Ethnic school segregation in England: measurement, change and prediction



Peter Maurice Mitchell

DPhil in Education

Lady Margaret Hall, University of Oxford

Supervised by:

Professor Steve Strand

Professor Lars-Erik Malmberg

Acknowledgements

I would like to express my deep gratitude for all the support provided to me by my two supervisors, Professor Steve Strand and Professor Lars Malmberg. They have given me a constant stream of support, via their feedback on the various drafts of this thesis and their words of encouragement during our supervision meetings. I deeply appreciate all the time they have spent helping me complete this project. Thank you.

I would also like to thank the UK Economic and Social Research Council (ESRC) and All Souls College, Oxford for providing me with the funding required to complete this PhD. Without their financial support, I would not have been able to complete this research.

Finally, I would like to thank my wife, Tisha Verma, without whom I would be lost.

Abstract

As England becomes increasingly ethnically diverse there have been concerns over perceived ethnic divisions within society (Cantle, 2013; Casey, 2016; May, 2016). Schools have been a focal point for these discussions, and the government is now committed to ensuring state schools become places where students create "relationships with others from different ethnic... groups" (HM Government, 2018a, p. 26). Unfortunately, academic literature on ethnic school segregation suffers from three issues. First, it is fairly small and often dated. Only a limited number of studies have measured ethnic school segregation across England, and many rely on data from the turn of the century. Second, no national study has measured segregation within local educational markets across England and many frame ethnic identity as "White" or "Non-White" which does not reflect the diversity of the school population. Finally, there is limited evidence on the predictors of ethnic school segregation, which hinders our ability to understand why different patterns exist.

Using school census data, this study provides the most in-depth, up-to date analysis of ethnic school segregation in England. It maps patterns in detail, investigates how segregation varies between areas with different characteristics and explores whether certain types of state schools contribute more to school segregation than others. Additionally, using a bespoke extract from the National Pupil Database on the application choices of year 6 students, the study charts how segregation shifted during the 2016/17 school admissions cycle and explores how segregation may have differed, had students been allocated to state schools using alternative approaches. The thesis also helps pioneers three methodological techniques. It is the first national study to use a multi-group measure

of ethnic school segregation (*H* Index) to calculate the level of segregation between five major ethnic groups. Additionally, by grouping sets of roughly six schools into "pseudo-neighbourhoods" using a clustering algorithm the study provides a new measure of ethnic school segregation calculated within local educational markets. Finally, it used simulation analysis to explore the links between school admissions and ethnic school segregation.

The study found that between 2006 and 2019 the median level of ethnic school segregation for pseudo-neighbourhoods in England fell by 25%, from an *H* value of 6.8 to 5.1, which suggests that ethnic groups became more evenly spread across local state schools during this period. After controlling for the number of schools within each pseudo-neighbourhood, segregation was found to be higher, on average, in pseudo-neighbourhoods with lower average house prices and higher rates of entitlement to Free School Meals. Ethnic school segregation also tended to be higher in primary school pseudo-neighbourhoods with a greater proportion of religious schools and secondary school pseudo-neighbourhoods with a grammar school. Subsequent models found that faith schools tended to contribute more to ethnic school segregation, on average, than non-religious schools in the same pseudo-neighbourhood, with schools linked to minority religions contributing particularly disproportionate amounts. Finally, school segregation was found to rise between each stage of the admissions process and all seven simulations exhibited lower median H values than observed in practice. Simulations that allocated students to schools using academic banding within the current school choice system, or that used a strict home-school distance rule without any parental choice, produced some of the lowest median H values. The implications of these results for policy making, public debates and school practices are discussed.

Contents

Acknowledgements....2
Abstract....3
Contents....5
List of Figures....10
List of Tables....13
List of Equations....15
1 Introduction....16
1.1 The changing ethnic profile of England and its school age population....16
1.2 Ethnic divisions within the school system....17
1.3 Rationale for the study....22
2 Literature Review....26
2.1 Understanding "Ethnic School Segregation"....28
2.1.1 What conceptualisation of "segregation" is being studied....28
2.1.2 Where are you measuring segregation? – Defining the geographic area(s) of study 33
2.1.3 Who is being studied? – Defining the ethnic groups....40
2.1.4 How will you measure segregation? – Selecting the right approach to measurement....44
2.1.5 How will you describe the results? – the importance of clarity and considerate language....48
2.2 Research measuring ethnic school segregation in England....52
2.2.1 Previous academic studies that measure evenness ethnic school segregation in England....53
2.2.2 Other studies of ethnic "separateness" within the English school system....59
2.3 Research into the predictors of ethnic school segregation....65
2.3.1 Population density....66
2.3.2 Socio-economic status of the neighbourhood....67
2.3.3 Selective admissions policies....70
2.3.4 School Quality....77
2.4 England's National School Admissions System....80
2.4.1 Understanding England's National School Admission System....81
2.4.2 Past research into how the school admission system leads to the separation of ethnic groups within English schools....84

2.4.3 A new approach to investigating how school admissions link to patterns of ethnic school segregation 88
2.5 Summary of the literature review 94
3 Research Questions 95
4 Methodology Chapter 101
4.1 Data Sources 102
4.2 Population and samples 103
4.2.1 School Census Data 104
4.2.2 National Pupil Database 109
4.3 Defining the geographic areas of study using k-means clustering 111
4.4 Defining five ethnic groups 117
4.5 Outcome Variables 119
4.5.1 H Index – Ethnic School Segregation 119
4.5.2 Measuring the contributions of schools to ethnic school segregation (H_j) 127
4.5.3 Measuring ethnic isolation - Shared Peer Ethnicity (SPE) 130
4.6 Control variables 131
4.6.1 Assigning pseudo-neighbourhoods to regions and cities 132
4.6.2 Number of schools included in the pseudo-neighbourhood 133
4.7 Predictor variables 134
4.7.1 Population density 138
4.7.2 Free School Meal Eligibility (FSM) 138
4.7.3 Average house price 139
4.7.4 Academically selective admissions policies (grammar schools) 140
4.7.5 Religious schools 140
4.7.6 OFSTED rating 142
4.8 Variables used in the simulation 142
4.8.1 KS2 Academic Results 143
4.8.2 Estimated Published Admissions Numbers 143
4.8.3 Preference Data 144
4.9 Modelling and analysis techniques 146
4.9.1 Descriptive statistics and sensitivity checks 147
4.9.2 Meta Regression Modelling 148
4.9.3 Simulation Analyses 155
4.10 Ethical Considerations 162
5 Describing the ethnic profile of England's school age population, the patterns of ethnic school segregation and the level of ethnic isolation between 2006 – 2019. 165

5.1 The ethnic profile of England's school age population 165
5.1.1 The current ethnic profile of England's school age population.....165
5.1.2 Changes in the ethnic profile of England's school aged population between 2006 – 2019 174
5.2 Ethnic isolation.....182
5.2.1 The current levels of ethnic isolation in English schools. 182
5.2.2 Changes in the levels of ethnic isolation in English schools between 2006 - 2019 185
5.3 Ethnic school segregation 188
5.3.1 Current levels of ethnic school segregation in England 189
5.3.2 Changes in the level of ethnic school segregation between 2006 – 2019 193
6 Exploring the predictors of ethnic school segregation 198
6.1 Pseudo- neighbourhood characteristics 198
6.1.1 Control variables 202
6.1.2 Population Density.....202
6.1.3 Socio-economic profile 203
6.1.4 School Characteristics 205
6.2 School level factors 209
6.2.1 The contribution of academically selective schools to ethnic school segregation 209
6.2.2 The contribution of religious schools to ethnic school segregation.....213
6.2.3 The association between a school's OFSTED rating and its contribution towards ethnic school segregation.....222
7 Investigating how the national school admission system links to patterns of ethnic school segregation 228
7.1 The shifting patterns of ethnic school segregation at each stage of the national school admissions process.....228
7.2 Exploring patterns of ethnic school segregation within seven simulated school admissions systems.....234
8 Discussions.....242
8.1 England's changing ethnic profile.....246
8.2 Ethnic Isolation in England's schools 250
8.3 Ethnic school segregation in England 252
8.4 The predictors of ethnic school segregation in England.....256
8.4.1 Population Density (Research Question 4a and 4b).....256
8.4.2 Socio-economic profile of pseudo-neighbourhoods (Research Question 4c – 4f) 258

8.4.3 Academic Selection (Research Question 4d and 5a) 260
8.4.4 Religious Schooling (Research Question 4h, 4i and 5b - 5e) 262
8.4.5 School Quality (Research Question 4j, 4k, 5f and 5g) 268
8.5 The role of the school admission system in shaping patterns of ethnic school segregation in England 270
8.5.1 How did ethnic school segregation shift between each stage of the national school admissions cycle? 271
8.5.2 What the simulated school admissions systems tell us about ethnic school segregation in England? 274
8.6 Methodological Contributions of this study 279
8.6.1 The creation of "pseudo-neighbourhoods" 280
8.6.2 Using the multi-group H Index 283
8.7 Limitations and Future Research 285
8.7.1 Limitations 286
8.7.2 Future Research 287
8.8 Summary 290
9 Conclusion 292
10 Bibliography 302
11 Appendix 312
11.1 How this study conducted its Literature Review 312
11.2 Modelling Segregation 314
11.3 Evidence on the association between racial or ethnic school segregation and academic performance 317
11.4 Information on the samples used in different years of the study. 320
11.5 Exploring the distribution of pseudo-neighbourhood sizes when *k* = 2800 for primary school pseudo-neighbourhoods and *k* = 550 for secondary school pseudo-neighbourhoods. 327
11.6 Comparing the number of schools in rural and urban pseudo-neighbourhoods when using k-means clustering 328
11.7 Comparing an *H* Index calculated using 5 vs 18 ethnic groups. 329
11.8 Calculating the level of "random segregation" or "upward bias" in *H* when measuring ethnic school segregation in England's school system 331
11.9 Comparing the relationship between a multigroup *H* Index, a multigroup *D* Index and a two group *D* Index. 333
11.10 Breakdown of the *H* Index 334
11.11 The Two group D Index 335
11.12 The Isolation index (*I*) 335

11.13 Exploring the relationship between the number of clusters (*k*) and the level of segregation (*H*) on English school data....336

11.14 Diagrams to show the numeric and geographic distribution of each of the predictor variables....337

11.15 Evaluating FSM as a measure of socio-economic disadvantage....344

11.16 Comparing the correlation of a school and pseudo-neighbourhood's house price when using a LSOA data or a cluster approach....345

11.17 Exploring the change in *H* for pseudo-neighbourhoods which were sufficiently diverse to be included in the sample in both 2006 and 2019....347

11.18 Exploring whether changes in the level of ethnic school segregation reflect changes in the ethnic profile of pseudo-neighbourhoods....347

11.19 Exploring whether the level of "random segregation" changed between 2006 – 2019. 349

11.20 Additional diagrams from the results section....350

11.21 Full results for pseudo-neighbourhood level meta-regression model....362

11.22 Model checks for pseudo-neighbourhood level meta-regression analysis....377

11.23 Model checks for school level meta-regression analysis....379

11.24 Contrasting the raw and ranked contributions towards ethnic school segregation (H_j) of schools with selective schools, faith schools and schools with different OFSTED ratings....380

11.25 Tables of results from the Simulation Analysis....385

11.26 Additional diagrams from Chapter 7....388

11.27 Using different sized pseudo-neighbourhoods to measure ethnic school segregation at various geographic scales....390

11.28 Random Forest Models....393

11.28.1 Explaining random forest models....393

11.28.2 Applying random forest models to this study....395

11.28.3 Findings from the random forest models....396

List of Figures

Figure 1: Comparing two schools with high and low levels of ethnic isolation. 30
Figure 2: Evenness segregation 32
Figure 3: Explaining how evenness segregation indexes work 47
Figure 4: Concentration profiles for white and non-white students across 3060 English Secondary Schools in 2001 *(Johnston, Wilson and Burgess, 2004)* 55
Figure 5: Ethnic school segregation in England in 2001 (Burgess, Wilson and Lupton, 2005). 57
Figure 6: Mapping the location of pseudo-neighbourhoods included in Sample 2 108
Figure 7: The location of pseudo-neighbourhoods included in the Sample 3 and 4 109
Figure 8: Diagram to show how k-means clustering is used to make pseudo-neighbourhoods. 113
Figure 9: The relationship between the value of *k* and the average number of schools within a pseudo-neighbourhood (England, 2019). 114
Figure 10: The Principle of Transfers and Exchanges 121
Figure 11: Compositional Invariance 122
Figure 12: Explaining what a school's "contribution to ethnic school segregation" means .130
Figure 13: An ethnic breakdown of England's school age population by phase of schooling (2019) 166
Figure 14: An ethnic breakdown of England's school age population by region (2019).....167
Figure 15: Comparing a pseudo-neighbourhood's population density against its proportion of students from each ethnic group in 2019, split by phase of schooling.....169
Figure 16: Mapping the ethnic profile of pseudo-neighbourhoods in England (2019).....171
Figure 17: A map to show which ethnic group, if any, were in the absolute majority for each pseudo-neighbourhood in 2019 173
Figure 18: Change in the number of students from different ethnic backgrounds in the English school system between 2006 – 2019. 175
Figure 19: A scatter plot to show the changes in the number of White British and ethnic minority students in English pseudo-neighbourhoods between 2006 – 2019. 177
Figure 20: The changing ethnic profile of England's school age population by region (2006 – 2019). 178
Figure 21: Plotting the percentage point changes in each ethnic group against a pseudo-neighbourhood's population density by phase (2006 – 2019).....180
Figure 22: Bar chart to show the number of pseudo-neighbourhoods where each ethnic group was in the majority (2006 – 2019).....181
Figure 23: A density plot to show the distribution of SPE values for students from different ethnic background in the English state school system (2019). 183
Figure 24: Comparing the change in the median average Shared Peer Ethnicity (SPE) for student attending schools in England (2006 – 2019). 185
Figure 25: A density plot, with a boxplot overlayed, to show the distribution of *H* values for primary and secondary school pseudo-neighbourhoods in England (2019).....189
Figure 26: Mapping the geographic patterns of H for pseudo-neighbourhoods in England (2019). 191

Figure 27: The distribution of H values for ethnically diverse pseudo-neighbourhoods in England (2006 – 2019). 194
Figure 28: A line graph to show the changes in the average level of H for pseudo-neighbourhoods in different regions of England (2006 – 2019). 195
Figure 29: Comparing the ethnic profile of grammar and non-selective schools contained in Sample 3.....210
Figure 30: Comparing the contributions to ethnic school segregation made by grammar and non–selective schools in ethnically diverse pseudo-neighbourhoods (2019).....211
Figure 31: Bar chart to describe the religious denomination of schools in the sample. 216
Figure 32: The ethnic profile of religious and non-religious schools in the sample.....217
Figure 33: The ethnic profile of the schools contained in the sample by religious denomination (2019). 218
Figure 34: Breakdown of the OFSTED ratings recieved by schools in the sample (2019).....223
Figure 35: The collective ethnic profile of schools grouped by OFSTED rating (2019). 224
Figure 36: Distribution of Hj values for schools grouped by their OFSTED rating (2019) 225
Figure 37: Regression Output to compare the contributions towards ethnic school segregation (H_j) made by school based on their OFSTED ratings.....225
Figure 38: Boxplot to show how the distribution of *H* values shift at each stage of the national school admissions cycle for secondary school students (2016/2017). 229
Figure 39: A bar chart to show how the median H values for pseudo-neighbourhoods in each region of England shifts between the three stage of the national school admission cycle (2016/2017) 232
Figure 40: A bar chart to show how the median *H* values for pseudo-neighbourhoods located in different physical environments shift between the three stage of the national school admission cycle (2016/2017).....233
Figure 41: Boxplot to describe the distribution of H values for each simulation.....236
Figure 42: A bar chart to describe the median H value for pseudo-neighbourhoods located in each region of England, under each of the seven simulations conducted by this study. 239
Figure 43: A bar chart to describe the median H value for pseudo-neighbourhoods located in different physical environments, under each of the seven simulations conducted by this study.....240
Figure 44: Exploring the number of schools included in a pseudo-neighbourhood when *k = 2800* for primary school pseudo-neighbourhoods and *k = 550 for* secondary school pseudo-neighbourhoods. 328
Figure 45: Comparing the number of schools in rural and urban pseudo-neighbourhoods 329
Figure 46: Correlation between an H index using 5 vs 18 ethnic groups. 330
Figure 47: Exploring the relationship between the *H* index and multi-group *D* Index 334
Figure 48: A Breakdown of the *H* Index.....334
Figure 49: Evaluating the average pseudo-neighbourhood *H* value when using pseudo-neighbourhoods of different sizes. 336
Figure 50: Population Densitys of ethnically diverse pseudo-neighbourhoods in Sample 2 (2019). 337
Figure 51: Mapping the Population Density of ethnically diverse pseudo-neighbourhoods in Sample 2 (2019). 338
Figure 52: Rates of FSM for ethnically diverse pseudo-neighbourhoods in Sample 2 (2019). 338

Figure 53: Mapping the rates of FSM for ethnically diverse pseudo-neighbourhoods in England (2019). 338
Figure 54: The average price of a terrace house in ethnically diverse pseudo-neighbourhoods in Sample 2 (2019)....339
Figure 55: Mapping the average price of a terrace house for pseudo-neighbourhoods in Sample 2 (2019). 339
Figure 56: Density plot to show the proportion of schools that were religious in sample 2 (2019). 340
Figure 57: Mapping proportion of schools that were religious in pseudo-neighbourhoods within Sample 2 (2019). 341
Figure 58: Density plot to show the average OFSTED ratings of pseudo-neighbourhoods in sample 2 (2019). 341
Figure 59: Mapping the average OFSTED Rating of pseudo-neighbourhoods within Sample 2 (2019). 342
Figure 60: Density plot to show the level of ethnic diversity (*E*) in pseudo-neighbourhoods in sample 2 (2019). 342
Figure 61: Mapping the ethnic diversity (*E*) of pseudo-neighbourhoods included in sample 2 (2019). 343
Figure 62: Density plot to show the number of schools in the pseudo-neighbourhoods that were included in sample 2 (2019)....343
Figure 63: Mapping the number of schools in the pseudo-neighbourhoods that were included in sample 2 (2019)....344
Figure 64: Comparing the estimated house price surrounding a school using an LSOA dataset and a cluster approach 346
Figure 65 Comparing the estimated house price in a neighbourhood using an LSOA dataset and a cluster approach 346
Figure 66: Comparing the change in segregation with the change in the level of White British students at a pseudo-neighbourhood level (2006 – 2019) 348
Figure 67: The proportion of ethnic school segregation in English pseudo-neighbourhoods explained by random segregation between 2006 – 2019. 350
Figure 68: The ethnic profile of pseudo-neighbourhoods in England (2019). 351
Figure 69: The percentage of White British students being educated in English pseudo-neighbourhoods (2019). 351
Figure 70: The percentage of students from Asian, Black, Other White, and Other ethnic groups being educated across pseudo-neighbourhoods in England (2019). 351
Figure 71: Mapping the percentage of White British students being educated in English pseudo-neighbourhoods (2019) 352
Figure 72: The changing number of White British and ethnic minority students, at a regional level (2006 – 2019)....352
Figure 73: The changing proportion of White British students at a neighbourhood level (2006 – 2019). 353
Figure 74: The distribution of changes in the proportion of students from each ethnic group within pseudo-neighbourhoods in England (2006 - 2019) 353
Figure 75: Mapping the change in the percentage of students from different ethnic groups a neighbourhood level (2006 – 2019)....355
Figure 76: Comparing the proportion of students from each ethnic group within pseudo-neighbourhoods in 2006 and 2019 356

Figure 77: SPE values for students in 2019. 357
Figure 78: The distribution of SPE values for students attending school in England in 2006 and 2019. 357
Figure 79: The *H* values for pseudo-neighbourhoods in the different regions of England (2016). 358
Figure 80: Mapping the location of religious schools in ethnically diverse pseudo-neighbourhoods (2019). 359
Figure 81: The collective ethnic profile of schools with different religous denomination, split by phase (2019) 360
Figure 82: Comparing the distributions of H_j between religious and non-religious schools (2019). 361
Figure 83: Comparing the distributions of H_j for schools with different religious denominations (2019) 361
Figure 84: Map to show the geographic distribution of *H* values for secondary school pseudo-neighbourhoods at each stage of the 2017/2018 national school admissions cycle. 388
Figure 85: Map showing how *H* values for pseudo-neighbourhoods differed between stage 1 and 2, and stage 1 and 3 of the 2017/2018 national school admissions cycle. 389
Figure 86: Map showing how H values for pseudo-neighbourhoods differed between what was observed, and the values produced by the seven simulations national school admission systems. 389
Figure 87: The distribution of H values for pseudo-neighbourhoods with an average of 5, 10, 15 and 20 schools 391
Figure 88: Plotting the average level of ethnic school segregation within pseudo-neighbourhoods with an average of 5, 10, 15 and 20 schools between 2006-2019 392
Figure 89: A fictitious tree-based model to predict the level of segregation based on the level of FSM, house price and ethnic diversity (E). 394
Figure 90: Diagram to show how random forest models work. 395
Figure 91: Comparing the predictive power of each independent variable on a pseudo-neighbourhood's H value. 397

List of Tables

Table 1: Describing the distribution of schools across different sets of government boundaries (2019) 38
Table 2: Oversubscription admissions criteria in publicly-funded secondary schools in London between 2001-2015 (West and Hind, 2016) 83
Table 3: Overview of datasets 102
Table 4: Describing the criteria used to make each sample and the research question it was used to answer. 104
Table 5: Describing the number of students, schools, and pseudo-neighbourhoods in Sample 1 - 4 used by the study (2019) 105
Table 6: Describing the ethnic profile of Samples 1 - 4 (2019) 106

Table 7: The number students, schools and pseudo-neighbourhoods that were used when calculating ethnic school segregation in each simulation. 110
Table 8: Procedure used by k means clustering to create pseudo-neighbourhoods.....113
Table 9: The number of schools contained within pseudo-neighbourhoods in England in 2019, when k = 2800 for primary school pseudo-neighbourhoods and *k* = 550 for secondary school pseudo-neighbourhoods. 116
Table 10: Condensing 18 ethnic groups into 5 groups. 117
Table 11: Comparing the properties of the H and D Index 126
Table 12: Descriptive statistics for the study's control variables 131
Table 13: The number of pseudo-neighbourhoods and schools in each region in Sample 2 132
Table 14: Descriptive statistics of the *school level* variables used to make pseudo-neighbourhood level predictor variables 134
Table 15: Descriptive statistics *of pseudo-neighbourhood level* predictor variables.....135
Table 16: Correlation matrix for pseudo-neighbourhood level predictor variables.....136
Table 17: Process for estimating the house price around a school. 139
Table 18: Religious groupings used in this study (2019) 141
Table 19: Variables used in the simulation analysis 143
Table 20: Explaining data structure of the student preference data 145
Table 21: Describing the extent to which parents from different ethnic groups completed their child's application form in full. 146
Table 22: Formulas for the pseudo-neighbourhood level meta-regression models 150
Table 23: Formulas for school level meta-regression analysis.....153
Table 24: Model checks 155
Table 25: Procedure for charting how patterns of ethnic school segregation shift between each stage of the 2016/17 application process. 156
Table 26: The seven simulations for allocation year 6 students to their secondary school. 157
Table 27: Meta- regression analysis (Model 6_s) to compare the relative contribution towards ethnic school segregation of grammar and non-selective schools. 212
Table 28: Regression output to show the difference in the contribution to ethnic school segregation of schools with different religious affiliations. 219
Table 29: Table to describe the distribution of H values for each simulation 234
Table 30: Describing the terms used in the literature search conducted on 1st April 2021 312
Table 31: List of academics who were included in the literature search 314
Table 32: Describing the number of students, schools, and pseudo-neighbourhoods in Sample 1 and 2 between 2006 – 2019 320
Table 33: Describing the ethnic profile of Sample 1 and 2 between 2006 – 2019 323
Table 34: The procedure to find the level of "random" ethnic school segregation.....332
Table 35: Full meta-regression output for primary school psuedo-neighbourhood level predictors. 362
Table 36: Full meta-regression output for secondary school pseudo-neighbourhood level predictors. 367
Table 37: Full meta-regression output for Model 6_s 373
Table 38: Full meta-regression output for Model 7_p, 8_p, 7_s and 8_s. 374
Table 39: Full meta-regression output for Model 9_p and 9_s.....376
Table 40: Results from the model checks for the meta regression analysis on primary school pseudo-neighbourhood level models. 378

Table 41: Results from the model checks for the meta regression analysis on secondary school pseudo-neighbourhood level models 379
Table 42: Results from the model checks for the meta regression analysis on school level predictors (section 4.2) 380
Table 43: Formulas for school level meta-regression analysis 381
Table 44: Comparing the raw and ranked contributions towards ethnic school segregation (H_j) of grammar school and non-selective secondary school 383
Table 45: Comparing the raw and ranked contributions towards ethnic school segregation (Hj) of different religious schools 384
Table 46: Comparing the raw and ranked contributions towards ethnic school segregation (Hj) of schools with different OFSTED results 385
Table 47: Summary Statistics to describe the distribution of H values for each simulation 385
Table 48: Describing the number and percentage of students allocated to their 1^{st} – 6^{th} preference school in each simulation. 387

List of Equations

Equation 1: Formula for K means clustering (Jain, 2010) 112
Equation 2: The H Index (Theil and Finizza, 1971) 123
Equation 3: The D Index 335
Equation 4: The Isolation Index (I) 335

1 Introduction

1.1 The changing ethnic profile of England and its school age population

England is in the midst of a profound shift in the ethnic composition of its population. In 1991 the nation was overwhelmingly "white", with only 5.9% of residents identifying as part of any other ethnic group. By 2011, this figure had more than doubled, increasing to 14% of the population (Office for National Statistics, 2012).[1] All ethnic minority groups grew as a proportion of the population and over a million people identified as being part of a "mixed" ethnic category (Centre on Dynamics of Diversity, 2012).

England's school age population replicates these patterns in the extreme (Hamnett, 2012; Harris and Johnston, 2020). As of 2018, 33% of primary school pupils (aged 4 – 10) and 30% of secondary school pupils (aged 11 – 16) came from an ethnic minority background (Strand, 2015, p. 18; Department for Education, 2018a), which is twice the proportion of the wider UK population. Urban schools have become especially diverse, with ethnic minority pupils now constituting a majority in London (81%), Luton (80%), (76%), Leicester (72%), Birmingham (69%), Manchester (63%) and Bradford (57%).[2] However, as more families from ethnic minority groups move into rural and suburban areas, ethnically diverse neighbourhoods are being created throughout England (Harris and Johnston, 2020), often

[1] Unfortunately, it is hard to track the ethnic profile of the UK before the 1990s, as ethnicity was first recorded in the 1991 census and the current 18 ethnic groups were only introduced in the 2001 census (Centre on Dynamics of Diversity, 2012, p. 2).

[2] Based on the authors own calculations using published census data for all state-funded schools in England (Department for Education, 2018a)

for the first time in their history (Hamnett, 2012; Harris, 2016; Catney, 2017; Catney and Lloyd, 2020; Catney, Wright and Ellis, 2020).

These trends are no longer reliant on continued inward immigration. Given the higher fertility rate of women born outside the UK (Office for National Statistics, 2018), the older profile of White British residents (Shankley, Hannemann and Simpson, 2020) and the steady increase in the number of mixed ethnicity families (Centre on Dynamics of Diversity, 2012, p. 1) most experts predict that the ethnic diversification of England's school aged population will continue for the foreseeable future (Finney and Simpson, 2009).

1.2 Ethnic divisions within the school system

As the country becomes more ethnically diverse there has been ongoing concern over the extent to which students of different ethnic groups are educated in different schools. Various media outlets have sounded alarm bells over perceived ethnic divisions within the school system (BBC, 2017; The Guardian, 2017; The Independent, 2017), along with think tanks (Boliver, Burgess and Leunig, 2013; The Challenge, School Dash and The iCoCo Foundation, 2016), government reports (Cantle, 2001; Casey, 2016) and prominent politicians (Cameron, 2015; May, 2016). Consequently, the government have recently announced schools should aim to become places where students "mix and form lasting relationships with others from different ethnic, religious or socio-economic groups" (HM Government, 2018a, p. 26).

The government justifies their position by suggesting that an ethnically divided school system is likely to hinder the ability of young people to "understand" or "get along with people from different backgrounds" (HM Government, 2018a, p. 28). This draws on the work of Allport's "contact theory", which argues that contact between different groups of people, under the right conditions, offers an effective way of improving intergroup relations (Allport, 1954). The theory has fairly widespread acceptance following 50 years of academic evidence collected from around the world (Hewstone and Swart, 2011), with results suggesting that the effect of inter-group contact in building trust and reducing prejudice is fairly strong. A meta-analysis of 500 studies, for example, found that 22% of an individual's level of prejudice towards a group of people was, on average, explained by a lack of contact (Pettigrew and Tropp, 2006). A similar analysis looking at "515 studies involving a quarter of a million participants in 38 nations" found that in 94% of cases there was a negative relationship between contact and participants' level of prejudice (Pettigrew and Tropp, 2008, p. 922).

Studies also suggest schools are ideal places to promote positive inter-group contact. Researchers argue that young people are conscious of group boundaries but flexible in their ways of thinking (Khmelkov and Hallinan, 1999; Siegel-hawley, 2012) which allows them to change their opinions on out-group members more easily than adults (Hawley and Rist, 1975). Evidence tends to back up this position. A survey of 1000 students in Colorado found, for example, that 76% of students felt being part of a racially diverse classroom helped them understand different points of view (Orfield and Frankenberg, 2001). Another study of 163 students in Oldham, an ethnically mixed town in England with a history of ethnic tension, found that "students who attended the ethnically mixed elementary school had more

positive contact experiences and intergroup relations than students in the segregated elementary schools", with "higher contact (both quantity and quality…), higher contact intentions and greater intergroup trust" (Birtel *et al.*, 2020, p. 169). Another larger study, which analysed data on 3958 children in 96 state schools in England, produced similar results. It found that students who attended schools with a higher proportion of peers from another ethnic group had, in most cases, "greater warmth towards and higher shares of friends from that group", which led the authors to conclude that "most of what happens in [English] schools is 'good contact'" (Burgess and Platt, 2020).

Evidence also suggests schools provide a unique space to form inter-group friendships, which is the form of contact most associated with reduced levels of prejudice (Pettigrew, 1997; Pettigrew *et al.*, 2011; Munniksma *et al.*, 2013). These friendships seem to reduce the salience of ethnic identities (Hollingworth and Archer, 2010) by "blurring ethnic and racial boundaries" (Warikoo, 2010, p. 445), thereby lowering the likelihood of inter-group hostility. Viewed collectively, these studies provide fairly solid evidence that if the government successfully encouraged more inter-group contact in schools, inter-group prejudice between young people would fall.

However, this falls well short of "proving" that ethnic segregation within school systems is always a "bad" thing. The impacts of integrated school systems are complex and unlikely to be uniformly positive. Indeed, "you do not have to be a wild-eyed fanatic or an ethnic separatist to believe that some degree of clustering is good for you", or indeed wider society (Phillips, 2013, p. 12). Many parents, irrespective of their ethnic background, are likely to prefer to send their child to a school where a substantial proportion of the school

population looks like them.[3] They may feel this helps develop their child's cultural identity while also protecting them from inter-group hostility or discrimination. Unfortunately, in some parts of the country, these concerns seem vindicated by the experiences of local residents (Ouseley, 2001, p. 10).

Families from ethnic minority backgrounds may have particularly good reasons to be cautious about sending their child to a school where their ethnic group is poorly represented. Unfortunately there is a long history of ethnic minority communities being mistreated within the British education system (Bagley, 1996, p. 570; Peach, 1996; Chessum, 1997) and there continues to be an under-representation of ethnic minority teachers, particularly in senior leadership positions (Cantle, 2001; Elonga Mboyo, 2019). This can leave ethnic minority families feeling weary and/or unable to "effectively participate or contribute to the education agenda" (Ouseley, 2001, p. 11) which in turn creates a feeling of distrust or exclusion. The continued harassment that some minority students suffer presents yet another reason which may discourage mixing (Crozier and Davies, 2008), along with the common narrative in mainstream media outlets which implicitly blames segregation on minority communities (Miah, 2012) while leaving "white middle-class self-segregation unchallenged" (Amin, 2002; Hollingworth and Williams, 2010).[4]

It is also important to highlight that, unlike socio-economic segregation, evidence suggests that there is no association between ethnic segregation in English schools and academic

[3] This human tendency, known as social homophily, is well documented in social psychology literature (Mcpherson, Smith-lovin and Cook, 2001).
[4] Muslim families in particular are likely to suffer from suspicion and distrust (Ouseley, 2001, p. 10; Gale, 2013; Samad, 2013).

achievement.[5] For example, a 2008 study that used the national pupil database to examine the performances of students from different ethnic groups found there was "no evidence of a negative impact of ethnic segregation on absolute attainment levels" and that the "score gap between White and minority ethnic students [was] largely unaffected by segregation" (Burgess *et al.*, 2008, pp. 1 & 17). Such evidence suggests that an ethnically divided school system does not inherently produce low or unequal academic results.[6]

Creating successful diverse schools can also be challenging and riddled with risk, particularly in areas where ethnic tensions are high. While the failure of school busing programmes in Boston, USA is a much cited example (Angrist and Lang, 2004), attempts to integrate "Asian" and "White" schools in northern England were similarly destructive (Miah, 2012). Efforts led to "amplified racial attacks on Muslim pupils and their feelings of insecurity, while also fuelling a backlash against what is perceived by some members of the white working class as a form of social engineering" (Miah, 2012).

Even when schools do manage to create a population which reflects the ethnic diversity of their local neighbourhood, various factors can hinder the establishment of positive inter-group relationships. Academics, for example, have highlighted that greater diversity does not always lead to greater inter-group contact (Hewstone *et al.*, 2018). School policies such as academic tracking can, for example, lead to greater ethnic homogeneity within classes

[5] See Section 11.3 for a more detail discussion of the literature.
[6] This should not be misunderstood as meaning that there are no differences between the average attainment of different ethnic groups, which there are (Strand, 2015). Instead, this suggests that there is limited evidence to suggest that students who attend schools where their ethnic group is under/over-represented, relative to other groups in the same area, perform differently, on average, to those who attend ethnically representative/mixed schools.

than in the wider school population (Rivkin, 2000; Mickelson, 2001; McKeown *et al.*, 2012; Camina and Iannone, 2014). Research has also found students tend to prefer making friends with members of their own ethnic group, which can reduce the net-benefit of diverse school intakes (Mcpherson, Smith-lovin and Cook, 2001; Munniksma *et al.*, 2013). When inter-ethnic friendships do form, research suggests they tend to be of a lower quality than intra-ethnic ones, involving less trust and time spent together (Kisfalusi, 2016, p. 3), less intimacy (Aboud, Mendelson and Purdy, 2003) and less stability (Shook and Fazio, 2008; Rude and Herda, 2010).

Appreciating the challenges and costs of creating ethnically diverse schools does not invalidate the benefits that this type of education can offer young people, or indeed wider society. Instead, it helps to clarify that this is a complex field to operate in, which requires measured discussions and considerate policy making. To this end, high-quality, up-to-date research is essential to guide these efforts as the nation tries to build a school system that can successfully educate the most ethnically diverse generation of British children to date.

1.3 Rationale for the study

Unfortunately, there remains three important gaps in current academic literature on ethnic school segregation in England. Firstly, research is fairly small and often dated. A limited number of studies have measured ethnic school school segregation across the whole country, and many of them use data from the turn of the century (Johnston, Wilson and Burgess, 2004; Burgess, Wilson and Lupton, 2005; Gorard, Hordosy and See, 2013; Gorard,

2016b).[7] Following stark changes in the ethnic profile of England's school age population, many of these patterns are likely to have changed substantially.

Secondly, these articles use methodologies that offer a limited approach to measuring ethnic school segregation in England. Most calculate segregation within Local Authority boundaries, that are too large to be considered a "local educational market", and vary substantially in size, which introduces bias into their results. Additionally, many of the methodological approaches that researchers had previously used no longer offer an appealing approach to measuring segregation in England today.

Finally, there is a lack of research into which types of areas experience higher levels of ethnic school segregation than others, the kinds of schools that might contribute most to ethnic school segregation, or the links between school admissions and the patterns of ethnic school segregation we observe at a national level. Unfortunately, this leaves policy makers with scant evidence to understand the potential drivers of segregation and therefore design policy solutions to address divisions when they arise.

This study aims to address all three gaps. By analysing school census data from 2006 – 2019, the thesis provides the most up to date measurement of ethnic school segregation and ethnic isolation in England, thereby helping to address the general lack of research in this area. Secondly, by adopting a multigroup segregation index (*H*) and pioneering the use of a clustering algorithm to define "pseudo-neighbourhoods", the study provides a

[7] There are several UK studies that have measured different conceptualisation of "segregation" as described in Section 2.2.2.

measurement of segregation within local educational markets between five major ethnic groups using a single succinct measure. This provides an appealing measure of ethnic school segregation which more closely aligns with how ethnic identities are conceived of in England today.

The study also provides three analyses to explore potential predictors of ethnic school segregation in England. Firstly, it uses meta-regression models to investigate whether there are associations between a pseudo-neighbourhood's level of ethnic school segregation and its population density, average house price, rate of FSM, proportion of religious schools, average school quality and the availability of a grammar school. Secondly, it examines whether grammar schools, faith schools or those with different OFSTED ratings contribute more toward ethnic school segregation than would otherwise be expected. Finally, it uses a bespoke extract from the NPD to describe how patterns of ethnic school segregation shifted between each stage of the 2016/7 school admissions cycle, before using seven simulations to explore whether patterns would have differed, had students been allocated to schools using alternative methods.

By providing a plurality of substantive findings, as well as developing a range of new methodological approaches, this study offers an important contribution to academic literature in this field.

The study follows a conventional structure of a PhD thesis. It begins with a review of the academic literature in the following chapter, before presenting the study's research questions. The methodology chapter then explains how the study conducted its analysis,

with the results presented in the three subsequent chapters, as described above. Finally, the results are discussed in Chapter 8 to explain how the findings have changed current academic understanding, and possible routes of future work are then presented. Finally, the thesis ends with a brief conclusion, which draws together the major themes from the research to highlight the key messages of this work.

2 Literature Review

The following chapter provides an overview of the academic literature in this field and is split into four sections, each with its own particular focus. Section 2.1 explores the theoretical underpinnings of "segregation" and explains the two ways it was conceptualised in this thesis, ethnic isolation and evenness segregation. It then highlights five decisions researchers face when measuring evenness segregation within the school system and explains how the thesis engaged with each of these issues. This offers readers a solid theoretical grounding to help them engage with the study's evaluation of past research and understand the rationale for using the new methodological approaches offered by this thesis.

Section 2.2 will then describe previous research that has measured unevenness ethnic school segregation in England and other conceptions of ethnic "separateness". It argues that studies into evenness ethnic school segregation are now dated and that the methodological approaches they took no longer offer the best approach for measuring patterns today. It also suggests that the geographic scale at which these studies calculated their results and the analysis techniques they employed to measure segregation could be improved. Finally, it argues the lack of research into what factors are predictive of ethnic school segregation constitutes an important gap in the literature. This provides the justification for carrying out this study and explains the rationale for research question 1 – 3.

Section 2.3 then presents evidence on the variables that this thesis suspects are associated with ethnic school segregation. It argues that a neighbourhood's population density and socioeconomic profile are likely to be predictive of its ethnic school segregation, though the direction of these associations remains unclear. Evidence is then provided to suggest that religious, academically selective, and "low-quality" schools are all likely to be associated with higher levels of ethnic school segregation. By the end of this section readers will understand the rationale for why each predictor variable was included in the analysis, providing a clear justification for research questions 4 and 5.

Finally, Section 2.4 justifies the study's investigation into the links between England's school admissions system and its patterns of ethnic school segregation. It initially describes the complexities of school admissions in England before explaining how a lack of data or an appropriate "counterfactual" system have limited past research in this area. Based on these discussions, it explains how a newly released extract from the National Pupil Database presents a novel opportunity to explore how patterns of ethnic school segregation shift between each stage of the yearlong school admissions cycle in England. Additionally, it argues that by using simulation analysis we can investigate how patterns of ethnic school segregation may have differed, had students been allocated to schools using a variety of alternative approaches. This explains and justifies research question 6.

2.1 Understanding "Ethnic School Segregation"

Ethnic school segregation is a complex, hotly debated term which "despite its long history in sociology, remains theoretically underdeveloped" (Fiel, 2013). Broadly speaking most agree that it centres on measuring 'the degree to which two or more groups are separated from each other' within a school system (Allen and Vignoles, 2007, p. 645).[8] However how this "separateness" is defined varies from study to study, as does the ethnic groupings that are used, the area(s) that are investigated, the tools or methods that academics use and the approach they take to describing their results. This complexity makes it sensible to begin the literature review with a description about how this thesis approaches each of these choices. This will equip readers with the necessary understanding to engage with past research in the following sub-sections, as well as evaluate the methodological approaches taken by this study. It is also worth noting that, based on the academic research this study reviewed, this is the first piece of work to discuss these various considerations in a single concise description. Consequently, the following discussion not only frames this thesis, but also constitutes a contribution to the academic literature in its own right.

2.1.1 What conceptualisation of "segregation" is being studied

Before beginning any study of segregation academics must first decide which conceptualisation of "separateness" they intend to use. Most refer to a landmark article by

[8] Segregation or "separateness" can also be measured in relation to societal groupings other than ethnicity including, but not limited to, people of different genders, ages, races, religions, wealth, or social classes. It can also be measured in spaces other than schools, such as residential neighbourhoods (Finney and Simpson, 2009, p. 9), workplaces, universities (Gamsu, Donnelly and Harris, 2018). This makes research on "segregation" a broad church.

Massey and Denton to guide this decision (1988) that proposes five theoretical approaches to understanding "separateness", of which only *ethnic isolation* and *evenness segregation* are normally applied to the school system.[9] As explained in the following sub-sections, these two dimensions are conceptually distinct and do not have a uniform relationship with one another.

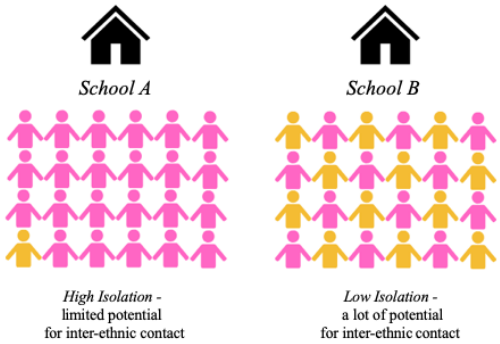
2.1.1.1 Ethnic Isolation

Ethnic isolation addresses how much potential there is for contact between students of different ethnic groups within a school system. Students who are ethnically isolated have limited space for inter-ethnic contact in their school while students with low levels of ethnic isolation have ample space for contact with peers from other ethnic groups.[10] Isolation is therefore an absolute measurement which is ultimately a student level metric, though findings are often aggregated to an ethnic group, a school, or a set of schools within a geographic area.

[9] The three not explored in this thesis are *Concentration* - The amount of physical space occupied by minority group(s); *Clustering* – How closely neighbourhoods/schools occupied by minority group(s) are to one another and *Centralisation* – How close areas/schools occupied by minority group(s) are to the centre of the city/town. As none of these dimensions are easily applicable to school segregation, or form part of this thesis, they will not be explored further.

[10] While research has found that schools with less ethnic isolation tend to have greater inter-ethnic contact, it is not always the case (Hewstone *et al.*, 2018).

Figure 1: Comparing two schools with high and low levels of ethnic isolation.[11]



Unfortunately, measures of ethnic isolation are heavily influenced by the ethnic diversity of the school system, with ethnically diverse neighbourhoods tending to have lower levels of ethnic isolation than more ethnically homogenous areas. This makes it hard for researchers to compare levels of ethnic isolation between students or schools operating in areas with different levels of ethnic diversity, or across periods of time when the ethnic profile of an area has changed substantially. Consequently this thesis, along with most other academic studies, tend to use measures of evenness segregation as the primary outcome variable, with ethnic isolation providing a supplementary perspective (Robinson, 1980; Lewis Mumford Center, 2001; Reardon et al., 2012; Fiel, 2013; Reardon and Ownes, 2014).[12]

[11] In School A most students have limited exposure to students from another ethnic group and so experience high level of "ethnic isolation", with the one orange students bring the exception. In contrast, in School B both purple and orange students can gain high levels of exposure to students from the other ethnic group, and so "ethnic isolation" would be very low. It is important to note that isolation for individual students can be asymmetric. For example, the orange student in School A has lots of exposure to other ethnic groups, in this case pink students, while pink students in the same school have very little exposure to orange students.

[12] Of the academic articles measuring "separateness" between ethnic groups within England's school system, three measured evenness segregation (Johnston, Wilson and Burgess, 2004; Gorard, Hordosy and See, 2013; Gorard, 2016a), one measured ethnic isolation (Johnston *et al.*, 2008) and one measured both isolation and segregation (Burgess, Wilson and Lupton, 2005). Lastly, some studies have measured ethnic "polarisation" which compares a school's ethnic profile to that of its surrounding area (Harris and Johnston, 2008).

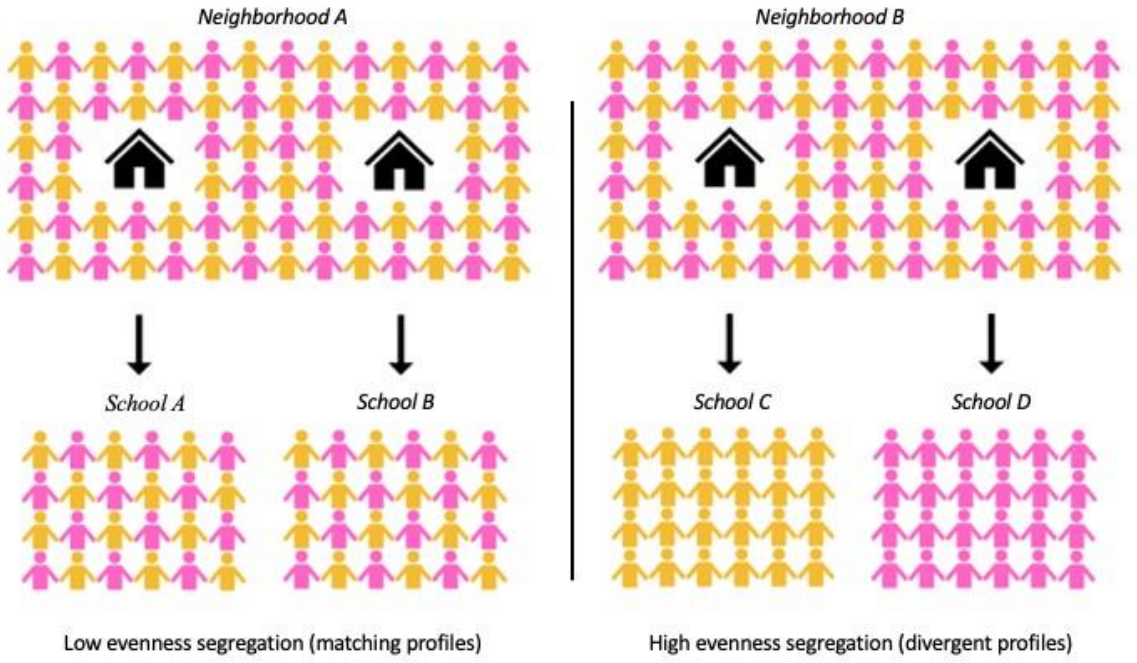
However, including a measure of ethnic isolation in an analysis still offers certain benefits. Firstly, it helps to avoid a study being wholly reliant on measures of evenness segregation, which suffer from their own methodological challenges and can often be hard to directly interpret. Additionally, given the importance of inter-group contact in relation to ethnic school segregation (Bifulco and Ladd, 2007; Croxford and Raffe, 2013) including a measure of isolation often helps the results have more direct relevance to the factors that policy makers remain most interested in.

2.1.1.2 Evenness segregation

Evenness segregation is the most common conceptualisation of "separateness" in the academic literature and is the primary outcome variable used in this thesis. Consequently, any reference to "ethnic school segregation" in this study should be interpreted in relation to *evenness* segregation. Evenness segregation measures the extent to which students from different ethnic groups are evenly spread across the schools in a predefined geographic area. As outlined in Figure 2 below, this can also be understood in relation to how closely the ethnic profile of individual schools matches the average ethnic profile of all schools within the predefined geographic area.[13] Alternatively, it can be interpreted as how closely the ethnic profile of schools within a predefined geographic area match one another.

[13] Research doesn't typically use the ethnic profile of the surrounding area for two reasons. Firstly, area statistics would typically include information on the adult population who often would have little engagement the school system and who, in the UK, have a different ethnic profile to the school aged population. Additionally, as students can attend schools outside their surrounding neighbourhood, creating the neighbourhood profile using school level statistics ensures a closed system of comparison, where each student informs the neighbourhood statistics in which then attend school, rather than the neighbourhood they live in.

Figure 2: Evenness segregation



Crucially, evenness segregation is a relational measurement between schools located within the same geographic area and therefore "segregation", as it is conceptualised by this thesis, is ascribed to a geographic area, not an individual school. Schools contribute to ethnic school segregation, based on how much the school's ethnic profile diverges from other schools in the same area, but a single school cannot be 'segregated' in and of itself. Additionally, "it is critical that [evenness] segregation is not conflated with diversity" (Catney, 2016a). A geographic area can be ethnically diverse and still have low levels of ethnic school segregation if the ethnic profile of schools matches one another, and vice versa.

This thesis decided to focus predominantly on measuring evenness segregation as it is the most commonly used conceptualisation of ethnic school segregation in the literature, with a

deep history of methodological development, particularly within the United States. This allows the results of the study to be contrasted with historical findings produced by other academics, and ensures that segregation indexes used by this study are methodologically rigorous, following decades of debate and scrutiny.[14] Secondly, the study argues that evenness segregation resembles the underlying idea that most people have in their mind when highlighting areas with high levels of ethnic school segregation, namely that the students in one school look dissimilar to the student population of a nearby school.

2.1.2 Where are you measuring segregation? – Defining the geographic area(s) of study

As evenness segregation is a comparison between the ethnic profiles of the schools within a predefined geographic area, how researchers define the geographic area(s) of study is very important. While there is no objectively "wrong" definition of this geographic area (Taylor, Gorard and Fitz, 2003; Jones *et al.*, 2015), their choice will have a substantial impact on the results.[15] For example, a study which measures ethnic school segregation within England as a single geographic area will produce drastically different results than a study which splits the nation into 3000 smaller geographic areas and measures ethnic school segregation within each of these smaller areas.

Researchers should therefore have a clear process for defining their geographic area(s) and a good justification for their chosen scale. Crucially, the findings of their research should be written in a way that continuously highlights the scale at which school segregation was

[14] See Section 4.5.1 for more details.
[15] This is known as the Modifiable Areal Unit Problem (Openshaw, 1984).

calculated and any conclusions drawn from the results should reflect the scale at which segregation was measured.

When defining their geographic area researchers must make two judgments (Reardon and O'Sullivan, 2004). Firstly, they must decide how large their geographic area(s) should be, a dilemma known as the scale problem. Secondly, they must decide where the exact boundaries of these area(s) should be drawn, a dilemma known as the zoning problem. Drawing on the limited number of previous articles on this topic (Gibson and Asthana, 2000a; Taylor, Gorard and Fitz, 2003; Harris and Johnston, 2008; Burgess *et al.*, 2015) this study suggests three desirable properties for defining geographic area(s) when measuring ethnic school segregation. These have heavily informed the methodology of this study and are regularly referred to throughout the thesis. The three properties are:

1. **Geographic areas should broadly reflect local educational markets so that schools which could feasibly educate the same pool of students are located in the same geographic area.**

   This is the most common scale at which researchers in England aim to operate (Gibson and Asthana, 2000a; Taylor, Gorard and Fitz, 2003; Harris and Johnston, 2008; Burgess *et al.*, 2015). It is conceptually clear and helps to avoid comparing schools which are spread across such a large geographic area that differences in their ethnic profiles primarily reflect ethnic residential patterns. Similarly, it avoids grouping schools into such small geographic areas that measures of segregation fail to detect dynamics occurring within local educational markets.

This thesis proposes that local education markets in England should be conceptualised as containing approximately six schools. This reflects the maximum number of schools parents can list on their child's school application form (HM Government, 2018b, p. 4) and broadly aligns with previous research that has investigated how local schools compete with one another. A study by Taylor, for example, estimated that "the majority of [secondary] schools competed with between five and nine other schools" (2001, p. 204) which chimes with a more recent study by Burgess *et al* that estimated 75% of students in England had six schools or less that they could realistically gain admission to.[16]

Of course, not all school neighbourhoods will contain exactly six schools. Those in densely populated urban areas often will contain more than six schools, while many in rural environments will contain fewer (Burgess *et al.*, 2011, p. 535). This is therefore a reasonable approximation to researchers working in this field, rather than a definitive answer to what scale ethnic school segregation should always be measured at.[17]

It is also important to recognise that it is impossible to define geographic areas that *perfectly* reflect a "local educational market" in England. While home-school distance is an important part of most schools' admissions policies (Pennell, West and Hind, 2006), students can apply to any school in the country, which means there are no discrete

[16] This was calculated using students' home post code data in the National Pupil Database. For each student the authors evaluated the distance between their home and other local schools. If the distance was less than the home school distance of 80% of the students who currently attend the school then it was added to the students "High-probability Feasible Choice Set" (Burgess *et al.*, 2015).

[17] Academic research in this field would stand to benefit from a more thorough debate over the ideal size in which to measure ethnic school segregation.

school markets in England. Additionally, while the majority of schools tend to compete with schools in their immediate vicinity, religious and grammar schools often compete with equivalent schools spread across a wider geographic area (Taylor, 2001; Johnston, Harris, *et al.*, 2016). This "weaken[s] the geographical criterion underpinning admissions" (Harris, Johnson and Burgess, 2016, p. 1682) and makes defining discrete school markets that bit harder. This principle therefore offers a guide to research in this field, rather than a straightforward rule to implement.

2. **When measuring ethnic school segregation within multiple geographic areas, each area should contain roughly the same number of schools.**

Ideally, if researchers split a large geographic area (e.g., a nation) into a series of smaller geographic areas (e.g., neighbourhoods) there should be minimal variation in the number of schools contained within each of the smaller geographic areas. As the number of schools within an geographic area influences measures of segregation, having a similar number of schools within each area ensures a fairer comparison (Harris, 2011, p. 484).[18] This also helps readers have a clear understanding of the scale at which they should interpret findings. As the size of educational markets vary across different parts of England, meeting this principle often involves a trade-off with meeting the first principle.

[18] See Section A6 for an analysis of the relationship between the size of pseudo-neighbourhoods and the level of ethnic school segregation (*H*)

3. **Ideally areas should be stable so that trends in ethnic school segregation can be measured across time.**

Lastly, geographic area(s) used by researchers would ideally remain stable throughout time so that researchers can measure changes in the patterns of ethnic school segregation. Of course, the opening or closing of new schools, any changes in local housing, or migration in and out of an area means that using stable geographic boundaries can sometimes involve a trade off with meeting the first principle.

Previous studies in England have typically used government boundaries to define the geographic area(s) for measuring school segregation, primarily due to the convenience it offers.[19]

[19] This is both in relation to measuring socioeconomic school segregation (Goldstein and Noden, 2003; Allen, 2007; Leckie *et al.*, 2012; Gorard, 2016b) and ethnic school segregation (Johnston, Wilson and Burgess, 2004; Johnston *et al.*, 2006, 2008)

Table 1: Describing the distribution of schools across different sets of government boundaries (2019)

Type of boundary	**Number of areas in dataset (& in England)**[20]	**Average number of schools in each government boundary (range)**	
		Primary	**Secondary**
Local Authorities (LA)	152 (152)	110.7 (1-480)	22.9 (1-98)
Middle Super Output Areas (MSOA)	6575 (6791)	1.2 (1-4)	1.1 (1-8)
Lower Super Output Areas (LSOA)	15512 (32,844)	2.6 (1-10)	1.2 (1-8)
Parliamentary Constituencies	534 (534)	31.3 (4-67)	6.5 (2-18)
Electoral Wards	6415 (9130)	2.6 (1-44)	1.3 (1-14)

Unfortunately, none of the available government boundaries provide a desirable set of geographic areas for measuring ethnic school segregation at a local level. All are inappropriately sized, either containing too many schools to be considered a proxy for local educational markets (LA and Parliamentary constituencies) or too few (Middle Super Output Areas, Lower Super Output Areas and electoral wards).[21] Additionally, as none of the

[20] There are differences between the number MSOAs, LSOAs and Electoral wards in the school census dataset and England, as a whole, as these areas do not consistently contain a school. In contrast, all LAs and Parliamentary Constituencies contain at least one school.

[21] The one exception to this is parliamentary constituencies in relation to secondary schools where the average of six schools makes them an appropriate size for an educational market.

boundaries are drawn in reference to schools, none are likely to reflect the choices available to local parents (Sandford, 2018). Furthermore, all of the boundaries exhibit considerable variation in the number of schools within each area. This introduces a hidden bias into measures of ethnic school segregation with larger geographic areas tending to experience higher levels of ethnic school segregation than smaller ones.[22] It also means that it is hard to have a clear understanding of the scale at which results are measured. Finally, "the boundaries of localities that are used for population statistics are reviewed and change regularly" (Simpson and Finney, 2006, p. 418) which means that government boundaries do not offer a stable set of areas for measuring school segregation across time.

Unfortunately, Local Authorities, the most commonly used set of boundaries, offers one of the worst options (Gibson and Asthana, 2000b, p. 139). While the historical role that LAs played in local schools systems may have meant that using LAs to measure segregation could have once helped research reflect local school admission policies (Taylor, Gorard and Fitz, 2003, p. 53) or translate findings into policy outputs by aligning neighbourhoods to governing bodies (Burgess, Wilson and Lupton, 2005, p. 1033) the weakening of LA control over local schooling means these arguments no longer carry much weight. Students can now apply to any school in England (Department for Education, 2018c) and 37% of primary and 78% of secondary schools are now academies or free schools, which typically can determine their own admissions policies (Office for National Statistics, 2021). This means

[22] As geographic areas become larger the ethnic profile of residents is likely to be less uniform due to ethnic groups clustering within the same residential area. This unevenness in residential patterns (ethnic residential segregation) then translates into divergent ethnic profiles between schools (ethnic school segregation). In contrast, smaller geographic areas are likely to have a more even spread of different ethnic groups, in terms of residents, which translates into less divergent ethnic profiles between different schools (less ethnic school segregation).

that LAs now represent a series of large, fairly arbitrary geographic areas (Harris and Johnston, 2020, p. p27) where findings are more likely to reflect differences in residential patterns than the dynamics occurring within local school eco-systems (Gibson and Asthana, 2000b, p. 139).[23] Additionally, the substantial variation in the number of schools contained within each Local Authority make them an inconsistent set of geographic areas (Harris and Johnston, 2020, p. 26). This introduces a substantial bias into any subsequent measure of segregation (Harris, Charlton and Brunsdon, 2018, p. 66), and makes it hard for readers to understand the scale at which results are calculated.[24]

The challenge of British research to define appropriate geographic areas for measuring school segregation is one of the most serious methodological problems facing this field and yet it has received limited attention in the literature (Harris, 2013). As discussed in later sections, the study's new approach to defining geographic areas to measure school segregation is therefore one of the most important contributions that this thesis makes to this field.

2.1.3 Who is being studied? – Defining the ethnic groups

Another factor underpinning the study of ethnic school segregation is how to define ethnic groups. As most studies use government data, researchers in England normally begin with the 18 official ethnic groups listed in the School Census dataset, before conducting some

[23] This was highlighted by Harris and Johnston (Harris and Johnston, 2008, p. 60).
[24] Larger LAs are more likely to report higher levels of ethnic school segregation than smaller LAs.

degree of aggregation.[25] While the groupings that a researcher uses is ultimately a subjective choice, it is a sufficiently important decision that researchers must explain and justify their decision. To contribute to this important step, this thesis proposes the following principles as a guide to approaching this aggregation process.

1. **Ethnic groupings should reflect how identities are commonly conceptualised within the area(s) of study**.

By meeting this principle, measures of segregation capture the level of separation between groups of people who are perceived to be from "different" ethnic groups. Crucially, this does not mean that researchers should inherently use all 18 ethnic groups. Grouping some ethnic groups into larger "meta-groups" may reflect how identities are perceived by local residents. Meeting this principle in practice, however, remains challenging as ethnic identities are "nebulous" (Harris *et al.*, 2013, p. 2283) and will be viewed differently by different members of society (Kaszycka and Strzałko, 2019). It becomes particularly difficult when measuring segregation across a large geographic area, as conceptions of ethnicity are more likely to vary across a larger region.

This thesis argues that if researchers are measuring ethnic school segregation at a national level, they should not use a binary categorisation of "ethnic minority" and "White British". There is now sufficient ethnic diversity, particularly within the school

[25] The ethnic groups are White British, Irish, Roma Gypsy or Irish Traveler, Any other White background, White and Black Caribbean, White and Black African, White and Asian, Any other mixed ethnic background, Indian Pakistani, Bangladeshi, Chinese, Any other Asian background, Black African, Black Caribbean, Any other Black background, Any other ethnic group.

age population, that it is both mathematically feasible and conceptually important that a multi-ethnic measurement is required. This also chimes with a growing feeling that a binary split between "White British" and "everyone else" is no longer appropriate.

2. Researchers should try and avoid using excessively large number of ethnic groups

While it may be tempting to use as many ethnic groups as possible, researchers should try and limit the number of ethnic groups to provide a clear narrative. As a methodologically complex field, which already requires readers to conceptualise mathematical concepts that most will be unfamiliar with, this is particularly important.

Additionally, as the number of ethnic groups increases, researchers risk creating small ethnic groups that can lead to measures of segregation becoming artificially high. As an ethnic group becomes increasingly small, it becomes increasingly unlikely that students from the group will be spread evenly across schools, even if they were randomly assigned to schools in the area.[26] By creating ethnic groupings that include a reasonable number of students in each grouping, researchers ensure this "upward bias" or "random segregation" remains minimal (Cortese, Falk and Cohen, 1976; Rathelot, 2012).

While researchers can technically adjust for random segregation by running simulation analyses or ammending the formulee of segregation indexes (Allen, Burgess and

[26] In fact, when there are less students from the group than there are schools in the geographic area, it becomes impossible for there to be an even distribution of ethnic groups between schools. For example, in a neighbourhood containing six schools and only four Asian students, it is impossible for all schools to have matching ethnic profiles.

Windmeijer, 2009) this introduces another layer of complexity to an already intricate methodological process. Therefore, if possible, researcher should try to avoid creating ethnic groupings which contain only a small number of students. Again, this may involve some trade off with the first principle.

It is also important to highlight that any analysis that uses government data on ethnic identities are only as solid as the categories upon which they are based. To this end, it is important for readers to be aware that the 18 ethnic categories used in the UK are subjective, man-made and structured in a fairly unsystematic way. Some ethnic groups include people with a certain skin colour from an entire continent (Black Africans) or region (Black Caribbean), while others include people with any skin tone from a given country (India, Pakistan, Bangladesh, China). Other ethnic groups include disparate groups of people who do not fit neatly into any ethnic group. For example, the "White Other" group, which is the second largest "ethnic" group after the White British majority, includes anyone who has "white" skin who is not "British", "Asian", "Irish", "Roma Gypsy" or "Irish Traveller". This catch-all group therefore includes "white" Australasians, Africans, Europeans, South and North Americans.

These groupings are also "far from neutral" (Phillips, 2007, p. 1153) and often are heavily influenced by the perspectives of the White British majority. It is worth highlighting, for example, that there are multiple ethnic groups for darker skinned people and "white" communities that have often been marginalised by the White British majority (Irish, Irish Traveller, Roma Gypsy), while one large nebulous ethnic category group contains all other "white" skinned people.

Additionally, while measuring ethnic school segregation requires the use of mutually exclusive groups, it "tends to transform nebulous categories into firmly bounded ones, and erase differences within them" (Phillips, 2007, p. 1153). Indeed, there can still be significant "diversity" within a single "ethnic group", particularly given the current typology of ethnic groupings in the UK (e.g., "Black Africans").

While these issues do not invalidate the use of official data to measure ethnic school segregation, readers need to nonetheless be aware of these points when interpreting the findings of this study.

2.1.4 How will you measure segregation? – Selecting the right approach to measurement

Researchers must also choose which technique they wish to use to measure ethnic school segregation. Luckily, when measuring eveness segregation all methods ultimately do the same thing; they calculate the extent to which the ethnic profiles of schools located in the same pre-defined geographic area diverge from one another. This involes comparing the proportion of each ethnic group in each school against the percentage of that ethnic group within the geographic area.

When working at a micro-scale, for example when studying a single neighbourhood, maps and other graphic approaches offer the clearest way of describing the small number of data points (James and Taeuber, 1985, p. 26). Readers can visually compare the proportions of each ethnic group across a limited number of schools to draw fairly accurate conclusions

about the extent to which the schools' ethnic profiles differ from one another, particularly when combined with the author's narrative. When schools in the same geographic area have very similar ethnic profiles (low segregation) or very distinct ethnic profiles (high segregation) this comparison becomes particularly straightforward. A good example of this comes from a recent study by Birtel *et al*, who used stacked bar charts to describe the level of ethnic school segregation between White British, Asian British and all other ethnic groups in a British town (2020, p. 163).

However, when working at a macro scale the number of comparisons becomes too large for a graphic approach to provide a feasible tool for measuring ethnic school segregation. This study, for example, compares segregation within 2227 pseudo-neighbourhoods, which would require researchers to describe 66,810 data points to readers.[27] Consequently, researchers normally aggregate these datapoints into a smaller set of numbers to create a clearer narrative of patterns.

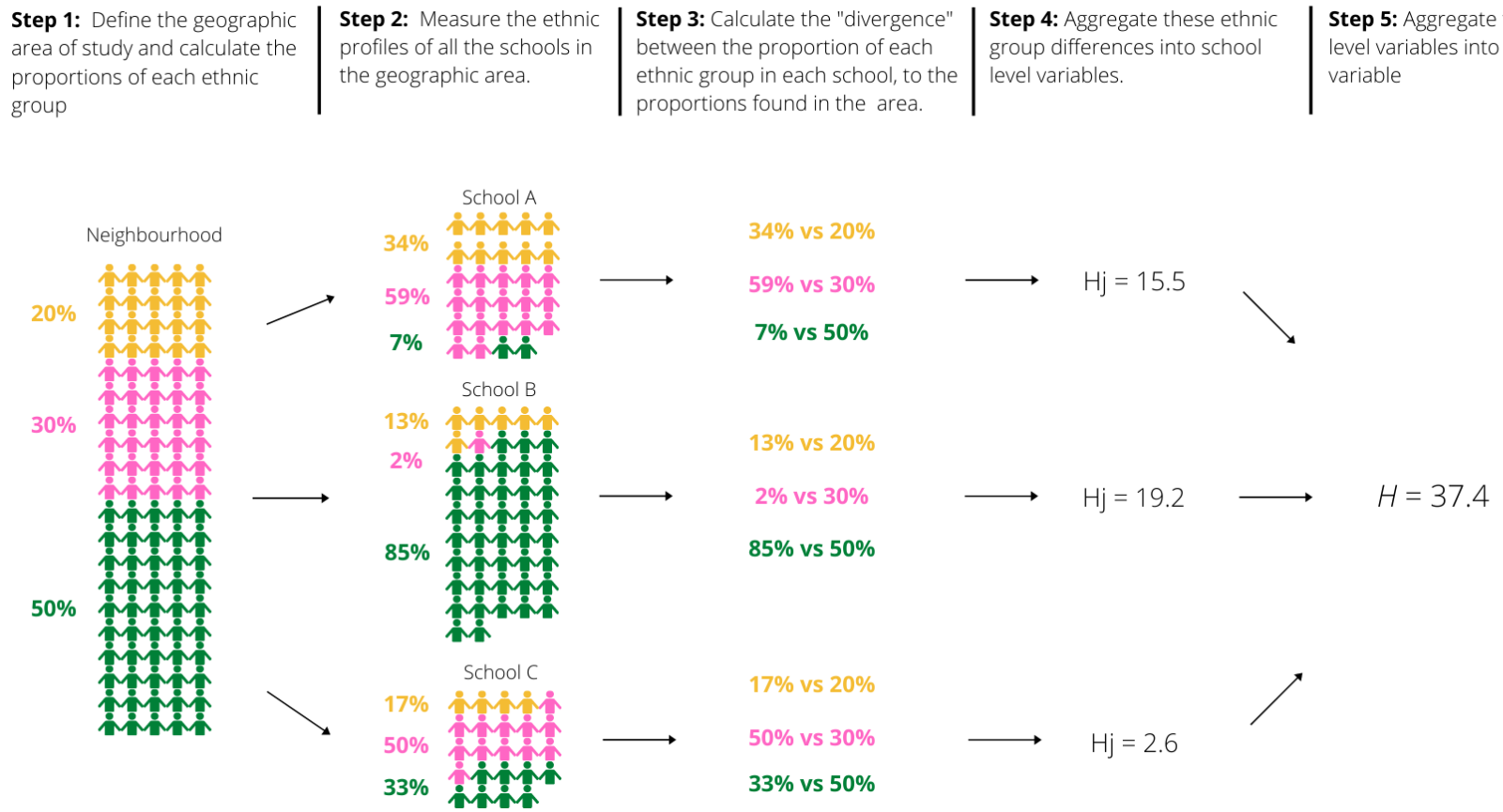
The most common tool used for this aggregation process is a segregation index, which reduces "huge data arrays into [a] simpler, more readily understandable number" (Grannis, 2002).[28] While the methodology chapter evaluates the mathematical properties of different segregation indexes in detail, Figure 3 on the following page, unpacks this process into five steps as it relates to most multi-group indexes. Firstly, researchers define the geographic area of analysis (step 1) and calculate the ethnic profile of each school (step 2). They then

[27] There are 30 data points per pseudo-neighbourhood, 2227 x 30 = 66,810. See Section 4.2.1for more detail.
[28] Some academics have advocated for using regression modelling to describe the distribution of these underlying datapoints. For a discussion of this approach, and an explanation as to why it was not adopted by this thesis, please refer to Section 11.2 in the appendix.

measure the extent to which the proportion of each ethnic group in each school diverges from the proportion of the same ethnic group within the area (step 3). The exact formula for measuring this divergence varies between different multi-groups indexes but are collectively known as an index's disproportionately function (Reardon and Firebaugh, 2002, p. 41)5. The following two steps then aggregate these divergences, first to the school level (Step 4), and then to the geographic area (step 5). This leaves the researcher with a single statistic to describe how evenly distributed all ethnic groups are, on average, across all schools within the geographic area in question.

As most research will split a large geographic area (i.e., a nation) into a plurality of smaller geographic areas (i.e., pseudo-neighbourhoods) most researchers will still have a large number of data points to describe to readers at this point in their calculations. Studies must then employ a range of summary statistics (mean, median, range, inter quartile range), plots (density plots, histograms, box and whisker plots) and maps to explain the numerical and geographic variation in these values.

Figure 3: Explaining how evenness segregation indexes work[29]



[29] This diagram focuses on the conceptual steps taken by most multi-group segregation indexes in the context of measuring ethnic school segregation (Reardon and Firebaugh, 2002, p. 41). Readers should note however, that binary segregation indexes, such as the commonly used D index, follow slightly different step, as described in Section 11.11. For more information on the H index, and the formulas used in step 4 and 5, see Section 4.5.1.

2.1.5 How will you describe the results? – the importance of clarity and considerate language

The final consideration for researchers is how to describe the results of their analysis in a clear and considered manner. While this may not seem strictly related to the *numeric* measurement of ethnic school segregation, it is a crucial challenge for researchers if they want readers to gain a balanced understanding of ethnic school segregation, where their results are framed within the context of wider societal dynamics.

The first challenge is to provide a clear description of patterns so that readers can follow the broad narrative of the results. As explained in the previous section, measuring school segregation is complex and ultimately requires readers to consider the "distribution" of ethnic groups across a set of schools. There will be few readers that have an immediate reference point for this type of mathematical thinking and so most people are likely to find it hard to initially interpret the results of a segregation index. Therefore, academics need to think especially hard about how they can make their findings clear and accessible, particularly when engaging with a broader range of academics, policy makers and members of the public.

Additionally, as ethnic school segregation is a sensitive topic, researchers must think about how to describe their results in a thoughtful and considered manner (Phillips, 2007, p. 1154). One change that would help in this respect would be for researchers to use mathematically neutral vocabulary to describe patterns of segregation rather than emotive

adjectives.[30] Previous quantitative research in the UK has used segregation indexes to label areas as being "dominated by ethnic minority pupils", which assumes a struggle between groups, or as having "extreme Muslim concentrations", which is a subjective term that would be better described using percentages.[31] Unfortunately, these types of phrasings can have serious consequences, even when used within academic contexts without any intended malice. For example, various UK newspapers ran headlines about worsening ethnic divisions in Britain after an academic gave a presentation on "Muslim ghettos" in Britain, in an Australian academic conference. The story gained substantial public and political attention in the UK, despite the substantial body of academic research that showed residential segregation was consistently falling across England (Peach, 2009; Johnston, Manley and Jones, 2016).

Another change advocated by this thesis would be for research to describe patterns of ethnic school segregation in relation to the geographic areas they are studying rather than in relation to ethnic groups. As described in the later sections, most research in England has

[30] There are plenty of examples where the same concept is called different things depending on whether it is applied to a majority or minority group. A good example is how academics and policy makers distinguish between "voluntary self-segregation", which relates to minority groups, and "white flight" which relates to White British/American population moving out of certain neighbourhoods. As previous academics have pointed out, both concepts are describing a fairly well-established human tendency called social homophily, that most people tend to prefer living and attending schools where many people look and sound like them (Simpson and Finney, 2006, p. 409/410). If we, as a society, can acknowledge that these types of behaviours are common across people from different ethnic backgrounds, then hopefully conversations about how we build a successful multi-ethnic society can move beyond comparative mistrust and blame towards building some type of mutual understanding.

[31] These quotes come from the following two articles (Hamnett, 2012; Johnston, Manley and Jones, 2016). As "there is no objective definition of 'intense' segregation or 'acceptable' segregation" (Phillips, 2007, p. 1147) academics should be careful when assigning value laden adjective to describe patterns of segregation. If academics do have a position on what the ideal level of segregation is, then they should make this position explicit from the beginning and be clear about where the "objective" measure of segregation ends, and their subjective interpretation of results start. Even in this instance, readers should be made aware that there is an ongoing discussion on what this "ideal" should be, and that the authors opinion is one side of this debate, rather than a description of an uncontested consensus viewpoint.

used binary measures of ethnic school segregation to assess how separate White British and ethnic minority students are from one another in schools and these trends are then described in terms of how segregated *ethnic minorities* are in English schools. As segregation is a relational measurement between ethnic groups, ethnic minority students are only as segregated from White British students as White British students are segregated from ethnic minorities. In other words, the same figures could legitimately describe how segregated *White British* students are in English schools. Describing trends in relation to one group places that group under undue scrutiny and implies that segregation is only relevant to them. In contrast, by describing trends for the geographic area of study, researchers present findings in the unit which segregation is measured, which should encourage readers to appreciate that patterns of ethnic school segregation reflect the collective actions of all groups.[32] Hopefully this would also lead towards a greater sense of shared responsibility.

Finally, researchers should also provide contextual information to support readers in forming nuanced interpretations. Readers need to be aware, for example, that certain ethnic groups face far higher levels of hostility than most White British families, or indeed other ethnic minority groups. As mentioned in the introduction, British Muslims suffer from particularly high levels of distrust and hate (Phillips, 2006, p. 26, 2015, p. 63; Miah, 2016), as do Traveller communities, who face some of the "last bastions of [perceived] acceptable racism in the UK" (Phillips, 2007, p. 1154). Secondly, academics should explain that many patterns of segregation are long running and heavily influenced by historical events. For example, Local Authorities in England forcibly bussed ethnic minority students to "white

[32] Using a multi-group measure of segregation, such as the H index, which captures the combined level of segregation for all ethnic groups in an area, would likely help in this regard.

suburban schools" as recently as 1986 (Esteves, 2018), while Birmingham City Council had an explicit rule that "no more than one in six properties in a block could be allocated to a black person' until 1975 (Phillips and Harrison, 2010, p. 224).[33] While such policies have since been retracted, the impacts are unlikely to have entirely faded away. Thirdly, readers should know that the majority of children from ethnic minority groups are likely to have been born in the UK, as where many of their parents. This means that, even under a restrictive notion of "Britishness" based on a child's place of birth, White British children have no more claim to being "British" than most of their peers from other ethnic backgrounds. Lastly, many of the dynamics underpinning ethnic school segregation are likely to be influenced by factors that are not explicitly related to ethnic identity (Carter and Fenton, 2010). For example, in a review of the academic literature, Finney and Simpson argue that "the shifting ethnic composition of British cities can… be understood through demographics of in situ natural growth and suburbanisation, without the need for any racially coloured interpretations of migration" (Simpson and Finney, 2009). Similarly, when applying for schools, parents are likely to be more interested in the academic quality of their child's school, and its distance from their home (Burgess, Greaves and Vignoles, 2019) rather than the school's ethnic profile, irrespective of their ethnic background. Remembering that ethnic school segregation need not be an entirely "ethnic" phenomena should encourage an appreciation that many of these trends are intertwined with other important societal and socio-economic dynamics (Harris and Johnston, 2020, p. 175).

[33] See also (Qureshi and Janmaat, 2014; Bebber, 2015)

Summary

Section 2.1 has explained the two different ways in which this thesis will conceptualise "separateness" in the school system and has highlighted the important differences between each concept. It then explored five important factors that researchers need to consider when measuring ethnic school segregation and explained the basis on which this study approached each choice. Hopefully this has provided readers with the necessary theoretical understanding to engage with this study's evaluation of previous research, as described in the following section of this thesis, and the new methodological approaches it adopts.

2.2 Research measuring ethnic school segregation in England

The next section of Chapter 2 describes previous studies that have measured evenness ethnic school segregation in England and argues that three gaps in the literature justify conducting this research. Firstly, it suggests that this body of academic literature is fairly small and often dated. While several studies have measured different types of ethnic "separateness" within English schools, only a limited number have measured ethnic school segregation in terms of evenness across England and the most recent analysis to do so in any detail used data from 2001. Secondly, past research has yet to measure ethnic school segregation within local educational markets or provide a multi-group measure of patterns. Lastly, academic literature provides limited insights into the predictors of ethnic school segregation which leaves policy makers and academics with limited empirical evidence to design policy solutions.

2.2.1 Previous academic studies that measure evenness ethnic school segregation in England

The two most recent national studies that measured evenness ethnic school segregation were released in 2014 and 2016 (Gorard, Hordosy and See, 2013; Gorard, 2016b). Using school census data, the articles both measured ethnic school segregation between 1997 – 2012 in relation to four different student characteristics, one of which was a binary categorisation of students' ethnicity ("White" vs "Non-White").[34] Using the Gorard Index, they found that ethnic school segregation had fallen substantially within England, when defined as a single geographic area. Ethnic school segregation for secondary schools fell by nearly a third, from a G value of 0.6 to just above 0.4, while primary schools experienced a smaller reduction from just under 0.3 in 1997 to around 0.25 by 2012.[35] Substantively these figures suggest that roughly 60% of ethnic minority students in secondary schools, and 30% of ethnic minority pupils in primary schools, would have had to change schools in 1997 for all schools in England to have matching ethnic profiles. However, by 2012, the same figures were approximately 40% for secondary schools and 25% for primary schools, suggesting that ethnic minority and White British students had become more evenly spread across the nation's schools over this period.[36] The findings also show that when measured within England as a whole, secondary schools typically experienced higher levels of ethnic school segregation than primary schools.

[34] The other three student characteristics were Free School Meal Eligibility (FSM), whether the student spoke English as a second language (ESL) and whether the student had a special educational need (SEN).
[35] These figures were gained by reading points of a figure in their article and are therefore not very precise.
[36] The latter article (Gorard, 2016b) built upon these results by exploring possible "determinates" of school segregation within 34 LAs, which provides a useful starting point for exploring potential predictors of ethnic school segregation, as discussed later in Section 2.3.

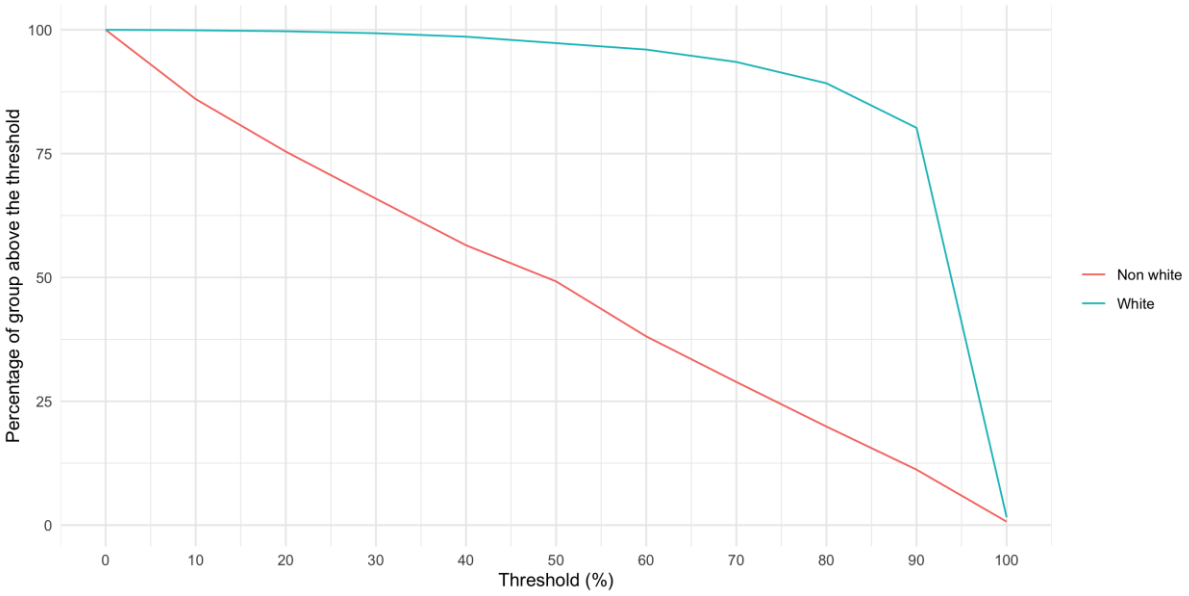
While these studies are interesting, a number of factors necessitate a new analysis. Both articles calculated segregation within England as a single geographic area, which means their findings were likely to be predominantly driven by residential patterns within England rather than dynamics within local educational markets. While the latter article (Gorard, 2016b) also measured ethnic school segregation within 34 Local Authorities, ethnic school segregation was not the focus of the article and so the numerical or geographical distribution of these results were not described in detail. A need for a more comprehensive description of how segregation varies across England therefore remains.

Additionally, both articles calculated their results using a binary segregation index with two ethnic groupings, “White” and “Non-White”, which does not capture the level of segregation between different ethnic minority groups. Given that ethnic minority groups in England have grown to constitute nearly a third of the student population, it is now mathematically feasible and conceptually important to use measures of segregation that also assess the level of separation between different ethnic minority groups. This also reflects changes in public discourse, which increasingly sees using a single heterogenous category of “Non-White”, or “ethnic minority” as inappropriate.

Another study of evenness ethnic school segregation was produced by Johnston, Wilson and Burgess (2004), who measured segregation within 3060 secondary schools in England across three different sets of geographic areas. They used a novel graphic approach called a “concentration profile”, as shown in Figure 4 below, which plots a series of threshold values between 0 and 100 to measure [the] cumulative percentage of an ethnic group (x%) who

attend a school where their ethnic group is equal to, or larger than, a given percentage threshold (y%).[37]

Figure 4: Concentration profiles for white and non-white students across 3060 English Secondary Schools in 2001 *(Johnston, Wilson and Burgess, 2004)*[38]



When measuring segregation within England as a whole they found that "whites [students were] much more segregated in all-white schools than non-whites [students] are into all-non-white schools" (2004, p. 244).[39] After delineating between ethnic minority groups, they found students from Asian backgrounds, particularly those of Pakistani and Bangladeshi heritage, where more segregated than students from other ethnic minority groups.

[37] This is conceptually like a Lorenz curve which is typically used to measure income or wealth inequality.
[38] This graph shows, for example that 96 percent of white students in 2001 attended schools at least 60 percent of their peers were white, while only 38 percent of non-white students were in schools where 60 percent or more of their peers were non-white (Johnston, Wilson and Burgess, 2004, p. 244).
[39] This was true even after accounting for the level of segregation that would occur if students were randomly allocated to schools across the whole country.

However, due to the scale at which these findings were measured, these national level findings primarily reflected ethnic residential patterns for young people across England rather than dynamics within England's school system.

To try and remedy this issue, the authors repeated their analysis within four smaller geographic areas that were created by grouping LAs together based on the percentage of the non-white students in each LA. [40] They then produced three separate concentration profiles for White, Black and Asian students in each of what this thesis would call a "meta-LA".[41] Unfortunately, while this may have historically been a useful approach to take when ethnic groups were clustered in a limited number of LAs, the diversification of England means that neighbourhoods increasingly sit along a spectrum of diversity. This means that categorizing areas based on the percentage of the non-white students no longer matches the reality of ethnic diversity in England today. Additionally, the meta-LAs varied substantially in size, with "30% non-white" meta-LA containing 118 LAs and the "75% non-white" containing only 16, which as explained in Section 2.1.2, introduces bias into measures of segregation and makes it hard understand the scale at which findings are calculated.

The fourth article to measure evenness ethnic school segregation across England was produced by Burgess, Wilson, and Lupton in 2005. Using the school census dataset from 2001, the authors measured segregation for secondary schools in relation to eight different ethnic groupings using Local Authorities as their geographic areas of study. The authors

[40] The cut-off points were 75%, 50% and 30% non-white.
[41] The authors also produced concentration profiles for five LAs which had a particularly large proportion of students from a given ethnic minority groups.

measured ethnic school segregation and ethnic isolation using a series of binary indexes, each which measured segregation or isolation for a given ethnic group in relation to students from all other ethnic backgrounds. The results of this analysis are outlined in the table below.[42]

Figure 5: Ethnic school segregation in England in 2001 (Burgess, Wilson and Lupton, 2005).

