logistic regression models predicting pupils' opportunities of attending grammar schools. As mentioned in the methods chapter, the most important outcome indicators are 1) the increase in the percentage correctness at each stage, which reveals how knowing certain sets of background variables increases the predictive ability of the model, and 2) the Exp (B) of each baseline variable in the right column, which provides the odds ratio of the probability of getting into grammar schools after accounting for other variables in the model.

Model 1 only includes pupil's background characteristics, without controlling for prior attainment. According to this model, boys are slightly more likely to attend grammar schools. However, as the total population, the number of grammar school places, and the number of single-sex grammar schools, are all similar for the two genders, the higher ratio of boys might be due to the differing proportion of missing data in background variables, which is higher overall for girls in grammar schools, but higher for boys in the general population. Therefore, more girls in grammar schools are excluded from the analysis, but more boys in the base population are deleted, which leads to a spurious 'bonus' for boys in Model 1. Meanwhile, older pupils within the year group have an advantage before accounting for prior attainment, which is consistent with findings in previous research. Children from poorer families are less likely to attend grammar schools, and the same trend also applies to FSM eligible pupils and SEN pupils, confirming the conclusions in the previous sections. For pupils from different ethnic groups, Chinese pupils are about 7 times as likely as white pupils to attend grammar schools. Unlike the national pattern revealed by Andrews, Hutchinson and Johnes (2016) that black pupils are believed to be the most disadvantaged, this research shows that in these 12 selective LAs with a relatively high proportion of grammar school places, white pupils are the least likely to attend grammar schools when other personal variables being equal. Furthermore, the chance to attend grammar schools for children staying in their own LAs for secondary schooling is only 28% of those who moved outside, which is also consistent with the conclusion in previous sections. Despite these differences, inputting all these personal variables into Model 1 only increases the accuracy of prediction by 3.4% over the null model which includes no

explanatory variables. This reveals that personal backgrounds account for a relatively small proportion of the variation in the opportunity to attend grammar schools.

Table 10.7: Logistic regression models of the opportunity to attend grammar schools

	Model 1		**Model 2**		**Model 3**		**Model 4**	
Variable	**B**	**Exp (B)**	**B**	**Exp (B)**	**B**	**Exp (B)**	**B**	**Exp (B)**
Girl vs. Boy	-0.096	0.908	-0.146	0.864	0.062	1.064	-0.006	0.995
Month Age	0.018	1.018	-0.065	0.937	-0.04	0.961	-0.054	0.947
Staying within Home LA	-1.266	0.282	-1.048	0.351	-0.717	0.488	-0.714	0.49
IDACI	-3.674	0.025	-2.974	0.051	-3.074	0.046	-2.946	0.053
FSM Eligible	-1.17	0.31	-0.867	0.42	-0.757	0.469	-0.727	0.483
SEN School Action	-2.307	0.1	-0.787	0.455	-0.529	0.589	-0.336	0.714
SEN School Action Plus	-2.267	0.104	-0.694	0.499	-0.465	0.628	-0.321	0.725
SEN Statement	-2.9	0.055	-0.757	0.469	-0.724	0.485	-0.554	0.574
Asian	1.164	3.202	1.67	5.313	1.139	3.124	1.263	3.536
Black	0.596	1.815	0.8	2.225	0.591	1.806	0.621	1.86
Chinese	2.001	7.393	2.082	8.018	1.402	4.065	1.506	4.51
Mixed	0.362	1.436	0.334	1.397	0.241	1.272	0.241	1.273
Unclassified	0.304	1.355	0.247	1.28	0.211	1.235	0.202	1.224
Other Ethnic Groups	0.826	2.285	1.234	3.434	0.649	1.914	0.77	2.161
KS1 Math Point Score	-	-	0.246	1.279	-	-	0.008	1.008
KS1 English Point Score	-	-	0.28	1.323	-	-	0.128	1.137
KS2 Math Fine Grade	-	-	-	-	2.766	15.89	2.604	13.524
KS2 English Fine Grade	-	-	-	-	1.674	5.335	1.255	3.509
No. of Observation	45,048		45,048		45,048		45,048	
Percentage Correctness	Increase 3.4%		Increase 10.2%		Increase 13.7%		Increase 13.8%	
	72.2%-75.6%		72.2%-82.4%		72.2%-85.9%		72.2%-86%	

Based on Model 1, Model 2 further adds KS1 attainment variables, which are KS1 maths point score and KS1 English point score. This leads to an additional 6.8% in the predictive correctness, reaching a total increase of 10.2% over that of the null model. In Model 2, adding KS1 attainment decreases the effect of moving outside the home LA, IDACI, FSM eligibility, ethnicity and most notably SEN in predicting the opportunity for grammar school participation. This underscores how considering attainment weakens the gaps between pupils with different backgrounds, especially for SEN pupils. Unlike the above variables, including KS1 attainment, there is a small increase in the effect of ethnicity for most groups, except the Mixed and Unclassified group. Meanwhile, it is not surprising that pupils with better KS1 results have a higher grammar school opportunity, and especially for those with higher KS1 English performance.

In Model 3, KS2 attainment rather than KS1 attainment is controlled for. Adding KS2 attainment based on pupil's background variables leads to an extra 10.3% increase in predictive accuracy, and has a total increase of 13.7% over that of the null model. The growth in the predictive accuracy of Model 3 demonstrates that KS2 attainment is not only more important than KS1 attainment in predicting pupils' grammar school opportunities, but also accounts for most of the explained variation in grammar school opportunities in these models. Model 3 presents a similar situation of Model 2 that older pupils in the year group are less likely to go to grammar schools once KS2 attainment variables are controlled for. As grammar school selection tests are usually standardised by age, this may reflect the inadequate standardisation of the test results during the selection process. However, it may also reflect the lack of age-standardisation in the KS2 results. Without considering age, the KS2 test is judging younger pupils to be less able than otherwise would be revealed by the 11+. Besides pupil's background variables, the odds ratio for KS2 attainment reveals that maths attainment is more important in predicting grammar school opportunities than English, as pupils with one grade higher in maths are about 15 times more likely to attend grammar schools, while pupils with equivalent advantages in English only have 5 times the difference. This is contrary to the pattern of KS1 result.

Model 4 combines all the aforementioned variables. However, after adding KS1 attainment back into the model again, there is almost no increase in the predictive correctness compared with Model 3 (only 0.1%). As KS2 and KS1 attainments are highly correlated, when KS2 attainment is controlled for, KS1 attainment no longer plays an important role in predicting

grammar school opportunities, and it can almost be fully explained by the difference in KS2 performance.

Besides the correlation between attainments at different key stages, it should also be noted that personal backgrounds are also related to attainment. In order to present whether the predictive ability of personal backgrounds overlaps with KS2 attainment, a reversed two-stage logistic regression is applied. KS2 attainment is put into the model first and then personal backgrounds—this is opposite the order of the biographical one in Model 3. The result of the reversed model shows that including KS2 attainment in the model constitutes 12.6% of the growth in predictive correctness, and leaves only 1.1% for personal backgrounds—smaller than the 3.4% demonstrated in Models 1 and Model 2. Looked at in this way, the influence of personal backgrounds in predicting grammar school opportunities is located between 1.1% to 3.4%. As most of the differences resulted from personal backgrounds overlap with KS2 attainment, the influence of personal backgrounds independent of attainment is small.

According to the above analysis, during the process of grammar school selection, attainment is still the most influential factor. Pupils with higher KS2 performance, especially those with high maths performance, have the highest probability of attending grammar schools, when other variables being equal. However, a strong indirect link between grammar school opportunities and family background still exists, which is realised through the close connection between early-age attainment and family background. The systematic underachievement of pupils from less advantaged families means although the selection is mainly based on academic ability, the distribution of the grammar school opportunity is still heavily dependent on family background. Moreover, even after accounting for prior attainment, the access to grammar schools is still unevenly distributed between pupil groups. While this may due to the imperfect control process as the available baseline variable is definitive, it also reveals the possibility of unfairness in the selection process which may further impede the opportunity for certain groups.

10.10 Conclusion

In sum, the opportunity to attend grammar schools varies across pupil groups. Since KS2 performance is the best predictor of pupils' grammar school opportunities, the underrepresentation of certain pupil groups in grammar schools is primarily due to their lower prior attainment. However, there is also evidence that even after accounting for prior attainment,

disadvantaged pupils are still less likely to attend grammar schools. One of the potential reasons for the low grammar school opportunities for disadvantaged pupils is their low levels of participation in the selection test. Moreover, even when disadvantaged pupils do sit the grammar school selection test, their success rates are lower than more advantaged pupils with equivalent KS2 performance levels. To conclude, based on the unbalanced opportunities of attending grammar schools whether or not prior attainment is controlled for, it is unlikely that the selection mechanism is equitable. After examining access to grammar schools, the next chapter discusses grammar schools' effectiveness in improving pupils' academic performance.

11 Findings about the effectiveness of grammar schools in improving pupils' academic performance

This chapter assesses the effectiveness grammar schools and their non-selective counterparts in selective LAs. Both traditional regression models which control for pre-existing differences between pupil groups, and the more innovative RDD approach, are applied to present the comparison between these two types of schools.

11.1 Result from OLS linear regression models

In this section, the effectiveness of grammar schools in raising pupils' academic performance is first evaluated through linear regression models. It starts with the descriptive results of the raw performance of grammar schools and non-selective schools in general. The analysis then turns to the estimation results of the effectiveness of grammar schools, which includes the national pattern and the patterns of individual LAs. It also discusses the stability of the estimation results, and the relationship between the effectiveness of grammar schools and the selectivity within the local area.

11.1.1 The general pattern of raw performance of grammar schools and non-selective schools

This section first compares the raw performance levels of pupils in grammar schools and non-selective schools in selective LAs. The analysis presents both KS4 and KS2 attainment results.

Since grammar schools are usually famous for their high performance, it is within reasonable expectations that their pupils' average KS4 attainments are higher than pupils from non-selective schools. According to Table 11.1, the difference between the two types of schools is most obvious in total GCSE point score. While grammar schools averaged point scores of 72, the result is under 50 for non-selective schools. The gap between the two groups decreases to only 15 point scores when capped GCSE is compared. The difference between these two indicators might be partly due to the number of KS4 exams taken, which is about one subject higher in grammar schools than in other schools. In terms of average GCSE results, grammar school pupils outperformed their counterparts by a point score of 1.7.

While grammar schools possess an obvious advantage in terms of raw KS4 performance, their pupils already have higher KS2 results before secondary school. The comparison of the KS2

total mark demonstrates that grammar school pupils averaged marks of 165 (out of 200) on their KS2 English and maths assessments. The corresponding results for non-selective schools is only 122. Therefore, while the advantage of grammar schools in raw KS4 performance is substantial, this may be largely a result of their advantaged intakes, due to selective admission. Grammar school pupils have already outperformed their counterparts since primary school. As the differences in the average performance between the two types of schools at both key stages are substantial, the comparison also implies that grammar school pupils might be very different from their counterparts in non-selective schools.

Table 11.1: Raw performance of grammar schools and non-selective schools in selective LAs

	Total GCSE point score	**Capped GCSE point score**	**Average GCSE point score**	**KS2 total mark**
Grammar schools	72	55	6.6	165
Non-selective schools	48	41	4.9	122

11.1.2 Raw performance of grammar schools and non-selective schools in individual LAs

Similar to the general results which showed higher raw KS4 performance for grammar schools than non-selective schools, the pattern is consistent when each selective LA is considered.

For the raw total GCSE results, the smallest advantage for grammar schools in comparison with non-selective schools within a single selective LA is 16.7 point scores (Stoke-on-Trent). The biggest advantage for grammar schools in terms of raw performance is 48.5 point scores in total GCSE. This is in Liverpool, where their pupils achieved twice as high as pupils in non-selective schools in the local area (92.8 vs. 44.2). The difference in capped GCSE attainment between the two types of schools also demonstrates the advantaged raw KS4 performance of grammar schools in each LA. The lowest raw advantage of grammar schools is in Stoke-on-Trent again, with their pupils on average possessing 11.1 more point scores than others. Reading and Liverpool produce the largest gaps between grammar schools and non-selective state-funded schools, which are 23.6 and 20.3 respectively. Meanwhile, the calculation of average GCSE results also confirmed grammar schools' better raw performance in each selective LA.

Overall, the higher raw performance for grammar schools is stable for three GCSE outcomes, and is the most pronounced in terms of total GCSE. Despite the scope of difference in each LA, grammar school pupils' higher raw KS4 performance is consistent across all selective LAs.

11.1.3 Individual schools' raw performance

In addition to presenting the overall pattern and the picture of each LA, the raw performance of individual schools is also evaluated to assess the internal differences within each type of school.

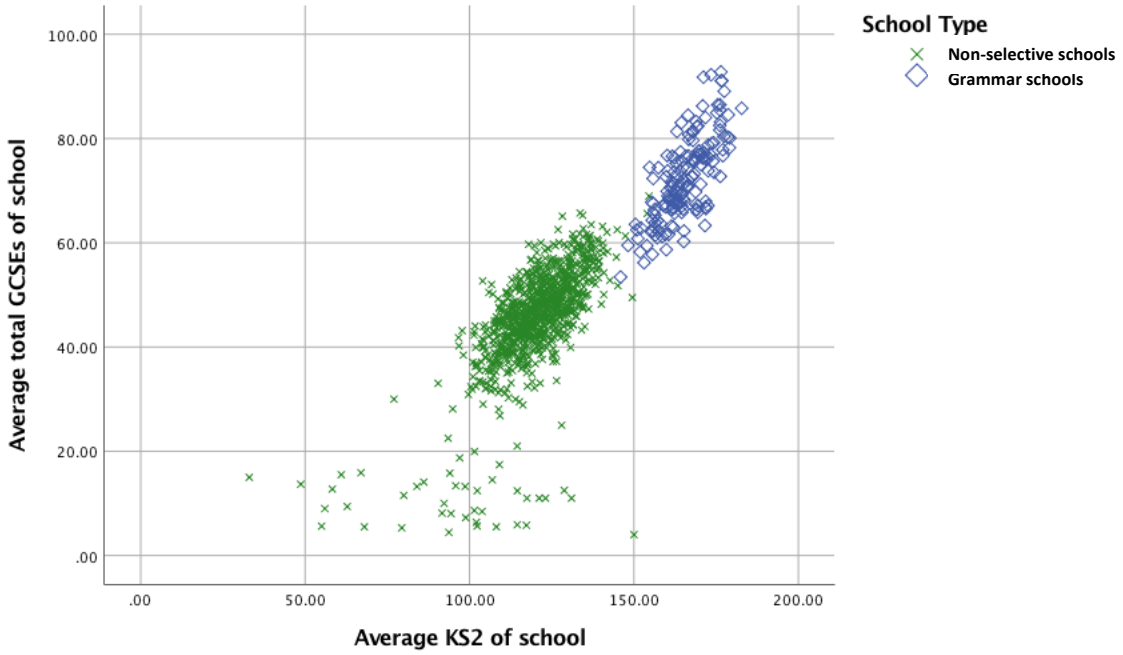

Figure 11.1: Relationship between school-level KS2 total mark and total GCSE point score in selective LAs

To demonstrate the performance of individual schools, Figure 11.1 presents the relationship between school average KS2 mark and total GCSE point score. As expected, there is a positive relationship between school aggregated KS2 and KS4 performance. Schools admitting pupils with higher KS2 performance usually perform better at KS4 as well. While the dots for non-selective state-funded schools are located in the lower-left of the figure, grammar schools can be found in the upper-right, demonstrating their superior intakes and outcomes. Meanwhile, the figure also presents clear internal differences between grammar schools. While some grammar schools overlap with non-selective schools at KS4 attainment (first group), others do not (second group). For the first group of grammar schools, although they admitted pupils with

higher KS2 attainment than top-performing non-selective schools, their advantage vanished at KS4. The second group of grammar schools shares no similarity with any non-selective schools, as their pupils have higher attainment for both KS2 and KS4 measurements. Most of these grammar schools have average KS4 results of 65 GCSE point scores or above, and the highest is 93. In contrast, the maximum total GCSE point score for non-selective state-funded schools is 69. Pupils in the second group of grammar schools also had high KS2 marks, and the average results exceeded 150. The outstanding average performance at both stages in these grammar schools implies their systematically different pupil compositions.

To evaluate whether the wide between-school variance in terms of intakes and later performance is a national pattern, schools in non-selective LAs are also presented (Figure 11.2). The figure of non-selective LAs shows that the dichotomous pattern as revealed in selective LAs no longer exists. Although schools still have substantial differences in GCSE results, there is no clear-cut separation between high-performing and low-performing schools in non-selective LAs. The total GCSE point scores of most schools in non-selective LAs are between 36 and 65. The KS4 performance variation in comprehensive areas is thus smaller than that of selective LAs, with the range being 30-60 for most non-selective schools and 60-80 for most grammar schools. While some schools in comprehensive LAs perform similar to low-performing grammar schools at KS4, no school in a non-selective LA has KS4 results comparable to those of top-performing grammar schools.

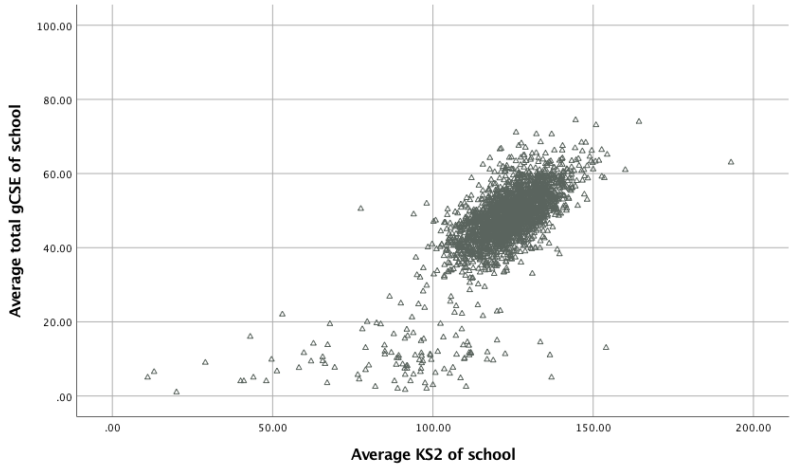

Figure 11.2: Relationship between school- level KS2 total mark and total GCSE point score in non-selective LAs

In addition to the between-school variance in KS4 performance, school-level KS2 performance in non-selective LAs is also more balanced than in selective LAs. In selective areas, the average KS2 performance range for most non-selective schools is 100-140. Meanwhile, most grammar schools in selective LAs have KS2 performances over 150, with the highest approaching 180. In contrast, the KS2 performance range for most schools in comprehensive LAs is 100-150. Despite the smaller school-level variation in KS2 attainment in comprehensive LAs, a minority of schools in comprehensive LAs have exceptional average KS2 performance, even exceeding that of some grammar schools. While the lowest school-level KS2 result for grammar schools is 146, there are 27 schools in comprehensive LAs which have reached this threshold, constituting 1.2% of the total number of schools in non-selective areas. This confirms previous research that some comprehensive schools can also be highly selective, even without overt academic selection (Coe et al., 2008).

Comparing the raw performance of selective and non-selective LAs reveals the larger between-school variance in the former, both in terms of intakes and later performance. In selective LAs, grammar schools present a salient difference from non-selective schools within the same areas, implying the more differentiated school-level performance of the selective system.

11.1.4 The effectiveness of grammar schools: The national pattern

As presented above, the unbalanced raw performance across schools reveals how pupils at certain types of schools may differ from others. This underscores the need to control for pre-existing differences between pupils in order to present a fair evaluation of school effectiveness. Thus, this section uses multi-stage OLS regression models to control for different sets of baseline variables and estimate the difference in academic progress associated with school type.

Four-sets of baseline variables are added in the multi-stage linear regression models, as described in the methods chapter. Table 11.2 presents the changes in the predictive ability of the model at each stage. For total GCSE and capped GCSE results (Panel A), the growth in the predictive accuracy of the model at each stage is the same. Adding all the pupil background variables at the first stage accounts for 17% of the variation in pupils' KS4 attainment. The percentage spikes to 57% when KS2 attainment is added, revealing the strong correlation between prior attainment and pupils' future academic performance. At this stage, more than half of the variation in pupils' KS4 attainment is explained by included baseline variables.

After accounting for all of the pupil-level variables, adding school compositional variables leads only to slight growth in the fitness of the model, with the overall R-square reaching 0.59. The final step of including 'whether attending grammar school or not' does not change the predictive ability of the model, with the final R-square remaining the same. This means that after knowing all of the pupil-level and school-level characteristics, school type is not an important predictor of KS4 performance. The results of the average GCSE point scores are slightly different, but still present the same tendency (Panel B in Table 11.2).

Table 11.2: Changes in R-square in OLS models predicting KS4 attainment

Baseline variables	**R-square of multi-stage linear regression**
Panel A: Total GCSE point score/Capped GCSE point score	
Pupil background	0.17
Pupil background, KS2 attainment	0.57
Pupil background, KS2 attainment, School composition	0.59
Pupil background, KS2 attainment, School composition, School type	0.59
Panel B: Average GCSE point score	
Pupil background	0.18
Pupil background, KS2 attainment	0.60
Pupil background, KS2 attainment, School composition	0.62
Pupil background, KS2 attainment, School composition, School type	0.62

Following the changes in R-square, the coefficients for school type are also presented. After accounting for all of the pupil-level and school-level baseline variables, the conclusion is consistent that grammar schools are only slightly better than other state-funded schools. While

the scope of difference in total GCSE is large, which is around 3 point scores, it is limited in capped GCSE and average GCSE, which are 0.3 and 0.02 respectively.

The detailed results of each stage of the regression models are given to present changes in coefficients when different sets of baseline variables are controlled for. Since the patterns are similar between the three GCSE outcome variables, Table 11.3 only presents the capped GCSE models. Results of total GCSE and average GCSE are attached in Appendixes 4 and 5.

Table 11.3: Coefficients in OLS models predicting capped GCSE point score

	Capped GCSE			
Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Girls vs. Boys	2.512	2.332	2.300	2.302
Month Age	0.155	-0.116	-0.107	-0.107
IDACI	-19.898	-9.412	-5.777	-5.818
FSM Eligible	-6.188	-3.039	-2.616	-2.615
SEN School Action Plus or SEN Statement	-11.292	-1.618	-1.896	-1.898
EAL	1.759	2.735	2.803	2.807
Asian	3.278	1.272	1.127	1.124
Black	2.498	2.112	1.954	1.966
Others	2.166	1.137	0.959	0.966
KS2 total mark of English and maths	-	0.25	0.23	0.23
Mean KS2 Total Mark in Secondary School	-	-	0.078	0.073
Mean FSM Proportion in Secondary School	-	-	-0.051	-0.053
Grammar School	-	-	-	0.308
No. of Observation	149,072	149,072	149,072	149,072

In addition to the difference in KS4 performance associated with school type, the regression models also reveal the unbalanced performance patterns across social groups. According to Models 1 and 2 (Table 11.3), after accounting for all of the pupil-level baseline variables, pupils with higher KS2 attainment, girls, EAL pupils, non-white pupils, and pupils from wealthier families are advantaged in capped GCSE. Meanwhile, FSM and SEN pupils are making less

progress from KS2 to KS4 when all the other variables are controlled for. The results for month age vary between Models 1 and 2. Without controlling for KS2 attainment, older pupils are expected to have better GCSE results, according to Model 1. The systematic underperformance of summer-born pupils has long been noted, and it has even triggered political attempts to give parents of summer-born pupils the option to delay school entrance one year. The results in Model 1 correspond to this pattern. However, according to Model 2, which also controls for pupils' KS2 attainment, the previous pattern is reversed—younger pupils are expected to make even more progress than older pupils within the year group.

The spectrum of the effects of the school-level variables is narrower than the pupil-level characteristics. For mainstream state-funded schools in selective LAs, a 1-mark increase in average school-level KS2 attainment is associated with a 0.07-point increase in capped GCSE for individual pupils. As over 95% of pupils were educated in schools with a mean KS2 mark ranging from 105 to 171, the maximum difference originating from school average performance for the vast majority would be about 4.6 point scores for the capped GCSE. Meanwhile, the proportion of FSM pupils is negatively associated with the KS4 results, as a 1 percent increase in the proportion is expected to decrease capped GCSE results by 0.05 of a point score for each pupil at the school. Since 95% of pupils are educated at schools where 0% to 32% of the pupils receive FSM, the greatest effect of school-level FSM proportion for the vast majority is roughly 1.6 capped GCSE point scores.

Adding school compositional variables increases the overall fitness of the model only slightly (Table 11.2). However, whether or not school compositional variables are controlled for has a major influence on the school type coefficient. If school compositional variables are excluded from the model, the coefficients for grammar school pupils increase substantially (Panel B in Table 11.4). This is consistent with the discussion in the methods chapter which showed that including school compositional variables lowers the estimation of grammar school effectiveness due to several factors such as measurement errors and pre-existing differences between pupil groups.

Table 11.4: The effectiveness of grammar schools before and after controlling for school compositional variables in OLS models

	Coefficient for grammar schools
Total GCSE point score	
Panel A: Models controlling for pupil-level and school-level variables	2.67
Panel B: Models controlling for pupil-level variables only	7.59
Capped GCSE point score	
Panel A: Models controlling for pupil-level and school-level variables	0.33
Panel B: Models controlling for pupil-level variables only	3.52
Average GCSE point score	
Panel A: Models controlling for pupil-level and school-level variables	0.016
Panel B: Models controlling for pupil-level variables only	0.46

In addition to the OLS models, fixed slope ML models and random slope ML models are also conducted on the same cohort in selective LAs. The detailed results of these two models are included in Appendix 6. In general, the OLS regression models produce better results for grammar schools than both types of ML models do. Meanwhile, the results of the fixed slope ML models are also better than those of the random slope ML models. Despite differences between model types in terms of estimated results, the conclusions based on all the three models are consistent. While there is some evidence that grammar schools do better in terms of total GCSE, the results of capped GCSE and average GCSE do not evidence grammar schools' academic superiority.

11.1.5 The effectiveness of grammar schools: Patterns of individual LAs

As the difficulty of grammar school selection varies across LAs, this may influence the patterns of school effectiveness in each LA. Therefore, in addition to the general pattern of the effectiveness of grammar schools and non-selective schools, this section also presents the individual patterns for each selective LA.

After taking pupil-level and school-level characteristics into account, the coefficients for grammar schools in the OLS regression models are not consistently positive across selective LAs. For total GCSE result, 23 LAs present positive patterns that grammar schools are outperforming non-selective state-funded schools after the better-off demographic feature is controlled for. The results of the capped GCSE tell a similar story. There are 22 selective LAs in which grammar schools outperform non-selective schools within the same LAs, and the results in the other 14 LAs are reversed. In terms of average GCSE point scores, the number of selective LAs with positive grammar school coefficients falls to 20.

The three GCSE outcome indicators present the same signal that after considering pupil-level and school-level baseline differences, grammar schools in about two-thirds of the selective LAs outperform non-selective schools for each GCSE outcome indicator. In addition, if the results of all three GCSE indicators are combined, only 17 out of the 36 selective LAs present consistently positive results for grammar schools. In addition to the OLS models, the pattern of the grammar school effectiveness in each LA is also analysed with fixed and random slope ML models. According to these two types of ML models, the number of selective LAs with positive results for grammar schools is even lower (Table 11.5). This comparison reveals how the choice of regression models, and outcome indicators, may alter conclusions of grammar school effectiveness in individual LAs.

Table 11.5: Number of selective LAs with a positive grammar school effect

	Number of selective LAs with a positive grammar school effect
OLS	
Total GCSE	23
Capped GCSE	22
Average GCSE	20
All three indicators	17
Fixed slope ML	
Total GCSE	21
Capped GCSE	18
Average GCSE	20
All three indicators	14
Random slope ML	
Total GCSE	20
Capped GCSE	20
Average GCSE	15
All three indicators	13

11.1.6 Stability of the estimates for individual LAs

Since the above analysis reveals the dissimilar estimation results when different outcome variables and regression models are applied, this section combines all of the GCSE indicator outcomes (total, capped and average GCSE) in three types of models (OLS, fixed and random slope ML) to assess the stability of the grammar school coefficient.

Surprisingly, there are only three selective LAs in which grammar schools consistently outperform non-selective schools in their LAs, when all three outcome indicators for the three types of models are combined. These LAs are Devon, Essex, and Cumbria. However, there are 32 LAs which at least have one positive result for their grammar schools when different outcome indicators and models are applied. Meanwhile, there is only 1 LA, Plymouth, where grammar schools invariably underperform their non-selective counterparts, regardless of outcome indicator or model applied. However, there are 30 LAs with at least one negative result for grammar schools. Therefore, most selective LAs (89%) show positive results under some circumstances, but present negative results for others. Thus, the evidence is mixed. This confirms that the stability of the estimation is low and the results for individual LAs are sensitive to outcome variables and models applied. Therefore, different choices in the statistical analysis may influence conclusions on grammar schools' effectiveness. This implies the limitations of the regression approach in controlling for pre-existing differences among pupil groups.

11.1.7 Characteristics of LAs with consistently positive results for grammar schools

Based on the evaluation above, only a handful of selective LAs produce consistent results when different outcome indicators and regression models are used. In order to assess whether there are any systematic similarities between these LAs, this section pays attention to their characteristics.