Ethnic Group	**Average *D* Value for Local Authorities**	**Average *I* value for Local Authorities**
Bangladeshi	0.630	0.102
Pakistani	0.610	0.099
Black Caribbean	0.530	0.062
Black other	0.505	0.048
Black African	0.497	0.044
Indian	0.481	0.054
Other ethnic minority	0.416	0.057
Chinese	0.396	0.026

The study had several strengths. Firstly, they maintained a consistent geographic unit of analysis throughout the study, which makes it clear to readers the scale at which they should interpret results. They also provided an explanation for their choice of geographic

[42] D is a measure of evenness segregation and is bounded between 0 and 1. Each figure listed in the table represents the proportion of students from a given ethnic group who would have to change schools for there to be an even spread of their ethnic group across schools in their area of interest. For a more detailed discussion on how D is calculated see Section 4.5.1.2. The I index is a measure of ethnic isolation or exposure, where values "ranges from 0 to 1, with a value of 0.3 being considered high" (Burgess, Wilson and Lupton, 2005, p. 1034). Unfortunately, these figures do not have a substantive interpretation.

area, which recognises the importance of this decision. By including a measure for ethnic isolation/exposure they also provide a broader evaluation that can help readers to understand the different dimensions of ethnic "separateness" in schools.

However there remains a need for a new analysis of trends in England. Firstly, the article's findings are based on data collected in 2001 which means these findings are unlikely to reflect the current patterns of ethnic school segregation nearly two decades later. Secondly, by calculating a series of binary segregation indexes, each which measures ethnic school segregation between students in each ethnic group against all other students, readers are provided with eight overlapping narratives of ethnic school segregation. To comprehend the overall level of segregation, readers must weave these various narratives together, which can be challenging. By adopting a multi-group measure, like the *H* index used in this study, academics can provide a single concise measure of ethnic school segregation which captures segregation between all ethnic groups within a geographic area. This also encourages readers to view ethnic school segregation as a relational measurement between *all* ethnic groups within a geographic area of study, rather than a pattern that relates to a single ethnic group. As discussed in Section 2.1.5, ethnic group X can only be as "segregated" from ethnic group Y, as ethnic group Y is from X, describing patterns of segregation in relation to one ethnic group implicitly hides the equal "segregation" being experienced by students in the other ethnic group(s).[43] Finally, the results provide no insight into the level of ethnic

[43] Readers should be aware that multi-group indexes come with certain downsides. Unfortunately, none of these indexes have particularly intuitive interpretations, which means academics must describe their results in relational, rather than absolute terms. This can make it hard for readers to conceptualise what the findings mean. Second, by collating segregation for all ethnic groups in a single measure, certain ethnic groups could be comparatively more segregated relative to all other ethnic groups, though as the H index can be decomposed, the authors could choose to present findings in relation to certain ethnic groups if they so wished.

school segregation occurring within the primary school system, which judging by Gorard's results (Gorard, Hordosy and See, 2013; Gorard, 2016b), are likely to have distinct patterns from those in the secondary school sector.

More broadly, this body of research is yet to offer a detailed investigation into why certain areas experience higher levels of evenness ethnic school segregation than others. While two of the articles provide some insights (Burgess, Wilson and Lupton, 2005; Gorard, 2016b), as described in Section 2.3, there remains a substantial gap in the literature for a more detailed and concerted investigation into the predictors of ethnic school segregation. This problem is not confined to British literature. Even in the United States, which has a far deeper literature on racial school segregation, research to explain patterns of school segregation remain sparse. As Fiel, a prominent American academic, writes:

> *Why are schools segregated? For a question so fundamental to a topic, so enduring in social science research, answers are remarkably superficial.*
>
> (Fiel, 2015)

By providing a rigorous investigation in different patterns of ethnic school segregation, within local educational markets, using modern measures of school segregation, this thesis aims to make an important contribution to both domestic and international research alike.

2.2.2 Other studies of ethnic "separateness" within the English school system

While the aforementioned studies are the only studies to investigate evenness segregation across England there are other pieces of research worth highlighting. Firstly, an article by

Leckie and Goldstein measured evenness ethnic school segregation within state-funded secondary schools in London between 2001- 2011 (Leckie and Goldstein, 2015).[44] Using a multi-level model approach, which is described in more detail in Section 11.2 of the appendix, the authors found that while "there [was often] substantial variation in the proportions of white, black and Asian students between schools" across the city (p421) but "patterns of segregation… remained largely stable" (p423) during the time period. They also found that "Asian and white students appear[ed] most segregated [from one another], followed by Asian and black students, and lastly black and white students" (p419).

Several other studies have measured other conceptualisations of ethnic "separateness" in English schools which, while distinct from evenness ethnic school segregation, provide a range of interesting findings that contextualize the work of this thesis. A book by Harris and Johnston's, released halfway through the completion of this thesis in 2020, is a prime example. Across seven chapters the authors measure various different conceptualizations of ethnic "separateness" within English schools between 2011 – 2017 to offer an insightful picture of trends within the English school system (Harris and Johnston, 2020). In Chapter 2 they mapped the ethnic composition of the school aged population within Local Authorities to show readers "a picture of increasingly [ethnic] diversification …across large parts of the country" (p61). In Chapter 3 they found that the ethnic composition of Local Authorities had typically converged with one another during this period, particularly in relation to the primary school aged population, and that the ethnic diversification of the school aged population had meant that "the potential exposure of the eight minority groups to the

[44] See Section 11.2 for a discussion on the use of multi-level models to measure ethnic school segregation.

White British ha[d] declined [at a local authority level while], that of the White British to other ethnic groups… increased by about one third" (p73- 4). In Chapter 4 they then measured the level of ethnic "concentration" within the school system by assigning each school to one of five groups based on the school's percentage of White British students. Using this method they showed a 12.8 percentage point decrease in the number of schools where White British students constituted more than 90% of the school population between 2011 – 2017, along with a 3-percentage point increase in the number of schools where less than 25% of the school population were White British. In addition, they found a growing percentage of ethnic minority pupils (52.8% - 58.3%) attended secondary schools where White British students were not in the absolute majority (> 50%) over the same period. After condensing the five categories into three, and focusing on individual ethnic groups, they highlighted that students from Bangladeshi and Pakistani heritage were far more likely to attend schools with a low proportion of White British students, and that it was rare for students from any ethnic group, other than White British, to attend a school where their co-ethnics formed an absolute majority (> 50%). Indeed, most ethnic minority students did not attend a school where their ethnic group was in even the *relative* majority. There were only 42 schools in England where an ethnic minority group formed over 90% of the school population, which represented less than 0.3% of the schools in England. This led them to conclude that while "schools with very high percentages of any one minority group do exist, they are exceedingly rare and do not warrant the disproportionate amount of attention they can receive in the media or in some political discussion" (2020, p. 18).

In Chapter 5 the authors measured what this thesis would term, "ethnic polarisation" by calculating the extent to which a school's ethnic profile diverged from the ethnic profile of

the surrounding area. Using simulations that allocated students to the secondary school closest to the Lower Super Output Area (LSOA) of their home address, the authors measured the ethnic composition of these simulated schools, and then compared the proportion of each ethnic group in the simulation to the proportion observed in each school in real life. Based on this comparison, they found that only a small number of schools educated a percentage of Asian or Black students that was dissimilar to what would have been expected, had students been allocated to the school closest to their home LSOA. This led to them to conclude that "rather than being disproportionately concerned by the exceptional cases where schools are not like neighbourhoods, we might be better reassured that in most cases they are" (p147). Drawing on these respective analyses the authors summarized their findings by concluding that:

> *While high levels of ethnic segregation do exist across schools between the majority White British population and some other ethnic groups such as the Bangladeshi and Pakistani, more so at the primary than secondary level of schooling, and more for those of greater affluence among the White British [,]… the general trend has been towards desegregation and greater ethnic diversity within local authority areas and their schools.*
>
> *(Harris and Johnston, 2020, p. 177)*

Three earlier studies that measured ethnic polarisation are also worth mentioning. The first was conducted by Johnston *et al* in 2006 who used data on all primary and secondary schools to create a "classification scheme" that assigned schools and geographic areas into five categories based on the percentage of non-white students/residents in the school/geographic area, as well as the size of the largest non-white ethnic group (Johnston

et al., 2006).[45] The authors then compared which categories each school and their surrounding geographic areas were assigned to in order to assess how closely each reflected one another (Johnston *et al.*, 2006, p. 980).[46] As many schools were assigned to a different category than their local areas, the authors suggested that residential patterns did not always reflect the ethnic profile of local schools. A later study measured ethnic polarisation in relation to primary schools in Birmingham in 2001 (Harris and Johnston, 2008). The authors created "core catchment areas" for each school by evaluating the home postcodes of current students, and then compared the ethnic profile of each primary school against the ethnic profile of students living in their "core catchment area". When the proportions differed, i.e., the proportion of a given ethnic group was larger or smaller in the school than found in the core catchment area, the authors classed this as ethnic "separation". The article mostly focuses on the 5,824 Black Caribbean students in the study who were found to be the most "separated" ethnic group. The third study compared the ethnic profile of primary schools in England to catchment areas the authors created by mapping the area in which current students at schools predominantly lived (Johnston, Harris, et al., 2016). Using this approach, they found that "for the great majority of schools the proportion of their pupils from South Asian or Black minorities is commensurate with the proportion in their model-defined catchment areas", with the exception of a few Voluntary Aided schools, who tended to over-recruit White British students and under recruit ethnic minority pupils (p22).

[45] Differences in the age profiles of ethnic groups make this a poor comparison.

[46] This approach has some overlap with more conventional approaches for measuring ethnic school segregation. Instead of numerically comparing the divergence between the proportion of each ethnic group in each school to the proportion of the same ethnic group in the neighbourhood (Step 3 in Figure 3), the authors make a single comparison between the school and neighbourhood based on their categorizations.

Finally, a think tank report, published by The Challenge, School Dash and iCoCo Foundation, adopted a novel methodology where each school's ethnic profile was compared to "those of the 10 schools nearest to them" (2016, p. 11). They then argued that any schools which had an ethnic group that was underrepresented by 50% or overrepresented by 100% was segregated. These "segregated" schools were then labelled as experiencing 'high', 'medium' or 'low' levels of segregation depending on the extent to which the school's expected ethnic profile diverged from the observed profile. Based on these categorisations the authors concluded that "26% of primary schools and 40.6% of secondary schools [in England] were found to be ethnically segregated or potentially contributing to segregation" (2016, p. 13), which garnered headlines from a number of major UK news outlets (BBC, 2017; The Guardian, 2017). While the report's findings are interesting, it is important to stress that it has not been peer reviewed and it does not explain their methodology in great detail.[47] This is problematic in a quantitative field of research where so much debate has focused on methodological disagreements (Gorard and Fitz, 2006; Allen, Burgess and Windmeijer, 2009; Harris *et al.*, 2013).

Summary

Section 2.2 has shown that past research on evenness ethnic school segregation in England is small, often dated and calculated within geographic areas that do not reflect local educational markets and use measurement techniques that no longer provide the best approach to measuring ethnic school segregation. The section also highlighted that research

[47] For a critique on this approach, readers can refer to a blog post by two prominent academics in this field (Burgess and Harris, 2018).

is yet to provide a thorough investigation into why certain areas experience higher levels of ethnic school segregation than others. These factors constitute important gaps in the academic literature. Consequently, by mapping the ethnic profile of England's school age population across various geographic scales (research question 1), exploring the level of ethnic isolation faced by its students (research question 2), describing in detail the current and changing patterns of ethnic school segregation (research question 3) and exploring the predictors of ethnic school segregation (research question 4 - 6) this thesis aims to make a substantial contribution to academic literature.

2.3 Research into the predictors of ethnic school segregation

The following section presents research to justify the study's predictor variables. Evidence is initially presented to suggest that population density (Section 2.3.1) and a neighbourhood's socio-economic status (Section 2.3.2) are likely to be predictive of ethnic school segregation. It then suggests that selective admissions policies, whether based on a child's academic ability or religious faith, are likely to be associated with higher levels of ethnic school segregation due to an implicit sorting of ethnic groups that these policies encourage. Finally, Section 2.3.4 argues that lower quality schooling is likely to be associated with higher levels of ethnic school segregation, due to an increased level of competition for places at higher quality schools.

2.3.1 Population density

Previous research suggests that an area's population density has an important association with the observed level of ethnic school segregation. In 2016, Gorard found that the population density of the 34 LA's he studied had the strongest correlation with their level of ethnic school segregation out of all the predictor variables that he explored (ρ = -0.60). This strong *negative* correlation led him to conclude that "areas with high population density have lower [ethnic school] segregation, presumably because families have feasible access to more schools than those in rural areas" (Gorard, 2016b, pp. 139–140).

The strength of this finding alone necessitates the inclusion of population density as a predictor variable in this study, though there is additional research to suggest that this variable has important links to ethnic "separateness" in schools. The article by Burges, Wilson and Lupton, for example, found that in 2001 an LA's population density was the strongest predictor of an LA's ratio of ethnic school segregation to residential segregation for all eight ethnic groups studied by the article (2005).[48] The results were that "in dense (urban) areas we find the school segregation index is higher relative to residential segregation indexes" (2005, p. 1043).

[48] Their estimated effects sizes ranged from beta = 0.439 ($p<0.01$) for measures relating to Black Caribbean students up to 0.978 for measures relating to Black African student ($p<0.01$). This led them to conclude that "in dense (urban) areas we find the school segregation index is higher relative to residential segregation indexes" (Burgess, Wilson and Lupton, 2005, p. 1043). Readers should note that the outcome variable they used has been labelled as a "woefully inadequate measures of the effect of parental choice" due to the fact that the authors did not "construct and compare equivalent catchment areas" when calculating this ratio (Simpson and Finney, 2006, p. 421/422). Unfortunately, it is outside the scope of this thesis to comment on the best methodological approach to understand the interaction between ethnic residential and school segregation, as this would require a detailed discussion on how to measure ethnic residential segregation, which is not the focus of the study.

These findings justify the following research questions:

4a) Is the population density of the area surrounding a primary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?

4b) Is the population density of the area surrounding a secondary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?

2.3.2 Socio-economic status of the neighbourhood

A neighbourhood's socio-economic status is also likely to be associated to its level of ethnic school segregation, though estimates on the strength and direction of this association vary depending on which measure of socio-economic status researchers have used. For example, Gorard's study from 2016 found that ethnic school segregation tended to be higher in LAs where residents had completed less formal education ($\rho = 0.35$), suggesting a *negative* relationship where LAs with lower SES had higher levels of ethnic school segregation. However, the study also found that ethnic school segregation was lower in LAs with higher poverty rates ($\rho = -0.28$) and rising unemployment ($\rho = -0.35$), which suggests a *positive* relationship where poorer neighbourhoods had lower levels of ethnic school segregation (Gorard, 2016b, p. 139). Strangely, the article also found no correlation between an LA's ethnic school segregation and the percentage of students eligible for FSM, which makes predicting the direction of the association even more difficult.

The results from Burgess, Wilson and Lupton's article paint a similarly mixed picture (2005). Their regression models found that the ratio of ethnic school segregation to residential segregation tended to be higher for LAs with higher levels of poverty when measuring ethnic school segregation for Black Caribbean ($\beta = 0.009$, $p<0.001$) and Bangladeshi students ($\beta = 0.013$, $p<0.001$). However, when measuring segregation in relation to Indian students the association was negative and, for all other ethnic groups there was no statistically significant association.

Johnston and Harris book also provides some useful findings in relation to ethnic exposure. Their analysis found that within local authorities, Johnston and Harris book also provides some useful findings in relation to ethnic exposure. Their analysis found that within local authorities, FSM students are typically less exposed to other ethnic groups, except for White British students, for whom for whom the opposite is true, with non-FSM student less exposed to those from other ethnic backgrounds (Harris and Johnston, 2020, p. 190).

When viewed collectively these various research findings suggest there remains an important empirical gap in the literature which necessitates further study. Consequently, this thesis proposes the following research questions:

> *4c) Is the average house price of the area surrounding a primary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?*

> *4d) Is the average house price of the area surrounding a secondary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?*

4e) Is the rate of Free School Meal eligibility (FSM) within a primary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?

4f) Is the rate of Free School Meal eligibility (FSM) within a secondary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?

By including these two types of research questions the study allows for different dimensions of a neighbourhood's socio-economic status to have different associations with ethnic school segregation, thereby providing a more flexible and robust investigation. The first two questions aim to explore whether there is an association between ethnic school segregation and the wider neighbourhood's socio-economic profile, as captured by the average neighbourhood house price, while the latter two questions investigates whether there is an association with the socio-economic profile of the students themselves, as captured by the rate of FSM.[49]

However, it is worth highlighting that while this thesis uses the area's average house price as a measure of socio-economic status, there is also likely to be a more direct association between housing markets and patterns of ethnic school segregation due to the continued importance of home-school distance in many school admissions policies. UK research has, for example, found that houses located near high performing schools cost, on average, 8% - 10% more than other local houses (Department for Education, 2017), and many academics have argued the housing market is a key arena in which parents compete for places in

[49] As discussed later in Section 4.7, these metrices are easier to incorporate into this thesis's methodology, which defines local neighbourhoods using k means clustering, when compared to other statistics that relate in government boundaries, such as local unemployment rates or indices of multiple deprivation.

oversubscribed schools (West and Hind, 2006, 2016). Unfortunately, it is beyond the scope of this thesis to investigate these complex interactions in detail, as this would require the use of spatial data, information on social housing and radically different modelling techniques than those used in this study. However, by investigating whether there is an overarching association between ethnic school segregation and house prices the study can provide a solid foundation for more targeted research in the future.

2.3.3 Selective admissions policies

Selective admissions policies, whether based on children's academic ability or religious faith, is another area which warrants investigation into possible associations with ethnic school segregation. As argued in the following two sub-sections, evidence suggests that both types of policies are likely to be associated with higher levels of ethnic school segregation due to the implicit sorting of ethnic groups that such policies facilitate.[50] With discussions ongoing over whether to increase the number of grammar schools (Long, Foster and Robert, 2018) and religious schools in England (Department for Education, 2018b), the findings of this study should also provide a timely insight to support the debate over these proposals.

[50] Private, fee-paying, schools is another important avenue of school choice and therefore should be included in this section. However, the government does not require private schools, who represent 7% of the school population, to release data on their student's ethnic profile which makes their inclusion in this quantitative study impossible.

2.3.3.1 Academically selective admissions policies

While most research in England has focused on showing that academically selective admission policies tend to be associated with higher levels of socio-economic school segregation (Goldstein and Noden, 2003; Allen, 2007; Gorard, 2016b), two studies provide evidence to suggest that there are similar associations with ethnic school segregation. The most compelling study was conducted by Gorard and Siddiqui, who used the 2015 school census data, to show that grammar schools educated a disproportionately large percentage of non-white students (25.1%) when compared to non-grammar schools, both nationally (19.3%) and in relation to schools operating within Local Authorities with selective schools (18.1%) (2018). A smaller study, using data from 2008 on students attending state schools in London, also found evidence that the ethnic profile of grammar schools diverged from non-selective schools. Specifically, they found that the proportion of Chinese students in academically selective schools (4%) was eight times the proportion in other, non-selective schools (0.5%) (Harris, 2013).

These results are likely to reflect differences in the average academic performance of students from different ethnic backgrounds at the end of Key Stade 2 (Strand, 2015) which mean that a disproportionate number of students from high achieving ethnic groups are likely to pass the entrance exams for grammar schools (Chinese, Indian, Irish) compared to students from lower achieving ethnic groups (Black Caribbean, Pakistani, White and Black Caribbean).[51]

[51] The end of KS2 is the point at which student sit the entrance exams for grammar schools. Note, differences in the overall size of the ethnic groups means that these proportional differences do not necessarily reflect absolute differences in the number of students from each ethnic group attending grammar schools. A good

There may also be socioeconomic factors that would suggest grammar schools contribute disproportionately to ethnic school segregation. Coldron, Cripps and Shipton, for example, argue that "the greater education and wealth of middle class parents combined with a greater motivation arising from the fear of downward mobility…lead[s] to a disproportionate number of better off children applying for a grammar school places and doing better in the tests" (Coldron, Cripps and Shipton, 2010). This, in turn, means that there is a likely to be overrepresentation of students from higher income/wealth ethnic groups, such as White British and Indian, and an underrepresentation of students from lower income/wealth ethnic groups, such as Black African and Bangladeshi (Office for National Statistics, 2020).

However, given that no study in England has directly investigated the association between selective admissions policies and evenness ethnic school segregation, there remains a clear gap in the literature for a targeted piece of research to fill. Instead the aforementioned studies compared the collective ethnic profile of grammar and non-grammar schools within London (Harris, 2014) or across an area comprised of all the Local Authorities with at least one grammar school (Gorard and Siddiqui, 2018).[52] As these comparisons were made using schools located across large geographic areas, the findings are likely to be driven by wider residential patterns and therefore tell us very little about the level of ethnic school segregation occurring within neighbourhoods.[53] Consequently, by providing an analysis

example of this is Chinese students, who are substantially over-represented in grammar schools but, due to their small size, only make up 4% of grammar school enrolment.

[52] This reflects the primary focus of the articles on exploring socio-economic issues.

[53] These approaches also suffer from a number of methodological problems. For example, the single grammar school in Liverpool will have introduced 56,432 students into Gorard and Siddiqui's analysis whose collective

which directly measures ethnic school segregation, at a geographic scale which reflects local educational markets, this study enhances existing knowledge in this area. Two questions are proposed to explore the association. These are:

> *4g) Do secondary school pseudo-neighbourhood's that contain a grammar school experience higher levels of ethnic school segregation, on average, than those that do not?*

> *5a) Do academically selective secondary schools contribute more to ethnic school segregation, on average, than non-selective secondary schools located in the same pseudo-neighbourhood?*

As explained later in Section 4.5.2 in the methodology chapter, the first question investigates possible associations occurring at a pseudo-neighbourhood level while the second question explores school level differences within the same pseudo-neighbourhood. This provides an analysis which is less susceptible to the ecological fallacy problem (Openshaw, 1984), where trends observed at higher geographic scales are assumed to be reflective of those occurring at lower levels.

ethnic profile will have been given equal weighting to students attending school in Kent, where 33 grammar schools were operating.

2.3.3.2 Religiously selective admissions policies

Evidence over whether religious schools have a positive or negative association with ethnic school segregation seems to vary depending on the scale at which is it measured. Gorard's article, which is the only national study in England to directly investigate the association between religious schools and evenness ethnic school segregation, found that LAs with more Voluntary Aided schools tended to experience lower levels of ethnic school segregation (ρ = -0.22) (Gorard, 2016b). As Voluntary Aided schools are overwhelmingly religious, this finding suggests that areas with more religious schools tend to have lower levels of ethnic school segregation.

In contrast, other studies that measured different types of ethnic "separateness" within schools, found evidence that religious schools tended to be associated with higher levels of ethnic "separateness". Harris and Johnston, for example, found that of the 149 schools in England that had a substantial underrepresentation of Asian pupils, relative to their surrounding area, 127, or 85%, were faith school(Harris and Johnston, 2020, p. 132). Similarly of the 77 primary schools with a substantial under-representation of black students, relative to their surrounding area, all but eight, or 90% were faith schools. An earlier study by Harris and Johnston, based on data from 2001 on all students attending state-funded primary schools in Birmingham, found that nine of the eleven schools which over-recruited Black Caribbean pupils, relative to the schools' catchment areas, were faith schools (Harris and Johnston, 2008, p. 78).[54] A study by Allen and West using data from 2005

[54] This measure "ethnic polarisation" which captures the divergence between each school's ethnic profile and the ethnic profile of student's living in the schools immediate surrounding area.

on year 11 students attending state funded schools in London found that pupils from Black African and Caribbean ethnic groups were over-represented in faith schools, while Pakistani and Bangladeshi students where under-represented, relative to the average ethnic profile of all students attending schools in London (Allen and West, 2009a). These findings were later corroborated by a study on year 6 students in London, which found that in 2008 the proportion of Black Caribbean (13.8%) and Black African (19.5%) pupils in Church of England schools was around twice the proportion found in non-religious schools in the city (6.4% / 11.8%) (Harris, 2013).

A number of articles have also found that religious schools in London tend to recruit pupils with higher academic abilities than their non-religious counterparts (Gibbons and Telhaj, 2007; Allen and West, 2009b; Harris, 2013). If this is indicative of a broader nationwide trend, then religious schools are likely to overrecruit students from high achieving ethnic groups, resulting in a similar implicit sorting of ethnic groups as the process described in relation to grammar schools.[55] Additionally, past research has found that faith schools sector tends to recruit a smaller proportion of students who are eligible for free school meals (Allen and West, 2009b; Harris, 2011).[56] This could mean that religious schools under-recruit students from lower-income ethnic groups, resulting in student populations whose ethnic profile differs from other local schools.[57] Finally, research that used data on the

[55] This is the same process outlined in relation to grammar schools.

[56] There is some variation between different types of religious schools. While Voluntary aided, Church of England and Roman Catholic schools all underrecruited FSM eligible pupils, voluntary controlled schools slightly over-recruited FSM pupils (Harris, 2011).

[57] These two cycles could also be compounded by what is known as the "Matthew effect", where good academic outcomes lead to the development of a "perception of excellence [which] leads more parents with higher social capital to select these schools… [which] only enriches the privileged status… these schools already enjoy" (Merry, 2015, p. 144).

application choices of children applying for secondary schools to compare the relative success of different ethnic groups found that children from ethnic minority backgrounds were less likely to gain a place at a church school than those from White British families (Weldon, 2016, p. 9). Such differences could mean that religious schools tend to over-recruit White British students, thereby contributing disproportionately to the local neighbourhood's level of ethnic school segregation.

Based on this evidence religious schooling is likely to have an association with ethnic school segregation and so the following research questions are posed:

4h) Is there an association between the proportion of schools in a primary school pseudo-neighbourhood that have a religious affiliation and its observed level of ethnic school segregation?

4i) Is there an association between the proportion of schools in a secondary school pseudo-neighbourhood that have a religious affiliation and its observed level of ethnic school segregation?

5b) Do primary schools with a religious denomination contribute more to ethnic school segregation, on average, than non-religious primary schools located in the same pseudo-neighbourhood?

5c) If so, do primary schools with particular religious denominations contribute more to ethnic school segregation, on average, than others?

5d) Do secondary schools with a religious denomination contribute more to ethnic school segregation, on average, than non-religious secondary schools located in the same pseudo-neighbourhood?

5e) If so, do secondary schools with particular religious denominations contribute more to ethnic school segregation, on average, than others?

As with the investigation into grammar schools, the study explores associations between pseudo-neighbourhoods and between schools located within the same pseudo-neighbourhood. The school level analysis is particularly helpful as it allows results to be delineated between faith schools with different religious denominations, thereby avoiding findings which are solely based on the single heterogenous grouping of "religious schools".

2.3.4 School Quality

The final factor which may have an association with ethnic school segregation is the quality of schools within an area[58]. As there were no studies that have directly explored this association in England, its inclusion is predominantly justified based on the theoretical work of competition theorists.[59]

[58] For a discussion on past research that explores how ethnic/racial school segregation is associated with students' academic performance see Section 11.3 in the appendix.

[59] It is worth noting that Harris and Johnston showed that "a greater percentage of the White British were in below good schools in 2017 than was any of the other groups" and that a "lower percentage [of White British students] were in outstanding schools. If such imbalances are reflected within local pseudo-neighbourhoods, or 'educational markets', then this school quality would be likely to be associated with differing levels of "school quality" (Harris and Johnston, 2020, p. 169)

Competition theorists suggest that when resources are scarce, competition tends to increase within an area, which leads to an increase in the salience of ethnic identities. This, in turn, makes shared spaces between ethnic groups increasingly fraught (Olzak, Shanahan and West, 1994; Goyette, Farrie and Freely, 2012) which encourages people to "hunker down" within their own ethnic group (Blalock, 1967; Putnam, 2007).

While these ideas are normally applied to competition over monetary resources or jobs, similar dynamics could well occur in relation to school places in areas where there are few spaces at high-quality schools. In neighbourhoods where schools are generally "good" there is less reason for families to compete for places at oversubscribed schools, as their child is likely to end up at a "good" school regardless of whether it is their first choice or not. In contrast, when there is a limited number of "good" schools, these places become scarcer and the incentives for families to compete increases. In this scenario, families with social or financial capital are likely to move to secure places in "good" schools for their children, most commonly through buying or renting houses near high performing schools, or through using the appeals system (Cheshire and Sheppard, 2004). This is likely to create an implicit sorting of students from different ethnic groups based on their average ability to financially compete within the housing market or navigate the appeals system. Wealthier ethnic groups, who are confident speaking English and have a better understanding of the school system, would be more likely to gain entry to higher performing schools, while students from poorer ethnic groups, or those who are less able to communicate in English and less

familiar with the local school system, would be more likely to end up in lower performing schools.[60] Based on this reasoning, the thesis proposes the following research questions:

4j) Is there an association between the average OFSTED rating of schools within a primary school pseudo-neighbourhood and its level of ethnic school segregation?

4k) Is there an association between the average OFSTED rating of schools within a secondary school pseudo-neighbourhood and its level of ethnic school segregation?

5f) Is there an association between a primary school's OFSTED rating and their contribution toward ethnic school segregation, relative to other primary schools located in the same pseudo-neighbourhood?

5g) Is there an association between a secondary school's OFSTED rating and their contribution toward ethnic school segregation, relative to other secondary schools located in the same pseudo-neighbourhood?

Section Summary

Section 2.3 has explained why each of the neighbourhood characteristics and school features included in this study were likely to have an association with ethnic school segregation, thereby justifying their inclusion in the analysis. These explanations, along with

[60] These ideas dovetail with debates over how school segregation, both in relation to ethnicity and socio-economic status, may be effected by greater school choice (Hamnett and Butler, 2013, p. 318)

the related research questions, should provide additional clarity on the nature of this study as well as offer key contextual information to help readers interpret the results in Chapter 6 and engage with the discussions produced in Chapter 8.

2.4 England's National School Admissions System

The final section of this chapter explains why the study decided to investigate the links between England's national school admissions system and ethnic school segregation. It begins by familiarising readers with the complexities of England's school admissions system, with a particular focus on explaining how places at oversubscribed schools are allocated to applicants. Section 2.4.2 then explains how a historical shortage of data, and the lack of a "counterfactual" system which England's school admissions can be compared against, has made it challenging to investigate the links between the school admissions system and ethnic school segregation at a national level.

Based on these discussions, Section 2.4.3 makes the case for two new analyses to improve current academic understanding. Firstly, it argues that a new bespoke extract from the National Pupil Database offers the chance for us to investigate how ethnic school segregation shifts between each stage of the yearly school admissions cycle. This would provide a detailed insight on the relative importance of each stage of the admissions cycle in setting patterns of ethnic school segregation. Secondly, it argues that simulation analysis offers a tool to help explore how patterns of segregation may have differed, had students been allocated to schools using alternative approaches. This would help to contextualise England's current approach as well as provide insights into how potential reforms could help

reduce ethnic school segregation in England. These discussions justify research questions 6a – 6e.

2.4.1 Understanding England's National School Admission System

To understand how England allocates students to state-funded schools it is useful to split the yearly admissions cycle into three stages.[61] In the first stage, parents are asked to rank between three to six schools that they would prefer their child to attend the following school year. [62] This is completed via an application form which they send to their local authority during the autumn term before their child is due to start primary or secondary school. In the second stage, Local Authorities collate these preferences and create a list of applicants for each school which does not consider how highly the applicant ranked the school on their application form. If a school has the capacity to offer all applicants on their list a place, in accordance with the school's Published Admissions Number (PAN), then all applicants are earmarked for an offer. If the school does not have the capacity to offer all applicants a place, then applicants are ranked based on the school's "oversubscription criteria" and a certain number of places are earmarked for higher ranked students.[63] Local Authorities then ensure that students who are earmarked for a place at more than one school only received one offer from whichever school was ranked highest on their original application form. Invariably this step creates space for other students at oversubscribed

[61] For more information see (Department for Education, 2018c). For more information on the development of school admissions in England please refer to the following articles (Pennell, West and Hind, 2006; Coldron *et al.*, 2008; Machin and West, 2013).
[62] The number of schools parents can list varies between different local authorities (Department for Education, 2018c).
[63] This is arranged with the Local Authority and is known as the school's Published Admissions Number (PAN).

schools, which in turn facilitates a degree of reallocation. Once this is completed, students who are not earmarked for an offer from any school on their application form are then earmarked for a place at a local school which still has capacity, typically based on their home-school distance. After this administrative procedure is completed, Local Authorities send all students a letter with a single school offer in the spring before the students are due to start their new school. This offer can be thought of as the second stage of the application cycle. Finally, in the third phase of the application cycle, students have the option to appeal against the Local Authority's decision. In practice, this means that some students do not attend the school they were initially offered.

Given that between 80% - 90% of children attend their first preference school (Roberts, 2017, p. 23), the expressed preferences of parents tend to translate directly into the patterns of school attendance that are observed in England, making it largely a "school choice" system. However, as a substantial proportion of students do not attend their preferred option, the mechanisms for allocating places at oversubscribed schools remains an important feature of the national school admissions system.

Unfortunately, it is hard to summarise how places at oversubscribed schools are allocated to students across England, due to the decentralised nature of the school admissions system. As mentioned previously, students applying to oversubscribed schools are placed on a school list and ranked based on the school's "oversubscription criteria". As these criteria are written by local bodies called Admissions Authorities for a single, or small group of schools, often in relative isolation, there is considerable variation in the criteria used by different

schools; both in terms of the factors they consider and the relative importance that each factor holds.[64] This makes it hard to describe these systems at a national level.

Table 2: Oversubscription admissions criteria in publicly-funded secondary schools in London between 2001-2015 (West and Hind, 2016)

Oversubscription Criteria	**2015 N = 410**	**2012 N = 392**	**2008 N = 398**	**2005 N = 362**	**2001 N = 382**
Sibling policies	91%	92%	95%	95%	86%
Home-school distance	90%	92%	95%	95%	86%
Statement of SEN	68%	55%	50%	47%	44%
Medical/Social Need	63%	65%	61%	70%	72%
Supplementary information form	33%	26%	-	-	-
Religion	25%	25%	24%	25%	27%
Banding	23%	21%	19%	-	20%
Catchment Area	18%	18%	11%	8%	6%
Child attends feeder school	12%	11%	20%	18%	15%
Partial selection by ability/aptitude	10%	8%	8%	7%	5%
Random allocation (zone/area)	6%	4%	-	-	-

Note: This is not an exhaustive list of criteria/practices. As 100% of schools prioritized children in care, in line with their legal obligation to do so, this factor was removed from the table

Key: - not recorded

[64] Admissions Authorities can write the oversubscription criteria for a single school or a large group of schools. All oversubscription criteria must be made public and abide by the statutory guidelines set out in The School Admissions Code (Department for Education, 2014). The code is fairly detailed, and ranges from specific requirements, for example that schools should give the "highest priority…to looked after children", to broader conditions like "oversubscription criteria must be reasonable, clear, objective, procedurally fair" (Department for Education, 2014, p. 10). It nonetheless leaves Admissions Authorities with a sizeable degree of autonomy over how they write their own criteria and, as a result, a broad range of methods are used to allocate places at oversubscribed schools, as shown in Table 2.

This challenge is only compounded by the lack of national-level data, which leaves researchers reliant on a small number of studies that have collected information on the oversubscription criteria of a sample of Admissions Authorities. The results from one recent study, which investigated 410 secondary schools in London, are outlined in the table above (West and Hind, 2016). Reassuringly, their results broadly align with those from previous studies which suggest they are fairly representative (Pennell, West and Hind, 2006; Coldron *et al.*, 2008; Allen, Coldron and West, 2012).

Unfortunately, these studies provide little insight into how important each of these factors are, relative to one another, or the exact procedure each school uses when ranking applicants. Given the decentralised nature of how oversubscription criteria are written, it is likely that there is considerable variety in this regard.

2.4.2 Past research into how the school admission system leads to the separation of ethnic groups within English schools

Given that most children attend their preferred school, most academic studies that explore whether schools admissions results in the separation of different ethnic groups have focussed on the first stage of the application process, investigating "how and what parents…choose" (Wilson and Bridge, 2019, p. 3198).[65] A recent study, for example, used

[65] In England, studies tend to find that a school's academic test scores, the home-school distance, the socio-economic profile of the school and the socio-economic status of the family are the most important factors in explaining the how parents rank schools in their area (Noden *et al.*, 1998; Benson, Bridge and Wilson, 2015; Burgess *et al.*, 2015; Wilson and Bridge, 2019)

random utility models to investigate whether the application choices parents made in three northern towns in England could be explained by the ethnic profile of the schools relative to the ethnic identity of each child (Weldon, 2016). The results suggested that "both white and Asian parents usually preferred to send their children to schools with a higher proportion of their own ethnicity" but that all groups tended to have an "aversion to sending one's child to a school overwhelmingly dominated by one's own ethnic group" (p173). This led the author to conclude that "parents do not want extreme ethnic segregation … they would rather their children went to an ethnically mixed school as long as their ethnic group constitutes the majority" (p173). Another study, which explored how families from different ethnic backgrounds completed their school application forms using descriptive statistics (Burgess, Greaves and Vignoles, 2019), found that "41% of White British households only make one choice, compared to 17% of Asian households and 12% of Black households". Additionally, families in more densely populated neighbourhoods tended to list more schools on their application form, presumably due to a greater number of available school places, and "Black and Asian households [were] less likely to have an offer from their first-choice school than White British households" (p702). Given that Black and Asian households tend to live in urban areas and rural areas tend to contain a higher proportion of White British students, these two findings could well have an interaction with one another, though more research would be needed to explore such associations in detail.

Unfortunately, far "fewer studies deal with the relationship between choice and the subsequent allocation of pupils to schools" (Wilson and Bridge, 2019, p. 3198) in stage two of the application process, due primarily to two methodological challenges. Firstly, until recently there had been no high-quality, national dataset on the application choices of

students. Instead, data was kept by each Local Authority, which prevented researchers from analysing the school admissions system across England as a whole. However, since 2016, the DfE has collated these datasets and released them as a bespoke extract as part of the National Pupil dataset. This dataset lists which school each parent wrote on their child's application form, the offers each child received, and the secondary schools they ended up attending the following school year. This under-utilised dataset therefore provides this thesis with the opportunity to be the first piece of research to investigate how England's national school admissions system links to patterns of ethnic school segregation.

A second, broader challenge is that there is no obvious "counterfactual system" which England's current admissions system could be compared against. Instead, academics have been forced to choose from two types of comparison, neither of which offers a particularly appealing way of understanding how England's *national* school admissions system relates to ethnic school segregation.[66] One approach is for researchers to investigate how the patterns of school segregation changed before and after a change in the national admissions system. For example, there was a heated academic debate in England over whether *socio-economic* school segregation rose following the introduction of school choice in England (Goldstein and Noden, 2004; Gorard and Fitz, 2006; Allen and Vignoles, 2007; Gorard, 2007).[67] Alternatively, they can carry out a study on one type of school or neighbourhood, and use the results to hypothesise about wider dynamics throughout the national system. For example a study released in 2015 found that 88 Free Schools in England had different ethnic

[66] One approach not included in this description would be to compare England's school admissions system and level of ethnic school segregation against other countries. However, this would be highly questionable, given how important the ethnic make-up and residential patterns of a nation's student population are in determining patterns of ethnic school segregation.

[67] In the end it was generally accepted that socio-economic school segregation did not change sig

profiles to their neighbouring schools (Green, Allen and Jenkins, 2015, p. 907). As Free Schools were a new type of school launched as part of a move to expand "school choice", these findings were used to argue that greater school choice was leading to more ethnic school segregation.

While both types of studies provide useful findings, they also suffer from methodological challenges. For example, when comparing trends across time, it is difficult to know whether a change in school segregation resulted from a change in the national school admissions system or merely coincided with it, particularly if key characteristics of the student population changed during the same time period (Harris, 2010, p. 6). For instance, this chapter has described research that suggests ethnic school segregation fell in England between 1997 – 2012 (Gorard, 2016b), which overlaps with the expansion of academies and other school choice apparatus.[68] Read in one way, this provides evidence to suggest that school choice is associated with lower levels of ethnic school segregation. However, the ethnic profile of the school aged population (Strand, 2015, p. 18; Department for Education, 2018a) and residential patterns of ethnic groups (Harris and Johnston, 2020) also changed during this time, which makes it hard to draw any firm conclusions from this association. Similarly, while localised results may offer a good insight into a particular feature of the national system, they do not necessarily provide reliable or accurate accounts when extrapolated to the system as a whole. While these issues do not invalidate the results of these types of studies, they suggest that the field could benefit from a broader range of methodological approaches being used to study these dynamics.

[68] For a description of this study see Section 2.1.5.

2.4.3 A new approach to investigating how school admissions link to patterns of ethnic school segregation

This study aims to help address these challenges by completing two new analyses that broaden current academic understanding in this area. Firstly, the recent collation of the dataset by the DfE creates the opportunity for this thesis to provide the first nationwide study into how patterns of school segregation shift at each stage of the school admissions cycle. Consequently, the thesis proposes the following research question:

> *6a) How do the likely patterns of ethnic school segregation shift between each stage of the national school admissions cycle in England?*

Secondly, the study uses simulation analysis to broaden the base of empirical evidence that academics and policymakers can consult to explore the links between the national school admissions system and patterns of ethnic school segregation.[69] Specifically, the study allocates students to schools using a range of school admissions procedures to create "counterfactual" admissions systems against which England's current system can be compared. While this is a novel approach, there are an increasing number of studies that have adopted similar methodologies. The most relevant example was produced by Harris and Johnston in 2020, who used NPD data to explore how the ethnic profile of schools

[69] While not discussed in detail in this section of the thesis, European researchers are increasingly using Agent Based modelling to explore segregation dynamics over time (Crooks, 2010; Mahdavi Ardestani, O'Sullivan and Davis, 2018). This could be an interesting way of extending the work of this thesis to offer a longitudinal perspective.

would differ if students in England were reallocated to their nearest school. By comparing the simulated pattern of school attendance to the patterns observed in real life they concluded that most schools educated a similar proportion of students from Asian and Black ethnic groups under the distance-based simulation and in real life (Harris and Johnston, 2020).[70] This built on an earlier piece of research by Allen, who used a similar type of simulation to explore how patterns of *socioeconomic* school segregation would differ under a strict home-school distance simulation. Her results concluded that socio-economic "school segregation [was] almost always lower in the proximity counterfactual than in the actual data" (Allen, 2007, p. 258).[71]

The insights gained from simulation analyses should provide a range of benefits for this field. Firstly, unlike localised studies, the simulations can be run at a national scale which removes the need for results to be extrapolated beyond the context in which they were produced. Secondly, it removes the restrictions on researchers to only study policy approaches that have already happened and instead allows them to explore what patterns of school segregation could occur if the system adopted a new design. For example, this thesis draws on two approaches to school admissions that have been used by a number of local areas to predict the patterns of ethnic school segregation that would be likely to occur if similar policies were rolled out across the country. Finally, the results from simulation

[70] The authors did not measure "ethnic school segregation", as conceptualized by this thesis, but "ethnic polarisation", which explores whether the ethnic profile of a school's surrounding area differs from the ethnic profile of the school. See Section 2.2.2 for more details.

[71] This is similar to a study by Bonal, Zancajo and Scandurra, who compared the levels of socio-economic school segregation in Barcelona with a counterfactual simulation where all students were allocated to their nearest school. They found lower levels of socio-economic segregation in the simulated proximity simulation which led them to conclude that there was a "trade-off between increasing families' choice and equity" in the school system (2021, p. 59).

analysis can often have a fairly direct application to policy debates, which increases the chances of the research having an impact on real-world decision making.

However, this does not mean that simulation studies offer a perfect tool. In particular, if the government changed the school admissions system, the behaviour of many families would also shift. For example, parents may choose to live in different areas or list different schools on their application form. This means that creating simulations based on data collected under one system may provide a dubious basis for simulating the outcomes under another system. Unfortunately, this is an unavoidable problem with this type of analysis which this thesis does not shy away from. However, given the difficulty of measuring these complex associations at a national scale, using a broader range of methodologies should still provide a more robust evidence base for researchers to draw on. Specifically, while simulations come with their own methodological issues, they can still contribute towards deepening current academic understanding.

Based on this reasoning, this thesis will use seven simulations to explore what patterns of ethnic school segregation might pertain in England under different national school admissions systems. The first five of these simulation work within the current school choice framework, while the latter two remove the ability of parents to express a direct preference over the school their child will attend. These simulations, which are discussed in more detail within the methodology chapter, will address the six research questions listed below: [72]

[72] Two simulations were used to investigate research question 6c. For more information see Section 4.9.3.

6b) What patterns of ethnic school segregation would be likely to arise in England if places at oversubscribed schools were allocated to applicants based on their home-school distance?

6c) What patterns of ethnic school segregation would be likely to arise in England if places at oversubscribed schools were allocated to applicants at random?

6d) What patterns of ethnic school segregation would be likely to arise in England if academic banding was used to allocate places at all oversubscribed schools?

6e) What patterns of ethnic school segregation would be likely to arise in England if places at oversubscribed schools were allocated to applicants with the highest Key Stage Two test results?

6f) What patterns of ethnic school segregation would be likely to arise in England if students were allocated to their nearest school?

6g) What patterns of ethnic school segregation would be likely to arise within the secondary school system in England if the government reintroduced the 1960's grammar school system?

It is important to stress that each simulation draws directly on an approach that has either been adopted by another country or used within a local area in England which helps to ground the research in "real world" policy. The first simulation extends one of the most

common approaches used in England today to all schools in a uniform manner. The second simulation draws on the novel "lottery" approach employed by Brighton, a city in southern England, which aims to reduce social stratification (Allen, Burgess and Mckenna, 2013). Similarly, the third simulation is designed around the "academic banding" approach used by some Local Authorities which also aims to reduce social or academic divisions within their local school eco-systems (Tower Hamlets, 2020).[73] Both these approaches have been championed by academics, who see these two approaches as a tool to reduce inequalities within the school system (Burgess, Vignoles and Greaves, 2014). By extending these two approaches to a national level, the study can explore whether either policy offers a way of reducing ethnic school segregation across the whole country. The fourth simulation aims to replicate the approach used in Stockholm, Sweden where students compete for places at their desired high schools based on their primary school grades (Söderström and Uusitalo, 2010). This will help us to explore how the use of academic selection within a school choice framework would impact patterns of ethnic school segregation in England. The fifth strict "home-school distance" simulation has many parallels with the admissions system that operated in England before the introduction of school choice, or the systems currently used in Finland and Germany (OECD, 2010, p. 428).[74] The findings from this simulation will also shed light on whether the English school age population is more or less ethnically segregated in schools than they are in their residential patterns. Finally, the sixth simulation explicitly copies the design of the grammar school system that existed in England during the 1960s, to investigate what patterns of ethnic school segregation would be likely to arise if this system was reintroduced. This should provide useful empirical evidence to inform

[73] For more information on school banding see (West, 2005; West and Hind, 2016).
[74] 83% of countries in the OCED using home-school distance as a factor when allocating students to publicly funded schools (OECD, 2010).

ongoing debates over whether to expand the grammar school sector in England (Long, Foster and Robert, 2018).

Collectively, these simulations provide a cache of empirical findings that offers a new perspective to deepen academic understanding on how national school admissions systems link to ethnic school segregation in England. This is a crucial gap in the academic literature, not just because of the ramifications on academic understanding, but also because school admissions are one of the most important policy levers available to address ethnic school segregation. Hopefully, by providing this new analysis, the thesis can help to address this gap.

Section Summary

Section 2.4 has provided an overview of England's national admissions system and explained why it remains so challenging to investigate the links it has with ethnic school segregation. Based on this discussion, it suggested that mapping patterns of ethnic school segregation through each stage of the admissions process will offer a useful longitudinal perspective to deepen current academic understandings on how the school admission system creates patterns of school segregation. Secondly, by simulating the patterns of ethnic school segregation that would be likely to arise if students were allocated to schools using a range of different approaches, the thesis will help contextualise England's current approach and explore the potential consequences that would arise from reforming the school admissions system, as it relates to ethnic school segregation.

2.5 Summary of the literature review

This chapter began by explaining how this thesis will define segregation and the important dilemmas researchers face when investigating this topic. It then evaluated the previous academic research which has measured evenness ethnic school segregation, and other forms of ethnic separateness in England schools, both to contextualise this research project as well as outline the gaps that this thesis aims to fill. Section 2.3 then drew on academic evidence to justify the variables this study will use to predict a pseudo-neighbourhood's level of ethnic school segregation and a school's contribution towards their pseudo-neighbourhood's level of ethnic school segregation. Finally, Section 2.4 explained how England's national school admissions system currently allocates students to schools before justifying the two new approaches that this study will use to investigates its links with ethnic school segregation. The following chapter will present the study's research questions before the methodology is outlined in Chapter 4.

3 Research Questions

This thesis will offer the most up to date measurement of ethnic "separateness" within the English school system in terms of ethnic isolation and evenness ethnic school segregation. In addition, it will describe how each of these trends have changed between 2006 - 2019, both geographically and numerically. It will then investigate which pseudo-neighbourhood characteristics are predictive of ethnic school segregation (*H*) and which school level features are predictive of a school's contribution towards ethnic school segregation (H_j). Finally, it explores how patterns of ethnic school segregation shift between each stage of the secondary school admissions process, before exploring what patterns of ethnic school segregation would be likely to occur if students were allocated to secondary schools via a range of different approaches. The thesis proposes 6 research questions, each with their respective sub-questions, as listed below:

Research Question 1: What is the ethnic profile of England's school age population and how has it changed between 2006 – 2019 across various geographic scales?

1a) What is the current ethnic profile of England's school aged population at a national, regional, and pseudo-neighbourhood level?

1b) How has the ethnic profile of England's school aged population changed at a national, regional, and pseudo-neighbourhood level?

Research Question 2: What is the current level of ethnic isolation within England's school system and how has it changed between 2006 – 2019?

2a) What are the current levels of ethnic isolation in English schools?

2b) How does the level of ethnic isolation in English schools vary between different ethnic groups?

2c) How has the level of ethnic isolation changed between 2006 - 2019?

Research Question 3: What is the current level of ethnic school segregation within England's pseudo-neighbourhoods and how has this changed between 2006 – 2019?

3a) What are the current levels of ethnic school segregation within England's primary and secondary school pseudo-neighbourhoods?

3b) How has ethnic school segregation within England's primary and secondary school pseudo-neighbourhoods changed between 2006 - 2019?

Research Question 4: What pseudo-neighbourhood characteristics are predictive of ethnic school segregation in England?

4a) Is the population density of the area surrounding a primary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?

4b) Is the population density of the area surrounding a secondary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?

4c) Is the average house price of the area surrounding a primary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?

4d) Is the average house price of the area surrounding a secondary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?

4e) Is the rate of Free School Meal eligibility (FSM) within a primary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?

4f) Is the rate of Free School Meal eligibility (FSM) within a secondary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?

4g) Do secondary school pseudo-neighbourhood's that contain a grammar school experience higher levels of ethnic school segregation, on average, than those that do not?

4h) Is there an association between the proportion of schools in a primary school pseudo-neighbourhood that have a religious affiliation and its observed level of ethnic school segregation?

4i) Is there an association between the proportion of schools in a secondary school pseudo-neighbourhood that have a religious affiliation and its observed level of ethnic school segregation?

4j) Is there an association between the average OFSTED rating of schools within a primary school pseudo-neighbourhood and its level of ethnic school segregation?

4k) Is there an association between the average OFSTED rating of schools within a secondary school pseudo-neighbourhood and its level of ethnic school segregation?

Research Question 5: What school features are predictive of a school's contribution towards their pseudo-neighbourhood's level of ethnic school segregation (H_j)?

5a) Do academically selective secondary schools contribute more to ethnic school segregation, on average, than non-selective secondary schools located in the same pseudo-neighbourhood?

5b) Do primary schools with a religious denomination contribute more to ethnic school segregation, on average, than non-religious primary schools located in the same pseudo-neighbourhood?

5c) If so, do primary schools with particular religious denominations contribute more to ethnic school segregation, on average, than others?

5d) Do secondary schools with a religious denomination contribute more to ethnic school segregation, on average, than non-religious secondary schools located in the same pseudo-neighbourhood?

5e) If so, do secondary schools with particular religious denominations contribute more to ethnic school segregation, on average, than others?

5f) Is there an association between a primary school's OFSTED rating and their contribution toward ethnic school segregation, relative to other primary schools located in the same pseudo-neighbourhood?

5g) Is there an association between a secondary school's OFSTED rating and their contribution toward ethnic school segregation, relative to other secondary schools located in the same pseudo-neighbourhood?

Research Question 6: How does the England's national school admissions system relate to patterns of ethnic school segregation within secondary school pseudo-neighbourhoods?

6a) How do patterns of ethnic school segregation within secondary school pseudo-neighbourhoods shift between each stage of the national school admissions cycle in England?

6b) What patterns of ethnic school segregation would be likely to arise in England if places at oversubscribed schools were allocated to applicants based on their home-school distance?

6c) What patterns of ethnic school segregation would be likely to arise in England if places at oversubscribed schools were allocated to applicants at random?

6d) What patterns of ethnic school segregation would be likely to arise in England if academic banding was used to allocate places at oversubscribed schools?

6e) What patterns of ethnic school segregation would be likely to arise in England if places at oversubscribed schools were allocated to applicants with the highest Key Stage Two test results?

6f) What patterns of ethnic school segregation would be likely to arise in England if students were allocated to their nearest school?

6g) What patterns of ethnic school segregation would be likely to arise within the secondary school system in England if the government reintroduced the 1960's grammar school system?

4 Methodology Chapter

The following methodology chapter provides a detailed description of the analysis that was carried out as part of this research project. The chapter begins by describing the sources of data and the four samples used by the study. Section 4.3 then explains how pseudo-neighbourhoods were used to create to measure ethnic school segregation within local educational markets. Section 4.4 then describes and justifies the five ethnic groups used by this study. Section 4.5 presents the multi-group H index, which was used to measure ethnic school segregation, and explains why this index was chosen over other multi-group measures. Section 4.5.3 then presents the study's secondary outcome variable, Shared Peer Ethnicity, which was used to measure ethnic isolation. The control and predictor variables are then outlined, in Section 4.6 and 4.7 respectively, before the analytical approaches used in this thesis are described in Section 4.9. Finally, the chapter ends with a discussion on the ethical considerations that arose from this study and the steps taken to minimalize any potential harm that could arise from the completion of this project.

4.1 Data Sources

The table below outlines the datasets and variables that were used in this study.

Table 3: Overview of datasets

Name	**Source**	**Level**	**Public / Restricted**	**Research Questions**	**Description of variables**
School Census Data	Department for Education	School	Public	1 - 5	URN Ethnic profile Percentage of students eligible for free school meals Phase of education Religious denomination Grammar school identifier
Price Paid Dataset (2019)	HM Land Registry	House Sale		4	Sale price for properties sold in England in 2019 Type of property sold Post code of property sold
Lower layer Super Output Area population density (2016)	Office for National Statistics	LSOA		4	Population Density
National Pupil Database extract	Department for Education	Student	Restricted	6	The list of schools that each student wrote on their application form, in order of preference (1-6).

					School offered to parents The school that students attended Student's home postcode Student's ethnicity

Access to the bespoke extract of the National Pupil Dataset was secured through an application to the DfE via the ONS (DR190514.02). As part this process, I completed the ONS's Secure Researcher Training as well as the Information Security Awareness Course provided by the University of Oxford. As the dataset contained information on students' ethnicity, which is "special category data" under GDPR regulations, all output was checked by staff at the ONS before being released.

4.2 Population and samples

This thesis completed two separate analyses and so will describe the populations and samples separately. Initially the section will describe data from the school census dataset that was used to answer research questions 1 - 5 before outlining the sample characteristics of the students in the National Pupil Database extract, which was used to answer research question 6.

4.2.1 School Census Data

The following tables provides information on the four samples from the School Census data used by this study to answer research questions 1 - 5. All figures are based on data from 2019, which was the most recent year analysed by this study. However, equivalent information for all other years can be found in Section 11.17 of the appendix.

Table 4: Describing the criteria used to make each sample and the research question it was used to answer.

Sample	Exclusions	Used in section	Research Question
1	Fee paying schools Specialist Schools Sixth form colleges Nurseries	5.1	What is the ethnic profile of England's school age population and how has it changed between 2006 – 2019?
		5.2	What is the current level of ethnic isolation within England's school system and how has it changed between 2006 – 2019?
2	Sample 1 + Pseudo-neighbourhoods where a single ethnic group constituted more than 90% of the student population.	5.3	What is the current level of ethnic school segregation within England's pseudo-neighbourhoods and how has it changed between 2006 – 2019?
		6.1	What pseudo-neighbourhood characteristics are predictive of ethnic school segregation in England?
		6.2.3	Is there an association between a school's OFSTED rating and their contribution toward ethnic school

			segregation, relative to other schools located in the same pseudo-neighbourhood?
3	Sample 2 + Pseudo-neighbourhoods without a grammar school.	6.2.1	Do academically selective schools contribute more to ethnic segregation than non-selective schools located in the same pseudo-neighbourhood?
4	Sample 2 + Pseudo-neighbourhoods without any religious schools.	6.2.2	Do schools with a religious denomination contribute more to ethnic school segregation than non-religious schools located in the same pseudo-neighbourhood? If so, do schools with particular religious denominations contribute more to ethnic school segregation, on average, than others?

Table 5: Describing the number of students, schools, and pseudo-neighbourhoods in Sample 1 - 4 used by the study (2019)

Sample	Phase	Number of Students	Number of Schools	Number of Pseudo-Neighbourhoods[75]
1	Primary	4,711,429	16,673	2758
	Secondary	3,357,602	3438	534
	Total	8,069,031	20,111	3292
2	Primary	3.790,644	11,974	1824
	Secondary	2,779,740	2768	403
	Total	6,570,384	14,742	2227
3	Secondary / Total	593,673	583	77
4	Primary	3,436,140	10,973	285

[75] The process for creating "pseudo-neighbourhoods" is explain in Section 4.3.

	Secondary	2202413	2210	1600
	Total	5,638,553	13,183	1885

Table 6: Describing the ethnic profile of Samples 1 - 4 (2019)

Sample	White British (%)	Asian (%)	Black (%)	White Other (%)	Other (%)
1	65.9	13.2	8.1	7.3	5.5
2	59.8	15.8	9.6	8.5	6.3
3	62.4	16.9	8	6.8	5.9
4	58.9	16.1	9.9	8.6	6.5

Sample 1 contained information on all mainstream, state-funded, primary and secondary schools in England. Any fee-paying school was removed from the data as the DfE does not hold information on the ethnic makeup of their student body.[76] Special schools and sixth form colleges were also removed to focus on schools that compete in mainstream primary or secondary school local educational markets.[77] The sample was used to answer Research

[76] According to a report released in 2019 by the Independent Schools Council, 7% of pupils in England attended a private school (Stevens, Parkes and Shun-Kai, 2019). Interestingly, a greater proportion of students attending the 1,364 independent schools that were surveyed came from an ethnic minority group than in the state school system. The report found that 66.1% of students attending a private school were White British and 33.9% belonged to an ethnic minority background, with the equivalent figures for state schools being 67.8% and 32.2%, respectively (2019, p. 17).

[77] The full list of schools types included in the sample are Community School, Academy Converter, Voluntary Aided School, Voluntary aided, Infant and Junior School 5-11, , Academy Sponsor Led, Academy Converters, Foundation School, Voluntary Controlled School, Comprehensive all-through 11-18, Comprehensive all-through 11-16, Junior School 7-11 or 8-11, Infant School 5-7 or 5-8, City Academy, Grammar School, Free Schools, First School 5-9, Modern School, Middle School 9-13 (deemed Secondary), Comprehensive Upper School 13-18, Other Secondary School, First and Middle School 5-12, Comprehensive 2 tier Junior 11-14 auto transfer, First School 5-10, Comprehensive 2 tier Senior 14-18 auto transfer, First School 5-8, Comprehensive Upper School 12-18, First School 5-7, Comprehensive Upper 14-18, Middle School 8-12 (deemed Primary), Middle School 10-14 (deemed Secondary), Academy Free Schools, Middle School 10-13 (deemed Secondary), Comprehensive 2 tier, Senior 13-18 opt transfer, Comprehensive 2 tier Senior 13-18 auto transfer,

Question 1 and 2, which described the ethnic profile of England's school aged population and measured ethnic isolation. The sample included 8,069,031 students attending 20,111 schools, nested within an average of 3292 pseudo-neighbourhoods for the year 2019.[78]

Sample 2 removed any schools from Sample 1 that were located in pseudo-neighbourhoods where a single ethnic group constituted more than 90% of the student population.[79] As explained in section 4.5.1.1, this ensured that ethnic school segregation was only measured in pseudo-neighbourhoods that were *These results offer the most up to date description of ethnic school segregation within local educational markets in England*.[80] On average, 1065 pseudo-neighbourhoods were removed from Sample 1 (33%), which, on average, left Sample 2 with 6,570,384 students attending 14,742 schools nested within 2227 pseudo-neighbourhoods in 2019.[81] On average, 99.9% of the excluded pseudo-neighbourhoods were removed because more than 90% of the student population were White British, while one pseudo-neighbourhood, on average, was removed due to the size of their Asian population.[82] The map below shows the locations of pseudo-neighbourhoods that were included or excluded in this second sample. The figure highlights that most excluded pseudo-neighbourhoods were located in rural environments.

Comprehensive Upper School 12-15 or 12-16, Comprehensive Middle and Upper School 10-16, Middle School 9-12 (deemed Primary), Comprehensive 2 tier Senior 14-18 opt transfer.

[78] 58 pseudo-neighbourhoods were excluded for only containing one school (< 1%).

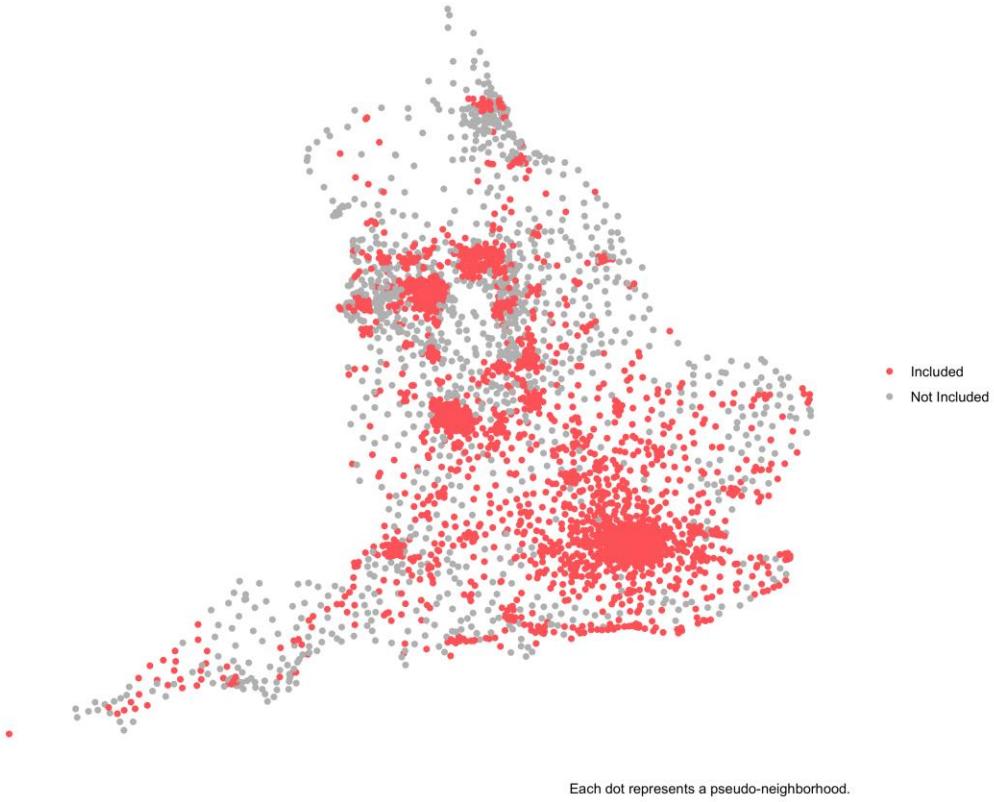
[79] See section 2.4.1b *Evaluating measures of evenness segregation* for more details.

[80] See Section 4.5.1.1 for a discussion on compositional invariance and the reasons for excluding areas with high levels of ethnic homogeneity.

[81] As this thesis does not use a fixed set of neighbourhoods (see Section 4.3) the characteristics of each sample varied slightly between each iteration. All figures reported therefore relate to the average statistics across 100 different sets of neighbourhood boundaries. Reader should note that these figures varied minimally between each iteration.

[82] These are average figures. There were no neighbourhood with a Black, White Other, or Other ethnic majority. The single Asian neighbourhood was not identified to ensure that the study met the ethical requirements, as laid out in Section 4.10.

Figure 6: Mapping the location of pseudo-neighbourhoods included in Sample 2[83]



Finally, Sample 3 and 4 were created to complete two of the school level analysis. Sample 3 removed any pseudo-neighbourhoods from Sample 2 which did not contain at least one grammar school. On average, this sample included 77 pseudo-neighbourhoods with 593,673 students in 583 secondary schools.[84] Similarly, Sample 4 removed any pseudo-neighbourhoods that did not contain any religious schools. This sample, on average, contained 1885 pseudo-neighbourhoods, with 5,638,553 students in 13,183 schools. The locations of pseudo-neighbourhoods in Samples 3 and 4 are outlined in the map below. The

[83] This

[84] The sample for Section 6.2.1 only included secondary schools as no primary schools in England use academic selection as part of their admissions process.

figure highlights that most excluded pseudo-neighbourhoods were located in rural environments.

Figure 7: The location of pseudo-neighbourhoods included in the Sample 3 and 4



4.2.2 National Pupil Database

The following section provides information on the bespoke extract from the National Pupil Database that was used to answer research questions 6a-6e. The original dataset contained information on all 564,488 year six students who had completed an application form in 2016/2017 to express a preference over which secondary school they would attend from Sept 2017. 26,287 students (4.7%) were no longer in the English school system the following school year and were removed from the sample. An additional 4823 (0.8%) students

attended a non-mainstream secondary school in the 2017/2018 school year and were also removed. Finally, 5563 students (1.1%) did not provide their ethnic identity and so were removed from the sample. This left a final sample of 527,775 students (93.5%) who were allocated to 3217 mainstream, state-funded secondary schools.[85]

The simulations were completed using all 527,775 students listed in the dataset. However, after this allocation procedure had been completed, ethnic school segregation was only measured in pseudo-neighbourhoods where no single ethnic group constituted more than 90% of the year 6/7 student population to ensure that ethnic school segregation was only measured in places with appropriate levels of ethnic diversity for the measurement of segregation to be appropriate.[86] The table below summarises the number of students, schools and pseudo-neighbourhoods that were used to *calculate* the level of ethnic school segregation in each simulation.

Table 7: The number students, schools and pseudo-neighbourhoods that were used when calculating ethnic school segregation in each simulation.

Stage of admissions cycle / Simulation	Number of students	Number of schools	Number of Pseudo-Neighbourhoods
Stage 1 of the admissions process - 1st Preference schools	388,704	2277	354
Stage 2 of the admissions process -	388,776	2260	354

[85] As with the school census data, this analysis only explored patterns of ethnic school segregation in relation to mainstream schools in England. For a full list of school types that were considered "mainstream" see Section 4.2.1.

[86] See Section 4.5.1 for more details.

Offers			
Stage 3 of the admissions process - Year 7 school attendance patterns	390,189	2248	352
Simulation 1 – Home-school distance at oversubscribed schools	391,524	2265	355
Simulation 2 – Random allocation at oversubscribed schools	393,269	2257	352
Simulation 3 – Banding with random allocation at oversubscribed schools	373,845	2239	349
Simulation 4 – Banding with distance-based allocation at oversubscribed schools	367,361	2235	347
Simulation 5 – Academic competition at oversubscribed schools	392,268	2262	352
Simulation 6 – Strict home-school Allocation	379,043	2242	345
Simulation 7 – Grammar school system	376,987	2184	354

4.3 Defining the geographic areas of study using k-means clustering

As discussed in Section 2.1.2, the geographic area(s) in which ethnic school segregation is measured is a crucial factor underpinning research in this field. After deciding that all pre-existing government boundaries in England did not offer an effective way of measuring ethnic school segregation within local educational markets, this thesis designed a new

approach.[87] The study used the k-means clustering algorithm (Hartigan and Wong, 1979) to group schools together which were geographically close to one another in order to form "pseudo-neighbourhoods".[88] The collective ethnic profile of the group of schools was then used as the ethnic profile of the geographic area. The formula for k means clustering is listed below, as is the procedure it follows in the context of this research and a diagram that shows how a group of schools would be assigned to a set of clusters.

Equation 1: Formula for K means clustering (Jain, 2010)

Let $X = \{x_i\}, i = 1, \dots, n$ be the set of *n d*- dimensional points to be clustered into a set of K clusters, $C = \{c_k, k = 1, \dots, K\}$. K-means algorithm find a partition such that the squared error between the empirical means of a cluster and the points in the cluster is minimized. Let μ_k be the mean of cluster c_k. The squared error between μ_k and the points in cluster c_k is defined as

$$J(c_k) = \sum_{x_i \epsilon c_k} ||x_i - \mu_k||^2$$

The goal of K-means is to minimize the sum of the squared error over all K clusters.[89]

$$J(c_k) = \sum_{k=1}^{K} \sum_{x_i \epsilon c_k} ||x_i - \mu_k||^2$$

[87] This study was only able to find one article where a clustering algorithm had been used to construct geographic areas for measuring segregation (Östh, Malmberg and Andersson, 2015). The Swedish study used k means clustering to construct "individualized neighbourhoods" to measure residential segregation between visible minorities and white Swedish families. Due to the differences between measuring residential and school segregation this article was not discussed in detail within the main body of the text.

[88] While alternative clustering algorithms exist, k- means clustering is currently "the most widely used" clustering algorithm (Celebi, Kingravi and Vela, 2013) and is comparatively simple (Jain, 2010). Hopefully, employing a simpler, more widely known approach will increase the method's accessibility and encourage its uptake. This does not prevent future research exploring the use of alternative algorithms, or the use of different versions of k-means clustering (Qi et al., 2017; Yu et al., 2018; Zhang, Zhang and Zhang, 2018).

[89] This is known as the "argument minimum", which is represented as $J(c_k)$ in Equation 1.

Table 8: Procedure used by k means clustering to create pseudo-neighbourhoods.

Step	Description
1	Researchers set how many clusters s/he would like the algorithm to make (*k)*.
2	*k* points (centroids) are randomly scattered across a map of England which contains the coordinates of all schools in the sample.
3	Schools are assigned to their nearest centroid, based on the Euclidian distance between the school and the centroid.
4	Each centroid is replotted as the midpoint of all schools assigned to it in step 3.
5	Steps 3 and 4 are then repeated until all schools are consistently assigned to the same centroid. These centroids are then used to define the pseudo-neighbourhoods.[90]

Figure 8: Diagram to show how k-means clustering is used to make pseudo-neighbourhoods.

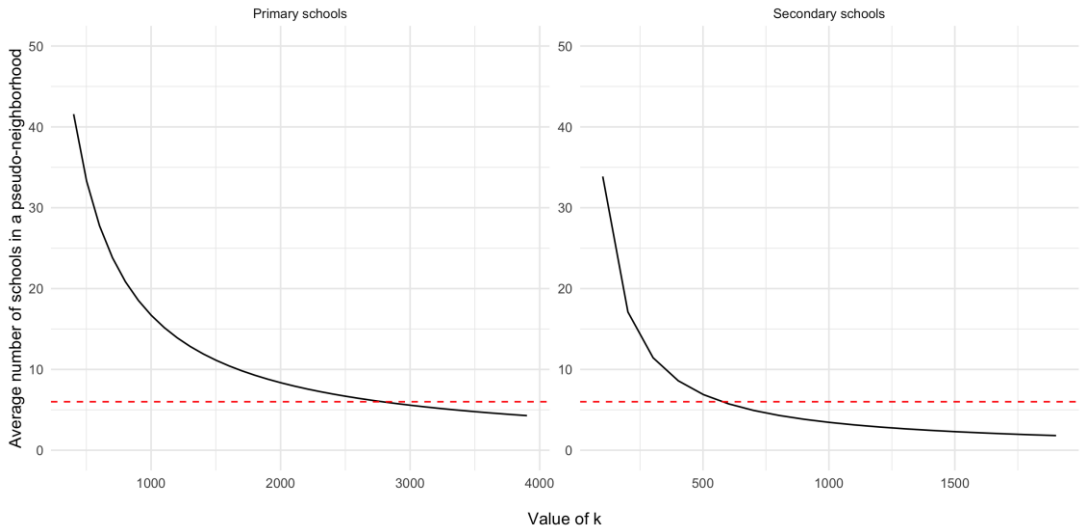


This approach has two main advantages over measuring school segregation within government boundaries. Firstly, by selecting the number of centroids (*k*) in step 1, researchers control the average number of schools in each pseudo-neighbourhood allowing

[90]Any pseudo-neighbourhood containing only one school was removed from the sample. This resulted in 59 schools / pseudo-neighbourhoods, on average, being removed from the sample. This represented a 1.8% reduction in the number of pseudo-neighbourhoods or the removal of 0.3% of schools from Sample 1.

them to make groups of schools which are of a adaptable yet similar size.[91] As shown in Figure 9, the greater the number of centroids (*k)*, the fewer schools each pseudo-neighbourhood will contain and vice versa. This not only gives researchers control over the size of pseudo-neighbourhoods but also forces her/him to justify their choice of *k* which alerts the readers to the scale of the geographic area in which segregation is being measured. As discussed in Section 2.1.2, this is an important feature underpinning any measurement of segregation which can often remain hidden when using government boundaries.[92] Hopefully, this will support the ongoing debate over the ideal scale at which to measure school segregation in England.

Figure 9: The relationship between the value of *k* and the average number of schools within a pseudo-neighbourhood (England, 2019).



Note: The dotted line shows the value of *k* where pseudo-neighbourhoods contain an average of 6 schools.

[91] Thereby meeting the two desirable properties for defining a geographic area in which to measure school segregation. See Section 2.1.2.

[92] See Section 11.13 in the appendix for the results of a simulation analysis which explored the average *H* value when neighbourhoods were formed using different values of *k*.

The second advantage of k-means clustering is that it addresses the zoning problem. By creating multiple sets of pseudo-neighbourhoods, running analyses on each set of boundaries, and then averaging results across each iteration, k-means can produce findings that are insensitive to any single set of pseudo-neighbourhood groupings. As the initial distribution of centroids is random (step 2) each time the algorithm is run a different combination of pseudo-neighbourhoods is created. As *k* is consistent across iterations, the average size of pseudo-neighbourhoods stays constant while different sets of schools are grouped together to form different pseudo-neighbourhoods. Crucially, however, when results are averaged across several iterations the findings are stable which allows researchers to validate previous studies as well as establish trends across time.

This study set k = 2800 for the analysis of primary schools and k = 550 for the analysis of secondary schools as this delivers pseudo-neighbourhoods that, on average, contain six schools.[93] As discussed in Section 2.1.2, this reflects the maximum number of schools parents can list on their child's school application form (HM Government, 2018, P4), which offers a reasonable estimation for the size of local educational markets. As the table below shows, pseudo-neighbourhoods contain a similar number of schools, which suggests that these values of *k* also meet the second desirable property for defining a geographic area for measuring school segregation outlined in Section 2.1.2.[94]

[93] This approach differs to conventional cluster analysis, which normally determines the number of clusters posteriori, based on their predictive power in relation to a given outcome variable. Instead, this study determines the number of clusters a priori, based on a subjective judgement about the size of educational markets. Consequently, the study did not run a conventional "goodness of fit" analyses on these clusters.

[94] See Section A7 in the appendix for more detail. Note, due to the higher density of schools in towns and cities than in the countryside, k-means clustering naturally makes neighbourhoods with a larger number of schools

Table 9: The number of schools contained within pseudo-neighbourhoods in England in 2019, when k = 2800 for primary school pseudo-neighbourhoods and *k* = 550 for secondary school pseudo-neighbourhoods.

Phase of schooling	Minimum	Lower Quartile	Median	Mean	Upper Quartile	Maximum	Total Number
Primary	2	4	5.3	6	8	26.1	2757
Secondary	2	4	6	6.4	8	23.6	534
Total	2	4	6	6.1	8	28	3294

Note: all figures in this table are averaged across 100 iterations. Any pseudo-neighbourhoods with only one school were excluded from the analysis.

Finally, pseudo-neighbourhoods were kept constant when conducting longitudinal analysis. The midpoint of schools located within each pseudo-neighbourhood was calculated based on the groupings using 2019 data and stored. Schools in previous years were then assigned to whichever midpoint/pseudo-neighbourhood they were closest to. This approach ensured that each school was consistently assigned to the same pseudo-neighbourhood, while schools which had closed before 2019 were still included in the analysis for earlier years.[95]

in urban environments. This creates neighbourhoods which reflects the natural differences in sizes of educational markets in across parts of England.

[95] While 55% of schools were part of the school census data for all 10 years of the study, 45% of all schools had one, or more, years of data missing. However, this figure is likely drastically overstate the number of schools which opened/closed as many change their URN number after changing their name or academy status.

4.4 Defining five ethnic groups

This study condensed the 18 ethnic groups listed on the UK census dataset into five broader ethnic groups – White British, White Other, Black, Asian, and Other – as outlined in the table below.[96]

Table 10: Condensing 18 ethnic groups into 5 groups.

Ethnic Grouping (5)	Ethnic group (18)	Percentage of the larger ethnic grouping (%)
White British	White British	100
White Other	White Other	90.6
	White Irish	3.8
	Gypsy or Irish Traveller	4.6
Black	Black African	48.2
	White and Black Caribbean	18.9
	Black Caribbean	13.4
	Black other	9.4
	White and Black African	10.1
Asian	Pakistani	33.9
	Indian	24
	Bangladeshi	13.4
	Other Asian	14.3
	Chinese	3.4
Other	Other Mixed	40.3

[96] See Section 2.1.3.

	Other	35.9
	Unclassified	23.8

Note: Percentages may not sum to 100% due to rounding.

These groups broadly align with how ethnic identities are regularly conceived of at a national level, make results easier to describe and help to reduce the risk of small unit bias. They also align with how two previous UK research articles conceptualised ethnic groupings (Jones *et al.*, 2015; Leckie and Goldstein, 2015).[97] Additionally, a sensitivity analysis conducted by this thesis found that measures of segregation based on 5 or 18 ethnic groups had a correlation of 0.90, suggesting that the aggregation of ethnic groups had a minimal impact on the overall results.[98]

However, this decision is ultimately subjective, and these five groups will not reflect important local dynamics within parts of England, where for example, distinctions between Indian and Pakistani communities, or African and Caribbean students, may be important. Unfortunately, this issue is symptomatic of working at a national scale where there is no set of ethnic groupings that are relevant across the whole country. Indeed, while using 5 large ethnic groups will be inappropriate in some parts of England, using all 18 would be inappropriate in others.

[97] See Section 11.8 in the appendix for details on the level of upward bias found in the English school system.
[98] See Section 11.7 in the appendix for further details.

4.5 Outcome Variables

The following section presents the outcome variables used by this study. Section 4.5.1 presents the multi-group *H* index which this study used to measure evenness segregation. Section 4.5.2 then explains how *H* was decomposed to create a school level outcome variable which measured each school's relative contribution towards their pseudo-neighbourhood's level of ethnic school segregation (H_j). Finally, Section 4.5.3 presents this study's measure of ethnic isolation, Shared Peer Ethnicity, and explains why this measure was chosen over more traditional area-based measures.

4.5.1 H Index – Ethnic School Segregation

This study is the first British research to use a multi-group measure of ethnic school segregation to map patterns of ethnic school segregation across England.[99] Starting from the premise that a multigroup measure is required, for the reasons outlined in Section 2.1.3, the following section presents a list of desirable properties for a segregation index to exhibit before demonstrating that *H* is the index that best meets these criteria. However, readers should note that different segregation indexes often produce similar results in practice. For example, an analysis conducted by this study found that a multigroup H index had an average correlation of 0.93 with a multigroup *D* index and 0.62 with a two-group *D* Index when using English school data.[100]

[99] Readers should note that multi-group measure of segregation have been regularly used to evaluate ethnic residential segregation (Jones *et al.*, 2015; Johnston, Jones, *et al.*, 2016; Jones, Johnston, *et al.*, 2018)
[100] See Section 11.9 in the appendix for further details. It is fairly common for researcher to repeat their analysis with more than one index (Fiel, 2015)

4.5.1.1 Five desirable properties for a segregation index to exhibit

Unfortunately, despite over 70 years of debate, discussions over which index offers the best measure of segregation remains contested (Peach, 1975, p. 3). Over 25 segregation indexes exist (Grannis, 2002) and none are considered "ideal" (Gorard, 2007, p. 674). Thankfully, academics normally agree on five properties which should be used to evaluate indexes (James and Taeuber, 1985; Grannis, 2002; Allen and Vignoles, 2007). When applied to ethnic school segregation, these properties are:

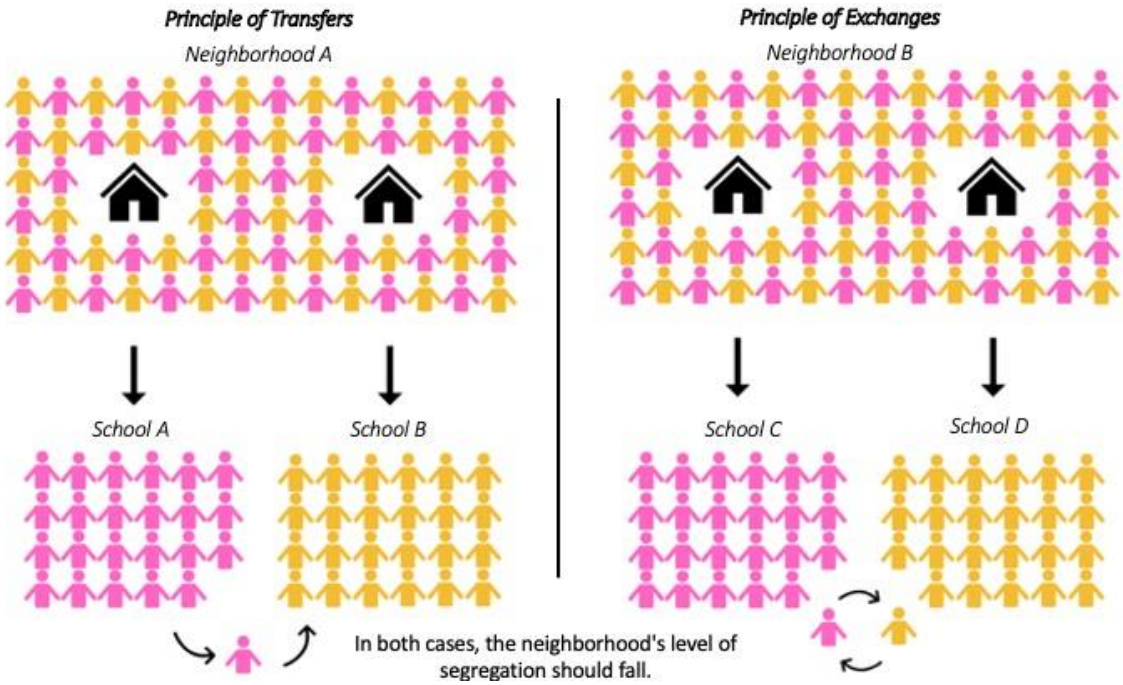
1). **Organisational Equivalence** – Measures of segregation should not be influenced by the number of schools in a neighbourhood.

2). **Size Invariance** – Measures of segregation should not be influenced by the total number of students in a neighbourhood.

3). **The principle of transfers** – If a student moves from a school where their ethnic group is over-represented, to one where their group is under-represented, or simply less over-represented, a neighbourhood's segregation should fall.

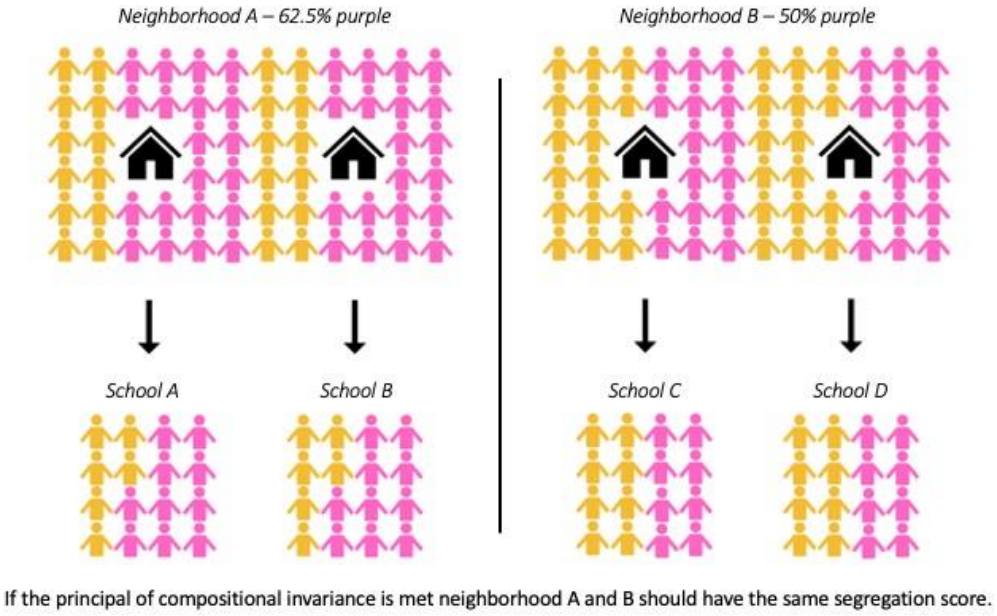
4). **The principal of exchanges** – Similar to the principle of transfers; if a student who attends a school where their ethnic group is overrepresented is exchanged with a student from a different ethnic group, who also attends a school where their ethnic group is over-represented, then a neighbourhood's level of segregation should fall.