For the three LAs in which grammar schools outperform non-selective counterparts, the most obvious common trait is their small proportion of grammar school places. Among all selective LAs, Devon has the lowest proportion of grammar school pupils, which is only 1.4% in 2011. Cumbria has the second-lowest proportion among the selective LAs, with only 2.2% of pupils attending grammar school in 2011. For Essex, although the rate is slightly higher than the

previous two LAs, which is 2.8%, it is still well below those of most selective LAs. Consistent with the low proportion of grammar school places, the number of grammar schools is also small in these 3 LAs. For Devon and Cumbria, there is only one isolated grammar school in the local area. While there are 4 grammar schools in Essex, this number is still small compared with other selective LAs (e.g. 32 in Kent and 15 in Lincolnshire).

Since these 3 LAs only accept a small proportion of pupils into grammar schools, it is possible that they are recruiting pupils with higher KS2 attainment than in other selective LAs. Therefore, the lowest KS2 marks for grammar school pupils in these 3 LAs are examined first. For pupils in Devon, the KS2 threshold for grammar school pupils is the highest among the selective LAs, which is 145. Similarly, pupils in Essex LA also need to achieve KS2 marks of at least 141 in order to have any chance of attending grammar school. This is the third-highest among selective LAs. However, in Cumbria, the KS2 threshold is only 91. This means that the high prior attainment for grammar school pupils is no longer true in Cumbria. In addition to the lowest KS2 marks for grammar school pupils in these 3 LAs, the average KS2 marks in grammar schools are also compared. Once again, the evaluation demonstrates that while the average prior attainment for grammar schools in Devon and Essex topped the rankings, Cumbria's mark was still low. Thus, no consistent pattern was found for intakes' prior attainment in these 3 LAs.

The demographic characteristics for pupils in these 3 LAs are also examined (Table 11.6). While the average IDACI in selective LAs is already below the national average, these 3 LAs are more advantaged than other selective LAs, in terms of IDACI. The index in Devon is the lowest of the three, revealing the high economic status of the area. The results are similar for Essex, with an index of 0.14. While the index for Cumbria is higher than the previous two LAs (0.17), it is still more advantaged than the average IDACI of selective LAs (0.19). Similar to the results for IDACI, there are also lower proportions of FSM pupils in these 3 LAs, in contrast to the 11.7% average for selective LAs, with the highest being 10.7% (Devon) and the lowest being 8% (Cumbria). These LAs also have lower proportions of EAL pupils. In contrast to the average EAL proportion for selective LAs (11.7%), the rate is only 1% for Cumbria and 1.9% for Devon. Although the proportion is higher for Essex (4.7%), EAL groups are still underrepresented in these 3 LAs. Unlike the above-mentioned aspects, the proportion of SEN-PS pupils in these LAs deviates little from the average of selective LAs, presenting no special patterns.

As a result of the advantaged pupil backgrounds in these 3 LAs, the composition of their grammar schools also demonstrates more advantaged status over the other selective LAs. Grammar schools in Devon, Essex and Cumbria enroll pupils from wealthier areas. Their IDACIs are below the national average for grammar school pupils (0.13). The index is around 0.1 for Devon and Cumbria, and 0.12 for Essex. The proportions of FSM pupils in grammar schools are also lower for these three areas. While 2.6% of pupils are eligible for FSM in English grammar schools, the rate is 2.1% in Devon, and is below 1% in Essex (0.3%) and Cumbria (0.9%). The gap in the proportion of SEN-PS pupils between these 3 LAs and the average is also salient. There are no SEN-PS pupils enrolled at Devon's grammar school, and the proportions are below 1% in Essex and Cumbria, which are lower than the average rate of 1.3%.

Table 11.6: Pupil characteristics in LAs with consistently positive results for grammar schools (Devon, Essex and Cumbria)

	Average IDACI	**FSM proportion**	**EAL proportion**	**SEN-PS proportion**	**Average KS2 total mark**
Devon	0.136	10.7	1.9	8.7	126
Essex	0.168	8.4	4.7	7.3	125
Cumbria	0.144	8	1	7.5	128
Selective LAs	0.185	11.7	14.3	7.2	128
Grammar school in Devon	0.095	2.1	5.2	0	175
Grammar schools in Essex	0.119	0.3	20.5	0.3	176
Grammar school in Cumbria	0.074	0.9	1.9	0.9	161
Grammar schools in selective LAs	0.129	2.6	17.6	1.3	165

These 3 LAs in which grammar schools consistently outperform their non-selective counterparts have several similarities. First, they only enroll a small proportion of grammar school pupils, and there are only a few grammar schools within each local area. Second, compared with other selective LAs, these three LAs are wealthier, and have proportionally fewer FSM and EAL pupils. Third, pupil compositions in these LAs' grammar schools are also more advantaged. Despite the better demographic characteristics of these three LAs, there is no consistent pattern in the prior attainment of their grammar school pupils. While grammar

school pupils' KS2 performances are exceptional in Devon and Essex, the results show little similarity in Cumbria.

11.1.8 Characteristics of the LA with consistently negative results for grammar schools

Among the 36 selective LAs, only Plymouth has negative results for their grammar schools regardless of outcome variables and models applied. The proportion of grammar school pupils and the total number of grammar schools in Plymouth present no special pattern. About 14.4% of pupils in this LA were enrolled in its 3 grammar schools in 2011. Similar to the pattern for these two factors, the KS2 prior attainment for grammar school pupils in Plymouth is also close to the average of selective LAs (164 vs. 165).

Contrary to the advantaged status of pupils in Devon, Essex and Cumbria, the IDACI in Plymouth is higher than the average of selective LAs (0.2 vs. 0.185). This demonstrates the poorer status of pupils in this LA. However, the index in Plymouth is not the worst among all selective LAs, and it is still better than that of 11 other selective LAs. The proportion of FSM pupils in Plymouth is also higher than the average in selective LAs (13.6% vs. 11.7%). But again, this high rate in Plymouth presents no substantial difference from other selective LAs, and there are 8 LAs with higher proportions of FSM pupils than Plymouth. The proportion of SEN-PS pupils in Plymouth is also within a reasonable range, which is just above the average. Unlike average IDACI, FSM and SEN-PS proportions, the proportion of EAL pupils in this LA is lower than the rate for selective LAs (4.1% vs.14.3%). When the characteristics of grammar school pupils in Plymouth are compared to the figures for all the grammar school pupils, the results are the same. While their pupils are more disadvantaged in terms of IDACI and FSM proportion, the situation diverges little from that of other selective LAs.

The attainment data and demographic features for pupils in Plymouth show no difference from those in other selective LAs, even though it is the only LA in which grammar schools consistently have worse estimation results than non-selective schools. Therefore, the reason for the low effectiveness of grammar schools in this LA is likely be due to other factors which are difficult to demonstrate through surface characteristics.

11.1.9 The effectiveness of grammar schools: Results from GCSE English and maths

While previous analysis has focused on pupils' overall GCSE performance, this section presents separate results for the two most fundamental subjects—English and maths. The analysis presents both a general pattern for selective LAs, and pictures of each individual LA.

In 2016, the average GCSE English point score was 6.5 for grammar schools and 5 for non-selective state-funded schools within the selective areas. The GCSE maths result is 7 point scores for grammar schools and 5 for other mainstream state-funded schools in selective LAs. While pupils in non-selective state-funded schools do equally well in English and maths, pupils in grammar schools score better in maths than English. As a result, the gap in KS4 performance between grammar schools and other schools is more pronounced in maths than in English.

After accounting for pupil-level and school-level characteristics, grammar schools still do better in both subjects, but the difference is small. The average pattern in selective LAs reveals that grammar schools score 0.14 point score higher than other state-funded schools in GCSE English and 0.09 higher in maths. Therefore, grammar schools' advantage after accounting for baseline variables is equivalent to 1/7 of a grade in GCSE English and about 1/10 of a grade in maths.

When individual LAs are considered, grammar schools do not consistently outperform non-selective schools in selective LAs. For GCSE English, 14 LAs have negative results for grammar school, and the remaining 22 LAs have positive coefficients for grammar school. For KS4 maths results, 16 LAs have negative coefficients for grammar school, and the results are positive for the remaining 20 LAs. The lower number of LAs with positive results in maths confirms the pattern that grammar schools are slightly more effective at improving English performance than maths in general. Meanwhile, the pattern for individual LAs also reveals that grammar schools in some LAs may be particularly effective in one subject, but not another. The correlation between the coefficients for grammar schools in terms of English and maths results is only 0.4 after accounting for background variables. This means that while grammar schools in some LAs are more effective at improving maths scores, this does not guarantee their effectiveness in English, or vice versa. For example, grammar schools in Kingston upon Thames placed within the top 5 in terms of their effectiveness in English. However, their results

for maths were negative. Due to the weak correlation between these two subjects, in only 13 of the selective LAs did grammar schools have positive estimations for both maths and English. This accounts for roughly 1/4 of the total selective LAs. Therefore, grammar school effectiveness not only differs across LAs, but also fluctuates by subject.

11.1.10 School effectiveness and the degree of selectivity

The internal differences between grammar schools across LAs have been exhibited in previous sections. Grammar schools in each selective LA not only have distinct selection difficulty, but also have dissimilar effectiveness patterns. In order to assess whether there is a trade-off between grammar school effectiveness and the degree of selectivity in each LA, the relationship between these two factors is examined.

11.1.10.1 Raw performance of grammar school pupils and the selectivity of LA

First, the correlation between the raw GCSE performance of grammar schools and the selectivity of each LA is evaluated. Based on previous evidence of the internal differences of grammar schools, it is within the expectation that the KS4 performance of grammar schools in each LA differ. Among selective LAs, the point score of total GCSE in grammar schools ranges from 62.98 to 92.1, and the range for capped GCSE is 50.63 to 61.35. In both outcomes, the raw performance of grammar schools differs substantially from LA to LA. When the relationship between grammar schools' raw performance and the degree of selectivity is evaluated, there is no systematic pattern found. A higher likelihood of attending grammar schools at the local level does not necessarily lead to lower KS4 results in grammar schools, and vice versa (Figure 11.3). The correlation between the proportion of grammar school places and total GCSE of grammar schools in each LA is only -0.29, and the rate is -0.28 for capped GCSE result. Meanwhile, the correlation figures between the lowest KS2 marks of grammar school pupils and KS4 results for grammar schools in each LA are also weak—0.52 for total GCSE and 0.65 for capped GCSE results.

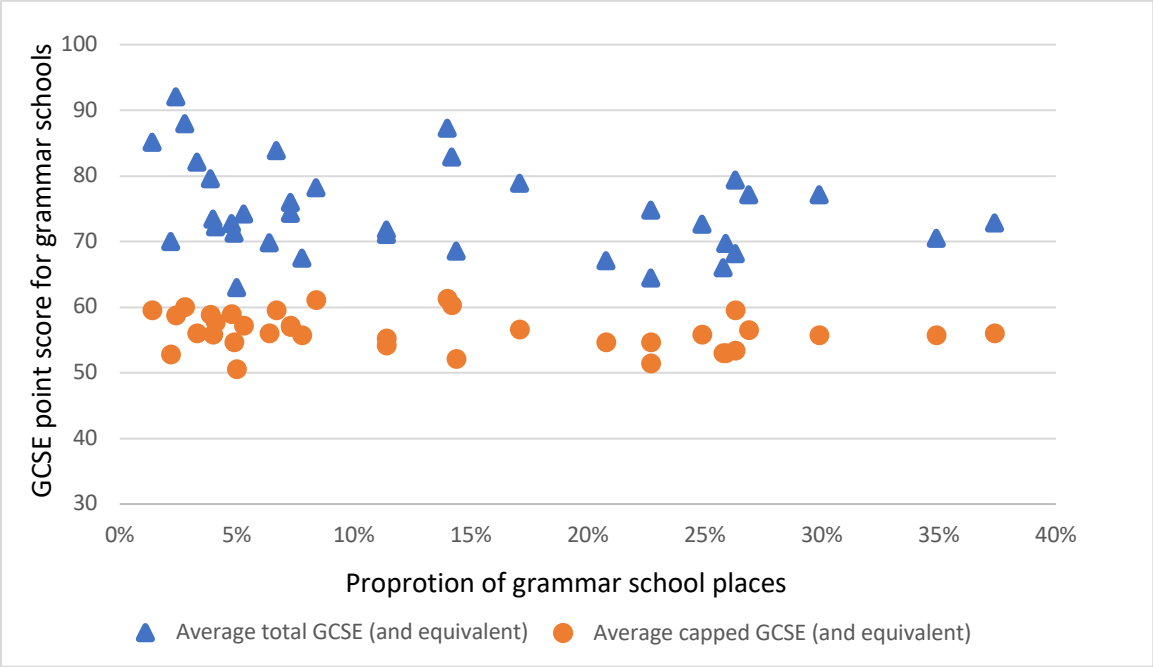

Figure 11.3: Proportion of grammar school places and GCSE point score for grammar schools in each LA

11.1.10.2 The effectiveness of grammar schools and the lowest KS2 score

In addition to evaluating raw GCSE performance for grammar schools, a detailed analysis between grammar school's effectiveness and the selectivity is conducted. The grammar school coefficients in linear regression models are used as the indicators of their effectiveness.

The coefficient for grammar schools and the lowest KS2 score for grammar school pupils in each LA are presented first, revealing how grammar schools perform in comparison with non-selective schools in their LAs when the selection difficulty varies. According to Table 11.7, these two factors are weakly correlated (Column A). Based on results in the OLS models, the correlation rate is only -0.02 for total GCSE, 0.007 for capped GCSE and 0.142 for average GCSE. The rates are similar when the coefficients in the ML models are applied as the indicator of grammar school effectiveness.

In sum, while there are slight fluctuations when estimation results from different models are applied, there is no evidence that grammar schools in LAs with easier selection are less effective.

Table 11.7: Correlation between grammar school effectiveness and the selectivity in each LA

Coefficient of grammar schools	Column A: Lowest KS2 mark of grammar school pupils	Column B: Proportion of grammar school places
Total GCSE		
OLS	-0.022	-0.189
Fixed slope ML	0.016	-0.09
Random slope ML	0.104	-0.128
Capped GCSE		
OLS	0.007	-0.126
Fixed slope ML	0.063	-0.082
Random slope ML	0.104	-0.066
Average GCSE		
OLS	0.142	-0.325
Fixed slope ML	0.187	-0.226
Random slope ML	-0.141	0.026

11.1.10.3 The effectiveness of grammar schools and the proportion of grammar school places

In addition to using grammar school selection difficulty as a proxy for selectivity, the proportion of grammar school places is also used to elucidate its relationship with grammar school effectiveness of each LA. The correlation between grammar school effectiveness and the proportion of grammar school places in each LA is negative in most of the presented results (Column B in Table 11.7). Although the correlation figures increase a bit, the correlation is still weak. The overall pattern shows no evidence of a trade-off between the proportion of grammar school places within an LA and the relative effectiveness of grammar schools in relation to their non-selective counterparts.

11.1.11 Conclusion

In sum, section 11.1 has used linear regression models to evaluate the effectiveness of grammar schools in comparison with other mainstream state-funded schools in selective LAs. The results reveal that grammar schools are only slightly more effective than their non-selective counterparts, after accounting for pupil-level and school-level characteristics. The general difference is about 1/3 of a grade in capped GCSE. This effect is tiny when considering that it is the cumulative effect over five years. In addition to the general pattern, the pictures of individual LAs underscore the dissimilarity between grammar schools in different areas. While in some LAs, grammar schools are more effective than non-selective schools, contradictory patterns are found in others. The comparison between OLS and ML models also demonstrates the instability of the estimated grammar school effect when different GCSE outcome variables

and models are applied. Most LAs have both positive and negative results for grammar schools under different circumstances. This reveals the sensitivity of the conclusion to statistical choices. When all of the outcome variables from different models are combined, only three LAs have consistently positive coefficients for grammar school. These LAs each have a low proportion of grammar school places, and more advantaged status than most selective LAs. However, the evaluation of grammar school effectiveness and LA selectivity does not find a systematic relationship between these two elements. Therefore, the differed effectiveness of grammar schools is not attributable to the degree of selectivity alone. Presenting internal differences for grammar schools, the pattern for individual LAs calls for a more detailed analysis of each LA in future studies.

11.2 Results from logistic regression models

In addition to using linear regression models, this section applies logistic regression models to estimate the academic effectiveness of grammar schools. It assesses whether grammar schools are more effective than their non-selective counterparts, in terms of achieving both basic and top KS4 grades.

11.2.1 Descriptive results for achieving 5 GCSE and equivalent qualifications at grades A-C (including English and maths) and A*-A*

For the 2015/2016 KS4 cohort in selective LAs, the proportion of pupils who achieved 5 A*-C grades is 61.7%. Based on the higher raw performance of grammar school pupils, it is within expectations that the proportion reaching this threshold would be higher for grammar schools than for other schools in selective LAs. While most grammar school pupils surpassed this threshold (97%), the proportion is just over half for other schools (56.2%). The gap between the two types of schools is even larger when the rates of achieving 5 A*-A grades are compared, with the proportion in grammar schools being about 5 times as high as in non-selective schools (63.1% vs. 12.1%). Furthermore, for grammar school pupils, the average number of subjects at grades A*-A is 6, but for other schools, it is only 1. Despite the gap, for all the pupils who have achieved 5 top GCSE grades, the average number of A*-A is similar between the two types of schools, which is 8 in grammar schools and 7 in non-selective schools. Therefore, although the overall gap in the rate of achieving 5 top GCSE grades between grammar schools and non-selective schools is immense, the difference between top-achieving pupils in different types of schools is smaller.

For pupils with equivalent performance, those who take more exams will have a higher probability of achieving 5 A*-C and 5 A*-A grades. Therefore, although grammar school pupils are more likely to reach these two thresholds, this may be due to the higher number of exams they take. On average, grammar school pupils took 10.8 subjects on GCSE or an equivalent, and the number for non-selective schools in selective LAs is 9.6. Regardless of the difference in KS4 performance between pupils in grammar school and non-selective schools, the former already have a higher probability of achieving 5 A*-C and 5 A*-A grades, as their pupils take more tests than the latter do. However, the difference in the number of tests is small, and thus cannot fully explain the substantial performance gap between grammar school pupils and others.

11.2.2 Logistic regression model estimation results

11.2.2.1 Achieving 5 GCSE and equivalent qualifications at grades A-C (including English and maths)*

As 61.5% of pupils in selective LAs had achieved 5 A*-C grades, it is the predictive accuracy of the null model before accounting for any baseline variables. After adding pupil's backgrounds at the first stage, the predictive accuracy of the model increases to 72.5%. When pupil's prior KS2 attainment is included at the second stage, the total accuracy of the prediction reaches 80.8%. The following stage, at which school-level variables are added, did not change the predictive accuracy. After including all of the aforementioned variables into the model from the previous three stages, adding school type in the last step does not bring any obvious changes to the final result as well, with the prediction accuracy stabilising at 80.9% (Table 11.8). The predictive accuracy of the logistic regression models shows that pupil's background is a very influential element in predicting their KS4 attainment, and the growth in the first stage is the largest. Meanwhile, although the predictive ability of KS2 attainment may already be accounted for by pupil's backgrounds in the first stage (as revealed in the findings about the opportunity to attend grammar schools), there is still an 8.3% increase in predictive ability when KS2 attainment is added.

Table 11.8: Predictive accuracy of the logistic regression models of 5 GCSE and equivalent qualifications at grades A*-C (including English and maths)

Variables	**Percentage correctness**	**Percentage of remaining variation explained**
Base figure	61.5	-
Personal backgrounds	72.5	11
KS2 total mark	80.8	8.3
School characteristics	80.8	0
School type (grammar school or not)	80.9	0.1
Overall	80.9	19.4

According to Column A in Table 11.9, after accounting for all of the baseline variables, the probability of achieving 5 good passes for grammar school pupils is still twice as high as pupils in other state-funded schools. Meanwhile, the probability of achieving this outcome is still unbalanced across pupil groups. While girls and ethnic minorities (with the exception of the Mixed group) have some advantages, older pupils within the year group, those from poorer areas, FSM pupils and the SEN group, are less likely to achieve this outcome. After accounting for pupil-level baseline variables, a more advantaged school composition is also associated with better academic results at KS4. Pupils in secondary schools which have higher average KS2 marks and a lower proportion of FSM pupils, are usually more likely to achieve 5 good passes at KS4.

Although school-level variables improve the predictive accuracy of the model only slightly, adding these variables alters the odds ratio for school type substantially. In order to present the difference in estimation results before and after accounting for school-level variables, the results of the model without school compositional variables are also presented in Column B in Table 11.9. In this model with only pupil-level characteristics and school type, the Exp(B) of grammar schools can be as high as 3. Compared with the results in Column A, in which the Exp(B) for grammar schools is 2, roughly 1/3 of the difference between the two school types in Column B is explained when average KS2 result and the proportion of FSM pupils of secondary schools are included. However, adding more school-level compositional variables no longer lowers the estimation results. For example, if the average school-level IDACI score or the proportion of SEN-PS pupils is added besides the previous two school compositional variables, the Exp(B) for school type increases. Therefore, some unexplained variation between grammar schools and non-selective schools remained after accounting for all of these baseline

characteristics. Based on the best available background data, the results imply that grammar school pupils have higher chances of achieving 5 A*-C grades, after controlling for pupil-level and school-level variables.

Table 11.9: Logistic regression models of 5 GCSE and equivalent qualifications at grades A*-C (including English and maths) with and without school compositional variables

	Column A: Model with school composition		**Column B: Model without school composition**	
Variable	**B**	**Exp (B)**	**B**	**Exp (B)**
Girls vs. Boys	0.449	1.567	0.454	1.574
Month Age	-0.031	0.969	-0.032	0.969
IDACI	-1.535	0.215	-1.936	0.144
FSM Eligible	-0.479	0.619	-0.532	0.588
SEN School Action	-0.432	0.649	-0.432	0.649
SEN School Action Plus	-0.53	0.589	-0.532	0.587
SEN Statement	-0.33	0.719	-0.314	0.731
EAL	0.507	1.66	0.488	1.628
Asian	0.159	1.172	0.098	1.103
Black	0.38	1.462	0.364	1.439
Chinese	1.011	2.748	1.016	2.761
Mixed	-0.003	0.997	-0.004	0.996
Unclassified	0.257	1.293	0.266	1.305
Other Ethnic Groups	0.496	1.642	0.477	1.611
KS2 total mark	0.052	1.054	0.053	1.054
Mean KS2 total mark in secondary school	0.007	1.007	-	-
Mean FSM proportion in secondary school	-0.009	0.991	-	-
Grammar school	0.75	2.117	1.083	2.953
No. of Observation	149,072		149,072	

11.2.2.2 Achieving 5 GCSE and equivalent qualifications at grades A-A*

In addition to the results for achieving good passes, the pattern for pupils with 5 A*-A grades is also examined to test whether attending grammar schools is associated with higher chances of earning high grades at KS4.

The four-stage logistic models control for pupil's characteristics, pupils' KS2 attainment, school-level characteristics and school type. Unlike in the prediction for 5 A*-C, in which personal backgrounds explain the largest proportion of the variation, including personal backgrounds in this model only led to a 1.9% increase in predictive accuracy (Table 11.10). In contrast, pupil's KS2 attainment is the most important factor in predicting the outcome, as the predictive accuracy of the model increased 11.3% when attainment variables were added. Compared with the previous 5 A*-C model, in which school characteristics only account for 0.3% of the growth in prediction correctness, the influence of school-level characteristics is now 1.1%. However, adding school type in addition to the above variables does not impact prediction accuracy, which is similar to the results of the 5 A*-C model.

Table 11.10: Predictive accuracy of the logistic regression models of 5 GCSE and equivalent qualifications at grades A*-A

Variables	**Percentage correctness**	**Percentage of remaining variation explained**
Base figure	62.7	-
Personal backgrounds	64.6	1.9
KS2 total mark	75.9	11.3
School characteristics	77	1.1
School type (grammar school or not)	77	0.0
Overall	77	14.3

In addition to the prediction accuracy, the Exp(B) for grammar schools is also examined. In the previous models, which set 5 A*-C as the outcome, the results showed that grammar school pupils are more likely to reach the target, either with or without accounting for school-level variables. However, when 5 A*-A is set as the outcome variable, the evidence is mixed. In the three-stage model with no school-level characteristics, the results show that grammar school pupils are twice as likely as others to achieve high levels on the GCSE (Column B in Table 11.11). However, once school-level variables are entered, grammar school pupils' advantage, as revealed in the three-stage model, is eliminated, with the odds ratio dropping to 0.6 (Column A in Table 11.11). This demonstrates that grammar school pupils have a lower probability of achieving 5 A*-A grades than pupils in non-selective schools do, once personal and school compositional differences between these two groups are taken into account. The comparison between the two models shows that the assumed grammar school 'benefit' vanishes when their advantaged compositional characteristics are also controlled for.

Table 11.11: Logistic regression models of 5 GCSE and equivalent qualifications at grades A*-A with and without school compositional variables

	Column A: Model with school composition		**Column B: Model without school composition**	
Variable	**B**	**Exp (B)**	**B**	**Exp (B)**
Girls vs. Boys	0.529	1.697	0.531	1.701
Month Age	-0.026	0.974	-0.027	0.973
IDACI	-2.055	0.128	-2.77	0.063
FSM Eligible	-0.385	0.681	-0.465	0.628
SEN School Action	-0.202	0.817	-0.22	0.803
SEN School Action Plus	-0.253	0.776	-0.273	0.761
SEN Statement	-0.132	0.876	-0.107	0.899
EAL	0.539	1.714	0.533	1.704
Asian	0.388	1.474	0.376	1.456
Black	0.474	1.606	0.536	1.709
Chinese	0.861	2.366	0.979	2.661
Mixed	0.159	1.173	0.21	1.233
Unclassified	0.171	1.187	0.22	1.246
Other Ethnic Groups	0.811	2.251	0.865	2.375
KS2 total mark of English and maths	0.077	1.081	0.081	1.084
Mean KS2 Total Mark in Secondary School	0.032	1.033	-	-
Mean FSM Proportion in Secondary School	-0.007	0.993	-	-
Grammar school	-0.48	0.619	0.759	2.137
No. of Observation	70,683		70,683	

To conclude, logistic regression models predicting 5 A*-C and 5 A*-A grades on GCSE and the equivalent qualifications demonstrate that knowing whether a pupil went to grammar school does not increase the accuracy of the prediction for KS4 results in either case. This means that once pupil-level and school-level characteristics are known, school type does not have predict ability for pupil's KS4 performance, either for basic or top levels. Unlike the unimportance of school type in prediction accuracy, the odds ratios for grammar schools

remain obvious for both outcomes. After accounting for pupil-level and school-level characteristics, pupils in grammar schools are twice as likely to get 5 A*-C results. But no similar advantage is found in terms of 5 A*-A grades, and the probability of reaching this outcome is even lower in grammar schools than in equivalent non-selective schools. Therefore, while grammar schools may be good at the basics, benefiting their pupils at the lower end of the performance distribution, there is no evidence that they are more helpful in reaching high levels than non-selective schools when pupil and school-level variables are controlled for. The lower rate in grammar schools might partly result from underestimation of the grammar school effect based on the positive correlation between advantaged composition characteristics and good school practices such as more effective teaching (as mentioned in the methods chapter). However, the result still does not provide evidence that grammar schools help high-performing pupils reach high academic grades, which is contrary to the conventional perception. While it is usually asserted that grammar schools are more suitable for high-performing pupils, as grammar schools respond to their needs better than other state-funded schools do, the evidence shows that for pupils with high performance, grammar schools are not more beneficial than equivalent mixed-ability schools. According to previous research, this may be partly due to the negative influence of high-performing peer groups on academic self-perception (Marsh & Hau, 2003). Thus, comparing the two GCSE outcome thresholds reveals the possibility of a differential grammar school effectiveness for pupils at varying points on the performance distribution.

11.2.2.3 Logistic regression models including the number of GCSE (and equivalent) entered

As mentioned, the number of exams entered at KS4 is higher among grammar school pupils than among others. Therefore, this variable is added into the logistic regression models to assess whether the varied chances of achieving 5 or more GCSE and equivalent qualifications at grades A*-C and A*-A correlate with the number of exams entered.

For the results of 5 A*-C grades, adding the number of exams based on the previous four-stage model improves the overall predictive accuracy by 1.6% (Panel A of Table 11.12). Similarly, in the 5 A*-A model, adding this variable increases the total predictive accuracy by 1.3%, with final results reached 78.3% (Panel B in Table 11.12). This confirms that knowing the number of GCSE and equivalent qualifications entered is already predictive of the probability of reaching both outcomes.

Table 11.12: Predictive accuracy of logistic regression models of 5 GCSE and equivalent qualifications at grades A*-C and A*-A (including the number of exams)

Variables	Percentage correctness	Percentage of remaining variation explained
Panel A: 5 A*-C grades		
Base figure	61.5	-
Personal backgrounds	72.5	11
KS2 total mark	80.8	8.3
School characteristics	80.8	0
Number of exams entered	82.4	1.6
School type (grammar school or not)	82.4	0
Overall	82.4	20.9
Panel B: 5 A*-A grades		
Base figure	62.7	-
Personal backgrounds	64.6	1.9
KS2 total mark	75.9	11.3
School characteristics	77	1.1
Number of exams entered	78.3	1.3
School type (grammar school or not)	78.3	0.0
Overall	78.3	15.6

In addition to the predictive accuracy, adding the number of exams entered also modifies the odds ratios for school types in both models. In terms of 5 A*-C grades, the odds ratio for school type decreased from 2.1 to 1.9 (Column A in Table 11.13). The lower results for grammar school pupils shows that the higher number of tests entered in grammar schools explains some of their advantages, but only to a mild extent. However, adding the number of exams into the 5 A*-A model brings even smaller changes to the odds ratio for school type, which increased from 0.62 (Column A in Table 11.13) to 0.64 (Column B in Table 11.13). This is partly because that the 5 A*-A model only includes pupils with high KS2 attainment. For pupils who scored above 134 at KS2, the number of exams entered in grammar schools and non-selective schools is similar. Therefore, although adding this variable has increased the prediction accuracy, it does not bring substantial changes to the estimated difference between pupils in grammar schools and others, in terms of their varied chances of achieving 5 GCSE and equivalent qualifications at A*-A.