Figure 10: The Principle of Transfers and Exchanges



5). **Compositional invariance** – Segregation scores should only reflect the relative distribution of ethnic groups between different schools, not the ethnic composition of the neighbourhood. For example, a fall in the percentage of White British students in every school within a neighbourhood from 62.5% to 50%, as shown in Figure 11, should not change the measure of school segregation. If a segregation index meets this criterion, researchers can compare patterns between areas with different ethnic profiles.

Figure 11: Compositional Invariance



Of the five properties, compositional invariance remains the biggest hurdle for segregation indexes. All indexes are influenced by the proportions of ethnic groups within the neighbourhood, with segregation normally higher in areas with greater ethnic diversity. This makes it challenging to compare areas with different ethnic profiles or to plot changes in segregation through time when an areas ethnic profile has changed (Harris *et al.*, 2013, p. 2282).

Academics need to engage more with the issue of compositional invariance. Currently most research only excludes areas where 95% of the population belongs to the same ethnic group (Zoloth, 1976; James and Taeuber, 1985; Reardon, Yun and Eitle, 2000), while some only removes areas where a group constitutes more than 99% of the population (Fiel, 2015). This still leaves researchers comparing neighbourhood's with radically different ethnic profiles. While this may have been permissible when using binary measures of segregation, an

increasingly diverse population makes this approach dubious. Consequently, this thesis is more conservative; filtering out any pseudo-neighbourhoods where a single ethnic group is larger than 90% of the student body, as outlined in Section 4.2.1.

4.5.1.2 The H Index

The *H* index will be used as the study's measure of ethnic segregation as it meets more desirable properties than any other multigroup index (Reardon and Firebaugh, 2002). The formula for *H* is listed below.[101]

Equation 2: The H Index (Theil and Finizza, 1971)

$$H = 100 * \sum_{j=1}^{J} \sum_{m=1}^{M} \frac{1}{E} \frac{t_j}{T} P_{jm} \ln \left(\frac{P_{jm}}{P_m}\right)$$

Where:

$j = a\ given\ schoool$

$m = a\ given\ ethnic\ group$

$P_{jm} = proportion\ of\ school\ j\ which\ belong\ to\ ethnic\ group\ m$

$P_m = proportion\ of\ students\ in\ the\ pseudo - neighborhood\ which\ belong\ to\ ethnic\ group\ m$

$t_j = number\ of\ students\ in\ the\ pseudo - neighborhood\ which\ attend\ school\ j$

$T = total\ number\ of\ students\ in\ the\ area$

$E = Theil's\ diversity\ (Entropy)\ index\ where\ E = \sum_{m=1}^{M} P_m \ln \left(\frac{1}{P_m}\right)$

[101] For a breakdown of *H* see section A4 in the appendix.

The most important feature of the H index is its disproportionality function - $\ln(\frac{P_{jm}}{P_m})$. This measures the difference of logs of proportions between the proportion of ethnic group *m* in school *j* to the proportion of ethnic group *m* in the pseudo-neighbourhood. For example, if 80% of school *j*'s students were White British and 40% of the students in the pseudo-neighbourhood were White British then the disproportionality function would be calculated as $\ln\left(\frac{80}{40}\right) = 0.69$.[102] The result of the disproportionality function is then weighted to reflect the proportion of the school which belongs to ethnic group *m* (P_{jm}) and the proportion of students in the pseudo-neighbourhood who are educated by school *j* $(\frac{t_j}{T})$. Lastly, the index adds a control for the ethnic diversity of the pseudo-neighbourhood$(\frac{1}{E})$ to avoid places with greater diversity inherently being measured as more segregated.[103]

This calculation is repeated for each ethnic group and then added together to create a school level measure (H_j).[104] This captures the total divergence between the ethnic profile of the school and the ethnic profile of other schools in the same pseudo-neighbourhood. To calculate the final *H* index, researchers simply add all these school-level divergences together ($H = \sum_{j=1}^{J} H_j$).[105]

Initially formulated in the 1970s (Theil and Finizza, 1971) *H* is based upon a branch of mathematics called Information Theory, which models the flow of information (Alencar, 2015). Bounded between 0 and 1, or in this thesis 0 and 100, its output can be interpreted

[102] If the proportion in the school and neighbourhood are equal, then the disproportionality function would be $\ln(1)$ which equals zero.
[103] For more details on Theil' s diversity (Entropy) index see Section 11.10.
[104] This is step 3 in Figure 3.
[105] This is step 4 in Figure 3.

as “the proportional increase in expected information about [a student’s ethnicity] that occurs when learning about the school that the student attends” (Mora and Castillo, 2011, p. 172). While a number of articles in the 1980s commended its mathematical properties (James and Taeuber, 1985; White, 1986) it’s complicated formula and interpretation meant that it was not until the turn of the century, following the racial diversification of the USA, that *H* began to be used more frequently (Reardon, Yun and Eitle, 2000, p. 352).

The binary Index of Dissimilarity (*D*), which measures segregation between two ethnic groups, has traditionally been preferred (Duncan and Duncan, 1955). [106] *D* has a straightforward formula and a comparatively easy interpretation. Bounded between 0 – 100, its output measures the percentage of minority students who would need to move schools in order to achieve an even spread of students across a neighbourhood’s schools. It also meets most of the criterions for a good measure of segregation (Hutchens, 2004; Allen and Vignoles, 2007). In particular, its compositional invariance marks it out against other indexes and its enduring popularity means that measuring trends across time was easier (Lieberson and Carter, 1982, p. 296).[107]

Unfortunately, when adapted to a multigroup index, *D* becomes far less appealing. Its interpretation changes and its compositional invariance, the main advantage it holds over other two-group indexes, does not hold in a multigroup setting. It also fails to meet the principle of exchanges, transfers, additive organisational and group decomposability

[106] See Section A5 in the appendix for the formula of a two group D index.

[107] *D* is not a perfect measure of segregation because it only satisfies a “weak” principle of transfers (Reardon and Firebaugh, 2002). A transfer or exchange of a student from a school where their ethnic group is over-represented to one where they are still over-represented, but to a lesser degree, will cause no change in *D*.

(Reardon and Firebaugh, 2002). While *H* is not perfect, it meets more of the desired properties listed in the table below.[108]

Table 11: Comparing the properties of the H and D Index

Properties	H Index	D Index
Organisational equivalence	✓	✓
Size Invariance	✓	✓
Transfers		
2 Group	✓	✗
Multi Group	✓	✗
Exchanges		
2 Group	✓	✗
Multi Group	✓	✗
Compositional Invariance		
2 Group	✗	✓
Multi Group	✗	✗
Additive Organisational Decomposability		
2 Group	✓	✗
Multi Group	✓	✗
Additive Group Decomposability	✓	✗

Additionally, as *H* is based on a logarithmic scale, transfers or exchanges in situations of extreme segregation result in larger changes in H values than the same transfer or exchange

[108] See Section 8.6.2 for a discussion of the limitations of using H.

in environments with lower levels of segregation. *H* therefore incorporates "the notion of diminishing marginal payoffs to desegregation" (Zoloth, 1976, p. 291), which reflects how most academics view the cost of segregation to society. Consequently, *H* was used as this study's measure of ethnic school segregation.

4.5.2 Measuring the contributions of schools to ethnic school segregation (H_j)

The study also created a school-level outcome variable to explore whether certain types of schools contributed more towards ethnic school segregation than others (RQ 5a – 5d). More specifically, the *H* index was decomposed to create a series of school level statistics which measured the extent to which each school's ethnic profile diverged from their pseudo-neighbourhood's average ethnic profile (H_j). The following section describes how these figures were calculated, as well as explaining why it was important to conduct a school level analysis.

As explained previously, the *H* index is calculated by finding the extent to which each school's ethnic profile diverges from the neighbourhood's ethnic profile (H_j) and summing these divergences together to form a neighbourhood level measure (*H*). Each school level divergence (H_j) can therefore be used as an outcome variable to evaluate the school's "contribution" towards the level of ethnic school segregation in their neighbourhood (*H*).[109] The formula to calculate a school's H_j value is listed below. This formula replicates the core part of the H index except that the school's contribution is no longer weighted in relation to

[109] This relates to step 3 in Figure 3.

how many students it educates $\left(\frac{t_j}{T}\right)$. This ensures that schools with a greater number of students do not inherently contribute more to ethnic school segregation than those with fewer students. The formula is listed below.

Equation 2: H_j

$$H_j = \sum_{m=1}^{M} \frac{1}{E} P_{jm} \ln \left(\frac{P_{jm}}{P_m}\right)$$

Where:

$j = a\ given\ schoool$

$m = a\ given\ ethnic\ group$

$P_j = proportion\ of\ school\ j\ which\ belong\ to\ ethnic\ group\ m$

$T = total\ number\ of\ students\ in\ the\ pseudo - neighborhood$

$E = Theil's\ diversity\ (Entropy)\ index$

As these figures do not have a direct interpretation and are influenced by the characteristics of the pseudo-neighbourhood, the study standardised each value of H_j in relation to the other schools located in the same pseudo-neighbourhood using a Z score normalization.[110] This brought two benefits. Firstly, it gave H_j a substantive interpretation, where values can be understood as the standardized contribution of school *j* to their pseudo-neighbourhood's level of ethnic school segregation, relative to other schools in the same pseudo-neighbourhood. A school with an $H_j = 1$, for example, makes a contribution to their pseudo-

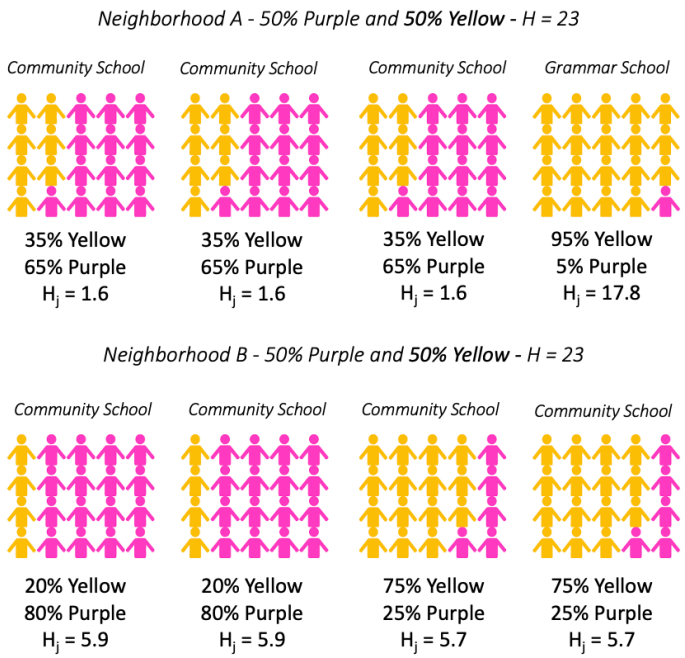
[110] This was completed the scale() function in R's base package.

neighbourhood's ethnic school segregation that is one standard deviation larger than the average contribution of other schools in the same pseudo-neighbourhood. In contrast, a school with an H_j of -2 makes a contribution which is two standard deviations lower than the average contribution made by schools in the same pseudo-neighbourhood. Secondly, by standardising the values of H_j, the study removed the risk that pseudo-neighbourhood differences drive school level figures, thereby creating a more robust measure.[111]

Completing this additional, school level analysis also reduced the risk that the study's findings would suffer from the ecological fallacy problem, where trends observed at higher levels of analysis (pseudo-neighbourhoods) are assumed to reflect those occurring at lower levels (schools). The figure below provides an illustration of this problem in the context of this research. As Neighbourhood A and B have the same level of ethnic school segregation (*H* = 23), a neighbourhood level regression analysis would conclude that there is no difference in ethnic school segregation between neighbourhoods that contain a grammar school vs those that do not. However, this would overlook the fact that the ethnic profile of the grammar school in Neighbourhood A diverges from the ethnic profile of the neighbourhood by a far greater extent than any of the community schools located in the same neighbourhood. In contrast, if a researcher had compared the contributions made by schools to their neighbourhood's level of ethnic school segregation (H_j), this substantial difference would have been identified.

[111] To ensure that the standardization of H_j did not unduly influence the school level analysis, the thesis reran the meta-regression models with two different versions of the H_j values as outcome variable. The first set of models used the untransformed raw H_j values while the second set ranked each school's H_j value, relative to other schools in the same pseudo-neighbourhood, and used this as its outcome variable. Reassuringly the results were substantively similar to those reported in the main body of the text, as described in Section 11.24 in the appendix.

Figure 12: Explaining what a school's "contribution to ethnic school segregation" means



4.5.3 Measuring ethnic isolation - Shared Peer Ethnicity (SPE)

This thesis also employed a measure of ethnic isolation to provide a supplementary description on the level of ethnic "separateness" in the English school system, as is common practice in the field (Harris and Johnston, 2020). As mentioned in Section 2.1.1.1, measures of ethnic isolation assess the potential space for inter-group contact within schools. This thesis described levels of isolation using a student level variable called Shared Peer Ethnicity (SPE) (Mitchell, 2017). SPE calculates the percentage of a student's school that comes from the same ethnic group. For example, a White British student attending a school where 80% of the school are also White British and 20% are Black, would have an SPE of 0.8. By

extension, a Black student attending the same school would have a SPE of 0.2. The advantage of SPE is that it offers a clear insight into the likely experience of students in schools across the country. By comparing SPE for students across a range of demographic characteristics, researchers can compare the level of inter-group contact that student from different ethnic groups are likely to experience. This insight would not be possible by traditional measures of Isolation, which aggregates statistics to geographic areas.[112]

4.6 Control variables

The following section describes two control variables that were used by this thesis. Details of how each variable was created, along with the justification for their inclusion, are discussed in the subsections below.

Table 12: Descriptive statistics for the study's control variables

Control Variable	Min	Lower Quartile	Mean	Median	Upper Quartile	Max	Missing/loss of data (%)
Number of schools in pseudo-neighbourhood	2	4	6.64	6	8.03	27.3	0
Region	See Table 13 for the equivalent figures						< 0.5

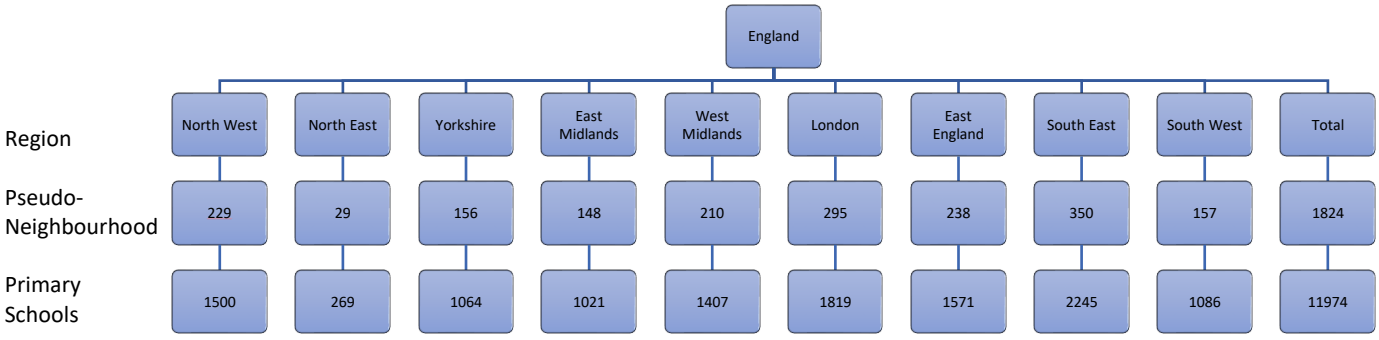
[112] The most commonly used Isolation Index is listed below (Lieberson and Carter, 1982), though others do exist (Wong and Shaw, 2011). See A9 in the appendix for the formula.

4.6.1 Assigning pseudo-neighbourhoods to regions and cities

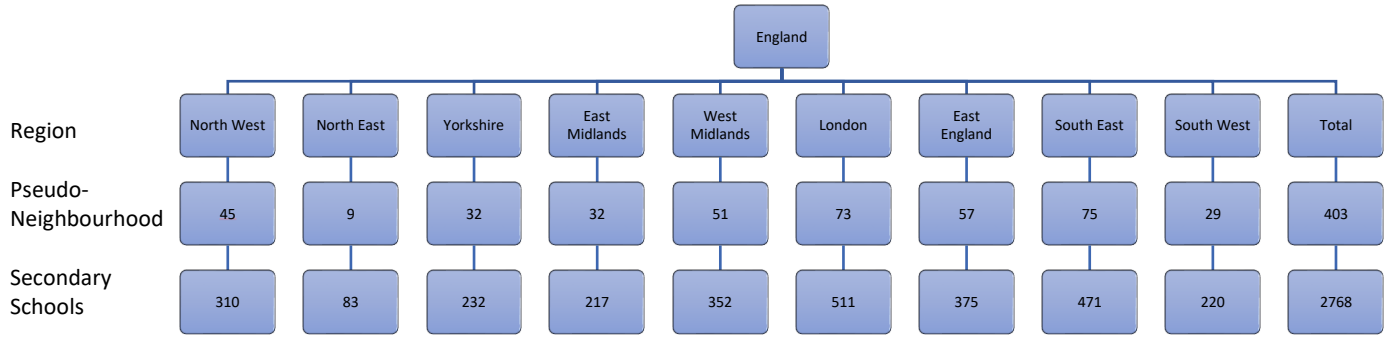
This thesis regularly refers to the nine regions of England, as defined by the Office for National Statistics. These are the South West, South East, London, East of England, West Midlands, East Midlands, North West, North East and Yorkshire and the Humber. The study determined which region a pseudo-neighbourhood was located in based on whichever region the majority of its schools were in. As regions are geographically large, 94% of pseudo-neighbourhoods only had schools located within a single region, which made the aggregation process relatively straightforward. Any pseudo-neighbourhood where schools were evenly split between two or more regions were removed from the sample, which resulted in a small loss of data (<0.5%). The diagram below outlines the average number of pseudo-neighbourhoods and schools within each region across 100 iterations based on sample 2.

Table 13: The number of pseudo-neighbourhoods and schools in each region in Sample 2

Primary School Sector



Secondary School Sector



Note: These figures are averaged over 100 iterations and only include ethnically diverse pseudo-neighbourhoods.

On several occasions, the study assigned pseudo-neighbourhoods to their city or urban dwelling, to help describe more localised trends. The thesis used the same method to link pseudo-neighbourhoods to cities as was used to assign pseudo-neighbourhoods to regions.

Assigning pseudo-neighbourhoods to regions and cities made it far easier to describe geographic trends. Additionally, regional control variables were added to pseudo-neighbourhood level regression models to ensure that national findings were robust to any regional differences.

4.6.2 Number of schools included in the pseudo-neighbourhood

As measures of segregation are influenced by the number of schools within each geographic area the study included a control variable for the number of schools within each pseudo-

neighbourhood.[113] The variable was standardised using a normal score transformation to address a positive skew in the data (Conover, Tercero-Gómez and Cordero-Franco, 2017).

4.7 Predictor variables

The following section describes the predictor variables used in this study. Table 14 outlines the distribution of the school level variables that were used to make the pseudo-neighbourhood level predictor variables, while Table 15 describes the distribution of the pseudo-neighbourhood level variables themselves. The section then provides a description of how the study calculated a pseudo-neighbourhood's population density, free school meal eligibility (FSM), average house price, average Ofsted rating and the proportion of schools which were religious as well as whether a pseudo-neighbourhood contained a grammar school.

Table 14: Descriptive statistics of the *school level* variables used to make pseudo-neighbourhood level predictor variables

Variable	Phase	Min Value	Lower Quartile	Mean	Median	Upper Quartile	Max Value	Missing Data
Population density	Secondary school	9.31	1566	3832	3046	4689	30801	0
	Primary school	8.21	1247	4313	3554	5727	44059	0
House price (£ Thousands)	Secondary school	34	141	298	226	350	4575	0.42
	Primary school	30	138	290	223	339	7563	6.45

[113] See section 4.9.2 for more details.

	Secondary school	0	0.07	0.15	0.13	0.21	0.62	0
FSM (%)	Primary school	0	0.07	0.16	0.13	0.22	0.69	0
	Secondary school	0	2	1.93	2	2	3	13.9
Ofsted rating	Primary school	0	2	2.02	2	2	3	10.8
Grammar school status (1 = Grammar)	Secondary school	0	0	0.05	0	0	1	0
Religious denomination of a school	See Table **18** for the equivalent figures							0

Note: All figures are mean averages across 100 iterations and were based on Sample 2.

Table 15: Descriptive statistics *of pseudo-neighbourhood level* predictor variables

Variable	Phase	Min Value	Lower Quartile	Mean	Median	Upper Quartile	Max Value	Missing Data (%)
Population density	Secondary school	167	2100	3630	2820	4000	21400	0
	Primary school	12	1630	4050	3340	5060	25500	0
House price (£ Thousands)	Secondary school	62	151	305	235	362	2860	0
	Primary school	49	150	307	239	358	4000	0
Contain a grammar school	Secondary school	0	0	0.192	0	0	1	0
FSM (%)	Secondary school	3.01	8.48	13.8	12.5	17.5	40.6	0
	Primary school	1.26	8.63	15	13.3	19.8	50.3	0

Average Ofsted rating	Secondary school	0.29	1.72	1.94	2	2.19	3	6.7
	Primary school	0.346	1.87	2.03	2	2.2	3	26.3
Percentage of religious schools in the pseudo-neighbourhood (%)	Secondary school	0	0	18.9	16.7	29.6	89.8	0
	Primary school	0	19.3	35.3	33.3	50	100	0
Ethnic diversity	Secondary school	27.6	39.2	58.1	55.4	75.2	99.8	0
	Primary school	25	36.1	56.4	52.6	75.6	99.4	0
Number of schools in the pseudo-neighbourhood	Secondary school	2	4.49	6.91	6.16	8.77	23.2	0
	Primary school	2	4	6.58	6	8	26.1	0

Note: All figures are mean averages across 100 iterations and were based on Sample 2.

Table 16: Correlation matrix for pseudo-neighbourhood level predictor variables

Variable	Phase	H Index	Population density	House price	FSM	Religious schools in pseudo-neighbourhood	Average Ofsted rating	Number of schools in the pseudo-neighbourhood	Ethnic Diversity Index
H Index	Primary School	1	0.07	-0.13	0.17	0.08	-0.08	0.19	0.07
Population density		0.07	1	0.53	0.46	-0.19	0.15	0.16	0.64
House price		-0.13	0.53	1	-0.04	0.08	0.27	-0.09	0.4
FSM		0.17	0.46	-0.04	1	-0.15	-0.21	0.14	0.32
Percentage of religious schools in the		0.08	-0.19	0.08	-0.15	1	0.05	-0.13	-0.23

pseudo-neighbourhood (%)									
Average Ofsted rating		-0.08	0.15	0.27	-0.21	0.05	1	-0.05	0.13
Number of schools in the pseudo-neighbourhood		0.19	0.16	-0.09	0.14	-0.13	-0.05	1	0.11
Ethnic Diversity Index		0.07	0.64	0.4	0.32	-0.23	0.13	0.11	1
H Index	Secondary School	1	0.11	-0.14	0.33	0.16	-0.11	0.34	0.29
Population density		0.11	1	0.7	0.57	0.24	0.25	0.13	0.59
House price (£ Thousands)		-0.14	0.7	1	0.16	0.21	0.37	-0.05	0.49
FSM (%)		0.33	0.57	0.16	1	0.21	-0.18	0.18	0.4
Percentage of religious schools in the pseudo-neighbourhood (%)		-0.11	0.25	0.37	-0.18	0.08	1	-0.04	0.25
Average Ofsted rating		0.16	0.24	0.21	0.21	1	0.08	0.12	0.26
Number of schools in the pseudo-neighbourhood		0.34	0.13	-0.05	0.18	0.12	-0.04	1	0.19
Ethnic Diversity Index		0.29	0.59	0.49	0.4	0.26	0.25	0.19	1

Note: These figures are average results over 100 iterations

4.7.1 Population density

The population density of each neighbourhood was calculated using LSOA level data produced by the Office for National Statistics. Each school was assigned a population density, based on whichever LSOA they were located in. Pseudo-neighbourhood level data was then calculated using the mean average of all schools located within the same pseudo-neighbourhood. The variable was standardised using a normal score transformation due to a positive skew in the data.

4.7.2 Free School Meal Eligibility (FSM)

The rate of FSM in a neighbourhood was calculated as the mean FSM rate for the schools in a neighbourhood, weighted by the size of the school population. To be eligible for free school meals a student's parent must be enrolled in a government scheme to support people on low incomes, which makes it a commonly used measure of socio-economic disadvantage. [114] The variable was standardised using a normal score transformation due to a positive skew in the data

[114] This includes the following programmes - Income Support, Income-based Jobseeker's Allowance, Income-related Employment and Support Allowance, the guaranteed element of Pension Credit, Child Tax Credit and/or Working Tax Credit run-on. For a discussion on the merits of FSM as a measure of socio-economic disadvantage, see Section 11.15 in the appendix.

4.7.3 Average house price

Table 17 outlines how this study estimated the average house price within a pseudo-neighbourhood.[115] The variable was then standardised using a normal score transformation due to a positive skew in the data.

Table 17: Process for estimating the house price around a school.

Step	Process
1	Download the Price Paid Dataset from the UK land registry containing the price paid for every property sold in England in the most recent year (HM Land Registry, 2019).[116]
2	Filter the dataset to only include terraced houses.[117]
3	Link each terrace house to the nearest primary and secondary school based on the shortest Euclidean distance.
4	Calculate the median price paid for all the terrace houses assigned to each school.[118] These statistics provide an estimate for the average house price in the area surrounding each school.[119]

[115] It is worth highlighting that the figures produced using this approach had a correlation of 0.84 with the method used in the pilot study of this thesis, which estimated house prices by assigning each school an average house price based on an LSOA level dataset before averaging findings across all schools in a pseudo-neighbourhood ($p<0.001$). Hopefully this should reassure readers of its validity. For further details, see Section 11.14 in the appendix.

[116] This dataset is publicly available on the HM Land Registry website.

[117] The types of residential properties listed in the dataset were flat, terrace, semi-detached and detached. Terraced houses were chosen to calculate the average house price of a pseudo-neighborhood because, out of the four property types, they were the most common properties being bought and sold. This meant that the pseudo-neighborhood level estimates were calculated using the largest number of comparative data points with no missing data at a pseudo-neighbourhood level, see Table 15. Second, the thesis hypothesized that terrace housing was likely to be a house type that is most commonly bought by young families with school aged children, which hopefully makes it particularly relevant to dynamics within the school system. However, as this is a hypothesis, further research would be needed to validate this claim

[118] The median was preferred over the mean, to ensure that positive skews in the distribution of house price data did not lead to upwardly biased estimates of housing costs surrounding a school.

[119] The median was preferred over the mean as the data was often positively skewed.

5	To find the average housing costs for a neighbourhood, calculate the mean average house price of each school's surrounding area for the schools included in the neighbourhood.[120]

4.7.4 Academically selective admissions policies (grammar schools)

Schools were defined as having an academically selective admissions policy if they were listed as a "Grammar School" in the "Type of Establishment" column in the school census dataset. For the pseudo-neighbourhood level analysis in Section 6.1, the study created a binary variable to distinguish between pseudo-neighbourhoods that contained a grammar school and those that did not. This approach was adopted instead of using a continuous measure as there were only a limited number of pseudo-neighbourhoods which contained a grammar school (n = 77).

4.7.5 Religious schools

The study condensed the 45 religious affiliations listed in the "Religious Denomination" column into seven lager groupings, as described in the table below. This aggregation allowed for a more parsimonious description of trends, particularly in relation to the school level analysis described in Section 6.2.

[120] The mean was preferred over the median as the data was normally distributed.

Table 18: Religious groupings used in this study (2019)

Religious Grouping	Number of schools	Percentage of schools	Denominations included within the religious groupings
None	8570	65.0	None, does not apply
Church of England	2866	21.7	Church of England, Anglican, Church of England/Christian, Anglican/Christian, Anglican/Church of England
Catholic	1538	11.7	Roman Catholic
Other Christian	104	0.8	Christian, Methodist, Free Church, Seventh Day Adventist, Church of England/Methodist, Congregational Church, Church of England/Free Church, Christian/Evangelical, Protestant/Evangelical, Church of England/Roman Catholic, Moravian, Quaker, United Reformed Church, Anglican/Evangelical, Church of England/Evangelical, Christian/Methodist, Roman Catholic/Church of England, Roman Catholic/Anglican, Christian/non-denominational, Protestant, Christian Science, Church of England/United Reformed Church, Reformed Baptist, Greek Orthodox, Inter-non- denominational, Plymouth Brethren, Christian Church, Church of England/Methodist/United Reform Church/Baptist
Jewish	49	0.4	Jewish Haredi, Jewish Orthodox, Jewish
Muslim	34	0.3	Islam, Sunni, Deobandi
Other	25	0.2	Multi-faith, Buddhist, Sikh, Hindu

For the pseudo-neighbourhood level analysis, the study calculated the proportion of schools within each pseudo-neighbourhood that were religious before standardising the scores using a normal score transformation. Unlike the subsequent school level analysis, this did not make any distinction between schools with different religious denominations.

4.7.6 OFSTED rating

To calculate an average OFSTED rating for each pseudo-neighbourhood the study first converted each school's rating into a numeric value. Schools that received an "Outstanding" rating were assigned a 3, "Good" a 2, "Requires improvement" a 1, "Serious Weaknesses" or "Special Measures" a 0. Pseudo-neighbourhood level statistics were then calculated by finding the mean value across all schools located within each pseudo-neighbourhood. The values were then standardised in relation to all pseudo-neighbourhoods included in the sample.

For the school level analysis completed in Section 6.2.3 the study used each rating as a binary variable. This created four variables – "Outstanding", "Good", "Requires Improvement" and "Special Measures or Serious Weaknesses"

4.8 Variables used in the simulation

The following sub-section outlines the distribution of the two variables used in the simulation analysis, described in Section 4.9.3, before explaining how each variable was created. It then describes how parents from different ethnic groups completed their child's application form.

Table 19: Variables used in the simulation analysis

Variable	Min	Lower Quartile	Mean	Median	Upper Quartile	Max	Missing data (%)
A School's Estimated Published Admission Number (PAN)	18	134	179.6	178	220	426	0
A Student 's KS2 Academic Results	80	100.8	104.8	105.2	109.5	120.0	2.1

4.8.1 KS2 Academic Results

Simulation 5 and 7, as described later in Section 4.9.3, considered a student's academic abilities at the end of year 6 when allocating them to different secondary schools. The study estimated students' academic abilities at the end of year 6 by firstly calculating their English KS2 Performance, using the average of their "Reading" and "Grammar, Punctuation and Spelling Scores" as listed in the National Pupil Database. The study then combined this English score with the student's Maths score using a simple mean average, to create a KS2 Academic Performance score which was equally weighted between English and Maths.

4.8.2 Estimated Published Admissions Numbers

The study needed to estimate the number of pupils that each school could feasibly accommodate. Unfortunately, there is no national level database that collected these figures, known as Published Admissions Numbers (PAN), for schools in England.

Consequently, the study estimated them based on the mean number of Year 7 students that each school had enrolled in the preceding two school years for which this study had data for (2016/17 and 2017/18).

4.8.3 Preference Data

The simulations were built using data collected from the application forms of parents applying to secondary schools in 2016/2017. Two fictious rows from dataset are provided in Table 20 below, so readers can visualise how the dataset is structured, while

Table 21 outlines how parents from different ethnic groups completed their forms. The number of schools that parents were allowed to express ranged from 3 to 6 depending on which Local Authority they lived in.

Table 20: Explaining data structure of the student preference data

Student ID	**1st Choice**	**2nd Choice**	**3rd Choice**	**4th Choice**	**5th Choice**	**6th Choice**	**School Offered**	**School Attended**
Stu 1	Sch 2	Sch 5	Sch 1	Sch 9	Sch 6	Sch 8	Sch 5	Sch 2
Stu 2	Sch 7	Sch 1	Sch 3	Sch 6	Sch 8	Sch 11	Sch 12	Sch 1

Note: this data is entirely fictious and is only intended to give readers an idea of how the dataset is structured.

Table 21: Describing the extent to which parents from different ethnic groups completed their child's application form in full.

Ethnic Group	Percentage of parents that used their child's application to express a…					
	1st Preference School	2nd Preference School	3rd Preference School	4th Preference School	5th Preference School	6th Preference School
Asian	99.8	81.0	64.2	34.7	23.5	15.7
Black	99.8	83.2	67.9	40.8	26.9	19
Other	99.7	76.7	58.7	32.0	20.5	14.1
White British	99.7	58.7	35.7	10.2	4.2	2.3
White Other	99.5	72.9	53.9	26.7	15.5	10.5
All	99.7	65.1	44.0	17.6	9.8	6.4

4.9 Modelling and analysis techniques

The following section describes how the thesis conducted its analysis. Section 4.9.1 outlines the types of statistics used to describe the ethnic profile of England's school age population, the levels of ethnic isolation and patterns of ethnic school segregation in Chapter 5. This section also explains three sensitivity checks it used to ensure that ethnic school segregation was measured reliably over time. Section 4.9.2 then outlines the meta regression models it built to identify possible associations between the predictor variables and outcome variables, both at a pseudo-neighbourhood and school level in Chapter 6. Finally, Section 4.9.3 explains how the simulation analyses were conducted, as reported in Chapter 7.

4.9.1 Descriptive statistics and sensitivity checks

Common descriptive tools, such as median averages, graphs and maps were used to describe the different ethnic profiles of pseudo-neighbourhoods, the levels of ethnic isolation and evenness segregation in Section 5.1 - 5.3.

As segregation indexes are influenced by the ethnic profile of a neighbourhood, three sensitivity checks were conducted to ensure that the measurement of segregation over time completed in Section 5.1.2 was done reliably for each phase of schooling. The first check aimed to ensure that measures of segregation were not unduly influenced by changes in the sample of pseudo-neighbourhoods between each year. As England diversified between 2006 - 2019, the sample of ethnically diverse pseudo-neighbourhoods grew from 1400 pseudo-neighbourhoods in 2006 to 2224 pseudo-neighbourhoods by 2019. These 824 "newly diverse" pseudo-neighbourhoods constitute a 63% growth in the sample and their introduction could have been responsible for changes in the median level of segregation. This thesis therefore calculated a separate set of median H values for pseudo-neighbourhoods that had been consistently included in the sample since 2006 and compared it to the median H values for the expanding sample of ethnically diverse pseudo-neighbourhoods, to ensure that the two patterns broadly aligned.

The second sensitivity analysis checked whether there was an association between the change in the proportion of White British students in a pseudo-neighbourhood and a change in ethnic school segregation (*H*) between 2006 - 2019. This helped to ensure that observed

changes in segregation did not directly reflect changes in the ethnic profile of a given pseudo-neighbourhood.[121]

The final sensitivity analysis checked whether the level of "random" segregation was consistent between 2006 - 2019. [122] The study ran a simulation analysis which randomly assigned students to a school in their current neighbourhood and measured the level of segregation based on these random school attendance patterns. This process was repeated 100 times, with each iteration using a different set of neighbourhood boundaries. The level of random segregation was then compared between the years of 2006 - 2019 to ensure that there were no substantial differences that could have been responsible for the change in the median level of segregation in England's school pseudo-neighbourhoods.

4.9.2 Meta Regression Modelling

Five pairs of meta-regression models, one for each phase of schooling, were used to investigate linear associations between ethnic school segregation (*H*) and the pseudo-neighbourhood level predictor variables. For each model, this study used k means clustering to define 100 different sets of pseudo-neighbourhood boundaries, ran an analysis on each set of pseudo-neighbourhoods and stored the coefficients from the 100 models. The study then calculated an unweighted mean average of each of the outputs from the 100 models to create "summary" results that were insensitive to how a single set of pseudo-neighbourhood boundaries were defined.[123] This is akin to the procedure used by meta-

[121] This analysis could only be completed on neighbourhoods which had an *H* value for both 2006 and 2019.
[122] For more information on random segregation, see the Section 2.1.3.
[123] This approach was also used to calculate the significance levels of models.

analyses, where the estimate for the "summary effect" of an intervention is gained by combining different "effect estimates" from separate, related studies (Borenstein *et al.*, 2009). One key difference is that the estimates from each of the 100 models in this thesis were assumed to be equally valid and are weighted accordingly, whereas in most meta-analysis each "effect estimate" is weighted to reflect the author's judgement on the reliability and precision of the study from which it was drawn.[124]

The thesis considered using a fixed or random effects model to describe the distribution of these 100 model estimates, but after discovering that there was minimal variation in these estimates, it was decided that summary statistics provided a simpler, more accessible approach that would still produce reliable results (Borenstein et al., 2009, p. 7).[125]

All regression models were run using the lm() function in R (R Core Team, 2019) and are listed in the table below. For each model, the subscript p or s was added to indicate whether the analysis was conducted using primary or secondary school pseudo-neighbourhoods. Readers should also be aware that *H* was log transformed in all models to ensure that the assumption of linearity was met.

[124] For example, an "effect estimate" from a study with a large population would be assigned a higher weighting when calculating the meta-analysis's "summary effect" than an "effect estimate" from a study with a smaller population size.

[125] For a full description of the variation in the coefficient estimates, and other model outputs, from the pair of underlying 100 regression models see Table 35 in the appendix.

Table 22: Formulas for the pseudo-neighbourhood level meta-regression models

Model 1_p & 1_s (Control variables)	$Log(H_n) = \beta_0 + \beta_1 NUMSCH_n + \varepsilon_n$

Model 2_p & 2_s	$Log(H_n) = \beta_0 + \beta_1 POPDENSITY_n + \beta_2 NUMSCH_n + \varepsilon_n$
Related Research Questions	4a) Is the population density of the area surrounding a primary school pseudo-neighbourhood associated with its observed level of ethnic school segregation? 4b) Is the population density of the area surrounding a secondary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?
Null Hypothesis	$\beta_1 = 0$

Model 3_p & 3_s	$Log(H_n) = \beta_0 + \beta_1 HOUSE_n + \beta_2 FSM_n + \beta_3 POPDENSITY_n + \beta_4 NUMSCH_n + \varepsilon_n$
Related Research Questions	4c) Is the average house price of the area surrounding a primary school pseudo-neighbourhood associated with its observed level of ethnic school segregation? 4d) Is the average house price of the area surrounding a secondary school pseudo-neighbourhood associated with its observed level of ethnic school segregation? 4e) Is the rate of Free School Meal eligibility (FSM) within a primary school pseudo-neighbourhood associated with its observed level of ethnic school segregation? 4f) Is the rate of Free School Meal eligibility (FSM) within a secondary school pseudo-neighbourhood associated with its observed level of ethnic school segregation?

Null Hypotheses	$\beta_1 = 0$
	$\beta_2 = 0$

Model 4_p & 4_s	$Log(H_n) = \beta_0 + \beta_1 RELIGOUS_n + \beta_2 GRAMMAR_n + \beta_3 AVG_OFSTED_n + \beta_4 HOUSE_n + \beta_5 FSM_n + \beta_6 POPDENSITY_n + \beta_7 NUMSCH_n + \varepsilon_n$
Related Research Questions	4g) Do secondary school pseudo-neighbourhood's that contain a grammar school experience higher levels of ethnic school segregation, on average, than those that do not? 4h) Is there an association between the proportion of schools in a primary school pseudo-neighbourhood that have a religious affiliation and its observed level of ethnic school segregation? 4i) Is there an association between the proportion of schools in a secondary school pseudo-neighbourhood that have a religious affiliation and its observed level of ethnic school segregation? 4j) Is there an association between the average OFSTED rating of schools within a primary school pseudo-neighbourhood and its level of ethnic school segregation? 4k) Is there an association between the average OFSTED rating of schools within a secondary school pseudo-neighbourhood and its level of ethnic school segregation?
Null Hypotheses	$\beta_1 = 0$
	$\beta_2 = 0$
	$\beta_3 = 0$

Model 5_p & 5_s	$Log(H_{nr}) = \beta_0 + \beta_1 RELIGOUS_{nr} + \beta_2 GRAMMAR_{nr} + \beta_3 AVG_OFSTED_{nr} + \beta_4 HOUSE_{nr} + \beta_5 FSM_{nr} + \beta_6 POPDENSITY_{nr} + \beta_7 NUMSCH_{nr} + REGION_r + \varepsilon_{nr}$

Where:

H_n	H index score for pseudo-neighbourhood *n*
$POPDENSITY_n$	Average Population Density in pseudo-neighbourhood *n* in the previous year.
$HOUSE_n$	Average price paid to buy a terrace house in pseudo-neighbourhood *n* in the previous year (£).
FSM_n	Percentage of students in pseudo-neighbourhood *n* who are eligible for free school meals.
$RELIGOUS_n$	Proportion of schools in pseudo-neighbourhood *n* which have a religious affiliation
$GRAMMAR_n$	Binary variable describing whether pseudo-neighbourhood *n* contains a grammar school where 0 = no and 1 = yes.
AVG_OFSTED_n	Average OFSTED rating provides to schools contained within pseudo-neighbourhood *n*
$NUMSCH_n$	Number of schools grouped within pseudo-neighbourhood *n.*
$REGION_r$	Level 2 variable which describes a pseudo-neighbourhood's region.

Seven meta-regression models were subsequently used in Section 6.2 to investigate linear associations between the school level predictor variables and a school's contribution towards its pseudo-neighbourhoods' level of ethnic school segregation (H_j). As with the previous analysis, the subscript p or s was added to the model number to indicate whether the analysis was conducted on primary or secondary school pseudo-neighbourhoods. These models are listed in the table below.

Table 23: Formulas for school level meta-regression analysis

Model 6_s	$H_j = \beta_0 + \beta_1 SELECTIVE_j + \varepsilon_j$
Related Research Question	5a) Do academically selective secondary schools contribute more to ethnic school segregation, on average, than non-selective secondary schools located in the same pseudo-neighbourhood?
Null Hypothesis	$\beta_1 = 0$

Model 7_p & 7_s	$H_j = \beta_0 + \beta_1 RELIGOUS_j + \varepsilon_j$
Related Research Question	5b) Do primary schools with a religious denomination contribute more to ethnic school segregation, on average, than non-religious primary schools located in the same pseudo-neighbourhood? 5d) Do secondary schools with a religious denomination contribute more to ethnic school segregation, on average, than non-religious secondary schools located in the same pseudo-neighbourhood?
Null Hypothesis	$\beta_1 = 0$

Model 8_p & 8_s	$H_j = \beta_0 + \beta_1 CHURCHENG_j + \beta_2 CATH_j + \beta_3 CHRISTOTHER_j + \beta_4 JEWISH_j + \beta_5 MUSLIM_j + \beta_6 OTHER_j + \varepsilon_j$
Related Research Question	5c) If so, do primary schools with particular religious denominations contribute more to ethnic school segregation, on average, than others? 5e) If so, do secondary schools with particular religious denominations contribute more to ethnic school segregation, on average, than others?
Null Hypothesis	$\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6$

Model 9_p & 9_s	$H_j = \beta_0 + \beta_1 OUTSTANDING_j + \beta_2 GOOD_j + \beta_3 REQIMPROVE_j + \varepsilon_j$
Related Research Question	5f) Is there an association between a primary school's OFSTED rating and their contribution toward ethnic school segregation, relative to other primary schools located in the same pseudo-neighbourhood? 5g) Is there an association between a secondary school's OFSTED rating and their contribution toward ethnic school segregation, relative to other secondary schools located in the same pseudo-neighbourhood?
Null Hypothesis	$\beta_1 = \beta_2 = \beta_3 = 0$

Where:

H_j	The standardised contribution of school *j* toward their pseudo-neighbourhood's ethnic school segregation.
$SELECTIVE_j$	Selective admission policy (1 = Selective)
$RELIGOUS_j$	Religious school (1 = school *j* has a religious denomination)
$CHURCHENG_j$	1 = school *j* has an affiliation to the Church of England
$CATH_j$	1 = school *j* has an affiliation to the Catholic faith
$CHRISTOTHER_j$	1 = school *j* has an affiliation to another Christian Faith
$JEWISH_j$	1 = school *j* has an affiliation to a Jewish faith
$MUSLIM_j$	1 = school *j* has an affiliation to an Islamic faith
$OTHER_j$	1 = school *j* has an affiliation to another non-Christian Faith
$OUTSTANDING_j$	1 = school *j* received a "Outstanding" OFSTED rating in its last review
$GOOD_j$	1 = school *j* received a "Good" OFSTED rating in its last review
$REQIMPROVE_j$	1 = school *j* received a "Requires Improvement" OFSTED rating in its last review

The following model checks were conducted on all regression models to ensure that they were correctly specified, and the full results are listed in Section 11.23 of the Appendix.

Table 24: Model checks

Test Name/Description	Purpose
Scatter plots of predictor vs outcome variables	Check that the relationships between predictor and outcome variables were linear.
Calculate the mean average of the residuals	Check the formula described the line of "best fit".
Breusch–Pagan Test	Checking for heteroskedasticity in the models
Correlational analysis	Checking for collinearity in the predictor variables
VIF Scores	Checking for multi-collinearity in the models
Shapiro-Wilk test	Checking that the residuals of each model were normally distributed.
Cook's distance	Checking that the findings were not unduly influenced by a single data point

4.9.3 Simulation Analyses

The thesis used student level data to chart how patterns of ethnic school segregation shifted between each stage of the 2016/2017 application process before using seven simulations to explore what patterns of ethnic school segregation might have arisen if students were allocated to schools using a range of different procedures. The tables below outline the various patterns of school attendance that were used to complete this analysis.

Table 25: Procedure for charting how patterns of ethnic school segregation shift between each stage of the 2016/17 application process.

Stage of the application Process	**Procedure**
Stage 1 - 1st Preference	Year 6 students were assigned to the school that was listed as their first preference. Schools' Estimated Published Admission Numbers (PANs) were ignored for this step.
Stage 2 - Offers	Year 6 students were assigned to whichever secondary school they were initially offered in the Spring of 2017.
Stage 3 – Real Attendance Patterns	Ethnic school segregation is measured based on which schools the year 7 students end up attending in the Spring of 2018.

Once each of the patterns of school attendance had been created, patterns of ethnic school segregation were calculated *for year seven students* via the same procedure that was used in earlier parts of the thesis. Readers should note that the values of *H* created as part of these analyses are not directly comparable to the results reported in other chapters of the thesis, which were based on the entire school aged population rather than a single year cohort.

The study also ran seven simulations to explore the likely levels of ethnic school segregation within pseudo-neighbourhoods under different national school admission systems. The first five adapted the current school choice model and tried to keep the procedure as similar as possible to the system currently employed in England. In contrast, Simulation 6 and 7 allocated students to schools without considering the expressed preferences of parents and

operated a very different type of admissions system. The specific procedures used to create these allocations are described in the table below. It is important to stress beforehand that the simulations were designed to be simple so they could be easily understood. Consequently, each should be viewed as a series of thought experiments based on empirical data rather than an attempt to create "ready-made" national admission systems that could easily be implemented. The study openly acknowledges, for example, that none of the simulations consider the gender or religious policies of different schools, the legal requirement of schools to offer places to students in care or with special needs, or the impracticalities of not offering places to applicants with siblings at a given school.

Table 26: The seven simulations for allocation year 6 students to their secondary school.

Simulation Number	**Description**	**School Choice**	**Allocation Procedure**
1	School Choice with home – school distance for oversubscribed schools	Yes	1. Year 6 students were assigned to the applicant list of all schools that were listed on their application form. 2. Students on each applicant list were ranked based on their home-school distance. 3. Places were earmarked for whichever students lived closest to the school, in accordance with the school's estimated PAN number. 4. If there were less applicants than the school's estimated PAN number, then all students were earmarked for an offer. 5. Students who were earmarked to receive an offer from more than one school were removed from the applicant lists of all schools that were lower down on their application form than the highest school that they were earmarked to receive an offer from. 6. Step 2 – 4 were then repeated until students were consistently earmarked for an offer from the same school.

			7. Any student that was not earmarked to receive an offer from any of the schools listed on their application form was then allocated to a nearby school that still had space. The process for this allocation replicated the procedure used in Simulation 6.
2	School choice with random allocation for oversubscribed schools	Yes	1. Year 6 students were assigned to the applicant list of all schools that were listed on their application form. 2. Places were earmarked to students at random, in accordance with the school's estimated PAN number. 3. If there were less applicants than the school's estimated PAN number, then all students were earmarked for an offer. 4. Students who were earmarked to receive an offer from their 1st preference school were allocated to this school and removed from all other school's applicant lists. 5. Estimated PAN numbers were updated to reflect the allocation of these places to reflect the number of remaining places that each school had to offer. 6. Step 2 – 5 were then repeated until no more students could be allocated to their 1st preference school. 7. Step 2 – 6 were then repeated until no more students were allocated to their 2nd preference school, and in turn, their 3rd, 4th, 5th, and 6th preference. 8. Any student that was still not allocated to a school was then allocated to a nearby school that still had space. The process replicated the procedure used in Simulation 6.
3	School choice with academic banding followed by a distance-based allocation for oversubscribed schools	Yes	1. Year 6 students were ranked nationally based on their KS2 results and grouped into four equally sized bands based on the quartile that their ranking fell into. 2. Year 6 students were assigned to the applicant list of any school that was listed on their application form. 3. Students on each applicant list were grouped with students that were in the same academic band. 4. Students *within each of these banded applicant lists* were ranked based on the distance between their home and school, with places earmarked for students that lived closest to the school. 5. Students ranked highest each *banded* applicant list were earmarked for a place in accordance with a quarter of the school's estimated PAN number.

			6. If there were less applicants on a *banded* applicant list than a quarter of school's estimated PAN number, then all students on the banded applicant list were earmarked for an offer. 7. Students who were earmarked to receive an offer from more than one school were removed from the applicant lists of schools that were lower down on their application form than the highest school that they were earmarked to receive an offer from. 8. Step 2 – 4 were then repeated until students were consistently earmarked for an offer from the same school. 9. Any student that was not allocated to any school listed on their application form was then allocated to a nearby school that still had space. The process for allocating these students replicated the procedure of Simulation 6 but incorporated the use of banding.
4	School choice with academic banding followed by a random allocation for oversubscribed school	Yes	1. Year 6 students were ranked nationally based on their KS2 results and grouped into four equally sized bands based on which quartile their results fell into. 2. Year 6 students were assigned to the applicant list of any school that was listed on their application form. 3. Students on each applicant list were grouped with students that were in the same academic band. 4. Students within each of these *banded* applicant lists were earmarked for an offer at random, in accordance with a quarter of the school's estimated PAN number. 5. If there were less applicants than a quarter of the school's estimated PAN number on the *banded* applicant list, then all students on the *banded* applicant list were earmarked for an offer. 6. Students who were earmarked to receive an offer from their 1st preference school were allocated to this school and removed from all other school's applicant lists. 7. Estimated PAN numbers for each quartile of each school were updated to reflect the allocation of these places so that they described the number of remaining places that each quartile within each school had to offer. 8. Step 2 – 5 were then repeated until no more students could be allocated to their 1st preference school. 9. Step 2 – 6 were then repeated until no more students were allocated to their 2nd preference school, and in turn, their 3rd, 4th, 5th, and 6th preference.

			10. Any student that was still not allocated to a school was then allocated to a nearby school that still had space. The process replicated the procedure used in Simulation 6 but incorporated the use of banding.
5	School choice with academic competition followed by a distance-based allocation for oversubscribed schools	Yes	1. Year 6 students were assigned to the applicant list of all schools that were listed on their application form. 2. Students on each applicant list were ranked based on their KS2 test performance scores, with places earmarked for whichever students had the highest attainment. The number of students earmarked for an offer corresponded to the schools' PAN. 3. If there were less applicants than the school's PAN, then all students were earmarked for an offer. 4. Students who were earmarked to receive an offer from more than one school were removed from the applicant lists of schools that were lower down on their application form than the highest school that they were earmarked to receive an offer from. 5. Step 2 – 4 were then repeated until students were consistently earmarked for an offer from the same school. 6. Any student that was not earmarked to receive an offer from any school on their application form was then allocated a nearby school that still had space based on their home-school distance. The process for this secondary allocation replicated Simulation 6 except that the ranking described in step 2 were based on the student's KS4 results, not their home-school distance.
6	Strict home – school distance without any school choice	No	1. Students were assigned to the applicant lists of the 15 schools that were closest to their home. 15 was chosen to ensure that over 98% of students could be ultimately allocated to a school, without students being regularly allocated to a school that was far away from their home postcode.[126] 2. Students on each applicant list were ranked, based on their home-school distance, and places were earmarked for whichever students lived closest to the school. The number of places that were earmarked for offer corresponded to the schools' PANs. 3. Students who were earmarked to receive an offer from more than one school were removed from the applicant lists of schools that were further away from their home.

[126] See Table 48 for more detail.

			4. Step 2 – 4 were then repeated until students were consistently earmarked for an offer from the same school.
7	Reintroducing the grammar school system without any school choice.	No	1. Any secondary school that had a GCSE pass rate in the top 25% of schools nationally was labelled a "pseudo-grammar school" while all other schools were labelled a "pseudo-secondary modern".[127] 2. Students were ranked nationally based on their KS2 performance and any student which fell within the top quartile of students was labelled a "pseudo-grammar student", while all other students were labelled a "pseudo-secondary modern student". 3. Students were assigned to the waiting lists of 15 compatible schools that were closest to their home. Specifically, all pseudo-grammar students were assigned to the applicant lists of 15 pseudo-grammar schools and pseudo-secondary modern students were assigned to the waiting lists of 15 pseudo-secondary modern schools. 15 schools were chosen to ensure that over 97% of students could be allocated to a school, without students being regularly allocated to a school that was far away from their home postcode. 4. Students were ranked based on their home-school distance and places were earmarked for whichever students lived closest to each school in accordance with their PAN numbers. 5. Students that were earmarked to receive an offer from more than one school were removed from the applicant list of whichever schools that were further away from their home. 6. Step 2 – 4 were then rerun until students were consistently earmarked for an offer from the same school.

[127] 25% was chosen to reflect the proportion of secondary school that were grammar schools in the UK at their peak during the min 1960s (Danechi, 2017, p. 1).

4.10 Ethical Considerations

The following section explains the range of steps the study took to engage with any ethical concerns that could arise from conducting this analysis, or any later publication of the results.

Firstly, the research project passed the CUREC review process carried out by the Departmental Research Ethics Committee (ED- C1A-19-244) and it agreed to several practical steps that help reduce the risk of any harm occurring because of this study. The study promised to:

- Adhere to all data privacy laws, notably the UK's Data Protection Act 2018 and the EU's General Data Protection Regulation.
- Aggregate all research findings to a unit of analysis that prevents the identification of any individual or any small group of individuals. All outputs will be reviewed by myself, staff at The Office for National Statistics and my supervisors before being published. Any findings which are at risk of disclosing information on any individual will not be allowed to leave the ONS's secure environment.
- Remind readers throughout the thesis that findings are based on observational data and so relationships cannot be assumed to be causal. Causal language, or language which implies causality, will be avoided.
- List all sources of funding, and any person(s) who contributes to the research, at the beginning of the thesis to offer transparency to the reader.

A number of additional steps were taken following the renewed focus on equity and ethnic issues after the Black Lives Matter protests. The study is conscious that it will be publishing results into a national environment which is fraught, and that the release of any related empirical work has the potential to increase tensions between different ethnic groups. In particular, there is a small but real danger that the findings of this thesis could be used to encourage inter-ethnic tensions or hostilities. Findings that Jewish, Muslim, and Other Non-Christian schools contribute disproportionately to ethnic school segregation is particularly relevant in this respect, given the ongoing harassment faced by these religious communities.[128]

This study will therefore take some additional steps to ensure that the risk of this study causing harm to any person(s) within the current context is minimalised. These are:

- Research findings will be written in a way that are sensitive to current debates over equity and ethnicity in the UK.
- Any language which attributes blame to a particular ethnic group(s) will be stringently avoided.
- Results which relate to White British students will be written in the same manner as those relating to other ethnic groups.
- The research will actively highlight different viewpoints to stress how contested these topics remain.

[128] Annual UK crime statistics released in 2019, for example, show that 48% of religious hate crimes targeted Muslims, 18% targeted Jews and 12% targeting members of other non-Christian minority religions (UK Home Office, 2019). These findings are particularly shocking given that these religious groups constitute only a small fraction of the overall UK population.

Summary of Methodology Chapter

The methodology chapter has described the sources of data used by this study, along with the characteristics of the samples it used. It then explained how pseudo-neighbourhoods were created and justified the five ethnic groups used by this study. The outcome variables of this study were introduced (H index, H_j and Shared Peer Ethnicity) along with the various control and predictor variables. Finally, the different analytical techniques used by the thesis were described, before ending with a discussion on the ethical issues that were considered as part of this study.

5 Describing the ethnic profile of England's school age population, the patterns of ethnic school segregation and the level of ethnic isolation between 2006 – 2019.

The following chapter presents results which address the research questions 1 – 3. Section 5.1 describes the ethnic profile of England's school aged population in state funded education (5-18 years old) and explains how it has changed between 2006 – 2019, at a national, regional, and pseudo-neighbourhood level. The changes are described both in relation to group proportions as well as the total number of pupils from each ethnic group, split by phase. Section 5.2 then describes the average level of ethnic isolation for students in state-funded English schools, both collectively and for each ethnic group for primary and secondary schools. It also briefly considers how these patterns have changed between 2006 – 2019. Section 5.3 then describes the levels of ethnic school segregation for pseudo-neighbourhoods in England (*H*), both numerically and geographically before describing how these patterns have changed between 2006 – 2019.

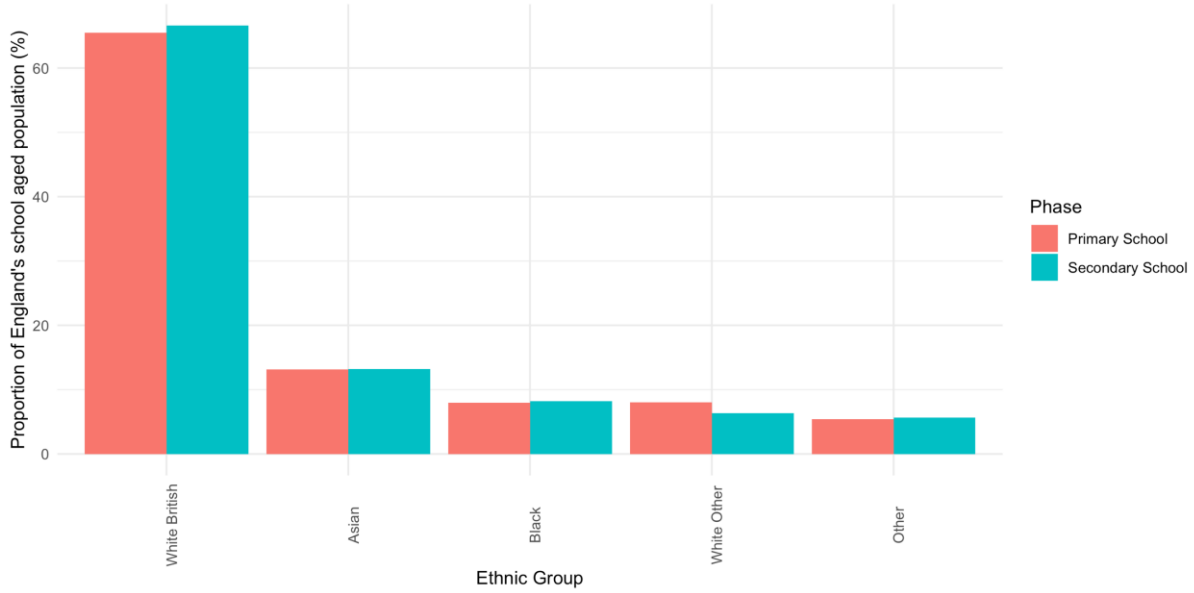
5.1 The ethnic profile of England's school age population

5.1.1 The current ethnic profile of England's school age population

As of 2019, there were 8.06 million students attending a mainstream state school in England, with 4.7 million primary school students (58%) and 3.4 million secondary school students (42%). Around two thirds of students were White British (66%), with 13% from

Asian ethnic groups, 8% from Black ethnic groups and 7% from Other White ethnic groups. The remaining 6% were from an "other" ethnic background.[129] Students from ethnic minority backgrounds constituted a slightly greater proportion of pupils in primary state schools (33%), though the difference was small (34%), as outlined in the figure below.

Figure 13: An ethnic breakdown of England's school age population by phase of schooling (2019)[130]



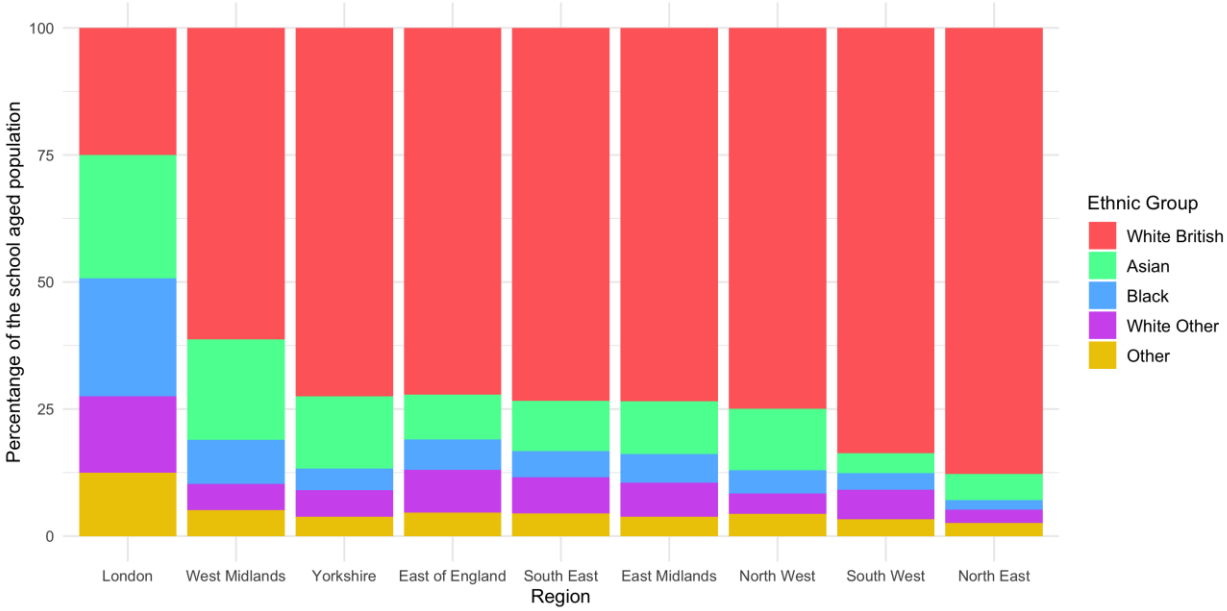
As ethnic groups were not spread evenly across the country, each region in England had a different ethnic profile. London was the most ethnically diverse region with near equal numbers of White British, Asian and Black students (≈25%). The West Midlands was also comparatively diverse, due in large part to its substantial Asian population and the ethnic

[129] As it is difficult to know what students are in this "other" group, the group is not discussed regularly in this chapter.

[130] There were only very small differences between primary and secondary school population and so these figures were jointly reported.

diversity of Birmingham. The regions of Yorkshire and the Humber, East of England, South East, East Midlands and North West all contained similar amounts of ethnic diversity; with around 25% of students coming from an ethnic minority group.[131] In contrast, the South West and North East were the least diverse regions, with White British students making up 83.7% and 87.8% of the student population, respectively.

Figure 14: An ethnic breakdown of England's school age population by region (2019)



There was substantial variation in the ethnic profile of pseudo-neighbourhoods across England. Some pseudo-neighbourhoods were exclusively White British while others were "super-diverse", with roughly equal numbers of students from all five ethnic groups. Given that there were over 3292 pseudo-neighbourhoods in England, describing these patterns in

[131] Which ethnic minority groups make up the remaining 20% does, however, differ between each region. This is particularly relevant in the case of Yorkshire and the Humber which educates a larger proportion of Asian students than other regions with similar levels of ethnic diversity.

detail is challenging, particularly when split by phase of schooling.[132] Initially the section describes the distribution of each ethnic group separately. While this is relatively clear, it requires the reader to piece together five separate narratives if they wish to gain a holistic picture of England's pseudo-neighbourhoods. Therefore, the thesis also describes England's pseudo-neighbourhoods in relation to which ethnic group, if any, was in the absolute majority (>50%). This approach provides a secondary, somewhat crude, description of the ethnic profile of pseudo-neighbourhoods across the country which is perhaps easier to interpret.

5.1.1.1 Describing the ethnic profile of pseudo-neighbourhoods in relation to each ethnic group

On average, White British students constituted around 86% of students within primary school pseudo-neighbourhoods and 81% in secondary school pseudo-neighbourhoods, though these figures were smaller in many pseudo-neighbourhoods in England.[133] By definition, the opposite was true for students from ethnic minority backgrounds. Most pseudo-neighbourhoods tended to have a small percentage of students from Asian (primary = 3.4%, secondary = 4.8%), Black (primary = 2.3% secondary = 3.0%) or other white ethnic backgrounds (primary = 4.4%, secondary = 4.5%), while some pseudo-neighbourhoods educated a substantially greater proportion of students from each of these ethnic groups.[134]

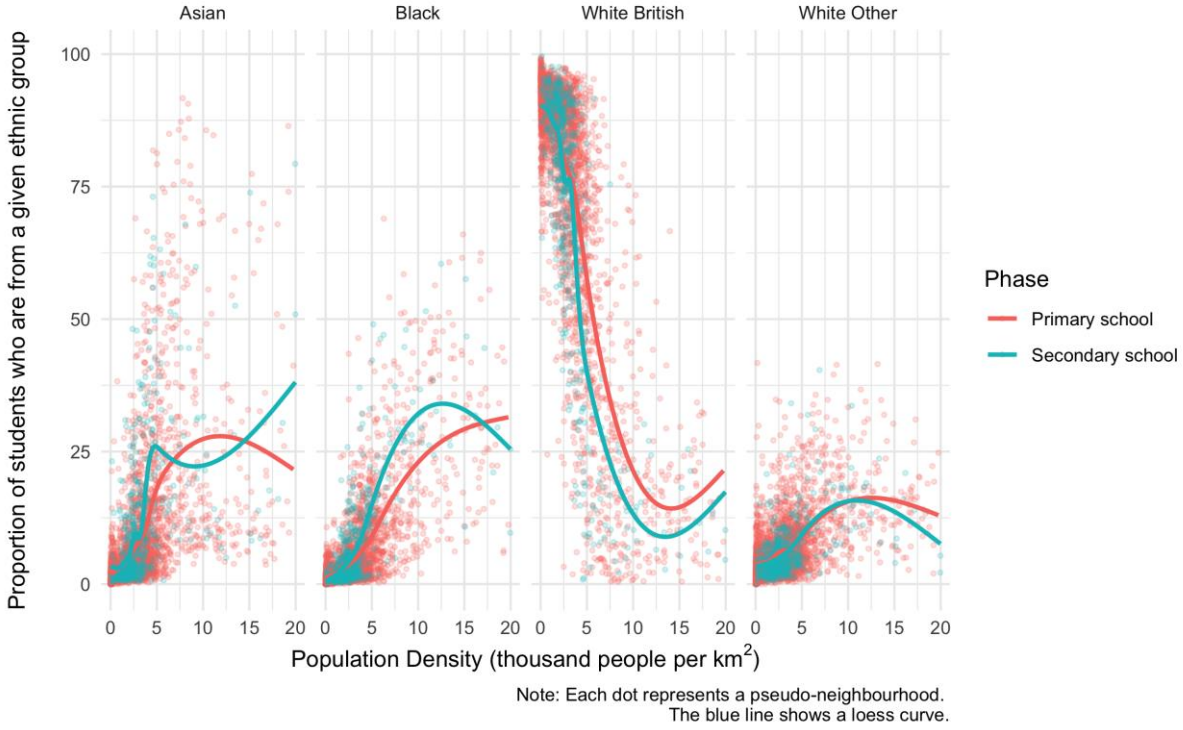
[132] See

Figure **68** in A11 for a stacked bar chart.

[133] These figures are based on median averages. See Figure 69 in Section 11.20 for a density plot on the distribution of White British students in English pseudo-neighbourhoods.

[134] This is due to a strong positive skew in the data. See Figure 70 in Section 11.20 for a density plot of these distributions.

Figure 15: Comparing a pseudo-neighbourhood's population density against its proportion of students from each ethnic group in 2019, split by phase of schooling.



As outlined in Figure 15, urban pseudo-neighbourhoods with higher population densities tended to educate a smaller percentage of White British students and a greater proportion of students from all other ethnic groups, both for primary and secondary school pseudo-neighbourhoods. Pseudo-neighbourhoods in cities such as London, Birmingham, Manchester, Leeds, and Bradford, in particular, educated large numbers of ethnic minority students. However, not all diverse urban pseudo-neighbourhoods educated similar proportions of each ethnic minority group. Pseudo-neighbourhoods in the Midlands and the North West, for example, tended to educate a higher proportion of students from Asian backgrounds, while pseudo-neighbourhoods in the East of England educated a greater

proportion of students from other white ethnic groups. These differences are outlined in the map on the following page.

Figure 16: Mapping the ethnic profile of pseudo-neighbourhoods in England (2019).[135]



[135] White British students were not included on this map as it would disrupt the scale and make the pattern hard to visualize. See
Figure **71** in Section 11.20 for a map to show the attendance patterns of White British students.

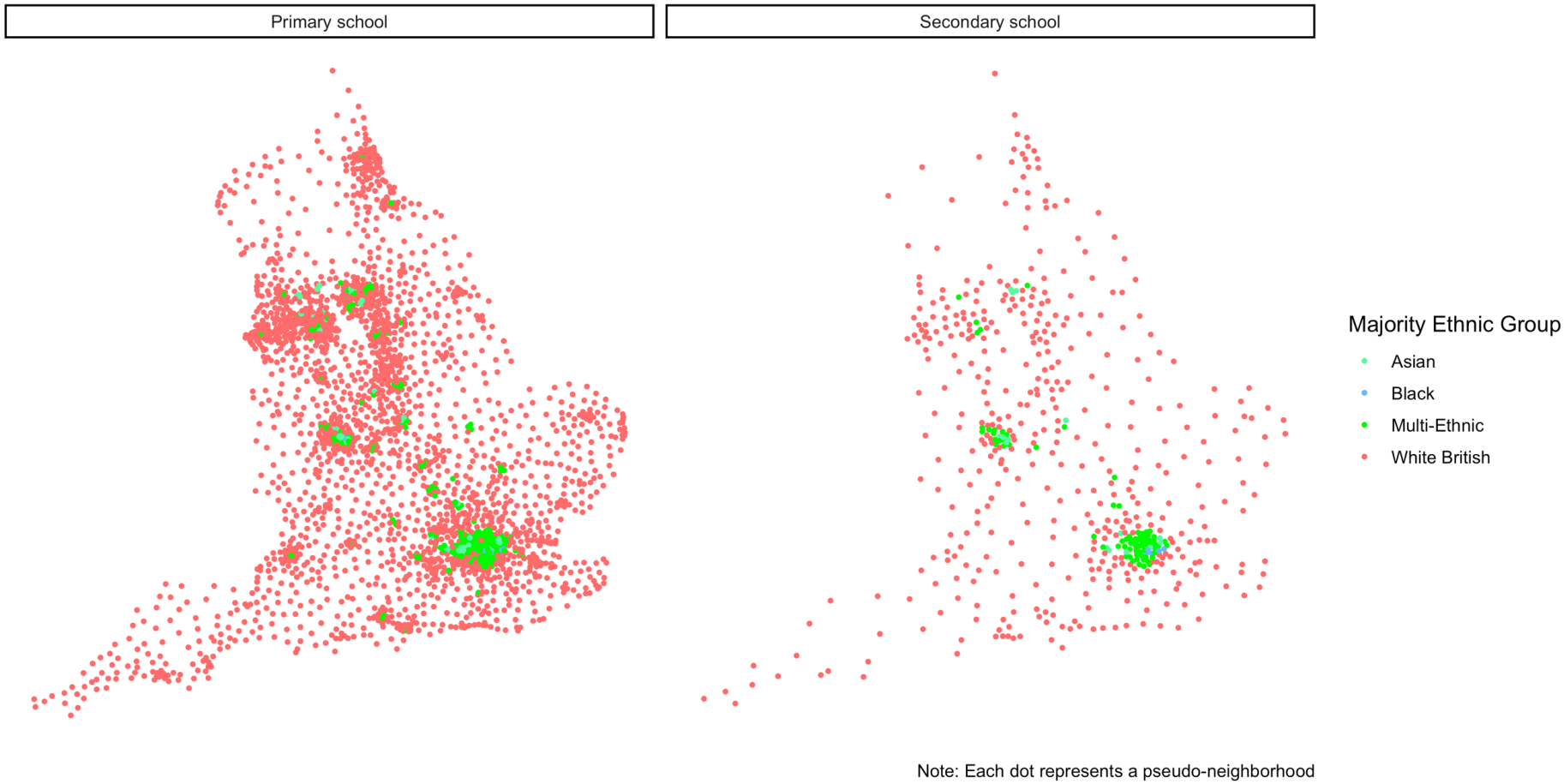
5.1.1.2 Exploring the ethnic profile of pseudo-neighbourhoods by describing which ethnic group is in the majority

Categorizing pseudo-neighbourhoods based on which ethnic group, if any, forms an absolute majority (>50%) provides another way of exploring the ethnic profile of England's pseudo-neighbourhoods.[136] In 2019, White British students formed an absolute majority in 83% of primary school pseudo-neighbourhoods with 3.4% having an Asian student majority and 0.5% having a Black student majority. The remaining 13.3% of primary school pseudo-neighbourhoods were multi-ethnic, with no single ethnic group constituting more than 50% of the student population. The figures were similar for secondary school pseudo-neighbourhoods. White British students formed a majority in 79.5% of cases, with 15.8% being multi-ethnic, 3.9% having an Asian student majority and 0.7% having a black student majority [137]

[136] This broadly aligns with how other UK research has defined multi-ethnic spaces in the past (Catney, Wright and Ellis, 2020, p. 1)

[137] There are no neighbourhoods in England were students from other white ethnic groups or other ethnic backgrounds form the majority.

Figure 17: A map to show which ethnic group, if any, were in the absolute majority for each pseudo-neighbourhood in 2019



Pseudo-neighbourhoods with a white British majority were particularly prevalent in rural environments while, multi-ethnic pseudo-neighbourhoods, or pseudo-neighbourhoods with an Asian or Black majority, were exclusively found in urban settings. The cities of London, Birmingham, Manchester, and the West Yorkshire Conurbanisation contained a particularly large number of pseudo-neighbourhoods where an ethnic minority group constituted an absolute majority, whereas pseudo-neighbourhoods with a White British majority were more spread across England.

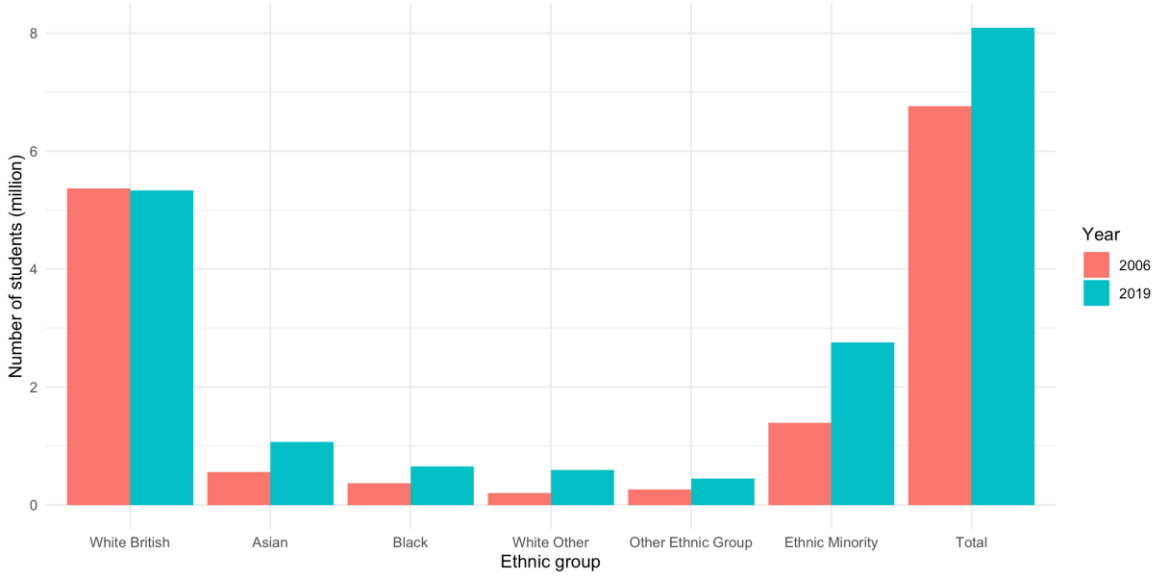
5.1.2 Changes in the ethnic profile of England's school aged population between 2006 – 2019

Between 2006 – 2019 the ethnic profile of England's school aged population in state funded education changed substantially. Despite an increase in the number of school aged students, following the lengthening of the period of compulsory schooling that young people in England must complete, the number of White British students fell by 30,000, from 5.37 million students to 5.34 million. During the same period, the number of students from ethnic minority groups rose by 1.36 million, from 1.39 million to 2.76 million.

This led to a 13.4 percentage-point reduction in the proportion of White British students at a national level, from 79% to 66%, and a growth in the percentage of students from all other ethnic backgrounds. The proportion of students from Asian backgrounds grew fastest, rising 4.9 percentage points from 8.3% to 13.2%. Students from other white ethnic groups also increased substantially, more than doubling from 3% of the student population to 7.3%

within the 13-year period. Black ethnic groups also increased, though less considerably, rising from 5.5% of the student population in 2006 to 8.1% by 2019.

Figure 18: Change in the number of students from different ethnic backgrounds in the English school system between 2006 – 2019.



These national changes were broadly reflected at a regional level. Between 2006 – 2019, all nine regions experienced an increase in the number of ethnic minority students while just over half experienced a fall in the number of White British students.[138] The West Midland's (+15.7%), London (+15%), and the East of England (+13.1%), in particular, experienced significant growth in the proportion of students from ethnic minority backgrounds. These changes were not driven by a single ethnic group with students from Asian, Black, and other White backgrounds *all* growing as a proportion of student in *all* nine regions. These results are displayed in

[138] These results should be read in the context of a significant growth in the size of the UK school population.

Figure 20.

Individual pseudo-neighbourhoods have also tended to become more ethnically diverse. Between 2006 - 2019 more than 95% of pseudo-neighbourhoods saw the number of ethnic minority students increase and 28% experienced a fall in the number of White British students. A quarter experienced both trends simultaneously (26%).

These trends differed slightly between primary and secondary school pseudo-neighbourhoods, as shown in Figure 19 . While both primary and secondary pseudo-neighbourhoods tended to educate more ethnic minority students in 2019 than they did in 2006, 88% of secondary school pseudo-neighbourhoods taught fewer White British students, whereas 82.8% of primary school pseudo-neighbourhoods educated more.

Despite this difference, 92% of primary school pseudo-neighbourhoods and 95% of secondary school pseudo-neighbourhoods in England experienced a fall in the proportion of White British students between 2006 – 2019. The median reduction was -6.27 and -8.25 percentage points for primary and secondary school pseudo-neighbourhoods, respectively. or many pseudo-neighbourhoods, the change was considerably greater.[139]

[139] See
Figure **73** in the appendix for a density plot of these pseudo-neighbourhood level changes.

Figure 19: A scatter plot to show the changes in the number of White British and ethnic minority students in English pseudo-neighbourhoods between 2006 – 2019.

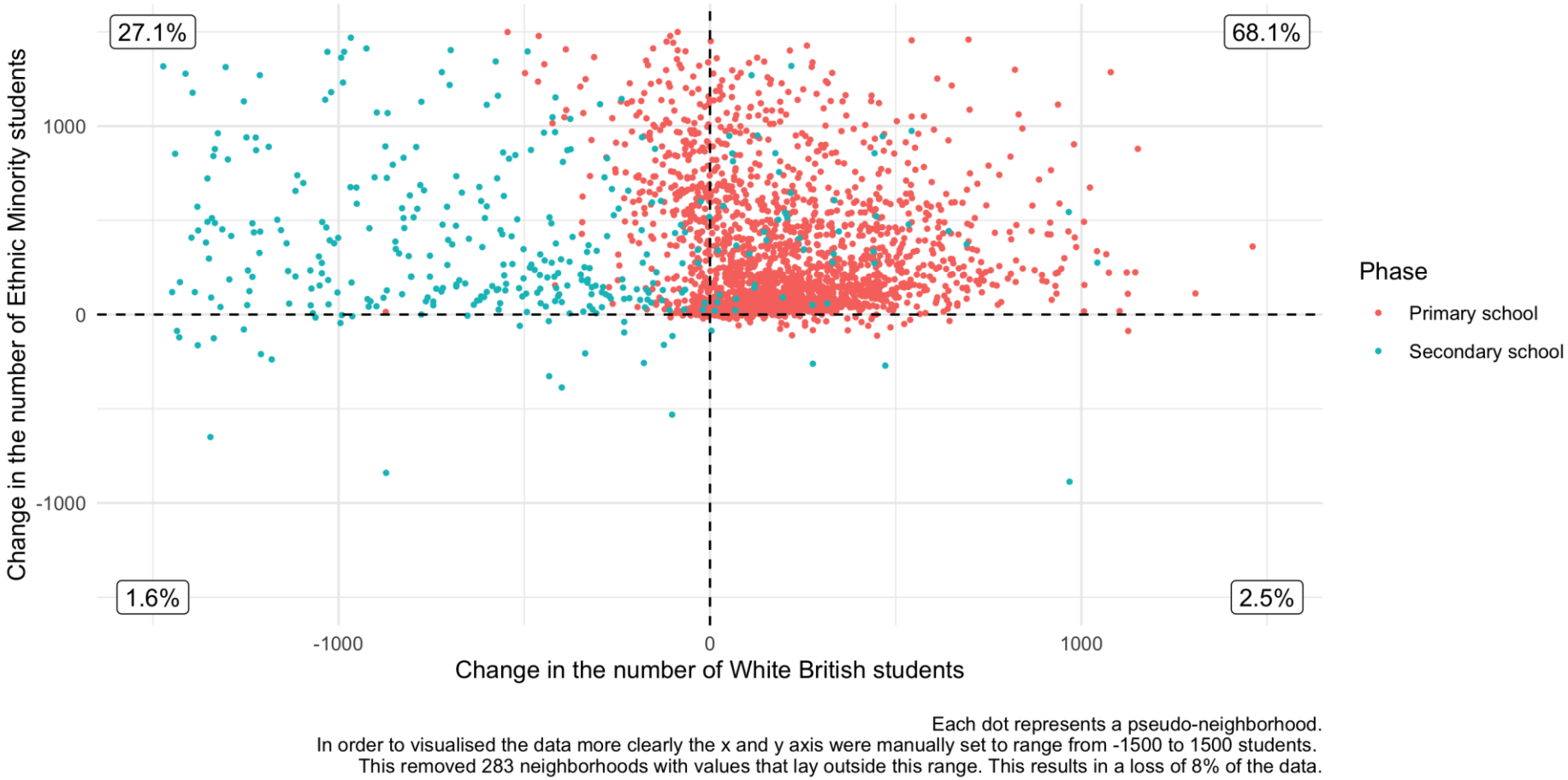
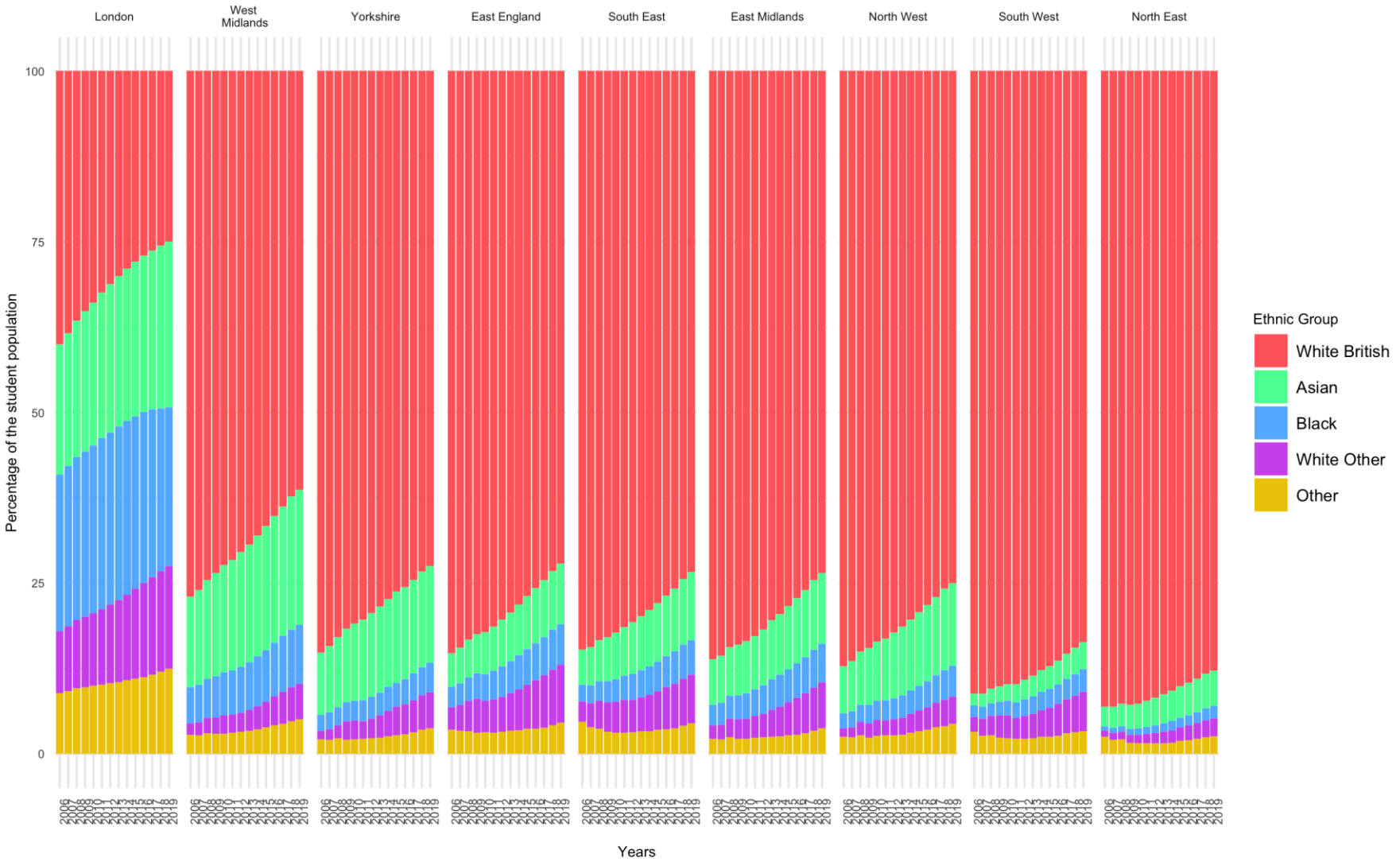


Figure 20: The changing ethnic profile of England's school age population by region (2006 – 2019).