Table 11.13: Logistic regression models of 5 GCSE and equivalent qualifications at grades A*-C and A*-A (including the number of exams)

	Column A: 5 A*-C		**Column B: 5 A*-A**	
Variable	**B**	**Exp (B)**	**B**	**Exp (B)**
Girls vs. Boys	0.463	1.589	0.518	1.678
Month Age	-0.029	0.971	-0.026	0.975
IDACI	-1.614	0.199	-2.029	0.132
FSM Eligible	-0.378	0.685	-0.313	0.731
SEN School Action	-0.373	0.689	-0.135	0.874
SEN School Action Plus	-0.38	0.684	-0.165	0.848
SEN Statement	-0.063	0.939	0.03	1.03
EAL	0.357	1.429	0.416	1.516
Asian	0.202	1.224	0.455	1.575
Black	0.363	1.438	0.435	1.545
Chinese	0.891	2.437	0.777	2.175
Mixed	-0.02	0.98	0.156	1.169
Unclassified	0.269	1.308	0.143	1.154
Other Ethnic Groups	0.417	1.518	0.744	2.103
KS2 total mark of English and maths	0.048	1.049	0.075	1.078
Mean KS2 Total Mark in Secondary School	0.009	1.009	0.029	1.03
Mean FSM Proportion in Secondary School	-0.002	0.998	-0.001	0.999
Number of exams entered	0.54	1.716	0.481	1.618
Grammar school	0.618	1.855	-0.448	0.639
No. of Observation	149,072		70,683	

In the models for both 5 A*-C and 5 A*-A, the odds ratios for the number of exams taken are greater than 1. This confirms that those who took more exams at KS4 have a higher probability of reaching high levels than equivalent pupils. However, it can also be interpreted that pupils who took exams for more subjects are those with higher abilities, when all other factors are equal. Therefore, the higher number of exams taken in grammar schools is likely due to their pupils' better performance, as pupils with higher performance are likely to sit more tests.

Meanwhile, the difference may be also related to the stricter restrictions on the total number of exams in some non-selective schools.

In sum, while the number of exams taken is predictive of pupils' chances of achieving basic and top levels at KS4, this variable only explains some of the differences between grammar school pupils and those in other schools.

11.2.3 Conclusion

Evidence from logistic regression models controlling for pupil-level and school-level characteristics has revealed that grammar schools may be more effective in helping pupils get good passes at KS4 than non-selective schools are, but not in reaching top GCSE grades. Grammar school pupils are more likely than pupils from other schools in selective LAs to achieve 5 A*-C grades at KS4. This difference remains mostly intact even when their higher number of exams taken is factored into account. However, similar evidence for reaching 5 top GCSE grades was not found. This reveals that there may be positive practices in grammar schools which produce satisfactory basic KS4 levels for borderline pupils, but are less effective at helping high performing pupils achieve top levels.

11.3 The regression discontinuity design of the effectiveness of grammar schools

Unlike previous analysis using traditional regression models to control for pre-existing differences between grammar school pupils and others, this part evaluates the effectiveness of grammar schools using the RDD approach. While the previous analysis discusses the national pattern, the RDD part only covers one selective LA where the data of the 11+ is available.

11.3.1 Descriptive results

As mentioned in the methods chapter, many valid cases in the 11+ files were not included for analysis. Therefore, the characteristics of the selected sample are first contrasted with the original cohort in the 11+ file. As presented in Table 11.14, the characteristics of the sample group are similar to the population data in the 11+ file. This is apparent in terms of academic performance, as the KS2 performance and the 11+ test scores are nearly identical between the two groups. However, the population in the 11+ file has a more advantaged average IDACI score and a lower proportion of FSM pupils. In order to minimise the influence of the difference in IDACI and FSM between the sample and the population, analysis has also been conducted

to randomly delete cases from the sample to keep the average IDACI score and the proportion of FSM pupils consistent with the 11+ file. However, the two sample sets yield similar results and lead to a same conclusion. Therefore, the sample group is not trimmed further in order to keep as many valid cases as possible.

Table 11.14: Characteristics of pupils in grammar schools (GS) and non-selective schools in the RDD

	KS2 average points		**Total score on the 11+**		**IDACI**		**FSM (%)**	
	GS	Others	GS	Others	GS	Others	GS	Others
Cases in the RDD	32.5	29	392	326	0.14	0.18	5	19
Cases in the 11+ file	32.7	29.2	392	327	0.13	0.17	3	12

Among the sample group, 40% (1,043) of the local pupils who sat the 11+ in 2011 were assessed as suitable to attend grammar schools. This included 51 pupils who did not cross the lowest threshold for total score. Meanwhile, there are also 170 pupils who reached the threshold for total score but still did not passed the selection, mainly due to inadequate scores in individual subjects. The difference between those who passed the selection and those who did not is clear, with the former having more advantageous results, both in terms of their academic performance at two key stages, and their demographic characteristics (Table 11.14).

Despite the pre-existing differences between grammar school pupils and their counterparts, crossing the cut-off point does not cause simultaneous discontinuities in baseline variables. Figure 11.4 is an example of the similar demographic features of pupils just above and below the cut-off point. When the average IDACI score at each point of the assignment variable is plotted, neither the binned average value, nor the fitted regression lines, presents discontinuity at the cut-off point. This means there is no systematic difference in IDACI scores between pupils who just reached the threshold and those who just missed it. Other baseline variables, FSM and KS2 performance, are also similar to Figure 11.4, with the results just above and below the cut-off point being very close. This proves the irrelevance of background variables in estimating the treatment effect at the cut-off point in this study.

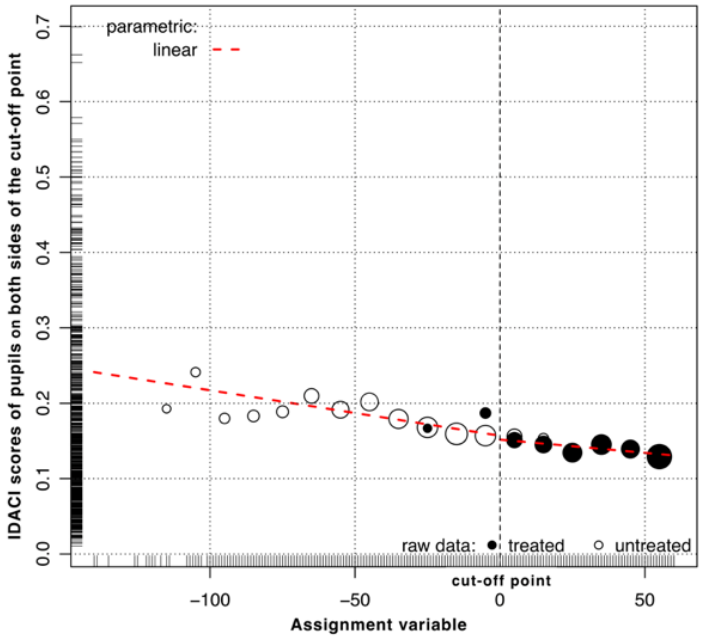

Figure 11.4: IDACI scores of pupils in grammar schools and non-selective schools in the RDD

In order to test the internal validity of the RDD further, the frequency of the assignment variable is also checked. If pupils can have accurate manipulation over the assignment variable, we could anticipate that the frequency just above the cut-off point would be higher than below the cut-off point. The frequency at each score is plotted in Figure 11.5. The graph shows a ceiling effect at the right end. Since 60 is the highest possible value in the assignment variable, it also contains pupils who might have achieved higher scores otherwise. Despite this outlier, the distribution of the assignment variable is smooth. There is no evidence that pupils can have full control over their test scores, with the frequencies just above and below the cut-off point being similar.

Although the assignment variable is not the only deciding factor of pupils' eligibility to attend grammar school, the jump in the probability of treatment is still strong at the cut-off point. As presented in Figure 11.6, the probability of passing the selection is near zero before point -10. The rate grows from point -10 and increases to about 0.4 at the cut-off point. The probability reaches 1 at point 13 and stabilises after point 30. Overall, the probability of going to grammar schools increased from near 0 to 1 within the small interval of -10 to 13.

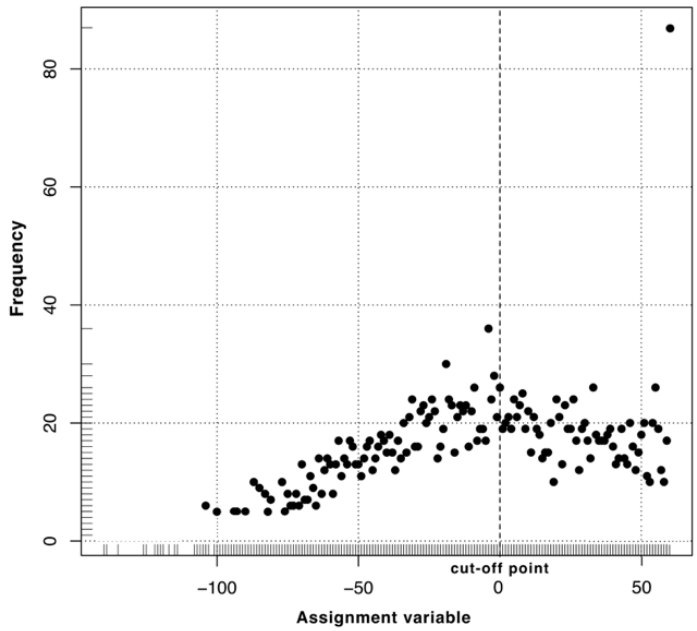

Figure 11.5: Frequency of the assignment variable in the RDD

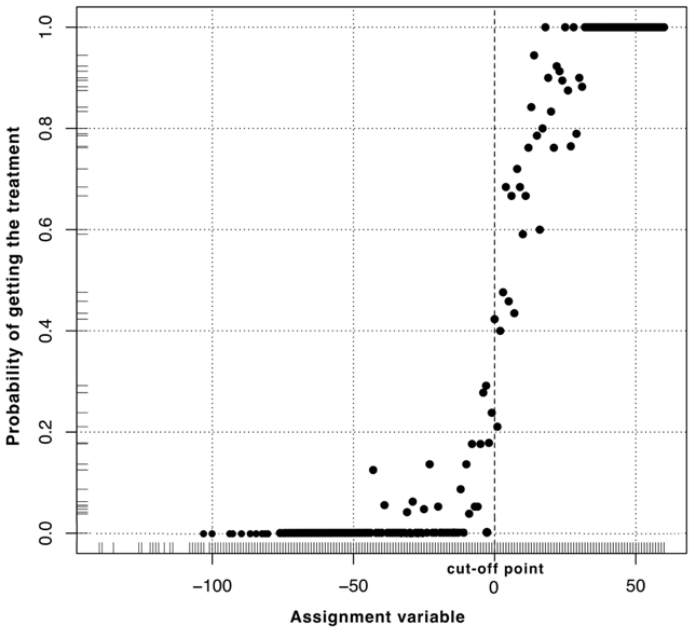

Figure 11.6: Probability of getting the treatment in the RDD

The relationship between the outcome variable and the assignment variable is also depicted. A visual inspection of the average GCSE results at each score of the assignment variable shows that these two variables are positively correlated (Figure 11.7). The distribution on the left lower side is irregular, which is the result of the dearth of observations at these scores. Due to the concentrated points, a visual inspection of the raw data reveals little discontinuity in the outcome variable at the cut-off point. However, when observations are grouped into bins and the number of observations within each bin is represented by the size of the dot, there is a

discontinuity at the cut-off point (Figure 11.8). The following sections test whether the graphic discontinuity can be regarded as the treatment effect of grammar schools.

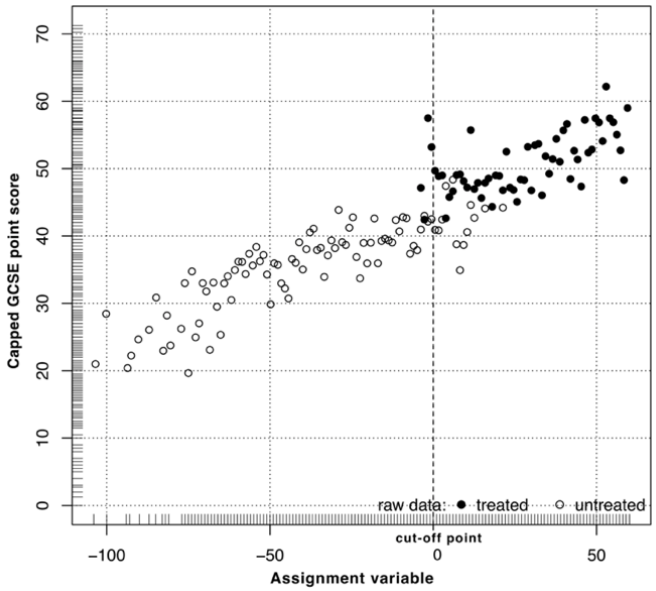

Figure 11.7: The relationship between the assignment variable and the outcome variable in the RDD (raw data)

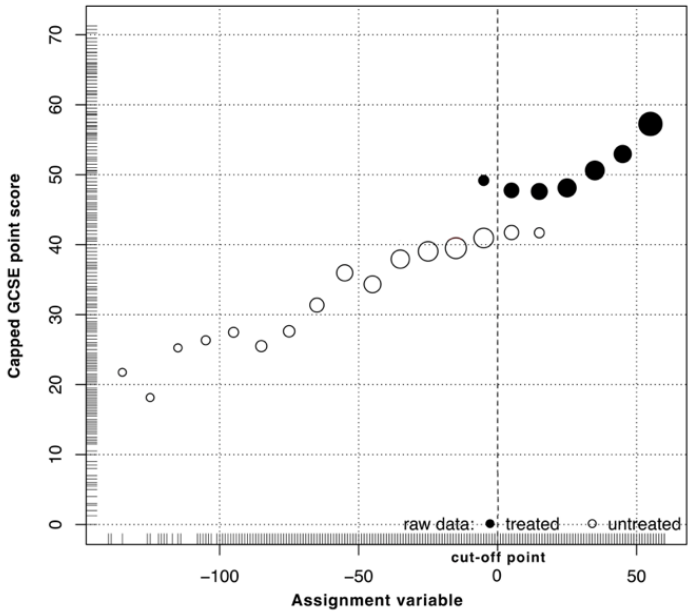

Figure 11.8: The relationship between the assignment variable and the outcome variable in the RDD (grouped data)

11.3.2 The parametric approach of RDD

After confirming the similar characteristics between pupils who just made the threshold and those who did not, the following analysis estimates the treatment effect of grammar schools based on the parametric approach. Table 11.15 presents the detailed estimation results. The first model calculates the treatment effect by fitting linear functions on both sides of the cut-off point, and in both stages of the regression in the 2SLS. The functional form used in the second model also includes interaction terms to allow slopes to vary on both side of the cut-off point (as described in the methods chapter). In the third and fourth models, quadratic and quadratic interaction functions are fitted.

Table 11.15: The parametric estimation of the grammar school effect in the RDD (without controlling for pupil-level baseline variables)

	R-square of the model	**Treatment effect**
Model 1 (Linear)	0.363	0.538
Model 2 (Linear interaction)	0.365	-0.395
Model 3 (Quadratic)	0.363	-1.180
Model 4 (Quadratic interaction)	0.370	4.572

As shown in Table 11.15, the estimates of the treatment effect vary when different functional forms are used. Based on the calculation of the first three models, the treatment effect is small. It is in fact negative in Model 2 and Model 3. Meanwhile, Model 4 not only has a much larger effect size, it also reveals that attending grammar schools is beneficial, which is about four GCSE point scores. This is equivalent to 0.57 C-E grade or 0.38 A*-C grade per GCSE subject.

The real effect of grammar schools depends on which model presents a more convincing result. As shown in Table 11.15 above, Model 4 has the highest R-square value, but the difference is subtle. However, when using the specification test suggested by Lee and Lemieux (2010), Model 4 is the only one that passes. This suggests that there is unexplained variability missing from Model 1-3. Meanwhile, Model 4 also yields the best result evaluated by the Akaike information criterion (AIC). While it is possible that when a functional form gives more

parameters, the fitness of the model will inevitably increase, a graphic presentation also suggests the fitness of a quadratic interaction function (Figure 11.9). The estimate is supported by the robustness check and the non-parametric approach in later sections as well.

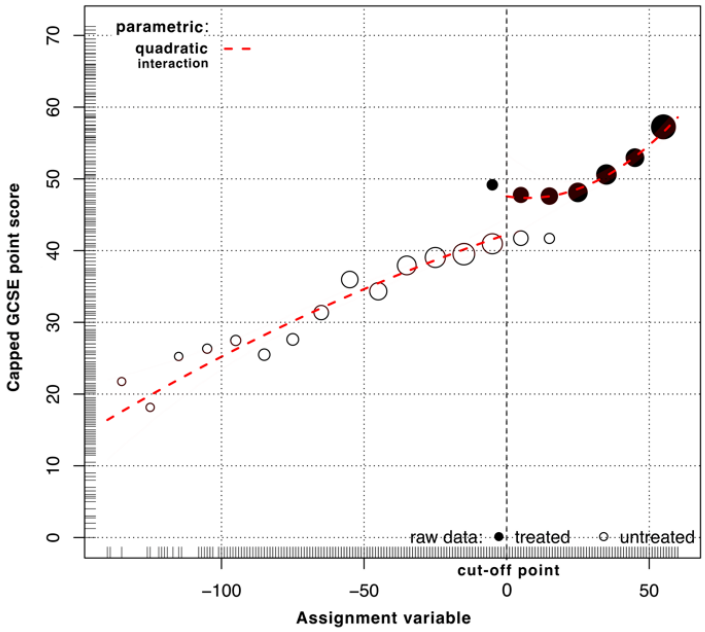

Figure 11.9: The grammar school effect in the RDD parametric approach (quadratic interaction regression lines superimposed)

Table 11.16: The parametric estimation of the grammar school effect in the RDD (including pupil-level baseline variables, FSM, IDACI, and EAL)

	R-square of the model	**Treatment effect**
Model 5 (Linear)	0.377	0.595
Model 6 (Linear interaction)	0.378	-0.429
Model 7 (Quadratic)	0.376	-1.224
Model 8 (Quadratic interaction)	0.384	5.209

In Model 5-8 (Table 11.16), pupils' FSM eligibility, IDACI score and EAL status are included as baseline covariates. The distinct estimates between models are still clear when different functional forms are used, and including these baseline variables increases the effect size in

each model without changing the direction. Overall, the results in Model 5-8 correspond to each of the model in Table 11.15. This means that whether or not the baseline variables in the RDD are included, the results are similar when the same functional form is applied.

11.3.3 Robustness check of the parametric approach of RDD

In order to assess whether the estimated treatment effect is sensitive to the changes in the data (especially cases with extremely high and low values in the assignment variable), a robustness check trimmed the 10% outmost data points at the two farthest ends of the assignment variable.

Table 11.17: The robustness check of the parametric estimation in the RDD (trimmed the outermost 10% on both ends of the cut-off point)

	R-square of the model	**Treatment effect**
Panel A: Models without pupil-level baseline variables		
Model 1 (Linear)	0.202	2.847
Model 2 (Linear interaction)	0.203	2.684
Model 3 (Quadratic)	0.203	2.212
Model 4 (Quadratic interaction)	0.203	5.039
Panel B: Models with pupil-level baseline variables (FSM, IDACI and EAL)		
Model 5 (Linear)	0.218	3.022
Model 6 (Linear interaction)	0.219	2.824
Model 7 (Quadratic)	0.219	2.276
Model 8 (Quadratic interaction)	0.220	5.554

Overall, the treatment estimates experienced some changes when observations with the highest and lowest values in the assignment variable were excluded (Table 11.17). The estimated treatment effects increased to about 2 point scores in Model 1-3, which was much larger than in the original models in Table 11.15. Meanwhile, the direction of the coefficient was also

altered in Models 2 and Model 3. The results in Model 4 also grew, but only to a mild extent. After trimming 10% of the cases, the treatment effect in Model 4 equals 5 point scores. This is close to the estimated treatment effect in the original model, which is 4.57 (Table 11.15). In Panel B of Table 11.17, the robustness check was conducted on models with demographic variables. The estimated treatment effects remained close to Panel A of Table 11.17, revealing again the irrelevance of including baseline variables in a valid RDD. However, the treatment effects in Model 5-7 are inconsistent with the original results in Table 11.16, and Model 8 is the only one remaining close to its untrimmed result. Therefore, the unstable results in Model 1-3/ Model 5-7 and the similar results in Model 4/ Model 8 before and after the data trimming reveals the better fit of a quadratic interaction function in depicting the sample data again. Among all the functional forms, it is the least sensitive to the changes in the data.

11.3.4 The non-parametric approach of RDD

To confirm whether the treatment effect in the parametric approach is convincing, a non-parametric approach using local data points within a bandwidth on both sides of the cut-off point was also applied. As mentioned previously, the optimal bandwidth was decided based on the calculation proposed by Imbens and Kalyanaraman (2012). According to this principle, the optimal bandwidth in this study is 41.3. Based on the non-parametric estimation, the treatment effect is 4.32. This is close to the parametric results in Model 4 in Table 11.15, which is 4.57. When baseline covariates are included in the non-parametric approach, the estimated treatment effect grows slightly to 4.62. This is also similar to the corresponding parametric results in Model 8 in Table 11.16, which is 5.21. Therefore, the non-parametric approach yields results similar to those of the parametric approach using quadratic interaction functional forms.

A sensitivity test of the non-parametric estimation was also conducted to assess how stable the results are when different bandwidths are used. As revealed in Figure 11.10, the non-parametric estimation is negative when the bandwidth is below 20, which is the interval where the probability of getting the treatment spikes. The estimated treatment effect grows as the bandwidth widens, and stabilises around 4 point scores within a bandwidth of 30 to 50. The result decreases slightly after bandwidth 50, but remains positive until 60. Results of bandwidth larger than 60 are not presented, as it is already the largest possible size (the same bandwidth is selected on both sides of the cut-off point, and 60 is the maximum assignment value on the

right side). Overall, the treatment effect based on the given optimal bandwidth is stable within a large interval.

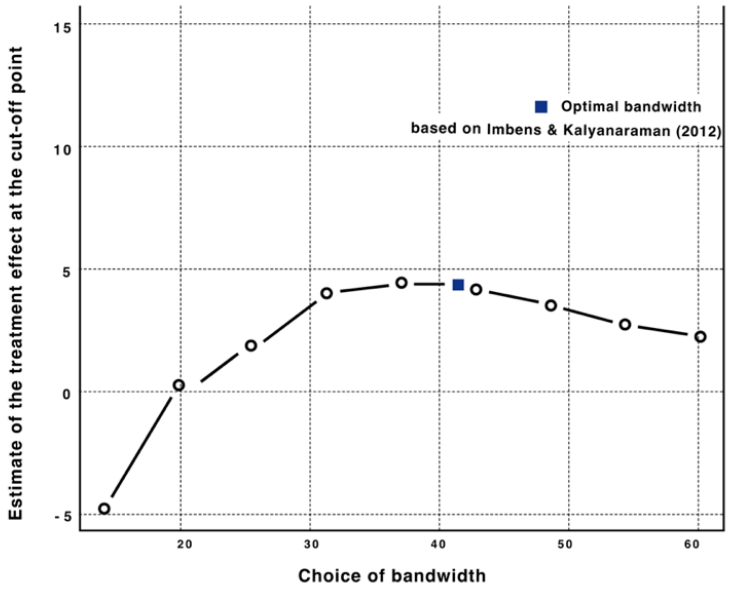

Figure 11.10: Bandwidth selection and the non-parametric estimation in the RDD

11.3.5 Generalisability of the RDD estimates

As discussed in the literature chapters, the degree of generalisability of the result at the cut-off point is largely related to the quality of the assignment variable. While it is not possible to know how close the estimation is from the overall average treatment effect, it is practical to examine the characteristics of pupils in the neighbourhood of the cut-off point. If the composition near the cut-off point is heterogeneous, the subgroup will be more similar to the population in the full data range, and the result at the cut-off can be applied to a wider scope. In order to see whether this subgroup is similar to the whole population, the analysis compared the KS2 test results between two groups. For most of the pupils who scored right at the cut-off point in this study, their average KS2 points were 30 (42%) and 33 (54%). Meanwhile, the proportion of pupils in this LA with these two KS2 points is 47%. For pupils whose scores in the assignment variable are within a ±10 interval of the cut-off point, the average KS2 points for the majority are 27 (10%), 30 (48%) and 33 (40%). This overlaps with the KS2 performance level of 77% of the pupils in this LA. While this may have been due to the low discriminative ability of KS2 points, the results are consistent when KS2 average marks are used for comparison. Therefore, the estimated treatment effect is at least relevant to pupils with the above mentioned KS2 academic levels. Meanwhile, if a differential grammar school effect

exists, the grammar school effect could be less pronounced for pupils with high prior attainment (Atkinson, Gregg, & McConnell, 2006; Levačić & Marsh, 2007). In this case, the treatment effect estimated at the cut-off point in this study would be larger than the effect for pupils with higher KS2 attainment, especially for those who scored above 33.

Besides the generalisability of the estimates within this selective LA, it is also meaningful to discuss whether the treatment effect of grammar schools in this LA is relevant to the national pattern in England. Based on the evaluation of pupils' attainment and background characteristics, the intake of grammar schools in this LA is similar to other selective LAs. However, due to the imbalance in grammar school places and different selection processes in each area, the nature of grammar schools in this LA may still differ from others. Since the content and threshold of the selection test varies across LAs, the discontinuity gap in this LA would not be informative about the grammar school effect elsewhere if the selection aims to pick the highest-performing 5% or 10% of pupils in a given year group. Additionally, the broader social context of each area may also influence the effectiveness of grammar schools and non-selective alternatives for pupils not in grammar schools. This means that the evidence in this study may diverge from the national pattern.

11.3.6 Conclusion

In sum, applying the RDD approach in one single LA, the analysis reveals a positive grammar school effect for their borderline pupils in terms of KS4 attainment. The conclusion of the RDD in this LA is similar to the national pattern of achieving 5 good passes at KS4 as revealed in the logistic regression models previously, suggesting that grammar schools might be beneficial for their borderline pupils. However, due to the limitation of the 11+ data and the very different social context of each LA in England, the evidence in this single LA is unable to provide a definite answer to the effectiveness of grammar schools. It is more about the feasibility of applying a stronger research design to estimate the grammar school effect.

12 Findings about the differential effectiveness of grammar schools for FSM pupils

The smaller attainment gap in grammar schools is sometimes used to support the claim of grammar schools' role in narrowing academic performance between pupils from high and low SES backgrounds (Andrews, Hutchinson & Johnes, 2016). Therefore, this analysis addresses grammar schools' differential effectiveness for FSM pupils.

12.1 A raw comparison of FSM and non-FSM pupils' performance

In order to test whether grammar schools are working well for disadvantaged pupils, the capped GCSE point scores of FSM and non-FSM pupils are compared first. A comparison of raw test scores confirms that the attainment gap between FSM and non-FSM pupils is small in grammar schools—2.6 point scores in capped GCSE. This is lower than both the between-group gap in non-selective schools in selective LAs, and the average figure in comprehensive LAs, which are 7.3 and 7.7 respectively. However, due to academic selection, the gap in KS2 attainment between FSM and non-FSM pupils is already narrow in grammar schools. While the difference in KS2 performance between these two groups in grammar schools is only 3 marks (162 vs. 165), the equivalent comparison in other schools presents a considerable gap—16 marks for non-selective schools in selective LAs, and 17 marks in comprehensive LAs.

If we choose pupils in non-selective LAs who have KS2 performance levels similar to those of grammar school pupils, the gap in GCSE results between FSM and non-FSM pupils drops. Limiting the comparison to a group of FSM pupils with an average of 162 KS2 marks, and a group of non-FSM pupils with an average of 165 KS2 marks in non-selective LAs, the result shows that the gap in the capped GCSE point scores between FSM and non-FSM shrank, with the difference now only being 4 point scores. This is close to the attainment gap in grammar schools. Since pupils with similar KS2 performance also tend to reach similar KS4 levels regardless of the type of school attended, the small KS4 attainment gap in grammar schools is largely a result of the homogeneous KS2 performance level of their intakes. Additionally, the comparison also reveals that for high-performing pupils, eligibility for FSM does not have substantial influence on their academic results, at least in secondary school. The findings are similar for different types of schools.

In addition to the raw comparison of the performance between FSM and non-FSM pupils in different types of schools, regression models are also applied to test whether the conclusion is

consistent. The analysis applied linear regression and logistic regression models to assess whether there is evidence that grammar schools are especially effective for FSM pupils.

12.2 Evidence from OLS linear regression models

The OLS models testing grammar school effectiveness for FSM and non-FSM pupils apply the capped GCSE point scores as the outcome variable, and control for a set of pupil-level and school-level baseline variables that are the same as in the previous general models of grammar schools' effectiveness. The analysis is first conducted on FSM and non-FSM pupils separately (Models 1 and 2 in Table 12.1). Adding school type does not increase the R-square for either group (thus, the classification table is not included), which is the same as the general pattern for grammar school effectiveness. Meanwhile, the coefficient for school type is also examined. In the model for FSM pupils, the coefficient for school type is 1.7, demonstrating an advantage for grammar school pupils when other factors are equal. This is higher than the coefficient in models for non-FSM pupils in which the difference between grammar schools and non-selective schools is only 0.32.