More than 99% of both primary and secondary school pseudo-neighbourhoods experienced a growth in at least one ethnic minority group. 91% of primary school pseudo-neighbourhoods saw an increase in the proportion of students from other white ethnic groups, 88% experienced an increase in the proportion of Asian students and 76% experienced an increase in the proportion of Black students. For secondary school pseudo-neighbourhoods, the change was even greater. 97% experienced an increase in the proportion of students from other white ethnic groups, 96% experienced a growth in the percentage of Asian students and 92% experienced a rise in the proportion of students from black ethnic groups.

The proportion of students from other White ethnic groups typically grew fastest within primary school pseudo-neighbourhoods, with a median increase of 2.5 percentage points, followed by Asian students (+1.5%) and students from Black ethnic groups (+0.8%).[140] The figures for secondary school pseudo-neighbourhoods was similar, with White ethnic groups growing 2.4 percentage points, Asian students 2.1 percentage points and black students 1.2 percentage points.

Densely populated urban pseudo-neighbourhoods tended to diversify at a faster rate than their rural counterparts.[141] As shown in

[140] See Figure 74 in the appendix for a box plot which outlines the average change in the proportions of each ethnic group between 2006 – 2019.

[141] Consequently, there were strong associations between the ethnic profile of pseudo-neighbourhoods in 2006 and the subsequent changes that occurred in their ethnic profile between 2006 - 2019. Pseudo-neighbourhoods where White British students constituted around 60% of the student body in 2006 tended to experience the largest percentage point reduction in the size of the White British population. In contrast, pseudo-neighbourhoods where roughly 30% of their student population were Asian in 2006 tended to experience the largest increase in the proportion of Asian students, with the equivalent figures for Black and Other White ethnic groups both smaller at around 8%. See

Figure **21** below, pseudo-neighbourhoods with population densities of around 5000 people per km^2 tended to experience the greatest change, with the steepest reductions in White British students and the sharpest increases in the proportion of students from Asian and Black ethnic backgrounds.[142] This meant that the most rapid ethnic diversification occurred in major cities and towns, such as London, Birmingham, Manchester, Bradford, and Leeds.[143]

Figure 21: Plotting the percentage point changes in each ethnic group against a pseudo-neighbourhood's population density by phase (2006 – 2019).

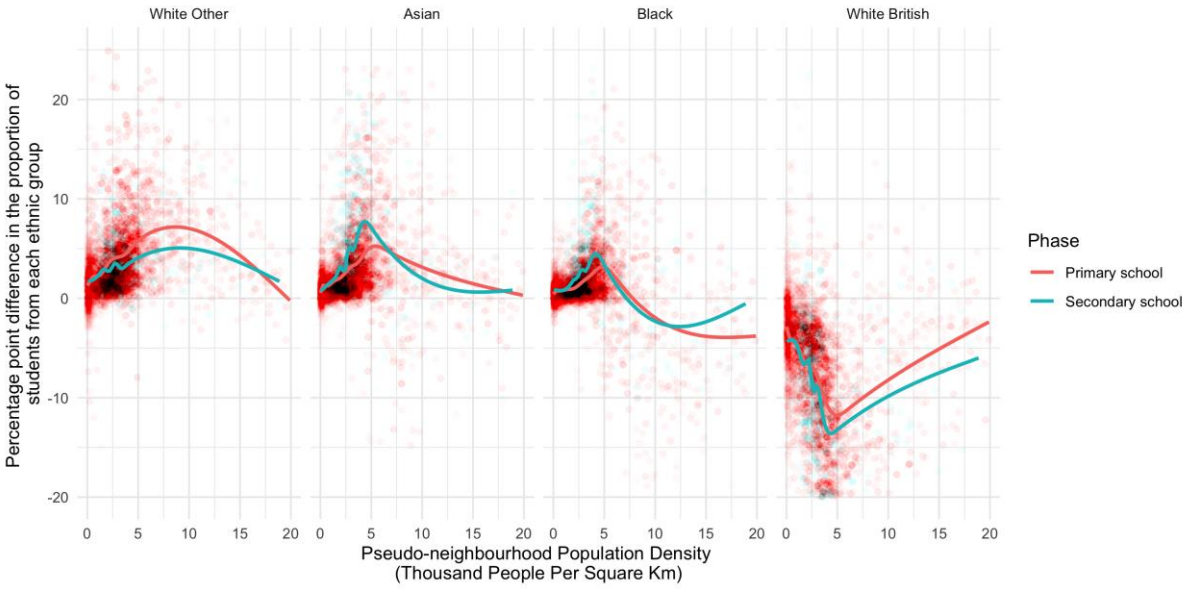


Figure 76 in the appendix for more detail.

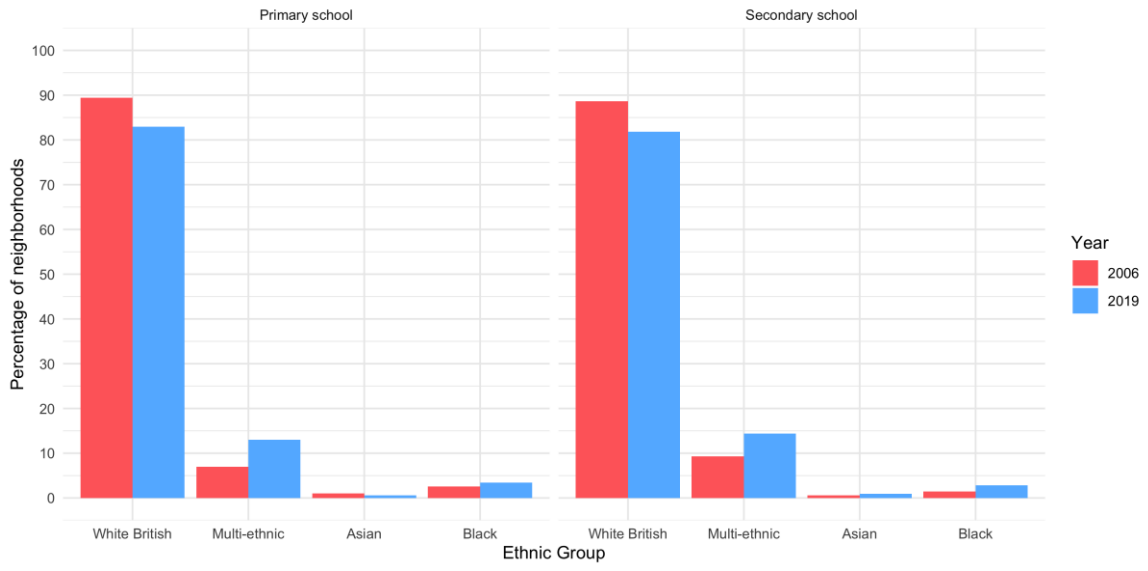
[142] Pseudo-neighbourhoods with 10,000 people per km^2 tended to experience the largest increase in students from other white ethnic groups.

[143] See

Figure **75** in the appendix for more detail.

The increased ethnic diversity of these urban environments led to a falling number of pseudo-neighbourhoods where one ethnic group formed an absolute majority (>50%). Between 2006 and 2019, the percentage of multi-ethnic pseudo-neighbourhoods increased from 6% to 14%, predominantly because of 208 urban pseudo-neighbourhoods losing their White British majority. As the number of pseudo-neighbourhoods with an Asian, White Other or Black student majority remained broadly stable, ethnic "ghettoization", as understood by this rudimentary measure, fell during this period. These trends were consistent between primary and secondary school pseudo-neighbourhoods, as shown in Figure 24 below.

Figure 22: Bar chart to show the number of pseudo-neighbourhoods where each ethnic group was in the majority (2006 – 2019).



Summary of Section 5.1

Section 5.1 has mapped the ethnic profile of England's school age population in state funded education and described how group compositions differed across parts of the country, thereby addressing 1a of the research questions. It then explained how a rising number of ethnic minority students and a falling number of White British students led to a more ethnically diverse population between 2006 - 2019. These changes were found to be wide ranging, with more than 9 out of every 10 pseudo-neighbourhoods educating a greater proportion of students from ethnic minority backgrounds in 2019 than they did in 2006. It also contrasted the changes occurring in urban and rural pseudo-neighbourhoods, and between different regions of the country, to show how the diversification has varied across different parts of England. This answered 1b of the research questions.

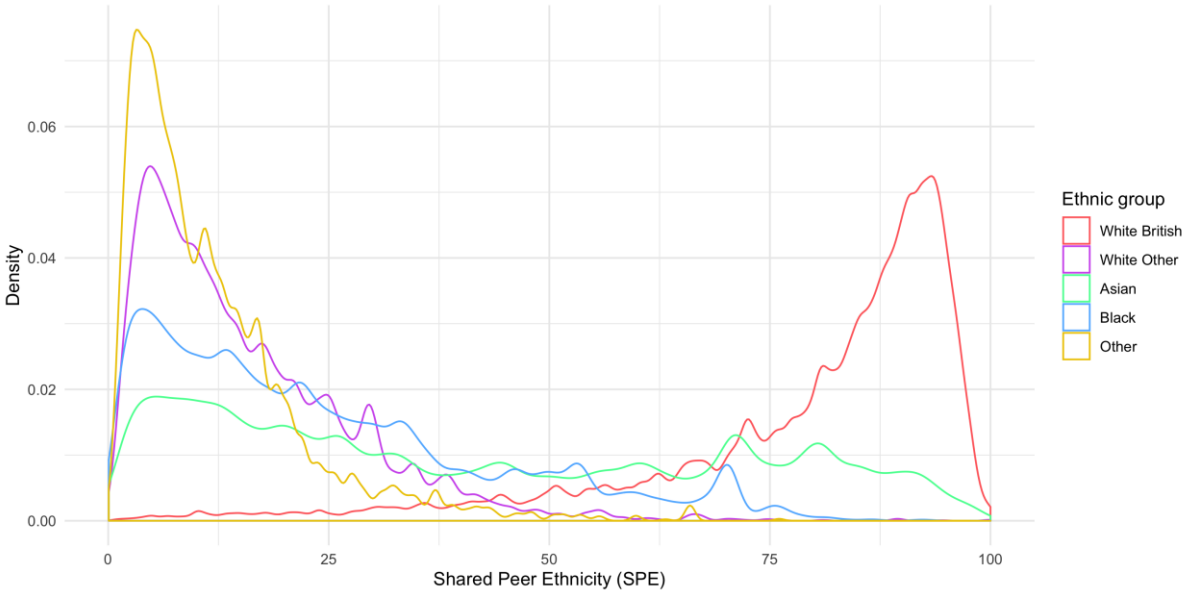
5.2 Ethnic isolation

The following section describes patterns of ethnic isolation for students attending state schools in England. Readers should be reminded that measures of isolation are heavily influenced by the relative proportions of ethnic groups in the pseudo-neighbourhood and so, as it is not possible to control for these differences, any results should be interpreted cautiously with crude comparisons avoided. Section 5.2.1 will initially explore the average level of Shared Peer Ethnicity (SPE) for all students, before delineating these values for students from different ethnic backgrounds. Section 5.2.2 then examines how the level of SPE has changed between 2006 – 2019.

5.2.1 The current levels of ethnic isolation in English schools.

In 2019 students attended state schools where, on average, 75% of their peers came from the same ethnic group, with 25% coming from a different ethnic background.[144] However, very few students have an SPE which matches the average figure. Instead, students either typically attend state schools were 95% of their peers come from the same ethnic background or they attend a school where only 5% of their peers come from the same ethnic group. [145]

Figure 23: A density plot to show the distribution of SPE values for students from different ethnic background in the English state school system (2019).



A student's SPE was heavily influenced by their ethnic background. White British students normally attended state schools where most of their peers came from the same ethnic

[144] These figures were similar for primary (74.9%) and secondary school students (72.4%)
[145] See Figure 77 in the appendix for a density plot of SPE for all students in England.

group. On average, both primary and secondary school White British students shared their ethnic identity with 86% of their peers.[146]

In contrast, students from all other ethnic groups attended schools where most of their peers came from a different ethnic background. This is particularly true for students from other white ethnic groups, who, on average, attended primary schools where only 13.8% of their peers came from the same ethnic background and secondary schools were 11.9% of the population were also from an Other White ethnic background. Similarly, students from Black ethnic groups attended primary schools where, on average 18.3% of their peers were also from a black ethnic group, with the median SPE for black students in secondary schools was 22.1%.

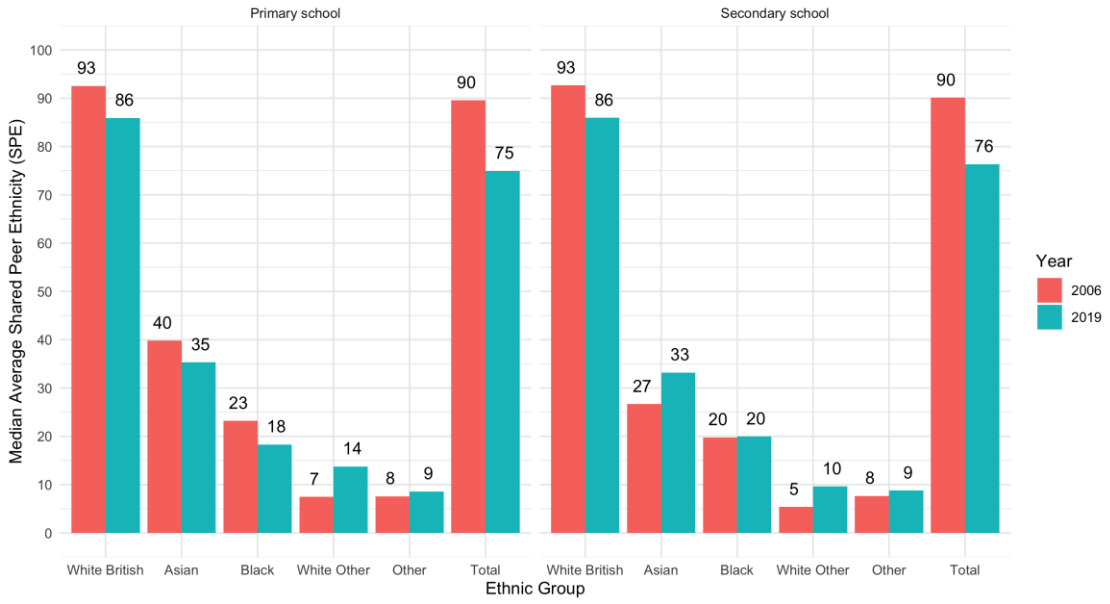
Asian students had a comparatively broad range of SPE values, both for primary and secondary state school students, with figures more closely resembling a uniform distribution. On average, Asian students attended primary schools where around 35.3% of the school were also from an Asian background and those attending secondary school shared their ethnic background with 41.0% of their peers. However, as Figure 23 suggests, many Asian students attended schools where the percentage of Asian students was either far higher or lower than the average SPE suggests. This wider range of SPE values makes it comparatively difficult to articulate what was a "typical" level of ethnic isolation for students with an Asian background.

[146] As SPE values for White British students are negatively distributed, a sizeable proportion of them also attend schools with a greater percentage of students from other ethnic groups.

5.2.2 Changes in the levels of ethnic isolation in English schools between 2006 - 2019

As England's school age population has diversified, students increasingly attend state schools with a larger proportion of students from other ethnic backgrounds. In 2006, students attended state schools were around 90% of their peers came from the same ethnic background. By 2019 this had fallen by 15 percentage points, with students attending schools where, on average, around 75% of their school was from the same ethnic group.[147]

Figure 24: Comparing the change in the median average Shared Peer Ethnicity (SPE) for student attending schools in England (2006 – 2019).



This resulted in a sharp reduction in the proportion of students who attended state schools where their ethnic group formed an overwhelming majority (>90%). In 2006 around half of all state school students in England (49.5%) attended a school where they shared their

[147] These figures vary minimally between primary and secondary school age students. See Figure 24.

ethnic background with at least 9 out of 10 of their peers. By 2019 this figure had fallen to 23.2%.[148]

Changes in ethnic isolation varied between students from different ethnic groups. Students from other white ethnic groups experienced a rise in the median SPE, both in primary and secondary schools. This meant that students from other white ethnic backgrounds, on average, attended state schools in 2019 where a greater proportion of their peers were also from other white groups than was the case in 2006. For primary school students, the median SPE rose from 7% to 14% from 2006 - 2019, while for secondary school students the median SPE rose from 5% to 10% over the same time period. Asian secondary school students also tended to attend schools with a higher proportion of Asian peers, with median SPEs rising from 27% to 33%, though the opposite finding was observed for Asian primary school students, with the median SPE value falling from 40% to 35% between 2006 – 2019.

Primary and secondary school students from White British backgrounds both experienced a reduction in their median SPE values, which meant White British students typically attended primary and secondary schools in 2019 where a smaller percentage of their peers were White British, on average, than was the case in 2006. In 2019, White British primary and secondary school students, on average, attended schools where 86% of their peers were also White British, down from 93% in 2006.[144]

[148] These figures were similar for primary and secondary school students.

For Black primary school students, the median SPE also fell from 23% in 2006 to 18% in 2019, while median SPE figures for black secondary school student remained consistent at 20% for both 2006 and 2019.

As stressed at the beginning of this section, values of SPE should be interpreted cautiously as they heavily reflect the relative size of each ethnic group. Without additional data, and a series of complex geographical models, it is hard to unpick the underlying drivers of these patterns or compare values between groups fairly. This does not, however, render this analysis useless. Having an estimate for the level of inter-group contact occurring in schools, and the differences between students from different groups, is important even if research struggles to explain such differences. Furthermore, given the mathematical and conceptual complexity of measuring evenness school segregation, using a broader set of metrics offers a more rounded evaluation of ethnic "separateness" in state schools.

Summary of Section 5.2

Section 5.2 has explored the levels of Shared Peer Ethnicity for students in 2019 and found that, on average, students attended state schools where around 75% of their peers came from the same ethnic group. This figure, however, did not capture the distinct experiences of students from different ethnic backgrounds. While students from ethnic minority groups typically attended schools where between 10% – 30% of their peers came from the same ethnic group, White British students attended schools where, on average, 85% of the students were also White British. These results answered question 2a and 2b of the research questions.

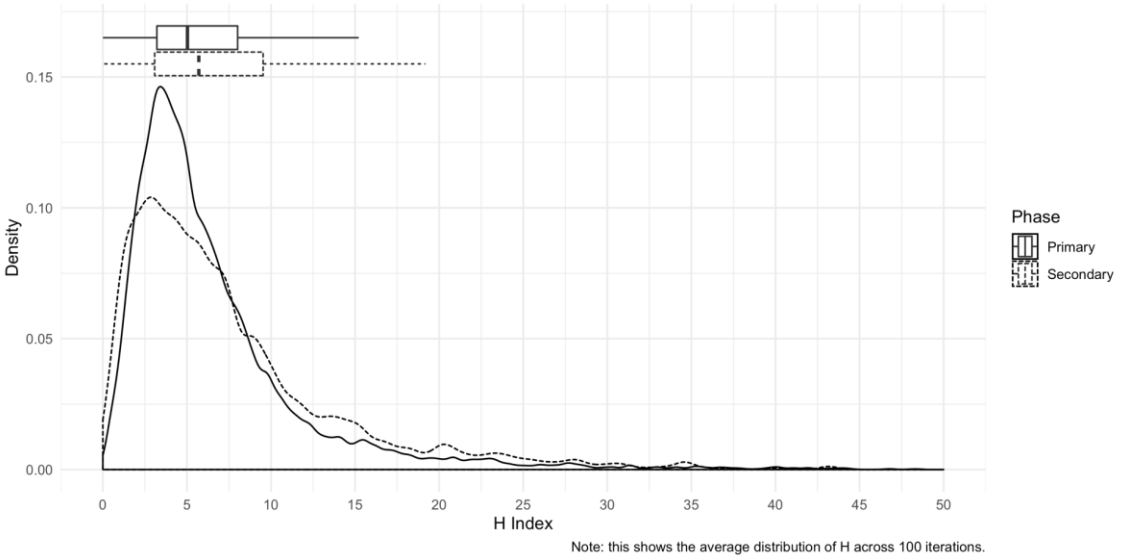
The section also described how trends changed between 2006 and 2019. It showed that students in 2019 typically shared their school with a greater proportion of students from different ethnic backgrounds than was the case in 2006. This was particularly true for White British students, with the trends varying for Asian and black students depending on their phase of education. In contrast, students from other white ethnic groups tended to share their schools with a greater proportion of students from the same ethnic groups in 2019 than was the case in 2006, both for those attending primary and secondary schools. These results answered question 2c.

5.3 Ethnic school segregation

Section 5.3 describes patterns of ethnic school segregation in England using the *H* index within pseudo-neighbourhoods. It initially focuses on describing the level of segregation in 2019 before investigating how trends have changed since 2006. As mentioned in the methodology section, only ethnically diverse pseudo-neighbourhoods, where no single ethnic group comprised of more than 90% of the pseudo-neighbourhood's state school population, was included in the following analysis. This ensured that ethnic school segregation was only calculated within pseudo-neighbourhoods that were sufficiently diverse to allow for a reliable measurement.

5.3.1 Current levels of ethnic school segregation in England

Figure 25: A density plot, with a boxplot overlayed, to show the distribution of *H* values for primary and secondary school pseudo-neighbourhoods in England (2019).



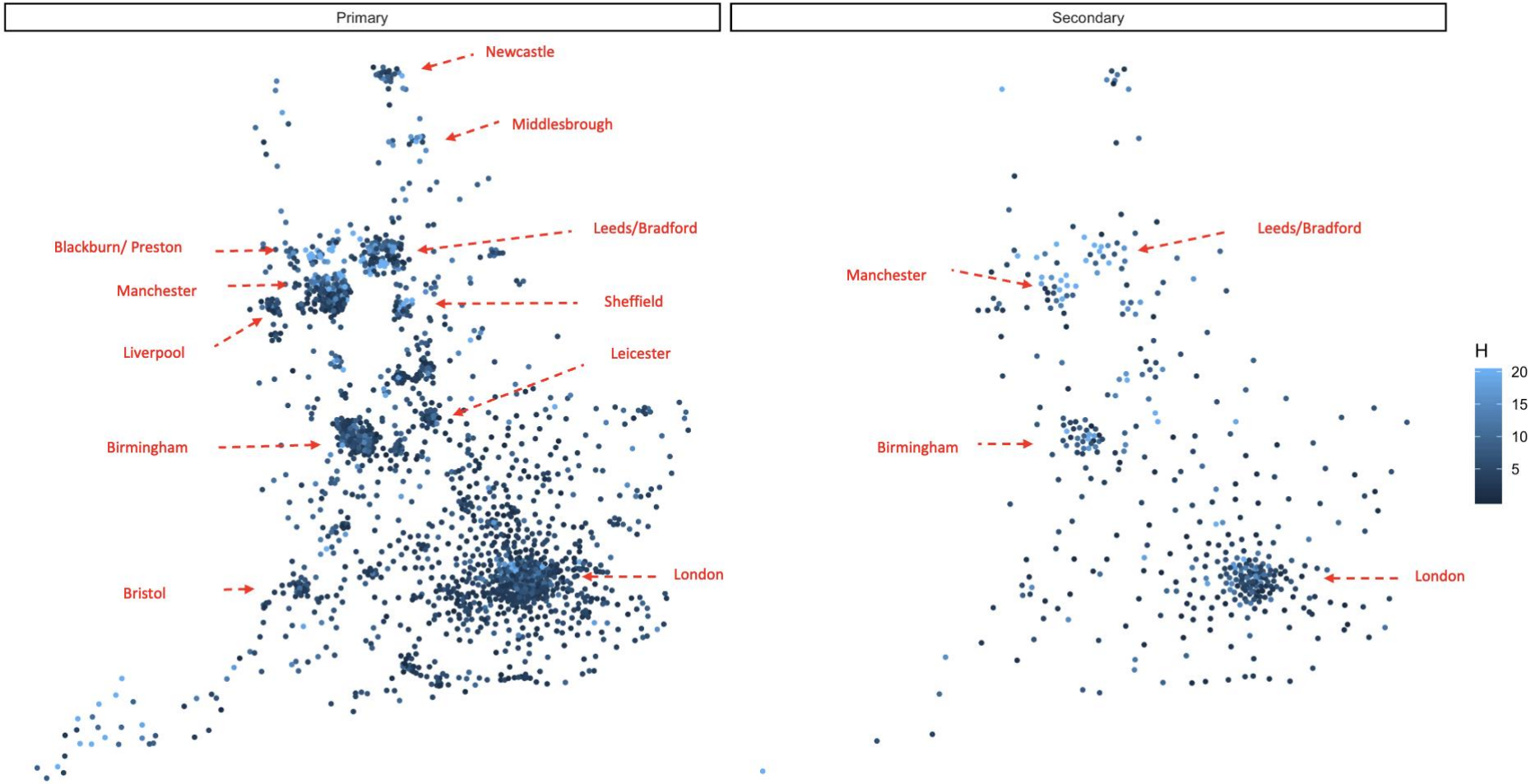
The analysis determined that the median H value for all pseudo-neighbourhoods in England for 2019 was 5.1, with a lower quartile of 3.2 and an upper quartile of 8.2. Primary school pseudo-neighbourhoods tended to exhibit slightly lower levels of ethnic school segregation than pseudo-neighbourhoods composed of secondary schools.[149] Primary school pseudo-neighbourhoods had a median H value of 5.0, which was 14% lower than the median H value for pseudo-neighbourhoods that were made up of secondary schools (H = 5.7). As

[149] This difference is most likely to reflect the fact that primary schools within the same pseudo-neighbourhood will tend to be closer to one another than secondary schools in the same pseudo-neighbourhood. This means that primary schools are more likely to recruit students living in the same geographic area, which due to patterns of ethnic residential segregation, mean that, on average, primary schools in the same pseudo-neighbourhood are likely to recruit from pools of students with more similar ethnic profiles. Further research would however be needed to confirm this hypothesis.

shown in the density plot below, H values for both types of pseudo-neighbourhoods had similar shaped distributions, with a strong positive skew. This meant that many pseudo-neighbourhoods experienced far higher levels of ethnic school segregation than the summary statistics would suggest.[150]

[150] All of the figures reported in this paragraph are average findings across 100 different sets of pseudo-neighbourhood boundaries. There was minimal variation across 100 iterations in the summary statistics. The Lower quartile ranged from 3.1 – 3.3, the median ranged from 5.0 – 5.3, the mean ranged from 6.7 –7.0 and the upper quartile ranged from 8.0 – 8.5. This suggests that, when investigating the level of segregation for of pseudo-neighbourhood across England the location (or zoning) of pseudo-neighbourhood boundaries does not have a significant impact on results.

Figure 26: Mapping the geographic patterns of H for pseudo-neighbourhoods in England (2019).



Each dot represents a pseud-neighbourhood. Note: each dot represents a neighborhood. Any neighborhood where H > 20 was recoded to 20 to create a clearer visual interpretation of the scale. Readers should be reminded that higher values of H represent higher levels of segregation

Pseudo-neighbourhoods in the north of England tended to experience higher levels of ethnic school segregation than those in the south. The median H values for primary school pseudo-neighbourhoods in the North East region (8.0), Yorkshire and the Humber (7.6) and the North West (6.7) were all higher than for pseudo-neighbourhoods in London (4.6), the East of England (4.0) and the South East (3.9). Similar trends were found for secondary school pseudo-neighbourhoods, with median H values for pseudo-neighbourhoods in Yorkshire (10.1), the North West (8.8) and the North East (7.6), higher than equivalent figures for pseudo-neighbourhoods in the East of England (4.8) South West (4.5), and the South East (2.9) [151]

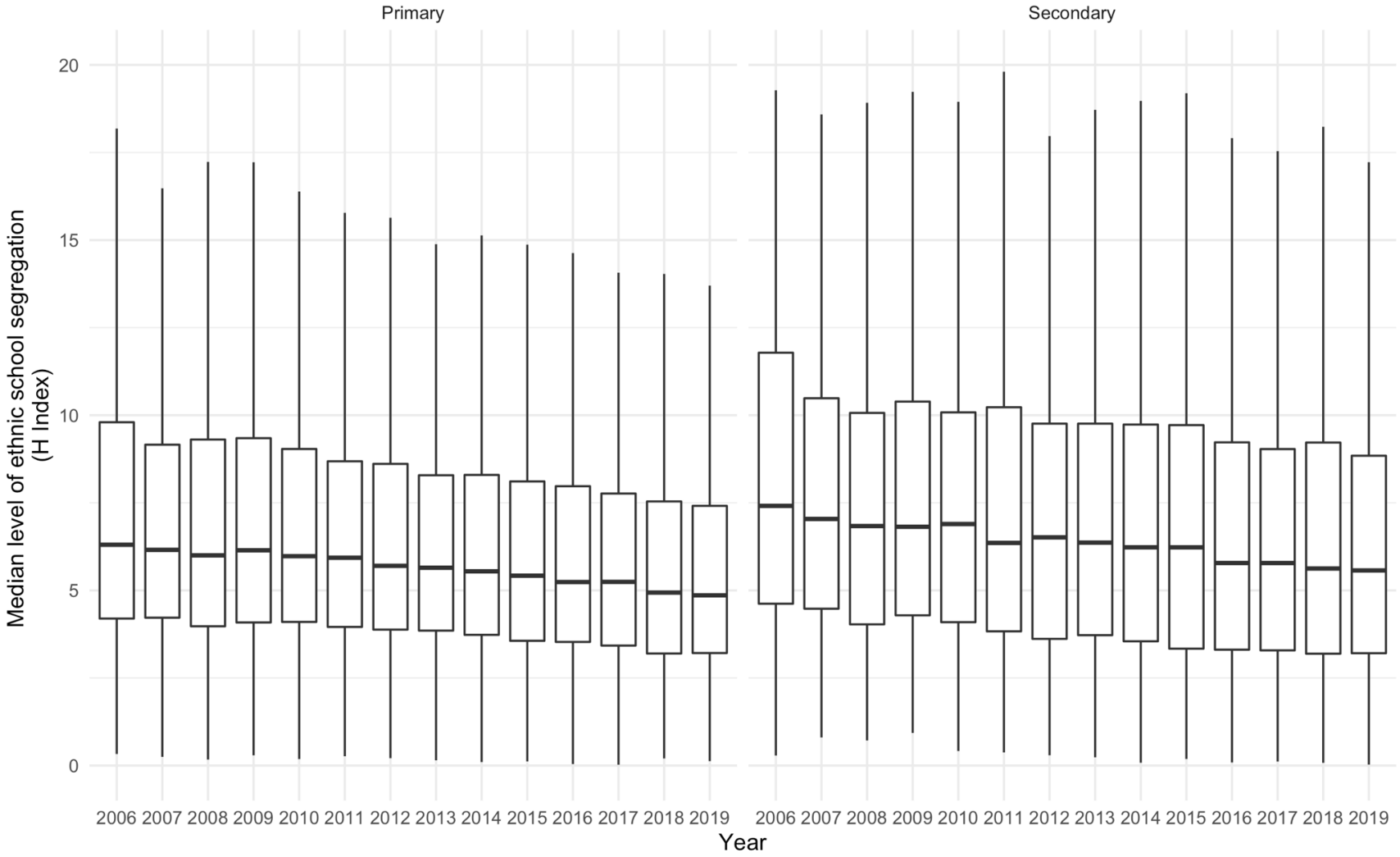
This north-south divide was particularly acute with regards to the number of pseudo-neighbourhoods with the highest levels of segregation, as defined by pseudo-neighbourhoods with in the top 5% of H values (H > 20). For example, more than 10% of primary school pseudo-neighbourhoods and 20% of secondary school pseudo-neighbourhoods in the region of the North West and Yorkshire and the Humber had H values above 20, compared to less than 3% in London, the East of England, and the South East. Remarkably, these two regions alone contained more pseudo-neighbourhoods where H was greater than 20 than the rest of the country combined.

[151] See Figure 79 in the appendix for a box plot of pseudo-neighbourhood H values by region.

5.3.2 Changes in the level of ethnic school segregation between 2006 – 2019

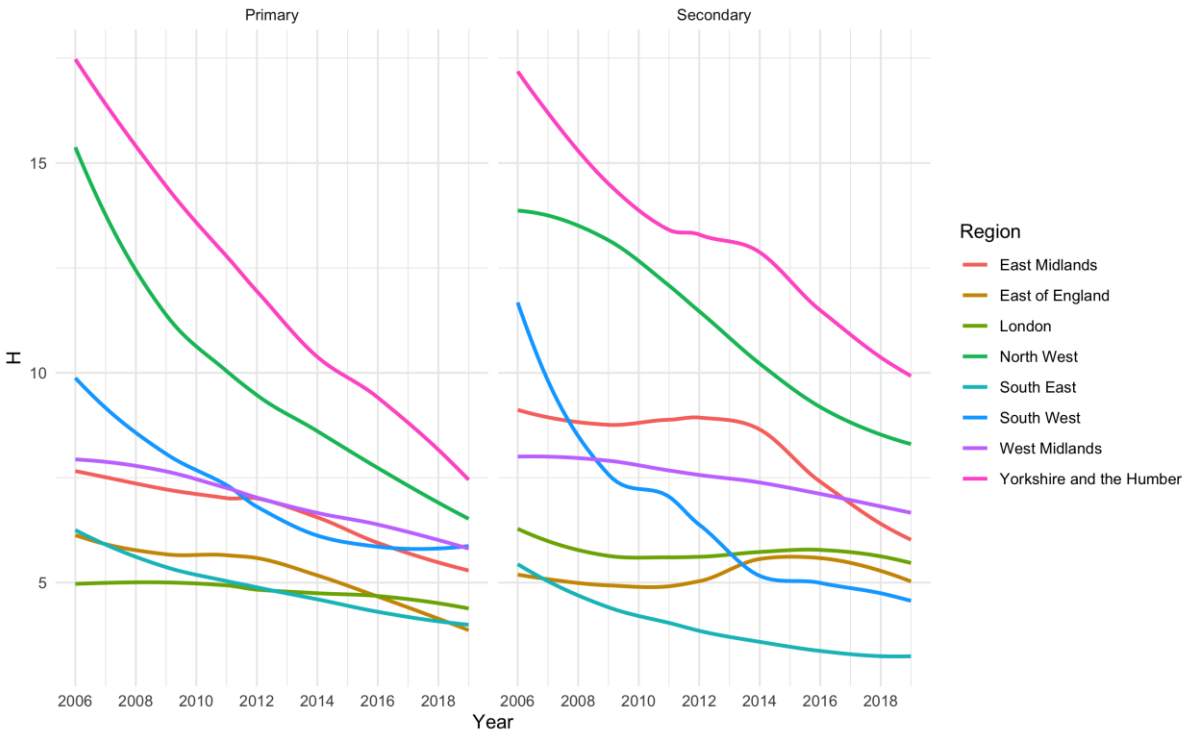
Between 2006 and 2019 the level of ethnic school segregation fell consistently from a median H value of 6.8 in 2006 to a median H value of 5.1 by 2019, representing a 25% reduction within the 13-year period. These trends were consistent for both primary and secondary school pseudo-neighbourhoods, which fell from a median H value of 6.8 to 5.0 (-26%) and 7.6 to 5.8 (-24%) respectively. Additionally, as shown in Figure 27, the lower quartile, the mean, the upper quartile, and the maximum H values all decreased, suggesting a reduction across the whole distribution of pseudo-neighbourhoods. These various statistics point to a substantial and consistent reduction in the average level of ethnic school segregation within local educational markets in England during this period.

Figure 27: The distribution of H values for ethnically diverse pseudo-neighbourhoods in England (2006 – 2019).



These national trends were also consistent for pseudo-neighbourhoods located in different parts of the country. Median H values for primary and secondary school pseudo-neighbourhoods fell across all nine regions between 2006 – 2019. As shown in Figure 6 below, pseudo-neighbourhoods in the northern regions, which had the highest average levels of segregation in 2006, tended to see the largest reductions in average segregation, which facilitated a degree of equalization between different parts of the country.

Figure 28: A line graph to show the changes in the average level of H for pseudo-neighbourhoods in different regions of England (2006 – 2019).



The three sensitivity checks ran by this study found no evidence to suggest that the changes in segregation were driven by changes in the sample of pseudo-neighbourhoods, the

percentage of White British students, or the level of random segregation. Ethnic school segregation fell by 1.4 in consistently diverse pseudo-neighbourhoods, which broadly aligns with the observed 1.7 reduction in the median *H* value for all ethnically diverse pseudo-neighbourhoods across England.[152] This suggests that, while the inclusion of newly diverse neighbourhoods was likely to have encouraged a greater reduction in the median *H* value, segregation also fell in pseudo-neighbourhoods that were consistently included in the sample. Secondly, the study found limited evidence to suggest an association between a change in the level of ethnic school segregation and White British students for pseudo-neighbourhoods for either phase.[153] Lastly there was no evidence to suggest that the level of "random" segregation changed significantly between 2006 – 2019 with it consistently accounting for around 12% of the observed median *H* values for primary school pseudo-neighbourhoods and 3% for secondary school pseud-neighbourhoods.[154] These findings provide welcome reassurance that changes in the level of segregation are reliable and reflect genuine changes the patterns of school attendance within the school system.

Summary of Section 5.3

These results offer the most up to date description of ethnic school segregation within local educational markets in England. The analysis found that pseudo-neighbourhoods in England had an average *H* value of 5.1, with an IQR of 3.2 - 8.2. Segregation was found to vary significantly within England, with primary and secondary school pseudo-neighbourhoods in the north of the country tending to experience higher levels of ethnic school segregation

[152] See Section 11.17 in the appendix.
[153] See Section 11.18 in the appendix.
[154] See Section 11.19 in the appendix.

than pseudo-neighbourhoods in the south. This was both in relation to average pseudo-neighbourhood *H* values and the number of pseudo-neighbourhoods with the highest levels of segregation (*H* > 20). In particular, the regions of Yorkshire and the Humber and the North West had large numbers of pseudo-neighbourhoods with exceptionally high levels of ethnic school segregation. These findings address research question 3a.

Section 5.3.2 then showed that ethnic school segregation in England fell from a median *H* value of 6.8 to 5.2 between 2006 - 2019. This represents a 25% reduction in the average level of segregation within a 13-year period, with similar changes observed within primary and secondary school pseudo-neighbourhoods. Pseudo-neighbourhoods in the north of England tended to experience larger reductions in ethnic school segregation than those in the south, again for both phases. Three sensitivity checks also found that changes in the sample of pseudo-neighbourhoods, the percentage of White British students and the level of random segregation were unlikely to be responsible for the observed changes. This section answered 3b of the research questions.

6 Exploring the predictors of ethnic school segregation

Chapter 6 moves beyond descriptive statistics to investigate the predictors of ethnic school segregation for primary and secondary school pseudo-neighbourhoods. Section 6.1 investigates whether a pseudo-neighbourhood's level of ethnic school segregation is associated with its population density, socio-economic profile, proportion of religious schools and its average OFSTED rating. It also investigates whether ethnic school segregation differs between secondary school pseudo-neighbourhoods that do, or do not, contain a grammar school, as well as considering whether these findings are consistent once regional differences in ethnic school segregation are accounted for. These results address research question 4a – 4k.

Section 6.2 then builds upon this investigation by comparing the contributions towards ethnic school segregation made by schools located within the same pseudo-neighbourhood. It explores whether academically selective schools or those with a religious denomination disproportionately contribute to ethnic school segregation. Additionally, it investigates whether there is an association between a school's OFSTED rating and their contribution towards ethnic school segregation. These results answer research question 5a – 5g.

6.1 Pseudo- neighbourhood characteristics

The following sub section describes the results from the pseudo-neighbourhood level meta regression analysis. As described in the methodology chapter, the analysis was conducted

on all ethnically diverse pseudo-neighbourhoods in England, as defined by any pseudo-neighbourhood where the largest ethnic group constitutes less than 90% of the student population. Two sets of models were used that either used primary or secondary school data with models identified using p or s subscripts (e.g., Model 1_p and Model 1_s).

As a *meta*-regression analysis, all figures in the output are mean averages across 100 different models, where each model is based on a different set of pseudo-neighbourhoods. Readers should be reminded that the outcome variable (*H*) was log transformed and, with the exception the binary variables, all independent variables were standardised. Results are therefore described in relation to "the percentage changes in H" that were associated with "a one standard deviation difference" in any given predictor variable.[155] Lastly, as all findings are based on observational data, none of the following associations should be assumed to be causal.

[155] To discuss the results in this format the thesis exponentiated each coefficient, subtracted one from the result, and then multiplied the answer by 100. For a more in depth explanation see (UCLA Institute for Digital Research and Education, 2020).

Primary school meta regression output

Predictor Variables	Dependent variable: Standardised log(H Index)				
	Control Variables (Model 1_p)	**+ Population Density (Model 2_p)**	**+ Socio-Economic Characteristics (Model 3_p)**	**+ School Characteristics (Model 4_p)**	**+ Regional Differences (Model 5_p)**
Population Density†		-0.018 (0.018)	-0.013 (0.022)	0.03 (0.023)	-0.053 (0.025)
House Price†			-0.169*** (0.02)	-0.182*** (0.019)	-0.150*** (0.032)
FSM†			0.068** (0.022)	0.065** (0.023)	0.052* (0.023)
Proportion of schools that are religious†				0.145*** (0.017)	0.132*** (0.018)
Average Ofsted Rating†				0.003 (0.017)	-0.012 (0.017)
Control Variables					
Number of schools	0.271*** (0.017)	0.277*** (0.018)	0.234*** (0.018)	0.234*** (0.017)	0.261*** (0.018)
East of England					-0.114 (0.073)
London					0.282** (0.095)
North East					0.283** (0.123)
North West					0.164** (0.073)
South East					-0.169** (0.071)
South West					0.139 (0.078)
West Midlands					0.042 (0.071)
Yorkshire and the Humber					0.298*** (0.076)
Intercept	1.615*** (0.017)	1.615*** (0.017)	1.616*** (0.016)	1.615*** (0.016)	1.549*** (0.057)
Observations	1820.28	1820.28	1820.28	1815.48	1815.48
R^2	0.122	0.123	0.193	0.224	0.265
Adjusted R^2	0.122	0.122	0.191	0.221	0.259
Adjusted S.E	0.707	0.707	0.679	0.665	0.649
F Statistic	254.536	128.131	108.563	87.089	46.465

Note: *p<0.1; **p<0.05; ***p<0.01

† Variable is standardised

Secondary school meta regression output

	Dependent variable: Standardised log(H Index)				
Predictor Variables	**Control Variables (Model 1_s)**	**+ Population Density (Model 2_s)**	**+ Socio-Economic Characteristics (Model 3_s)**	**+ School Characteristics (Model 4_s)**	**+ Regional Differences (Model 5_s)**
Population Density†		0.100**	0.101	0.062	0.026
		(0.043)	(0.056)	(0.057)	(0.058)
House Price†			-0.228***	-0.229***	-0.229**
			(0.048)	(0.049)	(0.08)
FSM†			0.176***	0.214***	0.133*
			(0.052)	(0.054)	(0.057)
Proportion of schools that are religious†				0.075	0.059
				(0.043)	(0.043)
Contains a grammar school (Yes = 1)				0.240**	0.264**
				(0.095)	(0.094)
Average Ofsted Rating†				0.039	0.005
				(0.042)	(0.041)
Control Variables					
Number of schools	0.456***	0.424***	0.361***	0.339***	0.36***
	(0.041)	(0.044)	(0.041)	(0.042)	(0.042)
East of England					-0.031
					(0.173)
London					0.261
					(0.228)
North East					-0.322
					(0.277)
North West					0.049
					(0.17)
South East					-0.375*
					(0.173)
South West					-0.281
					(0.192)
West Midlands					-0.017
					(0.163)
Yorkshire and the Humber					0.272
					(0.179)
Intercept	1.655***	1.656***	1.656***	1.607***	1.632***
	(0.04)	(0.04)	(0.037)	(0.041)	(0.139)
Observations	397.5	397.5	397.5	397.18	397.18
R^2	0.236	0.248	0.364	0.383	0.436
Adjusted R^2	0.235	0.244	0.357	0.372	0.414
Adjusted S.E	0.795	0.79	0.728	0.72	0.695
F Statistic	124.242	65.767	56.478	34.72	19.823

Note: *p<0.1; **p<0.05; ***p<0.01

† Variable is standardised

6.1.1 Control variables

Model 1_p and 1_s begin by regressing *H* on the number of schools contained within the pseudo-neighbourhood to provide useful reference points to interpret later results. As expected, ethnic school segregation tended to be higher in both primary and secondary school pseudo-neighbourhoods that contained a greater number of schools ($p<0.001$) and the substantive strength, direction, and statistical significance of these associations remained reasonably consistent across both sets of models. The primary school model explained 12.2% of the variation in H values, while 23.6% of the variation between the H values of secondary school pseudo-neighbourhoods was explained in Model 1_s.

6.1.2 Population Density

The thesis found no evidence to suggest that the population density of the area surrounding a primary school pseudo-neighbourhood had an association with its observed level of ethnic school segregation. None of the models identified a statistically significant association, the estimated regression coefficients for the population density variable were consistently small, and the direction of the predicted associations varied between models. Furthermore, the inclusion of the population density variable in Model 2_s only marginally increased the adjusted R^2 value of Model 3_s, which suggests its inclusion only made a minimal improvement to the models' predictive power. Consequently, the thesis fails to reject the null hypothesis of question 4a.

In contrast, the analysis found some evidence to suggest that the population density of the area surrounding a secondary school pseudo-neighbourhood *may* have an association with its level of ethnic school segregation. Model 2_s predicted that a one standard deviation difference in the population density of a secondary school pseudo-neighbourhood was associated with a level of ethnic school segregation that was 10.5% higher, after controlling for the number of schools within the pseudo-neighbourhood (p=0.047). While this statistically significant result provides sufficient evidence to reject the null hypothesis of question 4b, a strong degree of caution is required when interpreting the finding. The variables inclusion only increased the adjusted R^2 value by 0.012 between Model 1_s to 2_s, which suggests it offered limited additional explanatory power, and none of the subsequent models identified a statistically significant association. This indicates that the variation in *H* values for secondary school pseudo-neighbourhoods is better explained by the predictor variables of later models. Consequently, the thesis concludes that while there is evidence to suggest that the population density of the area surrounding a secondary school pseudo-neighbourhood is associated with its observed level of ethnic school segregation, it is unlikely to be a substantial or robust relationship, once other factors are controlled for.

6.1.3 Socio-economic profile

Including variables that describe the socio-economic profile of pseudo-neighbourhoods substantially improved the predictive power of both sets of meta-regression models. The adjusted R^2 value of Model 3_p increased by 7 percentage points from 12.3 to 19.3, while the adjusted R^2 value of secondary school pseudo-neighbourhood Model 3_s improved by 11.6 percentage points from 24.8 to 36.4.

Both sets of models found strong evidence to suggest that pseudo-neighbourhoods with higher average house prices tended to experience lower levels of ethnic school segregation, after accounting for other independent variables included in the models. Model 3_p, for example, estimated that a one standard deviation increase in the average house price of the area surrounding a primary school pseudo-neighbourhood was associated with an H value that was 15.5% lower, once other variables had been accounted for ($p<0.001$).[156] The association was found to be even stronger for secondary school pseudo-neighbourhoods. Model 3_s estimated that for each standard deviation increase in a secondary school pseudo-neighbourhood's average house price, the predicted level of ethnic school segregation was 20.4% lower ($p<0.001$), after controlling for other variables in the model.[157] The strength, direction and statistical significance of the house price coefficients were broadly consistent in later models, which suggests these estimates are robust, even after controlling for regional difference in model 5_p and 5_s. Consequently, the thesis finds sufficient evidence to reject the null hypothesis of question 4c and 4d and conclude that there is a statistically significant, negative association between a pseudo-neighbourhood's average house price and the level of ethnic school segregation for both primary and secondary school pseudo-neighbourhoods.[158]

Similarly, the analysis found robust, statistically significant associations between ethnic school segregation and the percentage of students who were eligible for free school meals

[156] $(e^{-0.169} - 1) \times 100 = -15.54$

[157] $(e^{-0.228} - 1) \times 100 = -20.38$

[158] This lends supportive evidence to the work of competition theorists who argue that neighbourhoods where residents have greater economic resources are more able to share spaces with other ethnic groups than places where families have fewer economic resource.

(FSM) in primary and secondary school pseudo-neighbourhoods. Model 3_p, for example, predicted that for every standard deviation increase in the rate of FSM, the predicted level of ethnic school segregation within primary school pseudo-neighbourhoods increased by 7.0% (p=0.017), once other variables in the model had been controlled for.[159] The association for secondary school pseudo-neighbourhoods was comparatively stronger. Model 3_s predicted that a one standard deviation difference in the rate of FSM was associated with a secondary school H value that was 19.2% higher, on average (p = 0.006).[160] Readers should note however that once regional differences were accounted for in Model 5_p and 5_s the p values and beta coefficients for both sets of models fell, which suggest that some of this associations are explained by regional differences. Nonetheless, when viewed collectively, the thesis found sufficient evidence to reject the null hypotheses of question 4e and 4f and conclude that there is a robust positive association between the level of ethnic school segregation for primary and secondary school pseudo-neighbourhoods and the observed rate of FSM.

6.1.4 School Characteristics

Model 4_p and 4_s included additional variables to explore the associations between the characteristics of a pseudo-neighbourhood's schools and its level of ethnic school segregation. By including these additional variables, the adjusted R^2 value for Model 4_p increased by 3.1 percentage points with 22.4% of the variation in *H* explained while the

[159] $(e^{0.068} - 1) \times 100 = 7.03$
[160] $(e^{0.176} - 1) \times 100 = 19.24$

adjusted R^2 value for Model 4_s rose by slightly less, increasing 1.9 percentage points to 38.3%.

In contrast to earlier findings, the analysis found that the association between ethnic school segregation and the proportion of religious schools in a pseudo-neighbourhood varied between primary and secondary school pseudo-neighbourhoods. Model 4_p provided strong evidence to suggest that primary school pseudo-neighbourhoods with a greater proportion of religious schools, on average, experienced higher levels of ethnic school segregation. It predicted that for each standard deviation increase in the proportion of religious schools within a primary school pseudo-neighbourhood, the expected level of ethnic school segregation rose by 15.6%, after controlling for other variables in the model ($p<0.001$).[161] The strength, direction and statistical significance of this association was broadly consistent after regional variation in H values were accounted for in Model 5_p, which further suggests an independent and robust relationship with ethnic school segregation. Consequently, the thesis rejects the null hypothesis of question 4h and concludes that there is a positive association between the proportion of schools in a primary school pseudo-neighbourhood that have a religious affiliation and its observed level of ethnic school segregation.[162]

However, the thesis found no evidence to suggest that levels of ethnic school segregation varied between secondary school pseudo-neighbourhoods with different proportions of religious schools. Neither model 4_s nor 5_s found a statistically significant association between these variables and the estimated strengths of both associations were small.

[161] $(e^{0.145} - 1) \times 100 = 15.60$

[162] As explored in Section 2.3.3.2, this could reflect overlapping ethnic and religious identities, which mean that recruiting from a single religious community also results in more ethnically homogenous school cohorts.

Consequently, the thesis failed to reject the null hypothesis of question 4i and therefore concludes there is insufficient evidence to suggest an association between the proportion of schools in a secondary school pseudo-neighbourhood that have a religious affiliation and its observed level of ethnic school segregation.[163]

The analysis did, however, find a statistically significant difference in the level of ethnic school segregation between secondary school pseudo-neighbourhoods that contained a grammar school and those that did not. Model 4_s predicted that ethnic school segregation was 27.1% higher, on average, in pseudo-neighbourhoods that contained a grammar school than in pseudo-neighbourhoods that did not ($p = 0.026$), after controlling for other variables in the model.[164] Additionally, as the strength, direction and statistical significance of this association remained consistent once regional differences in H were accounted for in Model 5_s, the thesis concludes that this difference is robust. Therefore, the study rejects the null hypothesis of question 4g and concludes that secondary school pseudo-neighbourhoods that contain grammar schools, on average, exhibit higher levels of ethnic school segregation.

Lastly, none of the models found a statistically significant relationship between a pseudo-neighbourhood's level of ethnic school segregation and their average OFSTED result for either primary ($p=0.552$) or secondary school pseudo-neighbourhoods ($p=0.359$), and the size of the coefficients were consistently small. The thesis therefore fails to reject the null

[163] As discussed later in Section 8.4.4, this could reflect a difference in parental perceptions of religious primary and secondary schools, with the former being considered more "religious" than the latter, and therefore less palatable to families from a broad range of ethnic backgrounds.
[164] $(e^{0.240} - 1) \times 100 = 27.12$

hypotheses of question 4j and 4k and instead concludes that there is insufficient evidence to suggest there is an association between the average OFSTED rating of schools within a primary or secondary school pseudo-neighbourhood and its level of ethnic school segregation.

Summary of Section 6.1

This section has presented evidence to suggest that both primary and secondary school pseudo-neighbourhoods tend to have higher levels of ethnic school segregation if the surrounding areas have a lower socio-economic profile, where local house prices are lower, and more students are eligible for free school meals. Ethnic school segregation also tends to be higher in primary school pseudo-neighbourhoods with a greater proportion of religious schools and in secondary school pseudo-neighbourhoods where at least one of their schools is a grammar school. In contrast, there was insufficient evidence to suggest any association between a pseudo-neighbourhood's level of ethnic school segregation and its population density or the average OFSTED rating of its schools, either for primary or secondary school pseudo-neighbourhoods. Similarly, no association was found between the proportion of schools with a religious affiliation and the level of ethnic school segregation for secondary school pseudo-neighbourhoods.

6.2 School level factors

The following section investigates whether different types of schools tend to make different sized contributions towards ethnic school segregation. Initially it investigates whether grammar schools contribute disproportionately to ethnic school segregation before studying the relative contributions of schools with different religious denominations. It then finishes by exploring whether there is an association between a school's OFTSED rating and their contribution towards ethnic school segregation.

All models included in this section use a school's standardized H_j value as the outcome variable. As discussed in Section 4.5.2, this variable describes the contribution made by school *j* to their pseudo-neighbourhood's level of ethnic school segregation, relative to other schools in the same pseudo-neighbourhood. Each of the three analyses used different samples.[165] Section 6.2.1 used Sample 3, which includes all ethnically diverse pseudo-neighbourhoods which contained at least one grammar school. Section 6.2.2 used Sample 4, which includes all ethnically diverse pseudo-neighbourhoods where at least one of the schools had a religious denomination. Finally, Section 6.2.3 used Sample 2, which includes all ethnically diverse pseudo-neighbourhoods. As the three samples did not overlap, the study conducted each analysis separately without fitting a combined model.

6.2.1 The contribution of academically selective schools to ethnic school segregation

[165] See Section 4.2

The following section explores the differences between the contributions towards ethnic school segregation made by academically selective (grammar) and non-selective schools. As shown in Figure 29, the grammar and non-selective schools included within the sample educated students with different collective ethnic profiles. 25% of students in grammar schools were from an Asian background, compared to only 13% in non-selective schools, which meant grammar schools educated a smaller proportion of students from all other ethnic groups. 59% of students who attended grammar schools were White British, compared to 66% who attended non-selective schools, 6% were from a Black ethnic group, compared with 8% of students in non-selective schools and 4% came from another white ethnic group, compared to 7% in non-selective schools. These differences suggest that grammar schools, as a school sector, educated a set of students whose ethnic profile differed from non-selective schools operating in the same set of pseudo-neighbourhoods.

Figure 29: Comparing the ethnic profile of grammar and non-selective schools contained in Sample 3.

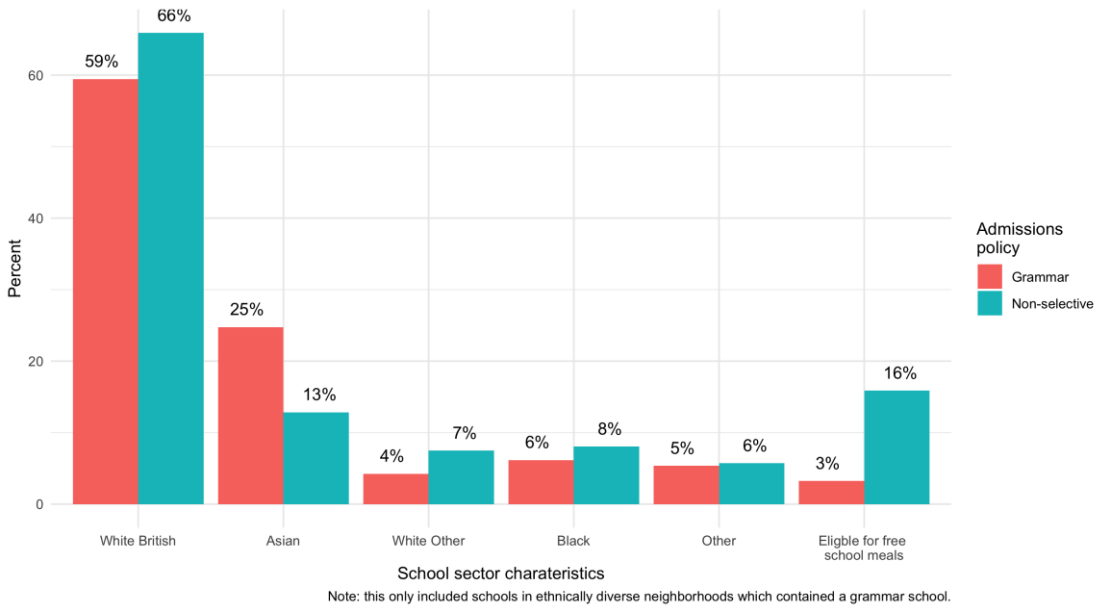
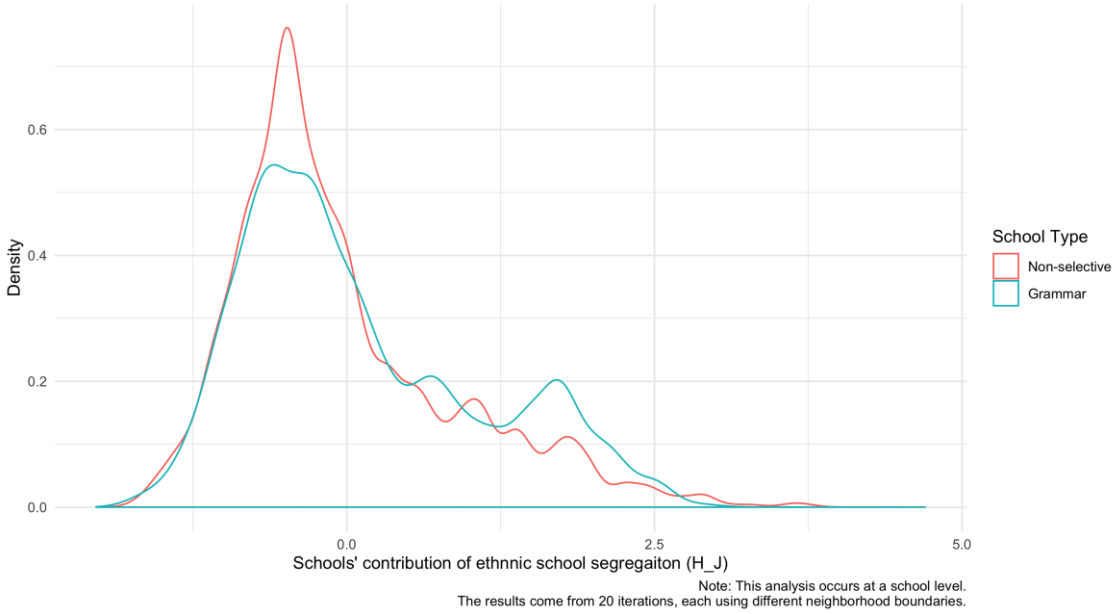


Figure 30: Comparing the contributions to ethnic school segregation made by grammar and non–selective schools in ethnically diverse pseudo-neighbourhoods (2019).



However, as the density plot above shows, most grammar and non-selective schools tended to make similar contributions to their neighbourhood's level of ethnic school segregation in 2019. However, as shown by the divergence between the H_j values of 0.8 and 2.5, there were a greater number of grammar schools that made particularly large contributions to the level of ethnic school segregation of their neighbourhood ($H_j > 1$).[166] For example, 23% of grammar schools made a contribution to ethnic school segregation which was more than one standard deviation greater than the average contribution of other schools in their neighbourhood ($H_j > 1$) compared to 18% for non-selective schools.

[166] These "high contributors" are marked by a bulge in the density plot between $1 < H_j < 2.5$.

Table 27: Meta- regression analysis (Model 6_s) to compare the relative contribution towards ethnic school segregation of grammar and non-selective schools.[167]

	Outcome variable: Contribution to ethnic school segregation (H_j)
	(Model 6_s)
Grammar School	0.142
	(0.088)
Constant	-0.036
	(0.044)
Observations	589.55
R Squared Value	0.005
Adjusted R Squared	0.003
Adjusted Standard Error	0.931
F Statistic	3.078
Note:	*p<0.1; **p<0.05; ***p<0.01

However, the results from the meta regression analysis found there was no statistically significant difference between the contribution of grammar and non-selective schools towards their pseudo-neighbourhood's level of ethnic school segregation ($\theta = 0.142$, $p = 0.17$). The model also failed to explain any substantial amount of the variation in H_j values with the R^2 value only 0.005. This means that the thesis failed to find sufficient evidence to reject the null hypothesis and this study therefore found no evidence to suggest that

[167] Following negative results of the Breusch–Pagan test, the standard errors reported are heteroscedasticity consistent, calculated using the robustbase r library (Maechler *et al.*, 2020). See Section 11.23 in the appendix for further details.

schools with academically selective admissions policies contributed more to ethnic school segregation than non-selective schools. This answers 5a of the research questions.

6.2.2 The contribution of religious schools to ethnic school segregation

The following subsection explores the differences between the contribution towards ethnic school segregation made by religious and non-religious schools. As Figure 31 shows, the majority of religious schools included in the sample were Christian. 22% of schools were linked to the Church of England, 12% were affiliated to the Catholic church and 0.8% were linked to another Christian faith. Only 0.8% of schools in the sample had a non-Christian affiliation. 0.4% were linked to a Jewish faith, 0.3% were linked to an Islamic faith and 0.2% were linked to another non-Christian religion, such as Buddhism, Hinduism, or Sikhism.[168]

There were also important differences in the religious make-up of the primary and secondary school sectors. Firstly, religious schools constituted a larger proportion of the primary school sector (33.0%) compared to the secondary school system (20.1%). Church of England schools were also more common in the primary school sector, while Catholic schools, and those linked to smaller minority religions such as Jewish, Muslim, or Other Non-Christian faiths, were more common in the secondary school system.

[168] For a map which outlines the location of schools with different religious denominations see Figure **80**.

The ethnic composition of students attending religious schools tended to differ from the ethnic profile of students attending non-religious schools, and the nature of these differences varied between the primary and secondary school sectors, as captured in Figure 32. Religious primary schools educated a larger proportion of White British students (67% vs 63%) and a smaller proportion from Asian backgrounds (15% vs 10%), while the proportion of students from Black and other White ethnic groups were broadly similar. In contrast, religious secondary schools educated fewer White British students (54% vs 60%), and more students from Black (14% vs 10%) and other White ethnic backgrounds (9% vs 7%), with similar numbers of Asian students to non-religious secondary schools.

Differences in the collective ethnic composition of schools was more pronounced once the study split the religious school sector into their respective denominations; with schools linked to minority religions (Jewish, Muslim, Other non-Christian faith) educating students with particularly distinct ethnic profiles. [169] Jewish schools, for example, educated substantially more White British students (75% vs 59%) and fewer Asian students (4% vs 18%) than non-religious schools, on average while 75% of students in Muslim schools came from an Asian background compared to 18% in the non-religious school sector.[170]

There were also differences between the collective ethnic profiles of Christian and non-religious schools, but these were less acute. Church of England schools educated a greater proportion of White British students (67% vs 59%) and a smaller percentage of Asian

[169] There are no substantively important differences in the ethnic profile of different religious schools between primary and secondary school sectors and so the profiles are described collectively. However, readers can consult Figure 81 in Section 11.20 in the appendix for more detail.
[170] The figures for other non-Christian schools are broadly similar to Muslim schools.

students (11% vs 18%) than non-religious schools. In contrast, Catholic schools educated more students from black ethnic groups (16% vs 9%) and from other white ethnic groups (14% vs 8%) than non-religious schools. Additionally, they educated fewer White British students (52% vs 59%).[171]

[171] Schools with another Christian faith educated a student body whose ethnic profile more closely resembles non-religious schools. Like with CoE schools, they educate a greater proportion of White British students and a lower proportion of Asian students than non-religious schools, though these differences are less pronounced. For Black, White Other and Other ethnic groups the proportions educated were similar.

Figure 31: Bar chart to describe the religious denomination of schools in the sample.

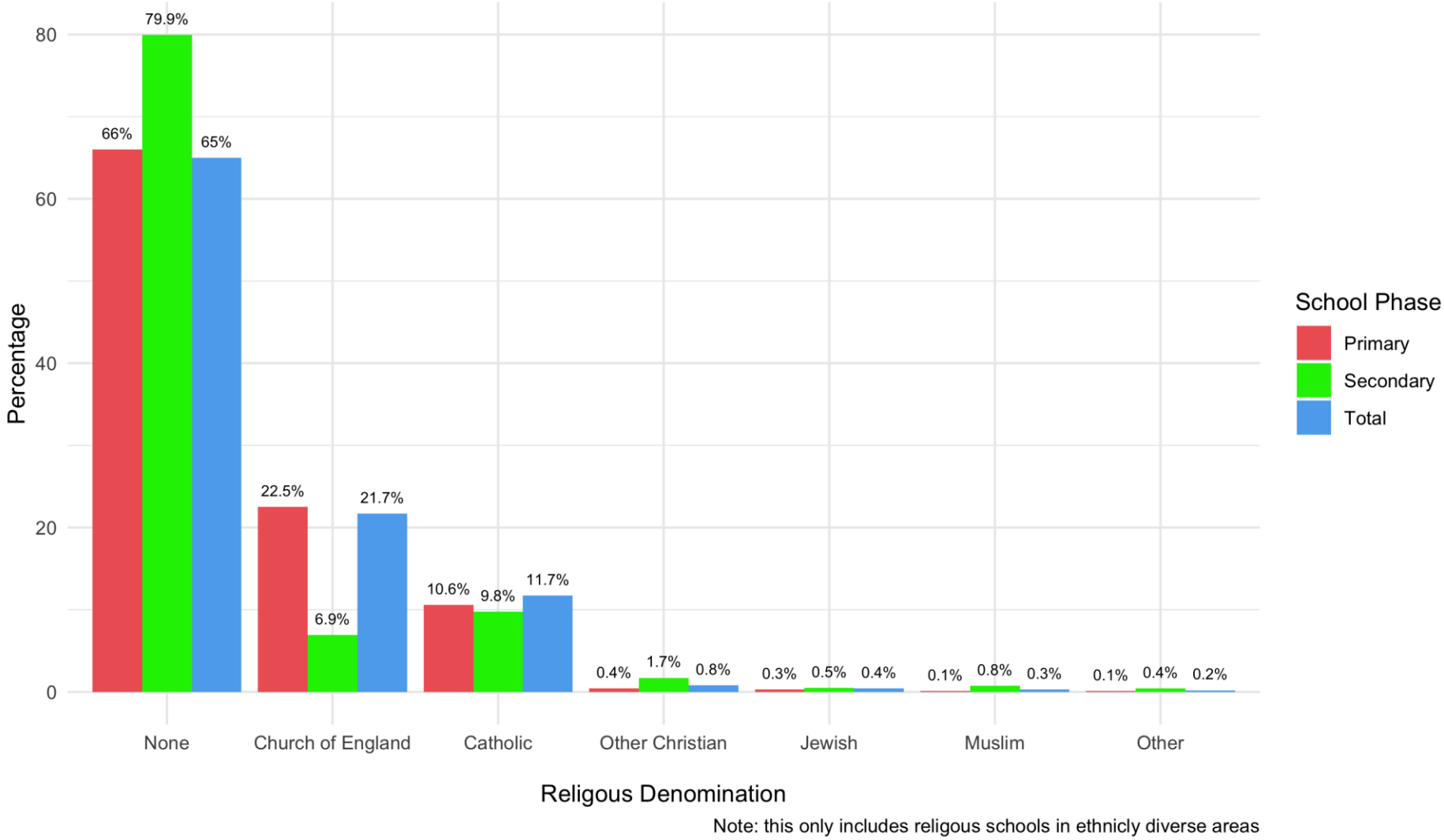


Figure 32: The ethnic profile of religious and non-religious schools in the sample

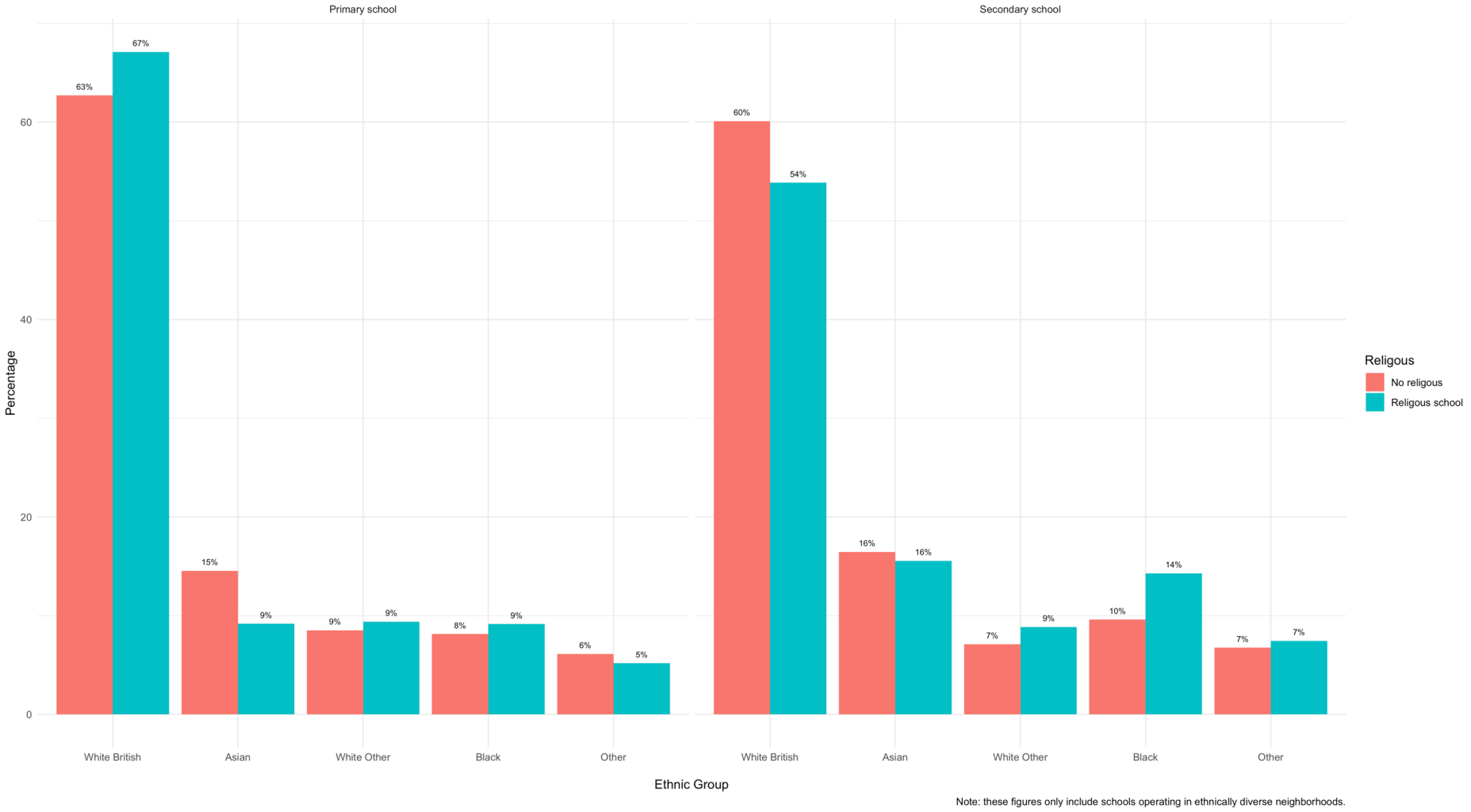
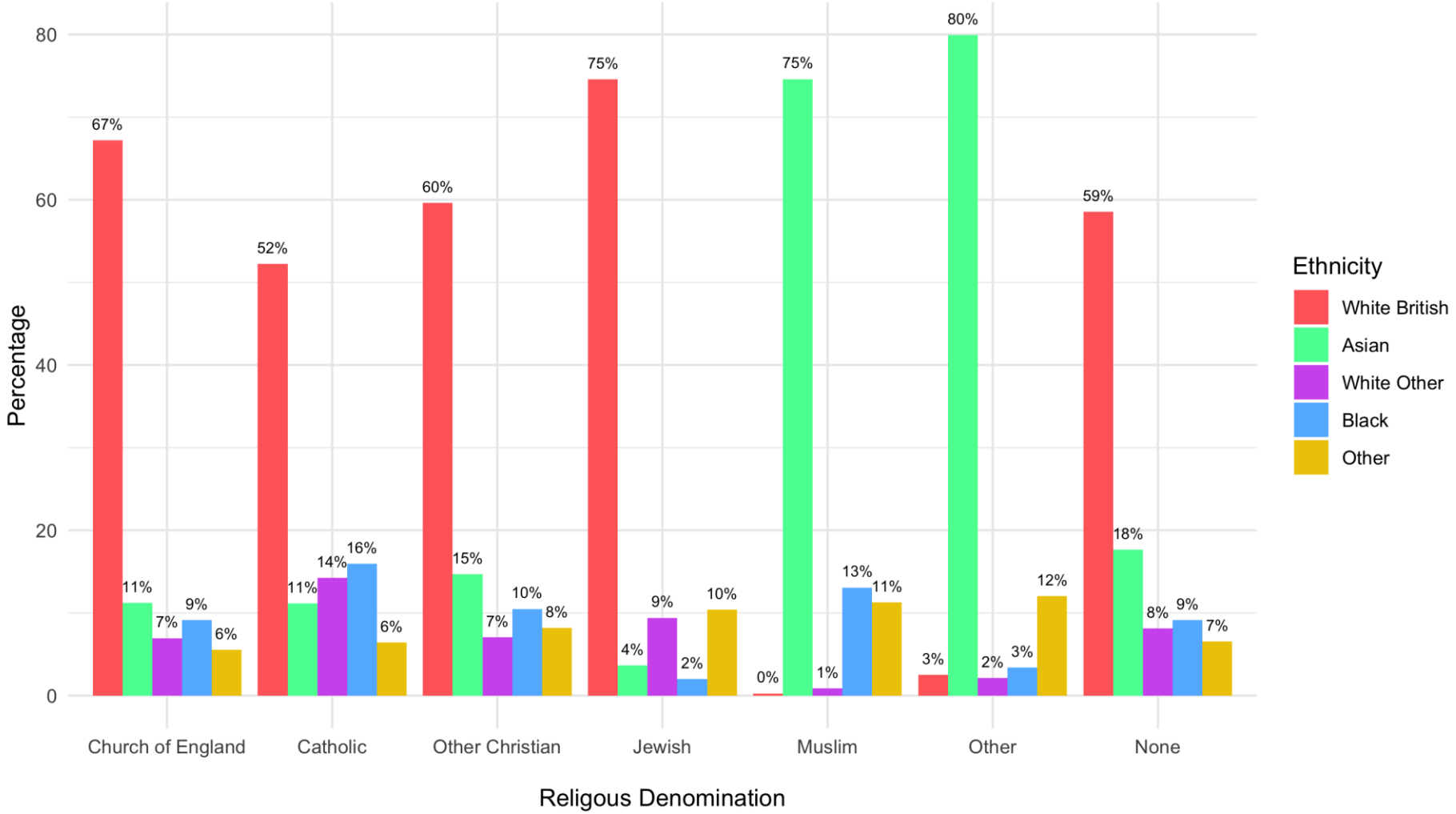


Figure 33: The ethnic profile of the schools contained in the sample by religious denomination (2019).



Note: these figures are for the collective student body of schools with different religious affiliations operating in ethnically diverse neighborhoods.

Table 28: Regression output to show the difference in the contribution to ethnic school segregation of schools with different religious affiliations.[172]

	Outcome variable: Contribution to ethnic school segregation (H_j)			
Variable	(Model 7_p)	(Model 8_p)	(Model 7_s)	(Model 8_s)
Religious	0.51*** (0.018)		0.37*** (0.045)	
Church of England		0.28*** (0.02)		0.14* (0.069)
Catholic		0.94*** (0.027)		0.43*** (0.06)
Another Christian Faith		0.49*** (0.116)		0.11 (0.131)
Jewish		1.65*** (0.145)		1.18*** (0.252)
Muslim		1.00*** (0.242)		1.34*** (0.199)
Other		1.68*** (0.241)		1.42*** (0.263)
Constant	-0.189*** (0.011)	-0.19*** (0.01)	-0.09*** (0.023)	-0.09*** (0.022)
Observations	10972.31	10972.31	2210.59	2210.59
R Squared Value	0.071	0.119	0.03	0.06
Adjusted R Squared	0.071	0.118	0.03	0.057
Adjusted Standard Error	0.891	0.868	0.92	0.906
F Statistic	843.329	246.844	68.713	23.416

Results from the meta-regression analysis, shown in Table 28, found that both primary and secondary schools with a religious affiliation made larger contributions to ethnic school segregation, on average, than non-religious schools in the same pseudo-neighbourhood. Model 7_p, for example, found that primary schools with a religious denomination contributed 0.51 standard deviations more to ethnic school segregation, on average, than non-religious schools in the same pseudo-neighbourhood ($p<0.001$). This provides sufficient evidence to reject the null hypothesis of question 5b. Similarly, Model 7_s found that religious secondary schools contributed 0.37 standard deviations more to ethnic school

[172] Following negative results of the Breusch–Pagan test in both models, the standard errors reported are heteroscedasticity consistent, calculated using the robustbase r library (Maechler *et al.*, 2020). See Section 11.23 in the appendix for further details.

segregation, on average, than non-religious schools in the same pseudo-neighbourhood ($p<0.001$). This provides sufficient evidence to reject the null hypothesis of question 5c.

The analysis also found strong evidence to suggest that primary and secondary schools with particular religious denominations contributed more to ethnic school segregation than others, with schools linked to minority religious groups tending to make larger contributions relative to Church of England schools. Jewish primary schools, for example, contributed 1.65 standard deviations more to ethnic school segregation, on average, than non-religious primary schools in the same pseudo-neighbourhood. Similarly, Jewish secondary schools contributed 1.176 standard deviations more to ethnic school segregation than non-religious secondary schools in the same pseudo-neighbourhood ($p<0.001$).

The figures for Muslim schools are similar. Model 8_p found that Muslim primary schools contributed 1.00 standard deviations more, on average, to ethnic school segregation than non-religious primary schools in the same pseudo-neighbourhood ($p < 0.01$). Similarly, Muslim secondary school contributed 1.34 standard deviation more to ethnic school segregation than non-religious secondary schools in the same pseudo-neighbourhood ($p < 0.01$).[173]

Catholic schools made the largest average contribution to ethnic school segregation of all the various types of Christian schools. On average, Catholic primary schools contributed 0.944 standard deviations more to ethnic school segregation than would be expected of

[173] The contribution of other non-Christian faith schools was broadly similar to that of Muslim schools ($\theta = 1.504$, $p < 0.001$). However, this is based on a very small sample of schools ($n = 17$) and so is not reported in the main body of the text.

non-religious primary schools in the same pseudo-neighbourhood ($p < 0.01$), while Catholic secondary schools contributed 0.425 standard deviations more to ethnic school segregation than would be expected of non-religious secondary schools in the same pseudo-neighbourhood ($p < 0.01$).

In contrast, Church of England primary schools contributed 0.283 standard deviations more, on average, than non-religious primary schools in the same pseudo-neighbourhood ($p<0.01$) and Model 8_s found no statistically significant difference between the contribution of Church of England secondary schools and non-religious schools in the same pseudo-neighbourhood ($p = 0.065$). These results provide sufficient evidence to reject the null hypothesis of question 5c and 5e and conclude that certain religious denominations contribute more to ethnic school segregation than others, both within the primary and secondary school sectors.

However, it is worth highlighting that the religious denomination of a school fails to explain comprehensively why schools contribute more to ethnic school segregation than others, with a substantial proportion of the variation in H_j was left unexplained in both sets of models (Model 8_p: $R^2 = 0.118$, Model 8_s: $R^2 = 0.057$). This partly reflects the fact that schools with the religious denominations that made the largest average contribution to ethnic school segregation were fewest in number (Jewish, Muslim, Other non-Christian schools) while those which were most common only contributed slightly more to ethnic school segregation than non-religious schools in the same pseudo-neighbourhood (Church of England). In fact, only Catholic schools contributed substantially more to ethnic school

segregation than non-religious schools and were also common in ethnically diverse pseudo-neighbourhoods in England.[174]

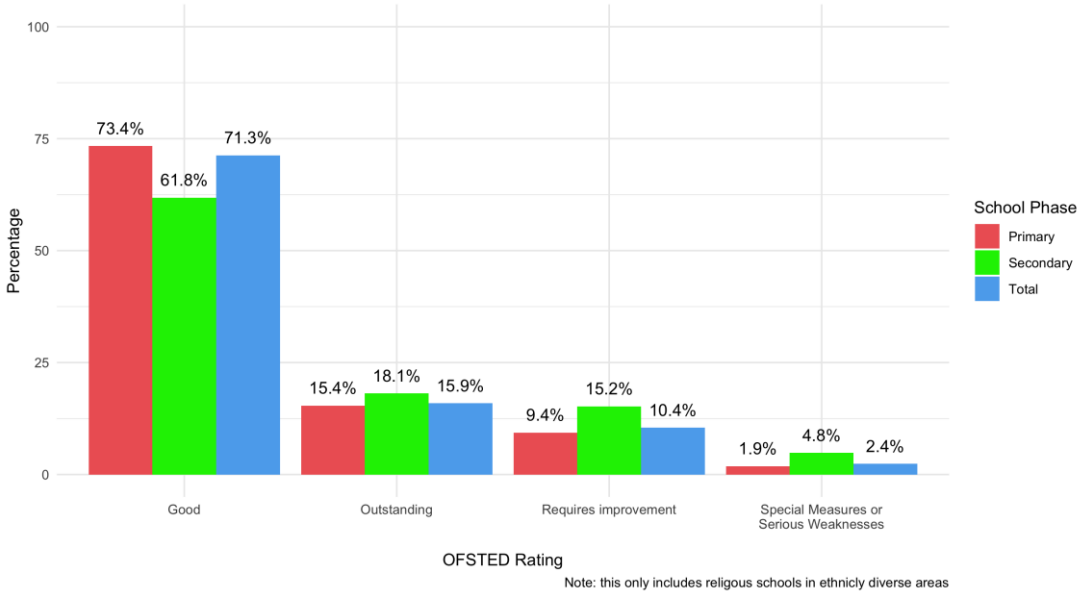
6.2.3 The association between a school's OFSTED rating and its contribution towards ethnic school segregation

The final section of this chapter investigates whether the contributions of schools towards ethnic school segregation varies between schools that received different OFSTED ratings. As summarized in the figure below, most schools in the sample received a "Good" OFSTED rating (71%), with 16% of schools being awarded an Outstanding rating, 10% designated as requiring improvement, and 2% being placed under "Special Measures" or deemed to have "Serious Weaknesses".[175]

[174] Catholic schools made up 11% of the sample.

[175] For detailed of how these differ between primary and secondary schools see the Figure **34**.

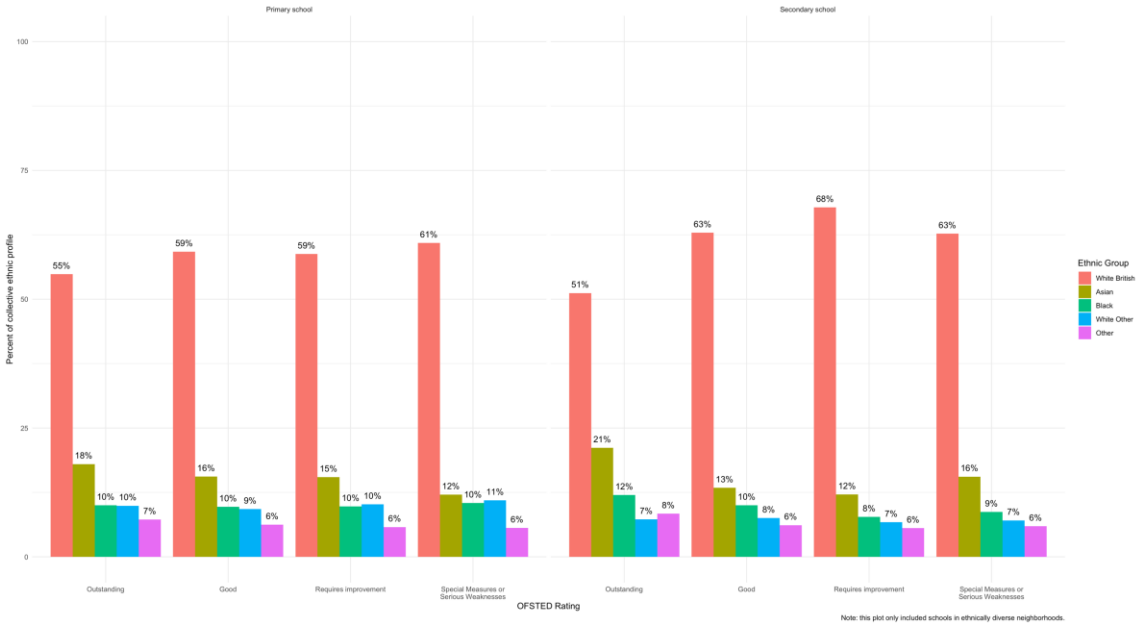
Figure 34: Breakdown of the OFSTED ratings recieved by schools in the sample (2019).