Instead of treating the two groups of pupils separately, the second approach adds an interaction term, FSM*Grammar School, to the OLS model (Model 3 in Table 12.1). Including this interaction variable does not increase the fitness of the model either, which remains 0.59. Meanwhile, after adding this interaction variable, the coefficient for school type becomes 0.28. In contrast to the small figure, the coefficient for FSM*Grammar School is considerable—0.9 in capped GCSE. This indicates that the advantage of attending grammar schools for FSM pupils is 0.9 point score higher than for non-FSM pupils. Therefore, for pupils eligible for FSM, the average advantage of attending grammar schools is about 1.2 point scores over equivalent FSM pupils in non-selective state-funded schools in these 36 LAs. Like the difference between Models 1 and 2 (Table 12.1), when the interaction term is used, the estimation result still demonstrates that attending grammar school matters more for FSM than for non-FSM pupils.

Table 12.1: OLS models predicting capped GCSE point scores (FSM vs. non-FSM)

	Model 1: FSM only	**Model 2: Non-FSM only**	**Model 3: All pupils**
Variable			
Girls vs. Boys	2.563	2.273	2.303
Month Age	-0.082	-0.11	-0.107
IDACI	-3.124	-5.895	-5.823
FSM Eligible	-	-	-2.647
SEN School Action Plus or SEN Statement	-2.182	-1.799	-1.897
EAL	4.321	2.5	2.806
Asian	2.135	1.027	1.121
Black	3.595	1.574	1.967
Others	1.992	0.802	0.964
KS2 total mark of English and maths	0.219	0.231	0.23
Mean KS2 Total Mark in Secondary School	0.068	0.07	0.073
Mean FSM Proportion in Secondary School	-0.063	-0.058	-0.053
Grammar School	1.718	0.324	0.283
FSM* Grammar School	-	-	0.905
No. of Observation	17,482	131,590	149,072

12.3 Evidence from logistic regression models

In addition to the linear regression models, logistic regression models accounting for the same sets of pupil and school characteristics are also applied. According to the overall pattern for selective LAs, there is only a grammar school effect for the 5 A*-C results. Therefore, the 5 A*-A results are not discussed further in this section.

Separate analysis is still first conducted on FSM and non-FSM pupils. The results show that adding school type to the logistic regression models does not improve the overall predictive accuracy of achieving 5 A*-C grades, despite FSM eligible status (thus, the classification tables are not included). Meanwhile, the Exp (B) of school type shows that the difference linked with

attending grammar schools is more pronounced for FSM pupils than for non-FSM pupils (Models 1 and 2 in Table 12.2). The odds ratio for school type is 3.3 for FSM pupils, and 2.1 for non-FSM pupils. Meanwhile, the evidence using the interaction term for FSM*Grammar School produces the same results (Model 3 in Table 12.2).

Table 12.2: Logistic regression models of 5 GCSE and equivalent qualifications at grades A*-C (including English and maths) for FSM and non-FSM pupils

	Model 1: FSM only		**Model 2: Non-FSM only**		**Model 3: All pupils**	
Variable	**B**	**Exp (B)**	**B**	**Exp (B)**	**B**	**Exp (B)**
Girls vs. Boys	0.428	1.534	0.455	1.576	0.449	1.567
Month Age	-0.024	0.976	-0.032	0.968	-0.031	0.969
IDACI	-0.82	0.44	-1.585	0.205	-1.536	0.215
FSM Eligible	-	-	-	-	-0.481	0.618
SEN School Action	-0.527	0.59	-0.414	0.661	-0.432	0.649
SEN School Action Plus	-0.5	0.607	-0.531	0.588	-0.53	0.589
SEN Statement	-0.178	0.837	-0.352	0.703	-0.33	0.719
EAL	0.873	2.394	0.424	1.527	0.507	1.66
Asian	0.164	1.178	0.17	1.185	0.159	1.172
Black	0.554	1.74	0.322	1.379	0.38	1.462
Chinese	1.507	4.511	0.993	2.7	1.011	2.747
Mixed	0.166	1.181	-0.029	0.971	-0.003	0.997
Unclassified	0.969	2.635	0.14	1.15	0.257	1.293
Other Ethnic Groups	0.45	1.569	0.48	1.615	0.496	1.642
KS2 Total Mark of English and maths	0.049	1.05	0.053	1.054	0.052	1.054
Mean KS2 Total Mark in Secondary School	0.005	1.005	0.006	1.006	0.007	1.007
Mean FSM Proportion in Secondary School	-0.01	0.99	-0.011	0.989	-0.009	0.991
Grammar School	1.182	3.262	0.733	2.082	0.74	2.095
FSM* Grammar School	-	-	-	-	0.223	1.249
No. of Observation	17,482		131,590		149,072	

12.4 Conclusion

In sum, the estimation results reveal that the positive coefficients for grammar school in different models are always greater for FSM than for non-FSM pupils. However, it should be noted that when analysing the differential grammar school effectiveness for FSM and non-FSM pupils, the results are the relative performance of grammar school pupils in relation to their peers in non-selective schools. Therefore, while the high estimates for FSM pupils at grammar schools may be evidence that grammar schools work especially well for this group of disadvantaged pupils, it is likely a result of the lower effectiveness of non-selective schools attended by FSM pupils. It has been reported that pupils from poorer backgrounds are less likely to attend high-performing schools, but schools rated as 'inadequate' by Ofsted usually have more FSM pupils (Tes, 2018). This means that for all the pupils in non-selective schools, FSM pupils are more likely to be enrolled in schools of lower quality than non-FSM pupils do. Therefore, even if grammar schools are equally effective for all their pupils regardless of FSM eligibility, the attainment gap between grammar schools and non-selective schools may still be more pronounced for the FSM group. Furthermore, it has also been mentioned that this tiny group of FSM pupils in grammar schools are systematically different from most FSM pupils, even after controlling for prior attainment and background. Therefore, the favourable statistical results for FSM pupils in grammar schools should not be simply regarded as the differential effectiveness of grammar schools, as these confounding factors need to be considered. After assessing the effectiveness of grammar schools for FSM pupils, the next chapter turns to the influence of the entire selective system.

13 Findings about the effectiveness of selective LAs in raising pupils' academic performance

The existence of grammar schools not only influences the academic trajectories of their own pupils, but also pupils in surrounding schools. Therefore, evaluation of grammar schools' effectiveness alone does not present a complete picture. Thus, the potential impact of the selective system on the overall academic performance of the local area is also examined.

13.1 Results from OLS models

This chapter focuses on the effectiveness of selective LAs in comparison with non-selective LAs, based on the results of OLS and logistic regression models. The effectiveness of selective LAs and non-selective LAs is first compared using OLS models.

The regression models of the effectiveness of selective LAs are very similar to the previous models of grammar schools' effectiveness. Apart from the new variable, LA type, and the categorical variable, school type, the other baseline variables and the outcome variable are the same as those in the previous chapters on grammar schools' effectiveness. According to this new set of OLS model, adding LA type does not change the overall R-square, revealing the irrelevance of this variable in predicting pupils' KS4 performance. Meanwhile, the coefficients in the OLS models also reveal that for equivalent pupils, whether they attend secondary schools in selective LAs or in non-selective LAs does not influence their capped GCSE results (Model 1 in Table 13.1). Therefore, the effectiveness of the two systems is similar. As the coefficient for selective LAs is just below zero (-0.1), these areas may be less effective than comprehensive areas.

The second model focuses on the internal differences within selective LAs. This reveals the effectiveness of two types of schools within selective LAs in comparison with schools in comprehensive LAs. Adding the categorical variable of school type (grammar school / non-selective schools / schools in comprehensive areas), Model 2 in Table 13.1 compares the results of grammar schools and non-selective schools in selective areas, against schools in comprehensive areas (the reference group). Adding this new variable, there is no increase in predictive accuracy. According to Model 2, grammar school pupils do slightly better than equivalent pupils in comprehensive LAs, but pupils in selective LAs who failed to get into grammar school do a little worse than those in comprehensive LAs. When combining these

two mixed effects, the overall effectiveness of selective LAs demonstrates no difference from comprehensive areas. Therefore, the results of Model 2 correspond to those of Model 1.

Table 13.1: The OLS models of capped GCSE point score in selective and non-selective LAs

	Capped GCSE point score	
Variable	**Model 1 (LA type)**	**Model 2 (School type)**
Girls vs. Boys	2.434	2.434
Month Age	-0.11	-0.11
IDACI	-6.533	-6.557
FSM Eligible	-2.706	-2.705
SEN School Action Plus and Statement	-1.785	-1.786
EAL	3.287	3.289
Asian	1.201	1.196
Black	1.672	1.676
Other Ethnic Groups	0.946	0.949
KS2 total mark of English and maths	0.234	0.234
Mean KS2 Total Mark in Secondary School	0.075	0.072
Mean FSM Proportion in Secondary School	-3.108	-3.2
Selective LA	-0.103	-
Grammar school	-	0.146
Non-selective school in selective LA	-	-0.131
No. of Observation	481,681	481,681

13.2 Results from logistic regression models

In addition to OLS models, logistic regression models are also used to assess the difference between selective and non-selective LAs. The analysis presents two outcome variables, which are achieving 5 GCSE (and equivalent qualifications) at grades A*-C, and grades A*-A.

The analysis first presents the classification tables of both outcomes. According to Table 13.2, the patterns for the two outcome variables are consistent. Thus, knowing LA type or school type has little bearing on the predictive ability of the models, either in terms of achieving 5 good passes or 5 top grades at KS4.

Table 13.2: Predictive accuracy of the logistic regression models of KS4 attainment in selective and non-selective LAs

Variables	**Percentage correctness**	**Percentage of remaining variation explained**
Panel A: 5 A*-C grades		
Base figure	60	-
Personal backgrounds	71.3	11.3
KS2 total mark	79.7	8.4
School characteristics	79.8	0.1
LA type / School type	79.8	0.0
Overall	79.8	19.8
Panel B: 5 A*-A grades (KS2>134)		
Base figure	66	-
Personal backgrounds	66.6	0.6
KS2 total mark	76	9.4
School characteristics	76.5	0.5
LA type / School type	76.6 / 76.5	0.1 / 0
Overall	76.6 / 76.5	10.6 / 10.5

In addition to the predictive accuracy, the odds ratios for the two models are also presented. In terms of reaching 5 A*-C grades, the results demonstrate the two systems' similar effects with the odds ratio of selective LAs compared to comprehensive LAs being 0.991 (Model 1 in Table 13.3). When the categorical variable of school type is used to contrast grammar schools and non-selective schools in selective LAs against schools in comprehensive LAs, the results are similar to those of the previous OLS models. While grammar school attendance is associated with advantages over equivalent schools in the comprehensive system, there are negative consequences associated with attending non-selective schools in selective LAs. This corresponds to a zero-sum situation for the overall effectiveness of the selective system, as presented in Model 1 (Table 13.3). While the general rates of achieving good passes at KS4 are similar in both systems, the internal variance should be smaller and the performance distribution is likely to be fairer in comprehensive areas.

Table 13.3: Logistic regression models of KS4 attainment in selective and non-selective LAs

	5 A*-C grades at KS4				5 A*-A grades at KS4			
	Model 1		Model 2		Model 3		Model 4	
Variable	**B**	**Exp (B)**	**B**	**Exp (B)**	**B**	**Exp (B)**	**B**	**Exp (B)**
Girls vs. Boys	0.454	1.574	0.455	1.576	0.534	1.706	0.532	1.702
Month Age	-0.03	0.971	-0.03	0.971	-0.028	0.972	-0.028	0.972
IDACI	-1.635	0.195	-1.661	0.19	-2.263	0.104	-2.227	0.108
FSM Eligible	-0.496	0.609	-0.494	0.61	-0.392	0.676	-0.394	0.674
SEN School Action	-0.434	0.648	-0.432	0.649	-0.212	0.809	-0.212	0.809
SEN School Action Plus	-0.498	0.607	-0.497	0.608	-0.16	0.852	-0.16	0.852
SEN Statement	-0.408	0.665	-0.403	0.668	-0.198	0.82	-0.202	0.817
EAL	0.543	1.721	0.544	1.723	0.675	1.965	0.672	1.958
Asian	0.247	1.28	0.242	1.274	0.425	1.53	0.432	1.541
Black	0.348	1.416	0.352	1.422	0.432	1.54	0.427	1.533
Chinese	0.715	2.044	0.708	2.03	0.928	2.531	0.93	2.535
Mixed	0.042	1.043	0.045	1.046	0.195	1.215	0.19	1.21
Unclassified	0.162	1.176	0.165	1.18	0.129	1.138	0.122	1.13
Other Ethnic Groups	0.508	1.661	0.511	1.667	0.712	2.038	0.708	2.03
KS2 Total Mark of English and maths	0.052	1.054	0.052	1.054	0.08	1.083	0.08	1.083
Mean KS2 Total Mark in Secondary School	0.011	1.011	0.007	1.007	0.022	1.022	0.026	1.026
Mean FSM Proportion in Secondary School	-0.389	0.678	-0.487	0.614	-0.762	0.467	-0.642	0.526
Selective LAs	-0.009	0.991	-	-	-0.135	0.874	-	-
Grammar School	-	-	0.722	2.06	-	-	-0.353	0.703
Non-selective Schools in Selective LAs	-	-	-0.041	0.96	-	-	-0.099	0.906
No. of Observation	481,681		481,681		216,099		216,099	

Unlike the conclusion of achieving 5 A*-C grades at KS4, which reveals the equal status between the selective and non-selective systems, the models for 5 A*-A grades present a negative result, demonstrating that the selective system may harm the overall chances of reaching high levels at KS4. According to Models 3 and 4 in Table 13.3, the odds ratios for selective LAs are lower than 1. The overall effectiveness of the selective system is about 90% of the non-selective system in achieving top grades at KS4. Meanwhile, pupils either in grammar schools or non-selective schools in selective LAs have a lower probability of reaching 5 A*-A than equivalent pupils in comprehensive LAs do. These results are consistent whether restricting the sample to those with high KS2 performance (as presented in Table 13.3), or using an alternative approach to include the whole year group.

13.3 Conclusion

In sum, the OLS and logistic regression models found no evidence of superior academic effectiveness in the selective system. All possible benefits associated with grammar school attendance are wiped out by the lesser progress of pupils in selective areas who did not attend grammar school. While the selective system does not influence the overall chances of getting satisfactory basic grades at KS4, there is evidence that the selective system may be less effective in getting top grades. For pupils in selective areas, regardless of whether they attended grammar school, their chances of getting top GCSE grades are lower than equivalent pupils in comprehensive LAs. This may be due to the high attainment in grammar schools which is negatively correlated with the self-perception of high-achieving pupils, according to the big-fish-little-pond effect (Marsh, 1984; Marsh & Hau, 2003). Meanwhile, for pupils who failed to get into grammar schools in selective LAs, a lack of role models and the sense of failure may also impede their progress. Therefore, while the general result of the effectiveness of the two systems does not reveal any superiority, the internal pattern implies the potential negative results of the selective system from which high-performers suffer. Overall, the pattern in this study finds no superiority of the selective system in raising academic standards.

14 Findings about the relationship between grammar school attendance and participation in HE

While previous chapters have revealed evidence of the relationship between grammar school attendance and academic outcomes at KS4, this chapter focuses on whether attending grammar school is predictive of pupil's post-18 destination. It starts with descriptive results of the patterns of HE participation and subject choice for grammar school pupils and those in non-selective schools. This is followed by statistical results of the logistic regression models predicting the relationship between the probability of attending HE institutions/ the Russell Group universities, and pupils' baseline variables at the compulsory education stage.

14.1 Descriptive result

14.1.1 HE participation patterns

Since the NPD data extract applied in this chapter is different from that of previous chapters, this section first presents the descriptive results of the 2007/2008 KS2 cohort in selective LAs. As mentioned in the methods chapter, only pupils with valid records from KS2 to KS5 are kept in the analysis, which includes 117,506 cases. Among these pupils, 51,016 (43.4%) can be matched to 2015/2016 HESA data, indicating that they attended HE institutions after secondary education. Among pupils who have HE records for 2015, about 30% (15,722) of them are at Russell Group universities. These pupils only account for 13.4% of the total cases included for analysis, revealing the selective nature of the Russell Group universities. For grammar school pupils, their rate of HE participation is higher than the average in selective LAs, as 66.7% of them attended universities in 2015, and 35.4% were admitted into Russell Group universities (Table 14.1).

Table 14.1: HE participation patterns in selective LAs

	The percentage of HE participation	**The percentage of attending the Russell Group universities**
Grammar school pupils	66.7	35.4
Average of selective LAs	43.4	13.4

14.1.2 Patterns of subject choice in HE

In addition to the general patterns of HE participation, detailed subject choice is also examined. In order to present a clearer picture, subjects are grouped into several major categories based on the definition created by Walker and Zhu in 2011, which was also applied in a later study

by Sullivan et al. (2016). The category includes STEM (Science, Technology, Engineering and Mathematics), LEM (Law, Economics and Management) and OSSAH (other social sciences, arts and humanities).

The proportion of each HE subject group for grammar school pupils is similar to the average rate of selective LAs (Table 14.2). In selective LAs, most pupils have chosen STEM, which is followed by OSSAH and then LEM. For grammar school pupils, the proportion taking STEM is higher than the average rate in selective LAs. Accordingly, the proportions of grammar school pupils studying OSSAH and LEM are lower than the average rate. Despite these small differences, grammar school pupils' subject preferences are close to those of other pupils in selective LAs.

Table 14.2: The two most frequently chosen HE subject groups in selective LAs

	Most frequently chosen subject	**Boys' most frequently chosen subject**	**Girls' most frequently chosen subject**
Grammar school pupils	STEM (52.4%)	STEM (57%)	STEM (48%)
Average of selective LAs	STEM (46.8%)	STEM (54%)	OSSAH (45%)

14.1.3 Pupil characteristics and patterns of HE participation

In this section, the analysis presents the characteristics of pupils with different HE participation patterns (Table 14.3). The favorable background and attainment of pupils in more advantaged post-18 destinations are clear between non-participants, HE participants, and participants of Russell Group universities. Additionally, for pupils with the same pattern of post-18 participation, the status of those who attended grammar schools is superior to their counterparts in non-selective schools. The patterns are consistent both in terms of family background and attainment at both key stages.

Based on the high raw attainment of grammar school pupils, it is reasonable that they would have a higher opportunity of participating in HE, and have higher rates of attending Russell Group universities. The following models explore whether the favourable HE participation results for grammar school pupils can be explained by the type of secondary school attended. Logistic regression models are applied to present the relationship between the HE participation pattern, academic performance, family background and school type.

Table 14.3: Characteristics of pupils with different HE participation patterns in selective LAs

		IDACI	FSM	SEN-PS	EAL	KS2 total mark	Capped GCSE point score (pre-2014 scale)*
Grammar schools	Non-participants	0.14	3%	1.3%	9%	159	402
	HE participants	0.12	2%	0.7%	12%	165	430
	The Russell Group participants	0.11	2%	0.6%	12%	170	446
Non-selective schools	Non-participants	0.20	12%	5.7%	11%	126	358
	HE participants	0.18	9%	2.7%	17%	140	393
	The Russell Group participants	0.15	6%	1.6%	16%	156	427

(*Note: grade G equals 16 point scores and the interval between each grade is 6 point scores)

14.2 Results of HE participation

In this section, the multi-stage logistic regression models predicting pupils' probability of attending HE institutions are presented first.

Table 14.4: Predictive accuracy of logistic regression models of HE participation

Variables	Percentage correctness	Percentage of remaining variation explained
Base figure	56.5	-
Personal background and attainment at KS2	65.8	9.3
School characteristics	66.2	0.4
School type (grammar school or not)	66.2	0
Personal background and attainment at KS4	69.4	3.2
Overall	69.4	12.9

According to the classification table (Table 14.4), the most pronounced growth in predictive accuracy is when KS2 pupil-level background and attainment variables are entered into the model, which is 9.3%. In contrast, including school compositional variables only slightly changes the overall accuracy, which is 0.4%. Meanwhile, knowing whether a pupil attended grammar school is also not informative about their HE participation rate, with no growth in

predictive accuracy at this stage. After controlling for all the above-mentioned baseline variables, there is still substantial growth resulting from including KS4 pupil variables, with an additional 3.2% variation in outcome explained.

14.2.1 HE participation and KS2 pupil characteristics

The first logistic regression model only includes KS2 pupil background and attainment variables, which is not intended to present how attending grammar schools is associated with the opportunity of HE participation. Rather, it is done to demonstrate the unbalanced HE participation rates for different pupil groups (Model 1 in Table 14.5).

Unsurprisingly, the results reveal that after accounting for attainment at KS2, pupils from more advantaged families have a higher probability of attending HE institutions. Those from the poorest areas are only half as likely as the wealthiest to go to universities, with the odds ratio for IDACI score being 0.43. Similarly, pupils eligible for FSM are about 80% as likely as others to attend HE institutions. However, pupils with SEN do not necessarily have lower HE participate rates, after accounting for their KS2 attainment. The rates of HE participation for pupils with all three SEN provisions are similar to non-SEN pupils when other factors are equal at KS2. Similar to the advantaged status of the EAL group and most ethnic minority groups in KS4 attainment, these pupils also enjoy HE participation rates higher than the equivalent white pupils. Meanwhile, girls are more likely to go to universities than boys, with an odds ratio of 1.3. These results also show that younger pupils have an advantage over older pupils within the same year group, with the Exp (B) of month age being 0.98. Therefore, pupils born in August only have about 80% of the rate of participating in HE as those born in September, after controlling for KS2 attainment and backgrounds. However, as mentioned before, this may be because KS2 attainment is not age-standardised. For pupils with similar family backgrounds, those with higher KS2 attainment enjoy better chances of HE participation. The odds ratio for KS2 attainment shows that performance at the end of KS2 is already correlated with pupils' HE participation rates. Although the Exp (B) of the total mark is small, the wide range of marks makes the difference between pupils with high and low KS2 performance salient.

Table 14.5: Logistic regression models of HE participation

	Model 1		Model 2		Model 3		Model 4	
Variable	**B**	**Exp (B)**	**B**	**Exp (B)**	**B**	**Exp (B)**	**B**	**Exp (B)**
Girls vs. Boys	0.315	1.371	0.317	1.373	0.317	1.373	0.105	1.111
Month Age	-0.017	0.983	-0.015	0.985	-0.015	0.985	-0.014	0.986
IDACI	-0.855	0.425	-0.736	0.479	-0.737	0.478	-0.464	0.629
KS2 FSM Eligible	-0.225	0.798	-0.241	0.786	-0.24	0.786	-0.148	0.863
SEN School Action	0	1	-0.023	0.978	-0.023	0.977	-0.043	0.958
SEN School Action Plus	0.037	1.037	0.008	1.008	0.008	1.008	-0.038	0.963
SEN Statement	0.1	1.106	0.066	1.069	0.067	1.069	0.065	1.067
EAL	0.363	1.437	0.357	1.43	0.358	1.43	0.229	1.257
Asian	0.686	1.985	0.6	1.822	0.599	1.821	0.524	1.689
Black	0.923	2.517	0.854	2.35	0.855	2.35	0.815	2.259
Chinese	1.209	3.35	1.056	2.874	1.055	2.871	0.825	2.282
Mixed	-0.002	0.998	-0.027	0.973	-0.028	0.973	-0.007	0.993
Unclassified	0.303	1.353	0.267	1.306	0.267	1.307	0.225	1.252
Other Ethnic Groups	0.544	1.723	0.449	1.567	0.45	1.568	0.276	1.318
KS2 Total Mark of English and Maths	0.025	1.025	0.02	1.02	0.02	1.02	0.006	1.006
Mean KS2 Total Mark in Secondary School	-	-	0.019	1.019	0.019	1.019	0.018	1.018
Mean FSM Proportion in Secondary School	-	-	-0.009	0.991	-0.009	0.991	-0.007	0.993
Grammar School	-	-	-	-	0.018	1.018	-0.199	0.819
KS4 FSM Eligible	-	-	-	-	-	-	-0.047	0.954
KS4 Capped GCSE Point Score	-	-	-	-	-	-	0.016	1.016
No. of Observation	117,506		117,506		117,506		117,506	
Percentage Correctness	Increase 9.3%		Increase 0.4%		Increase 0%		Increase 3.2%	
	56.5%-65.8%		65.8%-66.2%		66.2%-66.2%		66.2%-69.4%	

14.2.2 HE participation, KS2 pupil characteristics and school composition

The second model includes school compositional variables in addition to KS2 pupil-level variables, which are school-level KS2 attainment and FSM proportion (Model 2 in Table 14.5).

The small increase in the model's accuracy reveals the low predictive ability of these two school-level variables. Taking school characteristics into account, the difference in the probability of HE participation between pupil groups decreased slightly. The odds ratios of these compositional variables are consistent with the general perception that a more advantaged school context is associated with better outcomes at later stages. While higher school-level KS2 performance positively predicts HE participation rates, a large proportion of FSM pupils in schools is expected to decrease them.

14.2.3 HE participation, KS2 pupil characteristics, school composition and school type

The third stage finally includes school type into the model to test whether grammar school attendance is associated with a higher HE participation rate for equivalent pupils at KS2 (Model 3 in Table 14.5). Consistent with the weak predictive ability of school type as presented in Table 14.4, the odds ratio for school type is also close to 1. The participation rate among grammar school pupils is only 1.8% higher than that of others with equivalent backgrounds and attainment at KS2. This means the probability of HE participation varies little for equivalent pupils in different types of schools, even though there is a substantial gap in the raw rates of participation between grammar schools and non-selective schools. Therefore, the evaluation reveals that for pupils with equivalent KS2 backgrounds and attainment, attending grammar schools is not expected to bring any major advantage in HE participation.

14.2.4 HE participation, KS2/KS4 pupil characteristics, school composition and school type

In Model 4, KS4 pupil-level variables are also controlled for (Table 14.5). The results present whether attending grammar schools is associated with a difference in HE participation for pupils with equivalent attainment and characteristics at KS2 and KS4.

Introducing GCSE capped point score and KS4 FSM eligibility into the model, the results surprisingly present that pupils in grammar schools for secondary education are now slightly disadvantaged in their opportunities of HE participation, compared to equivalent peers in non-selective state-funded schools, with the odds ratio for grammar schools dropping to 0.83. This may partly be the result of the Widening Participation initiative, which attempts to raise HE participation rates for underrepresented groups, who usually come from more disadvantaged backgrounds (DfE, 2017). Since the targeted pupils are more likely to be educated in non-

selective schools, the Widening Participation initiative might create a negative bias against grammar schools.

It is within expectations that pupils with better GCSE results are more likely to attend universities, with its odds ratio being 1.02. While the figure seems small, when the cohort's complete GCSE score range is considered, pupils ranked in the top decile will have 12 times the probability of the bottom 10% of attending HE institutions. These results confirm that GCSE result is a strong predictor of a pupil's HE participation. Meanwhile, as pupils' performances at different stages are usually correlated (the correlation figure between the total KS2 mark and capped GCSE point score is 0.6), total KS2 mark becomes less influential in predicting HE participation once GCSE point score is controlled for. However, considering the wide range of KS2 marks, there is still a difference between high and low performing pupils, and the early-age attainment in primary schools retains some of its influence on the opportunity of HE participation, which is not fully captured by pupils' later performance at KS4.

In addition to the differentials correlated with school type and attainment, gaps between pupil groups still exist, even though they are smaller than in the previous three models in Table 14.5. While Chinese pupils are still the most likely to participate in HE, the odds ratio in comparison to white pupils decreases from 3.4 (Model 1) to 2.3 (Model 4). Pupils of mixed ethnicity remained the group least likely to attend HE institutions, but the results for this group are similar to that of the majority white group, with the odds ratio now being 0.99. Consistent with the higher participation rate among most ethnic minorities, non-native English speakers are also more likely to attend HE institutions. Although the odds ratio for the EAL group dropped in Model 4, this group's rate is still 26% higher than that of native pupils. The gap between pupils born in different months narrows, and the gap between the two genders diminishes, with about 2/3 of the initial difference in Model 1 being cancelled out. Meanwhile, the disadvantage for pupils from poorer areas is also mediated, with the IDACI odds ratio increasing to 0.63. While pupils eligible for FSM at the end of KS2 are only 86% as likely to go to universities as others, the KS4 FSM eligibility odds ratio is weaker, which is 0.95. This demonstrates that eligibility for FSM at earlier school stages is more harmful to pupils' opportunity of HE participation. Combining these two ratios, pupils eligible for FSM at both KS2 and KS4 are about 0.82 times as likely as others to participate in HE. This reveals the more disadvantaged status of pupils eligible for FSM for more years. Overall, the differentiated HE participation rates between pupil groups remain stubborn even after accounting for all of the baseline

variables from KS2 to KS4. While some pupil groups have initial disadvantages resulting from inadequate family support or poor health conditions and thus require extra effort to catch up with others, their opportunities of HE participation are still lower, even if they achieve an academic level equivalent to others at KS4. This reveals the long-term relationship between pupil backgrounds and academic trajectories.

14.2.5 Conclusion

The results from the four models in Table 14.5 demonstrate that for pupils with equivalent KS2 backgrounds and attainment, attending grammar schools almost brings no difference to HE participation rates. For equivalent pupils at KS4, those in grammar schools might even be slightly disadvantaged in terms of HE participation. Therefore, there is no evidence that grammar schools are better at promoting HE participation, either through improving academic performance, or through offering extra long-term bonus, such as boosting pupils' academic aspirations or providing adequate help with HE applications.

In addition to the connection between the opportunity of HE participation and school type, the gap in HE participation between pupil groups persists, even if they have the same attainment in primary schools and secondary schools. While the gaps are smaller between the two genders, between pupils from high and low SES groups, between SEN pupils and others, and between autumn- and summer-born pupils, it is the widest between different ethnic groups. White pupils are less likely to attend universities than most of their equivalent peers from ethnic minority groups. This pattern is consistent in all four models controlling for different sets of baseline variables.

14.3 Results of the opportunity to attend the Russell group universities

In addition to calculating pupils' chances of attending HE institutions in general, their probabilities of attending the Russell Group universities are also calculated. All the variables included in the four-stage models are the same as those in the previous models of HE participation. As mentioned in the methods chapter, unlike in the logistic regression models predicting 5 A*-A grades which only include high-performers, all valid cases are included for analysis in this chapter to make the Exp (B) between the models of HE participation and Russell Group participation more comparable. Due to the small proportion of pupils in the Russell

Group universities, the predictive accuracy of the base model is already high, leaving little space for growth when different sets of explanatory variables are added.

After controlling for KS2 pupil-level variables, the prediction accuracy of the model increases from 87% to 87.6%. Including school compositional variables adds another 0.2%, and there is still no growth when school type is included. The increase is slightly more pronounced when KS4 variables are added, with the model's final prediction accuracy reaching 89% (Table 14.6).

Table 14.6: Predictive accuracy of logistic regression models of the Russell Group participation

Variables	**Percentage correctness**	**Percentage of remaining variation explained**
Base figure	87	-
Personal background and attainment at KS2	87.6	0.6
School characteristics	87.8	0.2
School type (grammar school or not)	87.8	0
Personal background and attainment at KS4	89	1.2
Overall	89	2

14.3.1 The opportunity to attend the Russell Group universities and KS2 pupil characteristics

The first model only includes KS2 pupil-level variables to present the differences between pupil groups. According to Model 5 in Table 14.7, KS2 performance has become more influential in predicting the outcome than in the HE participation model (Model 1 in Table 14.5), with the odds ratio for KS2 total mark reaching 1.061. Meanwhile, the patterns of pupil characteristics remain mostly the same as in the models of HE participation. Girls are still more likely to attend the Russell Group, but the advantage has diminished, with them now possessing 12% higher chances than boys. Similarly, although the rate of attending elite universities remains the lowest for white pupils, the advantages of some ethnic minority groups (Asia, black and Chinese) also decreased. In contrast, unlike the narrower gap between ethnic groups, the division between pupils from wealthy and impoverished areas becomes wider than the gap in HE participation. These results demonstrate that pupils from areas with the worst IDACI scores only have 16% the probability of those in the most affluent areas of attending the Russell Group universities. Therefore, although the negative effect of living in impoverished areas is already salient in terms of HE participation rates (Model 1 in Table 14.5), it is even more pronounced in pupils' success rates when applying for elite universities.

Table 14.7: Logistic regression models of the Russell Group participation

	Model 5		Model 6		Model 7		Model 8	
Variable	**B**	**Exp (B)**	**B**	**Exp (B)**	**B**	**Exp (B)**	**B**	**Exp (B)**
Girls vs. Boys	0.154	1.166	0.163	1.177	0.163	1.177	-0.175	0.839
Month Age	-0.031	0.969	-0.028	0.972	-0.028	0.972	-0.023	0.977
IDACI	-1.845	0.158	-1.556	0.211	-1.51	0.221	-0.965	0.381
KS2 FSM Eligible	-0.274	0.76	-0.253	0.777	-0.258	0.772	-0.121	0.886
SEN School Action	-0.048	0.954	-0.069	0.933	-0.053	0.948	-0.053	0.948
SEN School Action Plus	0.033	1.034	0.029	1.029	0.032	1.033	0.004	1.004
SEN Statement	0.385	1.47	0.37	1.448	0.363	1.437	0.397	1.488
EAL	0.396	1.486	0.414	1.513	0.407	1.502	0.135	1.144
Asian	0.541	1.718	0.365	1.44	0.384	1.467	0.172	1.188
Black	0.669	1.952	0.559	1.749	0.55	1.733	0.417	1.517
Chinese	1.118	3.059	0.842	2.321	0.843	2.324	0.436	1.547
Mixed	0.129	1.138	0.078	1.081	0.088	1.092	0.133	1.143
Unclassified	0.354	1.425	0.288	1.334	0.282	1.326	0.21	1.233
Other Ethnic Groups	0.582	1.789	0.415	1.515	0.392	1.48	0.043	1.044
KS2 Total Mark of English and Maths	0.06	1.061	0.05	1.051	0.05	1.051	0.017	1.017
Mean KS2 Total Mark in Secondary School	-	-	0.026	1.027	0.04	1.041	0.033	1.034
Mean FSM Proportion in Secondary School	-	-	-0.012	0.988	-0.014	0.986	-0.01	0.99
Grammar School	-	-	-	-	-0.488	0.614	-0.631	0.532
KS4 FSM Eligible	-	-	-	-	-	-	-0.027	0.973
KS4 Capped GCSE Point Score	-	-	-	-	-	-	0.034	1.035
No. of Observation	117,506		117,506		117,506		117,506	

Similarly, the negative correlation between poorer family background and the pattern of post-18 destination is also solidified in terms of FSM eligibility. Pupils eligible for FSM at the end of KS2 are 20% less likely to attend elite universities than equivalent non-FSM pupils. Therefore, the neighbourhood index and the individual index demonstrate the same pattern,

revealing the deeper negative influence of poverty on the probability of attending elite universities than on the general HE participation rate. Another indicator of individual disadvantage, SEN, presents a mixed pattern within the category. While pupils with SEN action and SEN action plus usually have chances similar to those of others, the SEN statement group has a higher rate of attending the Russell Group universities when other factors are equal at KS2. This may partly be due to the evidence in previous studies questioning the validity of SEN category, as it is interrelated with other cofounding factors (Gorard et al., 2019). The pattern might also be the result of the tiny proportion of this group of SEN pupils at the Russell Group universities, thus implying their systematically varied characteristics.

14.3.2 The opportunity to attend the Russell Group universities, KS2 pupil characteristics and school composition

After accounting for pupils' prior differences at KS2, school compositional variables are also controlled for (Model 6 in Table 14.7). The results of these school-level variables in predicting the probability of attending the Russell Group universities are consistent with the HE participation models, revealing a positive correlation between participation rate and high-performing student intakes and a negative correlation between participation rate and the proportion of FSM pupils. Meanwhile, the patterns for both variables of school characteristics are stronger in the Russell Group model, with the odds ratios more pronounced than in the previous HE participation models, which implies that school composition may have stronger influence on the rate of attending elite universities.

14.3.3 The opportunity to attend the Russell Group universities, KS2 pupil characteristics, school composition and school type

After accounting for pupils' prior differences at KS2 and school compositional variables, whether they went to grammar school is added into the model (Model 7 in Table 14.7). Although adding school type does not increase the overall prediction accuracy, the odds ratio for school type shows that after accounting for KS2 pupil-level baseline variables and compositional variables of secondary schools, grammar school attendance is associated with a lower probability of attending elite universities (Exp (B)=0.61). While grammar school pupils have a higher raw rate of attending the Russell Group, the pattern is reversed after accounting for KS2 pupil-level characteristics and school composition.

There are concerns over the validity of logistic regression models when the base figure of the outcome variable is high, such as this model in which the prediction accuracy starts at 87%. However, when the analysis is only conducted on pupils with high KS2 levels (KS2 total marks higher than 134, as was done in the logistic model for 5 A*-A grades), the results present the same pattern. Furthermore, if the same number of pupils not in the Russell Group are randomly selected to make the base figure exactly 50%, the odds ratio for grammar schools is still similar. Therefore, the surprising result of school type in this model should not be due to modelling errors. The results for attending elite universities correspond to the 5 A*-A model for GCSE, which not only failed to find evidence of grammar schools' positive effects on producing top outcomes, but also presented the possibility that grammar schools are less successful than non-selective schools after accounting for pre-existing differences. However, similar to the issue of underestimation in the 5 A*-A model, the dramatic negative estimation of grammar schools might partly be due to the correlation between positive school practices and more advantaged school composition. This may attribute grammar schools' raw performance too much to their advantaged composition and dilute their real effect.

14.3.4 The opportunity to attend the Russell Group universities, KS2/KS4 pupil characteristics, school composition and school type

In Model 8 (Table 14.7), the KS4 pupil-level variables are also included, in addition to above-mentioned baseline variables. The changes in school type resulting from including KS4 pupil-level variables are similar to those of previous HE participation models, with grammar schools performing even lower than non-selective schools. According to the results of Model 8, the odds ratio for grammar schools dropped to 0.53 when KS4 variables were controlled for. This means that for those with the same attainment level and family background in primary schools and secondary schools, grammar school attendance is associated with lower probabilities of going to the Russell Group universities. However, as previously stated, this might be an underestimation of the real effect of grammar schools. The possible reasons for this low estimation for the grammar school effect are explored in more detail in the next section, after presenting the results of Model 8.

After accounting for KS2 and KS4 baseline variables, the unbalanced probabilities of attending elite universities between different pupil groups remain. Ethnic minorities are still more likely to go to elite universities, but the extent is milder than in the HE participation model. The

results for SEN pupils are consistent with Models 5-7, revealing that they are not disadvantaged once attainment and other background variables are considered. Meanwhile, the difference associated with IDACI scores is mitigated when KS4 pupil-level variables are added, with the odds ratio increasing to 0.38. However, it is still much lower than the result for HE participation, which shows the stronger negative influence of IDACI score on the rate of attending elite universities. This uncovers a disparity in the type of institution attended by pupils from wealthier and poorer areas, as the former are more likely to attend high-ranking universities than the latter, even after accounting for academic performance and other background characteristics from KS2 to KS4. For FSM pupils, the odds ratio for FSM eligibility at KS2 shrinks to 0.86, and the odds ratio for the KS4 FSM indicator is now 0.95. Therefore, pupils eligible for FSM at both KS2 and KS4 have 82% the probability of non-FSM pupils of attending elite universities. These results also confirm that eligibility for FSM at KS2 is slightly more harmful than KS4 FSM eligibility, which is the same as in the HE participation model. Among the background variables, only gender changes dramatically in Model 8. While in previous models (Models 1-4 in Table 14.5 and Models 5-7 in Table 14.7), girls are more likely to reach the outcome, regardless of the explanatory variables, they have only 84% the probability of boys of attending the Russell Group universities, after accounting for background and attainment at KS2 and KS4.

Meanwhile, it is within expectations that pupils with higher GCSE attainment are more likely to attend the Russell Group universities (Model 8 in Table 14.7). The division between pupils in the top and bottom GCSE quintiles is stronger than that of HE participation, which implies the higher academic requirements of the Russell Group. Meanwhile, introducing GCSE results into the model also weakens the link between KS2 attainment and the rate of attending elite universities, but KS2 attainment is still expected to influence the opportunity of attending elite universities.

14.3.5 Conclusion

The results of the four models in Table 14.7 demonstrate that for pupils with equivalent backgrounds and attainment at KS2 or KS4, attending grammar schools does not increase the opportunity of attending the Russell Group universities. On the contrary, grammar school pupils might even be disadvantaged, after accounting for pre-existing pupil-level and school-level differences.

In addition to the unbalanced access to the Russell Group universities for pupils in different types of schools, the gap between other pupil groups also remains salient even if they have the same attainment in primary and secondary school. Certain ethnic groups, as well as those from advantaged social backgrounds, are more likely to attend elite universities. The pattern is consistent in all the four models controlling for different sets of baseline variables.

14.4 Alternative models of HE participation patterns with no school compositional variables (for comparison only)

While the reasons for adding school compositional variables have been explained in the methods chapter, this section presents the results for HE and Russell Group participation without controlling for school-level variables. This helps explain the low estimation results for grammar schools that was presented in previous sections. The discussion of the differences between the models with and without school compositional variables is also relevant to the effectiveness of grammar schools in the GCSE results discussed in Chapter 11.

14.4.1 The opportunity of HE participation

When school compositional variables are removed from the model, the predictive accuracy changes little (classification table in Appendix 8). However, the odds ratios for school type surpass those of Table 14.5. After controlling for KS2 pupil variables, grammar school pupils are 60% more likely than others to attend HE institutions (Model 2 in Table 14.8). If KS4 variables are added into the model, the odds ratios for grammar school attendance drop from 1.6 to 1.3, with grammar school pupils still being 30% more likely than their equivalent peers in non-selective schools to attend HE institutions (Model 3 in Table 14.8). This demonstrates that although the advantage of grammar schools has been partially explained by their better GCSE results, grammar school pupils still have a higher opportunity of HE participation than equivalent pupils in non-selective schools, when school compositional variables are excluded.

Table 14.8 Logistic regression models of HE participation (without school compositional variables)

	Model 1		Model 2		Model 3	
Variable	**B**	**Exp (B)**	**B**	**Exp (B)**	**B**	**Exp (B)**
Girls vs. Boys	0.26	1.295	0.26	1.296	0.06	1.061
Month Age	-0.02	0.982	-0.02	0.983	-0.02	0.985
IDACI	-0.85	0.429	-0.74	0.475	-0.52	0.597
KS2 FSM Eligible	-0.22	0.802	-0.20	0.816	-0.12	0.885
SEN School Action	0.01	1.013	-0.01	0.989	-0.03	0.969
SEN School Action Plus	0.05	1.047	0.02	1.020	-0.02	0.981
SEN Statement	0.12	1.130	0.11	1.114	0.10	1.109
EAL	0.39	1.479	0.41	1.503	0.26	1.300
Asian	0.70	2.009	0.64	1.898	0.55	1.734
Black	0.92	2.500	0.89	2.434	0.84	2.322
Chinese	1.27	3.564	1.16	3.184	0.90	2.470
Mixed	0.30	1.342	0.28	1.322	0.24	1.268
Unclassified	0.00	1.000	-0.02	0.977	-0.01	0.994
Other Ethnic Groups	0.57	1.759	0.52	1.689	0.34	1.398
KS2 Total Mark	0.025	1.025	0.022	1.022	0.008	1.008
Grammar Schools	-	-	0.478	1.612	0.248	1.281
KS4 FSM Eligible	-	-	-	-	-0.05	0.949
KS4 Capped GCSE Point Score	-	-	-	-	0.02	1.016
No. of Observation	107,506		107,506		107,506	

14.4.2 The opportunity to attend the Russell Group universities

The tendency is the same in the models of the Russell Group universities. Without controlling for school-level variables, the predictive accuracy of the model changes little (Appendix 9). However, according to the odds ratios in Table 14.9, grammar school pupils possess a higher rate of attending the Russell Group universities when school compositional variables are not controlled for, both before and after KS4 pupil characteristics are added.

Table 14.9: Logistic regression models of the Russell Group participation (without school compositional variables)

	Model 4		Model 5		Model 6	
Variable	**B**	**Exp (B)**	**B**	**Exp (B)**	**B**	**Exp (B)**
Girls Vs. Boys	0.13	1.143	0.13	1.140	-0.21	0.810
Month Age	-0.03	0.968	-0.03	0.970	-0.03	0.975
IDACI	-1.84	0.159	-1.66	0.191	-1.12	0.327
KS2 FSM Eligible	-0.27	0.761	-0.23	0.795	-0.10	0.907
Sen School Action	-0.05	0.952	-0.07	0.929	-0.07	0.935
Sen School Action Plus	0.00	1.001	0.00	0.995	-0.06	0.994
Sen Statement	0.36	1.436	0.37	1.442	0.40	1.488
EAL	0.41	1.502	0.44	1.548	0.16	1.178
Asian	0.58	1.788	0.49	1.615	0.23	1.253
Black	0.69	1.986	0.64	1.903	0.48	1.612
Chinese	1.187	3.277	1.03	2.808	0.58	1.777
Mixed	0.35	1.419	0.32	1.380	0.247	1.280
Unclassified	0.15	1.157	0.11	1.111	0.142	1.153
Other Ethnic Groups	0.62	1.864	0.56	1.754	0.17	1.181
KS2 Total Mark	0.060	1.061	0.055	1.056	0.021	1.021
Grammar School	-	-	0.495	1.641	0.199	1.221
KS4 FSM Eligible	-	-	-	-	-0.05	0.951
KS4 Capped GCSE Point Score	-	-	-	-	0.03	1.035
No. of Observation	117,506		117,506		117,506	

In order to present a clearer comparison between models with and without school compositional variables, the odds ratios for school type in these two types of models are combined in Table 14.10. The estimation results for grammar schools are higher than non-selective schools in models which exclude school compositional variables. Adding school compositional variables diminishes the estimated grammar school effect in all of the presented models, regardless of the outcome indicators applied.

The decrease in the grammar school effect resulting from adding school compositional variables could be interpreted as the controversial ‘peer effect’, which suggests that more advantaged student composition drives school effectiveness (Jennings et al., 2015).

Furthermore, as discussed in the literature chapters, adding compositional variables may eliminate the influence of the cofounding sorting effect of more advantaged pupils in grammar schools, who are still systematically different from others even after accounting for surface pupil-level variables (Castellano et al., 2014). The mitigated grammar school effect also likely has a close relationship with the measurement errors pointed out by Perry (2019). He mentioned that models only controlling for pupil-level variables provide unfair estimations to less advantaged schools, and thus presenting a 'phantom' grammar school effect.

Table 14.10: Odds ratios for grammar schools in logistic regression models of HE participation patterns with and without school compositional variables

	Odds ratio for grammar schools
HE participation - control for KS2 pupil variables	
With school composition	1.02
Without school composition	1.63
HE participation- control for KS2 and KS4 pupil variables	
With school composition	0.82
Without school composition	1.30
The Russell Group participation - control for KS2 pupil variables	
With school composition	0.62
Without school composition	1.67
The Russell Group participation - control for KS2 and KS4 pupil variables	
With school composition	0.53
Without school composition	1.28

These supporting evidence shows that models with school-level variables are more reliable, and that the lower estimation results for grammar school should be more accurate. However, there are also concerns resulted from including school compositional variables, especially if an advantaged school context is associated with beneficial school practices (Raudenbush & Willms, 1995). For example, schools with better SES status may have more aspirational norms, higher academic expectations, and lower anxiety levels (Evans, Brooks-Gunn & Klebanov, 2011; Nelson & DeBacker, 2008). Schools with more advantaged intakes and prestigious reputations may also attract more high-qualified teachers and managers than other schools, especially since there is a shortage of secondary school teachers in England (See & Gorard, 2019). As these positive factors correlated with school composition all contribute to academic

success, accounting for school compositional variables may cover the impacts of these positive actions and underestimate the real effect of grammar schools.

As there may be both benefits and drawbacks of adding school compositional variables, it is difficult to distinguish them from each other. Although it is widely accepted that merely accounting for pupil-level variables may be inadequate to evaluate school effectiveness, the estimated results of models controlling for both pupil-level and school-level variables should not be simply regarded as the real grammar school effect. It is generally accepted that the estimated grammar school effect, after accounting for pupil-level and school-level characteristics, may provide a lower bound for the possible range of the real effect (Coe et al., 2008).

14.5 Conclusion

After accounting for KS2 pupil-level pre-existing differences and school compositions, grammar schools almost present no difference from equivalent non-selective schools in terms of HE participation rates. Meanwhile, grammar school pupils even have a lower opportunity of attending the Russell Group universities than similar pupils from equivalent non-selective schools do. While the low estimation results for grammar schools may be due to statistical factors such as the negative bias against grammar schools when their advantageous school-level variables are controlled for, and policy reasons such as initiatives of Widening Participation, the results in this chapter still failed to support that grammar schools are more successful in terms of HE participation patterns. After presenting the relationship between grammar school attendance and HE participation patterns in this chapter, the next chapter discusses the link between pupil's family background and post-18 destination in selective and non-selective LAs.

15 Findings about the link between pupil's family background and post-18 destination

Previous chapters have evaluated the unbalanced access to grammar schools for different pupil groups, and the dissimilar outcomes of attending grammar schools and non-selective schools. This chapter conducts a more direct evaluation of the link between family background and pupil's post-18 destination in different types of LAs.

The analysis first presents the correlation between pupils' attainment at different key stages (Table 15.1). For pupils in non-selective LAs, the correlation figure between KS2 and KS4 attainment is 0.56. It is weaker for KS2 and KS5 performance, which is 0.48. In contrast, the patterns in selective LAs reveal a closer connection between attainment at different stages. The correlation figure is 0.59 between KS2 and KS4 performance, and 0.53 between KS2 and KS5 performance. Overall, the correlation between any two stages is stronger in selective LAs than in non-selective areas. This means that pupils who perform higher at earlier ages in selective LAs are also more likely to have better attainment at the end of secondary education than those in non-selective LAs. The stronger correlation between different stages means that pupils at an initial disadvantage and with low early-age performance are less likely to catch up at later stages. Although the difference between the two types of LAs is miniscule, this is likely due to the low proportion of grammar school places, which is only around 12%, even in selective LAs.

Table 15.1: Correlation of attainment at different key stages in selective and non-selective LAs

	KS2 attainment	KS4 attainment	KS5 attainment
Non-selective LAs			
KS2 attainment	1	0.56	0.48
KS4 attainment	-	1	0.59
KS5 attainment	-	-	1
Selective LAs			
KS2 attainment	1	0.59	0.53
KS4 attainment	-	1	0.62
KS5 attainment	-	-	1

In addition to revealing the correlation figures, the results of the logistic regression models predicting pupils' HE participation patterns are also presented (Table 15.2). While in non-selective LAs, knowing pupils' family background and attainment at KS2 increases the prediction accuracy of the model by 8.3, the figure for selective LAs is slightly higher, which is 8.9. When the total unexplained variance is taken into account, KS2 pupil-level variables in non-selective LAs constitute 17.6% of the unexplained part. In contrast, the rate in selective

LAs is 19.6%. Therefore, the models indicate that early-age background has a stronger role in predicting the rate of HE participation in selective LAs than in comprehensive LAs.

Table 15.2: Predictive accuracy of logistic regression models of HE participation in selective and non-selective LAs

Variables	**Percentage correctness**	**Percentage of remaining variation explained**
Selective LAs		
Base figure	54.7	-
KS2 pupil variables	63.6	8.9
KS4 pupil variables	66.1	2.5
KS5 pupil variables	72.8	6.7
Overall	72.8	18.1
Non-selective LAs		
Base figure	52.8	-
KS2 pupil variables	61.1	8.3
KS4 pupil variables	64.3	3.2
KS5 pupil variables	72.3	8
Overall	72.3	19.5

Although the difference in HE participation between the two types of LAs seems minor, the results of attending the Russell Group universities are more obvious (Table 15.3). Adding KS2 pupil-level variables increases the prediction accuracy by only 0.2 in non-selective LAs, but the rate for selective LAs is 1.2. Considering the total unexplained variation in both areas, KS2 pupil-level variables account for 1.4% of the unexplained part in non-selective LAs and 6.6% in selective LAs. Therefore, in selective LAs, the link between pupils' KS2 baseline variables and post-18 destination is stronger than in non-selective LAs. The patterns are consistent both in terms of the general HE participation rate, and the rate of elite university attendance.

To conclude, the evidence above has demonstrated that the connection between background and later destination is closer in selective LAs than in non-selective areas. Previous chapters found only a small academic benefit associated with grammar school attendance (for their borderline pupils) and no advantages of the selective system. However, separating pupils into different secondary schools according to their early-age academic ability may tighten the connection between earlier and later attainment, and the link between early-age background and later destination. Although the differentiated patterns between the two types of LAs may be influenced by the broader social contexts within each area, the evidence did not suggest that

the selective system is more equitable at redistributing educational resources than a comprehensive system.

Table 15.3: Predictive accuracy of logistic regression models of attending the Russell Group universities in selective and non-selective LAs

Variables	**Percentage correctness**	**Percentage of remaining variation explained**
Selective LAs		
Base figure	81.9	-
KS2 pupil variables	83.1	1.2
KS4 pupil variables	85.1	2
KS5 pupil variables	86.9	1.8
Overall	86.9	5
Non-selective LAs		
Base figure	85.5	-
KS2 pupil variables	85.7	0.2
KS4 pupil variables	87.1	1.4
KS5 pupil variables	89	1.9
Overall	89	3.5

So far, all the findings of this study have been presented. The following chapters start to discuss the limitations of this study, implications for future research, summary of findings, and implications for policy.

16 Limitations of the study

While attempts have been made to evaluate the effectiveness and equity of grammar schools as accurately as possible, this study still has some limitations which might negatively impact the quality of its evidence. This chapter discusses these limitations.

16.1 The application of traditional regression models

Much of the evidence in this study is produced through traditional regression models controlling for pre-existing differences between pupil groups. Using the rich NPD database in England, many baseline variables which are believed to be influential for later outcomes are included in the regression models. However, despite the high data quality, the conclusion cannot rule out threats such as data collection errors, missing data, coding inconsistencies, or unavailable background variables. More importantly, relevant to all the regression models controlling for baseline differences, any finite explanatory variables exclude some unavailable or unmeasurable characteristics, leaving space for omitted variable bias. This means that regression approaches cannot eliminate the possibility that the estimated differences associated with variables of interest (i.e., school type and LA type in this study) are driven by imperfect modelling processes.

Decisions on what baseline variables should be included in the regression models for school effectiveness are difficult. Without accounting for school-level variables, estimations may suffer from grouping bias between grammar school pupils and pupils in other schools, which is not sufficiently controlled for by pupil-level surface variables (Coe et al., 2008; Zimmer, 2003). The special historical status of grammar schools, compounded by the fact that academic selection in England only applies to a small group of pupils, means that the grouping bias may be even stronger. Besides, models not accounting for school-level prior attainment are threatened by measurement errors which upwardly bias the results of more advantaged schools. This creates a ‘phantom’ grammar school effect (Perry, 2019). While regression models accounting for both pupil-level and school-level variables (as applied in this study) mitigate previous issues, they may suffer from over-controlling. This would underestimate grammar school effectiveness, especially if beneficial actions in these schools are correlated with their advantaged pupil compositions. The best balance is still unclear, and until the issue is resolved, all effectiveness claims using such methods remain in doubt.

16.2 The application of the RDD approach

Although the RDD approach is strong to make causal inference, the estimated effectiveness of grammar schools in this study does not provide a definitive answer. The findings are limited by the imperfect 11+ data which lacks both information on pupils' later academic performance, and identifiers consistent with any major databases in the UK. Although every precaution was made to deal with the 11+ data in this study, many cases in the original 11+ file had to be omitted. This process may have negatively impacted the quality of the evidence, which cannot be compensated for by research design. Therefore, the value of the RDD part of this study lies more in testing the feasibility of applying a stronger approach to generate robust causal evidence of grammar schools' effectiveness, rather than in the actual estimation results.

The evidence in the RDD part only covers a single LA. Although the characteristics of grammar school pupils in this LA do not present obvious deviations from the pattern of selective LAs, the unbalanced chances of attending grammar schools, the varied selection difficulty in each area, and the broader social context may create dissimilarities in school effectiveness. Therefore, any indication of grammar school effectiveness revealed in the RDD is mostly relevant to this LA, and may diverge from the national pattern, even when the same design is applied. According to traditional school effectiveness models such as OLS, grammar school effectiveness in this LA is above the national average. Thus, it is reasonable to anticipate that the RDD estimate of the grammar school effect on the national level would also be even less pronounced than the effect in this LA. Based on the internal differences among grammar schools in England, additional RDD studies in other LAs are needed to present a complete picture of the real effect of grammar schools.

Lastly, as the treatment effect in the RDD appears through a comparison between grammar schools and non-selective mainstream state-funded schools, the estimated grammar school effect is not an absolute academic level; rather, it is the benefit in relation to non-selective schools in the same LA. While the results may indicate the stronger effectiveness of grammar schools in improving academic achievement in the chosen area, they may also be a signal of penalties for other schools in LAs with high proportions of grammar school places.

16.3 Analysis of the link between pupil's family background and post-18 destination

In the evaluation of the link between pupil's family background and future destination, the indicator of family background is the combination of different variables of pupil characteristics, and the outcome variables are HE participation patterns which include no information on degree completion. While considerable previous research has included parental education and occupation status as the indicator of social origin, and has applied the occupation/wage data of the cohort member as the indicator of future destination, these variables are unavailable in this study. This limits its conclusions. However, despite the difference in indicators applied, the stronger link between early-age background and later destination in a selective system is still consistent with previous research.

17 Implications for future research

Limited by the scope, several meaningful issues that emerged during the analysis are not further discussed in this study. However, future research focusing on these issues may help us understand this topic.

17.1 Unbalanced grammar school access

If the current system of selecting a few high-performing pupils into grammar schools remains, the low proportion of FSM pupils taking the grammar school selection test would require more detailed analysis of the underlying reasons. Accordingly, more attention should be paid to FSM pupils' real-life difficulties in approaching grammar school. The issue may be relevant to the underrepresentation of disadvantaged pupils in high-performing comprehensive schools as well.