Interestingly, there are very small differences in the collective ethnic profiles of schools grouped by their OFSTED rating, except with respect to the proportion of White British students in secondary schools. As Figure 35 below shows, "Outstanding" secondary schools educate far fewer White British students (51%) than "Good" schools (63%) or those "Requiring Improvements" or in "Special Measures" (68%). In contrast, Asian students made up a larger proportion of students in "Outstanding" secondary schools (21%) than they did in secondary schools with lower OFSTED ratings (13%-12%).[176]

[176] This provides an interesting insight into how evenly distributed students from different ethnic groups are across schools of different "qualities".

Figure 35: The collective ethnic profile of schools grouped by OFSTED rating (2019).



The distribution of H_j values between schools that received different OFSTED ratings were remarkably similar for both primary and secondary schools, as shown in the density plot below. All distributions were positively skewed, with a similar modal point (-0.5), which provides strong evidence that there is no association between a school's OFSTED rating and their level of ethnic school segregation.

Figure 36: Distribution of Hj values for schools grouped by their OFSTED rating (2019)

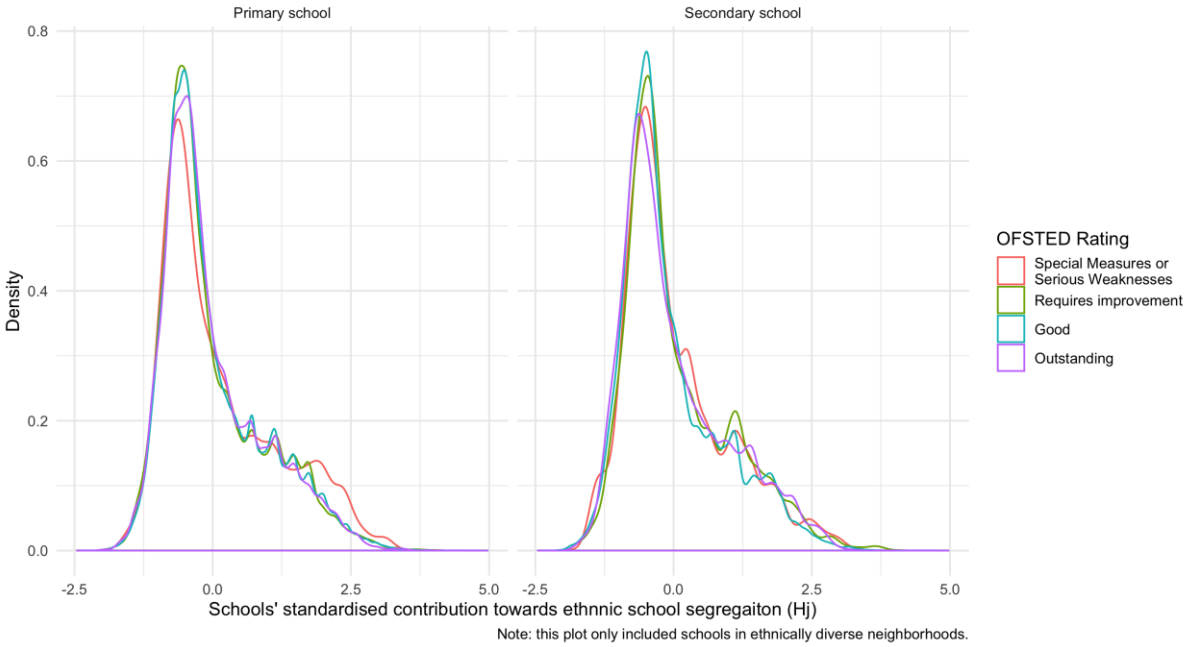


Figure 37: Regression Output to compare the contributions towards ethnic school segregation (H_j) made by school based on their OFSTED ratings

	Outcome variable:	
	Standardised contribution towards ethnic school segregation (H_j)	
	Model 9_p	Model 9_s
Outstanding	-0.029	-0.135*
	(0.096)	(0.069)
Good	-0.088	-0.117
	(0.089)	(0.066)
Requires improvement	0.026	-0.127
	(0.098)	(0.072)
(Intercept)	0.039	0.123*
	(0.086)	(0.066)
Observations	2386	10688
R Squared Value	0.003	0
Adjusted R Squared	0.002	0
Adjusted Standard Error	0.918	0.923
F Statistic	2.286	1.735
Note:	*p<0.1; **p<0.05; ***p<0.01	

This interpretation is corroborated by the findings of Model 9_p and 9_s, which are listed in the table above. Neither model found a statistically significant association between a school's OFSTED rating and the contribution they made to their pseudo-neighbourhood's level of ethnic school segregation.[177] Additionally, the adjusted R^2 value of both models was either 0.003 or 0.000 to three significant figures which suggests that knowing a school's OFSTED rating provides little to no insight into the likely contribution it makes towards ethnic school segregation (H_j). Therefore, this thesis fails to reject the null hypothesis of question 5f and 5g and concludes that the study found no evidence to suggest an association between a school's OFSTED rating and their relative contribution towards ethnic school segregation (H_j) for either primary or secondary schools.

Summary of Section 6.2

This section investigated whether the contributions made by schools towards their pseudo-neighborhood's level of ethnic school segregation varied depending on whether it used an academically selective admissions policy, had a religious affiliation or what the school's most recent OFSTED rating was. It found no evidence to suggest that grammar schools contributed more to ethnic school segregation than non-selective secondary schools, or that there was an association between a school's OFSTED rating and its contribution towards ethnic school segregation. The thesis therefore failed to reject the null hypotheses of question 5a, 5f and 5g. In contrast, both primary and secondary schools with religious

[177] P values ranged from 0.085 to 0.715 for the three predictor variables. For more information, please consult Table 39 in the Appendix.

affiliations were found to contribute more to ethnic school segregation, on average, than non-religious schools in the same neighbourhood ($p<0.01$). Additionally, Jewish, Muslim, and Catholic schools were found to contribute substantially more to ethnic school segregation than Church of England schools. The thesis therefore found sufficient evidence to reject the null hypotheses of question 5b - 5f.

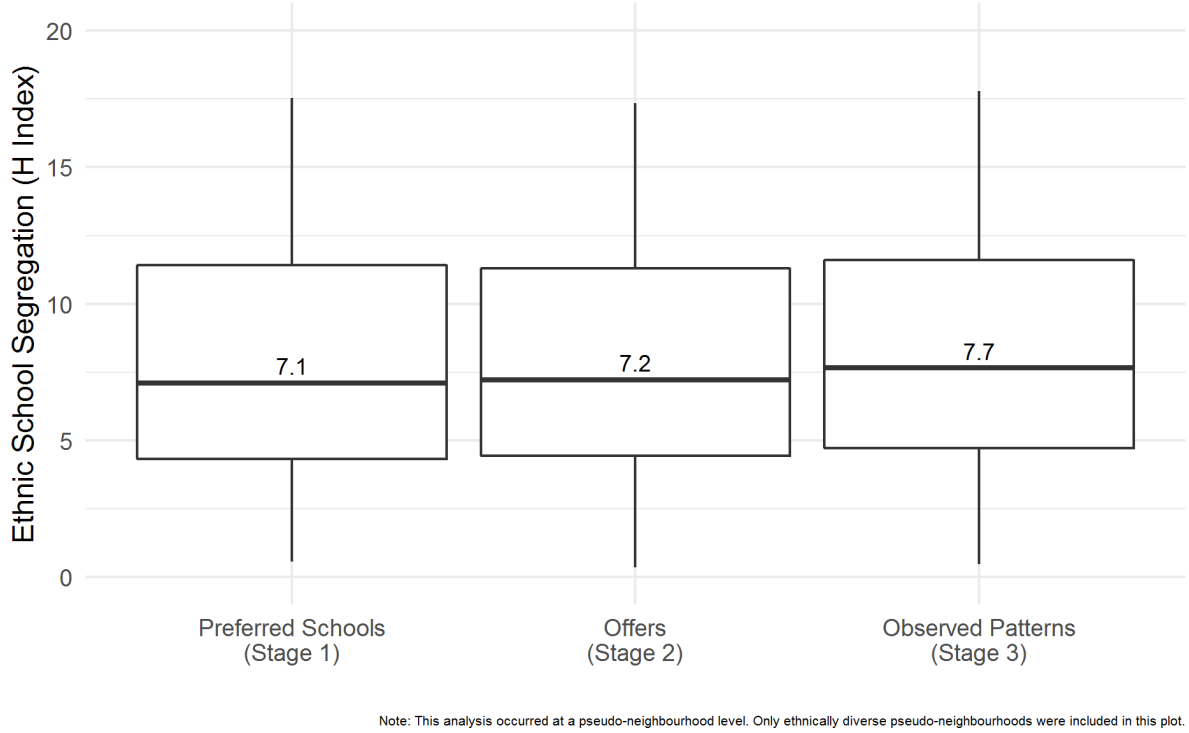
7 Investigating how the national school admission system links to patterns of ethnic school segregation

Chapter 7, which is the final results chapter, uses student level data to explore how the national school admissions system links to patterns of ethnic school segregation in England. It begins by describing how patterns of ethnic school segregation shifted at each stage of the 2016/2017 admissions cycle for year 6 students to address question 6a of the research questions. Section 7.2 then uses a range of simulations to predict what patterns of ethnic school segregation might have occurred if year 6 students had been allocated to secondary schools using seven different approaches in 2016/2017. These results address 6b- 6g of the research questions.

7.1 The shifting patterns of ethnic school segregation at each stage of the national school admissions process

As shown in the box plot below, the level of ethnic school segregation for the year six cohort rose at each stage of England's 2016/2017 admissions process. If all students ended up attending the school they listed as their 1st choice on their application form, in the first stage of the admission process, the median H value for pseudo-neighbourhoods would have been 7.1, with an interquartile range between 4.4 and 11.3. While the costs of providing this "pure school choice" would be sufficiently high as to make it largely impossible, it nonetheless provides a useful benchmark against which to compare subsequent results.

Figure 38: Boxplot to show how the distribution of *H* values shift at each stage of the national school admissions cycle for secondary school students (2016/2017).



If all students ended up attending the school that they were offered in the Spring of 2017, in stage two of the application process, the median H pseudo-neighbourhood value would have been slightly higher than in stage one (2%). The median *H* value for year seven students would have been 7.2, with an interquartile range of 4.5 to 11.2 for pseudo-neighbourhoods in England. This small change suggests that the collective impact of the schools' oversubscription criterion had a minimal impact on ethnic school segregation, above and beyond what was explained by the expressed preference of year six parents.

Finally, the reallocation of students from the school they were initially offered, in stage two of the application cycle, to the school they ended up attending, in stage three, led to a

larger increase in the average level of ethnic school segregation in England. Specifically, the median H value in 2017/2018 for year seven students within pseudo-neighbourhoods was 7.7, which was 6.5% higher than what would have been observed if these students had accepted the school they were offered in the Spring of 2017. This suggests that the appeals system, in combination with any other channel by which students did not attend the school they were offered, led, on average, to more ethnically segregated local school systems across England.[178]

Two sensitivity checks suggest these findings reflect trends occurring throughout England. Firstly, as shown in

[178] As this study removed any students who were not attending state funded schools in the 2017/2018 school year, this change does not reflect the departure of students from the state funded primary school system to the privately funded secondary school system.

Figure **39**, the median level of ethnic school segregation for pseudo-neighbourhoods in eight out of nine regions of England rose between each stage of the application process. This suggests that the substantive interpretation of these findings reflects dynamics occurring in pseudo-neighbourhoods throughout the country. Additionally, the median level of ethnic school segregation was also found to rise throughout the application cycle for pseudo-neighbourhoods located in a range of urbanicities, as shown in Figure 40. Similarly, this indicates that these results were not driven by one "type" of pseudo-neighbourhood. [179]

[179] There do not seem to be any clear geographic patterns to explain these changes, as shown by Figure 84 and Figure 85 in Section 11.26.

Figure 39: A bar chart to show how the median H values for pseudo-neighbourhoods in each region of England shifts between the three stage of the national school admission cycle (2016/2017)

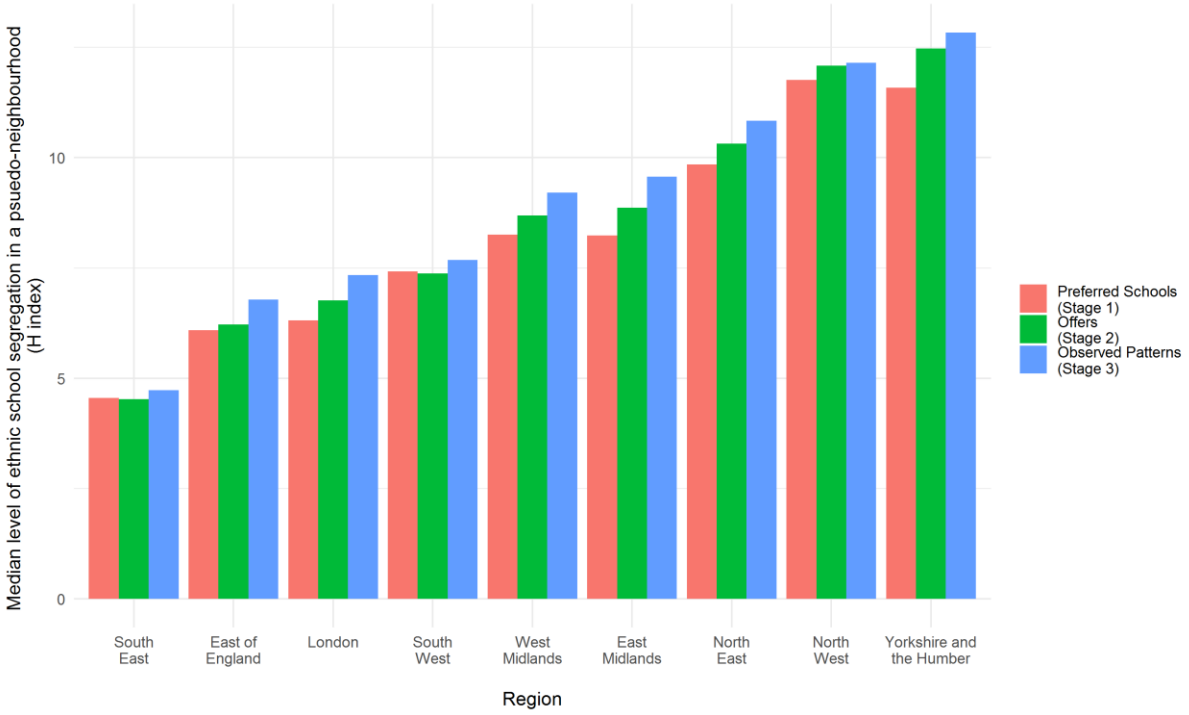
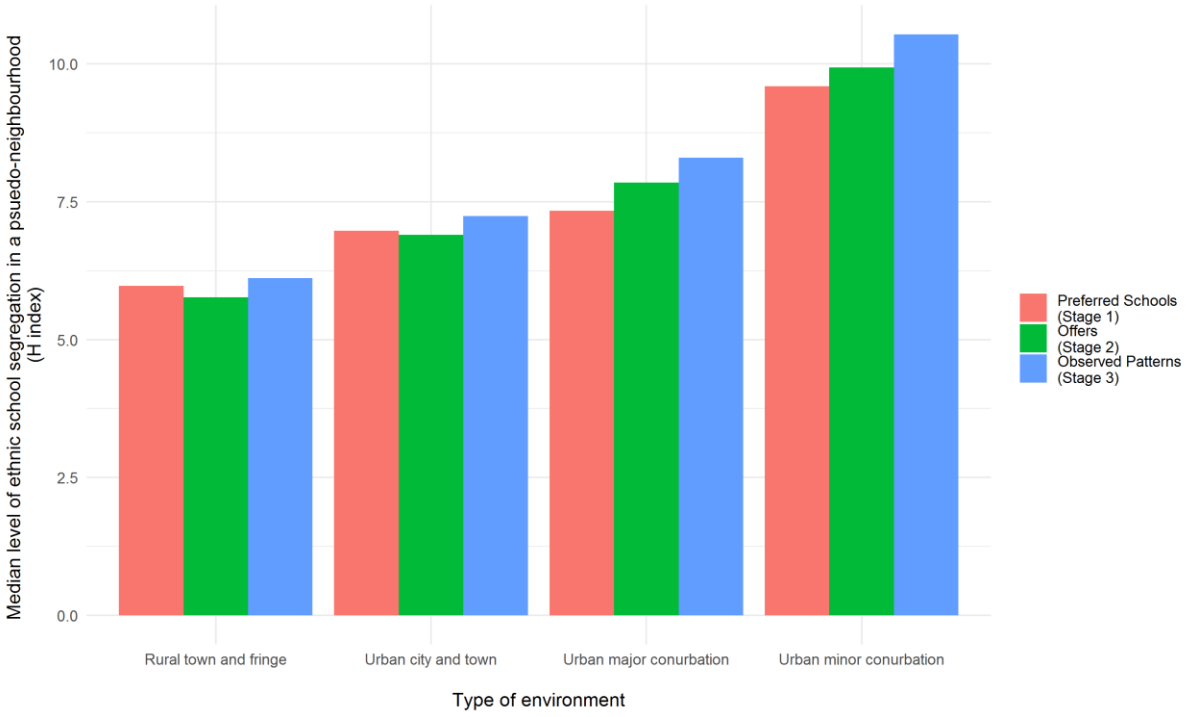


Figure 40: A bar chart to show how the median *H* values for pseudo-neighbourhoods located in different physical environments shift between the three stage of the national school admission cycle (2016/2017)



Summary of Section 7.1

Section 7.1 has shown that the likely levels of ethnic school segregation rose between each stage of the application process, with *H* values rising particularly fast between stages 2 and 3 of the school admission cycle. Specifically, the observed patterns of ethnic school segregation within pseudo-neighbourhoods for year 7 students in 2017/2018 was 6.5% higher than would have been observed if all students attended the school they had initially been offered in the Spring of 2017. In addition, the results suggests that these findings were

consistent for pseudo-neighbourhoods located throughout the country and for those operating in different physical environments. This answers research question 6a.

7.2 Exploring patterns of ethnic school segregation within seven simulated school admissions systems.

The following section explores the patterns of ethnic school segregation that might have occurred if students were allocated to schools under seven different school admissions systems.[180] It begins by comparing the numeric distribution of *H* values that were produced by each simulation, before examining whether these findings were consistent for pseudo-neighbourhood's located in different regions of England or those that operated in different physical environments.

Table 29: Table to describe the distribution of H values for each simulation

Simulation Number	Simulation Description	School Choice	Min	Lower Quartile	Median	Mean	Upper Quartile	Max
1	School Choice with home – school distance for oversubscribed schools	Yes	0.51	4.22	6.95	8.64	10.80	16.63
2	School choice with random allocation for oversubscribed schools	Yes	0.54	4.14	6.70	8.52	10.58	16.20
3	School Choice with academic banding followed by a distance-based allocation for oversubscribed schools	Yes	0.36	3.69	5.98	7.80	9.62	14.69

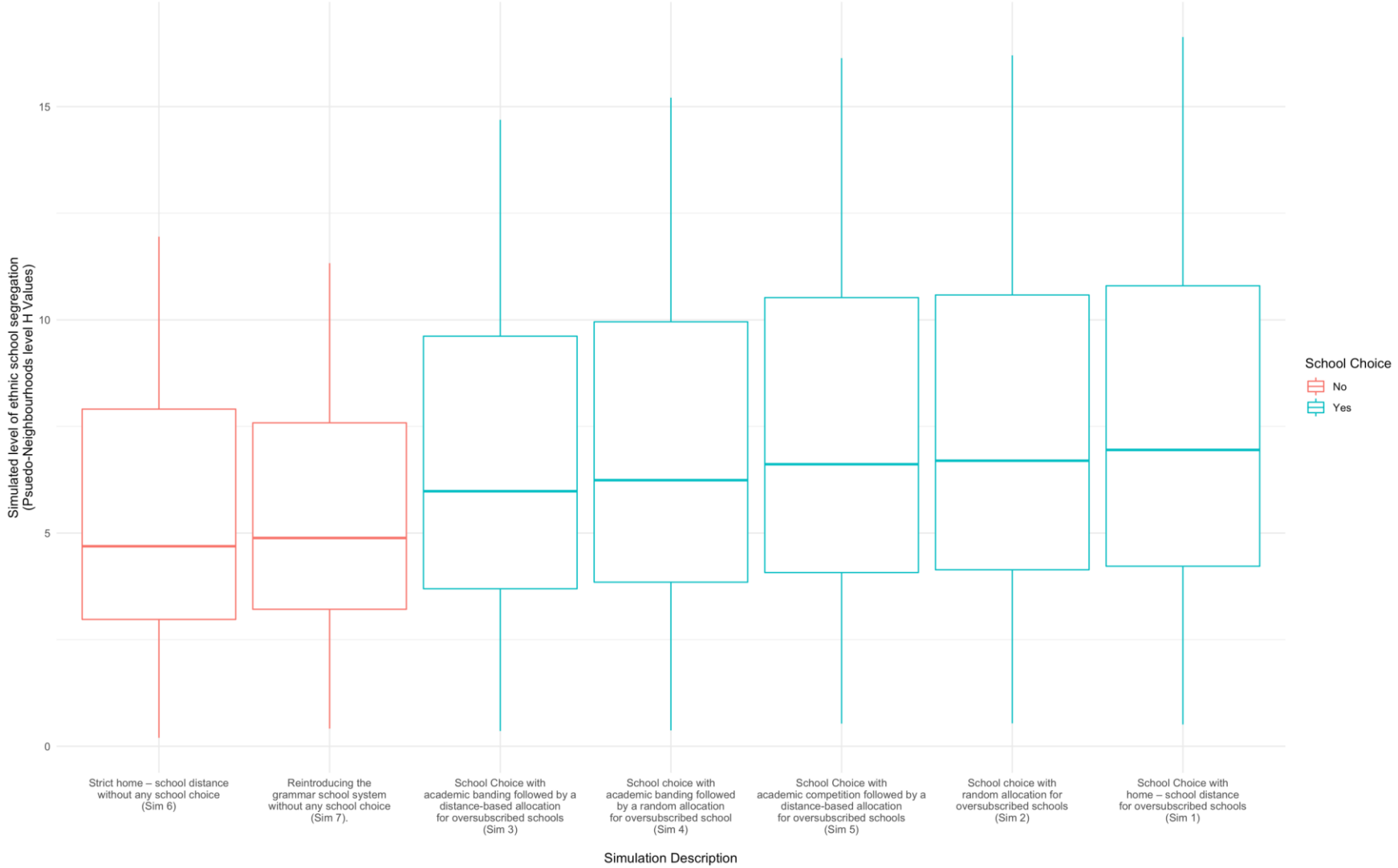
[180] The procedures used in each simulation are described in detail in Section 4.9.3 within the Methodology chapter. Data on the performance of these simulations is provided in Section 11.25 of the Appendix, which outlines the number and percentage of students who were assigned to their 1st – 6th preference school in each simulation.

4	School choice with academic banding followed by a random allocation for oversubscribed school	Yes	0.37	3.85	6.24	8.06	9.95	15.2
5	School Choice with academic competition followed by a distance-based allocation for oversubscribed schools	Yes	0.53	4.07	6.61	8.41	10.52	16.14
6	Strict home – school distance without any school choice	No	0.20	2.98	4.69	6.20	7.91	11.95
7	Reintroducing the grammar school system without any school choice.	No	0.42	3.21	4.89	6.17	7.59	11.33

The median level of ethnic school segregation was found to be lower for pseudo-neighbourhoods in all seven simulated school admissions systems than was observed in real life. The first simulation, which used home-school distance to allocate places at oversubscribed schools, had levels of ethnic school segregation that were most similar to those observed in real life. In this simulation pseudo-neighbourhoods exhibited a median *H* value of 6.9, with an inter quartile range of 4.2 - 8.6. These results suggest that if places at oversubscribed schools were exclusively allocated to students based on their home school distance, the average level of ethnic school segregation would have been likely to be around 10% lower than the observed median H value for year seven students in 2017/2018.[181] Given that most schools in England consider home-school distance, in some form or capacity, this 10% gap is likely to reflect the net impact of policies not considered by this simulation, such as sibling policies or requirements for schools to prioritise the applications of students with Special Educational Needs or those in care. This finding answers question 6b.

[181] This answers research question 6b.

Figure 41: Boxplot to describe the distribution of H values for each simulation



The simulation with the second highest median *H* value for pseudo-neighbourhoods allocated places at oversubscribed schools to applicants at random. In this simulation, pseudo-neighbourhoods exhibited a median *H* value of 6.7, with an inter quartile range of 4.1 to 10.6, for year 7 students in 2017/2018. While the median H value was still 13% lower than the average level of ethnic school segregation observed in practice, it nonetheless remained higher than most of the other simulations explored by this study. These results answer 6c of the research questions.

Surprisingly, the "academic competition" simulation, where places at oversubscribed schools were allocated to whichever students had higher KS2 results, exhibited lower levels of ethnic school segregation than the "random allocation" simulation. The median *H* value was found to be 6.6, which was 14% lower than was observed in practice, with an inter quartile range of 4.1 - 10.5. This suggest that, had students had been allowed to compete for places at oversubscribed schools using their KS2 test results, ethnic groups would have been more evenly distributed amongst local schools.[182] This result answers research question 6e.

The two simulations that used academic banding to allocate students to schools produced patterns of school attendance with the lowest median levels of ethnic school segregation out of the simulations that operated within the current school choice framework. In both simulations, the median level of ethnic school segregation was roughly 21% lower, on

[182] This could reflect the fact that allowing students to academically compete for places at oversubscribed schools would allow students to attend schools further away from their home, which could lead to greater ethnic mixing in areas with high levels of ethnic residential segregation. This theory would however need further research to validate.

average, then was observed for year seven students in 2017/2018. The simulation that used academic banding in conjunction with a random allocation produced an *H* value of 6.2, while the simulation that combined academic banding with a distance-based criterion had a median *H* value of 5.9. Both results suggest that efforts to encourage a more even distribution of academic abilities across secondary schools in England would be likely to also encourage a more equal distribution of ethnic groups. These results answer 6d of the research questions.

However, the two simulations that created a radically different national school admissions system produced the lowest average levels of ethnic school segregation. In both the grammar school system simulation and the strict home-school distance simulation the median *H* was around 38% lower than those observed in practice. The grammar school system simulation produced a predicted median H value of 4.9 with an inter quartile range of 3.2 - 6.2, while the strict home school distance simulation predicted a median H value of 4.7 with an inter quartile range of 3.0 - 6.2. These results suggest that if year six students had been allocated to schools based solely on their home address, or on their address in combination with a split based on their academic abilities, a more even distribution of ethnic groups would have been observed than any of the simulated approaches that worked within the current school choice framework. Additionally, the *H* values calculated using these two simulations had the smallest interquartile range of all seven simulations, suggesting that adopting these approaches would be likely to produce a more uniform level of ethnic school segregation between different pseudo-neighbourhoods in England. These results answer question 6f and 6g of the research questions.

The trends identified in this subsection were also found to be fairly consistent across pseudo-neighbourhoods located in different regions and for pseudo-neighbourhoods operating in different physical environments. For example, the median level of ethnic school segregation for pseudo-neighbourhoods in all nine regions of England was higher in reality than in all seven simulations as shown in Figure 42. Similarly, the relative size of the median *H* values for different simulations was similar for pseudo-neighbourhoods located within the same region. Additionally, these findings were also consistent for pseudo-neighbourhoods operating in different physical environments, as demonstrated in Figure 43.

Figure 42: A bar chart to describe the median H value for pseudo-neighbourhoods located in each region of England, under each of the seven simulations conducted by this study.

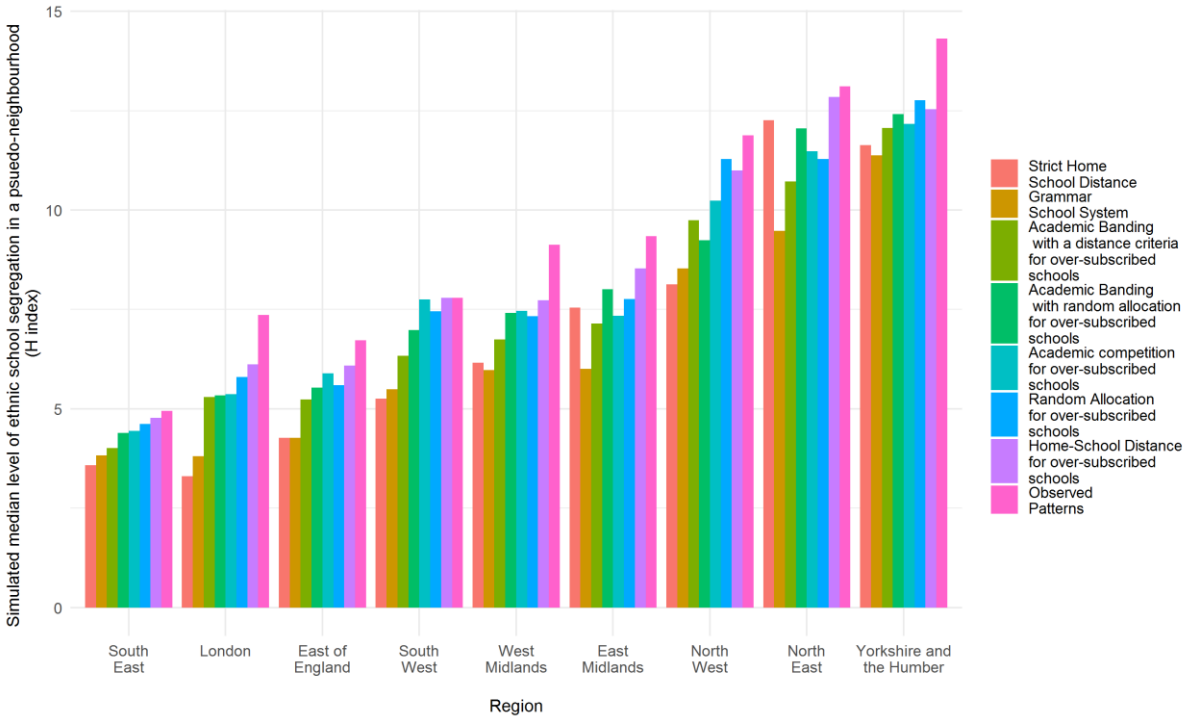
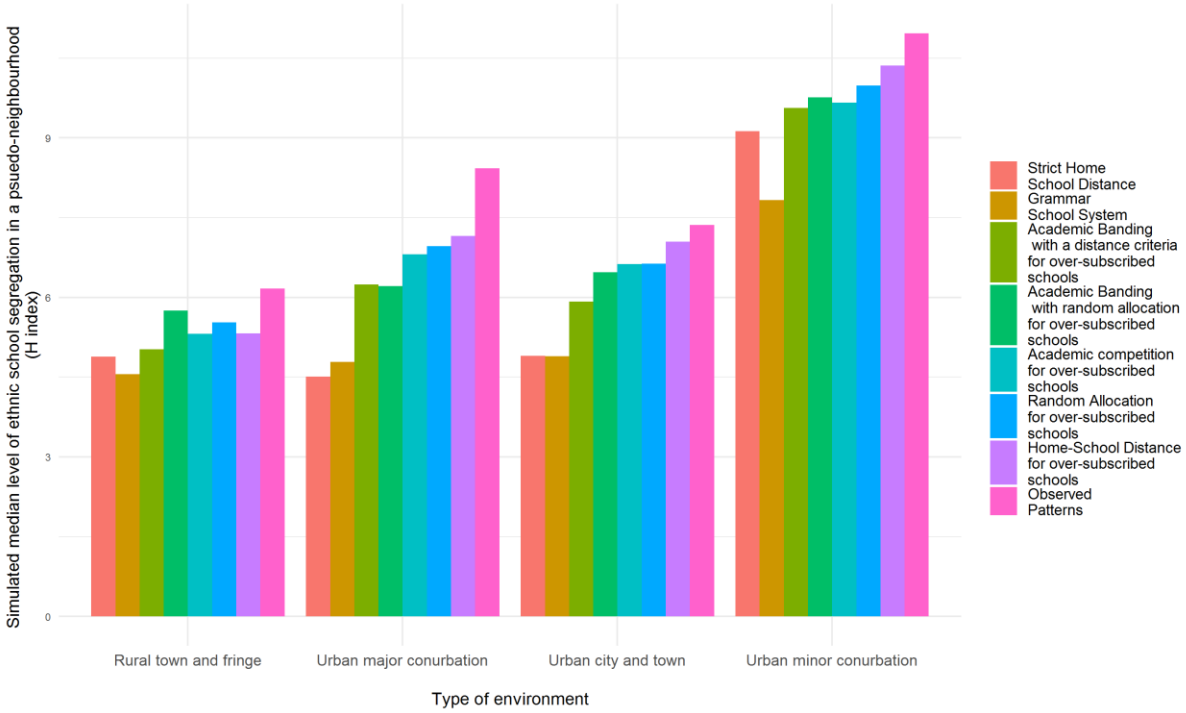


Figure 43: A bar chart to describe the median H value for pseudo-neighbourhoods located in different physical environments, under each of the seven simulations conducted by this study



Summary of 7.2

Section 7.2 has used seven simulations to explore the patterns of ethnic school segregation that might occur under different national school admissions systems. The results found that for all seven simulations ethnic school segregation would be lower than was observed for the year group in question. Simulations that used distance, random allocation, and academic competition to allocate places at oversubscribed schools produced patterns of school attendance where median H values were between 10% - 14% lower than those observed in practice. For the two simulations that used academic banding, ethnic school

segregation was lower still, with median H values around 21% lower than observed in practice. However, the two simulations that allocated students to schools based solely on their home-school distance, or their home-school distance in combination with their academic abilities, exhibited the lowest levels of ethnic school segregation. These simulations had median *H* values that were 38% lower than observed in real life. Finally, these results were found to be broadly consistent for pseudo-neighbourhoods in different regions of the country or for those operating in different physical environments, suggesting the findings are robust. Collectively, these results answer 6a – 6g of the research questions.

8 Discussions

The following chapter pulls together the results of this thesis to summarise the study's major findings and locate them within the wider academic literature and relevant public policy debates. In general, the findings are discussed in the same order which they appear in the three research chapters.

Section 8.1 begins by highlighting how the ethnic profile of students attending state funded schools in England changed between 2006 – 2019 and draws on different results to stress how widespread the ethnic diversification of this population has been. This provides key contextual information required to interpret the study's major results and answers research question 1a and 1b.[183]

Section 8.2 then links these demographic shifts to the level of ethnic isolation within English state schools. It explains how falling levels of ethnic isolation suggests a growing space for inter-ethnic contact within England's school system, while also highlighting that most White British students still attend schools where few of their peers come from different ethnic backgrounds. Based on these findings, it argues that to create state schools where young people build friendships that cross ethnic lines, reducing the level of ethnic isolation experienced by White British students must be the focus for the UK government, as the vast

[183] This asked, "What is the ethnic profile of England's school age population and how has it changed between 2006 – 2019 across various geographic scales?"

majority of students from ethnic minority background already attend ethnically diverse schools. These discussions engage with research question 2a – 2c.[184]

Section 8.3 then turns its attention to discussing ethnic school segregation, which was the study's main outcome variable. It argues that a 25% decline in the median level of ethnic school segregation for England's pseudo-neighbourhoods between 2006 – 2019 suggest a substantial reduction in the level of ethnic "separateness" within local educational markets in the state school system. This is an encouraging result that should reassure politicians, policy makers and the public that as England becomes more ethnically diverse, the ethnic composition of local schools increasingly resemble one another. At the same time, the study highlights that some local areas continue to experience comparatively high levels of ethnic school segregation. The regions of the North West and Yorkshire and the Humber, for example, contain more pseudo-neighbourhood's with exceptionally high levels of segregation ($H > 20$) than the rest of the country combined. Given the history of inter-ethnic tension in many of these areas, research should explore how support could be provided to try and encourage less ethnically divided school systems. This section addresses research question 3a and 3b.

Section 8.4 then discusses what types of primary and secondary school pseudo-neighbourhoods tend to experience higher levels of ethnic school segregation in England, as well as the kinds of schools that contribute most toward this "separateness". Taking one association at a time, the section reminds readers of the empirical basis on which each

[184] These asked "What is the current level of ethnic isolation within England's school system and how has it changed between 2006 – 2019?"

predictor was included in the study, the associations that were found by the thesis and the implications that these results had on their respective research question. It then discusses how the result(s) has changed current academic understanding, the relevance this has for policy makers and the possible routes that future research could take to deepen our understanding further.

Section 8.5 then discusses the findings from the first analysis into how the national school admissions system shapes patterns of ethnic school segregation in England. It begins by explaining that the average level of ethnic school segregation rose between each of the three stages of the application process, while also stressing the importance of parental preferences in determining school attendance patterns. It then examines findings from each of the seven simulations produced by this study, all of which allocated students to schools in a manner that resulted in lower average levels of ethnic school segregation. Based on these findings, the section argues that there seems to be a tradeoff between allowing parents to have greater choice over the secondary school their child attends and the ability of the government to reduce ethnic school segregation within the school system. At the same time, the study acknowledges the limitations of this methodological approach which means that such conclusions should be interpreted cautiously.

After the substantive results have been discussed, the chapter moves to the methodological contributions that this thesis has made to the academic literature. Section 8.6 reiterates the benefits of pseudo-neighbourhoods, before exploring some of the ways in which this technique could be refined and extended. It also argues that the use of a multi-group measure of ethnic school segregation has provided an important improvement in British

literature, and it explores ways in which the index may be applied to other fields. In both cases, the limitations of these approaches are also discussed.

The chapter ends with a discussion on the wider limitations of this study and two potential avenues for future research. It argues that in order to measure ethnic school segregation rigorously, this study has needed to adopt a tightly bound definition of "ethnic school segregation" that pays little attention to either ethnic residential segregation or socio-economic school segregation, despite their likely importance in determining patterns. It therefore encourages readers to explore research in these related fields to gain a more holistic understanding of the trends described in this thesis. It then suggests that public transport networks are likely to play an important role in creating patterns of ethnic school segregation and therefore warrant greater attention in future research. Additionally, it suggests that research on ethnic school segregation could benefit from a broader base of methodological approaches. In particular, it highlights the use of field experiments by Nobel prize winning economists, Banerjee and Duflo, as an interesting template for future research that could help to unpick the underlying dynamics that cause ethnic school segregation in England.

8.1 England's changing ethnic profile

This study has provided a detailed description of the ethnic composition of England's school age population at various geographic scales. These maps, diagrams and statistics provide important contextual knowledge to interpret the study's later findings and are likely to be useful for academics interested in the role that ethnicity plays within the English school system.

The study found that the ethnic profile of England's school aged population in 2019 varied substantially between different parts of the country. At a regional level, Greater London educated nearly equal numbers of students from White British, Asian and Black ethnic groups, while 87.8% of students in the North East, which was the least ethnically diverse region, came from a White British background. The variation grew further still when investigating patterns at a pseudo-neighbourhood level. While some pseudo-neighbourhoods were ethnically homogenous, and only educated White British students, others were "super diverse" with near equal numbers of students from all five ethnic groups. Generally speaking, ethnic diversity was higher in more densely populated, urban pseudo-neighbourhoods; with the cities of London, Birmingham, Manchester, Leeds and Bradford educating particularly large numbers of ethnic minority students. In contrast, rural pseudo-neighbourhoods tended to educate a far greater proportion of White British students. While these findings are unlikely to surprise many informed readers, the diagrams and maps provide welcome detail to flesh out these important residential patterns.[185]

[185] This set of findings answer question 1a of the research questions.

The study then outlined how England's ethnic profile had changed over the last 13 years at different geographic scales, offering an important longitudinal perspective. Set within the context of a growing ethnic diversity at a national level, the section underscored just how widespread changes had been, both in relation to England's nine regions as well as the 3292 pseudo-neighbourhoods used by the study. It found, for example, that *all* nine regions of England educated more students from *all* four ethnic minority backgrounds in 2019 than they did in 2006. Similarly, it revealed that 92% of primary school pseudo-neighbourhoods and 95% of secondary school pseudo-neighbourhoods in England educated a smaller proportion of White British students in 2019 than they did in 2006, and that 99% had experienced a growth in the proportion of students from *at least* one ethnic minority group within the same time period. Even pseudo-neighbourhoods that had previously been ethnically homogeneous were found to be educating students from an increasingly broad range of ethnic backgrounds, which supports the idea that "new spaces of diversity are emerging" throughout the country (Catney, 2016b, p. 750). While increased levels of immigration are likely to be a key driver for many of these changes, a range of other factors are likely to feed into the increasing diversity of England's school age population. These include the higher fertility rate of women born outside the UK (Office for National Statistics, 2018), the older age profile of White British residents (Shankley, Hannemann and Simpson, 2020) and the steady increase in the number of mixed ethnicity families (Centre on Dynamics of Diversity, 2012, p. 1).

At the same time, the rate of change was unequal across different parts of England. While most pseudo-neighbourhoods became more ethnically diverse, those in urban environments did so at a faster rate than pseudo-neighbourhoods in rural areas. This led to a growing proportion of "multi-ethnic" urban pseudo-neighbourhoods where no single ethnic group formed an absolute majority (>50%). Between 2006 – 2019 the proportion of these "multi-ethnic" pseudo-neighbourhoods more than doubled, from 6% to 14%, which represents a major shift within a short space of time. These findings, when read in combination with the various maps and diagrams presented in Section 5.1, point to a split in the ethnic profile of England's pseudo-neighbourhoods; between urban, ethnically diverse areas and rural pseudo-neighbourhoods where students remain predominantly White British. Of course, as the country continues to diversify, future studies could show that these positions may well converge.

Understanding the current profile of England's local areas, and how these have changed, provides important contextual information to support the interpretation of the study's major findings. At a foundational level, they highlight that England's school age population is in the midst of a fairly profound transformation and that these shifts are geographically diffuse. This suggest that patterns of school attendance, and the resulting levels of ethnic school segregation, are likely to be unstable and generally fluid, which aligns with the findings of previous research in this area (Harris and Johnston, 2020). Such change creates a considerable challenge as well as a unique opportunity for policy makers and academics. On one hand, stakeholders must work hard to keep up with the pace of change, both in terms of monitoring the shifts in ethnic school segregation, updating policies, and providing support for schools in ethnically diverse areas. At the same time, the fluidity of the current

situation presents a unique opportunity to create a healthier, more robust school system that brings young people from different ethnic backgrounds together. Seizing this moment is important. As demonstrated by the southern states of America, South Africa or indeed Northern Ireland, desegregating a school system after it becomes divided is far more difficult than avoiding such divisions in the first place. Governments and related stakeholders should therefore view this period of change, where people from different ethnic groups are moving into new areas, as an invaluable opportunity to build unified schools before any potential ethnic divisions arise or solidify.

These results also highlight the broad and multi-dimensional spectrum of ethnic diversity that local areas in England sit along which mean "ethnic school segregation" will look and feel differently across the country. Even within the sample of 2227 "ethnically diverse" pseudo-neighbourhoods studied by this thesis, the ethnic profiles of areas varied considerably. This should caution readers against any crude comparisons between pseudo-neighbourhoods that have radically different ethnic profiles and discourage any assumption that national level findings are necessarily reflected within local area. The breath of ethnic diversity should also serve as a reminder of the importance of localized studies which better fit the contexts in which schools operate.

Lastly, these findings will also be useful to a broad range of academics and policy makers who are interested in exploring the role that ethnic identity plays in England's education system. While this body of literature is diffuse, the geographic patterns described in this thesis should be applicable to a wide range of topics which would benefit from understanding the intricate nature of England's ethnic diversification. The maps, statistics

and diagrams also demonstrate that policy decisions relating to ethnicity in schools are increasingly relevant to broad swaths of the country, rather than just diverse inner-city neighbourhoods. Examples of such policy issues include, but are not limited to, the teaching of colonial British history, teaching English as an additional language, tackling racial bullying, hiring and retention of BAME teachers or the role of religion in schools that draw their student body from diverse religious, and non-religious, backgrounds.

8.2 Ethnic Isolation in England's schools

The second set of results mapped patterns of ethnic isolation which helped to link the observed changes in the ethnic profile of England's school aged population to a measure of ethnic "separateness". Two major findings emerged from this short analysis. First, the median level of ethnic isolation in English schools fell by around 15% between 2006-2019, which suggests that the space for inter-ethnic contact within English schools grew during this period.[186] Specifically, students in 2019 attended schools where, on average, around 75% of their peers came from the same ethnic background, down from 90% in 2006. While this change is likely to be heavily influenced by shifts in the ethnic makeup of England's school age population, it nonetheless provides the best current estimate for the level of inter-ethnic contact occurring within English schools.[187] If greater contact, on average, improves inter-group relations, as contact theory suggests it should, then this suggests that the government is managing to meet its commitment towards making schools places where

[186] As the findings of this analysis were calculated at a student level and were similar between students attending primary and secondary schools, the results are generally described collectively, without distinguishing between each phase.
[187] As mentioned in Section 4.5.3, measures of ethnic isolation are only an *estimation* of inter-ethnic contact.

students “mix and form lasting relationships with others from different ethnic… groups” (HM Government, 2018a, p. 26).

At the same time, the analysis also highlighted the stark differences in ethnic isolation experienced by students from different ethnic backgrounds. While White British students in 2019 attended schools where, on average, more than 8 out of 10 students were also White British, students from all other ethnic groups normally attended schools where only a fraction of their peers came from the same ethnic background. While these differences were driven, in large part, by the comparative size of each ethnic group and different residential patterns, they nevertheless offer an important insight into how the experience of attending schools in England differs for young people from different ethnic backgrounds.

Put simply, the results suggest that there are two common experiences for students in English schools. Either they attend a school where the vast majority of students are White British, or they attend a school were students come from a variety of different ethnic groups. By definition, White British students make up the bulk of students in the former group, while ethnic minority students are predominantly in the latter category. Crucially, the vast majority of ethnic minority students in 2019 attended schools where a substantial proportion of their peers were White British. This means that while most ethnic minority students had the opportunity to interact and form relationships with White British peers, the majority of White British students are unlikely to attend schools where they had the chance to meet young people from other ethnic backgrounds.

If society is truly concerned about students lacking the chance for inter-ethnic contact, then the UK government should focus its efforts on helping mono-ethnic White British schools attract more diverse intakes, for the sake of White British students. While this point has been made by previous academics (Cantle, 2013, p. 13), the results detailed in this thesis provide strong evidence to support this position. Additionally, it is important for academics and policy makers to broaden discussions on inter-ethnic contact to also include contact between different minority ethnic groups. Indeed, in many parts of the country encouraging more contact between smaller ethnic groups could play a key role in reducing ethnic tensions between local communities.

8.3 Ethnic school segregation in England

As ethnic isolation is highly influenced by the comparative size of ethnic groups, this study predominantly focused on exploring ethnic school segregation, as measured by the H index. As detailed in Section 5.3, it found that ethnic school segregation within England's primary and secondary school pseudo-neighbourhoods fell substantially between 2006 - 2019, reducing by around 25% from a median *H* value of 6.8 to 5.1. This meant that between 2006 - 2019 the ethnic profile of schools operating in the same local area tended to become increasingly similar, as students from different ethnic groups became more evenly spread across local schools. Alternatively, between 2006 - 2019 schools attracted intakes which increasingly reflected the ethnic diversity of their local area, intentionally or otherwise.

Reductions in ethnic school segregation were widespread. Between 2006 – 2019 pseudo-neighbourhoods located throughout the country experienced a fall in the average level of

ethnic school segregation, with the median *H* values for pseudo-neighbourhoods in all nine of England's regions decreasing during this time period. Pseudo-neighbourhoods in the north of England tended to experience larger reductions in ethnic school segregation than those in the south, which helped to reduce the gap between the average levels of ethnic school segregation in the north and south of England.

Unfortunately, it is hard to know definitively why these changes occurred at a national level. As *all* measures of segregation are influenced by the underlying proportions of different ethnic groups, albeit to a lesser extent than measures of ethnic isolation, it is possible that these trends partly reflect the increased ethnic diversity of England's school age population.[188] While the results of the three sensitivity checks conducted by this study reduce the likelihood of this, they do not entirely rule it out as a possibility. Future research is therefore required to understand these changes in more detail. It would be useful, for example, to investigate which variables are predictive of a pseudo-neighbourhood's change in ethnic school segregation over time. Given that previous research has highlighted that "the causes …[of ethnic school] segregation for any indicator will be different from the causes of changes in that level over time" (Gorard, 2016b, p. 131), this analysis could deepen our current understanding.[189] Predictors could include shifts in the ethnic composition of a local area or changes in local house prices, crime rates, political sentiments or transport links.

[188] Unfortunately, compositional invariance remains endemic to all measures of evenness segregation. See 4.5.1.1 for more details.

[189] This could even provide a type of "early warning system" for policy makers to identify neighbourhoods that are likely to need additional support to avoid the formation of ethnic segregation in the coming years.

These results broadly align with the findings of previous research which reenforces a growing body of evidence that suggests the association between a student's ethnicity and the school they attend is weakening within England across various geographic scales of measurement. Gorard, Hordosy and See's article (2013), who measured ethnic segregation within England as a single geographic area using a binary segregation index, found that the G index fell by around a third within the secondary school system and a sixth in the primary school system between 1997-2012. Similarly, the analyses conducted by Harris and Johnston in 2020 led them to conclude that "the general trend [across England] has been towards desegregation and greater ethnic diversity within local authority areas and their schools" (2020, p117).

However, these findings should not be taken to mean that ethnic school segregation is no longer a problem in England, or that if left alone ethnic school segregation will inherently fall. As this thesis has highlighted, many pseudo-neighbourhoods continue to experience high levels of ethnic school segregation and unfortunately, they are often located close to one another. The study found, for example, that pseudo-neighbourhoods in the North of the country tended to have far higher levels of ethnic school segregation than elsewhere in England. It was notable, for example, that in 2019 the North West region and Yorkshire and the Humber contained more neighbourhood's with exceptionally high levels of segregation ($H > 20$) than the rest of the country combined. Given the history of inter-ethnic violence in many of these areas (Cantle, 2001), these findings should act as a reminder to government officials, local authorities and schools operating in these neighbourhoods that many of the factors that had been highlighted as causes for previous tensions have yet to be addressed. The research community also has an important role to play in this regard. More needs to be

done to correct a widely acknowledged "London bias" in localised studies of ethnic school segregation (Taylor, 2009, p. P555), which does not reflect the comparatively low level of ethnic school segregation within the capital.[190] Instead focus should be diverted to areas with comparatively high levels of ethnic school segregation, which hopefully this thesis can help to identify.

More broadly, readers should be reminded that these results are average figures calculated using 3227 pseudo-neighbourhoods and that the ethnic dynamics within each local school system are unlikely to neatly align with these national findings. As highlighted in Section 8.1, pseudo-neighbourhoods sit along a spectrum of ethnic diversity and each will have their own unique history and local subtleties, both within school communities but also in other key areas such as the housing market. This means that while this study has presented an account of what is *generally* happening across England's local educational markets, the local schools surrounding your home may look and feel differently to those described in this thesis. This is an important limitation of this type of large-scale analysis, which underscore why localised studies also have an important role to play in helping academics and policy makers understand ethnic school segregation within England.

[190] Examples of localised studies surrounding issues of ethnic school segregation in London include (Hollingworth and Williams, 2010; Vowden, 2012; Hamnett and Butler, 2013; Hamnett, Butler and Ramsden, 2013; Harris, 2013; Leckie and Goldstein, 2015).

8.4 The predictors of ethnic school segregation in England

The study also examined which types of pseudo-neighbourhoods tended to exhibit higher levels of evenness ethnic school segregation than would otherwise be expected. The findings are discussed in the following five subsections in relation to each predictor variable included in the study and their respective research question(s). These variables include a pseudo-neighbourhood's population density, its socio-economic profile, the use of academic selection, the provision of religious schooling and the average quality of schools, as measured by their OFSTED rating. The section also discusses whether the latter three variables had an association to ethnic school segregation on a school level, as measured by a school's relative contribution towards ethnic school segregation (H_j). Readers should be reminded that these results are based on observational data and so the associations that these models identify should not be assumed to be causal. Consequently, great care has been taken to stress that any hypotheses on why these associations exist would require future research, particularly given the methodological complexities that measuring ethnic school segregation presents.

8.4.1 Population Density (Research Question 4a and 4b)

The study found limited evidence to suggest there was a robust association between a pseudo-neighbourhood's population density and its *H* value. None of the primary school pseudo-neighbourhood models found a statistically significant association and only one of the secondary school pseudo-neighbourhood models identified a statistically significant relationship. This model, Model 2_s, predicted that a one standard deviation difference in the

population density of a secondary school pseudo-neighbourhood was associated with a level of ethnic school segregation that was 10.5% higher, on average, once the number of schools had been controlled for (p=0.047). While this statistically significant result provides sufficient evidence to rejects the null hypothesis of question 4b, readers should view this with caution. Including the variable only increased the adjusted R^2 value between Model 1_s and 2_s by 0.012 and none of the subsequent models found a statistically significant association. When viewed collectively, these findings suggest there is unlikely to be a substantial or robust association between a secondary school pseudo-neighbourhood's population density and its level of ethnic school segregation.

These results contrast with previous studies, which found population density had a comparatively strong association with ethnic school segregation (Gorard, 2016b) or related concepts (Johnston, Wilson and Burgess, 2004).[191] A number of factors could explain these divergent results. As outlined in Table 16 in Section 4.7, a pseudo-neighbourhood's population density was correlated with the number of schools in a primary (r = 0.16) and secondary school pseudo-neighbourhood (r = 0.13). By controlling for this variable, the study is likely to have assigned a sizable proportion of the variation in *H* to the model's control variable that would have otherwise been assigned to a pseudo-neighbourhood's population density. The divergent results are also likely to be influenced by the scale at which the analyses were conducted. For example, associations measured between Local Authorities could well differ from associations measured between pseudo-neighbourhoods. Alternatively, the findings could reflect a genuine change in the nature of ethnic school segregation in England between 2019 and 2011, the year in which the most recent study

[191] This study used the ratio of ethnic residential segregation to ethnic school segregation.

was conducted (Gorard, 2016b). As the ethnic profile of urban areas changed dramatically within this period, it is not inconceivable that associations may have shifted in recent years, though on balance it seems less likely.

Substantively, this result suggests that there is a limited association between ethnic school segregation and urbanicity, at least when measured within pseudo-neighbourhoods or 'local educational markets'. This both underscores the methodological importance of the scale at which patterns of school segregation are measured and suggests that while ethnic school segregation is likely to look and feel differently in urban, sub-urban and rural neighbourhoods, there is not currently any substantial differences in the average level of ethnic school segregation that pseudo-neighbourhoods in these different physical environments experience.

8.4.2 Socio-economic profile of pseudo-neighbourhoods (Research Question 4c – 4f)

The second association explored by the study was between a pseudo-neighbourhood's socio-economic profile and its level of ethnic school segregation. Due to findings from previous research (Burgess, Wilson and Lupton, 2005; Gorard, 2016b; Harris and Johnston, 2020), the study postulated that while a pseudo-neighbourhood's socio-economic profile was likely to have an association with ethnic school segregation, the findings would be varied and fairly complex. In fact, the opposite was true, and the results identified by this thesis were surprisingly consistent.

Across nearly all relevant models, both for primary and secondary schools, the findings constantly suggested that ethnic school segregation tended to be higher in pseudo-neighbourhoods with lower socio-economic status, after other control variables had been accounted for. Model 3_p, for example, predicted that a one standard deviation difference in the average house price of a primary school pseudo-neighbourhood was associated with an H value which was 15.5% higher, once other variables had been accounted for ($p<0.001$). For secondary school pseudo-neighbourhoods, the association was estimated to be even stronger. Model 3_s predicted that for every standard deviation increase in a secondary school pseudo-neighbourhood's average house price, the forecast level of ethnic school segregation fell by 20.4%, after controlling for other variables in the model ($p<0.001$). This association, which was the strongest identified in both models, led to the rejection of the null hypotheses of question 4c and 4d.

Similarly, there was strong evidence to suggest that pseudo-neighbourhoods that educated a greater proportion of students who were eligible for free school meals tended to experience higher levels of ethnic school segregation, once other control variables were accounted for. Model 3_p, for example, predicted that for every standard deviation increase in the rate of FSM, the predicted level of ethnic school segregation within primary pseudo-neighbourhoods increased by 7.0% ($p=0.017$), once other variables in the model had been controlled for. The association for secondary school pseudo-neighbourhoods was comparatively stronger, with Model 3_s predicting that a one standard deviation difference in the rate of FSM was associated with a secondary school *H* value that was 19.2% higher, on average ($p = 0.006$).

This set of results provides greater clarity on an important association and serves as a useful starting point for future empirical work.[192] Future research could explore, for example, whether ethnic school segregation is higher in neighbourhoods with a lower socio-economic status due to practical issues, such as a higher relative cost of travel, or sociological reasons, such as a heightened desire to cluster with one's own ethnic group due to greater economic insecurity, as suggested by competition theorists (Blalock, 1967). Alternatively, the relationship could work in reverse. If ethnic divisions within a neighbourhood encourage those with financial or social capital to move away, then ethnic segregation could lower the socio-economic profile of an area, as suggested by proponents of White Flight. Longitudinal analysis comparing house prices, along with some measure of socioeconomic status, and an ethnic school segregation index could provide an interesting exploration of this potential association. As stressed earlier, the findings identified by this thesis should be largely viewed as a starting point for more targeted future research.

8.4.3 Academic Selection (Research Question 4d and 5a)

The third association explored by this thesis centred on the associations between grammar schools and ethnic school segregation via two separate research questions. The first explored whether pseudo-neighbourhoods that contained grammar schools had higher levels of ethnic school segregation than those without (4g), while the second investigated whether grammar schools contributed more to ethnic school segregation than non-selective schools in the same secondary school pseudo-neighbourhood (5a). For both questions,

[192] The association could be correlational, with no causal underlying mechanism.

there was fairly strong prior evidence that suggested they did. For one, there was a well-established relationship between socio-economic school segregation and grammar schools in the academic literature. While this did not inherently translate to ethnic school segregation it was suggestive that the school sector was associated with societal divisions that could well include ethnic groupings. Additionally, differences in the average academic performance of ethnic groups (Strand, 2015) made it likely that a disproportionate number of students from high achieving ethnic groups would attend grammar schools, while a disproportionately low number of students from lower achieving groups would not.

Interestingly, the results provided seemly opposite answers to each of these research questions. At a pseudo-neighbourhood level, places that contained a grammar school(s) tended to exhibit higher levels of ethnic school segregation. Specifically, Model 4_s found that ethnic school segregation was 27.1% higher, on average, in pseudo-neighbourhoods that contained a grammar school than in those that did not ($p<0.001$), once other variables in the model were accounted for.[193] This result, which was broadly consistent in model 5_s, which controlled for regional differences in H values, led to the rejection of the null hypothesis for question 4e. In contrast, school-level models found that grammar schools made similar sized contributions towards ethnic school segregation than other schools located in the same pseudo-neighbourhood. Specifically, model 6_s found no statistically significant difference between the contributions of grammar schools towards their pseudo-neighbourhood's level of ethnic school segregation and the contributions made by other schools located in the same pseudo-neighbourhood ($\theta = 0.142$, $p = 0.17$, $R^2 = 0.005$). This

[193] The findings were also similar to the results of Model 5_s, which controlled for regional differences.

meant that there was insufficient evidence to reject the null hypothesis of research question 5a.

These results suggest that readers should interpret the pseudo-neighbourhood level association with a degree of caution, as the higher average levels of segregation may have little to do with the grammar schools themselves. Specifically, while grammar schools were typically located in pseudo-neighbourhoods with unusually high levels of ethnic school segregation, the schools themselves did not have particularly divergent ethnic profiles relative to other local schools.

For many academics and policy makers, these results may be somewhat surprising. While this is the first study to explicitly investigate the association between academic selection and ethnic school segregation across England's local educational markets, grammar schools have long been viewed as socially divisive due to the substantial body of research that suggest they contribute disproportionately to socio-economic segregation. While these results in no way invalidate those links, they do suggest grammar schools are not as *ethnically* divisive as most would have assumed.

Such findings are likely to have important implications for ongoing debates over whether to expand grammar schools in England (Long, Foster and Robert, 2018). For one they offer the school sector a degree of political cover over their continued failure to recruit FSM students.

8.4.4 Religious Schooling (Research Question 4h, 4i and 5b - 5e)

This thesis also investigated the associations between religious schooling and ethnic school segregation in England via several related research questions. 4h and 4i asked whether pseudo-neighbourhoods with a greater proportion of religious schools tended to experience higher levels of ethnic school segregation for primary and secondary school pseudo-neighbourhoods. Secondly the thesis examined whether religious schools contributed more to ethnic school segregation than non-religious schools in the same pseudo-neighbourhoods via question 5b and 5d. Finally, the analysis also explored whether religious schools with specific religious denominations contributed to more ethnic school segregation than others in the same pseudo-neighbourhood via research questions 5c and 5e.

Previous literature suggested that the nature of this association varied depending on the scale at which segregation was measured. One study found that Local Authorities with more voluntary aided schools tended to have *lower* levels of ethnic school segregation (Gorard, 2016b), while various other analyses suggested religious schools were typically associated with *higher* levels of various types of ethnic "separateness" (Harris and Johnston, 2008, 2020; Allen and West, 2009a; Harris, 2013).[194] These latter studies also chime with the opinions of many residents living in neighbourhoods affected by the 2001 race riots, who told interviewers that they felt religious schools had contributed to the separation of ethnic groups within the school system in the preceding years (Cantle, 2001, p. 34).

The results from this study suggested that associations varied between primary and secondary school pseudo-neighbourhoods. The meta-regression models produced strong

[194] Voluntary Aided schools are almost entirely religious. The term "separateness" was used here instead of "ethnic school segregation" as the city level studies measured different constructs which, while related, do not align with what this study has defined as ethnic school segregation.

evidence to suggest that religious schooling was associated with higher levels of ethnic school segregation for primary school pseudo-neighbourhoods, while results for the secondary school sector were mixed. The analysis in Section 6.1 found that primary school pseudo-neighbourhoods that contained a greater proportion of religious schools tended to experience higher levels of ethnic school segregation. Model 4_p for example, predicted that for each standard deviation increase in the proportion of religious schools within a primary school pseudo-neighbourhood, the expected level of ethnic school segregation rose by 15.6%, after controlling for other variables in the model ($p<0.001$). The strength, direction and statistical significance of this association remain consistent after regional differences were controlled for in Model 5_p, suggesting a robust association. Consequently, the thesis found sufficient evidence to reject the null hypothesis of research question 4h and conclude that there is a positive association between the proportion of schools in a primary school pseudo-neighbourhood that have a religious affiliation and its observed level of ethnic school segregation.

These results aligned with the school level analysis in Section 6.2.2, that found primary schools with a religious affiliation tended to contribute more to ethnic school segregation than non-religious primary schools in the same pseudo-neighbourhood. Specifically, Model 7_p found that religious schools in 2019 contributed 0.51 standard deviations more, on average, to ethnic school segregation than non-religious primary schools in the same pseudo-neighbourhood ($p<0.001$). This led to the rejection of research questions 5b.[195] Later models also found that primary schools linked to smaller religious groups tended to

[195] This asked "Do primary schools with a religious denomination contribute more to ethnic school segregation, on average, than non-religious primary schools located in the same pseudo-neighbourhood?"

contribute substantially more to ethnic school segregation than schools linked to the Church of England, providing sufficient evidence to reject the null hypothesis of 5c. In particular, Jewish and Muslim primary schools were found to contribute disproportionately large amounts to the level of ethnic school segregation in their pseudo-neighbourhood, with Catholic primary schools also contributing more than Church of England, or indeed non-religious schools.

In contrast evidence pointed in opposing directions over whether there was an association between religious schooling and ethnic school segregation within the secondary school sector. Meta regression modelling in Section 6.1 found no evidence to suggest that levels of ethnic school segregation varied between secondary school pseudo-neighbourhoods with different proportions of religious schools. None of the regression coefficients in the models were statistically significant and the estimated strengths of these associations were consistently small. Consequently, the thesis failed to reject the null hypothesis of question 4i.

However, the school-level analyses in Section 6.2.2 found evidence to suggests that, just like primary schools, religious secondary schools contributed more towards ethnic school segregation, on average, than equivalent non-religious schools located in the same pseudo-neighbourhood. Specifically, Model 7_s found that religious secondary schools contributed 0.37 standard deviations more to ethnic school segregation within their pseudo-neighbourhood, on average, than non-religious schools in the same pseudo-neighbourhood. This finding, along with later results that showed that secondary schools linked to smaller

religious groups also tended to contribute disproportionate amounts to ethnic school segregation, led to the rejection of research question 5d and 5e.

It is difficult to know exactly why trends seem to vary between primary and secondary school pseudo-neighbourhoods. One plausible explanation is that pseudo-neighbourhoods containing six secondary schools are likely to recruit students from a far larger geographic region than pseudo-neighbourhoods of six primary schools and that the associations are dependent on the scale at which segregation is measured. As such, differences in the geographic size of primary and secondary school pseudo-neighbourhoods could translate into differences in these associations. It is possible, for example, that religious secondary schools find it easier to recruit students from different ethnic groups within their religious community living across a wider region, whereas religious primary schools may typically serve smaller, more ethnically homogenous localised religious communities, thereby leading to higher levels of ethnic school segregation. Alternatively religious primary and secondary schools may be viewed differently by non-religious families, with religious primary schools being viewed as more “religious” and therefore less attractive places to send their child when compared to larger secondary schools which could be seen as religious in name only. These hypotheses however would need further research to validate though these results provide the starting point for an interesting piece of research in the future.

Overall, this set of results still suggests that religious schooling is typically associated with higher levels of ethnic school segregation at a pseudo-neighbourhood level, though the associations for the secondary school sector are weaker and more complex. This finding should help direct ongoing debates over whether to expand the religious school sector in

England (Long, Danechi and Loft, 2019). In short, they suggest that expanding the school sector carries a risk of raising ethnic school segregation within local educational markets. Whether that is a risk worth taking, or whether a higher level of ethnic school segregation at this geographic scale would be a worthwhile cost to bear even if it were to arise, is a political question which encompasses a broad set of considerations that sit outside of the remit of this thesis.[196] However, by providing an empirical basis on which to have this debate, the study's results nonetheless contribute towards this important discussion, which should help decision making in this area.

Lastly is important this study provides some context to the findings that schools linked to minority religions tended to contribute particularly large amount to ethnic school segregation.[197] It is important for readers to remember that Muslim and Jewish communities have long face substantial hostility from wider British society, much of which continues to this day. Annual UK crime statistics released in 2019, for example, show that 48% of religious hate crimes targeted Muslims, 18% targeted Jews and 12% targeting members of other non-Christian minority religions (UK Home Office, 2019). While currently less acute, many members of the Catholic community, particularly those from Ireland or Eastern Europe, are also likely to have experienced enmity from other members of British society, which naturally would encourage a desire to cluster with members of your own religious or ethnic group. In other words, these schools could well provide a safe and secure space for members of these religious communities which means that their outsized contribution

[196] Readers who are interested in this debate can consult a broad range of literature (Burtonwood, 2000; Jackson, 2003; Gokulsing, 2006; Pring, 2007; Dwyer and Parutis, 2013)

[197] This draws on the ideas of Section 2.1.5 in the Literature Review and the ethical questions that were considered on how best to conduct this study, as described in Section 4.10.

towards ethnic school segregation could reflect the actions of wider society, not just the preferences of the individual religious or ethnic communities. More research would be needed to flesh these considerations out in more detail.

8.4.5 School Quality (Research Question 4j, 4k, 5f and 5g)

The final type of association investigated by this analysis was school's OFSTED ratings. As with the last two predictors, this association was explored at a pseudo-neighbourhood level, as well as between schools located within the same pseudo-neighbourhood. Two sets of research questions were posed. Firstly, the study asked whether there was an association between the average OFTSED rating received by primary and secondary schools in a pseudo-neighbourhood's and the pseudo-neighbourhood's level of ethnic school segregation. The second set of questions examined whether there was an association between a school's OFSTED rating and its contribution toward ethnic school segregation, relative to other schools located in the same pseudo-neighbourhood.

Unlike other predictor variables this association had yet to be empirically studied in England and so its inclusion was justified based on the theoretical work of competition theorists. Using this research, the thesis suggested that areas with lower school quality were likely to suffer from higher levels of ethnic school segregation as families would be incentivized to compete more fiercely to gain entry to the limited number of places in higher rated schools. This would lead to students being sorted into schools based on whether their family had the financial and social capital to compete for entry into highly demanded higher rated schools, either in the housing market or via the appeals process. As different ethnic groups in

England have different levels of financial and social capital, the study theorized that higher competition could ultimately lead to higher levels of ethnic school segregation.

However, the thesis found no evidence to suggest there was an association between ethnic school segregation and "school quality", as measured by OFSTED ratings, either for primary or secondary schools. None of the meta-regression models found an association between the average OFSTED rating of schools within a pseudo-neighbourhood and the pseudo-neighbourhood's level of ethnic school segregation once other control variables had been accounted for. This suggests that once other variables were accounted for, pseudo-neighbourhoods with "low quality" schools were no more likely to have high levels of ethnic school segregation than those with "high quality" schools. Similarly, no evidence was found in Section 6.2.3 to suggest that primary or secondary schools with different OFSTED ratings made different sized contributions towards ethnic school segregation (H_j) relative to other schools within the same pseudo-neighbourhood. In other words, "low quality" schools contributed similar amounts towards ethnic school segregation in 2019 than "high quality" schools did. Based on this evidence, the thesis failed to reject the null hypothesis for research question 4j and 4k, as well as for 5f and 5g.

These results suggest more segregated local educational markets do not offer lower quality schooling, on average, across England, at least as understood by OFSTED ratings. Specifically, these results suggest that pseudo-neighbourhoods with high levels of ethnic school segregation did not have "lower quality schools" than pseudo-neighbourhoods with low levels of segregation once other control variables were accounted for. Similarly, schools which had particularly divergent ethnic profiles, relative to other schools in their pseudo-

neighbourhood, were no less likely to be rated "outstanding" or "good" than schools who had ethnic profiles that reflected those of other local schools in the same pseudo-neighbourhood. This also suggests that neighbourhoods with ethnically segregated schools are no less likely to provide "high-quality" school experiences, as measured by OFSTED ratings, than neighbourhoods where schools educate students with similar ethnic profiles.

This finding could well remove a powerful argument for encouraging greater ethnic integration which would be likely to be viewed by most people as apolitical. Instead, advocates for an ethnically integrated school system will need to be reliant on more subjective positions built on the type of society that England "*should*" aim to build, rather than on more strictly "educational" basis. In the short term, this may not be particularly important, as the government is currently committed to reducing ethnic separations within schools and most voters are likely to agree that schools should be ethnically inclusive. However, in the longer term, if inter-ethnic tensions were to worsen or residential patterns shift, the loss of this argument may be more acutely felt.[198]

8.5 The role of the school admission system in shaping patterns of ethnic school segregation in England

The second analysis conducted by this thesis provides a detailed insight into how England's national school admission system helps to shape patterns of ethnic school segregation. The

[198] Future research could explore whether pseudo-neighbourhoods where schools have a greater range of OFSTED rating exhibit higher or lower levels of ethnic school segregation than those with more uniform OFSTED ratings. Readers should note however that the results from a pilot study conducted by the author would suggest little evidence for such an association.

following section begins by exploring how ethnic school segregation changed between each stage of the admission cycle, before reviewing the results of the study's seven simulations that explored how patterns of ethnic school segregation might have differed if students had been allocated to schools using a variety of different approaches. The two subsections initially reiterate why the analysis was carried out, what the major findings were and how these results addressed each of the research questions. It then explains how the results change current academic understanding, reviews the relevance for practitioners and policy makers, before suggesting future research that could improve our understanding. Collectively, these discussions engage with research questions 6a – 6g.

8.5.1 How did ethnic school segregation shift between each stage of the national school admissions cycle?

Due to a lack of national level data, no research in England had charted how patterns of ethnic school segregation shift between each stage of the yearlong national school admission process. This means it was impossible to gauge the extent to which ethnic school segregation was determined by how parents initially completed their child's application forms, the subsequent allocation of offers to students, or the latter appeals process which parents could use to challenge the offer they received. However, following the DfE's decision in 2016 to collect and release application data for all students entering secondary schools across England, there was a clear opportunity to describe these changes in detail to track how ethnic school segregation was shaped across an admissions cycle.

The major finding from this analysis was that the likely level of ethnic school segregation rose between each stage of the yearlong admissions process. If every student had attended the secondary school they listed as their first choice in their application form, the median pseudo-neighbourhood *H* value would have been 7.1 for these students in the following school year (2017/2018). While the cost of providing this "pure school choice" allocation would be sufficiently expensive to make it largely impossible, it nonetheless provides a useful benchmark to interpret later results. If every student had attended the school that they were offered 6 months later in the second stage of the application cycle, then ethnic school segregation for the 2017/2018 school year would have been marginally higher, with a median *H* value of 7.2 for pseudo-neighbourhoods (+ 1.5%).[199] This allocation can be thought of as the combined outcome of all school's oversubscription criteria. Finally, the study found that the average level of ethnic school segregation that was observed in 2017/2018 was 6.5% higher than the previous stage (H = 7.7), in what this study termed the third stage of the application cycle. This suggest that the appeals process, or any other process that allowed students to attend a secondary school they had not originally been offered, resulted in an increase in ethnic school segregation within local pseudo-neighbourhoods.

As there was only a minimal difference in the median *H* value between stage one and two, schools' oversubscription criterion seemly played a small role in determining ethnic school segregation for this cohort, above and beyond what was explained by the school preferences expressed by parents. In other words, the way in which the admission system

[199] These figures, along with all other mentions of segregation are only in reference to the year group studied, not the wider student population.

initially *offered* places to students in the Spring 2017 was not associated with a meaningful increase in ethnic school segregation, given how parents completed their child's application form the previous autumn. In contrast, the appeals process, which allowed parents to challenge the school that their child was offered in the Spring 2017, increased ethnic school segregation by around 6.5%.

These figures provide a useful guide as to where policy makers and academics should focus their efforts over any potential reforms of the school admission system. Specifically, they suggest that encouraging parents to change how they completed their child's application form, and to a lesser extent, investigating mechanisms within the appeals process that may be associated with increased rates of ethnic school segregation, hold the greatest potential for reducing ethnic school segregation within the current school choice model. The results do not, however, mean that reforming the second stage of application process could not also help reduce ethnic school segregation. Indeed, as shown by the simulation analysis in the following section, reforming how places are initially offered to students could *potentially* reduce ethnic school segregation to a median *H* value as low as 5.9, which would represent a 23% drop.

It is also worth highlighting that there remains plenty of room for future research to develop this approach further. For example, if researchers collected information on the oversubscription criteria of individual schools, they could offer a more fine-grained analysis that explores whether particular criteria were associated with higher levels of ethnic school segregation than others. This would provide further insight into how the admissions criteria of schools feed into patterns of ethnic school segregation.

8.5.2 What the simulated school admissions systems tell us about ethnic school segregation in England?

The study's final analysis created seven simulations that allocated year 6 students to secondary schools using a range of different approaches and then measured the level of ethnic school segregation within pseudo-neighbourhoods for each simulated pattern of school attendance. This provides a new cache of evidence that contextualises England's current approach and explores the possible implications that reforms to the school admissions system could have on patterns of ethnic school segregation. This subsection begins by summarising the major findings before discussing their relevance to academics and policy makers alike. It ends by highlighting the limitations of this method and suggests future research that could enhance our understanding further.

The major finding was that ethnic school segregation tended to be lower in all seven simulations than the median H value that was observed in practice. While the observed median *H* value for this cohort within pseudo-neighbourhoods was 7.7, the median H values for the simulation ranged from 7.0 to 4.7, which was 10% to 39% lower.

The five simulations that operated within the current school choice framework exhibited higher median levels of ethnic school segregation than the two simulations that did not take into consideration the expressed preferences of parents. The three simulations that relied on students' home-school distance, random allocation, or academic competition to allocate places at oversubscribed schools within the current school choice system had the highest

median *H* values of the simulations, ranging from 6.6 – 7.0. These values were between 10% - 14% lower than the median *H* value observed in practice. The two simulations that used academic banding within a school choice system had the lowest average levels of ethnic school segregation of the simulations that incorporated school choice, with median *H* values of 6 and 6.2, which was roughly 21% lower than those observed in practice. However, the two simulations that allocated students to schools based solely on their home-school distance, or using their home-school distance in combination with their KS2 test scores, exhibited the lowest levels of ethnic school segregation of all the seven simulations. These simulations had median *H* values of 4.7 and 4.9 for pseudo-neighbourhoods, respectively, which was around 39% lower than observed in practice. Collectively, these results addressed the research questions 6b – 6h.

These findings hold several implications for current academic understanding and the design of national education policy. Firstly, the results provide initial evidence to suggests that reforming the national school admissions system could reduce ethnic school segregation, indicating that school admissions offer the government a key policy lever that it could use to fulfil its stated objective of creating schools where students "form lasting relationships with others from different ethnic… groups" (HM Government, 2018a, p. 26). These results should therefore encourage the DfE, and other relevant authorities, to consider how admissions policies could be updated to encourage greater ethnic integration in local schools. Hopefully, other academics will contribute to this effort.

The simulations also provide the first "ballpark figures" on how different types of reforms could potentially influence patterns of ethnic school segregation across England. If the

government reformed how places at schools were offered to students within the current school choice model, these results suggests that the median level of ethnic school segregation would fall by around 10% for year 6 students entering secondary school. If these reforms included some form of academic banding, then this reduction would be likely to be higher, perhaps as much as 20%. However, to produce the most drastic reduction in the average level of ethnic school segregation, these results suggest that policy makers would need to ignore the preferences of parents and allocate students to schools based on their home-school distance. If they did this then these results suggest ethnic school segregation could potentially fall by as much as 40%, *all things being equal,* which is of course unlikely.

Broadly speaking, the results indicate a tradeoff between how much influence parents have over the secondary school their child attends and the ability of the government to reduce ethnic school segregation within the school system. All five simulations that considered parental preferences when allocating students to schools resulted in a higher median level of ethnic school segregation than the two simulations which did not. Perhaps this is not surprising. As discussed in the literature review, humans have a well-documented tendency to group together with people that are similar to us. If offered the chance, perhaps it is only "human" for parents to lean toward selecting schools where pupils look like their family, and come from groups with similar levels of social and economic capital. Indeed, historical conflicts over school (de)segregation have regularly centered on the contrasting desires of local parents and the aims of government officials.[200] Unfortunately, it is beyond the scope

[200] For example, the school busing programmes in Boston during the 1970-80s.

of this thesis to evaluate what admissions systems secures the best balance between these two positions. Instead, the study has provided evidence to suggest this trade off exists and has produced findings on a range of different school admission systems that each prioritize parental preferences to varying degrees. Hopefully this will offer a sufficient breath of options for future studies to debate the relative merits of each approach more thoroughly.

While the results from this analysis relate to England as a whole, the findings should also provide useful insights for policymakers and academics focused on local policy making. Several local areas in England are actively engaged in redesigning local school admission systems to promote greater social cohesion within their communities. With an ever-growing focus on the role that ethnicity plays in British society, it is likely more local authorities will evaluate their school admissions in the coming years to ensure students from different ethnic groups do not end up being implicitly sorted into different schools. As mentioned in the literature review, the city of Brighton and Hove pioneered the use of randomly allocating places at oversubscribed schools to reduce social stratification in their schools, and Tower Hamlets have used academic banding to similar effect. The results from this study offer supporting evidence to suggest that both approaches can be used to reduce ethnic school segregation in England, particularly in the case of academic banding. While any direct translation of these national level findings to a single local area should be done cautiously, they nonetheless provide a useful guide for areas interested in these types of reforms.

While these simulations have made an important contribution to current academic understanding, there are important methodological limitations of this approach that should

encourage readers to interpret the results with a degree of caution. For one, the simulations all use simple procedures to allocate students to schools that ignore many of the complexities of England's national school system. For example, none of the simulations distinguish between single sex or mixed gender schools, or between schools with religious affiliations. They also ignore the legal requirement for schools to offer places to students in care, or with special needs, and they sidestep the issue of whether to prioritise applicants with siblings at the school. While this simplicity means it is easier to grasp what the simulations are doing, it also means that each simulation may not provide an *accurate* prediction of how that policy would impact ethnic school segregation if implemented in reality given the "complex and diverse patterns of competition" (Taylor, 2001, p. 198) in England's school system.

In addition, the simulations implicitly assume that the actions of parents within the current system would be broadly similar under a radically different school admissions system. It assumes, for example, that parents would apply for the same schools, and list them in the same order, and that they would choose to live in the same home. For many parents, this assumption would not hold. If highly desired schools allocated places solely based on home-school distance, then many parents would move home to be closer to the school of their choice. Similarly, if places at oversubscribed schools were allocated at random then many parents would have applied to schools that they would otherwise have been unlikely to gain entry to.

While both of these concerns are important to recognize, the simulations nevertheless contribute to knowledge in this field by broadening the empirical evidence that academics

and policy makers can draw on. Ethnic school segregation is an incredibly complex phenomena, produced by millions of interacting decisions on the part of students, parents, teachers, education policy makers and successive governments. Therefore, no methodological approach could ever realistically capture all the various factors that underpin different patterns of ethnic school segregation, a point that previous academics have also made (Harris and Johnston, 2020). Instead, academics should aim to produce a range of different types of studies that can provide a solid understanding of what is likely to be happening across England. In this sense, these simulations contribute towards improving academic understanding in this field without claiming to provide *the definitive* answer to how school admissions relate to ethnic school segregation.

8.6 Methodological Contributions of this study

The methodological techniques developed by this thesis also offer an important contribution towards current academic literature. First, the study's discussion over what scale ethnic school segregation should be measured at, and the subsequent development of "pseudo-neighbourhoods", offers a useful tool for researcher with exciting applications for future work. Second, the study's use of a multi-group index to calculate ethnic school segregation between five ethnic groups in a single concise measure provides an approach which better reflects the growing diversity of England and its school age population. Both contributions are discussed in more detail within the following two sub-sections, along with their limitations and possible avenues for future use.

8.6.1 The creation of “pseudo-neighbourhoods”

As discussed in Section 2.1.2 of the literature review, the continued use of government boundaries to measure ethnic school segregation remains a weakness of British research. Not only are Local Authorities (LA), the government boundary of choice, too large to reflect the choices available to parents but they also varied substantially in size, making them an inconsistent unit of analysis. Based on these issues, the thesis argued that many segregation indexes calculated using LA boundaries risked being conceptually vague and heavily influenced by their relative size and ethnic residential patterns.

While many academics had acknowledged this issue, this thesis was the first in England to propose a new, practical approach to address the problem. This thesis argues that pseudo-neighbourhoods offer an innovative new solution that addresses many of these issues. Namely, the use of pseudo-neighbourhoods brought three main advantages. First, as the average number of schools within each pseudo-neighbourhood remains directly controlled by the researcher, pseudo-neighbourhoods allow researchers to measure ethnic school segregation at an appropriate geographic scale which must be justified and made explicit to the reader. Second, as pseudo-neighbourhoods remain similar in size, they offer a more consistent unit of analysis which avoids the inherent bias suffered by indexes that are calculated using government boundaries that vary considerably in size. Third, by creating multiple sets of pseudo-neighbourhoods and finding the average level of school segregation across each iteration, researchers can produce measures of segregation which are not reliant on a single set of area boundaries. This reduces an inherently arbitrary factor that underpins research that relies on the use of government boundaries.

There are likely to be researchers investigating segregation around the world who could benefit from using this approach. In many cases, the incentives to use pseudo-neighbourhoods will be the same as in this study. Government boundaries will either be inappropriately sized or will bear little resemblance to the school choices available to parents. However, in areas where government boundaries are drawn in a way where ethnic groups are located within the same area, unintentionally or by design, the benefits gained by using pseudo-neighbourhoods increase dramatically.[201] In these instances, pseudo-neighbourhoods offer a flexible way of measuring school segregation that avoids the need to map, in detail, where specific boundaries should be drawn. This is particularly useful when measuring segregation in areas where drawing boundaries in any shape or size would struggle to be viewed as impartial.[202]

The use of pseudo-neighbourhoods also offers the chance to measure patterns of school segregation within areas that encompasses two or more school systems, which are unlikely to have a uniform or appropriately sized set of government boundaries that span both jurisdictions. This could facilitate a more accurate measurement of school segregation within contested border regions, federal systems where education is organised at a state level or even across large supra-national blocs like the EU. In all three instances being able to produce a neutral, uniformly sized set of boundaries would be invaluable to offering a more accurate picture of trends.

[201] Examples include the boundaries within Northern Ireland or gerrymandering of school districts in the Southern States of America (Richards, 2014). If researchers use these boundaries, the segregation indexes will be artificially low and will fail to accurately detect the extent of divisions within the school system.
[202]Northern Ireland is a good example.