The low participation rate in the selection test, even after accounting for prior attainment, may not only happen to FSM pupils, but also be relevant to other pupil characteristics, such as gender and ethnicity. A complete picture is needed to present the rate of taking the grammar school selection test for other pupil groups. Similarly, FSM pupils' lower performance than equivalent non-FSM pupils on the 11+ implies the possibility that other social groups may be underestimated in the selection test as well. Since the 11+ data in this study only covers one LA, other areas adopting different selection criteria and procedures should be evaluated. More importantly, crucial to the future paths of pupils in selective LAs, the validity and reliability of these high-stakes selection tests also needs a formal, independent and transparent evaluation.

Meanwhile, policies aiming at helping FSM applicants navigate the process of grammar school selection have already been adopted in some areas, such as the quota system in The King Edward VI Foundation Schools in Birmingham. The real effects of these actions, as well as the practical problems encountered during their implementation, need further analysis.

17.2 Understanding the differential grammar school effectiveness for FSM pupils

The evaluation of the assumed role of grammar schools in narrowing the attainment gap suggests that the similar KS4 performance of FSM and non-FSM pupils in grammar schools is largely due to their equivalent KS2 performance. While regression models reveal that attending grammar school may matter more for the former than for the latter, this study is unable to

confirm that grammar schools are especially effective for disadvantaged pupils. This is due to cofounding elements such as the poorer conditions of non-selective schools attended by FSM pupils, as they are more likely to be enrolled into low-quality schools (Tes, 2018), and the systematic difference within this small group of FSM grammar school pupils. Future studies are needed to evaluate the above-mentioned factors. It would also be meaningful to explore reasons why this small group of FSM pupils attained high academic levels, despite inadequate family support, and whether they continue performing well at later stages. This will help elucidate the link between family disadvantage and attainment.

17.3 The effectiveness of the selective system in developing countries

The selection system in England has several unique features. First, grammar schools have only been retained in a few local areas, and they recruit only a small minority of pupils within the state system. Secondly, this group of selective schools are not only academically famous, but also have a prestigious social status originating from historical tradition. Therefore, more studies focusing on other countries would help identify a global pattern. While there has been plenty of research in developed countries, especially in the US and European countries, more research is needed to reflect contextual differences in developing countries. For example, evidence from Kenya presents substantial dissimilarities to this study and to other OECD countries (Duflo, Dupas & Kremer, 20011; Hanushek & Wößmann, 2006). While this study finds no evidence of an early-age academic selection benefit in England, the selection system might be helpful under other circumstances, such as when there is too much variation in pupils' performance levels. Similar research would produce better results in countries where the data on both entry selection and attainment at later stages are publicly available.

17.4 Difficulties in controlling for baseline variables in school effectiveness models

As mentioned in previous chapters, the choice of baseline variables in models controlling for pre-existing differences between pupil groups is never perfect. Regression models with or without school composition variables have their limitations. The differences between these two types of models can be a combination of measurement error, grouping bias, peer effects, positive school practices, or other confounding factors. Therefore, it is difficult to separate the net effect of schools and distinguish beneficial school actions from error and bias. While more research is needed to present detailed patterns of the inner mechanism of school effectiveness, the difficulty in choosing baseline variables also reveals the importance of applying stronger

research designs. Although there are non-experimental designs which can be used as alternatives to an RCT, such as the RDD in this study, an RCT is still the strongest design for causal inferences. Therefore, if we truly care about the effectiveness of certain type of schools, an RCT might still be the most robust choice, when it is ethically and practically appropriate.

18 Summary of findings

After discussing the limitations of the study and implications for future research, this chapter summarises the findings of the previous chapters, answering this study's research questions.

18.1 The opportunity to attend grammar schools, and its relationship with pupil's prior attainment, geographical location and family background

18.1.1 The unbalanced opportunity to attend grammar schools between pupil groups

The varied proportion of available grammar school places in each selective LA leads to an imbalance in grammar school opportunities. Due to this, the threshold of grammar school selection across LAs varies. Some LAs allow pupils from the bottom national quartile for KS2 performance to attend grammar schools; others only enrol pupils with well above-average performance, which leads to dissimilar student compositions in grammar schools across LAs. This questions the appropriateness of analysing grammar schools in different LAs as a single entity, thus masking the internal differences within. The inequality of opportunity between LAs means that sending children to another LA would influence their chances of getting into grammar schools. However, according to the characteristics of pupils who move to another LA for secondary education and those who do not, moving across LAs has become a shortcut for more affluent families to manoeuvre within the selective system, and it is not usually an option for the disadvantaged. Although for pupils living near the border of an LA, schools in their own LAs are not necessarily geographically closer than those in a nearby LA, for most pupils from families with fewer resources, the extra time and effort necessary to access application information in other LAs may present obstacles.

In addition to the differences in grammar school opportunities due to pupils' geographical locations, the gaps associated with other pupil background characteristics are also considerable. Since early-age inequality in average achievement between different social groups is substantial, selection by attainment is indirectly associated with family background. Therefore, the characteristics of pupils in grammar schools deviate from the national pattern. They are more advantaged than the general population of their local communities, not only in terms of academic performance, but also in most other social characteristics. However, after accounting for early-age performance, grammar schools are still not equally accessible to different social groups. For pupils with equivalent KS2 attainment, ethnic minorities usually have more

opportunities than white pupils, but pupils from poorer areas, those eligible for FSM, and those who have SEN, tend to have less opportunities to attend grammar schools.

18.1.2 Possible reasons for the underrepresentation of disadvantaged pupils in grammar schools

Focusing on the gap between FSM and non-FSM pupils as an example, this study addresses potential explanations for the low probability of disadvantaged pupils attending grammar schools. The 11+ file provides evidence that FSM pupils are less likely to take the selection test, even if they have reached the same attainment level as non-FSM peers at KS2. There is a gap in the rate of taking the selection test between FSM and non-FSM pupils for the entire performance distribution (except for those at the bottom), which is most pronounced among pupils who have reached the expected performance level of grammar schools. This means that many FSM pupils could have had the chance to attend grammar schools, if only they had taken the selection test. This finding confirms previous research on the differentiated patterns of school choice between social groups, and the possibility of the self-exclusion of disadvantaged families from top-ranked schools. This conclusion also calls for detailed analysis of why many FSM pupils (or their parents) are reluctant to take the selection test.

Disadvantaged pupils are not only less likely to take the selection test, they also have lower passing rates when they do, either before or after accounting for prior attainment at KS2. The evidence from the chosen LA in this study reveals that FSM pupils' scores on the 11+ are lower than those of non-FSM pupils with equivalent KS2 performance. So the selection test systematically judges FSM pupils as less able than the national assessment at KS2 does. As a consequence, the success rate in selection is systematically lower for FSM pupils than for non-FSM pupils, even after accounting for KS2 attainment. The difference in the overall success rate between FSM and non-FSM pupils is small at lower KS2 levels, as pupils with weaker performance are usually not academically prepared for grammar schools, regardless of their FSM eligibility status. However, the gap widens at higher KS2 levels, and the most pronounced difference between the two groups is among pupils who have reached the expected academic level of grammar schools. The underestimation of FSM pupils in the selection test based on their KS2 attainment levels may be the result of several factors, such as private coaching in rich families, the difference in measured construct in the 11+ and KS2 assessment, and biased

test content. However, lacking public information on the quality of the 11+ test, it is impossible to pinpoint the reason.

Despite the explanations above for the low probability of disadvantaged pupils attending grammar schools, which have easier and more direct solutions, the most fundamental reason for their underrepresentation in grammar schools is still their lower average achievement at the end of primary schools as a result of the influence of multiple social disadvantages (Gorard et al., 2006). Similar to international research which has revealed that academic selection in OECD countries is largely decided by pupils' ability, which leaves little space for family background to explain between-school variation (Marks, 2006), the analysis in this study also confirms that pupil's attainment accounts for most of the difference in grammar school opportunities. Therefore, while the gap in grammar school opportunities between pupil groups partly results from the imperfect selection process described above, broader social inequality which layered early-age attainment is a major issue.

In sum, this study finds that the opportunity to attend grammar schools is not equitably distributed across social groups. Although grammar schools select by 'ability', the lower rate of disadvantaged pupils taking and passing the selection test, even after accounting for prior attainment, means that the selection fails to offer equal opportunities for pupils with equivalent performance. This implies that the selection system does not achieve the principle of unbiased merit-based equity. Moreover, the status quo of the layered attainment between social groups means that if secondary schools are allowed to select by ability, they are actually selecting pupils from more advantaged backgrounds. Due to the link between attainment and family background, even a fair and valid selection process based purely on ability cannot eliminate the strong influence of family background on grammar school opportunities. Thus, the selection process is not only rewarding merit, but also family advantage. This does not help pupils with greater needs, and may in fact imped their chances. Therefore, the grammar school selection system has failed to achieve needs-based equity either.

18.2 The effectiveness of grammar schools in improving pupils' academic performance

18.2.1 Evidence from OLS and logistic regression models

After accounting for pupil-level and school-level background variables, the overall estimated grammar school effect is low—equivalent to 0.3 grades on capped GCSE results. This is a

small difference that accumulates from KS2 to KS4. Meanwhile, although pupils are more likely to achieve basic KS4 levels in grammar schools, high-achieving pupils may do even worse in grammar schools than equivalent pupils in non-selective schools. The pattern reveals that there may be some positive practices in grammar schools which help pupils achieve satisfactory basic levels, but not in top grades. The conclusion of the effectiveness of grammar schools, in terms of KS4 academic results, corresponds to previous findings that grammar school attendance is beneficial to borderline pupils, but not to pupils at the higher end of the performance distribution (Levačić & Marsh, 2007; Schagen & Schagen, 2003).

In addition to the overall results of grammar school effectiveness in comparison with non-selective schools, the evaluation of individual selective LAs reveals differentiated patterns. Similar to the unbalanced grammar school opportunities in each LA, grammar school effectiveness across LAs is also variegated. Based on OLS linear regression models, about 2/3 of selective LAs have positive results for grammar schools in each GCSE outcome. However, the estimated grammar school effect in each LA is sensitive to statistical choices. Approximately 90% of selective LAs have both positive and negative results when different outcome variables and models are applied. When the three GCSE outcome variables (total, capped and average point scores) in the presented models (OLS, fixed slope and random slope ML) are combined, only three LAs have consistently positive results for grammar schools. These three LAs each have low proportions of grammar school places, and also have more advantaged populations than most selective LAs. Meanwhile, the evaluation of the effectiveness of grammar schools and the degree of selectivity in each LA found no systematic relationship between these two elements. Therefore, the reason for any differences in grammar school effectiveness is more complicated than the degree of selectivity alone, which may be correlated with the wider historical, political, geographical and SES context of each area.

18.2.2 Evidence from the RDD approach

Due to the imperfect 11+ data, the estimation results in the RDD part are not definitive answers to the effectiveness of grammar schools. According to the RDD using different functional forms in the parametric approach, and testing different bandwidths in the non-parametric approach, the estimated treatment effect of grammar schools in the chosen selective LA is approximately 4.5 point scores on capped GCSE. This is equivalent to half a grade per GCSE subject, and is about 10% of the average attainment of pupils below the cut-off point. The

results thus imply a benefit from attending grammar schools for borderline pupils in the participating LA, similar to the 5 A*-C GCSE logistic regression models. The RDD part in this study also overlap with the only previous RDD research on the grammar school effect in England, which revealed that the treatment effect on the Year 9 test score was 7% of the average performance of pupils just below the cut-off point (Clark, 2010). Due to the two studies' differences in school-leaving age and school system, a comparison of the treatment effect at the same school stage is not feasible. The results of the RDD in this study are also within the range of the national pattern of possible grammar school advantage presented by Coe et al. (2008). However, as the effectiveness of grammar schools in this participating LA is already larger than the national average according to traditional school effectiveness models, we may expect the national pattern using the RDD to be smaller than the estimation in this LA as well.

18.2.3 The differential effectiveness of grammar schools for FSM pupils

One of the most cited benefits of grammar schools is the small achievement gap between their pupils of high and low SES groups (Andrews, Hutchinson & Johnes, 2016). However, the analysis in this study reveals that the small variation in KS4 attainment in grammar schools is largely a result of the homogeneous KS2 performance level of their intakes. The evaluation of pupils in comprehensive areas which have achieved similar KS2 performance for grammar school pupils reveals the equivalent KS4 performance between the FSM and non-FSM group as well.

While the regression models present evidence that attending grammar schools is slightly more beneficial for FSM pupils than for the vast majority of non-FSM pupils, there may be confounding factors. Since the estimation results present the performance of grammar school pupils in relation to their peers in non-selective schools, the higher coefficients for FSM pupils in grammar schools may be due to the worse conditions of non-selective schools attended by FSM pupils. Pupils from poorer backgrounds are less likely to attend high-performing schools and those eligible for FSM are usually enrolled in schools of inferior quality (Tes, 2018). Therefore, even if grammar schools are equally effective for FSM and non-FSM pupils, the difference between grammar schools and non-selective schools may still be larger for FSM pupils. Additionally, it has been mentioned that this tiny group of high-performing and highly-motivated FSM pupils in grammar schools systematically differ from their FSM peers, and their achievement in grammar school may have little relevance to that of others. Therefore, the

superior statistical results for FSM pupils in grammar schools should not be simply regarded as evidence of grammar schools' roles in promoting social mobility by narrowing academic performance between high and low SES groups.

18.3 The effectiveness of selective LAs in improving pupils' academic performance

While attending grammar schools is associated with a small positive impact on KS4 attainment for their borderline pupils, the presence of these academically selective schools does not improve the overall performance standards of the local area. After accounting for pupil-level and school-level characteristics, there is no obvious difference in effectiveness between selective and non-selective LAs. While grammar school pupils perform slightly better than equivalent pupils in comprehensive LAs, pupils in selective LAs who failed to get into grammar school perform worse than those in comprehensive LAs. Thus, these two contrary effects, if this is indeed what they are, lead to a zero-sum situation.

While the results of achieving 5 A*-C grades are the same as those for the capped GCSE point scores above, presenting no difference between the selective and comprehensive system, the rate of achieving 5 A*-A GCSE grades reveals that the selective system is associated with inferior results. The overall probability of achieving 5 A*-A grades in selective LAs is about 90% of non-selective LAs. Meanwhile, regardless of whether pupils attended grammar schools, high-performing pupils in selective LAs do not perform as well as those in comprehensive LAs in terms of the rate of getting 5 A*-A grades. Therefore, while the general result of the effectiveness of the two systems does not reveal superiority of any kind, the detailed pattern of KS4 grades suggests that the selective system may be detrimental to high-performers. This conclusion differs from previous research in developing countries which has reported a positive effect of the selective system, both for high-performing pupils and for those with lower attainment (Duflo, Dupas & Kremer, 2011). This may be because high-quality education is universally provided at the compulsory stage in England. The conclusion of this study is similar to that of Hanushek & Wößmann (2006), who applied the PISA and TIMSS test data and found no evidence that the tracking system promotes efficiency. They noticed that pupils at both ends of the performance distribution progress slower in a tracked system than in a comprehensive system. While there are ways by which academic selection may be helpful in countries with unbalanced education resources and diverse student compositions, there is no evidence that

academic selection promotes the overall performance standards in countries with mature state education systems.

18.4 The relationship between grammar school attendance and participation in HE

While there are some signs that grammar schools are slightly more effective in improving borderline pupils' KS4 attainment, the evaluation of HE participation reveals that pupils in grammar schools are at no advantage when pre-existing differences between pupils are taken into account. Furthermore, for pupils with similar backgrounds and attainment at KS2, attending grammar schools is even associated with a lower rate of attending the Russell Group universities. The pattern persists when pupils' KS4 attainment and background variables are taken into account. For pupils with equivalent attainment and characteristics at KS4, those in grammar school do not have higher chances of either attending HE institutions in general or attending elite universities. This may be a sign that grammar schools do not provide effective extra help which is beneficial to HE participation, such as assistance with HE applications. However, the low estimation results for grammar schools are also likely due to statistical bias when their advantageous school-level variables are controlled for. This is similar to the results of 5 A*-A GCSE grades. Meanwhile, the pattern may be relevant to the Widening Participation initiative, which aims to increase HE participation rates for disadvantaged pupil groups, such as those with low SES and ethnic minorities (DfE, 2017). As grammar schools have fewer disadvantaged pupils than other state-funded schools, the estimation results might indicate an admissions bias against grammar schools. Despite possible explanations for the statistical results, after accounting for pupil-level and school-level baseline variables, the evaluation in this study found no evidence that grammar schools are more successful in terms of HE participation patterns.

18.5 The link between family background and post-18 destination

The analysis of the effectiveness of grammar schools has revealed a small advantage at KS4 performance associated with grammar school attendance for borderline pupils under certain circumstances. Meanwhile, pupils in selective LAs who failed to get into grammar schools progressed slightly slower than equivalent pupils in comprehensive areas. According to the unbalanced opportunity to attend grammar schools and its link with pupils' family background, the selection by ability also acts as a form of social selection. As a result, the academic benefit of attending grammar schools, and the costs of failure to get into these schools, are not equally

distributed among pupils. While the benefit is concentrated in more advantaged pupil groups, pupils in selective LAs without sufficient family support, who thus perform worse than they would have otherwise at the age of 11, may lag further behind. For these reasons, academic selection at the age of 11 may reinforce the influence of family background on later destination. This pattern is confirmed by the stronger link between early-age background and HE participation patterns in selective LAs than in non-selective LAs. However, due to the small difference in effectiveness associated with school types (the two tracks within the selective system do not present highly differentiated academic outcomes), and the low overall proportion of grammar school places in England, the gap between the two types of LAs is small. However, the evidence still does not support the claim that grammar schools would narrow the attainment gap and promote equity. Inversely, the selective system may perpetuate or even reinforce social inequalities. This finding implies that policies relying on the expansion of grammar schools (or the selection system) are unlikely to promote social equity.

19 Implications for policy

After summarising this study's findings, this chapter provides implications for future policy. While some implications suggest radical reforms of the current system, such as closing grammar schools, others are more moderate, which are based on the premise that the current selection system will remain unchanged in the future.

19.1 The effectiveness of grammar schools and selective LAs

As there may be a small advantage associated with grammar school attendance for borderline pupils, it would be worthwhile to explore whether there are beneficial practices in grammar schools and whether they are transferable to other schools. Based on previous lessons from several unsuccessful attempts to import practices from advantaged schools (e.g. private schools) to others, schools with different intakes may not benefit from the same set of practices (Hick & Wrigley, 2009). However, the revealed effectiveness of grammar schools is presumed independent of their advantaged intakes. Therefore, if there are any positive practices in grammar schools, they might also be helpful in less advantaged schools as well.

When the effectiveness of the whole selective system was evaluated, the study found no evidence of its advantage in raising pupils' academic performance, and the selective system may even have negative results for high-performing pupils. This means that the expansion of grammar schools is unlikely to raise national academic standards. Based on the high costs of new grammar school places, the small number of potential participants in grammar schools, and the concurrent need to invest in basic educational areas in England (Weale, 2018), expanding grammar schools would not be a wise decision. The evidence also implies that maintaining current grammar schools is unlikely to generate any substantial academic gain. advising possible benefits of converting grammar schools into comprehensive schools. Meanwhile, similar to the situation of early-age academic selection, we may then expect policies advocating selection based on other characteristics which are indirect to academic performance (e.g. specialist, faith, gender) to be even less likely to bring any obvious academic benefits.

19.2 The link between family background and post-18 destination

While separating pupils into different tracks at some point is an unavoidable feature of the education systems of most countries (e.g. HE selection), the evidence in this study reveals that

an early-age selection system is unhelpful and even harmful. Based on the evidence that the opportunity to attend grammar schools is dependent on pupils' family backgrounds, and that the link between origin and destiny is stronger in selective LAs, policies relying on the academic selection system to promote social equity are unlikely to work. On the contrary, the selective system may even perpetuate or reinforce the attainment gap and hinder educational equity. In contrast, a more comprehensive compulsory education would allow the state education system more time to compensate for the initial disadvantages of pupils with low attainment at early-ages (Lucas, 2001). This implies that converting the academic selective system into a comprehensive system and closing grammar schools might create a more equitable secondary education system in England. However, as any substantial changes to the education system are difficult and time-consuming, the following sections discuss possible actions which would improve the current grammar school selection system. These suggestions assume that no other substantial changes are made to the selection system.

19.3 Geographical differences in grammar school opportunities

While it has been frequently mentioned that coaching gives affluent pupils an advantage in grammar school selection tests, the results of this study reveal that a simpler, but effective action for the more affluent would be to let their children sit the 11+ in other LAs with more grammar school opportunities. Therefore, if the current system of skewed grammar school opportunities and the freedom to apply across LAs is to remain, improvements are needed. For example, if the financial ability to resolve the geographical constraints should not become a deciding factor in pupils' grammar school opportunities, then a travel bursary for poor children living further away would be a responsible option. Although the government currently provides free school transport to some disadvantaged pupil groups (e.g. low-income family and SEN) under certain circumstances, it is not really relevant to the case of grammar schools. Moreover, as information about applying to grammar schools outside home LAs might not be available to parents from less advantaged backgrounds, more accessible information that can be accepted and valued by them should be provided to ensure grammar schools would be contained in their school choices.

19.4. The unbalanced rate to take the grammar school selection test

Since the patterns of decision-making in taking the grammar school selection test largely vary across social groups, an opt-out test system which automatically enters all eligible pupils for

the test, rather than an opt-in system asking parents to register for the test, could relieve the unbalanced grammar school opportunities across pupil groups. As the 11+ file in this study only covers one LA, more analysis of selective LAs which have adopted an opt-out selection system should be conducted to assess whether parents of disadvantaged pupils are still more likely to give up on the selection test, even if no extra effort is required to register for the test. A thorough understanding of their real-life difficulties in accessing more advantaged schools is also needed.

19.5 FSM pupils' lower average performance on the 11+

Lack of public information on the quality of the grammar school selection tests means that it is hard to determine the reason for the lower average performance of FSM pupils on the 11+, compared to non-FSM pupils with equivalent KS2 attainment. This apparent underperformance of FSM pupils on the 11+ might be because the two tests assess different aspects of 'ability'. However, it is also possible that the selection tests are biased, which systematically underestimate poorer pupils. Thus, the pattern calls for a reliable disclosure of the details on grammar schools' selection tests. Meanwhile, FSM pupils' inferior scores on the selection test might also be tied to the private coaching that more affluent pupils received. Although test designers have attempted to make grammar school selection tests 'tutor-proof', the unbalanced results between high and low SES pupils with equal KS2 attainment persist. Therefore, if the current system of selecting pupils into grammar schools is kept intact, primary schools in selective areas may need to compensate for poorer pupils' disadvantage. One way to do this would be to familiarise their pupils with the format and content of the grammar school selection tests and provide guidance on subjects which are less similar to the national curriculum (e.g. reasoning).

Meanwhile, as some pupil groups might encounter more challenges than others due to weaker family support, compensatory policies such as a quota system may be an option to enrol a proportion of disadvantaged pupils and alleviate the segregated pupil composition in grammar schools. Some researchers have also advocated that FSM pupils should be given extra marks on the 11+ in order to balance their disadvantages (Allen, Bartley & Nye, 2017). While these actions may provide simple and direct answers to the low grammar school opportunities for disadvantaged pupils, they may also raise philosophical, political, or practical difficulties in deciding who deserves extra help, the degree of compensation permitted, and how different

combinations of pupil characteristics should be treated. Therefore, an evaluation of the possible results of these policies is also needed before implementation.

Based on the varied opportunity to attend grammar schools between different pupil groups, both in terms of geographical location and pupil backgrounds after accounting for prior attainment, there is plenty of chances to improve the fairness of grammar school selection, if the current system is kept. The solutions to these issues should be easier and quicker than the solutions to the more fundamental issue of early-age layered attainment, as explained below.

19.6 Layered early-age attainment between social groups

Although the difference in grammar school opportunities between pupil groups is considerable, during the process of grammar school selection, attainment is still the most important deciding factor. Thus, the broader social inequality that engenders the attainment gap from an early-age is a more demanding issue. This cannot be solved by simply creating more grammar school places, as previously advocated by the Conservative Party (DfE, 2016). Because of the skewed performance between different pupil groups, new grammar school places will still be disproportionately filled by advantaged groups, according to the results of this study. The problem is unlikely to be solved by merely providing extra help to disadvantaged pupils during the selection process, and the solution should involve changing the structure to ameliorate the overall underachievement of certain groups from an early-age. As the difference appears before grammar school selection, more attention should be given to the primary school stage, or possibly even the pre-school age, such as early childhood education and care (Breen & Jonsson, 2005). However, as mentioned by previous researchers, education's role in narrowing the attainment gap is limited. The low attainment of certain social groups is a manifestation of multiple social disadvantages, and the inequalities in education are often reflections of more profound problems within the communities and societies in which schools function (Coleman et al., 1996; Gorard et al., 2006; Jencks et al., 1972). Based on the importance of the out-of-school context and broader social inequality, education alone cannot solve the attainment gap between pupil groups.

19.7 Disclosure of the 11+ data

Limited by the 11+ data, most evidence in this study is produced through traditional regression models controlling for pre-existing differences between pupils in different academic tracks.

Similar to all the regression models trying to control for baseline differences, the imperfect control process is always a threat to causal inferences. Despite the benefit of applying the RDD design, the implementation is constrained by the limited access to the 11+ data in England. The 11+ data in this study only covers a single LA, and difficulties in matching it to the national dataset have forced the analysis to exclude many valid cases.

The threats imposed by unavailable 11+ data are not unique to this study. They are also relevant to all future research attempting to present a fair and accurate evaluation of England's selective system. The results of this study reveal that a strong research design which bypasses previous limitations in grammar school evaluation is workable. However, a definitive answer to the grammar school effect cannot be reached without transparent and reliable disclosure of the national 11+ data linked to later achievement. Government policy requires the support of hard evidence based on robust research designs (See, 2018). However, this absence of data prevents accurate evaluation, even with the aid of effective research methods. Disclosure of the 11+ data is especially important at the present stage when the expansion of grammar schools is being promoted by multiple groups and actions are being taken to implement them, including spending scarce public money on this.

20 Conclusion

To conclude, this study has found only limited evidence that grammar schools are more effective than non-selective mainstream state-funded schools in terms of academic performance, which is most relevant to pupils with lower attainment in grammar schools, and does not apply to high-performing pupils. Meanwhile, the study found no advantage from grammar school attendance for HE participation. When the effectiveness of selective LAs is compared to that of non-selective LAs, the former is not associated with any academic benefit. On the contrary, the selective system might even have a negative influence on high-performing pupils, as the rate of achieving high grades at KS4 is lower in both types of schools in selective LAs than in comprehensive areas. This means the assumed advantages of grammar schools in academic effectiveness is very limited, and there is no overall positive gain associated with the selective system. The idea that grammar schools and the selective system raise national academic standards is unrealistic.

While the pattern of effectiveness does not support an expansion of grammar schools, the negative consequences on social equity do present evidence against the expansion, or even in favour of their removal. The layered early-age attainment between social groups compounded by the imperfect selection process reveals that grammar school selection is not fair, regardless of whether or not prior attainment is considered. Meanwhile, the unbalanced opportunity to attend grammar schools and the small advantage associated with grammar school attendance (as well as the negative influence on pupils who failed to get into grammar schools) amplify the attainment gap between pupils from high and low SES groups through the mechanism of selection and differentiation. Thus, the system has failed to comply with the principle of equity, in terms of either merit-based or needs-based standards. Moreover, separating pupils into different secondary schools according to their early-age performance has created homogeneous peer groups and segregated school compositions, both in terms of attainment and social backgrounds. This may endanger the coherence and long-term development of society as a whole. The results thus suggest that grammar schools should not be expanded. On the contrary, it might be beneficial if current grammar schools are converted into comprehensive ones.

Overall, the analysis in this study finds no substantial advantage associated with early-age academic selection in England. This conclusion is consistent with international evidence that selection during compulsory education is usually associated with a lower level of equity, and may also reduce performance standards (Hanushek & Wößmann, 2006). The conclusion is not

only relevant to grammar schools in England, but to the widespread global practice of separating pupils at an early-age based on academic ability. These results provide implications for countries with state-wide academic selection systems. They are also of interest to countries which have not adopted between-school selection, but do have within-school selective placement such as tracked curricula. Additionally, these results are relevant to schools adopting other forms of selection, such as single-sex and religious schools, as the underlying principles of selective education are similar.

Although education's role in tackling social inequality is limited, every little change in the education system can influence the life trajectory of generations. Inequalities in education are usually a manifestation of profound social inequalities. However, the unbalanced educational opportunities, conditions, and outcomes among social groups are not inevitable. Schools can be more than a microcosm of the larger society replicating broader social problems, and reforms within schools enacted prior to fundamental social change are indeed a realistic possibility (Goarad & Smith, 2010).

Appendix

Appendix 1: Methods and results of previous studies of the effectiveness of grammar schools in England

The table below includes detailed methods and findings of previous research on grammar schools' effectiveness discussed in Chapter 8.1.1. Since some studies have applied extensive sets of explanatory and outcome variables, it is not possible to include all their results in detail, and the table only presents general patterns for comparison. Studies are arranged in the same order as appeared in the main text.

The table excludes studies applying the National Child Development Study. In these studies, the sample group attended secondary schools from 1969 to 1974, when the reform of comprehensive schools occurred. Therefore, the results have limited relevance to today's schools.