There nonetheless remains room for further developments in this technique. As researchers now have a tool to make boundaries of whichever size they prefer, a more detailed discussion is required over what the ideal size of areas should be when measuring ethnic school segregation. This thesis has proposed three properties to aim for, and ultimately argues that pseudo-neighbourhoods in England should contain 6 schools, on average, based on the number of choices that parents can list on their child's application forms. This, of course, leaves plenty of room for a more robust and rigorous debate.[203] Secondly, the technical ways in which pseudo-neighbourhoods are created could also benefit from further scrutiny. Specifically, it would be worth exploring whether other clustering algorithms produced "better" school clusters than k-means, which would of course need to be accompanied by a discussion over what a "good" set of clusters was. Lastly, it would be interesting to see if there was a way of grouping schools together based on the time it takes to travel between schools. If possible, this would avoid schools being grouped in the same pseudo-neighbourhood that were geographically close to one another but difficult to travel between, such as two schools separated by a motorway, hill or river.[204]

There are also important limitations of this method that readers should be aware of. Firstly, the approach is more computationally challenging and more theoretically challenging than using preestablished government boundaries. Until a software packaged is released, researchers would need to be proficient at coding to create pseudo-neighbourhoods and

[203] Alternatively, researchers could avoid choosing one size and instead measure ethnic school segregation at multiple different scales. A brief pilot analysis was conducted to show how this approach may work. The results are described in Section 11.27 of the appendix.

[204] As this would add considerable complexity to the analysis, it would have to provide a noticeable difference to findings in order to be worth undertaking.

the statistical models they need to employ tend to require greater complexity than more traditional approaches. For example, this thesis needed to use meta-regression analysis to analyse 100 sets of pseudo-neighbourhoods, rather than a single set of regression models on a fixed set of government boundaries. Even basic descriptive statistics are trickier. For example, as none of the pseudo-neighbourhoods have place names, academics need to assign pseudo-neighbourhoods either to a region or a city which makes describing patterns more cumbersome. Finally, while Local Authorities or other government boundaries suffer from methodological issues, many readers will be familiar with their usage and feel comfortable interpreting results that are aggregated at this level. In contrast pseudo-neighbourhoods require readers to conceptualise findings over multiple sets of clusters, which is a more abstract idea. While these difficulties do not invalidate the approach, particularly given the weaknesses of using existing government boundaries, they nonetheless are worth highlighting.

8.6.2 Using the multi-group H Index

This thesis has been the first study in England to use a multi-group index to measure ethnic school segregation between multiple ethnic groups across the whole country. This differs from previous research that has either measured ethic school segregation using a binary index based on “White” and “Non White” students (Gorard, 2016b) or had produced multiple binary indexes to measure segregation for each ethnic group in relation to students from all other backgrounds (Burgess, Wilson and Lupton, 2005). This thesis argued that using binary categorisations were an increasingly inappropriate way of conceptualising ethnic identities in England, and that it was now mathematically feasible and conceptually

important to use a measure that captured "separateness" between different minority groups too. It also argued that, while using multiple binary indexes offered a better approach than a single "White" / "Non-White" index, the narrative can often be overwhelming, and can leave readers having to weave together different sets of results to gain a holistic picture of ethnic school segregation.

In contrast, this thesis argues that multi-group measures of segregation, such as the *H* index, offers a more concise measure of segregation that better reflects how ethnic identities are conceptualised in England today. As explained in Section 2.1.5, it argues that multi-ethnic measures of segregation are also likely to encourage a more balanced description of patterns, where the focus of results is on the geographic area being studied rather than any given ethnic group(s). When researchers have previously described segregation using a binary index, they tend to discuss how segregated *ethnic minority students* are across LAs in England. However, as segregation is a relational measurement between different ethnic groups, ethnic minority students are only as "segregated" from their White British peers, as the White British peers are from them. By presenting the narrative in relation to ethnic minority students, the focus has tended to be unduly placed on ethnic minority students and their actions. In contrast, by utilising a multi-group measure that combines the collective level of segregation for all ethnic groups in an area, the focus is shifted onto the schools within a geographic area, which is likely to be less provocative and also better reflects the unit at which segregation is measured. Furthermore, given that schools and local government organisations are better placed to engage with ethnic school segregation than individual ethnic groups, presenting research in relation to geographic areas should also help to translate findings into actionable policy.

However, there are several limitations that using a multi-group index brings. First, unlike binary indexes, none of the multi-group indexes provide results that have a particularly intuitive interpretation. This forces academics to describe findings in relational, rather than absolute terms, which can make it hard for readers to conceptualise what the results signify. In turn this is likely to limit public interest in these types of findings, which could hinder the reach of such research. Second, by reporting results for all ethnic groups in a single measure readers become unaware of whether certain ethnic groups are more segregated from other ethnic groups, which could obscure important underlying trends.[205]

8.7 Limitations and Future Research

The chapter finishes by discussing a wider limitation of the study before proposing two areas which future research may wish to explore. While the study has provided a rigorous investigation of ethnic school segregation, it has only referred to ethnic residential patterns or socio-economic segregation briefly, even though both are likely to play important roles in determining school attendance patterns. While the conceptual and methodological complexity of measuring ethnic school segregation necessitates a targeted methodological approach, readers are encouraged to read wider literature on different forms of segregation to build a more holistic understanding of the findings set out in this thesis.

[205] Readers should note the H index can be decomposed in a way to calculate the relative contribution of different ethnic groups. (Reardon and Firebaugh, 2002). While this study has avoided using utilising this feature of the index, for the reason outlined previously, it acknowledges that it could potentially lead to interesting findings being overlooked.

The following subsection then suggests two pieces of future work that could improve our current understanding on ethnic school segregation. First, future analysis should investigate how the availability of transport links relate to patterns of ethnic school segregation. It also argues that a broader range of methodological approaches is required to better understand the underlying dynamics of ethnic school segregation. Second, it argues that the fieldwork experiments pioneered by economists Banerjee and Duflo offer an interesting template for a new type of research within the field of ethnic school segregation.

8.7.1 Limitations

By focusing on how to rigorously measure ethnic school segregation, this thesis provides readers with a focused description of ethnic "separateness" within England's school system that dedicates little attention to the role played by ethnic residential segregation or socio-economic school segregation. While this approach is typical within the academic literature, which considers ethnic segregation separately to socio-economic segregation, and residential segregation separately to school segregation, all these processes are likely to be highly interlinked. Where children live and go to school clearly impact one another, and many of the factors underpinning ethnic school segregation are likely to overlap with socio-economic school segregation (Harris and Johnston, 2020). By describing patterns in relation to ethnic school segregation, it is therefore possible for undue focus to be placed on local school systems or the ethnicity of students, rather than important residential or socio-economic dynamics within local communities.

While the conceptual and methodological complexity of measuring each type of segregation necessitates this narrower, more targeted approach, readers should try and read research on other types of segregation to ensure they can contextualise the findings of this study. Fortunately, there is a large body of literature on ethnic residential segregation (Harris, 2016; Catney, 2018; Harris, Charlton and Brunsdon, 2018; Jones, Manley, *et al.*, 2018) and socio-economic school segregation in England (Gorard, Taylor and Fitz, 2001; Allen and Vignoles, 2007; Gorard and Siddiqui, 2018). This thesis should therefore be viewed as one part of a wider literature on segregation in England, that ideally would be read in tandem.

8.7.2 Future Research

The study also suggests two pieces of future research work could improve existing literature. While this study has provided a robust investigation into five important predictor variables, there remains additional factors which should be investigated for possible associations with ethnic school segregation. Future research should explore whether ethnic school segregation varies between areas with different types of public transport networks. Despite a number of academics highlighting the likely importance of transport networks to ethnic school segregation (Noreisch, 2007, p. 83; Gorard, 2016b, p. 138) it was deemed outside the scope of this thesis to accurately assess the scope or quality of a pseudo-neighbourhood's local transport network, in addition to the new methodological approaches used by this thesis. However, given that the UK government produces a dataset with all access points to the nation's bus, train, tram and tube network, including the coordinates of each stop or station (HM Government, 2021), future researchers would have all the required data to produce an interesting study. Furthermore, given the direct control

that the government has over transport systems, it is likely that any findings could help to inform future government decision making.

The field would also benefit from a wider variety of methodological approaches to studying ethnic school segregation. Currently there are two broad approaches used by British researchers. The largest body of work, to which this study belongs, focuses on the quantitative measurement of ethnic "separateness" in schools. While different studies conduct different calculations to measure different constructs, they are normally carried out remotely using secondary data on the proportion of each school that belongs to a given ethnic group.[206] The second, smaller body of work interviews parents about the factors that influence the choices they make regarding their child's schooling.[207]

While both bodies of research are useful, there are plenty of other ways of studying ethnic school segregation which can provide new and useful insights into this complex phenomenon. In particular, the field experiments pioneered by economists Banerjee and Duflo for investigating poverty offer an interesting template for future research to explore the dynamics driving school segregation (Banerjee and Duflo, 2013; Duflo and Banerjee, 2017). Academics could, for example, investigate whether information campaigns could be used to encourage ethnic groups to apply for schools where they are under-represented.

[206] This study also conducted an additional analysis that used Random Forest models, a machine learning tool, to estimate the combined predictive power of the study's independent variables on H and to explore which variables had the greatest importance in estimating a pseudo-neighbourhood H value. While these findings were ultimately not reported in the main body of the text, Section 11.28 in the Appendix explains this approach and presents the findings.

[207] While "Anglo-American culture tends to privilege quantitative evidence over other forms of knowledge" (Gibson and Asthana, 2000a) there is real "scope for further investigation into the processes of segregation, using qualitative research to reveal motivations and perceptions"(Farley and Blackman, 2014).

Similarly, studies could explore whether publishing a school's contribution to ethnic school segregation (H_j) could encourage the school to act towards attracting a more ethnically representative student cohort.[208] Alternatively, academics could capitalise on "natural experiments", where a change occurs that was not initiated by the researcher, but which nonetheless provides an opportunity for some type of comparison. For example, it would be useful to understand how changes in the ethnic make-up of a school's staff, particularly those in senior leadership positions, effects the ethnic profile of students applying to the school. This could help us to understand whether schools attract a more ethnically diverse applicant pool if their senior leadership team becomes more ethnically diverse.

There are two major advantages to these studies. Firstly, if designed carefully, they offer the closest thing to causal findings that this field is likely to offer. While the studies would still fall short of a Randomised Control Trial, they would nonetheless allow researchers to get closer to measuring the underlying dynamics of school segregation. Secondly, these studies would be geared towards finding practical solutions to reduce ethnic school segregation. This helps to correct a general bias in educational research where problems are often (re)identified, with far less attention paid to design potential solutions to address such problems. As Gorard argues "the field needs to move on from its simple laying bare approach and begin to find the underlying malleable causes of the manifested educational inequalities" (Gorard, 2014, p. 89).

[208] This mirrors similar approaches in the UK to promote equal pay between people of different genders and ethnicities.

8.8 Summary

The discussions chapter has summarised this study's results and incorporated them into the wider body of academic literature. Section 8.1 showed how national changes in the ethnic profile of England's school age population led to a growing level of ethnic diversity across the country, including in many rural areas which were previously ethnically homogenous. Section 8.2 then highlighted the substantial reduction in ethnic isolation between 2006 – 2019 and suggested that this provided a growing space for inter-ethnic contact between young people in English schools. This subsection also drew attention to the substantial variation in ethnic isolation between students of different ethnic backgrounds and highlighted that many White British students have far less space for inter-ethnic contact than those from other ethnic groups. Section 8.3 then presented statistics on ethnic school segregation within the English school system. It explained that between 2006 – 2019 the average *H* value for primary and secondary school pseudo-neighbourhoods declined by around 25% which suggests that ethnic "separateness" within local school systems fell substantially during this period. At the same time, it also stressed that there remain many pseudo-neighbourhoods which continue to experience high levels of ethnic school segregation, many of which are located in the north of the country.

Section 8.4 then discussed which types of neighbourhoods tended to experience higher levels of ethnic school segregation and the kinds of schools that contributed most toward such "separateness". Section 8.5 then discussed the findings from the first national analysis into how the school admissions system shaped patterns of ethnic school segregation in England. It explained that the average level of ethnic school segregation rose between each

of the three stages of the application process and examined findings from each of the seven simulations produced by this study.

After the substantive results had been discussed, the chapter outlined the methodological contributions that this thesis made to the academic literature by stressing the benefits of using “pseudo-neighbourhoods” and the multi-group H index, as well as discussing their limitations. Finally, the chapter ended with a discussion on the wider limitations of this study by encouraging readers to engage with research on other forms of residential and socio-economic segregation to contextualise the findings of this thesis. It then highlights some potential avenues for future research by suggesting that researchers explore how public transport networks shape patterns of ethnic school segregation and that the field could benefit from using a broader base of methodological approaches in the future.

9 Conclusion

England is in the middle of becoming an increasingly multi-ethnic society. Due to continued inward immigration, the higher fertility rates of women born outside the UK, the increased number of inter-ethnic families and the older age profile of the White British population, an exponential change in the ethnic profile of the UK is taking place, which requires public policy and research to adapt with it.

The state school system is at the vanguard of these shifts. Nearly a third of all students in England's state school system are from an ethnic minority background, more than double the proportion in the wider population. Driven by various demographic factors, the diversification of England's youth is unlikely to slow, which means the school system is poised to serve a more ethnically diverse group of people than any other government service, possibly for decades to come. How schools engage with this transition will therefore lay the foundation for other government services and will heavily influence ongoing conversations on the future of England's multi-ethnic society.

Within this changing landscape there have been growing concerns over perceived ethnic divisions within England's school system, with media outlets, think tanks, government reports and prominent politicians all sounding alarm bells (Casey, 2016; May, 2016; BBC, 2017). Despite such attention, there remains a growing need for more empirical evidence to support these important conversations. As described in the literature review, academic research on ethnic school segregation in England state school system is fairly small and

often dated. A limited number of academic studies have measured evenness ethnic school segregation at a national level and a number of these were based on data that was close to two decades old (Johnston, Wilson and Burgess, 2004; Burgess, Wilson and Lupton, 2005; Gorard, Hordosy and See, 2013; Gorard, 2016b; Harris and Johnston, 2020). Unfortunately, these studies also suffered from two methodological issues. Firstly, they calculated evenness segregation between schools within Local Authority boundaries, which are too large to be considered "local educational markets", and they vary substantially in size, which introduces a hidden bias into their results. Secondly, the approaches they took to measuring ethnic school segregation no longer offer a particularly appealing approach in an increasingly multi-ethnic society where Local authorities have a diminished role in directing local schools. Lastly, only a small number of studies offered insights into the predictors of segregation, which left policy makers and academics with limited evidence to design policy solutions that could address segregation, as and when it arose.

Unfortunately, the quality of public debate has suffered due to these gaps. Instead of using national level statistics, most public conversations on segregation focus on the small number of schools where nearly all of the children come from ethnic minority backgrounds, with these used as case studies to suggest a wider "divided" school system (Channel 4, 2019). Rarely is it acknowledged what poor quality evidence this approach offers for evaluating the entire school system.

In light of these issues, this thesis has contributed to addressing some of the gaps in the literature through three sets of analyses. The first, set out in Chapter 5, offered a detailed description of the ethnic profile of England's school age population, the amount of ethnic

isolation in English state schools, and the patterns of ethnic school segregation in England's local pseudo-neighbourhoods. The second, outlined in Chapter 6, explored whether certain types of pseudo-neighbourhoods experienced higher levels of ethnic school segregation than others, or whether certain types of state schools contributed more to segregation. The third, described in Chapter 7, investigated the links between ethnic school segregation and England's school admissions system by charting how segregation changed between each stage of the 2016/17 admission cycle as well as using simulations to evaluate whether ethnic school segregation would have been different if students had been allocated to schools using alternative approaches. As a collective body of work, this thesis offers one of the most in depth, rigorous analysis of ethnic school segregation in England which has led to a plurality of important findings. The key results are described below.

Chapter 5 explained how an increase in the number of ethnic minority students, coupled with a falling number of White British students, led to a substantial rise in the level of ethnic diversity within the school age population between 2006 - 2019. 9 out of 10 pseudo-neighbourhoods in England educated a greater proportion of students from ethnic minority backgrounds in 2019 than they did in 2006, with new areas of ethnic diversity being created throughout the country.

Ethnic isolation was then shown to have fallen which suggested a growing space for inter-ethnic contact for young people in English schools. The analysis found that students in 2019 attended schools where, on average, 75% of their peers were from the same ethnic group, down from 90% in 2006. This is likely to be welcomed news for the government, as they seek to foster positive inter-group relations between young people in schools. The thesis

stressed, however, that this result should be interpreted cautiously. As levels of ethnic isolation are heavily influenced by the ethnic profile of students, the reduction in isolation will be largely down to the diversification of the school age population rather than a change in patterns of school attendance. In addition, this "average" figure obscures two distinct levels of ethnic isolation experienced by students from different ethnic backgrounds. While students from ethnic minority groups typically went to schools where between 10% to 30% of their peers came from the same ethnic group, White British students attended schools where 85% of the students were typically White British. These differences suggests that if schools are to become spaces for young people to build friendships across ethnic lines, reducing the level of ethnic isolation experienced by White British students should remain a priority, as most students from ethnic minority backgrounds already attend ethnically diverse schools.

Chapter 5 ended by mapping patterns of ethnic school segregation within England's pseudo-neighbourhoods.[209] The results showed that pseudo-neighbourhoods in England had a median *H* value of 5.1 in 2019, with an inter quartile range between 3.2 and 8.2. Unfortunately, these figures do not have a straightforward substantive interpretation, but they do offer a useful benchmark for future studies. The section identified that segregation varied considerably between different parts of England, with pseudo-neighbourhoods in the north tending to experience higher levels of ethnic school segregation than those in the south. This was both in relation to the median average pseudo-neighbourhood *H* values, as well as the proportion of pseudo-neighbourhoods that experienced the highest levels of

[209] This provided a more robust measure of "separateness" which was less influenced by the ethnic proportions of the school aged population.

segregation (*H* > 20). For government agencies, as well as researchers in this area, this should provide an important signal as to where their effort should be focused. In short, while London remains the most diverse part of the country, the school system has far lower levels of ethnic school segregation than neighbourhoods outside the capital. Specifically, if there is a desire to reduce ethnic segregation within English schools, then energy and resources should be diverted to supporting areas in the north of the country.

At the same time, ethnic school segregation within primary and secondary school pseudo-neighbourhoods was found to have fallen substantially between 2006 – 2019. Median H values for primary school pseudo-neighbourhoods fell from 6.8 to 5.0 during this period, while median H values for secondary school pseudo-neighbourhoods dropped from 7.6 to 5.8. This constituted around a 25% reduction in the average level of segregation within a 13-year period for both school sectors, which suggests that the ethnic profile of local schools has tended to converge. In other words, young people seem to be attending local schools that increasingly educate groups of students with similar ethnic compositions to one another. The analysis also found that pseudo-neighbourhoods in the north of England tended to experience sharper falls in ethnic school segregation than pseudo-neighbourhoods in the south of the country, leading to a degree of greater similarity in segregation levels between different parts of the country.[210]

Chapter 6 then built upon these descriptions by evaluating whether ethnic school segregation varied between pseudo-neighbourhoods with different characteristics, or

[210] These results were also validated by three sensitivity checks which found changes in the sample of pseudo-neighbourhoods, the percentage of White British students and the level of random segregation were unlikely to be responsible for the observed changes in ethnic school segregation.

whether certain types of schools contributed more to ethnic school segregation than others. Using a series of meta regression models, the analysis found that ethnic school segregation tended to be higher for both primary and secondary school pseudo-neighbourhoods in areas with lower average house prices and higher rates of FSM. Additionally, ethnic school segregation tended to be higher in primary school pseudo-neighbourhoods with more religious schools, and secondary school pseudo-neighbourhoods with a grammar school. In contrast, the thesis found no evidence to suggest there was an association between a pseudo-neighbourhood's level of ethnic school segregation and its population density or the average OFSTED rating of its schools, either for primary or secondary school pseudo-neighbourhoods.

The second section of Chapter 6 then explored whether certain types of schools contributed more to ethnic school segregation than others. Of the three school characteristics explored, having a religious affiliation was the factor that explained the largest proportion of the variation in a school's contribution toward ethnic school segregation (H_j). In all four models, strong evidence was provided to suggest that faith schools, on averaged, contributed disproportionately to ethnic school segregation, when compared to non-religious schools for both primary and secondary schools. Later models found sufficient evidence to conclude that religious primary and secondary schools affiliated to smaller religious groups (Jewish and Muslim), or the Catholic church typically contributed more to ethnic school segregation than Church of England schools. In contrast, the models found no statistically significant difference between the average contribution of Church of England secondary schools to non-religious schools in the same pseudo-neighbourhood.

These results provide welcome clarity on the important association between ethnic school segregation and religious schools, that both expands current academic understanding and informs ongoing public debate over whether or not to expand the religious school sector in England (Long, Danechi and Loft, 2019). Namely these results suggest that increasing the number of religious schools carries a risk of encouraging higher levels of ethnic school segregation within local school neighbourhoods.

Chapter 6 also investigated whether grammar schools, or those with different OFSTED ratings, made different sized contributions towards their pseudo-neighbourhood's level of ethnic school segregation. In both cases there was no evidence to suggest they did. In other words, grammar schools seemed to contribute similar amounts towards ethnic school segregation than non-selective schools in the same pseudo-neighbourhood, as did schools with different OFSTED ratings. While the OFSTED result aligns with earlier findings, the grammar school result does not. When viewed collectively the results suggest that while pseudo-neighbourhoods with at least one grammar school tend to have higher levels of ethnic school segregation, the grammar schools themselves do not seem to contribute more to ethnic school segregation than non-selective schools located in the same pseudo-neighbourhood.[211]

[211] These contrasting findings are difficult to unpick without more targeted research into the grammar school system, as a range of possible factors could underpin these results. Qualitative research would be particularly helpful in this regard to provide a more concrete basis to validate any speculative rationale. However, one possible reason could be that the impact on the average contribution of a grammar school to the pseudo-neighbourhoods H value is spread across other local schools. In this way, pseudo-neighbourhoods with a grammar school, on average, exhibit higher H values while grammar schools themselves make similar sized contributions to their pseudo-neighbourhoods H value (H_j).

Finally, Chapter 7, the last of the results chapter, used a bespoke extract from the NPD on the application choices of year 6 students to explore the links between England's national school admissions system and patterns of ethnic school segregation, as of 2016/17. The initial analysis found that the median level of ethnic school segregation for pseudo-neighbourhoods rose between each stage of the application process, particularly between when students received their offers in stage 2 and when they started school in stage 3. The increase between these two stages alone meant that the observed patterns of ethnic school segregation for year 7 students in 2017/2018 was 6.5% higher than would have been observed if all students attended the school they were offered initially.

A set of subsequent simulations then explored what patterns of ethnic school segregation would have been likely to occur under various national school admissions systems. In all seven simulations the median level of ethnic school segregation was found to be lower than what was observed in practice, suggesting that reforming the admissions system could encourage lower levels of ethnic school segregation.

The results also provided an indication as to which type of school admissions system would be likely to produce the lowest median level of ethnic school segregation. The five simulations that kept the core feature of the school choice system, had median *H* values that were most similar to those observed in practice. For example, the simulations that allocated places at oversubscribed schools based on the applicants home-school distance, the child's academic abilities or simply at random exhibited median values between 10% - 14% lower than those observed in reality. Two later simulations, that employed academic

banding, exhibited lower levels still, with median *H* values that were 21% lower than observed in practice.

However, the two simulations that removed parental choice and instead allocated students to schools based on their home-school distance, or their home-school distance in combination with their academic abilities, exhibited the lowest levels of ethnic school segregation. These simulations had median *H* values that were 39% less than observed in practice. While the complexities of simulation studies require readers to interpret these findings cautiously, they nonetheless provide evidence of a tradeoff between the power of parents to influence the schools their children attend and the ability of governments to reduce ethnic school segregation.

Collectively, all three results chapters have provided welcome clarity on the nature of ethnic school segregation in England. The study has established that the diversification of England school age population is widespread, that children increasingly attend schools with peers from other ethnic backgrounds and that local schools have increasingly similar ethnic profiles. It also provided an insight into the areas of England where segregation is higher and the types of schools that contribute most to such trends. Finally, the results flesh out how patterns of ethnic school segregation are created each year within the school admission system, and the reforms that might help reduce ethnic school segregation in the future.

In addition to these substantive results, the thesis made two important contributions to improve how ethnic school segregation is measured. Its development of “pseudo-

neighbourhoods" provides a useful tool to future researchers by allowing patterns to be measured at a given scale within a broadly uniform sized set of areas. Secondly, the study was the first national study in England to use a multi-group index to measure ethnic school segregation, thereby capturing the "separateness" between multiple ethnic groups in a single concise measure. Hopefully this will help shift researchers away from using binary measures, that often have calculated segregation between "white" and "non-white" students, to provide a description of trends that better reflect how identities are conceived by society today.

There remains plenty of room for future work and debate. While this thesis has focused on what *is* happening in England, other researchers will need to wade into questions about what *should* be happening, or indeed *why* these patterns are observed. As England diversifies as a nation there will need to be a conversation about what type of school system England should build and researchers must actively contribute to this discussion. While the current government has committed to making schools places where students "form lasting relationships with others from different ethnic...groups" (HM Government, 2018a, p. 26), this question, along with many others, remains far from settled. This research therefore aims to be one small part of a growing body of academic work that supports the nation in its decision making, as it becomes an increasingly multi-ethnic society.

10 Bibliography

Aboud, F. E., Mendelson, M. J. and Purdy, K. T. (2003) 'Cross-race peer relations and friendship quality', *International Journal of Behavioral Development*, 27(2), pp. 165–173. doi: 10.1080/01650250244000164.

Alencar, M. S. (2015) *Information theory*. New York, [New York] (222 East 46th Street, New York, NY 10017): Momentum Press (Communications and signal processing collection).

Allen, R. (2007) 'Allocating pupils to their nearest secondary school: The consequences for social and ability stratification', *Urban Studies*, 44(4), pp. 751–770. doi: 10.1080/00420980601184737.

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

Allen, R., Burgess, S. and Windmeijer, F. (2009) *More Reliable Inference for Segregation Indices*. 09. Bristol. Available at: https://onlinelibrary.wiley.com/doi/full/10.1111/ectj.12039.

Allen, R., Coldron, J. and West, A. (2012) 'The effect of changes in published secondary school admissions on pupil composition', *Journal of Education Policy*, 27(3), pp. 349–366. doi: 10.1080/02680939.2011.604137.

Allen, R. and Vignoles, A. (2007) 'What should an index of school segregation measure?', *Oxford Review of Education*, pp. 643–668. doi: 10.1080/03054980701366306.

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

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

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

Amin, A. (2002) 'Ethnicity and the multicultural city: Living with diversity', *Environment and Planning*, 34, pp. 959–980. doi: 10.1068/a3537.

Angrist, J. D. and Lang, K. (2004) 'Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program', *The American Economic Review*, 94(5), pp. 1613–1634.

Bagley, C. (1996) 'Black and White Unite or Flight? The Racialised Dimension of Schooling and Parental Choice', *British Educational Research Journal*, 22(5), pp. 569–580.

Banerjee, A. V and Duflo, E. (2013) *Poor economics : rethinking poverty & the ways to end it*, *Poor economics : rethinking poverty and the ways to end it*. Gurgaon, India: Penguin Random House India Pvt. Ltd.

BBC (2017) *Warning over segregation in England's schools*.

Bebber, B. (2015) '"We Were Just Unwanted": Bussing, Migrant Dispersal, and South Asians in London', *Journal of Social History*, 48(3), pp. 635–661. doi: 10.1093/jsh/shul.

Benson, M., Bridge, G. and Wilson, D. (2015) 'School Choice in London and Paris - A Comparison of Middle-class Strategies', *Social Policy and Administration*, 49(1), pp. 24–43. doi: 10.1111/spol.12079.

Bifulco, R. and Ladd, H. F. (2007) 'School Choice, Racial Segregation, and Test-Score Gaps: Evidence from North Carolina's Charter School Program', *Journal of Policy Analysis and Management*, 26(1), pp. 31–56. doi: 10.1002/pam.20226.

Birtel, M. D. *et al.* (2020) 'Change in school ethnic diversity and intergroup relations: The transition from segregated elementary to mixed secondary school for majority and minority students', *European Journal of Social Psychology*, 50(1), pp. 160–176. doi: 10.1002/ejsp.2609.

Blalock, H. M. (1967) *Toward a theory of minority-group relations*. New York: Wiley.

Boliver, V., Burgess, S. and Leunig, T. (2013) *Education*, *Policy Exchange*. Available at: http://www.integrationhub.net/module/education/ (Accessed: 24 October 2018).

Bonal, X., Zancajo, A. and Scandurra, R. (2021) 'Student mobility and school segregation in an (un)controlled choice system: A counterfactual approach', *British Educational Research Journal*, 47(1), pp. 42–64. doi: 10.1002/berj.3694.

Borenstein, M. *et al.* (2009) *Introduction to Meta-Analysis*.

Burgess, S. *et al.* (2008) *Segregation and the attainment of minority ethnic pupils in England*, *CMPO Working Paper Series*.

Burgess, S. *et al.* (2011) 'Parental choice of primary school in England: What types of school do different types of family really have available to them?', *Policy Studies*, 32(5), pp. 531–547. doi: 10.1080/01442872.2011.601215.

Burgess, S. *et al.* (2015) 'What Parents Want: School Preferences and School Choice', *Economic Journal*, 125(587), pp. 1262–1289. doi: 10.1111/ecoj.12153.

Burgess, S., Greaves, E. and Vignoles, A. (2019) 'School choice in England: evidence from national

administrative data', *Oxford Review of Education*. Routledge, 45(5), pp. 690–710. doi: 10.1080/03054985.2019.1604332.
Burgess, S. and Harris, R. (2018) *(Mis-) understanding school segregation in Eng- land ? Comments on a new measure of segregation*. Available at: https://simonburgesseconomics.co.uk/mis-understanding-school-segregation-england-comments-new-measure-segregation/ (Accessed: 8 April 2021).
Burgess, S. and Platt, L. (2020) 'Inter-ethnic relations of teenagers in England's schools: the role of school and neighbourhood ethnic composition', *Journal of Ethnic and Migration Studies*. Taylor & Francis. doi: 10.1080/1369183X.2020.1717937.
Burgess, S., Vignoles, A. and Greaves, E. (2014) *Scrap school admissions policies based on postcode – they entrench inequality*, *The Conversation*. Available at: https://theconversation.com/scrap-school-admissions-policies-based-on-postcode-they-entrench-inequality-32579 (Accessed: 8 April 2021).
Burgess, S., Wilson, D. and Lupton, R. (2005) 'Parallel lives? Ethnic segregation in schools and neighbourhoods', *Urban Studies*, 42(7), pp. 1027–1056. doi: 10.1080/00420980500120741.
Burtonwood, N. (2000) 'Must Liberal Support for Separate Schools be Subject to a Condition of Individual Autonomy?', *British Journal of Educational Studies*, 48(3), pp. 269–284. doi: 10.1111/1467-8527.00147.
Cameron, D. (2015) *Extremism: PM speech*. Available at: https://www.gov.uk/government/speeches/extremism-pm-speech (Accessed: 26 March 2019).
Camina, M. M. and Iannone, P. (2014) 'Housing mix, school mix: barriers to success', *Journal of Education Policy*, 29(1), pp. 19–43. doi: 10.1080/02680939.2013.783933.
Cantle, T. (2001) *Community Cohesion: A report of The Independant Review Team*. London. Available at: https://dera.ioe.ac.uk/14146/1/communitycohesionreport.pdf.
Cantle, T. (2013) 'Segregation of schools – the impact on young people and their families and communities', in *Accord Coalition and All-Party Parliamentary Groups, February 2013*. Interculturalism Community Cohesion (iCoCo) Foundation, pp. 1–10.
Carter, B. and Fenton, S. (2010) 'Not thinking ethnicity: A critique of the ethnicity paradigm in an over-ethnicised sociology', *Journal for the Theory of Social Behaviour*, 40(1), pp. 1–18. doi: 10.1111/j.1468-5914.2009.00420.x.
Casella, G., Fienberg, S. and Olkin, I. (2015) *An Introduction to Statistical Learning with Applications in R*, *Springer*. doi: 10.1016/j.peva.2007.06.006.
Casey, L. (2016) *The Casey Review: A review into opportunity and integration*. London. doi: December 2016.
Catney, G. (2016a) *Britain is becoming more diverse , not more segregated*, *The Conversation*. Available at: https://theconversation.com/britain-is-becoming-more-diverse-not-more-segregated-68610 (Accessed: 8 April 2021).
Catney, G. (2016b) 'The Changing Geographies of Ethnic Diversity in England and Wales, 1991–2011', *Population, Space and Place*, 22(8), pp. 750–765. doi: 10.1002/psp.1954.
Catney, G. (2017) 'Towards an enhanced understanding of ethnic group geographies using measures of clustering and unevenness', *Geographical Journal*, 183(1), pp. 71–83. doi: 10.1111/geoj.12162.
Catney, G. (2018) 'The complex geographies of ethnic residential segregation: Using spatial and local measures to explore scale-dependency and spatial relationships', *Transactions of the Institute of British Geographers*, 43(1), pp. 137–152. doi: 10.1111/tran.12209.
Catney, G. and Lloyd, C. D. (2020) *Population Grids for Analysing Long-Term Change in Ethnic Diversity and Segregation*, *Spatial Demography*. Springer International Publishing. doi: 10.1007/s40980-020-00071-6.
Catney, G., Wright, R. and Ellis, M. (2020) 'The evolution and stability of multi-ethnic residential neighbourhoods in England', *Transactions of the Institute of British Geographers*, (September), pp. 1–17. doi: 10.1111/tran.12416.
Celebi, M. E., Kingravi, H. A. and Vela, P. A. (2013) 'A comparative study of efficient initialization methods for the k-means clustering algorithm', *Expert Systems With Applications*. Elsevier Ltd, 40(1), pp. 200–210. doi: 10.1016/j.eswa.2012.07.021.
Centre on Dynamics of Diversity (2012) *How has ethnic diversity grown 1991-2001-2011?* Manchester. Available at: http://www.ethnicity.ac.uk/medialibrary/briefings/dynamicsofdiversity/how-has-ethnic-diversity-grown-1991-2001-2011.pdf.
Channel 4 (2019) 'The Great British School Swap'. United Kingdom: Channel 4. Available at: https://www.channel4.com/programmes/the-great-british-school-swap.
Cheshire, P. and Sheppard, S. (2004) 'Capitalising the Value of Free Schools: The Impact of Supply Characteristics and Uncertainty', *The Economic Journal*, 114(499), pp. F397–F424.
Chessum, L. (1997) '"Sit down, you haven't reached that stage yet": African Caribbean children in Leicester schools 1960-74', *History of Education*, 26(4), pp. 409–429. doi: 10.1080/0046760970260405.

Coldron, J. *et al.* (2008) *Secondary School Admissions*. Available at: http://dera.ioe.ac.uk/8617/%5Cnhttp://dera.ioe.ac.uk/8617/1/DCSF-RR020.pdf.
Coldron, J., Cripps, C. and Shipton, L. (2010) 'Why are English secondary schools socially segregated?', *Journal of Education Policy*, 25(1), pp. 19–35. doi: 10.1080/02680930903314285.
Conover, W. J., Tercero-Gómez, V. G. and Cordero-Franco, A. E. (2017) 'The sequential normal scores transformation', *Sequential Analysis*. 2017, 36(3), pp. 397–414. doi: 10.1080/07474946.2017.1360091.
Cortese, C. F., Falk, R. F. and Cohen, J. K. (1976) 'Further Considerations on the Methodological Analysis of Segregation Indices', *American Sociological Review*, 41(4), pp. 630–637. doi: 10.2307/2094840.
Crain, R. and Mahard, R. (1978) 'Desegregation and black achievement: a review of the research', *Law and Contemporary Problems*. Edited by R. Crain, 42, pp. 17–56.
Crooks, A. T. (2010) 'Constructing and implementing an agent-based model of residential segregation through vector GIS', *International Journal of Geographical Information Science*, 24(5), pp. 661–675. doi: 10.1080/13658810903569572.
Croxford, L. and Raffe, D. (2013) 'Differentiation and social segregation of UK higher education, 1996-2010', *Oxford review of education*, 39(2), pp. 172–192. doi: 10.1080/03054985.2013.784193.
Crozier, G. and Davies, J. (2008) '"The trouble is they don't mix": Self-segregation or enforced exclusion?', *Race Ethnicity and Education*, 11(3), pp. 285–301. doi: 10.1080/13613320802291173.
Danechi, S. (2017) *Grammar school statistics*, *Commons Briefing Papers*. Available at: https://dera.ioe.ac.uk/28767/1/SN01398.pdf.
Department for Education (2014) *School Admissions Code Statutory guidance for admission authorities, governing bodies, local authorities, schools adjudicators and admission appeals panels*, *Department for Education*. United Kingdom. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/389388/School_Admissions_Code_2014_-_19_Dec.pdf.
Department for Education (2017) *House prices and schools: do houses close to the best-performing schools cost more?* London, England. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/600623/House_prices_and_schools.pdf.
Department for Education (2018a) *Schools, pupils and their characteristics: January 2018*, *Department for Education*. London. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/719226/Schools_Pupils_and_their_Characteristics_2018_Main_Text.pdf (Accessed: 26 March 2019).
Department for Education (2018b) *Schools That Work For Everyone*. London, England. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/706243/Schools_that_work_for_everyone-Government_consultation_response.pdf.
Department for Education (2018c) *Secondary and primary school applications and offers: methodology*. Available at: https://explore-education-statistics.service.gov.uk/methodology/secondary-and-primary-school-applications-and-offers-methodology (Accessed: 23 November 2020).
Duflo, E. and Banerjee, A. V (2017) *Handbook of field experiments. Volume 1*. Amsterdam: North Holland (Handbook of Economic Field Experiments Volume 1).
Duncan, O. D. and Duncan, B. (1955) 'A Methodological Analysis of Segregation Indexes', *American Sociological Review*, 20(2), pp. 210–217. doi: 10.2307/2088328.
Dwyer, C. and Parutis, V. (2013) '"Faith in the system?" State-funded faith schools in England and the contested parameters of community cohesion', *Transactions of the Institute of British Geographers*. Oxford, UK, 38(2), pp. 267–284. doi: 10.1111/j.1475-5661.2012.00518.x.
Elonga Mboyo, J. P. (2019) 'School leadership and Black and minority ethnic career prospects in England: The choice between being a group prototype or deviant head teacher', *Educational Management Administration & Leadership*. London, England, 47(1), pp. 110–128. doi: 10.1177/1741143217725326.
Esteves, O. (2018) 'Babylon by bus? The dispersal of immigrant children in England, race and urban space (1960s-1980s)', *Paedagogica Historica*. Routledge, 54(6), pp. 750–765. doi: 10.1080/00309230.2018.1521451.
van Ewijk, R. and Sleegers, P. (2010) 'Peer Ethnicity and Achievement: A Meta-Analysis into the Compositional Effect', *School Effectiveness and School Improvement*, 21(3), pp. 237–265. doi: 10.2139/ssrn.1402651.
Farley, K. and Blackman, T. (2014) 'Ethnic residential segregation stability in England, 1991–2001', *Policy and Politics*. London: Policy Press, 42(1), pp. 39–54. doi: http://dx.doi.org/10.1332/030557312X655558.
Fiel, J. (2015) 'Closing Ranks: Closure, Status Competition, and School Segregation', *American Journal of Sociology*, 121(1), pp. 126–170. doi: 10.1086/682027.
Fiel, J. E. (2013) 'Decomposing School Resegregation: Social Closure, Racial Imbalance, and Racial Isolation',

American Sociological Review, 78(5), pp. 828–848. doi: 10.1177/0003122413496252.
Finney, N. and Simpson, L. (2009) 'Population dynamics: The roles of natural change and migration in producing the ethnic mosaic', *Journal of Ethnic and Migration Studies*, 35(9), pp. 1479–1496. doi: 10.1080/13691830903125935.
Gale, R. (2013) 'Religious residential segregation and internal migration: The British Muslim case', *Environment and Planning*, 45(4), pp. 872–891. doi: 10.1068/a4515.
Gamsu, S., Donnelly, M. and Harris, R. (2018) 'The spatial dynamics of race in the transition to university: Diverse cities and White campuses in U.K. higher education', *Population, Space and Place*, (October 2018), pp. 1–15. doi: 10.1002/psp.2222.
Gibbons, S. and Telhaj, S. (2007) 'Are schools drifting apart? Intake stratification in English secondary schools', *Urban Studies*, 44(7), pp. 1281–1305. doi: 10.1080/00420980701302346.
Gibson, A. and Asthana, S. (2000a) 'Local markets and the polarization of public-sector schools in England and Wales', *Transactions of the Institute of British Geographers*, 25(3), pp. 303–319. doi: 10.1111/j.0020-2754.2000.00303.x.
Gibson, A. and Asthana, S. (2000b) 'What's in a number? Commentary on Gorard and Fitz's "Investigating the determinants of segregation between schools"', *Research Papers in Education*. Taylor & Francis, 15(2), pp. 133–153. doi: 10.1080/026715200402461.
Gokulsing, K. M. (2006) 'Without prejudice: An exploration of religious diversity, secularism and citizenship in England (with particular reference to the state funding of Muslim faith schools and multiculturalism)', *Journal of Education Policy*, 21(4), pp. 459–470. doi: 10.1080/02680930600731914.
Goldstein, H. and Noden, P. (2003) 'Modelling social segregation', *Oxford Review of Education*, 29(2), pp. 225–237. doi: 10.1080/0305498032000080693.
Goldstein, H. and Noden, P. (2004) 'A response to Gorard on social segregation', *Oxford Review of Education*, 30(3), pp. 441–442. doi: 10.1080/0305498042000260539.
Gorard, S. (2007) 'What does an index of school segregation measure? A commentary on Allen and Vignoles', *Oxford Review of Education*, 33(5), pp. 669–677. doi: 10.1080/03054980701451140.
Gorard, S. (2014) 'The link between Academies in England, pupil outcomes and local patterns of socio-economic segregation between schools', *Research Papers in Education*, 29(3), pp. 268–284. doi: 10.1080/02671522.2014.885726.
Gorard, S. (2016a) 'Damaging real lives through obstinacy: Reemphasising why significance testing is wrong', *Sociological Research Online*, 21(1), pp. 1–14. doi: 10.5153/sro.3857.
Gorard, S. (2016b) 'The complex determinants of school intake characteristics and segregation, England 1989 to 2014', *Cambridge Journal of Education*. Routledge, 46(1), pp. 131–146. doi: 10.1080/0305764X.2015.1045446.
Gorard, S. and Fitz, J. (2006) 'What Counts as Evidence in the School Choice Debate?', *British Educational Research Journal*, 32(6), pp. 797–816.
Gorard, S., Hordosy, R. and See, B. H. (2013) 'Narrowing Down the Determinants of Between-School Segregation: An Analysis of the Intake to All Schools in England, 1989-2011', *Journal of School Choice*, 7(2), pp. 182–195. doi: 10.1080/15582159.2013.791182.
Gorard, S. and Siddiqui, N. (2018) 'Grammar schools in England: a new analysis of social segregation and academic outcomes', *British Journal of Sociology of Education*. Routledge, 39(7), pp. 909–924. doi: 10.1080/01425692.2018.1443432.
Gorard, S., Taylor, C. and Fitz, J. (2001) 'Social exclusion and public policy: The relationship between local school admission arrangements and segregation by poverty', *The International Journal of Sociology and Social Policy*. Bingley: Emerald Group Publishing Limited, 21(4–6), pp. 10–36. doi: http://dx.doi.org/10.1108/01443330110789420.
Goyette, K. A., Farrie, D. and Freely, J. (2012) 'This School's Gone Downhill: Racial Change and Perceived School Quality among Whites', *Social Problems*, 59(2), pp. 155–176. doi: 10.1525/sp.2012.59.2.155.
Grannis, R. (2002) 'Discussion: Segregation indices and their functional inputs', *Sociological Methodology*, 32, pp. 69–84. doi: 10.1111/1467-9531.00111.
Green, F., Allen, R. and Jenkins, A. (2015) 'Are English free schools socially selective? A quantitative analysis', *British Educational Research Journal*, 41(6), pp. 907–924. doi: 10.1002/berj.3190.
Guryan, J. (2004) 'Desegregation and Black Dropout Rates', *The American Economic Review*, 94(4), pp. 919–943.
Hamnett, C. (2012) 'Concentration or Diffusion? The Changing Geography of Ethnic Minority Pupils in English Secondary Schools, 1999–2009', *Urban Studies*. London, England, 49(8), pp. 1741–1766. doi: 10.1177/0042098011422573.

Hamnett, C. and Butler, T. (2013) 'Distance, education and inequality', *Comparative Education*, 49(3), pp. 317–330. doi: 10.1080/03050068.2013.807165.
Hamnett, C., Butler, T. and Ramsden, M. (2013) '"I Wanted My Child to Go to a More Mixed School": Schooling and Ethnic Mix in East London', *Environment and Planning A: Economy and Space*, 45(3), pp. 553–574. doi: 10.1068/a45223.
Harris, R. (2010) *Segregation by choice ? The debate so far December 2010 Segregation by choice ? The debate so far*. Bristol, UK. Available at: https://www.bristol.ac.uk/media-library/sites/cmpo/migrated/documents/wp251.pdf.
Harris, R. (2011) *Measuring social segregation between London ' s secondary schools , 2003 – 2008 / 9*. 11. Bristol. Available at: http://www.bristol.ac.uk/media-library/sites/cmpo/migrated/documents/wp260.pdf.
Harris, R. *et al.* (2013) 'Are indices still useful for measuring socioeconomic segregation in UK schools? A response to watts', *Environment and Planning*, 45(10), pp. 2281–2289. doi: 10.1068/a130118c.
Harris, R. (2013) 'Geographies of transition and the separation of lower and higher attaining pupils in the move from primary to secondary school in London', *Transactions of the Institute of British Geographers*. Oxford, UK, 38(2), pp. 254–266. doi: 10.1111/j.1475-5661.2012.519.x.
Harris, R. (2014) *School choice , ethnic segregation and neighbourhood effects : a study of School of Geographical Sciences and Centre for Market and Public Organisation*. Bristol. Available at: http://www.bris.ac.uk/Depts/CMPO/.
Harris, R. (2016) 'Measuring segregation as a spatial optimisation problem, revisited: a case study of London, 1991–2011', *International Journal of Geographical Information Science*. Taylor & Francis, 30(3), pp. 474–493. doi: 10.1080/13658816.2015.1032973.
Harris, R., Charlton, M. and Brunsdon, C. (2018) 'Mapping the changing residential geography of white British secondary school children in England using visually balanced cartograms and hexagrams', *Journal of Maps*, 14(1), pp. 65–72. doi: 10.1080/17445647.2018.1478753.
Harris, R., Johnson, R. and Burgess, S. (2016) 'Tangled spaghetti: Modelling the core catchment areas of London's secondary schools', *Environment and Planning A*, 48(9), pp. 1681–1683. doi: 10.1177/0308518X15603987.
Harris, R. and Johnston, R. (2008) 'Primary Schools, Markets and Choice: Studying Polarization and the Core Catchment Areas of Schools', *Applied Spatial Analysis and Policy*, 1(1), pp. 59–84. doi: 10.1007/s12061-008-9002-8.
Harris, R. and Johnston, R. (2020) *Ethnic Segregreation: Between Schools Is it Increaing for Decresing in England*. Bristol, UK: Bristol University Press 2020.
Harris, R. and Owen, D. (2018) 'Implementing a Multilevel Index of Dissimilarity in R with a case study of the changing scales of residential ethnic segregation in England and Wales', *Environment and Planning B: Urban Analytics and City Science*, 45(6), pp. 1003–1021. doi: 10.1177/2399808317748328.
Hartigan, J. A. and Wong, M. A. (1979) 'A K-Means Clustering Algorithm', *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28(1), pp. 100–108. doi: 10.2307/2346830.
Hawley, W. D. and Rist, R. C. (1975) 'On The Future Implementation of School Desegregation: Some Considerations', *Law and Contemporary Problems*, 39(2), pp. 412–426. doi: 10.4324/9781351319164-15.
Hewstone, M. *et al.* (2018) 'Influence of segregation versus mixing: Intergroup contact and attitudes among White-British and Asian-British students in high schools in Oldham, England', *Theory and Research in Education*, 16(2), pp. 179–203. doi: 10.1177/1477878518779879.
Hewstone, M. and Swart, H. (2011) 'Fifty-odd years of inter-group contact: from hypothesis to integrated theory.', *The British journal of social psychology / the British Psychological Society*, 50(3), pp. 374–386. doi: 10.1111/j.2044-8309.2011.02047.x.
Higgins, J. P. T. *et al.* (2006) 'Measuring inconsistency in meta-analyses', *Journal of Intelligent Information Systems*, 27(2), pp. 159–184. doi: 10.1007/s10844-006-2974-4.
HM Government (2018a) *Integrated Communities Strategy Green Paper*, *UK Housing, Communities and Local Government*. London. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/696993/Integrated_Communities_Strategy.pdf.
HM Government (2018b) *Secondary and primary school applications and offers: March and April 2018*. London. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/715650/2018_School_Application_Offers_Text.pdf.
HM Government (2021) *National public transport access nodes*. Available at: https://www.gov.uk/government/publications/national-public-transport-access-node-schema (Accessed: 27

May 2021).
HM Land Registry (2019) 'Price Paid Data: 2019'. London: HM Land Registry. Available at: https://www.gov.uk/government/statistical-data-sets/price-paid-data-downloads#yearly-file.
Hollingworth, S. and Archer, L. (2010) 'Urban schools as urban places: School reputation, children's identities and engagement with education in London', *Urban Studies*, 47(3), pp. 584–603. doi: 10.1177/0042098009349774.
Hollingworth, S. and Williams, K. (2010) 'Multicultural mixing or middle-class reproduction? The white middle classes in London comprehensive schools', *Space and Polity*, 14(1), pp. 47–64. doi: 10.1080/13562571003737767.
Hutchens, R. (2004) 'One Measure of Segregation', *International Economic Review*, 45(2), pp. 555–578.
Ilie, S., Sutherland, A. and Vignoles, A. (2017) 'Revisiting free school meal eligibility as a proxy for pupil socio-economic deprivation', *British Educational Research Journal*, 43(2), pp. 253–274. doi: 10.1002/berj.3260.
Jackson, R. (2003) 'Should the State Fund Faith Based Schools? A Review of the Arguments', *British journal of religious education*, 25(2), pp. 89–102. doi: 10.1080/0141620030250202.
Jain, A. K. (2010) 'Data clustering: 50 years beyond K-means', *Pattern Recognition Letters*. Elsevier B.V., 31(8), pp. 651–666. doi: 10.1016/j.patrec.2009.09.011.
James, D. R. and Taeuber, K. E. (1985) 'Measures of Segregation', *Sociological Methodology*, 15, pp. 1–32.
Johnson, R. C. (2011) *Long-run Impacts of School Desegregation & School Quality on Adult Attainments*. Cambridge, Mass: National Bureau of Economic Research (Working paper series (National Bureau of Economic Research : Online)).
Johnston, R. *et al.* (2006) 'School and residential ethnic segregation: An analysis of variations across England's Local Education Authorities', *Regional Studies*, 40(9), pp. 973–990. doi: 10.1080/00343400601047390.
Johnston, R. *et al.* (2008) '"Sleep-walking towards segregation"? The changing ethnic composition of English schools, 1997-2003: An entry cohort analysis', *Transactions of the Institute of British Geographers*, 33(1), pp. 73–90. doi: 10.1111/j.1475-5661.2007.00285.x.
Johnston, R., Jones, K., *et al.* (2016) 'Macro-scale stability with micro-scale diversity: modelling changing ethnic minority residential segregation – London 2001–2011', *Transactions of the Institute of British Geographers*, 41(4), pp. 389–402. doi: 10.1111/tran.12142.
Johnston, R., Harris, R., *et al.* (2016) 'Segregation at school and at home: an English exploration', *BELGEO (Leuven)*, 2(2–3). doi: 10.4000/belgeo.18730.
Johnston, R., Manley, D. and Jones, K. (2016) 'In search of Britain's Muslim ghettoes', *Environment and Planning A*, 48(9), pp. 1684–1690. doi: 10.1177/0308518X16651873.
Johnston, R., Wilson, D. and Burgess, S. (2004) 'School segregation in multiethnic England', *Ethnicities*, 4(2), pp. 237–265.
Johnston, R., Wilson, D. and Burgess, S. (2007) 'Ethnic segregation and educational performance at secondary school in Bradford and Leicester', *Environment and Planning A*, 39(3), pp. 609–629. doi: 10.1068/a38155.
Jones, K. *et al.* (2015) 'Ethnic Residential Segregation: A Multilevel, Multigroup, Multiscale Approach Exemplified by London in 2011', *Demography*, 52(6), pp. 1995–2019. doi: 10.1007/s13524-015-0430-1.
Jones, K., Johnston, R., *et al.* (2018) 'Ethnic and class residential segregation: exploring their intersection – a multilevel analysis of ancestry and occupational class in Sydney', *Urban studies (Edinburgh, Scotland)*. London, England, 55(6), pp. 1163–1184. doi: 10.1177/0042098017730239.
Jones, K., Manley, D., *et al.* (2018) 'Modelling residential segregation as unevenness and clustering: A multilevel modelling approach incorporating spatial dependence and tackling the MAUP', *Environment and planning. B, Urban analytics and city science*. Edited by R. Harris and R. Johnston. London, England, 45(6), pp. 1122–1141. doi: 10.1177/2399808318782703.
Kam Ho, T. (1995) 'Random Decision Forests', *AT&T Bell Laboratories*.
Kaszycka, K. A. and Strzałko, J. (2019) 'Race and perception of human diversity: Five European case studies', *Journal of Anthropological Sciences*, 97, pp. 173–187. doi: 10.4436/jass.97006.
Khmelkov, V. T. and Hallinan, M. T. (1999) 'Organizational Effects on Race Relations in Schools', *Journal of Social Issues*, 55(4), pp. 627–645.
Kisfalusi, D. (2016) 'The quality of inter- and intra-ethnic friendships among roma and non-roma students in Hungary', *Corvinus Journal of Sociology and Social Policy*, 7(1), pp. 3–26. doi: 10.14267/cjssp.2016.01.01.
Leckie, G. *et al.* (2012) 'Multilevel Modeling of Social Segregation', *Journal of Educational and Behavioral Statistics*, 37(1), pp. 3–30. doi: 10.3102/1076998610394367.
Leckie, G. and Goldstein, H. (2015) 'A multilevel modelling approach to measuring changing patterns of ethnic composition and segregation among London secondary schools, 2001-2010', *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 178(2), pp. 405–424. doi: 10.1111/rssa.12066.

Lewis Mumford Center (2001) *Ethnic Diversity Grows, Neighborhood Integration Lags Behind Report*. Albany, USA. doi: 10.3389/frym.2018.00038.
Liaw, A. and Wiener, M. (2002) 'Classification and Regression by randomForest', *R News*, 2(3), pp. 18–22.
Lieberson, S. and Carter, D. K. (1982) 'Temporal Changes and Urban Differences in Residential Segregation: A Reconsideration', *American Journal of Sociology*, 88(2), pp. 296–310.
Long, R., Danechi, S. and Loft, P. (2019) *Faith Schools in England: FAQs*. Available at: http://researchbriefings.files.parliament.uk/documents/SN06972/SN06972.pdf.
Long, R., Foster, D. and Robert, N. (2018) *Grammar schools in England*. London. Available at: https://commonslibrary.parliament.uk/research-briefings/sn07070/.
Machin, S. and West, A. (2013) 'Admissions to Academy Schools in England : School Composition and House Prices', (August), pp. 0–30.
Maechler, M. *et al.* (2020) 'robustbase: Basic Robust Statistics'. R package.
Mahan, T. W. and Mahan, A. M. (1971) 'The impact of schools on learning: Inner-city children in suburban schools', *Journal of School Psychology*, 9(1), pp. 1–11. doi: 10.1016/0022-4405(71)90058-6.
Mahdavi Ardestani, B., O'Sullivan, D. and Davis, P. (2018) 'A multi-scaled agent-based model of residential segregation applied to a real metropolitan area', *Computers, Environment and Urban Systems*. Elsevier, 69, pp. 1–16. doi: 10.1016/j.compenvurbsys.2017.11.002.
Manley, D. *et al.* (2015) 'Macro-, Meso- and Microscale Segregation: Modeling Changing Ethnic Residential Patterns in Auckland, New Zealand, 2001–2013', *Annals of the Association of American Geographers*, 105(5), pp. 951–967. doi: 10.1080/00045608.2015.1066739.
Massey, D. S. and Denton, N. (1988) 'The Dimensions of Residential Segregation', *Social Forces*, 67(2), pp. 281–315.
May, T. (2016) *Britain, the great meritocracy: Prime Minister's speech*, *HM Government*. Available at: https://www.gov.uk/government/speeches/britain-the-great-meritocracy-prime-ministers-speech (Accessed: 26 March 2019).
McKeown, S. *et al.* (2012) 'Micro-ecological behavior and intergroup contact', *Journal of Social Psychology*, 152(3), pp. 340–358. doi: 10.1080/00224545.2011.614647.
Mcpherson, M., Smith-lovin, L. and Cook, J. M. (2001) 'Birds of a Feather: Homophily in Social Networks', *Annual Review of Sociology*, 27, pp. 415–444.
Merry, M. S. (2015) 'The conundrum of religious schools in twenty-first-century Europe', *Comparative Education*. doi: 10.1080/03050068.2014.935582.
Miah, S. (2012) 'School desegregation and the politics of 'forced integration', *Race and Class*. Springer Netherlands, 54(2), pp. 26–38. doi: 10.1177/0306396812454987.
Miah, S. (2016) 'The Muslim problematic: Muslims, state schools and security', *International Studies in Sociology of Education*. Routledge, 26(2), pp. 138–150. doi: 10.1080/09620214.2016.1200996.
Mickelson, R. A. (2001) 'Subverting Swann: First- and Second-Generation Segregation in the Charlotte-Mecklenburg Schools', *American Educational Research Journal*, 38(2), pp. 215–252.
Mitchell, P. (2017) *A student centered analysis of ethnic segregation in London's schools*. London. Available at: http://www.civitas.org.uk/content/files/astudentcenteredanalysisofsegregation.pdf.
Mora, R. and Castillo, J. R. (2011) 'Entropy- Based Segregation Indices', *Sociological Methodology*, 41, pp. 159–194.
Munniksma, A. *et al.* (2013) 'Extended intergroup friendships within social settings: The moderating role of initial outgroup attitudes', *Group Processes and Intergroup Relations*, 16(6), pp. 752–770. doi: 10.1177/1368430213486207.
Noden, P. *et al.* (1998) 'Choices and destinations at transfer to secondary schools in london', *Journal of Education Policy*, 13(2), pp. 221–236. doi: 10.1080/0268093980130204.
Noreisch, K. (2007) 'School catchment area evasion: The case of Berlin, Germany', *Journal of Education Policy*, 22(1), pp. 69–90. doi: 10.1080/02680930601065759.
OECD (2010) 'What school choices are available and what measures do countries use to promote or restrict school choice?', in, pp. 418–432. doi: 10.1787/eag-2010-29-en.
Office for National Statistics (2012) *Ethnicity and National Identity in England and Wales: 2011*, *Office of National Statistics*. London. doi: 10.1093/nq/s7-IV.96.349-a.
Office for National Statistics (2016) 'Lower layer Super Output Area population density'. London: Office for National Statistics.
Office for National Statistics (2018) *Births in England and Wales: 2017*, *Population Trends*. London. doi: 10.1057/pt.2009.51.
Office for National Statistics (2020) *Household wealth by ethnicity, Great Britain: April 2016 to March 2018*.

Available at:
https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/articles/householdwealthbyethnicitygreatbritain/april2016tomarch2018#:~:text=Only the White British and,across other ethnic groups differs. (Accessed: 21 May 2021).
Office for National Statistics (2021) *Schools, pupils and their characteristics*. Available at: https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics (Accessed: 3 January 2022).
Olzak, S., Shanahan, S. and West, E. (1994) 'School Desegregation, Interracial Exposure, and Antibusing Activity in Contemporary Urban America', *American Journal of Sociology*, 100(1), pp. 196–241.
Openshaw, S. (1984) 'Ecological Fallacies and the Analysis of Areal Census Data', *Environment and Planning A*. London, England, 16(1), pp. 17–31. doi: 10.1068/a160017.
Orfield, G. and Frankenberg, E. (2001) *Experiencing Integration in Louisville: How Parents and Students See the Gains and Challenges*. Louseville.
Östh, J., Malmberg, B. and Andersson, E. K. (2015) 'Analysing segregation using individualised neighbourhoods', in Lloyd, C. D., Shuttleworth, I. G., and Wong, D. W. (eds) *Social-spatial segregation*. Online. Policy Press Scholarship Online, pp. 1–18. doi: 10.1332/policypress/9781447301356.001.0001.
Ouseley, H. G. (2001) *Oldham Independant Review*, *Independent Review*. Oldham.
Peach, C. (1975) *Urban social segregation*. London: Longman.
Peach, C. (1996) 'Good segregation, bad segregation', *Planning Perspectives*, 11(4), pp. 379–398. doi: 10.1080/026654396364817.
Peach, C. (2009) 'Slippery segregation: Discovering or manufacturing ghettos?', *Journal of Ethnic and Migration Studies*, 35(9), pp. 1381–1395. doi: 10.1080/13691830903125885.
Pennell, H., West, A. and Hind, A. (2006) *Secondary school admissions*. London.
Pettigrew, T. F. (1997) 'Generalized Intergroup Contact Effects on Prejudice', *Personality and Social Psychology Bulletin*. Thousand Oaks, CA, 23(2), pp. 173–185. doi: 10.1177/0146167297232006.
Pettigrew, T. F. *et al.* (2011) 'Cross-Group Friendships and Intergroup Attitudes: A Meta-Analytic Review', *Personality and Social Psychology Review*, 15(4), pp. 332–351. doi: 10.1177/1088868311411103.
Pettigrew, T. F. and Tropp, L. R. (2006) 'A meta-analytic test of intergroup contact theory', *Journal of Personality and Social Psychology*, 90(5), pp. 751–783. doi: 10.1037/0022-3514.90.5.751.
Pettigrew, T. F. and Tropp, L. R. (2008) 'How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators', *European Journal of Social Psychology*, 38, pp. 922–934. doi: 10.1002/ejsp.
Phillips, D. (2006) 'Parallel lives? Challenging discourses of British Muslim self-segregation', *Environment and Planning D: Society and Space*, 24(1), pp. 25–40. doi: 10.1068/d60j.
Phillips, D. (2007) 'Ethnic and Racial Segregation: A Critical Perspective', *Geography Compass*, 1(5), pp. 1138–1159. doi: 10.1111/j.1749-8198.2007.00051.x.
Phillips, D. (2015) 'Claiming spaces: British Muslim negotiations of urban citizenship in an era of new migration', *Transactions of the Institute of British Geographers*, 40(1), pp. 62–74. doi: 10.1111/tran.12054.
Phillips, D. and Harrison, M. (2010) 'Constructing an integrated society: Historical lessons for tackling black and minority ethnic housing segregation in Britain', *Housing Studies*, 25(2), pp. 221–235. doi: 10.1080/02673030903561842.
Phillips, T. (2013) 'The story so far and why it matters', in Goodhart, D. (ed.) *Mapping Integration*. London: Demo, pp. 9–15.
Pring, R. (2007) 'The common school', *Journal of Philosophy of Education*, 41(4), pp. 503–522. doi: 10.1111/j.1467-9752.2007.00588.x.
Putnam, R. D. (2007) 'E Pluribus Unum : Diversity and Community in the Twenty-first Century The 2006 Johan Skytte Prize Lecture', *Scandinavian Political Studies*. Oxford, UK, 30(2), pp. 137–174. doi: 10.1111/j.1467-9477.2007.00176.x.
Qi, J. *et al.* (2017) 'An effective and efficient hierarchical K-means clustering algorithm', *International Journal of Distributed Sensor Networks*, 13(8), pp. 1–17. doi: 10.1177/1550147717728627.
Qureshi, Y. F. and Janmaat, J. G. (2014) 'Diverging or converging trends: An investigation of education policies concerning the incorporation of ethnic minority children in England, France and Germany', *Compare*. Routledge, 44(5), pp. 710–731. doi: 10.1080/03057925.2013.800782.
R Core Team (2019) 'R: A Language and Environment for Statistical Computing'. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.r-project.org/.
Rathelot, R. (2012) 'Measuring Segregation When Units are Small: A Parametric Approach', *Journal of Business and Economic Statistics*, 30(4), pp. 546–553. doi: 10.1080/07350015.2012.707586.
Reardon, S. F. *et al.* (2012) 'Brown Fades : The End of Desegregation and the Resegregation of American Public

Schools', *Journal of Policy Analysis and Management*, 31(4), pp. 876–904. doi: 10.1002/pam.
Reardon, S. F. and Firebaugh, G. (2002) 'Measures of Multigroup Segregation', *Sociological Methodology*, 32, pp. 33–67. Available at: http://www.ingentaconnect.com/content/bpl/some/2002/00000032/00000001/art00110.
Reardon, S. F. and O'Sullivan, D. (2004) 'Measures of Spatial Segregation', *Sociological Methodology*, 34, pp. 121–162.
Reardon, S. F., Yun, J. T. and Eitle, T. M. (2000) 'The Changing Structure of School Segregation: Measurement and Evidence of Multiracial Metropolitan-Area School Segregation, 1989-1995', *Demography*, 37(3), pp. 351–364. doi: 10.2307/2648047.
Reardon, S. and Ownes, A. (2014) '60 Years after Brown: Trends and Consequences of School Segregation', *The Annual Review ofSociology*, 40, pp. 199–218. doi: 10.1146/annurev-soc-071913-043152.
Reber, S. J. (2010) 'School Desegregation and Educational Attainment for Blacks', *The Journal of Human Resources*, 45(4), pp. 893–914.
Richards, M. P. (2014) 'The Gerrymandering of School Attendance Zones and the Segregation of Public Schools: A Geospatial Analysis', *American Educational Research Journal*, 51(6), pp. 1119–1157. doi: 10.3102/0002831214553652.
Rivkin, S. G. (2000) 'School Desegregation, Academic Attainment, and Earnings', *The Journal of Human Resources*, 35(2), pp. 333–346.
Roberts, N. (2017) *School places in England: applications, allocations and appeals*. Available at: https://commonslibrary.parliament.uk/research-briefings/sn07147/.
Robinson, V. (1980) 'Lieberson's Isolation Index; A Case Study Evaluation', *The Royal Geographical Society*, 12(4), pp. 307–312.
Rude, J. and Herda, D. (2010) 'Best Friends Forever? Race and the Stability of Adolescent Friendships', *Social Forces*, 89(2), pp. 585–607.
Samad, Y. (2013) 'Community cohesion without parallel lives in Bradford', *Patterns of Prejudice*, 47(3), pp. 269–287. doi: 10.1080/0031322X.2013.812343.
Sandford, M. (2018) 'Local Government in England: structures', *House of Commons Library*, (07104), pp. 1–20. doi: 10.2307/2139473.
Shankley, W., Hannemann, T. and Simpson, L. (2020) 'The Demography of Ethnic Minorities', in Byrne, B. et al. (eds) *Ethnicity and Race in the UK*, pp. 15–34. doi: 10.1017/S0021932000024895.
Shook, N. J. and Fazio, R. H. (2008) 'Roommate relationships: A comparison of interracial and same-race living situations', *Group Processes and Intergroup Relations*, 11(4), pp. 425–437. doi: 10.1177/1368430208095398.
Siegel-hawley, B. G. (2012) 'Research Brief How Non-Minority Students Also Benefit', in *The National Coalition on School Diversity*. National Coalition on School Diversity, pp. 1–8. Available at: www.school-diversity.org.
Simpson, L. and Finney, N. (2006) 'Do ethnic minorities exclude themselves?', *Radical Statistics*, (99), pp. 34–45.
Simpson, L. and Finney, N. (2009) 'Ethnic ghettos in Britain: A fact or a myth?', *Significance*, 6(2), pp. 72–75. doi: 10.1111/j.1740-9713.2009.00353.x.
Van Der Slik, F. W. P., Driessen, G. W. J. M. and De Bot, K. L. J. (2006) 'Ethnic and socioeconomic class composition and language proficiency: A longitudinal multilevel examination in Dutch elementary schools', *European Sociological Review*, 22(3), pp. 293–308. doi: 10.1093/esr/jci058.
Söderström, M. and Uusitalo, R. (2010) 'School Choice and Segregation: Evidence from an Admission Reform', *The Scandinavian Journal of Economics*, 112(1), pp. 55–76.
Stevens, D., Parkes, J. and Shun-Kai, C. (2019) *ISC Census and Annual Report 2019*. Available at: https://www.isc.co.uk/media/4890/isc_census_2018_report.pdf.
Strand, S. (2015) *Ethnicity, deprivation and educational achievement at age 16 in England: trends over time*. London. doi: 10.1080/00071005.2015.1076566.
Subramanian, S. V., Duncan, C. and Jones, K. (2001) 'Multilevel perspectives on modeling census data', *Environment and Planning*, 33(3), pp. 399–417. doi: 10.1068/a3357.
Sykes, B. and Kuyper, H. (2013) 'School Segregation and the Secondary-School Achievements of Youth in the Netherlands', *Journal of Ethnic and Migration Studies*. Taylor & Francis, 39(10), pp. 1699–1716. doi: 10.1080/1369183X.2013.833707.
Taylor, C. (2001) 'Hierarchies and "local" markets: The geography of the "lived" market place in secondary education provision', *Journal of Education Policy*, 16(3), pp. 197–214. doi: 10.1080/02680930110041024.
Taylor, C. (2009) 'Choice, Competition, and Segregation in a United Kingdom Urban Education Market', *American Journal of Education*, 115(4), pp. 549–568. doi: 10.1086/599781.
Taylor, C. (2018) 'The Reliability of Free School Meal Eligibility as a Measure of Socio-Economic Disadvantage:

Evidence from the Millennium Cohort Study in Wales', *British Journal of Educational Studies*. Routledge, 66(1), pp. 29–51. doi: 10.1080/00071005.2017.1330464.
Taylor, C., Gorard, S. and Fitz, J. (2003) 'The modifiable areal unit problem: Segregation between schools and levels of analysis', *International Journal of Social Research Methodology: Theory and Practice*, 6(1), pp. 41–60. doi: 10.1080/13645570305053.
The Challenge, School Dash and The iCoCo Foundation (2016) *Understanding Segregation in England: 2011-2016*. London. Available at: https://the-challenge.org/uploads/documents/TCN-Understanding-School-Segregation-in-England-2011-to-2016.pdf.
The Guardian (2017) *Quarter of English state primary schools are 'ethnically segregated*. Available at: https://www.theguardian.com/education/2017/mar/22/english-state-primary-schools-ethnically-segregated-white-british-children (Accessed: 28 April 2020).
The Independent (2017) *Faith schools and grammars the most ethnically and socially segregated in UK, warns charity report*. Available at: https://www.independent.co.uk/news/education/education-news/faith-schools-segregated-ethnic-social-most-uk-grammar-education-charity-schooldash-challenge-icoco-a7643621.html (Accessed: 28 April 2020).
Theil, H. and Finizza, A. J. (1971) 'A note on the measurement of racial integration of schools by means of informational concepts', *The Journal of Mathematical Sociology*, 1(2), pp. 187–193. doi: 10.1080/0022250X.1971.9989795.
Tower Hamlets (2020) *Frequently Asked Questions on Pupil Ability Banding What is Banding?* Available at: https://www.towerhamlets.gov.uk/Documents/Education-and-skills/Admissions-and-exclusions/FAQ-on-Pupil-Ability-Banding.pdf (Accessed: 24 September 2020).
UCLA Institute for Digital Research and Education (2020) *FAQ HOW DO I INTERPRET A REGRESSION MODEL WHEN SOME VARIABLES ARE LOG TRANSFORMED?* Available at: https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faqhow-do-i-interpret-a-regression-model-when-some-variables-are-log-transformed/ (Accessed: 11 September 2020).
Vowden, K. J. (2012) 'Safety in numbers? Middle-class parents and social mix in London primary schools', *Journal of Education Policy*, 27(6), pp. 731–745. doi: 10.1080/02680939.2012.664286.
Warikoo, N. K. (2010) 'Symbolic Boundaries and School Structure in New York and London Schools', *American Journal of Education*, 116(3), pp. 423–451. doi: 10.1086/651415.
Weldon, M. (2016) *Segregation in Lancashire : Evidence from structural models of two-sided matching*. Lancaster University. Available at: https://eprints.lancs.ac.uk/id/eprint/86791/1/2016WeldonPhD.pdf.
West, A. (2005) '"Banding" and secondary school admissions: 1972-2004', *British Journal of Educational Studies*, 53(1), pp. 19–33. doi: 10.1111/j.1467-8527.2005.00281.x.
West, A. and Hind, A. (2006) 'Selectivity, admissions and intakes to "comprehensive" schools in London, England', *Educational Studies*, 32(2), pp. 145–155. doi: 10.1080/03055690600631036.
West, A. and Hind, A. (2016) *Secondary school admissions in London 2001 to 2015: compliance, complexity and control*. London, England. Available at: http://eprints.lse.ac.uk/66368/.
White, M. J. (1986) 'Segregation and Diversity Measures in Population Distribution', *Population Index*, 52(2), pp. 198–221.
Wilson, D. and Bridge, G. (2019) 'School choice and the city: Geographies of allocation and segregation', *Urban Studies*, 56(15), pp. 3198–3215. doi: 10.1177/0042098019843481.
Wong, D. W. S. and Shaw, S. L. (2011) 'Measuring segregation: An activity space approach', *Journal of Geographical Systems*, 13(2), pp. 127–145. doi: 10.1007/s10109-010-0112-x.
Yu, S. S. *et al.* (2018) 'Two improved k-means algorithms', *Applied Soft Computing Journal*. Elsevier B.V., 68, pp. 747–755. doi: 10.1016/j.asoc.2017.08.032.
Zhang, G., Zhang, C. and Zhang, H. (2018) 'Improved K-means algorithm based on density Canopy', *Knowledge-Based Systems*. Elsevier B.V., 145, pp. 289–297. doi: 10.1016/j.knosys.2018.01.031.
Zoloth, B. S. (1976) 'Alternative Measures of School Segregation', *Land Economics*, 52(3), pp. 278–298. doi: 10.2307/3145527.

11 Appendix

11.1 How this study conducted its Literature Review

The thesis conducted two searches of the academic literature. In September 2017, at the start of the project, I identified articles using a literature search, with the help of the university's library team. These 200 or so articles were read, and their bibliographies used as a way of identifying other relevant research. Unfortunately, I did not record the search terms that were used in this initial search, nor the exact number of studies it identified.

A second literature search was conducted on 1st April 2021 to offer a more systematic review of the literature that helped to provide clarity for the reader and ensured that I did not overlook studies that had been released since 2017. With the help of the department's library team, I searched four databases; the British Education Index, ProQuest Social Science Premium Collection, SCOPUS, and the Web of Science using the search terms outlined in Table ***30***. In addition, I created a list of academics that had previously published in this field and reviewed their university, google scholar or sematic scholar pages to find any relevant articles that had not yet been identified. This secondary search was conducted on the 7th and 8th of April 2021, as outlined in Table 31 below.

Table 30: Describing the terms used in the literature search conducted on 1st April 2021

Name	Search terms	Number of results	Number downloaded

British Education Index	(Ethnic OR minority OR black OR white OR Asian OR race OR racial OR Majority OR Minority OR Black OR White OR "all white" OR "all black" OR "all Asian" OR Asian OR Indian OR Pakistani OR Caribbean OR African OR Bangladeshi OR "African American" OR "Asian American") AND (Segregat* OR Cluster* OR Divi* OR Separat*) AND (School*) AND (UK OR “United Kingdom” OR England)	226	26
ProQuest Social Science Premium Collection	ab(Ethnic OR minority OR black OR white OR Asian OR race OR racial OR Majority OR Minority OR Black OR White OR "all white" OR "all black" OR "all Asian" OR Asian OR Indian OR Pakistani OR Caribbean OR African OR Bangladeshi OR "African American" OR "Asian American") AND ti(Segregat* OR Cluster* OR Divi* OR Separat*) AND ab(School*) AND ab(UK OR “United Kingdom” OR England)	50	9
SCOPUS	TITLE-ABS-KEY ((ethnic OR minority OR black OR white OR Asian OR race OR racial OR majority OR minority OR black OR white OR "all white" OR "all black" OR "all Asian" OR Asian OR Indian OR Pakistani OR Caribbean OR African OR Bangladeshi OR "African American" OR "Asian American") AND (segregat* OR cluster* OR divi* OR separat*) AND (school*) AND (uk OR “united AND kingdom” OR England) 268 Documents 20 downloaded	168	8
Web of Science	B=(Ethnic OR minority OR black OR white OR Asian OR race OR racial OR Majority OR Minority OR Black OR White OR "all white" OR "all black" OR "all Asian" OR Asian OR Indian OR Pakistani OR Caribbean OR African OR Bangladeshi OR "African American" OR "Asian American") AND AB=(Segregat* OR Cluster* OR Divi* OR	125	7

	Separat*) AND AB=(School*) AND AB= (UK OR "United Kingdom" OR England)		

Table 31: List of academics who were included in the literature search

Academics Name	**Source**
Rebecca Allen	Google Scholar
Ron Johnston	Google Scholar & Semantic Scholar
Ellen Greaves	Workplace Webpage (IFS)
Anna Vignoles	Oxford Solo & Semantic scholar
Ted Cantle	Personal Website
Debroah Phillips, Ceri Peach, Audrey Hind, Philip Noden, Chris Hamnett, Rick Grannis	Semantic scholar
Stephen Gorard, Beng Huat See, Rita Hordosy, Simon Burgess, David J Manley, Kelvyn Jones, Richard Harris, Yiming Wang, Dewi A Owen, George Leckie, Harvey Goldstein, Shelley McKeown Jones, Deborah Wilson, Gemma Catney, Anne West, John Coldron, Chris Taylor, John Fitz, Ruth Lupton, Asthana Sheena, Alex Gibson, Shamin Miah, Sean Reardon, Glenn Firebaugh, Jeremy Fiel, Ludi Simpson	University page

11.2 Modelling Segregation

While not discussed at length in this thesis, a group of studies have argued that using parametric models offers a superior way of measuring evenness segregation over traditional segregation indexes (Goldstein and Noden, 2003; Leckie *et al.*, 2012; Leckie and Goldstein, 2015). These build off of a larger body of work that use multilevel models to measure

residential segregation (Manley *et al.*, 2015; Johnston, Jones, *et al.*, 2016; Harris and Owen, 2018; Jones, Johnston, *et al.*, 2018; Jones, Manley, *et al.*, 2018).

The most relevant study to this thesis, which measured ethnic school segregation, was written by Leckie and Goldstein (2015). They used multilevel multinomial regression models to explore patterns of ethnic school segregation for White, Black, and Asian students attending state-funded secondary schools in London between 2001-2010. They argued that the model's variance parameter provided a good estimate for the underlying degree of segregation between groups. They argued that their approach provided a measure of segregation which better accounts for "random" segregation, offers a way of measuring segregation simultaneously at different geographic scales of analysis (Subramanian, Duncan and Jones, 2001) and, by adding covariates into their models, could offer an easier way to examine the predictors of segregation.

Ultimately, this thesis did not use these models as it felt that the additional complexity it brought did not provide a sufficient improvement to the research, to make it worthwhile. As shown in their analysis, Leckie and Goldstein's substantive findings were broadly similar to results calculated using the multigroup *H* index. It also remains debatable as to whether working at various geographic scales of analysis is advantageous when measuring *school* segregation. Unlike residential segregation, the practicalities of travelling to and from school each day mean that there is a more "natural" geographic area for measuring school segregation (i.e., the local educational markets). While there remains a debate to be had over how best to practically define educational markets (Gibson and Asthana, 2000a; Harris and Johnston, 2008) the article did not provide any explanation as to why it was useful to

investigate patterns of *evenness* school segregation at a larger or smaller geographic scales. Lastly, as highlighted by Gorard and Fitz, using a model's variance parameter as the measure for school segregation remains contested (2006). The variance parameter could describe the level of variance in ethnic groupings across schools, but it also could describe the extent to which the parametric model fits the structure of the data. This means that you could lower the measure of segregation by improving the model's specification – which seems counter-intuitive.

More broadly, as ethnic school segregation is already a methodologically dense academic field, changing existing research practices to add complexity should only be done when it is clearly warranted. After 50 years of debate, segregation indexes are well researched and, following a series of rigorous evaluations (James and Taeuber, 1985; Reardon and Firebaugh, 2002; Allen and Vignoles, 2007), it is relatively straightforward for researchers to select an index which matches their intended purpose. Even if they did choose the "wrong" segregation index, the results are likely to be highly correlated with results they would have calculated if they had chosen the "right" one. Sticking with these conventional segregation indexes therefore seems sensible.

Using a conventional approach to measuring segregation also provided greater freedom to adopt more nuanced ways of defining the geographic area at which segregation is calculated. Given that the results from this study's simulation analysis showed that this choice had a far greater impact on the subsequent findings of the study, this seemed like a more appropriate area to introduce complexity.

11.3 Evidence on the association between racial or ethnic school segregation and academic performance

There is partial evidence to suggest that ethnic school segregation is associated with lower academic achievement for minority groups, though it remains unlikely that a strong association exists within the English school system. The strongest evidence that reducing ethnic/racial school segregation raises the educational performance of students comes from the USA, where over 100 studies have used longitudinal data to explored how the academic performance of African American students changed following the end of formal segregation (Crain and Mahard, 1978). In a metanalysis of 73 of these studies, 55% of research found a that academic performance for African American students rose following the introduction of desegregation orders, while only 16% found that performance fell (Crain and Mahard, 1978).

While a "effect size" is hard to calculate, due to the plurality of methodologies used across these studies, one piece of research that employed a methodology closely resembling a Randomised Control Trial gives some idea of a potential effect size. Conducted in 1971, researchers compared the change in IQ scores of 196 Black or Hispanic students who were randomly assigned either to an urban "segregated" or suburban "desegregated" school in Hartford between 1966 – 1967 (Mahan and Mahan, 1971). The authors found that students attending the suburban "desegregated" schools improved their IQ by 5.12 percentage points higher, on average, than students attending the urban "segregated" schools. While this study was relatively small, often dated and failed to control for differences in school quality, more recent studies have highlighted other educational benefits of school

desegregation in America. These include a "reduced high school dropout rate of blacks by two to three percentage points" (Guryan, 2004, p. 919) and an increase in per-pupil spending on black students by as much as "22.5 percent… [or] about $1,300" (Johnson, 2011). Given these types of positive findings, most academics in the US lean towards the idea that formal desegregation led to a small but positive impact on the test scores of African American students (Reardon and Ownes, 2014).

It would, however, be naïve to assume that results collected in 1960s America could be easily applied to understanding modern day England. Aside from the obvious wider societal differences, there are a number of important dissimilarities within the education system. For one, a number of American academics have argued that the educational gains experienced by African American students after desegregation orders were introduce resulted from an equalisation of school funding between school districts which educated white vs black students (Reber, 2010; Johnson, 2011) As these types of funding disparities do not exist in England, promoting greater ethnic integration is likely to have a more limited impact on academic performance. Furthermore, unlike in the USA, many ethnic minority groups in England have higher average academic performance than their White British counterparts (Strand, 2015)

Instead, a smaller body of European research provides a better insight into the most probable relationship, which tends to find a fairly weak association between the ethnic composition of schools and academic performance.[212] A metanalysis of 34 studies based on

[212] These studies have explored the association between student's academic achievement and the percentage of ethnic minority students in the school or classroom. While this "ethnic composition" is not the same thing

cross-sectional data, for example, found that a positive 10 percentage point difference in the number of ethnic minority students attending a school was associated with a test score that was 0.018 standard deviations lower for ethnic minority students ($p < 0.01$) (van Ewijk and Sleegers, 2010). While statistically significant, the effect size remains small. Other European studies found even weaker relationships. A study, using longitudinal data on 60,000 Dutch elementary school students, found that any relationship between the ethnic composition of a classroom and the students' test scores became non-significant after controlling for the socioeconomic status of the child's parent (Van Der Slik, Driessen and De Bot, 2006). Another longitudinal study, based on 19,200 secondary school students, found similar findings (Sykes and Kuyper, 2013). One of the few British studies on this topic found equally mixed results which led them to conclude that "no general conclusions should be drawn regarding the posited link between educational outcomes, ethnic background, and school composition" (Johnston, Wilson and Burgess, 2007, p. 609).

as ethnic segregation, there is likely to be a sufficient overlap that these findings provide the best insight currently available into any possible association.

11.4 Information on the samples used in different years of the study.

The table below provides information on Sample 1 and 2 for all the years in which they were used. Sample 3 and 4 are not described in this section as neither was used for any longitudinal analysis.

Table 32: Describing the number of students, schools, and pseudo-neighbourhoods in Sample 1 and 2 between 2006 – 2019

Sample	Year	Phase	Number of Students	Number of Schools	Number of Pseudo-neighbourhoods
1	2006	Primary school	3,341,556	17,437	2763
		Secondary school	3,401,262	3581	536
		Total	6,742,818	21,018	3299
	2007	Primary school	3,292,376	17,285	2762
		Secondary school	3,377,756	3576	536
		Total	6,670,132	20,861	3298
	2008	Primary school	3,261,328	17,125	2762
		Secondary school	3,360,819	3567	536
		Total	6,622,147	20,692	3298
	2009	Primary school	3,221,653	16,987	2762
		Secondary school	3,336,347	3548	536
		Total	6,558,001	20,535	3298
	2010	Primary school	3,222,404	16,883	2760
		Secondary school	3,316,565	3521	536
		Total	6,538,969	20,404	3296

	2011	Primary school	3,246,636	16,791	2759
		Secondary school	3,331,581	3506	536
		Total	6,578,217	20,297	3294
	2012	Primary school	3,296,891	16,721	2758
		Secondary school	3,284,616	3439	536
		Total	6,581,508	20,160	3294
	2013	Primary school	3,371,021	16,681	2758
		Secondary school	3,236,058	3402	535
		Total	6,607,079	20,083	3294
	2014	Primary school	3,476,223	16,663	2759
		Secondary school	3,210,273	3432	536
		Total	6,686,496	20,095	3295
	2015	Primary school	3,558,203	16,645	2758
		Secondary school	3,199,500	3444	536
		Total	6,757,702	20,089	3294
	2016	Primary school	3,641,572	16,645	2757
		Secondary school	3,201,342	3439	536
		Total	6,842,914	20,084	3293
	2017	Primary school	3,721,378	16,668	2758
		Secondary school	3,230,359	3434	535
		Total	6,951,736	20,101	3293
	2018	Primary school	4,701,170	16,681	2758
		Secondary school	3,291,091	3433	535
		Total	7,992,261	20,115	3293
	2019	Primary school	4,711,393	16,673	2758
		Secondary school	3,357,616	3438	535
		Total	8,069,010	20,111	3293

2	2006	Primary school	1,768,999	7,861	1133
		Secondary school	1,883,412	1897	272
		Total	3,652,410	9,759	1405
	2007	Primary school	1,806,559	8,022	1153
		Secondary school	1,899,294	1920	272
		Total	3,705,852	9,942	1425
	2008	Primary school	1,898,522	8,449	1220
		Secondary school	2,023,328	2061	290
		Total	3,921,850	10,510	1511
	2009	Primary school	1,913,465	8562	1254
		Secondary school	1,994,679	2040	286
		Total	3,908,144	10,601	1540
	2010	Primary school	1,953,981	8667	1271
		Secondary school	2,053,816	2091	296
		Total	4,007,797	10,757	1566
	2011	Primary school	2,021,434	8843	1307
		Secondary school	2,077,774	2100	299
		Total	4,099,207	10,943	1606
	2012	Primary school	2,131,585	9185	1365
		Secondary school	2,106,449	2113	305
		Total	4,238,033	11,298	1670
	2013	Primary school	2,264,223	9573	1433
		Secondary school	2,161,421	2190	317
		Total	4,425,644	11,764	1750
	2014	Primary school	2,413,465	9951	1496
		Secondary school	2,235,671	2301	329
		Total	4,649,136	12,252	1826

	2015	Primary school	2,547,554	10,282	1546
		Secondary school	2,293,401	2382	341
		Total	4,840,955	12,664	1888
	2016	Primary school	2,704,113	10,770	1625
		Secondary school	2,385,265	2476	355
		Total	5,089,379	13,246	1980
	2017	Primary school	2,835,412	11,131	1679
		Secondary school	2,501,896	2570	372
		Total	5,337,307	13,701	2050
	2018	Primary school	3,698,783	11,619	1761
		Secondary school	2,634,279	2660	386
		Total	6,333,062	14,279	2147
	2019	Primary school	3,790,610	11,978	1823
		Secondary school	2,781,181	2769	403
		Total	6,571,791	14,747	2226

Table 33: Describing the ethnic profile of Sample 1 and 2 between 2006 – 2019

Sample	Year	Phase	White British (%)	White Other (%)	Black (%)	Asian (%)	Other (%)
1	2006	Primary school	78	3.3	6.1	8.9	3.7
		Secondary school	80.7	2.8	4.9	7.6	4.1
		Total	79.4	3	5.5	8.3	3.9
	2007	Primary school	77.1	3.7	6.4	9.4	3.5
		Secondary school	80	3	5.1	8	3.8
		Total	78.6	3.4	5.7	8.7	3.7

	2008	Primary school	75.7	4.1	6.6	9.9	3.8
		Secondary school	78.7	3.5	5.4	8.4	3.9
		Total	77.2	3.8	6	9.1	3.8
	2009	Primary school	74.7	4.5	6.8	10.4	3.6
		Secondary school	78	3.9	5.7	8.8	3.7
		Total	76.4	4.2	6.2	9.6	3.6
	2010	Primary school	73.8	4.7	7	10.8	3.7
		Secondary school	77.3	3.9	6	9.1	3.7
		Total	75.6	4.3	6.5	10	3.7
	2011	Primary school	72.8	4.9	7.3	11.2	3.8
		Secondary school	76.5	4	6.3	9.5	3.7
		Total	74.7	4.5	6.8	10.3	3.8
	2012	Primary school	71.7	5.2	7.5	11.6	4
		Secondary school	75.6	4.1	6.5	9.9	3.8
		Total	73.7	4.7	7	10.8	3.9
	2013	Primary school	70.7	5.6	7.7	11.9	4.1
		Secondary school	74.4	4.4	6.8	10.4	3.9
		Total	72.5	5	7.3	11.2	4
	2014	Primary school	69.8	6.1	7.9	12	4.2
		Secondary school	73.1	4.8	7.1	10.9	4.1
		Total	71.4	5.4	7.5	11.5	4.2
	2015	Primary school	68.8	6.5	7.9	12.3	4.5
		Secondary school	71.8	5.1	7.3	11.4	4.4
		Total	70.2	5.8	7.6	11.9	4.4
	2016	Primary school	67.9	7.1	8	12.4	4.6
		Secondary school	70.5	5.5	7.6	11.9	4.5
		Total	69.1	6.3	7.8	12.2	4.6

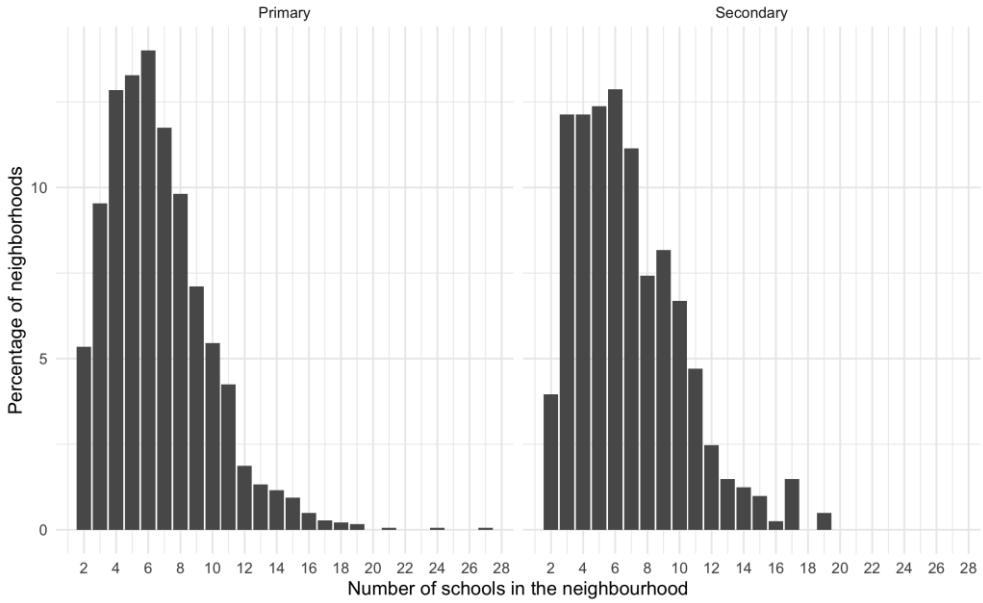
	2017	Primary school	67.2	7.5	8	12.5	4.7
		Secondary school	69.1	5.8	7.8	12.4	5
		Total	68	6.7	7.9	12.5	4.8
	2018	Primary school	66	7.8	8	13	5.1
		Secondary school	67.7	6.1	8.1	12.9	5.3
		Total	66.7	7.1	8	13	5.2
	2019	Primary school	65.4	8.1	7.9	13.2	5.4
		Secondary school	66.5	6.3	8.2	13.3	5.6
		Total	65.9	7.3	8.1	13.2	5.5
2	2006	Primary school	63.6	5	10.6	15.5	5.4
		Secondary school	69.6	3.9	8.1	12.6	5.7
		Total	66.7	4.4	9.3	14	5.5
	2007	Primary school	62.9	5.5	10.7	15.8	5.1
		Secondary school	69	4.2	8.4	13	5.3
		Total	66.1	4.8	9.5	14.4	5.2
	2008	Primary school	62.5	5.9	10.5	15.8	5.3
		Secondary school	68.7	4.8	8.4	13	5.1
		Total	65.7	5.3	9.4	14.3	5.2
	2009	Primary school	61.6	6.4	10.6	16.2	5.1
		Secondary school	67.3	5.3	8.9	13.6	4.9
		Total	64.5	5.8	9.8	14.9	5
	2010	Primary school	60.9	6.5	10.8	16.6	5.1
		Secondary school	67.1	5.2	9.1	13.7	4.9
		Total	64.1	5.9	9.9	15.1	5
	2011	Primary school	60.3	6.7	10.9	16.8	5.3
		Secondary school	66.1	5.2	9.5	14.2	4.9
		Total	63.2	6	10.2	15.5	5.1

	2012	Primary school	60	6.9	10.9	16.8	5.4
		Secondary school	65.6	5.3	9.6	14.6	4.9
		Total	62.8	6.1	10.3	15.7	5.1
	2013	Primary school	59.9	7.3	10.9	16.6	5.4
		Secondary school	64.9	5.6	9.7	14.8	5
		Total	62.3	6.4	10.3	15.7	5.2
	2014	Primary school	59.6	7.8	10.8	16.4	5.4
		Secondary school	64.3	5.9	9.7	15	5.2
		Total	61.8	6.8	10.3	15.7	5.3
	2015	Primary school	59.3	8.2	10.6	16.3	5.6
		Secondary school	63.4	6.2	9.8	15.3	5.4
		Total	61.2	7.2	10.2	15.8	5.5
	2016	Primary school	59.3	8.7	10.4	16	5.6
		Secondary school	62.8	6.5	9.8	15.4	5.4
		Total	60.9	7.7	10.1	15.7	5.5
	2017	Primary school	59.1	9.1	10.2	15.8	5.7
		Secondary school	62.1	6.7	9.8	15.5	5.9
		Total	60.5	8	10	15.7	5.8
	2018	Primary school	58.8	9.3	9.8	16	6
		Secondary school	61.5	7	9.8	15.7	6.1
		Total	59.9	8.3	9.8	15.9	6
	2019	Primary school	58.8	9.4	9.6	15.9	6.3
		Secondary school	61.1	7.2	9.7	15.6	6.4
		Total	59.8	8.5	9.6	15.8	6.3

11.5 Exploring the distribution of pseudo-neighbourhood sizes when *k* = 2800 for primary school pseudo-neighbourhoods and *k* = 550 for secondary school pseudo-neighbourhoods.

The two graphs below describe the distribution for the number of schools within each pseudo-neighbourhood when *k* = 2800 for primary schools and *k* = 550 for secondary schools. In both cases, the average number of schools in each pseudo-neighbourhood is 6 to reflect the number of schools that parents can list in their child's application form. As the diagrams show, there is a slight positive skew in the data, though most pseudo-neighbourhoods contain a similar number of schools. This suggests that k mean's clustering defines pseudo-neighbourhoods which are relatively uniform in size, albeit with a degree of variation which should reflect the different size of educational markets.

Figure 44: Exploring the number of schools included in a pseudo-neighbourhood when *k = 2800* for primary school pseudo-neighbourhoods and *k = 550 for* secondary school pseudo-neighbourhoods.



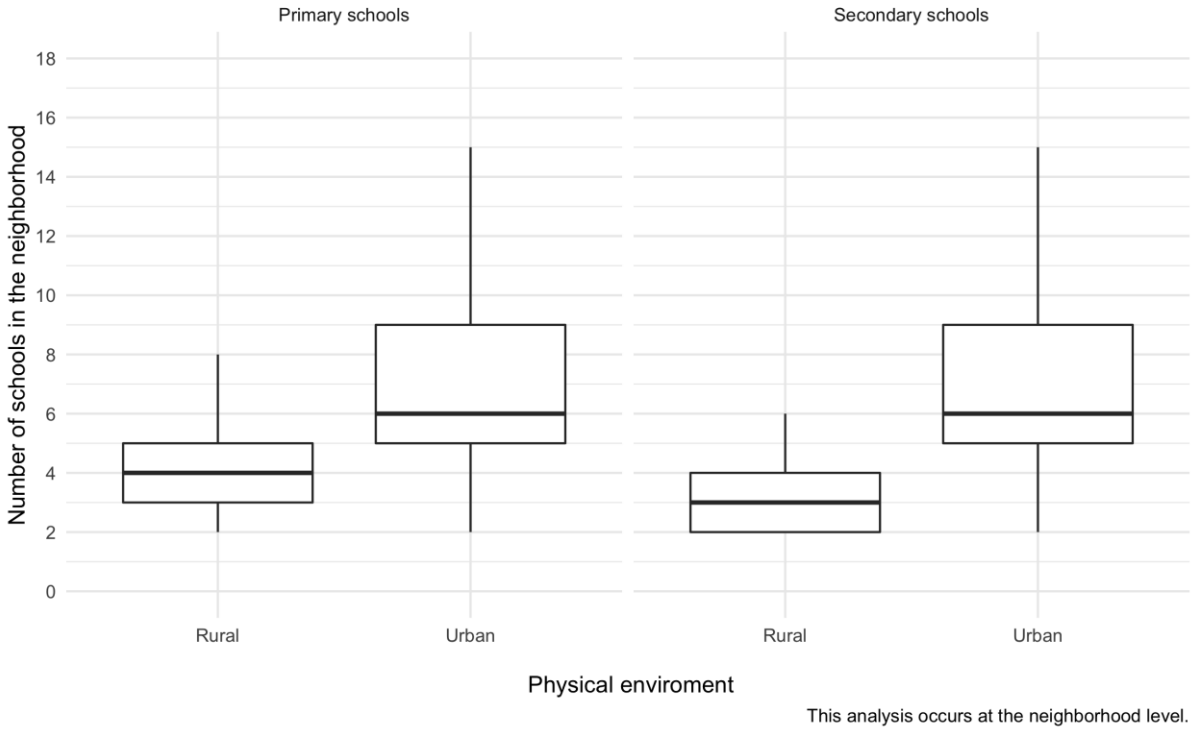
11.6 Comparing the number of schools in rural and urban pseudo-neighbourhoods when using k-means clustering

An analysis using Sample 2 found that rural pseudo-neighbourhoods, on average, contained less schools than pseudo-neighbourhoods in urban environments for both primary and secondary school pseudo-neighbourhoods.[213] As shown in Figure 45 below, the simulation found that rural primary pseudo-neighbourhoods had an average of 4.1 primary schools whereas urban pseudo-neighbourhoods had an average of 7.1 primary schools. Similarly,

[213] Pseudo- neighbourhoods were defined as rural or Urban depending on the variable "Urban rural" listed in the school census dataset.

rural secondary pseudo-neighbourhoods contained an average of 3.5 secondary schools whereas urban pseudo-neighbourhoods contained an average of 6.9 secondary schools. This suggests that k means clustering produces pseudo-neighbourhoods that reflect the different structures of rural and urban educational markets.

Figure 45: Comparing the number of schools in rural and urban pseudo-neighbourhoods



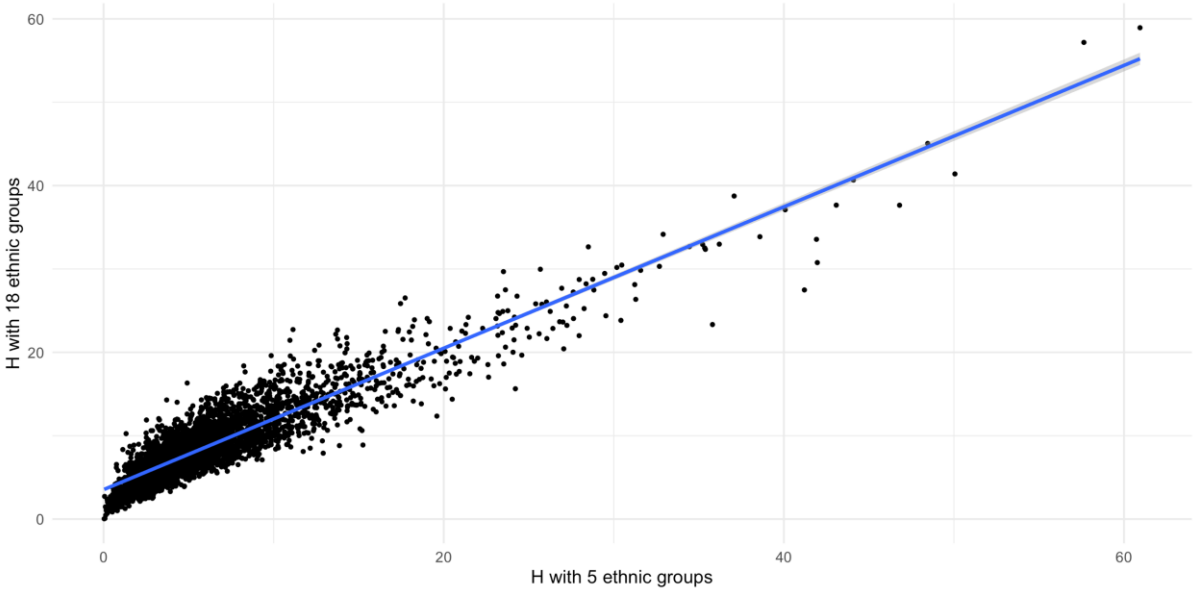
11.7 Comparing an *H* Index calculated using 5 vs 18 ethnic groups.

Using the 2019 school census data, the study calculated an *H* index using five ethnic groups (H_5) and then, using the same data and pseudo-neighbourhood groupings, calculated an *H*

index using 18 ethnic groups (H_{18}). This process was repeated 100 times with the median *H* values and the correlation between the two values (H_5 and H_{18}) being recorded.

The analysis found a substantial difference between the median values of H_5 and H_{18}, with an average H_{18} of 8.3 and an average H_5 of 5.2. This showed that using 18 ethnic groups would lead to measures of *H* being 3.1 units, or 60%, higher, on average, than when using 5 ethnic groups.

Figure 46: Correlation between an H index using 5 vs 18 ethnic groups.



However, these figures were highly correlated, with H_5 and H_{18} having an average correlation of 0.9 ($p<0.001$) across the 100 iterations. This suggests that the decision to use H_5 or H_{18} would only impact the direct values of *H*, instead of the shape of the distribution. As *H* does not have an intuitive interpretation, this was not particularly problematic.

Consequently, the study concluded that the added clarity of describing results for only 5 ethnic groups justified using H_5 over H_{18}.

11.8 Calculating the level of "random segregation" or "upward bias" in *H* when measuring ethnic school segregation in England's school system

Some researchers argue that segregation indexes are upwardly bias (i.e. measures are too large) as they fail to account for the segregation that occurs even when students are randomly assigned to schools (Cortese, Falk and Cohen, 1976, p. 361; Allen, Burgess and Windmeijer, 2009; Rathelot, 2012). To evaluate how sensitive *H* is to this "upward bias" or "random" segregation an analysis was conducted to measure the level of segregation that would be recorded after students in England were randomly assigned to a school in their pseudo-neighbourhood. The analysis used 2019 school data and only included ethnically diverse pseudo-neighbourhoods- as defined by any pseudo-neighbourhood where the largest ethnic group constituted less than 90% of the population. The analysis followed the steps listed below:

Table 34: The procedure to find the level of "random" ethnic school segregation

Step Number	Description
Step 1	Schools were grouped into pseudo-neighbourhoods using k-means clustering.
Step 2	The school-level dataset was transformed into a student level dataset, where each row represented a different student. This dataset contained three variables – the student's school ID (URN), the school's pseudo-neighbourhood ID and the student's ethnicity.
Step 3	A function was run which grouped student attending schools in the same pseudo-neighbourhood together and then shuffled the student URN column.[214] This mimics the process of randomly allocating students to a school within the pseudo-neighbourhood of their current school.
Step 4	This shuffled student data was then aggregated back to school level data.
Step 5	The *H* index was then calculated using this shuffled school level data. This measures the level of segregation which occurs when students are randomly allocated to a school in their pseudo-neighbourhood.
Step 6	The algorithm was then rerun 20 times and an average *H* score was found across all the iterations.

Unfortunately, this analysis was computationally expensive which limited the number of times that the algorithm could feasibly be repeated (step 6). Luckily, over 20 iterations the findings were highly stable. The results found that when students were randomly assigned to a school in their pseudo-neighbourhood, the median *H* value was 0.65 for primary schools and 0.19 for secondary schools, which accounts for around 12% of the real level of

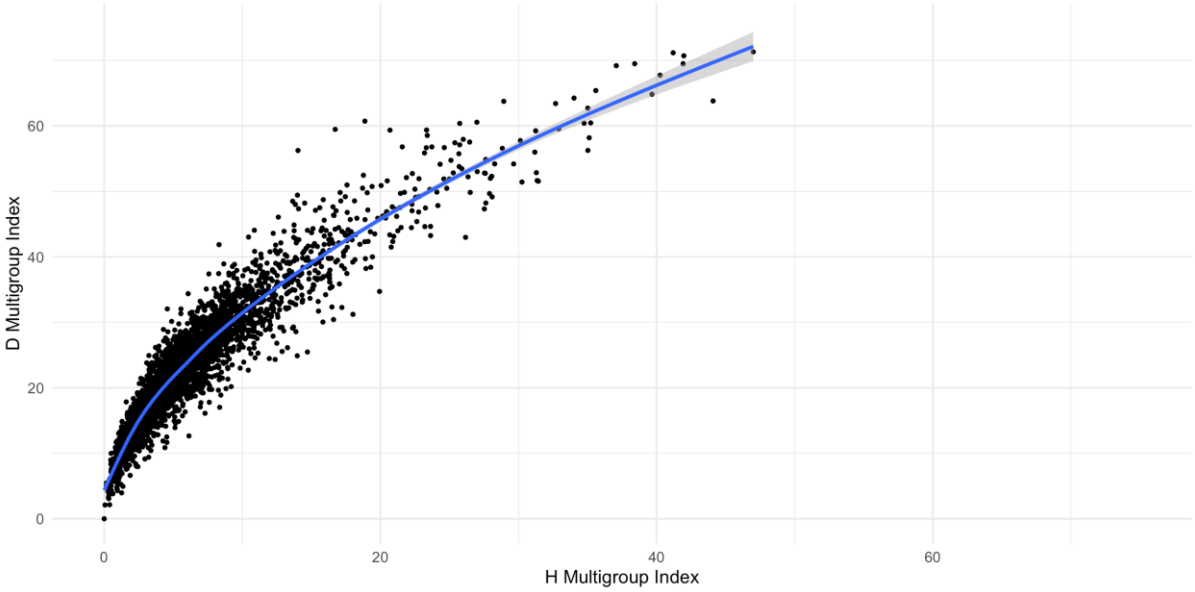
[214] Technically, the URN column is replaced by a vector which is created from drawing a random sample from the URN column without replacement. Conceptually this is the same as shuffling the column.

segregation currently found in England's primary school system and 3% for the secondary school system.

11.9 Comparing the relationship between a multigroup *H* Index, a multigroup *D* Index and a two group *D* Index.

To explore how *H* correlated with a binary (D_{Bin}) and multigroup *D* Index (D_{MG}), all three indexes were calculated using 2019 school data and then correlated with one another across 100 iterations. As shown in Figure 47 below, *H* had a strong association with multigroup *D*. Across 100 iterations, they had a correlation of 0.92 which suggest using different multigroup indexes on school data in England produces similar results. *H* also had a relatively strong correlation to a two group *D* Index (White British, Ethnic Minority), with an average correlation of 0.6 across all 100 iterations ($p<0.001$).

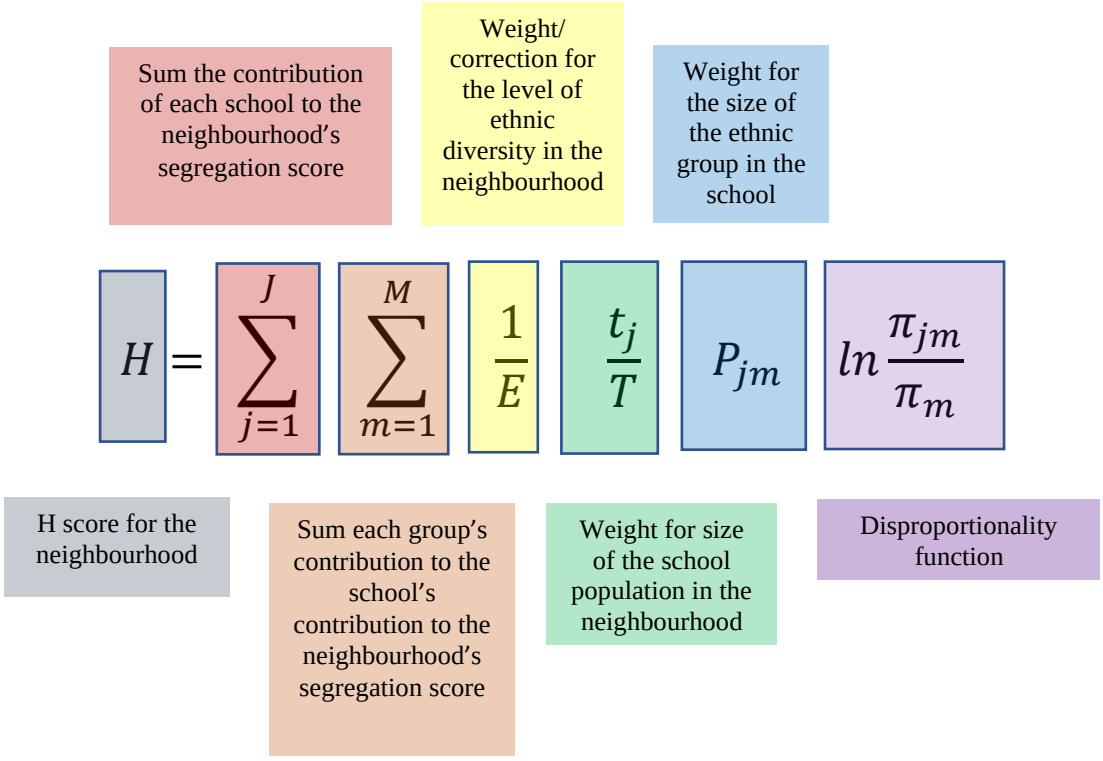
Figure 47: Exploring the relationship between the *H* index and multi-group *D* Index



11.10 Breakdown of the *H* Index

The figure below provides a diagrammatic explanation of the *H* index.

Figure 48: A Breakdown of the *H* Index



11.11 The Two group D Index

The formula for *D* is listed below.

Equation 3: The D Index

$$\mathrm{D} = \frac{1}{2}\sum_{school}\left|\frac{Group\ A_{school}}{Group\ A_{LA}} - \frac{Group\ B_{school}}{Group\ B_{LA}}\right|$$

Where:

$Group\ A_{school} = Number\ of\ students\ belonging\ to\ ethnic\ group\ A\ in\ a\ school_i$

$Group\ A_{LA} = Number\ of\ students\ belonging\ to\ ethnic\ group\ A\ in\ local\ authority_i$

$Group\ B_{school} = Number\ of\ students\ belonging\ to\ ethnic\ group\ B\ in\ school_i$

$Group\ B_{LA} = Number\ of\ students\ belonging\ to\ ethnic\ group\ B\ in\ local\ authority_i$

11.12 The Isolation index (*I*)

The formula for *I* is listed below.

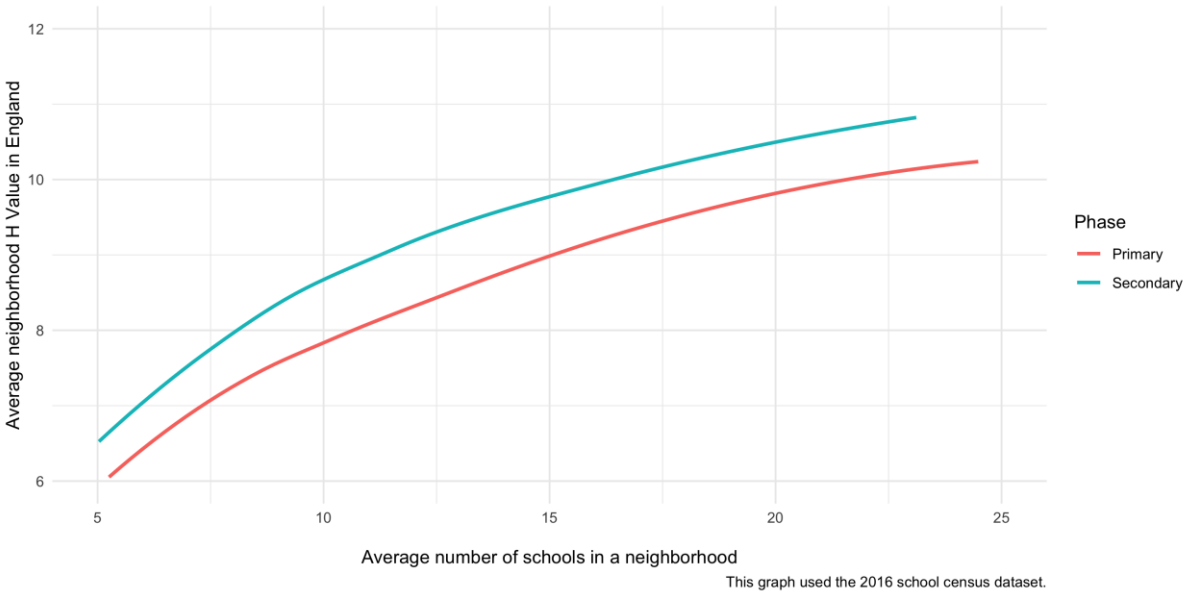
Equation 4: The Isolation Index (I)

$$\mathrm{I} = \sum\left(\frac{Group\ A_{school}}{GroupA_{neighborhood}}\right)\left(\frac{Group\ A_{school}}{Population_{school}}\right)$$

11.13 Exploring the relationship between the number of clusters (*k*) and the level of segregation (*H*) on English school data

As shown in Figure 49 below, there is a positive relationship between the size of a pseudo-neighbourhood and the observed level of ethnic school segregation (*H*). If England was split into pseudo-neighbourhoods containing, on average, 5 schools then the average *H* index would be 6.1. If, however, pseudo-neighbourhoods were created larger and contained, on average, 25 schools then the average *H* value would be 10.6, which is 73% higher. These results highlight how important the scale of pseudo-neighbourhood's is to the level of observed segregation.

Figure 49: Evaluating the average pseudo-neighbourhood *H* value when using pseudo-neighbourhoods of different sizes.



11.14 Diagrams to show the numeric and geographic distribution of each of the predictor variables

The following section provides readers with a density plot and map of each of the continuous predictor variables for the pseudo-neighbourhood level meta regression analysis so they can explore their numeric and geographic distributions.

Figure 50: Population Densitys of ethnically diverse pseudo-neighbourhoods in Sample 2 (2019).

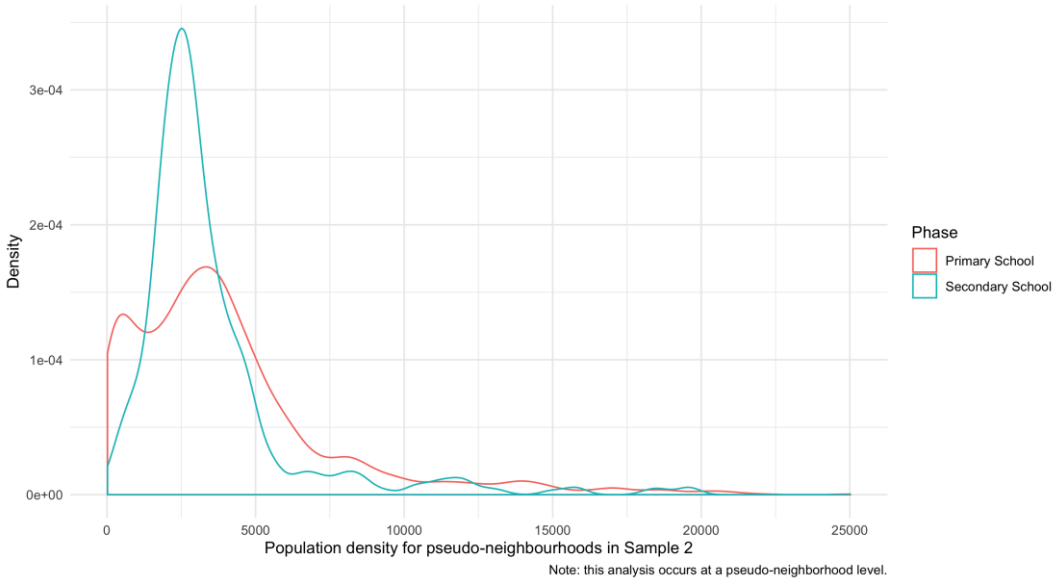


Figure 51: Mapping the Population Density of ethnically diverse pseudo-neighbourhoods in Sample 2 (2019).

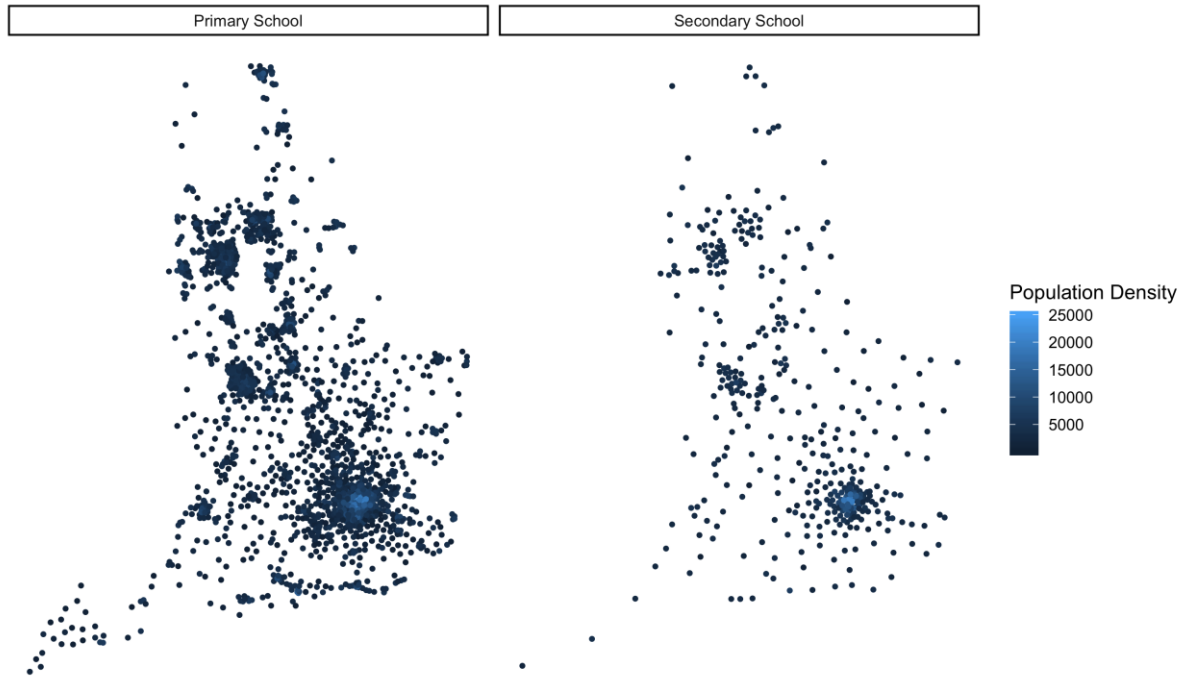


Figure 52: Rates of FSM for ethnically diverse pseudo-neighbourhoods in Sample 2 (2019).

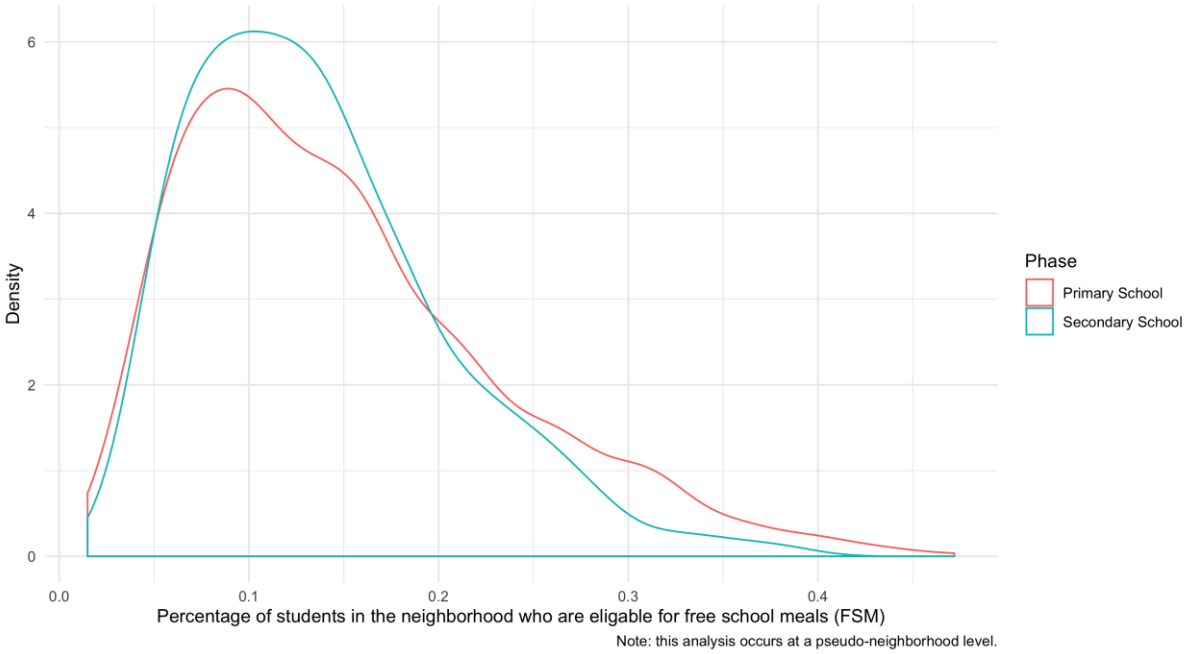


Figure 53: Mapping the rates of FSM for ethnically diverse pseudo-neighbourhoods in England (2019).

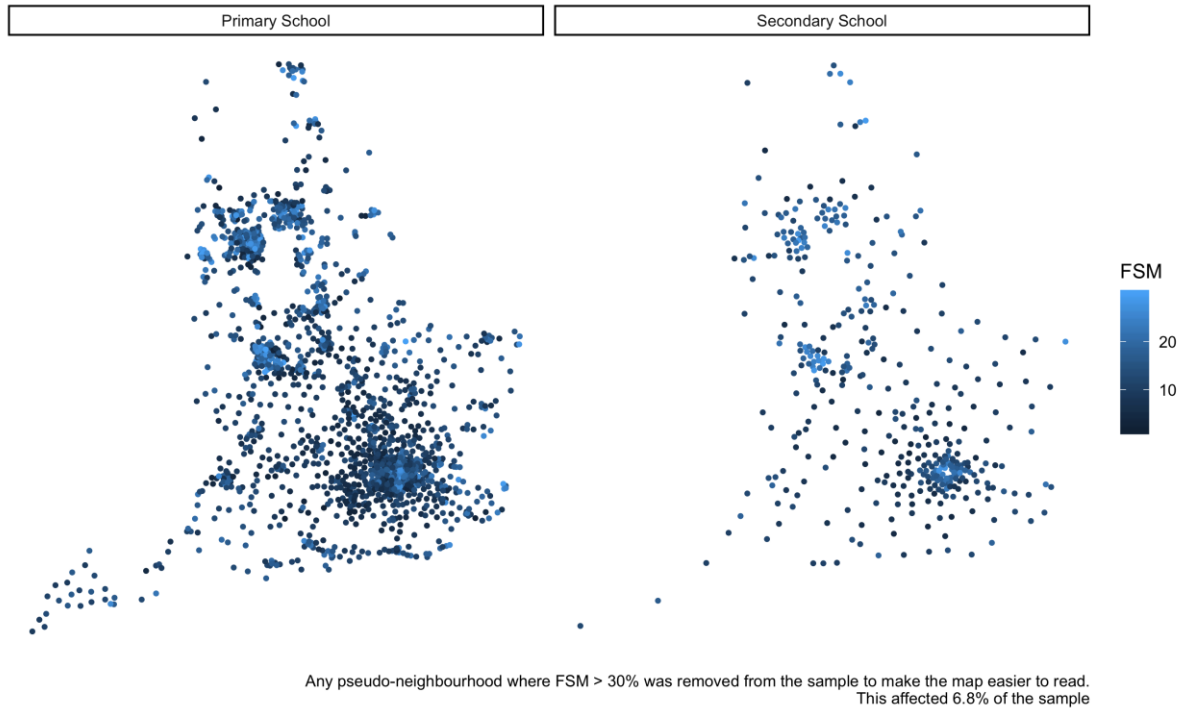


Figure 54: The average price of a terrace house in ethnically diverse pseudo-neighbourhoods in Sample 2 (2019)

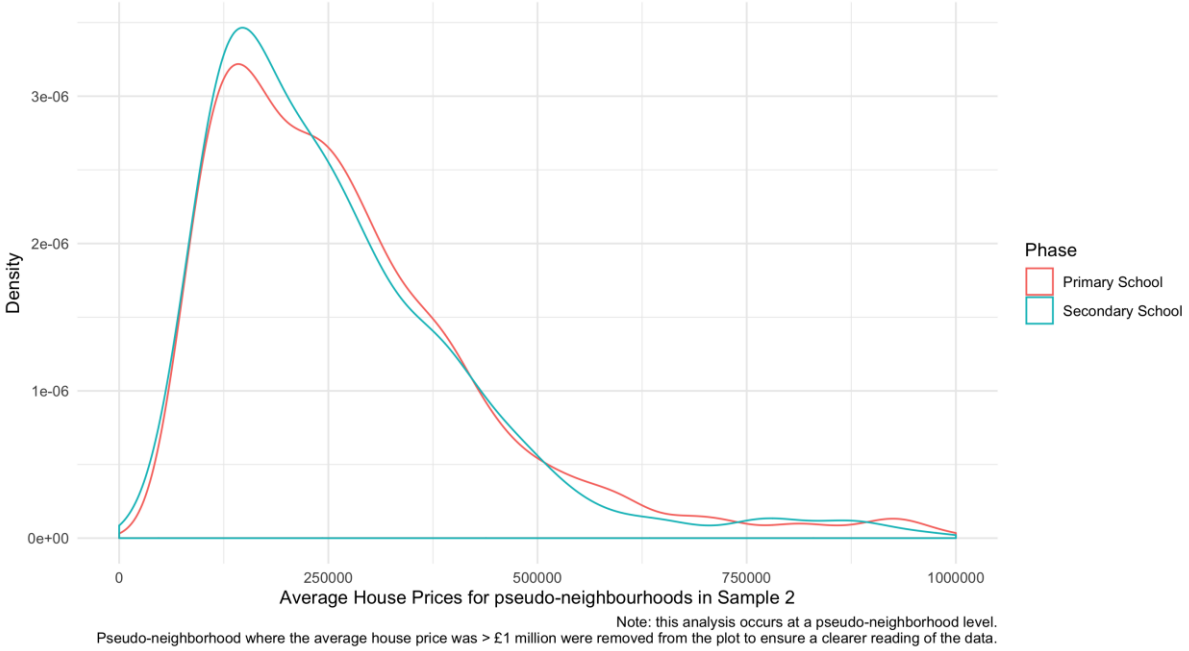


Figure 55: Mapping the average price of a terrace house for pseudo-neighbourhoods in Sample 2 (2019).

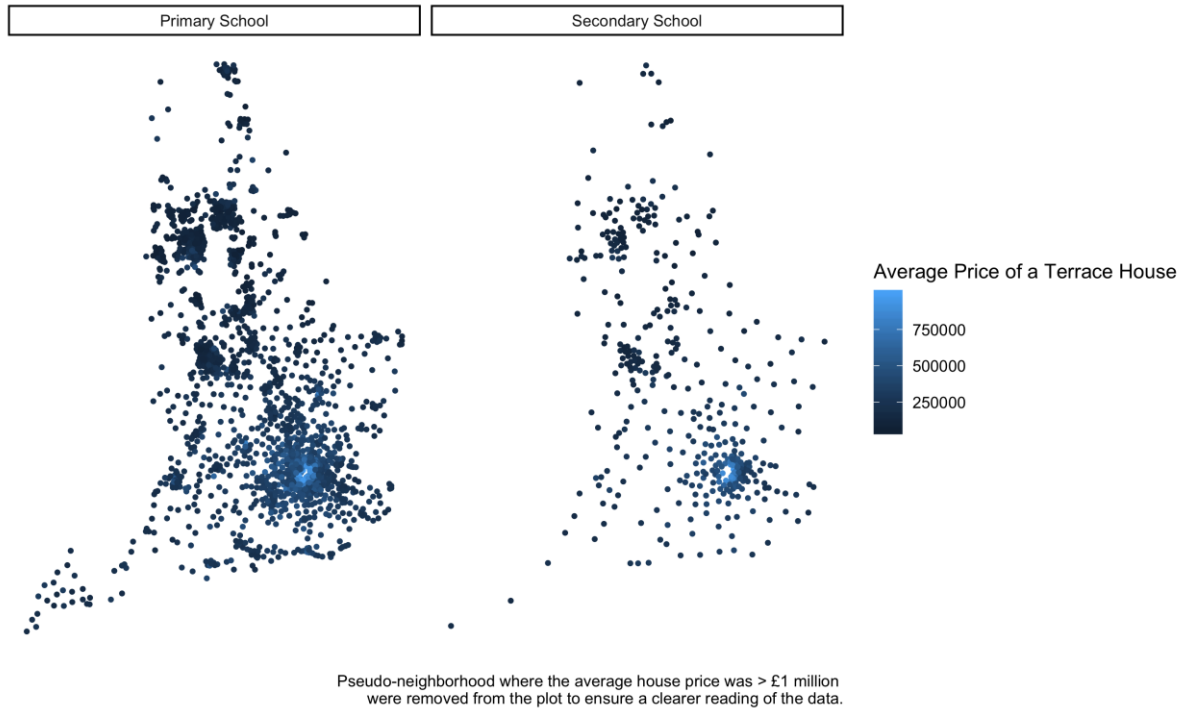


Figure 56: Density plot to show the proportion of schools that were religious in sample 2 (2019).

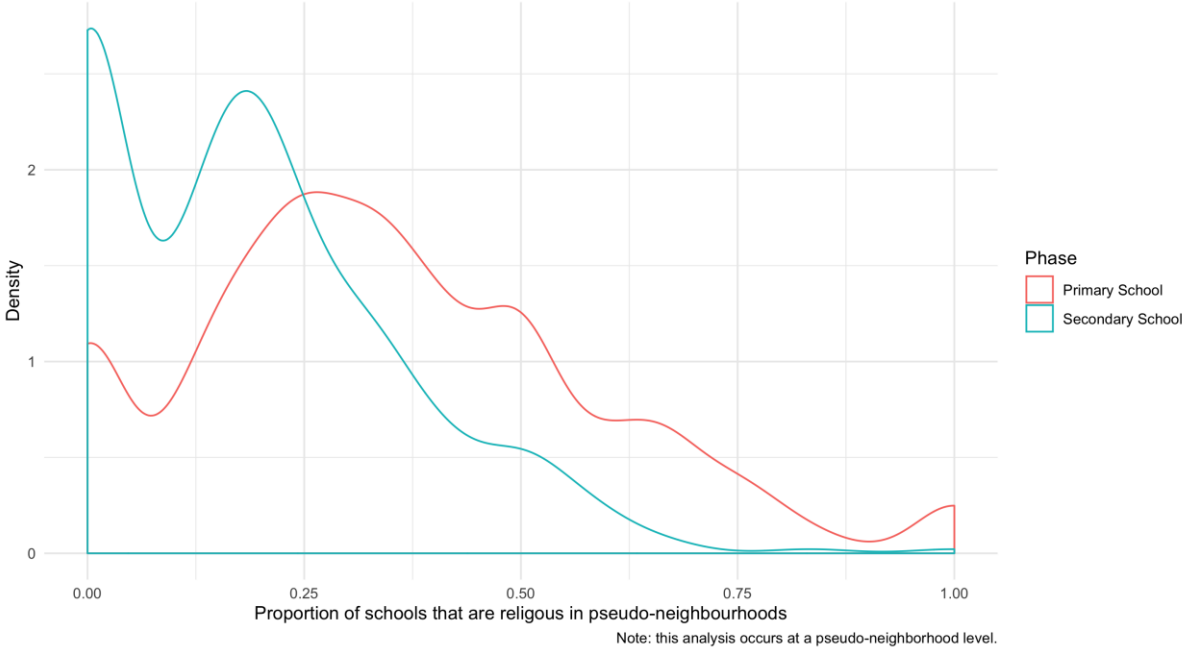


Figure 57: Mapping proportion of schools that were religious in pseudo-neighbourhoods within Sample 2 (2019).



Figure 58: Density plot to show the average OFSTED ratings of pseudo-neighbourhoods in sample 2 (2019).

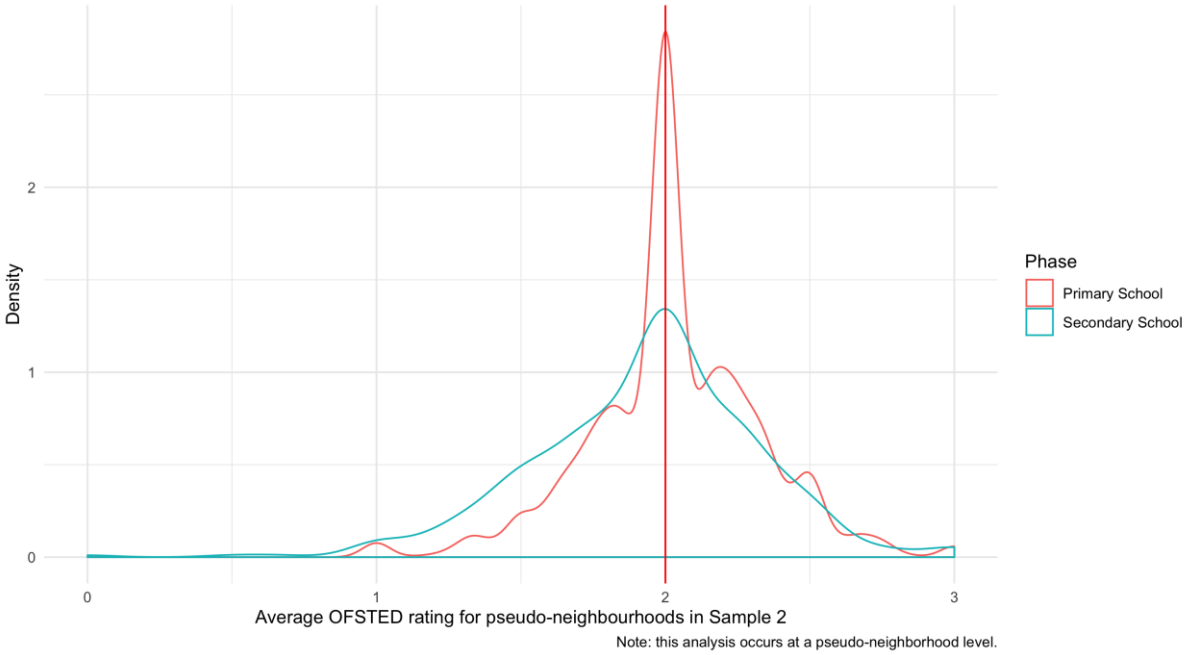


Figure 59: Mapping the average OFSTED Rating of pseudo-neighbourhoods within Sample 2 (2019).

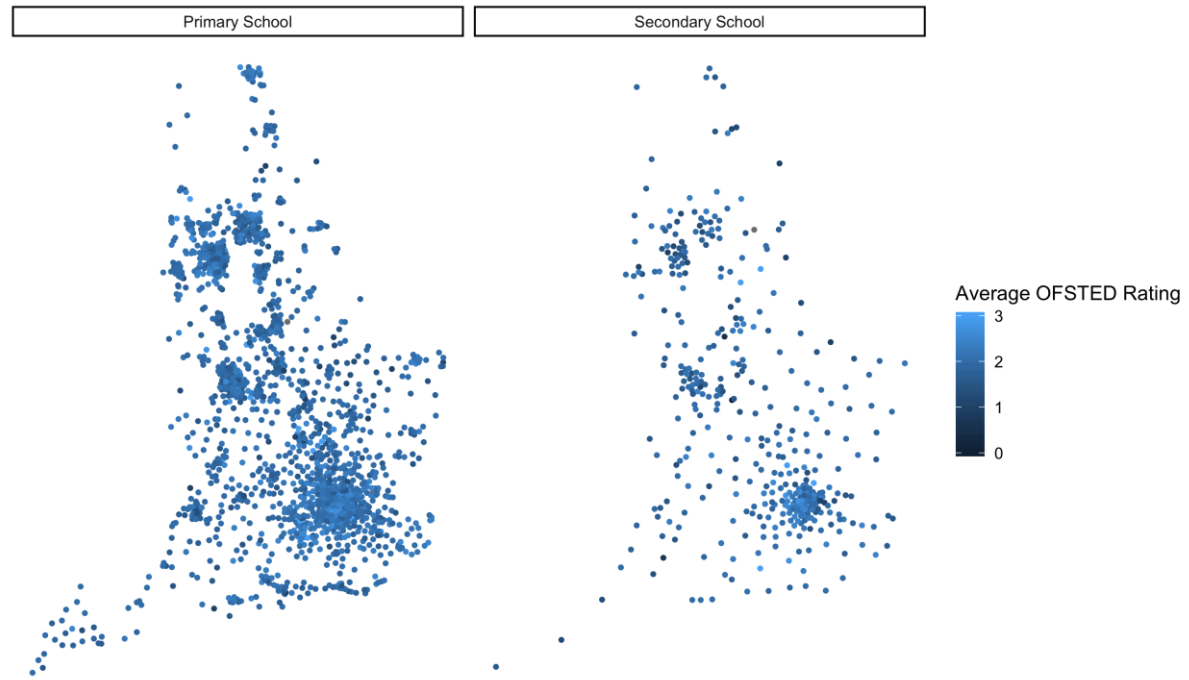


Figure 60: Density plot to show the level of ethnic diversity (*E*) in pseudo-neighbourhoods in sample 2 (2019).

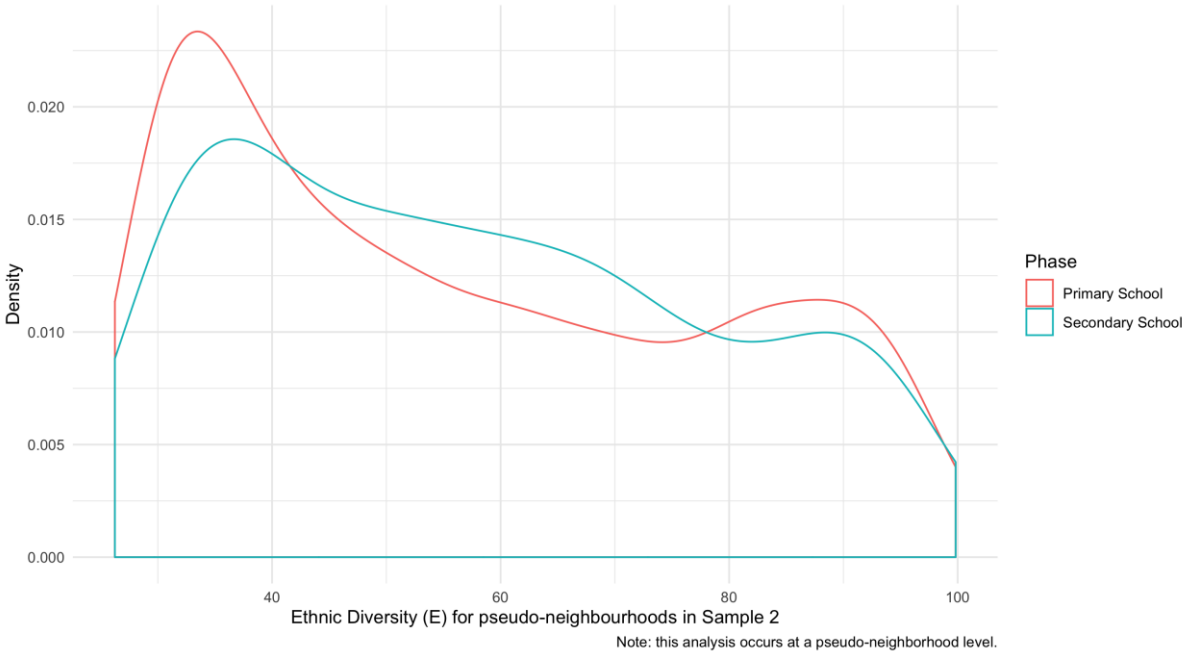


Figure 61: Mapping the ethnic diversity (*E*) of pseudo-neighbourhoods included in sample 2 (2019).

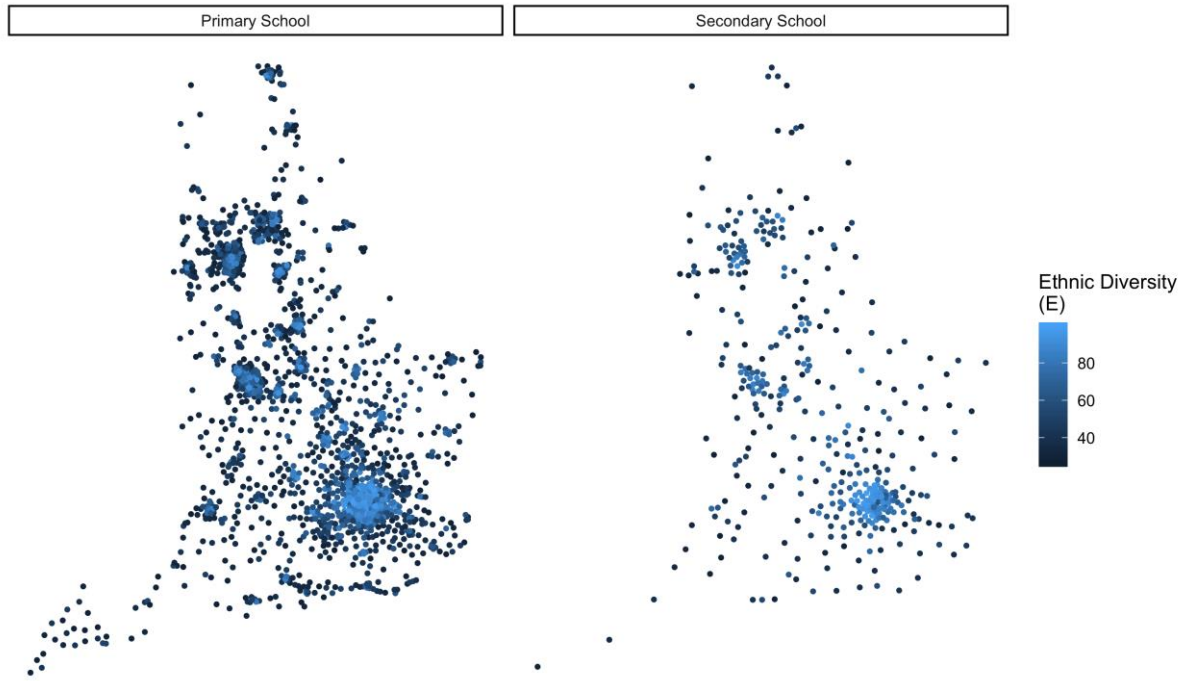


Figure 62: Density plot to show the number of schools in the pseudo-neighbourhoods that were included in sample 2 (2019).

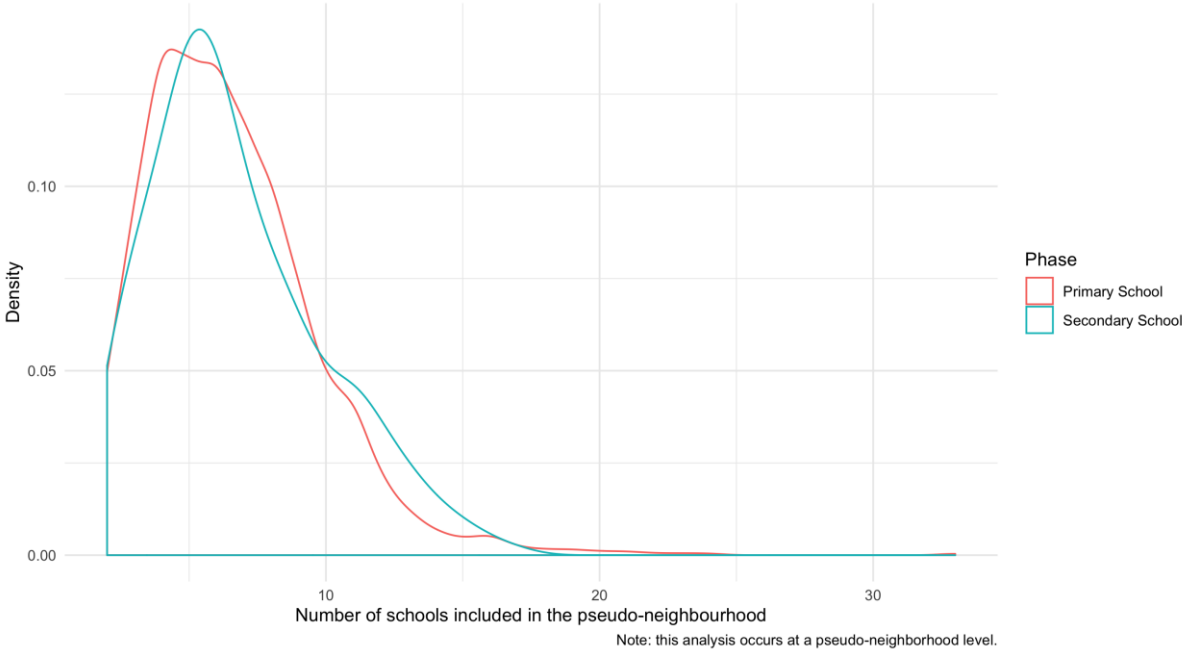
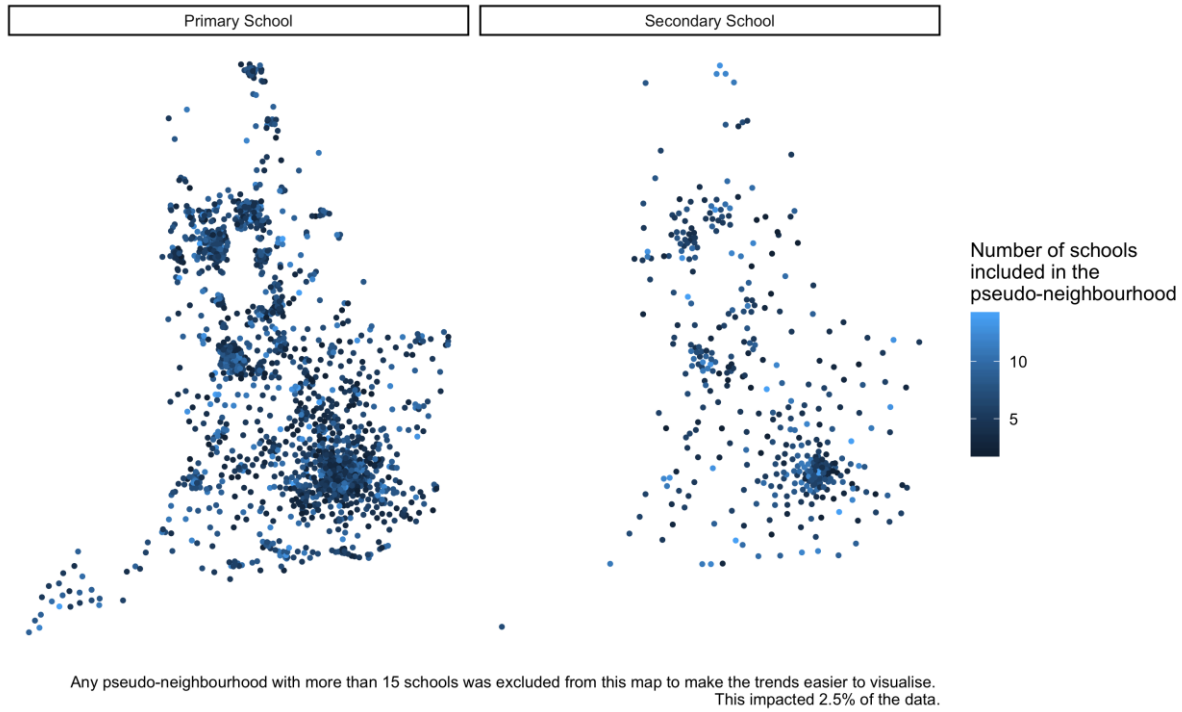


Figure 63: Mapping the number of schools in the pseudo-neighbourhoods that were included in sample 2 (2019).



11.15 Evaluating FSM as a measure of socio-economic disadvantage

As with any metric of socio-economic disadvantage, FSM is not perfect. Research has found that rates of FSM tend to underestimate the extent of disadvantage within a neighbourhood (Ilie, Sutherland and Vignoles, 2017; Taylor, 2018) and, as a binary metric, FSM is a crude measure of socio-economic status. In particular it offers little insight into the background of the 86% of students who are not FSM. Furthermore, as FSM is solely determined by the level of parental income, it ignores other factors which also impact socio-economic status such as their parent's level of formal education or their "social class".

Nonetheless FSM still provides the best measure for this thesis. It remains the most commonly used metric for measuring disadvantage in educational research. Given how many new methodological approaches this research adopts, using a commonly understood

metric should make the findings more accessible to readers and policy makers. Secondly, other measures, such as the Income deprivation affecting children index (IDACI), are calculated for geographic areas using government boundaries which makes them difficult to incorporate into pseudo-neighbourhood's that are defined by k mean's clustering. In contrast, rates of FSM are assigned to schools which is makes its inclusion easier. Lastly, FSM remains highly correlated with most metrics of socio-economic disadvantage which suggests that, despite its faults, rates of FSM still provide a reliable way of gauging socio-economic disadvantage.

11.16 Comparing the correlation of a school and pseudo-neighbourhood's house price when using a LSOA data or a cluster approach.

As the analysis below shows, estimates for the house price surrounding a school using LSOA data or a cluster approach are highly correlated. When comparing school level figures these estimates have a correlation of 0.77 for primary schools and 0.75 for secondary schools ($p<0.001$). For pseudo-neighbourhood averages the figures are even stronger, with the two values both having an average correlation of 0.82 ($p<0.001$) across 100 iterations.

Figure 64: Comparing the estimated house price surrounding a school using an LSOA dataset and a cluster approach

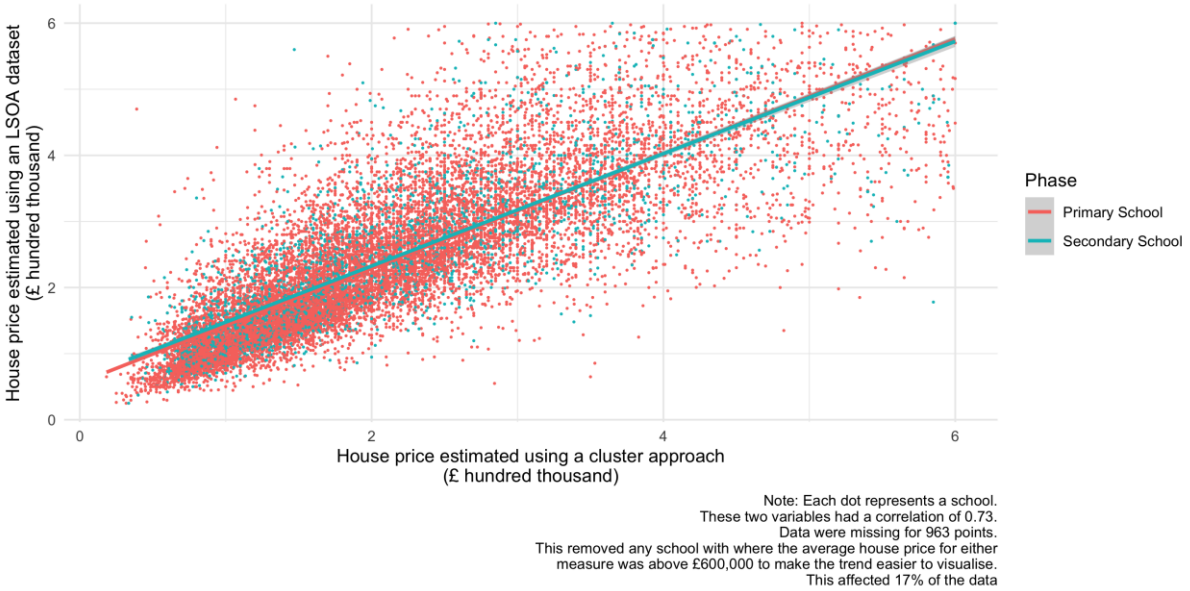
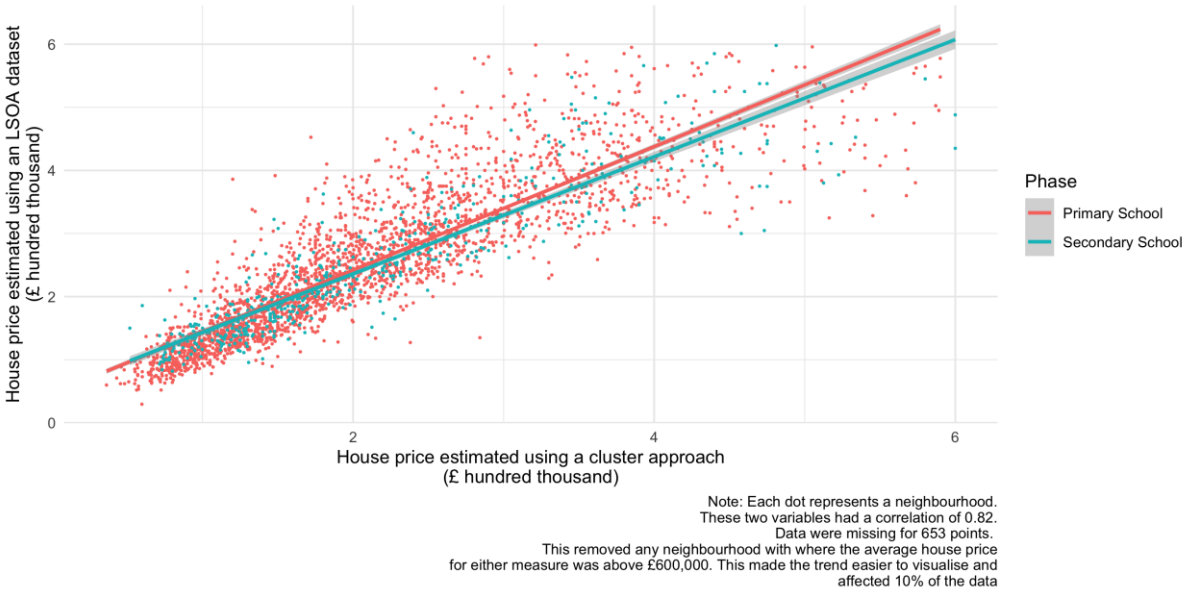


Figure 65 Comparing the estimated house price in a neighbourhood using an LSOA dataset and a cluster approach



11.17 Exploring the change in *H* for pseudo-neighbourhoods which were sufficiently diverse to be included in the sample in both 2006 and 2019

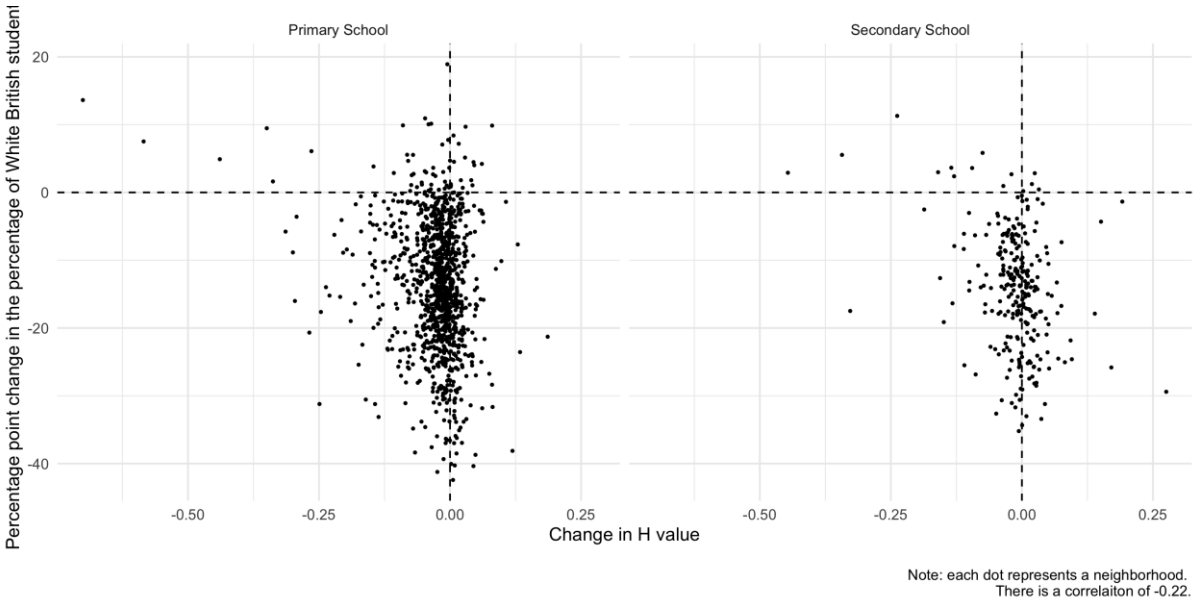
The study ran an analysis to check that changes in the average level of segregation were not unduly driven by changes in the sample between 2006 – 2019. The study compared the changes in median *H* values for ethnically diverse primary and secondary school pseudo-neighbourhoods that were included in the 2006 and 2019 sample. The changes in *H* values for pseudo-neighbourhoods was broadly similar between the two school sectors. Primary school pseudo-neighbourhoods who were sufficiently diverse to be included in the 2006 and 2019 sample experienced a median reduction in H of 1.64, while the equivariant figure for secondary school pseudo-neighbourhoods was 0.68. Both figures broadly align with the reduction in the median *H* value for all ethnically diverse primary school pseudo-neighbourhoods of 1.7 and the reduction in the median *H* value for all ethnically diverse secondary school pseudo-neighbourhoods 1.8. these figures suggests that, while the inclusion of newly diverse pseudo-neighbourhoods was likely to have encouraged a greater reduction in the median *H* value, segregation also fell in pseudo-neighbourhoods that were ethnically diverse in both 2006 and 2019.

11.18 Exploring whether changes in the level of ethnic school segregation reflect changes in the ethnic profile of pseudo-neighbourhoods

The study checked whether there was an association between the change in the proportion of White British students in a neighbourhood and its change in ethnic school segregation (*H*)

between 2006 – 2019. This was completed to explore whether observed changes in segregation directly reflect changes in the ethnic profile of a pseudo-neighbourhood. As this requires pseudo-neighbourhoods to have a *H* value for both 2006 and 2019, only pseudo-neighbourhoods who were ethnically diverse in both years were included in this sensitivity analysis. The analysis found weak evidence to suggest that there was an association, with a correlation of 0.341 for primary school pseudo-neighbourhoods and 0.149 for secondary school pseudo-neighbourhoods. However, a visual inspection of the data provides sufficient evidence to conclude that changes in segregation were not reflective of changes in the percentage of White British students, as shown in
Figure **66** below.

Figure 66: Comparing the change in segregation with the change in the level of White British students at a pseudo-neighbourhood level (2006 – 2019)

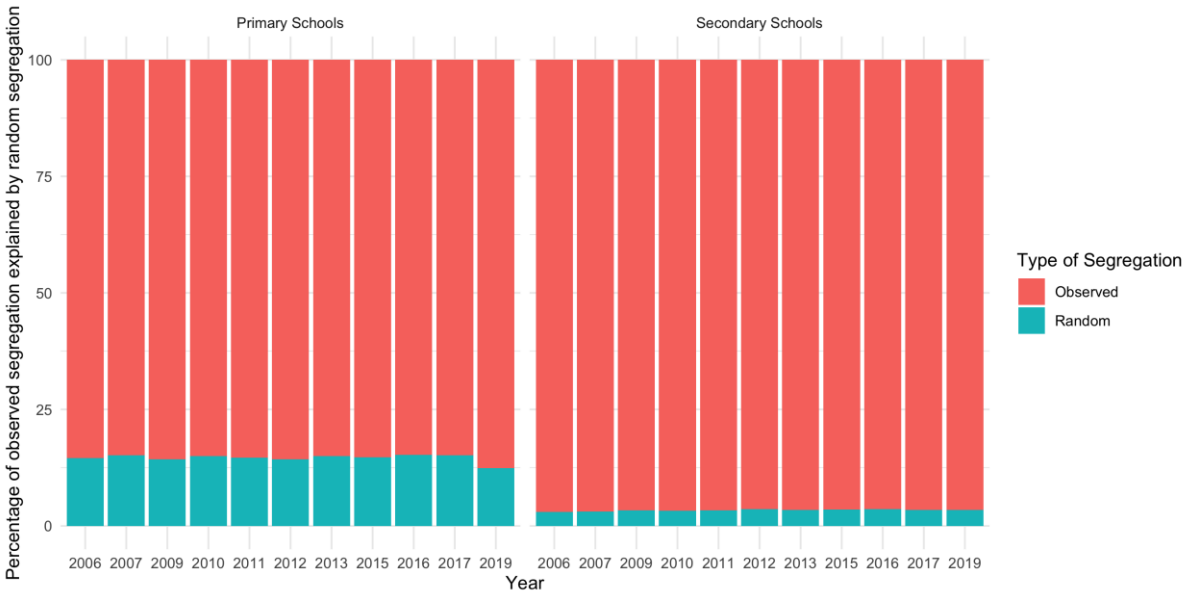


11.19 Exploring whether the level of "random segregation" changed between 2006 – 2019.

The study ran a simulation analysis which randomly assigned students to a school in their current pseudo-neighbourhood and measured the level of segregation based on these "random" school attendance patterns. This process was repeated 100 times, with each iteration using a different set of pseudo-neighbourhood boundaries. The level of random segregation was then compared across years to see if there were substantial differences that could be responsible for the observed changes in the median level of segregation.

As shown in Figure 67, the was no evidence to suggest that changes in the level of random segregation in the English school system explained changing patterns of ethnic school segregation. Random segregation consistently explained around 12% of the observed level of ethnic school segregation in England for primary school pseudo-neighbourhoods and 3% for secondary school pseudo-neighbourhoods, with no substantially important changes occurring between 2006 – 2019.

Figure 67: The proportion of ethnic school segregation in English pseudo-neighbourhoods explained by random segregation between 2006 – 2019.



11.20 Additional diagrams from the results section

The following section contains various figures which were not included in the main body of text but may still be of interest to readers.

Figure 68: The ethnic profile of pseudo-neighbourhoods in England (2019).

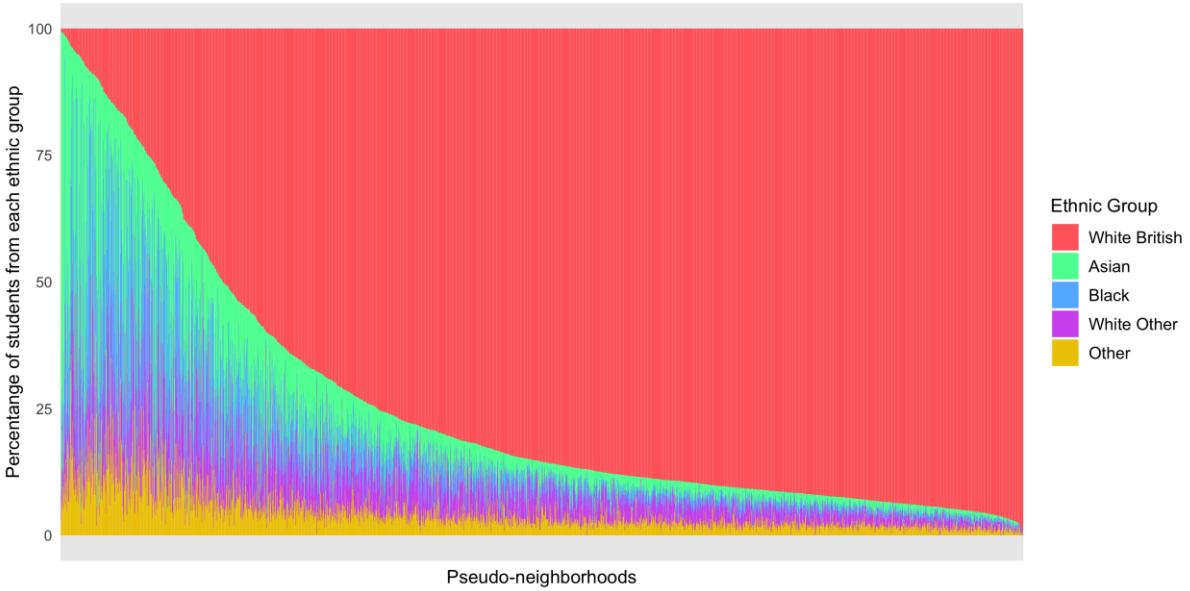


Figure 69: The percentage of White British students being educated in English pseudo-neighbourhoods (2019).

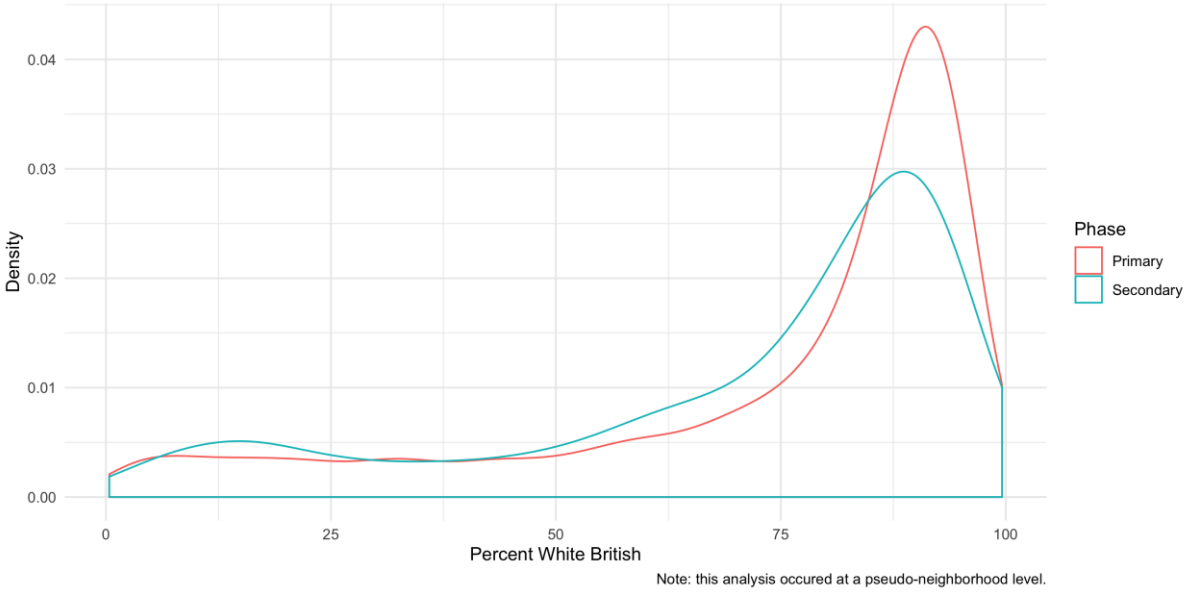


Figure 70: The percentage of students from Asian, Black, Other White, and Other ethnic groups being educated across pseudo-neighbourhoods in England (2019).