Schagen & Schagen, 2003	
Area	England
Data	1) 1998 KS3 to 2000 KS4
	2) 1995 KS2 to 1998 KS3
Comparison group	Grammar schools vs. comprehensive and secondary modern schools
Model	Multiple regression and multilevel modelling
Explanatory variables	Pupil level: prior attainment, gender, age
	School level: school type, size of school, proportion of FSM pupils
	LA level: proportion of grammar school places of LA
Outcome variable(s)	1) Total GCSE and average GCSE 2) KS3 attainment level
Indicator of outcome	Coefficients
Effect of grammar schools	1) 3 grades on total GCSE or 0.4 grade per GCSE subject for average pupils, no effect for high performers
	2) 0.8 level at KS3 for average pupils (about one year's progress), no effect for high performers
Atkinson, Gregg & McConnell, 2006	
Area	19 LAs with more than 10% grammar school places
Data	1997 KS2 to 2002 KS4
Comparison group	Grammar schools vs. comprehensive schools in non-selective LAs
Model	OLS multiple regression
Explanatory variables	Pupil level: prior attainment, gender, age, ethnicity, FSM, SEN, English as a second language
	School level: school type, school size, single sex
Outcome variable(s)	Total GCSE and capped GCSE

Indicator of outcome	Coefficients
Effect of grammar schools	4 grades on total GCSE or 3 grades on capped GCSE
Levačić & Marsh, 2007	
Area	England
Data	1996 KS2 to 2001 KS4
Comparison group	Grammar schools vs. comprehensive schools in non-selective LAs
Model	Multilevel modelling, logistic regression
Explanatory variables	Pupil level: prior attainment, gender, age
	School level: school type, proportion of FSM pupils, proportion of pupils with SEN statements, proportion of white pupils, pupil-teacher ratio (average for 1997-2001)
Outcome variable(s)	Total GCSE/GNVQ points score, The probability of obtaining 5 or more A*-C GCSE/GNVQ
Indicator of outcome	Coefficients
Effect of grammar schools	6 grades on total GCSE/GNVQ or 25% more likely to achieve 5 A*-C on GCSE/GNVQ
Harris & Rose, 2013	
Area	Buckinghamshire
Data	Borderline pupils in grammar schools and secondary modern schools, who took GCSE between 2007-2009
Comparison group	Grammar schools vs. secondary modern schools
Model	Logistic regression
Explanatory variables	Matched by pupils' prior attainment, gender, FSM, birth month, Pakistani or not, year of examination
Outcome variable(s)	The probability of obtaining 5 or more A*-C GCSE
Indicator of outcome	Coefficients
Effect of grammar schools	10% more likely to achieve 5 A*-C on GCSE/GNVQ
Andrews, Hutchinson & Johnes, 2016	
Area	England
Data	2009/2010 KS2 to 2014/2015 KS4
Comparison group	Grammar schools vs. comprehensive schools
Model	Propensity score matching
Explanatory variables (matched by)	Pupil level: prior attainment at KS2, progress between KS1-KS2, gender, ethnicity, autumn/spring/summer-born
	School level: proportion of FSM pupils, average IDACI,
Outcome variable(s)	Capped GCSE
Indicator of outcome	Coefficients

Effect of grammar schools	About 2.4 grades on capped GCSE for all pupils, and 3.9 grades for FSM pupils
Gorard & Siddiqui, 2018	
Area	England
Data	2010 KS2 to 2015 KS4 (same results for 2014 and 2016 KS4 cohorts)
Comparison group	1) Grammar schools vs. other state-funded schools 2) Grammar schools vs. other state-funded schools in selective LAs
Model	Multi-stage regression models
Explanatory variables	Pupil level: prior attainment, gender, age, ethnicity, KS4 FSM eligibility, whether a pupil has been eligible for FSM in any of the past six years, the number of years in total a pupil was eligible for FSM up to KS4, IDACI, SEN status, English as an additional or second language, whether the pupil moved to the school in the last two years
	School level: school type, segregation residual for FSM status (the amount by which each school's intake deviates from the national average)
Outcome variable(s)	Capped GCSE
Indicator of outcome	The R value of regression models
Effect of grammar schools	No effect (no increase in R value when school type is included in addition to other background variables)
Coe et al., 2008	
Area	England
Data	2001 KS2 to 2006 KS4
Comparison group	Grammar schools vs. other schools
Model	OLS and multilevel modelling
Explanatory variables	Pupil level: prior attainment, gender, ethnicity, FSM, IDACI,
	School level: school type, average KS2 level, proportion of FSM pupils, average IDACI, single sex
Outcome variable(s)	Total points score on GCSE and equivalents Capped points score on GCSE and equivalents
Indicator of outcome	Coefficient
Effect of grammar schools	0-3/4 grade per subject on GCSE and equivalents

Appendix 2: Identifying the treatment effect in the 'fuzzy' RDD

In an ideal 'sharp' RDD, all individuals who have passed the cut-off point would get the treatment and those who have missed it would not. The probability of treatment jumps from 0 to 1 at the cut-off point. However, in a 'fuzzy' RDD, as the jump in the probability of treatment at the cut-off point is lower than 1, the discontinuity at the outcome cannot be simply regarded as the treatment effect. The treatment effect in a 'fuzzy' RDD needs to be recovered by calculating the ratio of the gap in the outcome variable and the gap in the probability of the treatment at the cut-off point (Jacob et al., 2012; Lee & Lemieux, 2009). Therefore, both the outcome variable and the treatment probability on the two sides of the cut-off point need to be calculated. Accordingly, the estimation can be written as:

$$Y_i = \theta + \pi D_i + f_0(X_i) + \mu_{0_i},$$

$$T_i = \eta + \lambda D_i + g_0(X_i) + v_{0_i},$$

where Y_i is the outcome measure for each individual i; T_i is the treatment dummy; X_i is the assignment variable (X_i=0 is the cut-off point); D_i is the binary indicator of whether individual i reached the cut-off point (D_i=1 if $X_i \geq 0$); μ and ν are the random error for each individual. The coefficient π is the 'intend to treat' effect, and the real treatment effect equals the ratio of π/λ. This is the ratio of the discontinuity in the outcome and the discontinuity in the treatment at the cut-off point, as mentioned above. Analytically, the regression equations of the treatment effect can be transformed into equations which have been presented in the main text:

$$Y_i = \alpha + \beta T_i + f(X_i) + u_i, \quad (1)$$

$$T_i = \gamma + \delta D_i + g(X_i) + v_i, \quad (2)$$

where every variable is the same as the above, and the effect of attending grammar school which needs to be estimated now equals β.

As mentioned in the main text, when fitting functional forms in equation (1) and (2), the slopes of the regression lines are allowed to vary on two sides of the cut-off point. This is realised through including interaction terms between the assignment variable (*X*) and the treatment variable (*T*), as well as the interaction between the assignment variable (*X*) and the cut-off point variable (*D*).

Appendix 3: Alternative sampling strategies of the RDD

Instead of only keeping cases which can be uniquely matched, the figure below presents the difference between grammar schools and non-selective schools using all of the cases in the 11+ file (including those with the same combination of baseline variables). Pupil records in the 11+ file are still matched to the NPD based on characteristics, KS2 attainment, and school type. However, instead of linking to the individual capped GCSE result as the KS4 attainment indicator, the average capped GCSE result of each subgroup with the same combination of background variables is calculated for grammar school pupils and their counterparts. While it is impossible to elucidate the individual relationship between the 11+ result and the capped GCSE result in this way, it presents an overall picture of how each group of pupils with the same characteristics are doing in grammar schools and in non-selective schools. Figure A1 presents the average GCSE result at each 11+ score. On Figure A2, cases are grouped into bins, and the number of observations within each bin is represented by the size of the dot. In total, 6,732 records in the 11+ file are matched to the NPD data.

Another alternative for dealing with potential misspecification between the indicator of passing the selection (the 11+ file variable) and real attendance in grammar schools (the NPD variable) is to first match all grammar school pupils in the NPD to all pupils who passed the selection in the 11+ file. Then, the non-grammar school pupils in the NPD can be matched to the remaining unmatched cases in the 11+ file. However, after matching, the sample group is different from the original cohort in the 11+ file, with the FSM proportion in the former being twice as high as in the latter. Therefore, an analysis based on this sample group was not conducted.

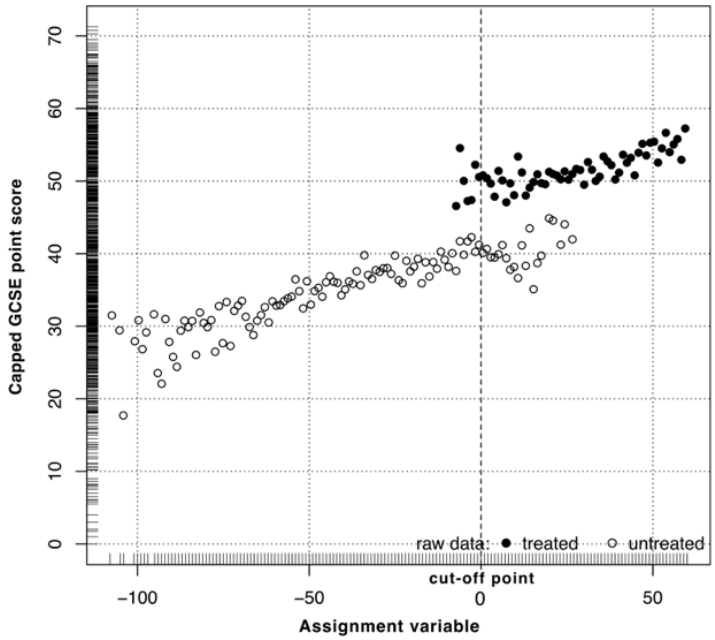

Figure A1: The relationship between the assignment variable and the average capped GCSE of each subgroup in the RDD (raw data)

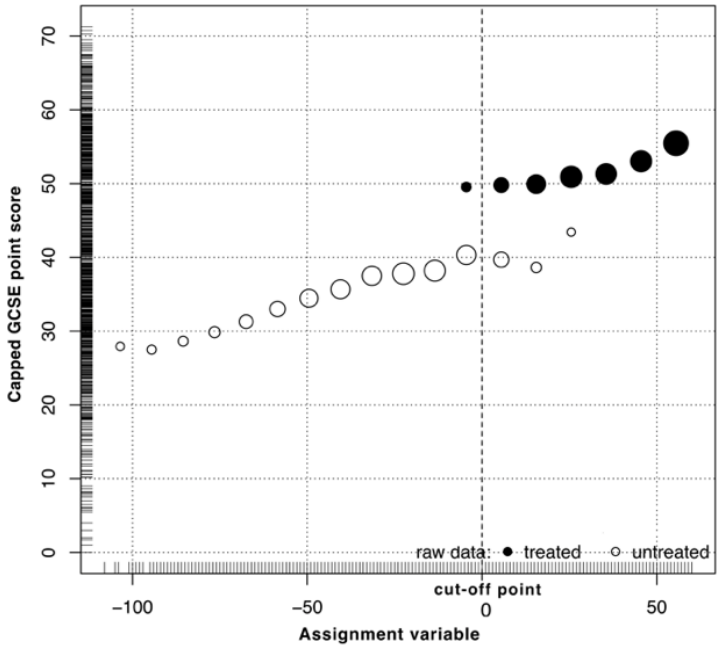

Figure A2: The relationship between the assignment variable and the average capped GCSE of each subgroup in the RDD (grouped data)

Appendix 4: Results of multi-stage OLS models predicting total GCSE point score

	Total GCSE			
Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Girls vs. Boys	3.421	3.148	3.087	3.109
Month Age	0.251	-0.16	-0.143	-0.143
IDACI	-30.997	-15.07	-8.482	-8.833
FSM Eligible	-9.078	-4.295	-3.574	-3.558
SEN School Action Plus or SEN Statement	-16.027	-1.334	-1.883	-1.903
EAL	3.093	4.575	4.666	4.693
Asian	4.634	1.587	1.193	1.164
Black	3.387	2.801	2.417	2.516
Others	3.402	1.839	1.452	1.513
KS2 Total Mark of English And Maths	-	0.379	0.34	0.34
Mean KS2 Total Mark in Secondary School	-	-	0.159	0.111
Mean FSM Proportion in Secondary School	-	-	-0.078	-0.09
Grammar School	-	-	-	2.638
No. of Observation	149,072	149,072	149,072	149,072

Appendix 5: Results of multi-stage OLS models predicting average GCSE point score

	Average GCSE			
Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Girls vs. Boys	0.338	0.316	0.313	0.313
Month Age	0.019	-0.013	-0.012	-0.012
IDACI	-2.493	-1.266	-0.867	-0.869
FSM Eligible	-0.683	-0.314	-0.275	-0.275
SEN School Action Plus or SEN Statement	-1.228	-0.095	-0.132	-0.132
EAL	0.199	0.313	0.316	0.316
Asian	0.401	0.166	0.13	0.13
Black	0.309	0.264	0.232	0.233
Others	0.271	0.151	0.122	0.122
KS2 Total Mark of English And Maths	-	0.029	0.027	0.027
Mean KS2 Total Mark in Secondary School	-	-	0.011	0.011
Mean FSM Proportion in Secondary School	-	-	-0.003	-0.003
Grammar School	-	-	-	0.015
No. of Observation	149,072	149,072	149,072	149,072

Appendix 6: Coefficients for grammar schools in ML regression models after accounting for pupil-level and school-level baseline variables

	Coefficient for grammar schools		
	Total GCSE	Capped GCSE	Average GCSE
Fixed slope ML model	2.18	-0.76	0.003
Random slope ML model	0.80	-2.3	-0.17

Appendix 7: Logistic regression models of HE/the Russell Group participation controlling for KS2-KS5 variables

	HE participation		The Russell group participation	
Variable	**B**	**Exp (B)**	**B**	**Exp (B)**
Girls vs. Boys	0.023	1.024	-0.04	0.961
Month Age	-0.003	0.997	-0.007	0.993
IDACI	-0.02	0.98	-0.553	0.575
KS2 FSM Eligible	-0.1	0.905	-0.062	0.94
SEN School Action	-0.058	0.944	-0.053	0.948
SEN School Action Plus	-0.081	0.923	0.004	1.004
SEN Statement	-0.014	0.986	0.322	1.38
EAL	0.098	1.103	0.009	1.009
Asian	0.532	1.703	0.155	1.168
Black	0.76	2.139	0.431	1.539
Chinese	0.787	2.197	0.449	1.567
Mixed	-0.061	0.941	0.124	1.132
Unclassified	0.231	1.259	0.212	1.236
Other Ethnic Groups	0.188	1.207	-0.062	0.939
KS2 Total Mark	-0.001	0.999	0.008	1.008
Mean KS2 Total Mark in Secondary School	-0.001	0.999	0.018	1.018
Mean FSM Proportion in Secondary School	-0.007	0.993	-0.012	0.988
Grammar School	0.034	1.035	-0.372	0.689
KS4 FSM eligible	-0.098	0.907	-0.005	0.995
KS4 Capped GCSE point score	0.006	1.006	0.016	1.016
KS5 FSM eligible	0.253	1.288	0.026	1.026
KS5 A Level Point Score	0.003	1.003	0.003	1.003
KS5 No. of A*-A in Facilitating Subjects	0.052	1.053	0.51	1.665
No. of Observation	107,506		107,506	

Appendix 8: Classification table of logistic regression models of HE participation without controlling for school compositional variables

Variables	Percentage correctness	Percentage of remaining variation explained
Base figure	56.5	-
Personal background and attainment at KS2	65.8	9.3
School type (grammar school or not)	66.1	0.3
Personal background and attainment at KS4	69.2	3.1
Overall	69.2	12.7

Appendix 9: Classification table of logistic regression models of the Russell Group participation without controlling for school compositional variables

Variables	Percentage correctness	Percentage of remaining variation explained
Base figure	87	-
Personal background and attainment at KS2	87.6	0.6
School type (grammar school or not)	87.7	0.1
Personal background and attainment at KS4	89	1.3
Overall	89	2

Appendix 10: Cohort member and data resources

Cohort member

	Cohort	**Geographical range**	**Linking point**
Grammar school opportunities	2010/2011 KS2	England	NA
Grammar school effectiveness: Traditional regression approach	2010/2011 KS2 to 2015/2016 KS4	England	NA
Grammar school effectiveness: RDD approach	2011/2012 KS2, matched for 2016/2017 KS4	One LA in England	KS2
LA effectiveness	2010/2011 KS2 to 2015/2016 KS4	England	NA
HE participation	2007/2008 KS2 to 2012/2013 KS4 to 2013/2014 KS5, linked with 2015/2016 HE participation records	England	KS5
The link between family background and post-18 destination	2007/2008 KS2 to 2012/2013 KS4 to 2013/2014 KS5, linked with 2015/2016 HE participation records	England	KS5

Data resources and analysis

	Dataset	**Analysis models**	**Outcomes**
Grammar school opportunities	NPD 2010/2011 KS2	Logistic regression models	Attended grammar schools or not
Grammar school effectiveness	NPD 2010/2011 KS2 to 2015/2016 KS4	Linear regression; logistic regression models	1. GCSE point scores 2. 5 or more GCSE at grades A*-C (including English and maths); 3. 5 or more GCSE at grades A*-A;
	The 11+ local file 2011/2012 KS2, matched with NPD 2011/2012 KS2 to 2016/2017 KS4	RDD models	Capped GCSE point scores
LA effectiveness	NPD 2010/2011 KS2 to 2015/2016 KS4	Linear regression; logistic regression models	1. GCSE point scores; 2. 5 or more GCSE at grades A*-C (including English and maths); 3. 5 or more GCSE at grades A*-A;
HE participation	NPD 2007/2008 KS2 to 2012/2013 KS4 to 2013/2014 KS5, linked with HESA 2015/2016	Logistic regression models	1. HE participants or not; 2. Participants of the Russell group universities or not
The link between family background and post-18 destinations	NPD 2007/2008 KS2 to 2012/2013 KS4 to 2013/2014 KS5, linked with HESA 2015/2016	Logistic regression models	1. HE participants or not; 2. Participants of the Russell group universities or not

Bibliography

Adams, R. (2017). Grammar schools may ask parents for donations to cover funding cuts. Retrieved from https://www.theguardian.com/education/2017/jan/25/grammar-schools-may-ask-parents-for-donations-to-cover-funding-cuts

Adams, R. (2018). Third of maintained English secondary schools in the red. Retrieved from https://www.theguardian.com/education/2018/dec/06/third-of-maintained-english-secondary-schools-in-the-red

Allan, A. (2010). Picturing success: young femininities and the (im)possibilities of academic achievement in selective, single-sex education. *International Studies in Sociology of Education, 20*(1), 39-54. doi:10.1080/09620211003655630

Allen, R., Bartley, J., & Nye., P. (2017). *The 11-plus is a Loaded Dice: Analysis of Kent 11-plus Data*. London: Education Datalab.

Allen, R., Mian, E., & Sims, S. (2016). Social inequalities in access to teachers. Retrieved from http://www.smf.co.uk/wp-content/uploads/2016/04/Social-Market-Foundation-Social-inequalities-in-access-to-teachers-Embargoed-0001-280416.pdf.

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

Allport, G. W. (1954). *The nature of prejudice*. MA: Addison-Wesley.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy, 113*(1), 151-184. doi:10.1086/426036

Anderson, C. A. (1961). A Sceptical Note on the Relation of Vertical Mobility to Education. *American Journal of Sociology, 66*(6), 560-570.

Andrews, J., Hutchinson, J., & Johnes, R. (2016). *Grammar Schools and Social Mobility*. London: Education Policy Institute.

Arnott, M., Bullock, A., & Thomas, H. (1992). *Consequences of Local Management: An Assessment by Head* Teachers. Paper presented at the ERA Research Network.

Arnot, M., & Gubb, J. (2001). *Adding value to boys' and girls' education: a gender and achievement project in West Sussex.* Chichester: West Sussex County Council.

Atkinson, A., Gregg, P., & McConnell, B. (2006). *The Result of 11 Plus Selection: An Investigation into Opportunities and Outcomes for Pupils in Selective LEAs*. Bristol: The Centre for Market and Public Organisation.

Au, W. (2009). The condition of equality in the US. In M. Cole (Ed.), *Equality in the secondary school*. London: Continuum.

Ball, S. (2002). *Class strategies and the education market: The middle classes and social advantage*. London: RoutledgeFalmer.

Barber, M., & McCallum, B. (1996). *City Technology Colleges and Technology Colleges: Review of Admissions Procedures at Year 7*. London: DfEE.

Bentham, J. (1948). *An introduction to the principles of morals and legislation*. New York: Hafner.

Bernstein, B. (1975). Class and pedagogies: Visible and invisible. *Educational Studies, 1*(1), 23-41.

Bloom, H., & Porter, K. (2012). *Assessing the Generalizability of Estimates of Causal Effects from Regression Discontinuity Designs. Society for Research on Educational Effectiveness*. Retrieved from https://files.eric.ed.gov/fulltext/ED530557.pdf

Boliver, V., & Swift, A. (2011). Do comprehensive schools reduce social mobility?. *The British Journal of Sociology, 62*(1), 89-110.

Bolton, P. (2017). *Grammar School Statistics*. Retrieved from http://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN01398#fullreport

Bourdieu, P., & Boltanski, L. (1979). Changes in social structure and changes in the demand for education. *Information Sur Les Sciences Sociales XII*, 61-113.

Bourne, J., & Moon, B. (1995). A question of ability? In B. Moon & A. S. Mayes (Eds.), *Teaching and learning in the secondary school* (pp. 25-37). London: Routledge.

Breen, R. (2005). Explaining Cross-National Variation in Youth Unemployment: Market and Institutional Factors. *European Sociological Review, 21*(2), 125-134.

Breen, R., & Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annual Review of Sociology, 31*, 223-243.

Brown, S. (1995). Assessment: a changing practice. In B. Moon & A. S. Mayes (Eds.), *Teaching and learning in the secondary school* (pp. 266-272). London: Routledge.

Brunello, G., Giannini, M., & Ariga, K. (2007). The optimal timing of school tracking: a general model with calibration for Germany. In L. Woessmann & P. E. Peterson (Eds.), *Schools and the equal opportunity problem* (pp. 129-156). Cambridge: The MIT Press.

Bullock, A., & Thomas, H. (1994). *The Impact of Local Management of Schools: Final Report*. Birmingham: University of Birmingham.

Burgess, S., & Thomson, D. (2013). *Key Stage 4 accountability: Progress measure and intervention trigger*. Bristol: Bristol University.

Caldas, S. J., & Bankston, C. L. (1999). Multilevel Examination of Student, School, and District-Level Effects on Academic Achievement. *The Journal of Educational Research, 93*(2), 91-100. doi:10.1080/00220679909597633

Carneiro, P., & Heckman, J. (2003). *Human capital policy NBER Working paper 9547*. Cambridge: NBER.

Castellano, K. E., Rabe-Hesketh, S., & Skrondal, A. (2014). Composition, Context, and Endogeneity in School and Teacher Comparisons. *Journal of Educational and Behavioral Statistics, 39*(5), 333-367. doi:10.3102/1076998614547576

Chowdry, H., Crawford, C., Dearden, L., Goodman, A., & Vignoles, A. (2013). Widening participation in higher education: analysis using linked administrative data. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 176*(2), 431-457. doi:10.1111/j.1467-985X.2012.01043.x

Clark, D. (2010). Selective Schools and Academic Achievement. *The B.E. Journal of Economic Analysis & Policy, 10*(1), 1-40.

Coe, R., Jones, K., Searle, J., Kokotsaki, D., Kosnin, A. M., & Skinner, P. (2008). *Evidence on the effects of selective educational systems: A report for the Sutton Trust.* Retrieved from https://www.suttontrust.com/wp-content/uploads/2008/10/SuttonTrustFullReportFinal-1.pdf

Coffey, M., & Whetton, C. (1996). *Aptitude Tests of Technology: an Investigation of Aptitude and its Relationship with GCSE Scores*. London: DfEE.

Cohen, E. G., & Lotan, R. A. (1995). Producing equal-status interaction in the heterogeneous classroom. *American Educational Research Journal, 32*(1), 99-120. doi:10.2307/1163215

Coldron, J., Willis, B., & Wolstenholme, C. (2009). Selection by attainment and aptitude in English secondary schools. *British Journal of Educational Studies, 57*(3), 245-264.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1996). *Equality of educational opportunity*. Washington: US Government Printing Office.

Condron, D. J. (2013). Affluence, inequality, and educational achievement: A structural analysis of 97 jurisdictions across the globe. *Sociological Spectrum, 33*(1), 73-97.
Cookson, P. (1994). *School Choice*. London: Yale University.
Cooper, D. (2010). *Illusions of Equality (International Library of the Philosophy of Education Volume 7).* London: Routledge.
Coughlan, S. (2016). The persistent appeal of grammar schools. Retrieved from http://www.bbc.co.uk/news/education-30483031.
Crawford, C. (2014). *The link between secondary school characteristics and university participation and outcomes*. Warwick: Institute for Fiscal Studies
Cribb, J., Jesson, D., Sibieta, L., Skipp, A., & Vignoles, A. (2013). *Poor grammar: Entry into Grammar Schools for Disadvantaged Pupils in England.* Retrieved from https://www.suttontrust.com/wp-content/uploads/2013/11/poorgrammarreport-2.pdf
Croxford, L., & Raffe, D. (2007). Education Markets and Social Class Inequality: A Comparison of Trends in England, Scotland and Wales. In R. Teese, S Lamb, M. Duru-Bellat, & S. Helme (Eds.), *International Studies in Educa- tional Inequality, Theory and Policy. Inequality, Educational Theory and Public Policy (Vol. 3)* (pp. 39–66). Berlin: Springer.
Cullinane, C., Hillary, J., Andrade, J., & McNamara, S. (2017). *Selective comprehensives 2017*. Retrieved from https://www.suttontrust.com/wp-content/uploads/2017/05/Selective-Comprehensives-2017.pdf
David, R., Teddlie, C., & Reynolds, D. (2000). *The international handbook of school effectiveness research*. London: Routledge.
Department for Education (DfE). (2016). *Schools that work for everyone: Government consultation*. Retrieved from https://consult.education.gov.uk/school-frameworks/schools-that-work-for-everyone/supporting_documents/SCHOOLS%20THAT%20WORK%20FOR%20EVERYONE%20%20FINAL.PDF
Department for Education (DfE). (2017). *Widening Participation in Higher Education, England, 2014/15 age cohort.* Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/635103/SFR39-2017-MainText.pdf
Department for Education (DfE). (2018). *Secondary accountability measures: Guide for maintained secondary schools, academies and free schools*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/772801/Secondary_accountability_measures_guidance.pdf
Department for Education and Employment (DfEE). (2000) *Statistics of education: schools in England.* London: Her Majesty's Stationery Office.
Department for Education and Department of Health. (2015). *Special Educational Needs and Disability Code of Practice: 0 to 25 Years*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/398815/SEND_Code_of_Practice_January_2015.pdf.
Department for Environment, Food & Rural Affairs. (2014). *The 2011 Rural-Urban Classification for Local Authority Districts in England*. Retrieved from https://www.gov.uk/government/statistics/2011-rural-urban-classification-of-local-authority-and-other-higher-level-geographies-for-statistical-purposes
Department of Education. (2000). *The Effects Of The Selective System Of Secondary Education In Northern Ireland*. Retrieved from https://www.ulster.ac.uk/__data/assets/pdf_file/0020/223409/2000_The_Effects_Of_The_Selective_System_Of_Secondary_Education_In_Northern_Ireland_Main_Report.pdf

Deutsch, M. (1975). Equity, Equality, and Need: What Determines Which Value Will Be Used as the Basis of Distributive Justice? *Journal of Social Issues, 31*(3), 137-149.
Dewey, J. (1916). *Democracy and education*. New York: Macmillan.
Duflo, E., Dupas, P., & Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review, 101*(5), 1739-1774.
Dustmann, C. (2004). Parental background, secondary school track choice, and wages. *Oxford Economic Papers, 56*, 209–230.
Echols, F. H., & Willms, J. D. (1995). Reasons for school choice in Scotland. *Journal of Education Policy, 10*(2), 143-156. doi:10.1080/0268093950100202
Ermisch, J., & Francesconi, M. (2001). Family Matters: Impacts of Family Background on Educational Attainments. *Economica, 68*(270), 137-156. doi:10.1111/1468-0335.00239
Espinoza, J., Lexi Finnigan, & Gurney-Read., J. (2015). 'Floodgates open' for new wave of grammar schools across England. Retrieved from http://www.telegraph.co.uk/education/11934799/Floodgates-open-for-new-wave-of-grammar-schools-across-England.html
Evans, G., Brooks-Gunn, J., & Klebanov, P. (2011). Stressing out the poor: Chronic physiological stress and the income-achievement gap. *Community Investments, 3*, 22-27.
Feinberg, J. (1970). *Doing & deserving; essays in the theory of responsibility*. Princeton: Princeton University Press.
Feinstein, L. (2003). Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica, 70*, 73-97.
Field, S., Kuczera, M., & Pont, B. (2007). *No More Failures: Ten steps to equity in education*. Retrieved from www.sourceoecd.org/education/9789264032590
Finch, L., McCreight, B., & McAleavy, G. (2010). The existence of school-related stress in two grammar schools in Northern Ireland: contributing factors and moderation. *Pastoral Care in Education, 28*(4), 311-329. doi:10.1080/02643944.2010.528016
Foreman-Peck, J. (2004). Spontaneous disorder? A very short history of British vocational education and training, 1563–1973. *Policy Futures in Education*, *2*(1), 72-101.
Galindo-Rueda, F., & Vignoles, A. (2005). *The Heterogeneous Effect of Selection in Secondary Schools: Understanding the Changing Role of Ability*. London: Centre for the Economics of Education.
Gardner, J., & Cowan, P. (2000). *Testing the test: A study of the reliability and validity of the Northern Ireland transfer procedure test in enabling the selection of pupils for grammar school places*. Retrieved from https://eric.ed.gov/?id=ED467432
George, M. (2018). Grammar school expansion returns backed with £200m fund. Retrieved from https://www.tes.com/news/grammar-school-expansion-returns-backed-ps200m-fund
Gewirtz, S., Ball, S. J., & Bowe, R. (1995). *Markets, choice, and equity in education*. London: Open University Press.
Gibbons, S., Machin, S., & Silva, O. (2013). Valuing school quality using boundary discontinuities. *Journal of Urban Economics, 75*, 15-28.
Gillborn, D., & Youdell, D. (2000). *Rationing education: policy practice, reform and equity*. Buckingham: Open University Press.
Glatter, R., Woods, P. E., & Bagley, C. (1997). Diversity, Differentiation and Hierarchy. In C. Bagley et al. (Eds.), *Choice and Diversity in Schooling: Perspectives and Prospects*. London: Routledge.

Goldstein, H. (2001). Using Pupil Performance Data for Judging Schools and Teachers: scope and limitations. *British Educational Research Journal, 27*(4), 433-442.
Gorard, S. (1997). *School choice in an established market.* London: Ashgate.
Gorard, S. (2006). Is there a school mix effect? *Educational Review, 58*(1), 87-94. doi:10.1080/00131910500352739
Gorard, S. (2010). Serious doubts about school effectiveness. *British Educational Research Journal, 36*(5), 745-766. doi:10.1080/01411920903144251
Gorard, S., Boliver, V., Siddiqui, N., & Banerjee, P. (2019). Which are the most suitable contextual indicators for use in widening participation to HE? *Research Papers in Education, 34*(1), 99-129. doi:10.1080/02671522.2017.1402083
Gorard, S., & Fitz, J. (1998). The More Things Change ... The Missing Impact of Marketisation? *British Journal of Sociology of Education, 19*(3), 365-376. doi:10.1080/0142569980190306
Gorard, S., & See, B. H. (2013). *Overcoming Disadvantage in Education*. Abingdon: Routledge.
Gorard, S., & Siddiqui, N. (2018). Grammar schools in England: a new analysis of social segregation and academic outcomes. *British Journal of Sociology of Education, 39*(7), 909-924. doi:10.1080/01425692.2018.1443432
Gorard, S., & Smith, E. (2010). *Equity in education: An international comparison of pupil perspectives*. London: Palgrave Macmillan UK.
Gorard, S., Smith, E., May, H., Thomas, L., Adnett, N., & Slack, K. (2006). *Review of widening participation research: addressing the barriers to participation in higher education*. York: HEFCE.
Gorard, S., & Taylor, C. (2001). The composition of specialist schools in England: track record and future prospect. *School Leadership & Management, 21*(4), 365-381.
Gorard, S., Taylor, C., & Fitz, J. (2002). Does school choice lead to 'spirals of decline'? *Journal of Education Policy, 17*(3), 367-384. doi:10.1080/02680930210127612
Gurin, P., Nagda, B. A., & Lopez, G. E. (2004). The Benefits of Diversity in Education for Democratic Citizenship. *Journal of Social Issues, 60*(1), 17-34.
Gurin, P. Y., Dey, E. L., Gurin, G., & Hurtado, S. (2003). How does racial/ethnic diversity promote education? *The Western Journal of Black Studies, 27*(1), 20-29.
Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica, 69*(1), 201-209.
Hammer, T. (2003). The Probability for Unemployed Young People to Re-Enter Education or Employment: A Comparative Study in Six Northern European Countries. *British Journal of Sociology of Education, 24*(2), 209-223.
Hanushek, E. A., & Wößmann, L. (2006). Does Educational Tracking Affect Performance and Inequality? Differences- in-Differences Evidence Across Countries. *The Economic Journal, 116*(510), C63-C76. doi:10.1111/j.1468-0297.2006.01076.x
Harding, E. (2017). We've not given up on grammars, says May: PM backs expansion of selective schools for first time since election. Retrieved from http://www.dailymail.co.uk/news/article-4990862/Theresa-backs-grammar-school-expansion.html.
Harker, R., & Tymms, P. (2004). The Effects of Student Composition on School Outcomes. *School Effectiveness and School Improvement, 15*(2), 177-199. doi:10.1076/sesi.15.2.177.30432
Harris, D., & Williams, J. (2012). The association of classroom interactions, year group and social class. *British Educational Research Journal, 38*(3), 373-397.

Harris, R., & Rose, S. (2013). Who benefits from grammar schools? A case study of Buckinghamshire, England. *Oxford Review of Education, 39*(2), 151-171. doi:10.1080/03054985.2013.776955
Harvey, J. (2018). Charter schools in the context of poverty, changing demographics, and segregation. In C. Rotberg & J. L. Glazer (Eds.), *Choosing charters: better schools or more segregation?* (pp. 24-40). New York: Columbia University.
Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
Haveman, R. H., & Wolfe, B. L. (1984). Schooling and Economic Well-Being: The Role of Nonmarket Effects. *The Journal of Human Resources, 19*(3), 377-407. doi:10.2307/145879
Herbert, D. (2000). School choice in the local environment: headteachers as gatekeepers on an uneven playing field. *School Leadership and Management, 20*(1), 79-98.
Hick, P., & Wrigley, T. (2009). Promoting equality: pedagogy and policy. In M. Cole (Ed.), *Equality in the secondary school: Promoting good practice across the curriculum* (pp. 36-56). London: Continuum.
Hofman, R. H., & Hofman, A. (2001). School Choice, Religious Traditions and School Effectiveness in Public and Private Schools. *International Journal of Education and Religion*, *2*(1), 144-164.
Horn, D. (2009). Age of selection counts: a cross-country analysis of educational institutions. *Educational Research and Evaluation, 15*(4), 343-366. doi:10.1080/13803610903087011
House of Common. (2004). *Secondary Education: School Admissions Volume 1*. Retrieved from https://publications.parliament.uk/pa/cm200304/cmselect/cmeduski/58/58.pdf
House of Common. (2017). *Evidence Check: Grammar schools*. Retrieved from https://publications.parliament.uk/pa/cm201617/cmselect/cmeduc/780/780.pdf
Howard, A. (2013). *Learning privilege: Lessons of power and identity in affluent schooling*. London: Routledge.
ILEA. (1995). Widening the achievement concept. In B. Moon & A. S. Mayes (Eds.), *Teaching and learning in the secondary school* (pp. 47-49). London: Routledge.
Iannelli, C. (2013). The role of the school curriculum in social mobility. *British Journal of Sociology of Education, 34*(5-6), 907-928. doi:10.1080/01425692.2013.816031
Imbens, G. W., & Kalyanaraman, K. (2012). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *Review of Economic Studies*, *79*(3), 933-959.
Imbens, G. W., & Lemieux, T. (2008). Regression Discontinuity Designs: A Guide to Practice. *Journal of Econometrics, 142*(2), 615-635.
Jackson, C. (2002). Can Single-sex Classes in Co-educational Schools Enhance the Learning Experiences of Girls and/or Boys? An Exploration of Pupils' Perceptions. *British Educational Research Journal, 28*(1), 37-48. doi:10.1080/01411920120109739
Jackson, C. K. (2013). Can higher-achieving peers explain the benefits to attending selective schools? Evidence from Trinidad and Tobago. *Journal of Public Economics, 108*, 63-77.
Jacob, R., Zhu, P., Somers, M. A. e., & Bloom, H. (2012). *A practical guide to regression discontinuity*. Retrieved from https://www.mdrc.org/sites/default/files/RDD%20Guide_Full%20rev%202016_0.pdf
Jacobs, N. (2013). Understanding school choice: Location as a determinant of charter school racial, economic, and linguistic segregation. *Education and Urban Society, 45*(4), 459-482.

Jencks, C. S., Smith, M., Ackland, H., Bane, M. J., Cohen, D., Gintis, H., & Michelson, S. (1972). *Inequality: A reassessment of the effect of the family and schooling in America*. New York: Basic Books.

Jennings, J. L., Deming, D., Jencks, C., Lopuch, M., & Schueler, B. E. (2015). Do differences in school quality matter more than we thought? New evidence on educational opportunity in the twenty-first century. *Sociology of Education, 88*(1), 56-82.

Jesson, D. (2013). *The creation, development and present state of grammar schools in England*. York: University of York.

Judson, R. (1998). Economic Growth and Investment in Education: How Allocation Matters. *Journal of Economic Growth, 3*(4), 337-359.

Kalogrides, D., & Loeb, S. (2013). Different teachers, different peers: The magnitude of student sorting within schools. *Educational Researcher, 42*(6), 304-316.

Kamin, L. J. (1981). *Intelligence: the battle for the mind: HJ Eysenck verses Leon Kamin*. London: Pan Books.

Kang, C., Park, C., & Lee, M. J. (2007). Effects of ability mixing in high school on adulthood earnings: Quasi-experimental evidence from South Korea. *Journal of Population Economics, 20*, 269-297.

Kentish, B. (2018). Tory Government's child literacy scheme 'will do nothing' to reverse billions in school budget cuts, Labour says. Retrieved from https://www.independent.co.uk/news/uk/politics/investment-literacy-schools-justine-greening-angela-rayner-reading-writing-reverse-budget-cuts-a8145386.html

Kerckhoff, A. C., Fogelman, K., Crook, D., & Reeder, D. (1998). *Going Comprehensive in England and Wales: A Study of Uneven Change*. London: Routledge.

Kiselica, M. S., Baker, S. B., Thomas, R. N., & Reedy, S. (1994). Effects of stress inoculation training on anxiety, stress, and academic performance among adolescents. *Journal of Counseling Psychology, 41*(3), 335.

Kitchener, D. (2013). What Price Free Schools? The Continued Insidious Privatisation of UK State Education. *Forum, 55*(3), 407-414.

Kurlaender, M., & Yun, J. T. (2006). Measuring school racial composition and student outcomes in a multiracial society. *American Journal of Education, 113*(2), 213-242.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational evaluation and policy analysis, 24*(1), 37-62.

Lareau, A. (1997). Social-class differences in family-school relationships: the importance of cultural capital. In H. L. A.H. Halsey, Phillip Brown, Amy Stuart Wells (Ed.), *Education, culture, economy, society* (pp. 703-717). Oxford: Oxford University.

Larkin, J., & Staton, P. (2001). Access, inclusion, climate, empowerment (AICE): A framework for gender equity in market-driven education. *Canadian Journal of Education/Revue canadienne de l'éducation*, 361-376.

Le Grand, J. (1991). Quasi-Markets and Social Policy. *The Economic Journal, 101*(408), 1256-1267.

Le Grand, J., & Bartlett, W. (1993). Quasi-markets and social policy: the way forward? In J. Le Grand & W. Bartlett (Eds.), *Quasi-Markets and Social Policy* (pp. 202-220). Basingstoke: Macmillan Press.

Leckie, G., & Goldstein, H. (2009). *The limitations of using school league tables to inform school choice, Working Paper 09/208.* Bristol: Centre for Market and Public Organisation.

Leckie, G., & Goldstein, H. (2017). The evolution of school league tables in England 1992–2016: 'Contextual value-added', 'expected progress' and 'progress 8'. *British Educational Research Journal, 43*(2), 193-212. doi:10.1002/berj.3264

Lee, D., & Lemieux, T. (2009). Regression discontinuity designs in economics. Retrieved from https://www.nber.org/papers/w14723.pdf

Lee, V. E., & Burkam, D. T. (2002). *Inequality at the starting gate*. Washington: Economic policy institute.

Levačić, R., & Marsh, A. J. (2007). Secondary modern schools: are their pupils disadvantaged?. *British Educational Research Journal, 33*(2), 155-178.

Levin, B., & Riffel, J. (1997). School System Responses to External Change: implications for school choice. In R. Glatter, P. Woods, & C. Bagley (Eds.), *Choice and Diversity in Schooling. Perspectives and Prospects*. London: Routledge.

Levin, H. M. (1990). The theory of choice applied to education. In J. Whitte & J. Clune (Eds.), *Choice and control in American education, Volume 1: The theory of choice decentralization and school restructuring*. Basingstoke: Falmer Press.

Levin, H. (1992). Market approaches to education: vouchers and school choice. *Economics of Education Review, 11*, 279-285.

Levin, H. (2018). Charter schools: rending or mending the nation. In C. Rotberg & J. L. Glazer (Eds.), *Choosing charters: better schools or more segregation?* (pp. 195-204). New York: Columbia University.

Lu, B. (2018). Selection on attainment? Local Authorities, pupil Backgrounds, attainment and grammar School Opportunities. *Educational Review*. doi:10.1080/00131911.2018.1483893

Lu, B. (2019). How can we evaluate the effectiveness of grammar schools in England? A regression discontinuity approach. *British Educational Research Journal*. doi:10.1002/berj.3581

Lucas, S. R. (2001). Effectively maintained inequality: Education transitions, track mobility, and social background effects. *American Journal of Sociology, 106*(6), 1642-1690.

Luyten, H. (2006). An empirical assessment of the absolute effect of schooling: regression-discontinuity applied to TIMSS-95. *Oxford Review of Education, 32*(3), 397-429. doi:10.1080/03054980600776589

Luyten, H., Tymms, P., & Jones, P. (2009). Assessing school effects without controlling for prior achievement. *School Effectiveness and School Improvement, 20*(2), 145-165.

Mansfield, I. (2019). *The impact of selective secondary education on progression to higher education*. Oxford: HEPI.

Marks, G. N. (2005). Cross-National Differences and Accounting for Social Class Inequalities in Education. *International Sociology, 20*(4), 483-505. doi:10.1177/0268580905058328

Marks, G. N. (2006). Are between- and within-school differences in student performance largely due to socio-economic background? Evidence from 30 countries. *Educational Research, 48*(1), 21-40. doi:10.1080/00131880500498396

Marsh, H.W. (1984). Self-concept, social comparison, and ability grouping: A reply to Kulik and Kulik. *American Educational Research Journal, 21*(4), 799-806.

Marsh, H.W., & Hau, K.T. (2003). Big-Fish–Little-Pond Effect on Academic Self-Concept. *American Psychological Association, 58*(5), 364-376.

Massey, D. S. (2006). Social background and academic performance differentials: White and minority students at selective colleges. *American Law and Economics Review, 8*(2), 390-409.

Massey, D. S., & Fischer, M. J. (2006). The effect of childhood segregation on minority academic performance at selective colleges. *Ethnic and Racial Studies, 29*(1), 1-26.

McCulloch, A. (2015). *Does Ability-based Selection Have to Increase Inequalities in Education?* . Retrieved from https://educationdatalab.org.uk/2015/11/does-ability-based-selection-have-to-increase-inequalities-in-education/
McGuinn, P., & Hess, F. (2000). *Business as usual: the minimal competitive effects of the Cleveland Voucher Program*. Paper presented at the AERA, New Orleans.
Mickelson, R. A., Nkomo, M., & Wimberly, G. L. (2012). Integrated Schooling, Life Course Outcomes, and Social Cohesion in Multiethnic Democratic Societies. *Review of Research in Education, 36*, 197-238.
Miller, D. (1999). *Principles of Social Justice*. Cambridge, MA: Harvard University Press.
Moore, D., & Davenport, S. (1990). School choice: The new, improved sorting machine. In W. Boyd & H. Walberg (Eds.), *Choice in education* (pp. 187-223). Berkeley: McCutchan.
Morris, A. B. (2001). Patterns of performance of Catholic schools in England. *Networking, 3*(1), 17-21.
Morris, R., & Perry, T. (2017). Reframing the English grammar schools debate. *Educational Review, 69*(1), 1-24.
Mortimore, P. (1997). Can effectiveness schools compensate for society?. In H. L. A.H. Halsey, Phillip Brown, Amy Stuart Wells (Ed.), *Education, culture, economy, society* (pp. 476-488). Oxford: Oxford University.
Mortimer, P., & Whitty, G. (2000). *Can school improvements overcome the effects of disadvantage?*. London: Institute of Education University of London.
Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School matters: The junior years*. Somerset: Open Books.
Nash, R. (2003). Is the School Composition Effect Real?: A Discussion With Evidence From the UK PISA Data. *School Effectiveness and School Improvement, 14*(4), 441-457. doi:10.1076/sesi.14.4.441.17153
Nelson, R. M., & DeBacker, T. K. (2008). Achievement motivation in adolescents: The role of peer climate and best friends. *The journal of experimental education, 76*(2), 170-189.
Nye, P. (2016). *Understanding Grammar Schools*. Retrieved from https://educationdatalab.org.uk/2016/12/understanding-grammar-schools/
OECD. (2010). *PISA 2009 Results: Overcoming Social Background: Equity in Learning Opportunities and Outcomes (Volume II)*. Paris: OECD.
OECD. (2015). *Universal Basic Skills: What Countries Stand to Gain*. Retrieved from http://www.oecd.org/edu/universal-basic-skills-9789264234833-en.htm
OECD. (2016). *Education in China: a snapshot*. Retrieved from http://www.oecd.org/china/Education-in-China-a-snapshot.pdf
Parsons, C., & Welsh, P. J. (2006). Public sector policies and practice, neo-liberal consumerism and freedom of choice in secondary education: a case study of one area in Kent. *Cambridge Journal of Education, 36*(2), 237-256. doi:10.1080/03057640600718661
Parsons, E., Chalkley, B., & Jones, A. (2000). School Catchments and Pupil Movements: A case study in parental choice. *Educational Studies, 26*(1), 33-48. doi:10.1080/03055690097727
Penney, D. (2004). Policy tensions being played out in practice. The Specialist Schools initiative in England. *Journal for Critical Education Policy Studies, 2*(1), 227-246.
Perry, T. (2016). English Value-Added Measures: Examining the Limitations of School Performance Measurement. *British Educational Research Journal, 42*(6), 1056-1080. doi:10.1002/berj.3247

Perry, T. (2019). 'Phantom' compositional effects in English school value-added measures: the consequences of random baseline measurement error. *Research Papers in Education, 34*(2), 239-262.
Peterson, P. E., & Woessmann, L. (2007). Introduction: schools and the equal opportunity problem. In L. Woessmann & P. E. Peterson (Eds.), *Schools and the equal opportunity problem* (pp. 3-28). Cambridge: The MIT Press.
Pfeffer, F. T. (2015). Equality and quality in education. A comparative study of 19 countries. *Social Science Research, 51*, 350-368.
Piaget, J. (1983). Piaget's theory. In P. H. Mussen (Ed.), *Handbook of child psychology: Vol. 1. History, theory, and methods* (pp. 103–126). New York: John Wiley.
Power, S., Edwards, T., Whitty, G., & Wigfall, V. (2003). *Education and the middle class*. Buckingham: Open University Press.
Power, S., Fitz, J., & Halpin, D. (1994). Parents, pupils and grant-maintained schools. *British Educational Research Journal, 20*(209-226).
Power, S., & Frandji, D. (2010). Education markets, the new politics of recognition and the increasing fatalism towards inequality. *Journal of Education Policy, 25*(3), 385-396. doi:10.1080/02680930903576404
Power, S., Halpin, D., & Whitty, G. (1997). Managing the state and the market: 'new' education management in five countries. *British Journal of Educational Studies, 45*(4), 342-362. doi:10.1111/1467-8527.00057
Power, S., & Whitty, G. (2015). Selective, Comprehensive and Diversified Secondary Schooling in England: A Brief History. In A. de Waal (Ed.), *The Ins and Outs of Selective Secondary Schools: A Debate* (pp. 1-8). London: Civitas.
Raffe, D., Brannen, K., Fairgrieve, J., & Martin, C. (2001). Participation, Inclusiveness, Academic Drift and Parity of Esteem: A Comparison of Post-Compulsory Education and Training in England, Wales, Scotland and Northern Ireland. *Oxford Review of Education, 27*(2), 173-203.
Randall, Jeff. (2009). How the Class War Backfired and Put Social Mobility into Retreat. Retrieved from https://www.telegraph.co.uk/finance/comment/jeffrandall/5941721/How-the-class-war-backfired-and-put-social-mobility-into-retreat.html.
Rasbash, J., Leckie, G., Pillinger, R., & Jenkins, J. (2010). Children's Educational Progress: Partitioning Family, School and Area Effects. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 173*(3), 657-682.
Raudenbush, S. W., & Willms, J. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics, 20*(4), 307-335.
Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.
Reichelt, M., Collischon, M., & Eberl, A. (2019). School tracking and its role in social reproduction: reinforcing educational inheritance and the direct effects of social origin. *The British Journal of Sociology*. doi:10.1111/1468-4446.12655
Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., & Stringfield, S. (2014). Educational effectiveness research (EER): a state-of-the-art review. *School Effectiveness and School Improvement, 25*(2), 197-230. doi:10.1080/09243453.2014.885450
Robertson, S., & Lauder, H. (2001). Restructuring the education/social class relation: a class choice. In R. Phillips & J. Furlong (Eds.), *Education, reform and the state: twenty-five years of politics, policy and practice* (pp. 222-236). London: RoutledgeFalmer.
Roemer, J. E. (1998). *Equality of opportunity*. Cambridge: MA: Harvard University.
Roemer, J. E. (2002). Equality of Opportunity: A Progress Report. *Social Choice and Welfare, 19*(2), 455-471.

Roemer, J. E., & Trannoy, A. (2016). Equality of opportunity: Theory and measurement. *Journal of Economic Literature, 54*(4), 1288-1332.
Rutter, M. (1979). *Developmental psychiatry*. London: Heinemann.
Rutter, M. (1983). School Effects on Pupil Progress: Research Findings and Policy Implications. *Child Development, 54*(1), 1-29. doi:10.2307/1129857
Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. London: Open Books.
Sacker, A., Schoon, I., & Bartley, M. (2002). Social inequality in educational achievement and psychosocial adjustment throughout childhood: magnitude and mechanisms. *Social Science & Medicine, 55*, 863–880.
Schagen, I., & Schagen, S. (2003). Analysis of National Value-added Datasets to Assess the Impact of Selection on Pupil Performance. *British Educational Research Journal, 29*(4), 561-582.
See, B.H. (2018). Evaluating the evidence in evidence-based policy and practice: Examples from systematic reviews of literature. *Research in Education, 102*(1), 37-61.
See, B.H. & Gorard, S. (2019). Why don't we have enough teachers?: A reconsideration of the available evidence. *Research Papers in Education.*
See, B.H., Gorard, S. & Siddiqui, N. (2017). Does participation in uniformed group activities in school improve young people's non-cognitive outcomes? *International Journal of Educational Research, 85*, 109-120.
Sellgren, K. (2018). The 16 grammars that have won funds to expand. Retrieved from https://www.bbc.com/news/education-46429040
Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generliazed causal inference*. New York: Houghton Mifflin Company.
Sibieta, L. (2016). *Can Grammar Schools Improve Social Mobility?* Retrieved from https://www.ifs.org.uk/publications/8469
Sibieta, L. (2018). *School funding per pupil falls faster in England than in Wales* Vol. 2018. Retrieved from https://www.ifs.org.uk/publications/13143
Siddiqui, N. (2017). Parental education as a determinant of school choice: A comparative study of school types in Pakistan. *Research in Education, 99*(1), 3-18.
Siddiqui, N., Gorard, S. & See, B.H. (2018). The importance of process evaluation for randomised control trials in education. *Educational Research, 60*(3), 357-370.
Siddiqui, N. & Ventista, O.M. (2018). A review of school-based interventions for the improvement of social emotional skills and wider outcomes of education. *International Journal of Educational Research, 90*, 117-132.
Smith, D. J., & Tomlinson, S. (1989). *The school effect: A study of multi-racial comprehensives*. London: Policy Studies Institute.
Smith, E. (2005). *Analysing underachievement in schools*. London: A&C Black.
Smith, J., & Naylor, R. (2001). Determinants of Degree Performance in UK Universities: A Statistical Analysis of the 1993 Student Cohort. *Oxford Bulletin of Economics and Statistics, 63*(1), 29-60. doi:10.1111/1468-0084.00208
Smithers, A., & Robinson, P. (2010). *Worlds Apart: social variation among schools*. Retrieved from https://www.suttontrust.com/wp-content/uploads/2010/03/Worlds_apart.pdf
Spielhofer, T., Benton, T., & Schagen, S. (2004). A study of the effects of school size and single-sex education in English schools. *Research Papers in Education, 19*(2), 133-159. doi:10.1080/02671520410001695407
Sternberg, R., Wagner, R., Williams, W., & Horvath, J. (1995). Testing common sense. *The American Psychologist, 42*, 912–926.

Strand, S. (2010). Do some schools narrow the gap? Differential school effectiveness by ethnicity, gender, poverty, and prior achievement. *School Effectiveness and School Improvement, 21*(3), 289-314. doi:10.1080/09243451003732651
Strike, K. A. (1979). The Role of Theories of Justice in Evaluation: Why a House Is Not a Home. *Educational Theory, 29*(1), 1-9. doi:10.1111/j.1741-5446.1979.tb00831.x
Stringer, N. (2008). Aptitude tests versus school exams as selection tools for higher education and the case for assessing educational achievement in context. *Research Papers in Education, 23*(1), 53-68. doi:10.1080/02671520701651771
Sullivan, A., Parsons, S., Green, F., & Wiggins, R. (2016). *Social origins, elite education and elite destinations*. London: IOE.
Teddlie, C., & Reynolds, D. (2000). *The international handbook of school effectiveness research*. London: Falmer Press.
Tes. (2018). Fewer poor pupils in high-performing schools. Retrieved from https://www.tes.com/news/fewer-poor-pupils-high-performing-schools
The Conservative Party. (2017). *Forward Together, Our Plan for a Stronger Britain and Prosperous Future: The Conservative and Unionist Party Manifesto 2017*. London: St. Ives PLC.
Thomas, S., & Mortimore, P. (1995). Comparison of value-added models for secondary-school effectiveness. *Research Papers in Education, 11*(1), 5-33. doi:10.1080/0267152960110103
Thomas, S., Sammons, P., Mortimore, P., & Smees, R. (1997). Differential Secondary School Effectiveness: comparing the performance of different pupil groups. *British Educational Research Journal, 23*(4), 451-469.
Tomlinson, S. (1997). Diversity, choice and ethnicity: The effects of educational markets on ethnic minorities. *Oxford Review of Education, 23*(1), 63-76.
Tomlinson, S. (2005). Race, ethnicity and education under New Labour. *Oxford Review of Education, 31*(1), 153-171.
Trochim, W. (1984). *Research design for program evaluation: the regression-discontinuity approach*. California: Sage Publications.
Tudge, J. R., Hogan, D. M., Snezhkova, I. A., Kulakova, N. N., & Etz, K. E. (2000). Parents' child-rearing values and beliefs in the United States and Russia: The impact of culture and social class. *Infant and Child Development: An International Journal of Research and Practice, 9*(2), 105-121.
UNESCO. (2018). *Handbook on Measuring Equity in Education*. Montreal: UNESCO Institute for Statistics.
Van de Werfhorst, H. G., & Mijs, J. J. B. (2010). *Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective*. Amsterdam: University of Amsterdam.
Van Houtte, M. (2004). Gender context of the school and study culture, or how the presence of girls affects the achievement of boys. *Educational Studies, 30*(4), 409-423. doi:10.1080/0305569042000310336
Vigdor, J., & Nechyba, T. (2007). Peer effects in North Carolina public schools In L. Woessmann & P. E. Peterson (Eds.), *Schools and the equal opportunity problem* (pp. 73-102). Cambridge: The MIT Press.
Visscher, A. J. (2001). Public school performance indicators: Problems and recommendations. *Studies in Educational Evaluation, 27*(3), 199-214.
Walford, G. (2001). Affirming and contesting the comprehensive ideal: from common schooling to selection? In R. Phillips & J. Furlong (Eds.), *Education, reform and the state: twenty-five years of politics, policy and practice* (pp. 45-57). London: RoutledgeFalmer.

Walker, I., & Zhu, Y. (2011). Differences by degree: Evidence of the net financial rates of return to undergraduate study for England and Wales. *Economics of Education Review, 30*(6), 1177-1186.
Walzer, M. (1984). *Spheres of Justice: A Defense of Pluralism and Equality*. New York: Basic Books.
Warnock, H. M. (1978). *Special educational needs: Report of the committee of enquiry into the education of handicapped children and young people* (Vol. 7212). London: Stationery Office Books (TSO).
Weale, S. (2018). Campaigners criticise £50m fund for grammar schools' expansion. Retrieved from https://www.theguardian.com/education/2018/dec/03/campaigners-criticise-50m-fund-for-grammar-schools-expansion
Weldon, M. (2018). *Secondary school choice and selection*. London: DfE.
West, A., & Hind, A. (2003). *Secondary Schools Admissions in England: Exploring the Extent of Overt and Covert Selection London School of Economics and Political Science*. Retrieved from www.risetrust.org.uk.
Whitty, G., Halpin, D., & Power, S. (1998). *Devolution and choice in education: the school, the state, and the market*. Buckingham: Open University Press.
Windle, J. (2009). The limits of school choice: some implications for accountability of selective practices and positional competition in Australian education. *Critical Studies in Education, 50*(3), 231-246. doi:10.1080/17508480903009566
World Bank. (2006). *World Development Report 2006: Equity and Development*. Washington DC and New York: World Bank and Oxford University Press.
Wylie, C. (1994). *Self-Managing Schools in New Zealand: The Fifth Year*. Wellington: New Zealand Council for Educational Research.
Wylie, C. (1995). Contrary Currents: The application of the Public Sector Reform Framework in Education. *New Zealand Journal of Educational Studies, 20*(2), 149–164.
Zimmer, R. (2003). A new twist in the educational tracking debate. *Economics of Education Review, 22*(3), 307-315.