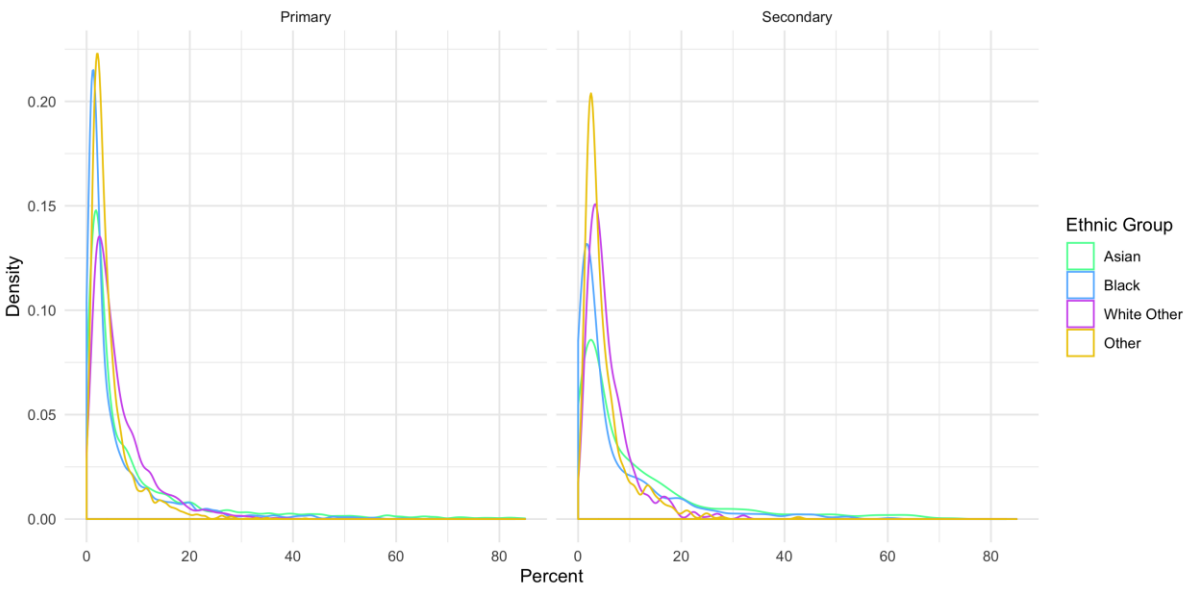


Figure 71: Mapping the percentage of White British students being educated in English pseudo-neighbourhoods (2019)

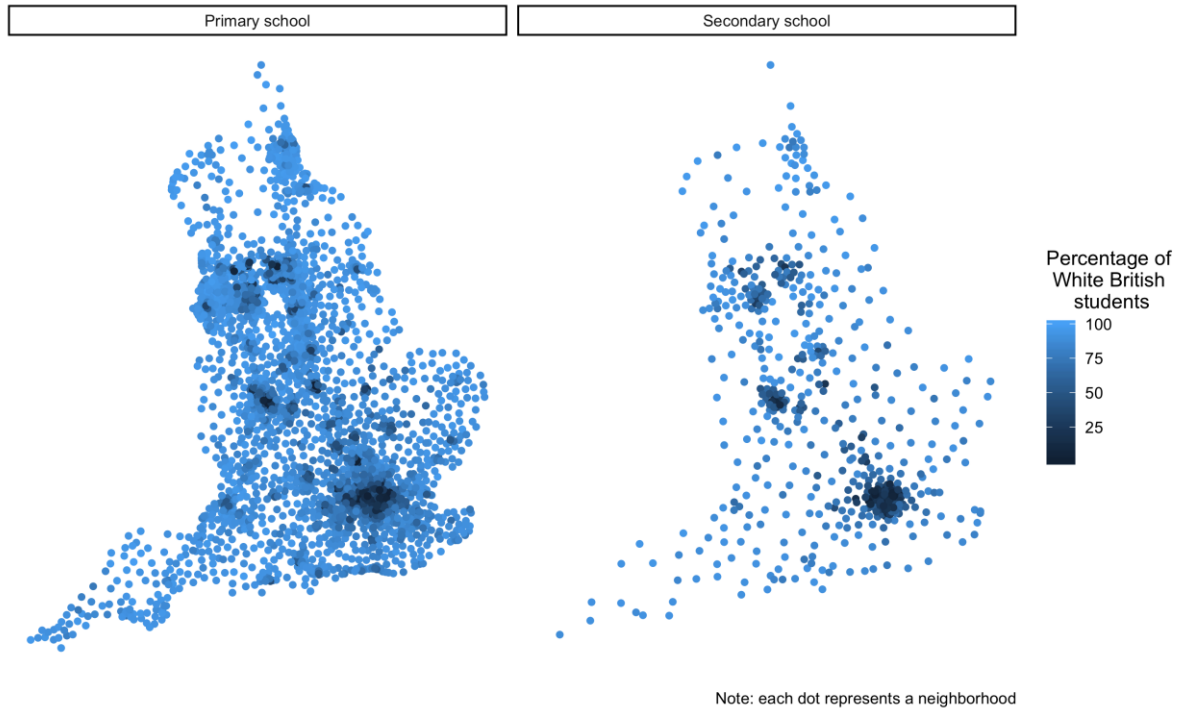


Figure 72: The changing number of White British and ethnic minority students, at a regional level (2006 – 2019).

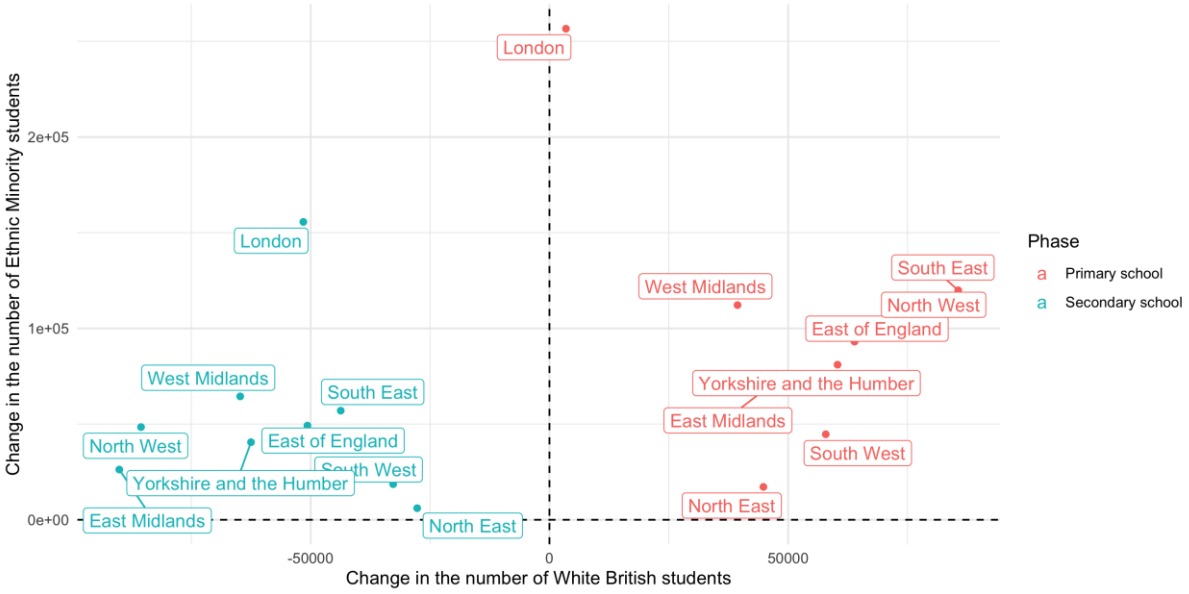


Figure 73: The changing proportion of White British students at a neighbourhood level (2006 – 2019).

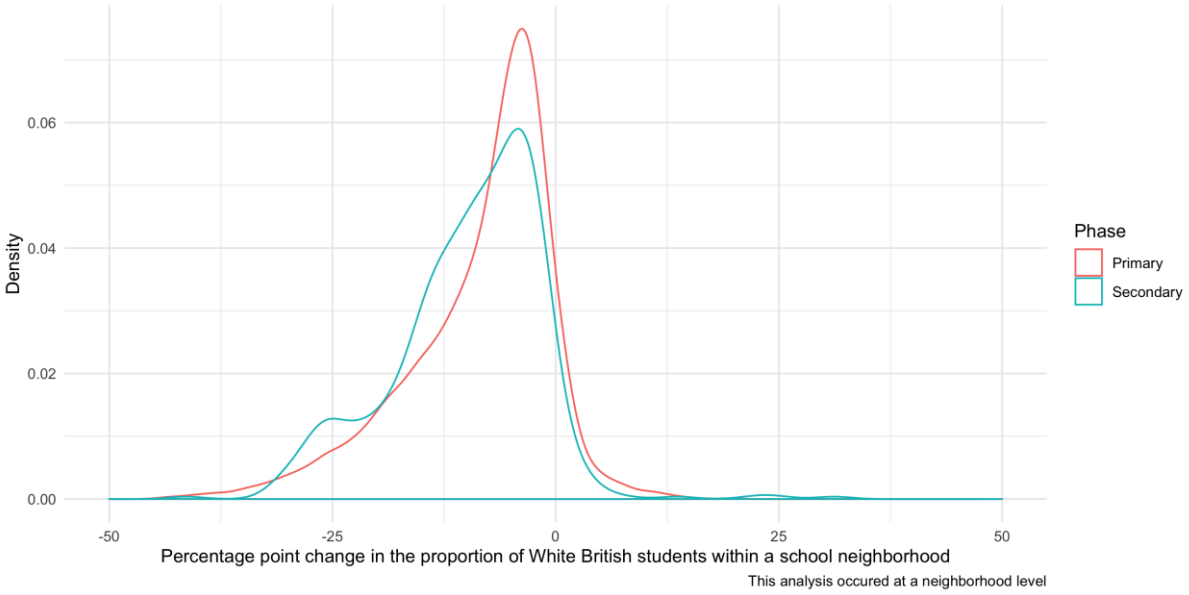


Figure 74: The distribution of changes in the proportion of students from each ethnic group within pseudo-neighbourhoods in England (2006 - 2019)

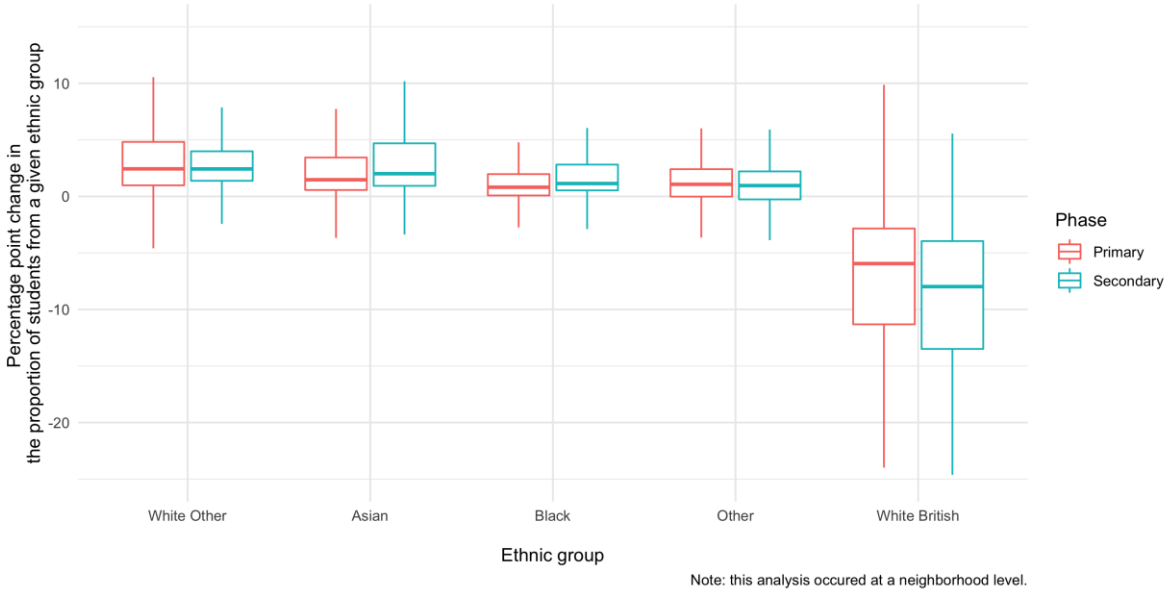
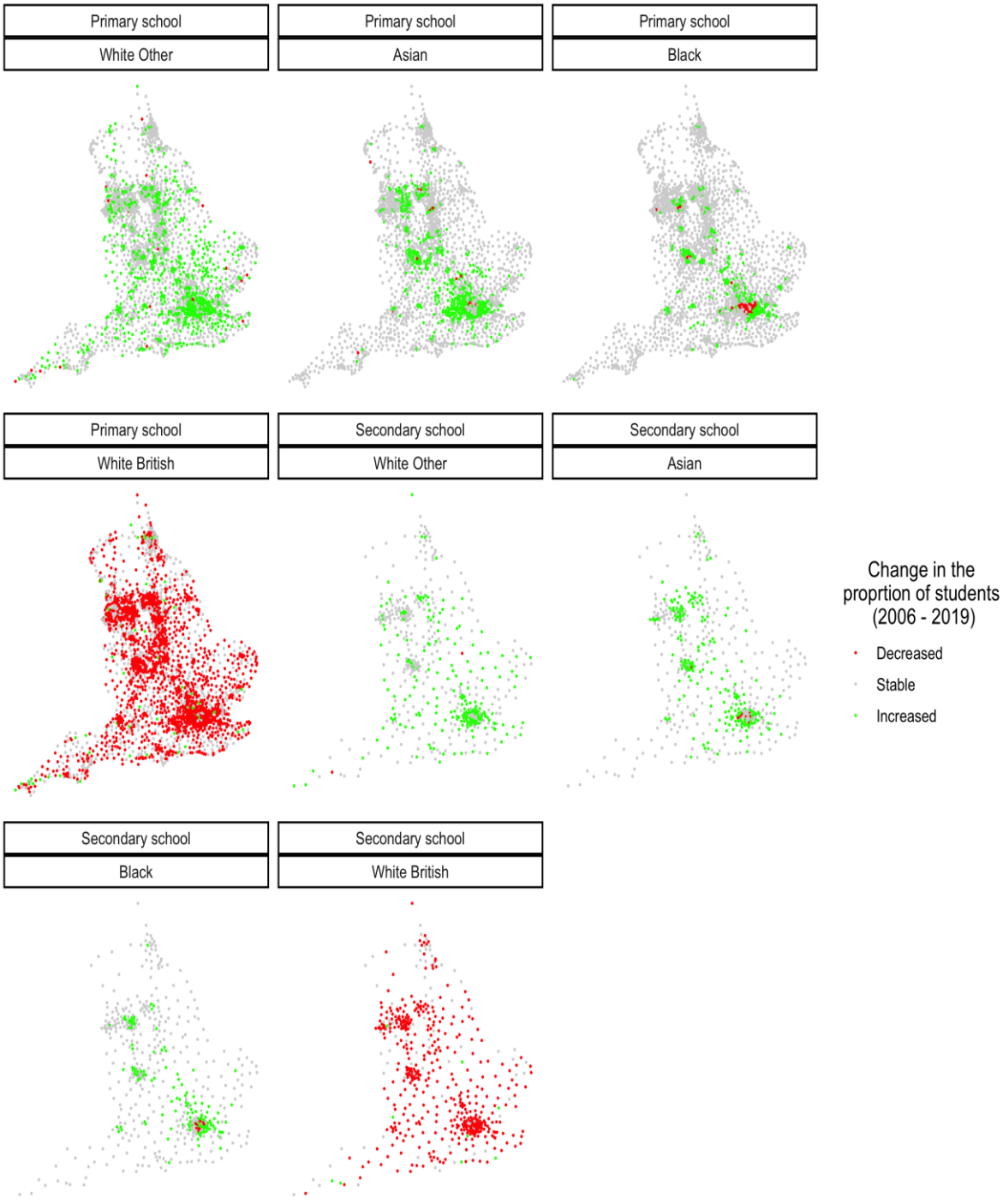

Percentage point change in the proportion of students from a given ethnic group
10
0
-10
-20
White Other
Asian
Black
Other
White British
Ethnic group
Phase
Primary
Secondary
Note: this analysis occured at a neighborhood level.

Figure 75: Mapping the change in the percentage of students from different ethnic groups a neighbourhood level (2006 – 2019)



Note: If the proportion of students from the ethnic group fell by 3 percentage points or more it was as "Decreased"
If the proportion did not change by more than 3 percentage points it was marked as "Stable"
If the proportion of rose by more than 3 percentage points it was marked as "Increased"

Figure 76: Comparing the proportion of students from each ethnic group within pseudo-neighbourhoods in 2006 and 2019

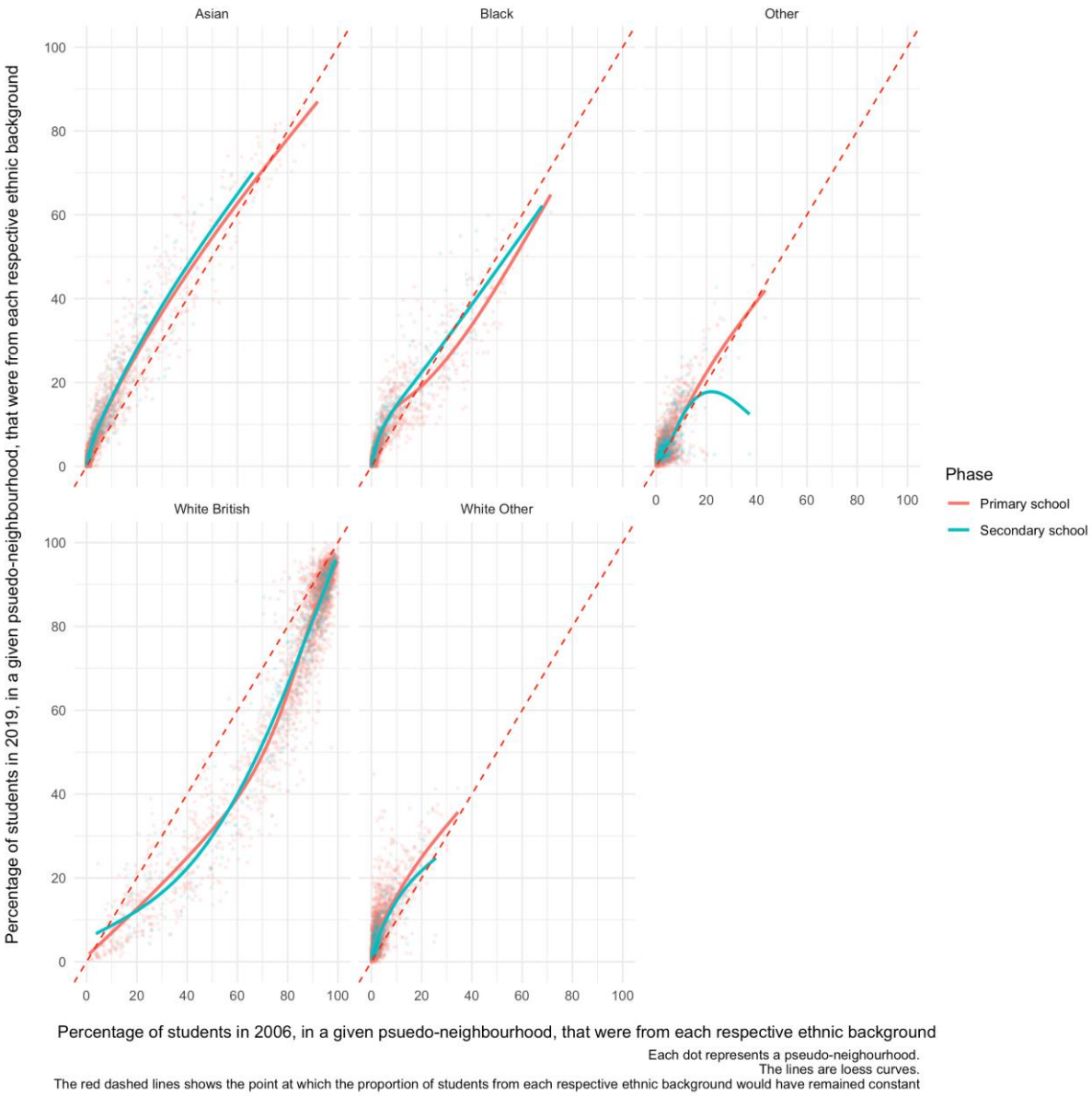


Figure 77: SPE values for students in 2019.

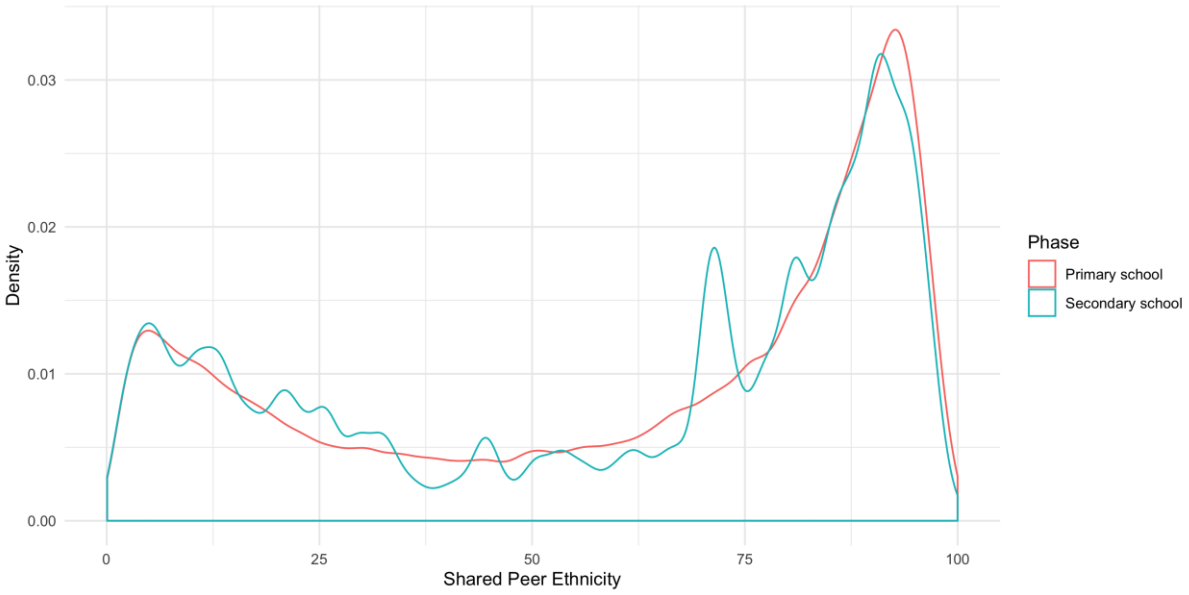


Figure 78: The distribution of SPE values for students attending school in England in 2006 and 2019.

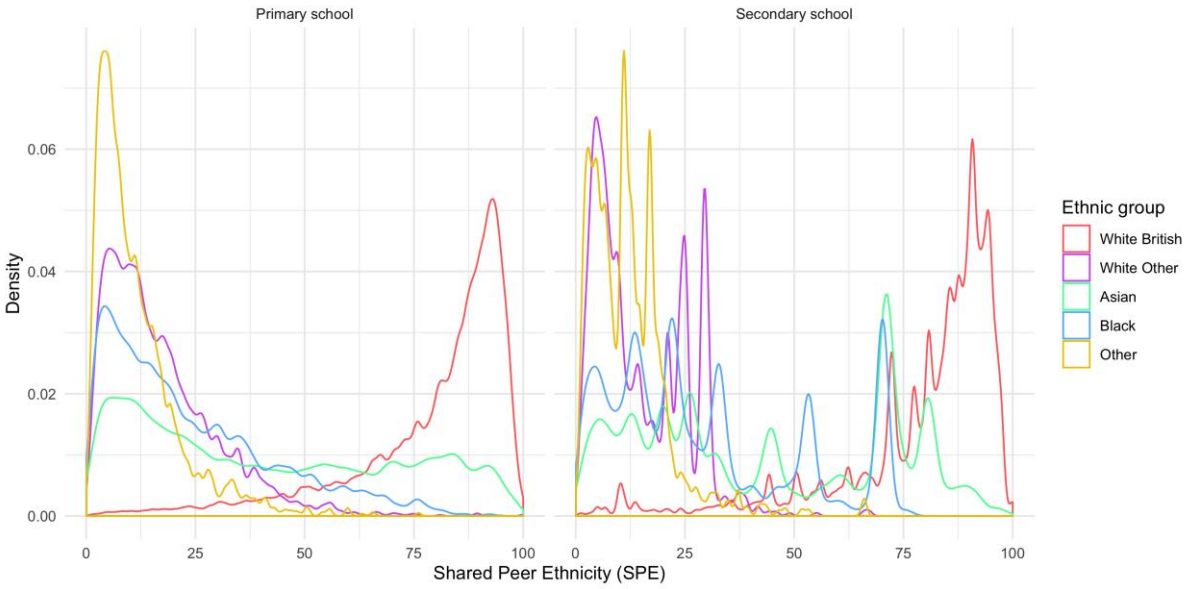


Figure 79: The *H* values for pseudo-neighbourhoods in the different regions of England (2016).

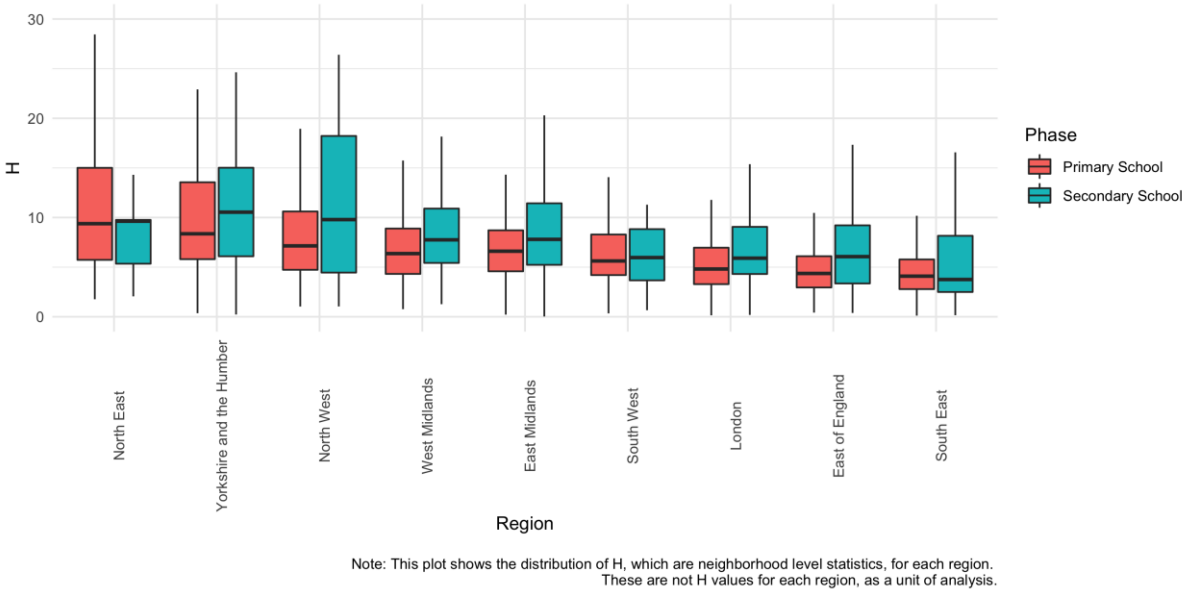
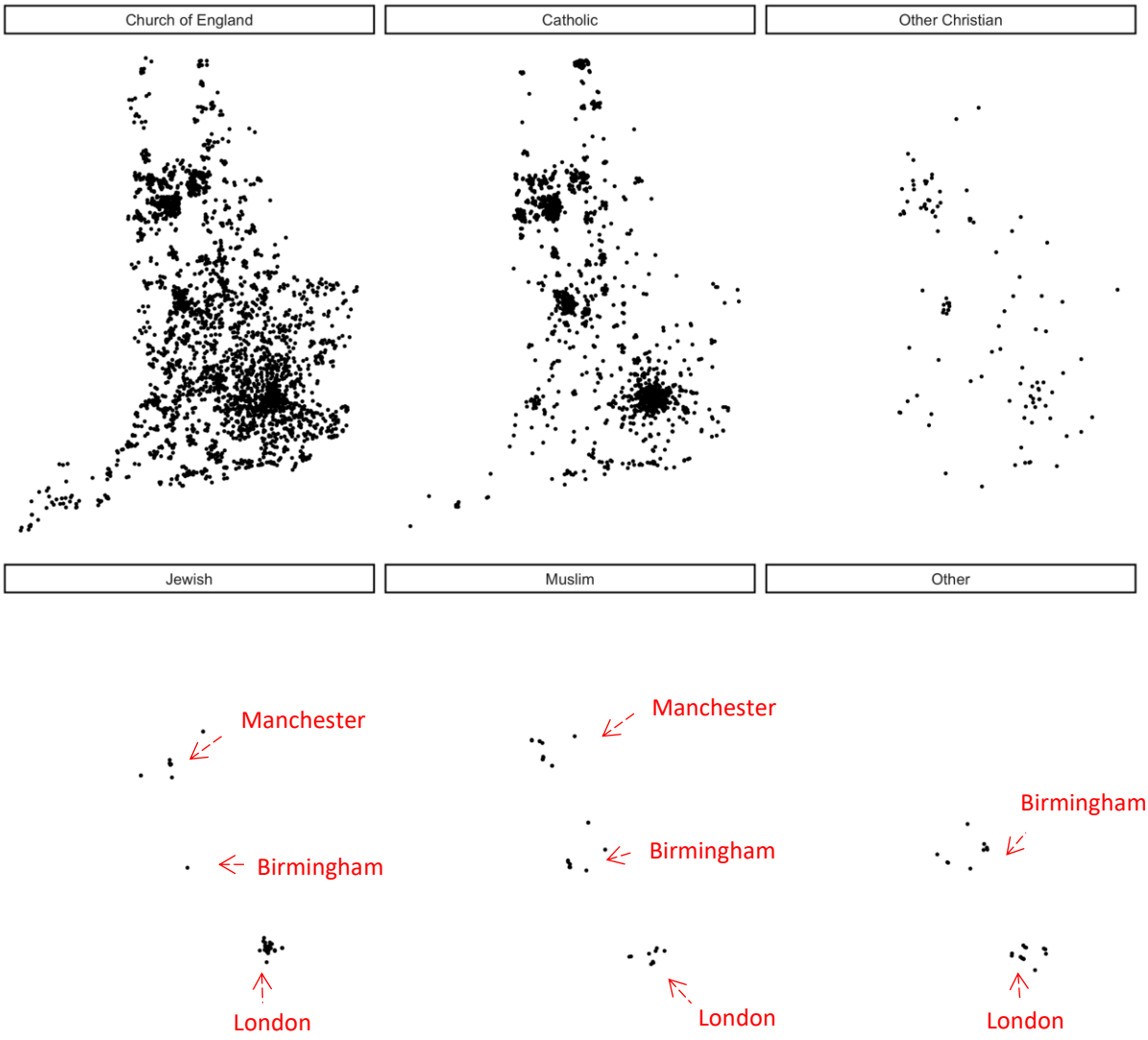
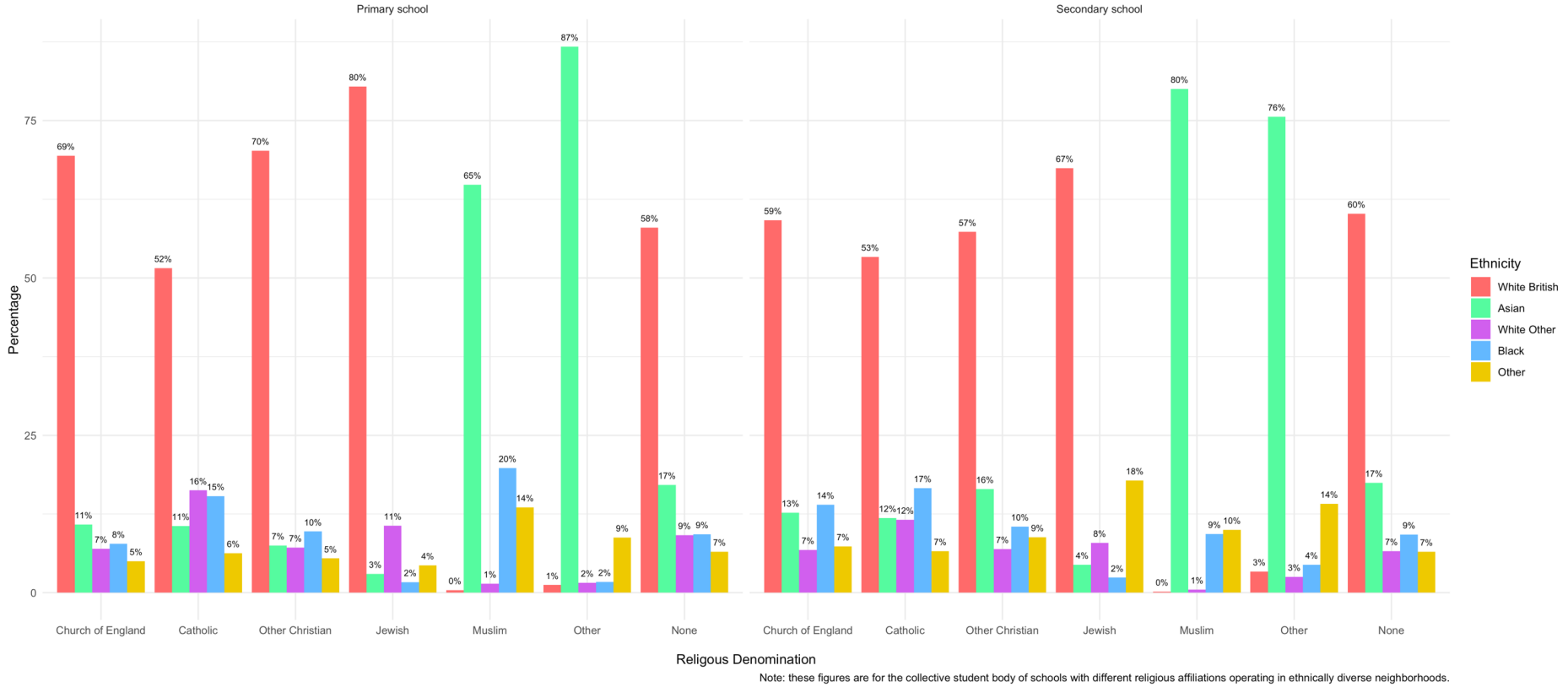


Figure 80: Mapping the location of religious schools in ethnically diverse pseudo-neighbourhoods (2019).



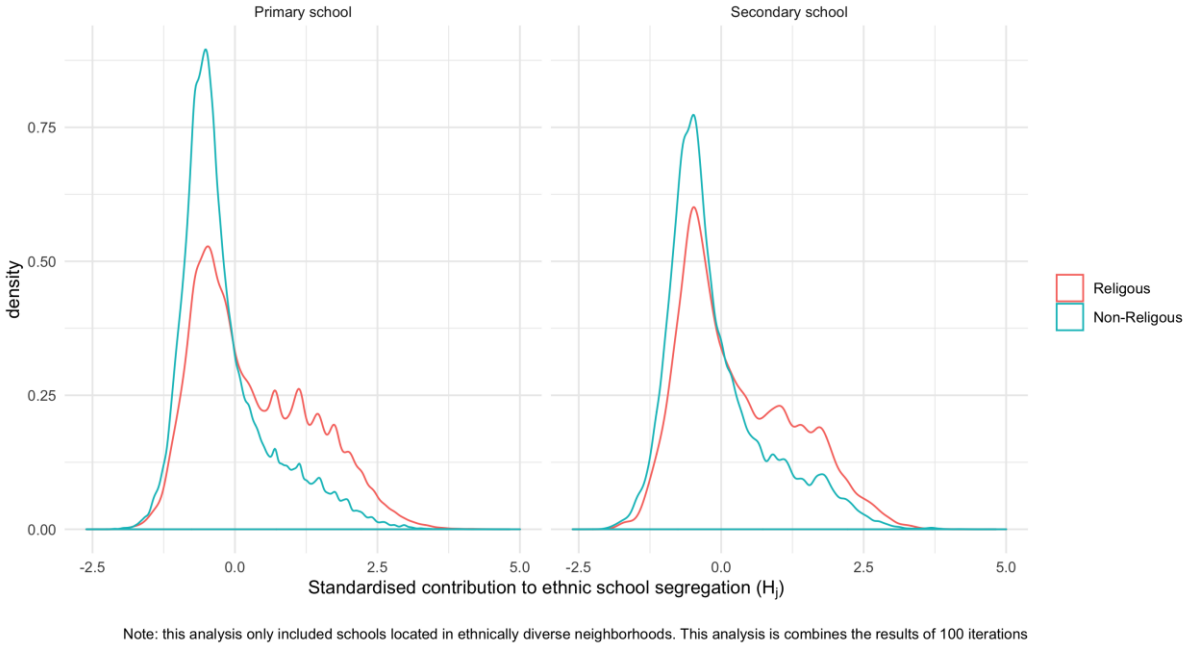
Note: this only includes religous schools in ethnicly diverse neighborhoods where at least one school has a religous denomination.

Figure 81: The collective ethnic profile of schools with different religous denomination, split by phase (2019)



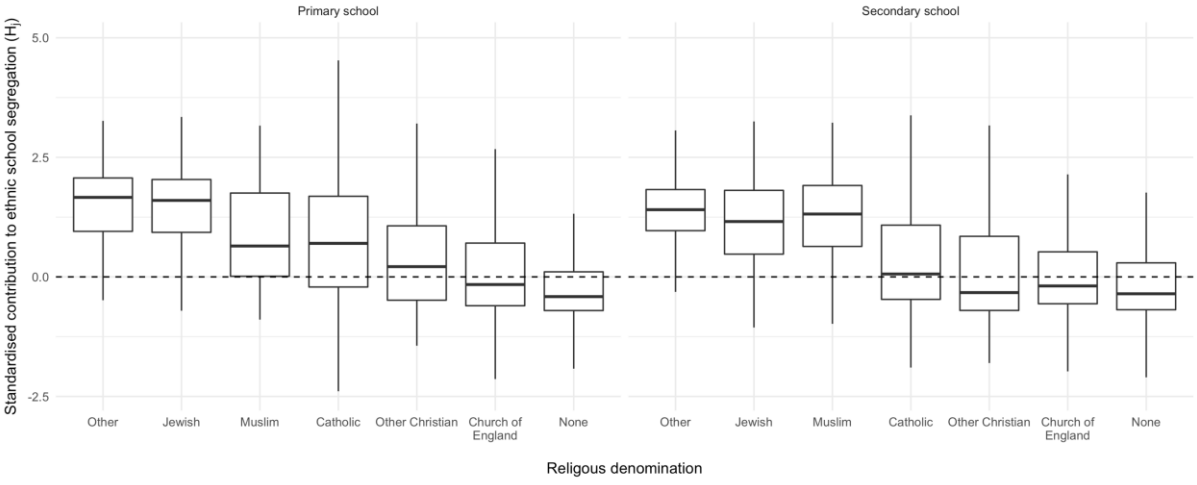
Note: these figures are for the collective student body of schools with different religious affiliations operating in ethnically diverse neighborhoods.

Figure 82: Comparing the distributions of H_j between religious and non-religious schools (2019).



Note: this analysis only included schools located in ethnically diverse neighborhoods. This analysis is combines the results of 100 iterations

Figure 83: Comparing the distributions of H_j for schools with different religious denominations (2019).



Note: this analysis only included schools located in ethnically diverse neighborhoods.
This analysis is combines the results of 100 iterations.
The dashed line shows the point at which schools make an average contribution to ethnic school segregation.

11.21 Full results for pseudo-neighbourhood level meta-regression model

The tables below detail the distribution of each statistic, in relation to each variable, that were included in the meta-regression models in Chapter 6. This allows readers to investigate how much variation was recorded in relation to each statistic, across the 100 different regression models used in the analysis. Substantively, this allows readers to understand the extent to which each statistic was impacted by the specific pseudo-neighbourhood grouping used. Fortunately, the statistics were both substantively consistent across all model parameters, and the I^2 value for all coefficients in models were 0%, which suggests a high degree of consistency across the findings of the 100 regression models (Higgins *et al.*, 2006; Borenstein *et al.*, 2009).

Table 35: Full meta-regression output for primary school psuedo-neighbourhood level predictors.

Variable	Statistic	Model 1_p (LQ - UQ)	Model 2_p (LQ - UQ)	Model 3_p (LQ - UQ)	Model 4_p (LQ - UQ)	Model 5_p (LQ - UQ)
Population Density	estimate		-0.018 (-0.028 to -0.01)	-0.013 (-0.027 to -0.001)	0.03 (0.017 to 0.043)	-0.053 (-0.068 to -0.034)
	std. error		0.018 (0.017 to 0.018)	0.022 (0.022 to 0.022)	0.023 (0.022 to 0.023)	0.025 (0.025 to 0.025)
	p.value		0.368 (0.115 to 0.545)	0.458 (0.191 to 0.791)	0.298 (0.063 to 0.442)	0.113 (0.007 to 0.174)

	VIF	1.119 (1.095 to 1.135)	1.911 (1.863 to 1.957)	2.107 (2.05 to 2.149)	2.673 (2.589 to 2.747)
	estimate		-0.169 (-0.179 to -0.157)	-0.182 (-0.192 to -0.174)	-0.15 (-0.165 to -0.135)
	std. error		0.02 (0.019 to 0.02)	0.019 (0.019 to 0.02)	0.032 (0.032 to 0.032)
House Price	p.value		0 (0 to 0)	0 (0 to 0)	0 (0 to 0)
	VIF		1.497 (1.471 to 1.52)	1.526 (1.499 to 1.551)	4.402 (4.338 to 4.473)
	estimate		0.068 (0.061 to 0.078)	0.065 (0.055 to 0.076)	0.052 (0.044 to 0.063)
	std. error		0.022 (0.022 to 0.022)	0.023 (0.022 to 0.023)	0.023 (0.022 to 0.023)
FSM	p.value		0.017 (0 to 0.006)	0.022 (0.001 to 0.014)	0.065 (0.005 to 0.055)
	VIF		1.932 (1.905 to 1.955)	2.058 (2.031 to 2.083)	2.185 (2.154 to 2.218)
	estimate			0.145 (0.137 to 0.154)	0.132 (0.124 to 0.142)
Proportion of Religious Schools	std. error			0.017 (0.017 to 0.017)	0.018 (0.018 to 0.018)
	p.value			0 (0 to 0)	0 (0 to 0)

					1.096	1.228
	VIF				(1.089 to 1.102)	(1.219 to 1.239)
	estimate				0.003 (-0.003 to 0.013)	-0.012 (-0.02 to -0.002)
	std. error				0.017 (0.017 to 0.017)	0.017 (0.017 to 0.017)
Average Ofsted Rating	p.value				0.552 (0.348 to 0.833)	0.467 (0.233 to 0.65)
	VIF				1.184 (1.173 to 1.195)	1.261 (1.248 to 1.274)
	estimate	0.271 (0.253 to 0.288)	0.277 (0.255 to 0.295)	0.234 (0.218 to 0.25)	0.234 (0.219 to 0.249)	0.261 (0.247 to 0.274)
Number of Schools in the Pseudo Neighbourhood	std. error	0.017 (0.017 to 0.017)	0.018 (0.018 to 0.018)	0.018 (0.018 to 0.018)	0.017 (0.017 to 0.018)	0.018 (0.017 to 0.018)
	p.value	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)
	VIF		1.119 (1.095 to 1.135)	1.177 (1.15 to 1.198)	1.18 (1.154 to 1.202)	1.26 (1.234 to 1.286)
	estimate					-0.114 (-0.147 to -0.078)
East of England	std. error					0.073 (0.072 to 0.074)
	p.value					0.189 (0.044 to 0.289)

London	estimate	0.282 (0.245 to 0.327)
	std. error	0.095 (0.094 to 0.096)
	p.value	0.012 (0.001 to 0.011)
Northeast	estimate	0.283 (0.244 to 0.345)
	std. error	0.123 (0.119 to 0.127)
	p.value	0.038 (0.007 to 0.05)
North West	estimate	0.164 (0.141 to 0.185)
	std. error	0.073 (0.072 to 0.074)
	p.value	0.045 (0.012 to 0.056)
South East	estimate	-0.169 (-0.196 to -0.134)
	std. error	0.071 (0.07 to 0.072)
	p.value	0.038 (0.006 to 0.058)

South West	estimate					0.139 (0.099 to 0.173)
	std. error					0.078 (0.077 to 0.079)
	p.value					0.131 (0.023 to 0.205)
West Midlands	estimate					0.042 (0.009 to 0.068)
	std. error					0.071 (0.07 to 0.072)
	p.value					0.533 (0.329 to 0.734)
Yorkshire and the Humber	estimate					0.298 (0.265 to 0.328)
	std. error					0.076 (0.075 to 0.077)
	p.value					0 (0 to 0)
Intercept	estimate	1.615 (1.607 to 1.625)	1.615 (1.607 to 1.625)	1.616 (1.608 to 1.625)	1.615 (1.606 to 1.624)	1.549 (1.524 to 1.576)
	std. error	0.017 (0.016 to 0.017)	0.017 (0.016 to 0.017)	0.016 (0.016 to 0.016)	0.016 (0.015 to 0.016)	0.057 (0.056 to 0.059)
	p.value	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)

Observations	estimate	1820.28 (1811 to 1831)	1820.28 (1811 to 1831)	1820.28 (1811 to 1831)	1815.48 (1806 to 1825.25)	1815.48 (1806 to 1825.25)
R Squared Value	estimate	0.122 (0.108 to 0.134)	0.123 (0.109 to 0.136)	0.193 (0.182 to 0.203)	0.224 (0.216 to 0.237)	0.265 (0.254 to 0.277)
Adjusted R Squared	estimate	0.122 (0.108 to 0.134)	0.122 (0.108 to 0.135)	0.191 (0.18 to 0.202)	0.221 (0.213 to 0.234)	0.259 (0.249 to 0.271)
Adjusted Standard Error	estimate	0.707 (0.701 to 0.712)	0.707 (0.701 to 0.712)	0.679 (0.672 to 0.683)	0.665 (0.658 to 0.672)	0.649 (0.643 to 0.656)
F Statistic	estimate	254.536 (220.326 to 282.444)	128.131 (110.23 to 143.585)	108.563 (100.371 to 115.289)	87.089 (82.254 to 93.079)	46.465 (43.547 to 49.277)
Breusch Pagan test	estimate	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)
Mean of the residuals	estimate	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)
Shapiro Wilks Test	estimate	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)
Cooks Distance	estimate	0.028 (0.025 to 0.029)	0.03 (0.021 to 0.03)	0.029 (0.023 to 0.032)	0.026 (0.02 to 0.03)	0.019 (0.016 to 0.019)

Table 36: Full meta-regression output for secondary school pseudo-neighbourhood level predictors.

Variable	**Statistic**	**Model 1_s (LQ - UQ)**	**Model 2_s (LQ - UQ)**	**Model 3_s (LQ - UQ)**	**Model 4_s (LQ - UQ)**	**Model 5_s (LQ - UQ)**
Population Density	estimate		0.1 (0.08 to 0.116)	0.101 (0.075 to 0.134)	0.062 (0.034 to 0.092)	0.026 (-0.013 to 0.055)

		0.043	0.056	0.057	0.058
	std. error	(0.042 to 0.044)	(0.054 to 0.058)	(0.056 to 0.059)	(0.056 to 0.059)
	p.value	0.047 (0.007 to 0.066)	0.134 (0.018 to 0.19)	0.365 (0.108 to 0.564)	0.531 (0.331 to 0.76)
	VIF	1.126 (1.087 to 1.158)	2.253 (2.175 to 2.332)	2.396 (2.322 to 2.474)	2.619 (2.527 to 2.7)
	estimate		-0.228 (-0.245 to -0.201)	-0.229 (-0.247 to -0.202)	-0.229 (-0.267 to -0.192)
	std. error		0.048 (0.046 to 0.049)	0.049 (0.047 to 0.05)	0.08 (0.078 to 0.082)
House Price	p.value		0.000 (0 to 0)	0.000 (0 to 0)	0.025 (0.001 to 0.016)
	VIF		1.621 (1.578 to 1.659)	1.747 (1.698 to 1.789)	5.034 (4.911 to 5.166)
	estimate		0.176 (0.149 to 0.204)	0.214 (0.188 to 0.238)	0.133 (0.106 to 0.156)
	std. error		0.052 (0.05 to 0.053)	0.054 (0.052 to 0.056)	0.057 (0.055 to 0.059)
FSM	p.value		0.006 (0 to 0.004)	0.001 (0 to 0.001)	0.056 (0.006 to 0.067)
	VIF		1.942 (1.879 to 2.011)	2.181 (2.124 to 2.249)	2.594 (2.513 to 2.663)

Proportion of Religious Schools	estimate	0.075 (0.055 to 0.093)	0.059 (0.04 to 0.079)
	std. error	0.043 (0.041 to 0.043)	0.043 (0.042 to 0.044)
	p.value	0.157 (0.03 to 0.181)	0.255 (0.071 to 0.377)
	VIF	1.124 (1.106 to 1.142)	1.253 (1.225 to 1.28)
Does the Pseudo-neighbourhood Contain a Grammar School? (Yes = 1)	estimate	0.24 (0.205 to 0.275)	0.264 (0.231 to 0.3)
	std. error	0.095 (0.093 to 0.098)	0.094 (0.091 to 0.097)
	p.value	0.026 (0.004 to 0.031)	0.015 (0.001 to 0.014)
	VIF	1.091 (1.081 to 1.102)	1.147 (1.135 to 1.158)
Average Ofsted Rating	estimate	0.039 (0.016 to 0.067)	0.005 (-0.016 to 0.031)
	std. error	0.042 (0.041 to 0.042)	0.041 (0.04 to 0.042)
	p.value	0.359 (0.11 to 0.559)	0.519 (0.269 to 0.767)

	VIF				1.308 (1.282 to 1.331)	1.39 (1.36 to 1.419)
Number of Schools in the Pseudo Neighbourhood	estimate	0.456 (0.429 to 0.49)	0.424 (0.393 to 0.453)	0.361 (0.335 to 0.387)	0.339 (0.311 to 0.364)	0.36 (0.33 to 0.386)
	std. error	0.041 (0.04 to 0.042)	0.044 (0.042 to 0.044)	0.041 (0.04 to 0.042)	0.042 (0.04 to 0.043)	0.042 (0.041 to 0.043)
	p.value	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)
East of England	VIF		1.126 (1.087 to 1.158)	1.193 (1.147 to 1.241)	1.24 (1.187 to 1.303)	1.362 (1.308 to 1.416)
	estimate					-0.031 (-0.092 to 0.041)
	std. error					0.173 (0.168 to 0.18)
	p.value					0.635 (0.488 to 0.821)
London	estimate					0.261 (0.197 to 0.336)
	std. error					0.228 (0.22 to 0.234)
	p.value					0.304 (0.129 to 0.413)
North East	estimate					-0.322 (-0.428 to -0.228)

	std. error	0.277 (0.256 to 0.292)
	p.value	0.29 (0.122 to 0.416)
	estimate	0.049 (-0.02 to 0.124)
North West	std. error	0.17 (0.164 to 0.176)
	p.value	0.612 (0.401 to 0.855)
	estimate	-0.375 (-0.439 to -0.313)
South East	std. error	0.173 (0.167 to 0.179)
	p.value	0.061 (0.012 to 0.078)
	estimate	-0.281 (-0.358 to -0.205)
South West	std. error	0.192 (0.184 to 0.2)
	p.value	0.192 (0.07 to 0.276)
West Midlands	estimate	-0.017 (-0.076 to 0.05)

	std. error					0.163 (0.158 to 0.168)
	p.value					0.637 (0.438 to 0.8)
	estimate					0.272 (0.182 to 0.353)
Yorkshire and the Humber	std. error					0.179 (0.172 to 0.186)
	p.value					0.21 (0.053 to 0.295)
Intercept	estimate	1.655 (1.638 to 1.676)	1.656 (1.639 to 1.678)	1.656 (1.639 to 1.68)	1.607 (1.588 to 1.632)	1.632 (1.578 to 1.688)
	std. error	0.04 (0.039 to 0.041)	0.04 (0.039 to 0.04)	0.037 (0.036 to 0.037)	0.041 (0.04 to 0.042)	0.139 (0.133 to 0.145)
	p.value	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)
Observations	estimate	397.5 (392 to 403)	397.5 (392 to 403)	397.5 (392 to 403)	397.18 (392 to 402.25)	397.18 (392 to 402.25)
R Squared Value	estimate	0.236 (0.211 to 0.264)	0.248 (0.225 to 0.27)	0.364 (0.341 to 0.386)	0.383 (0.358 to 0.404)	0.436 (0.415 to 0.455)
Adjusted R Squared	estimate	0.235 (0.209 to 0.262)	0.244 (0.221 to 0.266)	0.357 (0.334 to 0.38)	0.372 (0.347 to 0.394)	0.414 (0.392 to 0.433)
Adjusted Standard Error	estimate	0.795 (0.774 to 0.816)	0.79 (0.769 to 0.809)	0.728 (0.709 to 0.746)	0.72 (0.699 to 0.735)	0.695 (0.676 to 0.714)

F Statistic	estimate	124.242 (104.991 to 140.666)	65.767 (56.641 to 73.003)	56.478 (50.384 to 62.072)	34.72 (31.211 to 37.937)	19.823 (17.963 to 21.135)
Breusch Pagan test	estimate	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.001 (0 to 0)	0.011 (0 to 0.008)
Mean of the residuals	estimate	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)
Shapiro Wilks Test	estimate	0.481 (0.174 to 0.743)	0.485 (0.222 to 0.739)	0.392 (0.073 to 0.639)	0.391 (0.089 to 0.657)	0.375 (0.073 to 0.593)
Cooks Distance	estimate	0.09 (0.067 to 0.092)	0.066 (0.048 to 0.067)	0.085 (0.046 to 0.081)	0.109 (0.048 to 0.084)	0.077 (0.035 to 0.055)

Table 37: Full meta-regression output for Model 6_S

Variable	Coefficient	Standard Error	P Value
Grammar School	0.142 (0.105 to 0.176)	0.088 (0.087 to 0.088)	0.17 (0.046 to 0.228)
(Intercept)	-0.036 (-0.044 to -0.026)	0.044 (0.044 to 0.045)	0.435 (0.312 to 0.557)
Observations	589.55 (576 to 600)		
R Squared Value	0.005 (0.002 to 0.007)		
Adjusted R Squared	0.003 (0.001 to 0.005)		
Adjusted Standard Error	0.931 (0.928 to 0.933)		

F Statistic	3.078 (1.456 to 4.004)		
Breusch Pagan test	0.45 (0.109 to 0.681)		
Mean of the residuals	0 (0 to 0)		
Shapiro Wilks Test	0 (0 to 0)		
Cooks Distance	0.025 (0.022 to 0.028)		

Table 38: Full meta-regression output for Model 7_p, 8_p, 7_s and 8_s.

Variable	**Statistic**	**Model 7_p** **(LQ - UQ)**	**Model 8_p** **(LQ - UQ)**	**Model 7_s** **(LQ - UQ)**	**Model 8_s** **(LQ - UQ)**
Religious school	estimate	0.511 (0.504 to 0.517)		0.372 (0.353 to 0.391)	
	std. error	0.018 (0.018 to 0.018)		0.045 (0.045 to 0.045)	
	p.value	0.000 (0 to 0)		0.000 (0 to 0)	
Church of England	estimate		0.283 (0.275 to 0.29)		0.142 (0.115 to 0.167)
	std. error		0.02 (0.02 to 0.02)		0.069 (0.068 to 0.069)
	p.value		0.000 (0 to 0)		0.065 (0.015 to 0.095)
Catholic	estimate		0.944 (0.933 to 0.956)		0.425 (0.401 to 0.453)
	std. error		0.027 (0.026 to 0.027)		0.06 (0.059 to 0.06)

	p.value		0.000		0.000
			(0 to 0)		(0 to 0)
	estimate		0.49		0.105
			(0.437 to 0.539)		(0.049 to 0.153)
Other Christian	std. error		0.116		0.131
			(0.115 to 0.118)		(0.131 to 0.133)
	p.value		0.000		0.469
			(0 to 0)		(0.252 to 0.697)
	estimate		1.65		1.176
			(1.593 to 1.722)		(1.08 to 1.281)
Jewish	std. error		0.145		0.252
			(0.145 to 0.145)		(0.252 to 0.253)
	p.value		0.000		0.000
			(0 to 0)		(0 to 0)
	estimate		1.003		1.34
			(0.891 to 1.12)		(1.255 to 1.443)
Muslim	std. error		0.242		0.199
			(0.241 to 0.242)		(0.198 to 0.2)
	p.value		0.000		0.000
			(0 to 0)		(0 to 0)
	estimate		1.683		1.422
			(1.599 to 1.771)		(1.318 to 1.546)
Other Non-Christian	std. error		0.241		0.263
			(0.24 to 0.241)		(0.262 to 0.263)
	p.value		0.000		0.000
			(0 to 0)		(0 to 0)
	estimate	-0.189	-0.189	-0.094	-0.094
		(-0.192 to -0.187)	(-0.192 to -0.187)	(-0.099 to -0.089)	(-0.099 to -0.089)
	std. error	0.011	0.01	0.023	0.022
(Intercept)		(0.011 to 0.011)	(0.01 to 0.01)	(0.022 to 0.023)	(0.022 to 0.022)
	p.value	0	0	0	0
		(0 to 0)	(0 to 0)	(0 to 0)	(0 to 0)
	estimate	10972.31	10972.31	2210.59	2210.59
Observations		(10923 to	(10923 to	(2190.25 to	(2190.25 to
		11015.25)	11015.25)	2235.25)	2235.25)

R Squared Value	estimate	0.071 (0.069 to 0.073)	0.119 (0.116 to 0.121)	0.03 (0.027 to 0.033)	0.06 (0.055 to 0.064)
Adjusted R Squared	estimate	0.071 (0.069 to 0.073)	0.118 (0.116 to 0.121)	0.03 (0.027 to 0.033)	0.057 (0.053 to 0.061)
Adjusted Standard Error	estimate	0.891 (0.89 to 0.892)	0.868 (0.866 to 0.869)	0.92 (0.918 to 0.921)	0.906 (0.904 to 0.909)
F Statistic	estimate	843.329 (815.689 to 866.54)	246.844 (240.674 to 251.598)	68.713 (61.14 to 74.889)	23.416 (21.529 to 25.128)
Breusch Pagan test	estimate	0.000 (0 to 0)	0.000 (0 to 0)	0.003 (0 to 0.002)	0.144 (0.032 to 0.215)
Mean of the residuals	estimate	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)
Cooks Distance	estimate	0.002 (0.002 to 0.003)	0.046 (0.038 to 0.052)	0.01 (0.009 to 0.012)	0.071 (0.057 to 0.079)

Table 39: Full meta-regression output for Model 9_p and 9_s

Variable	**Statistic**	**Model 9_p**	**Model 9_s**
Outstanding	estimate	-0.029 (-0.065 to 0.001)	-0.135 (-0.159 to -0.112)
	std. error	0.096 (0.096 to 0.097)	0.069 (0.069 to 0.07)
	p.value	0.655 (0.487 to 0.87)	0.085 (0.022 to 0.107)
Good	estimate	-0.088 (-0.122 to -0.062)	-0.117 (-0.136 to -0.1)
	std. error	0.089 (0.088 to 0.09)	0.066 (0.066 to 0.067)
	p.value	0.363 (0.166 to 0.493)	0.114 (0.039 to 0.132)

Requires improvement	estimate	0.026 (-0.003 to 0.052)	-0.127 (-0.147 to -0.102)
	std. error	0.098 (0.098 to 0.099)	0.072 (0.072 to 0.072)
	p.value	0.715 (0.567 to 0.909)	0.112 (0.041 to 0.151)
(Intercept)	estimate	0.039 (0.014 to 0.068)	0.123 (0.104 to 0.143)
	std. error	0.086 (0.085 to 0.087)	0.066 (0.065 to 0.066)
	p.value	0.607 (0.423 to 0.789)	0.093 (0.028 to 0.109)
Observations	estimate	2386.64 (2375.75 to 2398.25)	10688.81 (10650.75 to 10718.25)
R Squared Value	estimate	0.003 (0.002 to 0.003)	0 (0 to 0.001)
Adjusted R Squared	estimate	0.001 (0.001 to 0.002)	0 (0 to 0)
Adjusted Standard Error	estimate	0.918 (0.916 to 0.921)	0.923 (0.922 to 0.925)
F Statistic	estimate	2.094 (1.484 to 2.655)	1.618 (1.076 to 2.225)
Breusch Pagan test	estimate	0.402 (0.225 to 0.585)	0.065 (0.007 to 0.077)
Mean of the residuals	estimate	0.000 (0 to 0)	0.000 (0 to 0)
Cooks Distance	estimate	0.021 (0.019 to 0.023)	0.014 (0.013 to 0.015)

11.22 Model checks for pseudo-neighbourhood level meta-regression analysis

The table below outlines the results from model checks completed as part of the analysis in Section 6.1. For all the pseudo-neighbourhood level models the mean of the residuals was 0

(3s.f) and the largest Cooks D Value was small (<0.2). Both of these results are reassuring. However, the Breusch–Pagan test and the Shapiro Wilks Test both were statistically significant, suggesting that the residuals were neither homoscedastic nor normally distributed. Consequently, the analysis employed heteroscedasticity-consistent standard errors to ensure that the models were correctly specified. Visual inspection of scatter plots and a pre-test correlation matrix raised no further concerns with the models' specification.

Table 40: Results from the model checks for the meta regression analysis on primary school pseudo-neighbourhood level models.

Variable	Model 1_p (LQ - UQ)	Model 2_p (LQ - UQ)	Model 3_p (LQ - UQ)	Model 4_p (LQ - UQ)	Model 5_p (LQ - UQ)
Breusch Pagan test	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)
Mean of the residuals	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)
Shapiro Wilks Test	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)
Cooks Distance	0.028 (0.025 to 0.029)	0.03 (0.021 to 0.03)	0.029 (0.023 to 0.032)	0.026 (0.02 to 0.03)	0.019 (0.016 to 0.019)

Note: all figures in this table are average results across 100 models.

Table 41: Results from the model checks for the meta regression analysis on secondary school pseudo-neighbourhood level models

Variable	Model 1_s (LQ - UQ)	Model 2_s (LQ - UQ)	Model 3_s (LQ - UQ)	Model 4_s (LQ - UQ)	Model 5_s (LQ - UQ)
Breusch Pagan test	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.001 (0 to 0)	0.011 (0 to 0.008)
Mean of the residuals	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)
Shapiro Wilks Test	0.481 (0.174 to 0.743)	0.485 (0.222 to 0.739)	0.392 (0.073 to 0.639)	0.391 (0.089 to 0.657)	0.375 (0.073 to 0.593)
Cooks Distance	0.09 (0.067 to 0.092)	0.066 (0.048 to 0.067)	0.085 (0.046 to 0.081)	0.109 (0.048 to 0.084)	0.077 (0.035 to 0.055)

Note: all figures in this table are average results across 100 models.

11.23 Model checks for school level meta-regression analysis

The table above outlines the results from model checks completed as part of the analysis in Section 6.1. For all models, the mean of the residuals was 0 (3s.f) and the largest Cooks D Value was comparatively small (<0.075). However, the Shapiro Wilks Test for all models was statistically significant, suggesting that the residuals were not normally distributed, and for models 7_p, 8_p and 7_s the Breusch–Pagan test was statistically significant which suggests that the residuals were not homoscedastic. Consequently, the analysis employed heteroscedasticity-consistent standard errors to ensure that the models were correctly specified. Visual inspection of graphs and pre-test correlation matrix raised no further concerns with the model specification.

Table 42: Results from the model checks for the meta regression analysis on school level predictors (section 4.2)

Variable	Model 7_p (LQ - UQ)	Model 8_p (LQ - UQ)	Model 7_s (LQ - UQ)	Model 8_s (LQ - UQ)	Model 9p (LQ - UQ)	Model 9s (LQ - UQ)
Breusch Pagan test	0.000 (0 to 0)	0.000 (0 to 0)	0.003 (0 to 0.002)	0.144 (0.032 to 0.215)	0.402 (0.225 to 0.585)	0.065 (0.007 to 0.077)
Mean of the residuals	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)	0.000 (0 to 0)
Cooks Distance	0.002 (0.002 to 0.003)	0.046 (0.038 to 0.052)	0.01 (0.009 to 0.012)	0.071 (0.057 to 0.079)	0.021 (0.019 to 0.023)	0.014 (0.013 to 0.015)

11.24 Contrasting the raw and ranked contributions towards ethnic school segregation (H_j) of schools with selective schools, faith schools and schools with different OFSTED ratings.

To ensure that the standardisation of H_j values did not unduly impact the substantive findings of the school level analysis, as outlined in Section 4.5.2, the thesis reran the school level meta-regression models using two different versions of H_j. The first set of models used the untransformed raw H_j values and the second ranked each school's H_j value, relative to other schools in the same pseudo-neighbourhood. In both instances, the study controlled for the number of schools in each pseudo-neighbourhood, as the size of pseudo-neighbourhoods influenced the result. The equations of these models are listed below and closely reflect those used in the main body of the thesis.

Table 43: Formulas for school level meta-regression analysis

Model 10	$H_RAW_{jn} = \beta_0 + \beta_1 SELECTIVE_{jn} + NUMSCH_n + \varepsilon_{jn}$
Model 11	$H_RANK_j = \beta_0 + \beta_1 SELECTIVE_{jn} + NUMSCH_n + \varepsilon_{jn}$
Null Hypothesis	$\beta_1 = 0$

Model 12	$H_RAW_{jn} = \beta_0 + \beta_1 RELIGOUS_{jn} + NUMSCH_{jn} + \varepsilon_{jn}$
Model 13	$H_RANK_{jn} = \beta_0 + \beta_1 RELIGOUS_{jn} + NUMSCH_n + \varepsilon_{jn}$
Null Hypothesis	$\beta_1 = 0$
Model 14	$H_RAW_{jn} = \beta_0 + \beta_1 CHURCHENG_{jn} + \beta_2 CATH_{jn} + \beta_3 CHRISTOTHER_{jn} + \beta_4 JEWISH_{jn} + \beta_5 MUSLIM_{jn} + \beta_6 OTHER_{jn} + NUMSCH_n + \varepsilon_{jn}$
Model 15	$H_RANK_{jn} = \beta_0 + \beta_1 CHURCHENG_{jn} + \beta_2 CATH_{jn} + \beta_3 CHRISTOTHER_{jn} + \beta_4 JEWISH_{jn} + \beta_5 MUSLIM_{jn} + \beta_6 OTHER_{jn} + NUMSCH_n + \varepsilon_{jn}$
Null Hypothesis	$\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6$

Model 16	$H_RAW_{jn} = \beta_0 + \beta_1 OUTSTANDING_{jn} + \beta_2 GOOD_{jn} + \beta_3 REQIMPROVE_{jn} + NUMSCH_n + \varepsilon_j$
Model 17	$H_RANK_{jn} = \beta_0 + \beta_1 OUTSTANDING_{jn} + \beta_2 GOOD_{jn} + \beta_3 REQIMPROVE_{jn} + NUMSCH_n + \varepsilon_j$
Null Hypothesis	$\beta_1 = \beta_2 = \beta_3 = 0$

Where:

H_RAW_{jn}	The raw contribution of school *j* toward their pseudo-neighbourhood's ethnic school segregation.
H_RANK_{jn}	The ranked contribution of school *j* toward their pseudo-neighbourhood's ethnic school segregation.
$SELECTIVE_j$	1 = school *j* has a selective admissions policy (Grammar School)
$RELIGOUS_j$	1 = school *j* has a religious denomination
$CHURCHENG_j$	1 = school *j* has an affiliation to the Church of England
$CATH_j$	1 = school *j* has an affiliation to the Catholic faith
$CHRISTOTHER_j$	1 = school *j* has an affiliation to another Christian Faith
$JEWISH_j$	1 = school *j* has an affiliation to a Jewish faith
$MUSLIM_j$	1 = school *j* has an affiliation to an Islamic faith
$OTHER_j$	1 = school *j* has an affiliation to another non-Christian Faith
$OUTSTANDING_j$	1 = school *j* received a "Outstanding" OFSTED rating in its last review
$GOOD_j$	1 = school *j* received a "Good" OFSTED rating in its last review
$REQIMPROVE_j$	1 = school *j* received a "Requires Improvement" OFSTED rating in its last review
$NUMSCH_n$	Number of schools in pseudo-neighbourhood *n*

The regression tables below report the models' findings. Fortunately, the majority the results from these meta-regression models reflect those outlined in the main body of the text, which suggests that the findings were reliable. Model while 10_s found a statistically significant difference between the ranked H_j values of grammar schools and non-selective secondary schools, but model 11_s showed no statistically significant difference between the Raw H_j values. Model 12_p, 13_p, 12_s and 13_s, all showed a statistically significant difference between the raw and ranked versions of Hj values of religious and non-religious schools. Similarly, Model 14_p 15_p, 14_s and 15_s all found that the raw and ranked H_j values of CoE schools tended to be smaller than those linked to minority religious groups. Finally, Model

16_p,17_p, 16_s and 17_s found no statistically significant difference between the raw or ranked H_j values of schools with different OFSTED ratings.

Table 44: Comparing the raw and ranked contributions towards ethnic school segregation (H_j) of grammar school and non-selective secondary school

	Model 10_s (Ranked Hj)	**Model 11_s (Raw Hj)**
Grammar School	0.679** (0.254)	0.002 (0.001)
Number of Schools	0.514*** (0.038)	-0.001*** (0)
(Intercept)	0.246 (0.352)	0.018*** (0.002)
Observations	541.9	541.9
R Squared Value	0.256	0.046
Adjusted R Squared	0.253	0.043
Adjusted Standard Error	2.584	0.014
F Statistic	93.153	13.143

Table 45: Comparing the raw and ranked contributions towards ethnic school segregation (Hj) of different religious schools.

Variable	Model 12_p (Ranked Hj)	Model 13_p (Raw Hj)	Model 12_s (Ranked Hj)	Model 13_s (Raw Hj)	Model 14_p (Ranked Hj)	Model 15_p (Raw Hj)	Model 14_s (Ranked Hj)	Model 15_s (Raw Hj)
Religious	0.664*** (0.048)	0.812*** (0.134)	0.003*** (0)	0.004*** (0.001)				
Church of England					0.203*** (0.055)	0.293 (0.206)	0.001** (0)	0 (0.001)
Catholic					1.515*** (0.073)	1.112*** (0.179)	0.007*** (0)	0.003*** (0.001)
Other Christian					0.589 (0.319)	0.299 (0.39)	0.002 (0.002)	0 (0.002)
Jewish					2.636*** (0.394)	1.483* (0.746)	0.043*** (0.002)	0.035*** (0.004)
Muslim					2.213*** (0.664)	1.538* (0.629)	0.021*** (0.004)	0.03*** (0.003)
Other Non-Christian					2.713*** (0.66)	2.638*** (0.787)	0.036*** (0.004)	0.015** (0.004)
Number of Schools	0.513*** (0.008)	0.512*** (0.021)	-0.001*** (0)	-0.001*** (0)	0.503*** (0.008)	0.506*** (0.021)	-0.001*** (0)	-0.001*** (0)
(Intercept)	0.153** (0.071)	0.376* (0.186)	0.016*** (0)	0.015*** (0.001)	0.233*** (0.07)	0.426* (0.186)	0.016*** (0)	0.015*** (0.001)
Observations	10459	2080	10459	2080	10459	2080	10459	2080
R Squared Value	0.284	0.234	0.043	0.026	0.304	0.243	0.086	0.092
Adjusted R Squared	0.284	0.234	0.043	0.025	0.304	0.24	0.085	0.089
Adjusted Standard Error	2.382	2.651	0.015	0.015	2.349	2.64	0.015	0.015
F Statistic	2075	318.2	233.4	28.15	652.4	94.88	140.5	30.25

Table 46: Comparing the raw and ranked contributions towards ethnic school segregation (Hj) of schools with different OFSTED results

Variables	Model 16_p (Ranked Hj)	Model 17_p (Raw Hj)	Model 16_s (Ranked Hj)	Model 17_s (Raw Hj)
	-0.121	0.681**		0.002
Outstanding	(0.181)	(0.27)	-0	(0.002)
	-0.2	0.29	-0.002	0
Good	(0.173)	(0.25)	(0.001)	(0.002)
Requires	-0.286	0.439	-0.001	0.001
improvement	(0.187)	(0.276)	(0.001)	(0.002)
	0.501***	0.506***	-0.001***	-0.001**
Number of Schools	(0.008)	(0.018)	(0)	(0)
	0.687***	0.237	0.019***	0.015***
(Intercept)	(0.182)	(0.279)	(0.001)	(0.002)
Observations	10232	2275	10232	2275
R Squared Value	0.284	0.266	0.031	0.014
Adjusted R Squared	0.284	0.265	0.031	0.012
Adjusted Standard Error	2.347	2.496	0.015	0.016
F Statistic	1014	205.9	82.3	7.814

11.25 Tables of results from the Simulation Analysis

The following section collates various outputs that shed greater details on the findings from the simulation analysis. Table 47 outlines summary statistics to describe the distribution of *H* values for each simulation while Table 48 explains the number and percentage of students who were assigned to their 1^{st} – 6^{th} preference school in each simulation.

Table 47: Summary Statistics to describe the distribution of H values for each simulation

Application Stage / Simulation		Minimum	Lower Quartile	Median	Mean	Upper Quartile	Maximum
Number	**Name**						
1	Preferred School	0.566	4.338	7.103	9.091	11.421	17.534
2	Offered Schools	0.348	4.447	7.229	8.979	11.302	17.342
3	Observed Patterns	0.477	4.726	7.660	9.422	11.610	17.797
1	Home-School Distance for over-subscribed schools	0.511	4.222	6.950	8.639	10.799	16.632
2	Random Allocation for over-subscribed schools	0.538	4.139	6.697	8.524	10.583	16.202
3	Academic Banding with a distance criterion for over-subscribed schools	0.358	3.694	5.981	7.800	9.617	14.691
4	Academic Banding with random allocation for over-subscribed schools	0.372	3.849	6.240	8.055	9.954	15.210
5	Academic competition for over-subscribed schools	0.532	4.074	6.614	8.414	10.522	16.138
6	Strict Home School Distance	0.200	2.976	4.691	6.195	7.906	11.950
7	Grammar School System	0.415	3.213	4.885	6.171	7.585	11.329

Table 48: Describing the number and percentage of students allocated to their 1st – 6th preference school in each simulation.

		Observed Patterns		Home school distance		Random Allocation		Academic Banding / distance criteria		Academic Banding / random allocation		Academic Competition		Strict Home School Distance		Grammar School System	
	Rank	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
Ranked Preference of school on application form	1	444924	84.3	448897	85.9	484300	92.2	444246	86.4	401929	80.2	466908	89.2	Not Applicable			
	2	42191	8	39644	7.6	15634	3	19095	3.7	45781	9.1	25573	4.9				
	3	13749	2.6	12332	2.4	5509	1	7315	1.4	15928	3.2	8605	1.6				
	4	4642	0.9	3853	0.7	1819	0.3	2362	0.5	5371	1.1	2757	0.5				
	5	1965	0.4	1460	0.3	878	0.2	1115	0.2	2418	0.5	1154	0.2				
	6	858	0.2	527	0.1	421	0.1	612	0.1	1346	0.3	550	0.1				
Rank of the school in relation to the student's home school distance	1	Not Applicable		11562	2.2	5756	1.1	8816	1.7	12682	2.5	4218	0.8	415850	78.8	406910	79.0
	2			2382	0.5	4926	0.9	6763	1.3	3694	0.7	4702	0.9	43399	8.2	38474	7.5
	3			954	0.2	2595	0.5	3833	0.7	1886	0.4	3382	0.6	17625	3.3	17530	3.4
	4			442	0.1	1257	0.2	2440	0.5	1084	0.2	2096	0.4	11265	2.1	10007	1.9
	5			198	0	723	0.1	1710	0.3	791	0.2	1398	0.3	7259	1.4	5816	1.1
	6			123	0	466	0.1	1146	0.2	614	0.1	855	0.2	5500	1.0	4294	0.8
	7			79	0	264	0.1	1016	0.2	421	0.1	448	0.1	3969	0.8	3220	0.6
	8			92	0	205	0	860	0.2	434	0.1	317	0.1	3106	0.6	2319	0.5
	9			19	0	125	0	666	0.1	305	0.1	149	0	2567	0.5	2021	0.4
	10			22	0	95	0	595	0.1	268	0.1	92	0	1854	0.4	1930	0.4
	11			13	0	67	0	511	0.1	211	0	52	0	1138	0.2	1590	0.3
	12			19	0	38	0	478	0.1	225	0	26	0	1136	0.2	942	0.2
	13			3	0	14	0	444	0.1	174	0	7	0	1037	0.2	997	0.2
	14			2	0	20	0	295	0.1	187	0	7	0	790	0.1	1077	0.2
	15			0	0	24	0	342	0.1	139	0	6	0	631	0.1	889	0.2
	Not Assigned			3	0	40	0	9622	1.9	5278	1.1	2	0	10587	2.0	17117	3.3
	Other	527786	3.6	Not Applicable													

11.26 Additional diagrams from Chapter 7

The following section outlines the geographic distribution of H values for each stage of the application process as well as the seven simulations that were conducted as part of the analysis in Chapter 7.

Figure 84: Map to show the geographic distribution of *H* values for secondary school pseudo-neighbourhoods at each stage of the 2017/2018 national school admissions cycle.

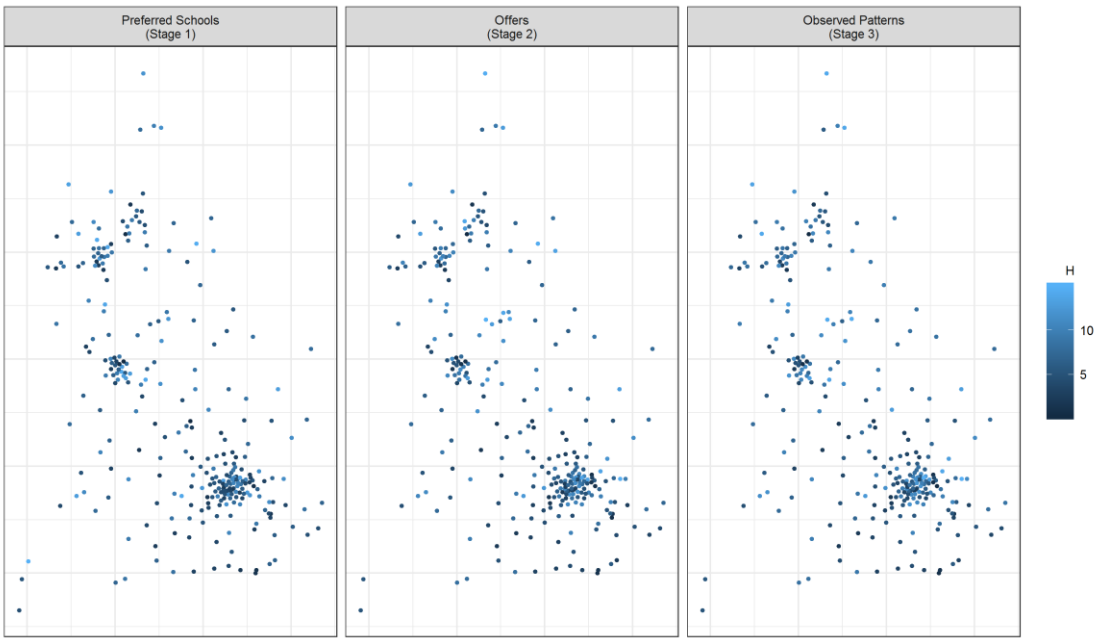


Figure 85: Map showing how *H* values for pseudo-neighbourhoods differed between stage 1 and 2, and stage 1 and 3 of the 2017/2018 national school admissions cycle.

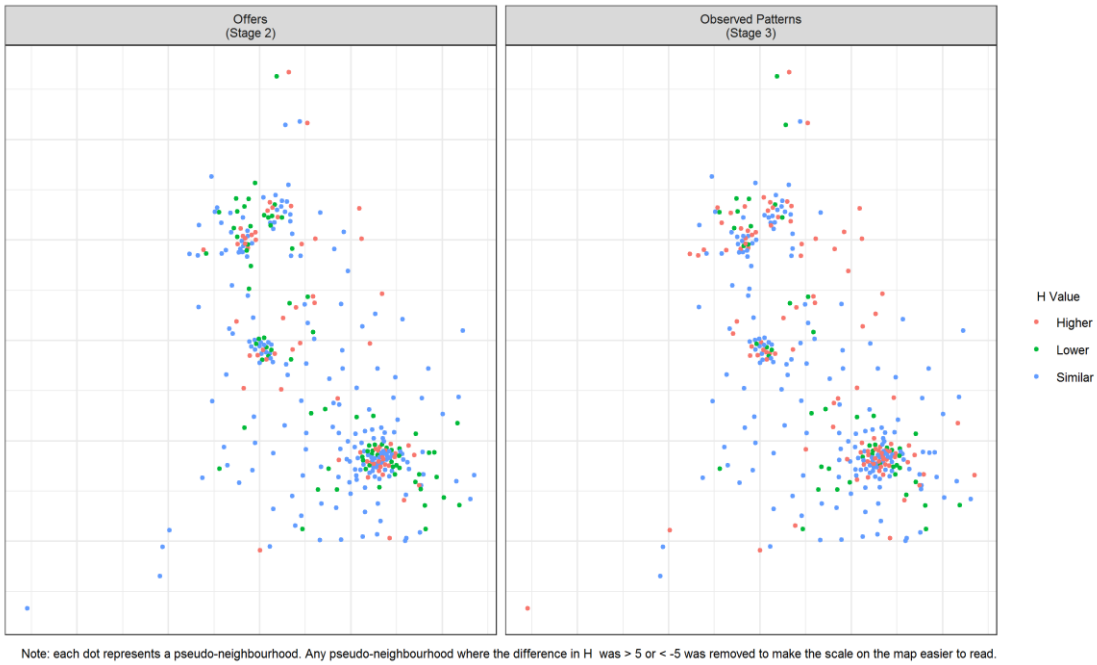
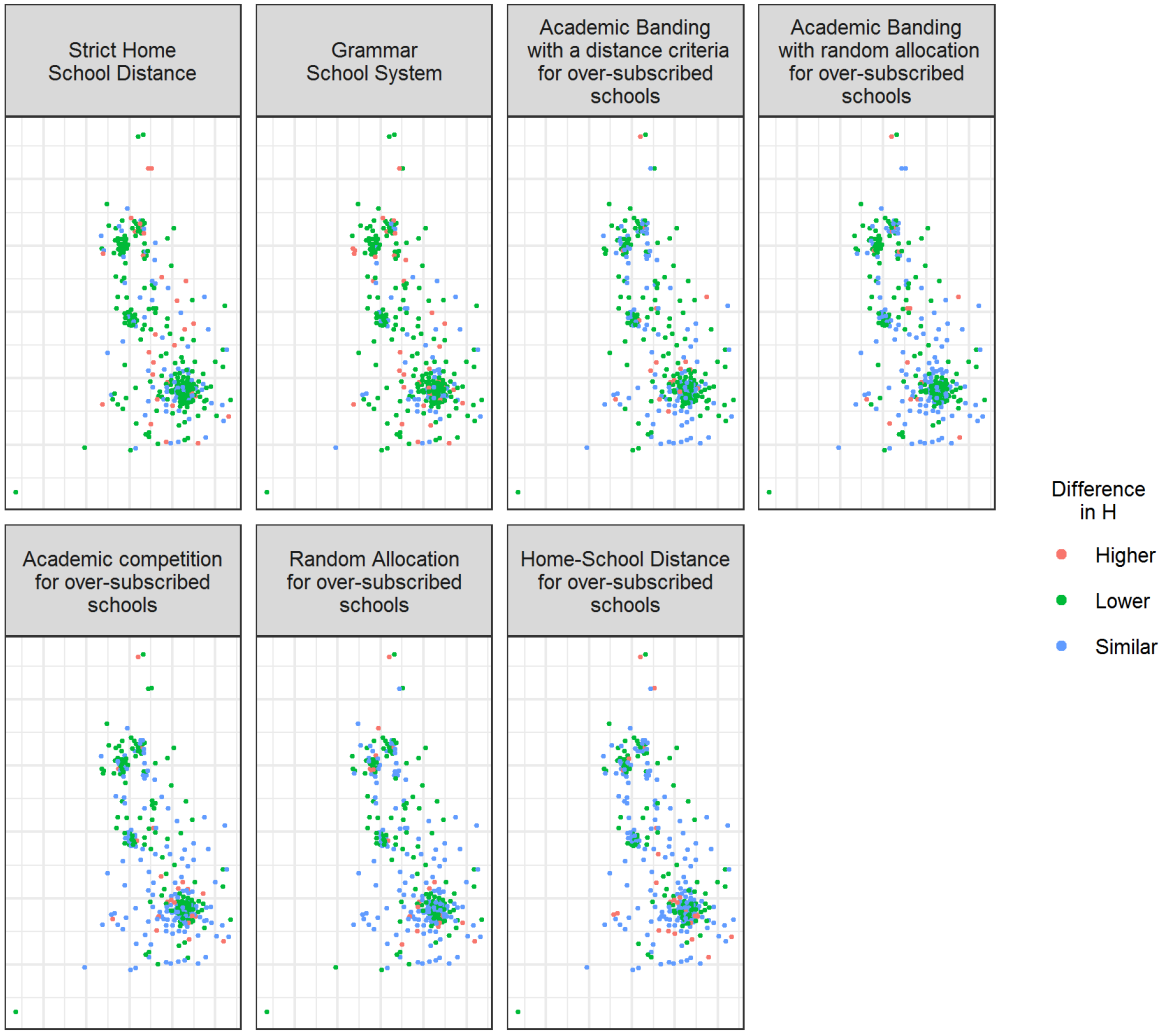


Figure 86: Map showing how H values for pseudo-neighbourhoods differed between what was observed, and the values produced by the seven simulations national school admission systems.

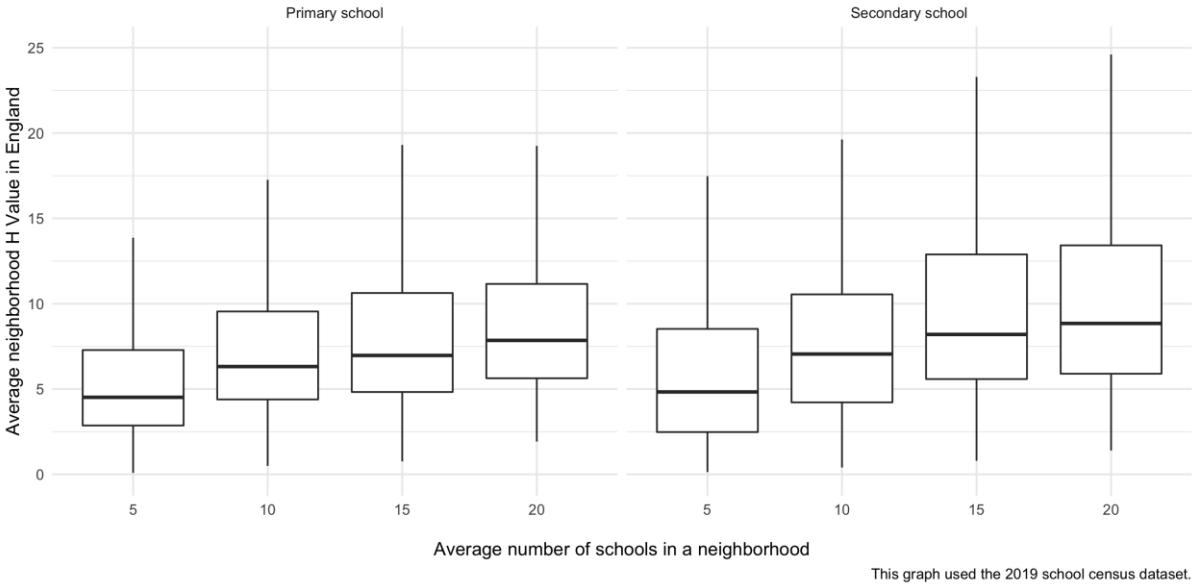


Note: each dot represents a pseudo-neighbourhood. Any pseudo-neighbourhood where H did not change by more than 1 was marked as "Similar".

11.27 Using different sized pseudo-neighbourhoods to measure ethnic school segregation at various geographic scales

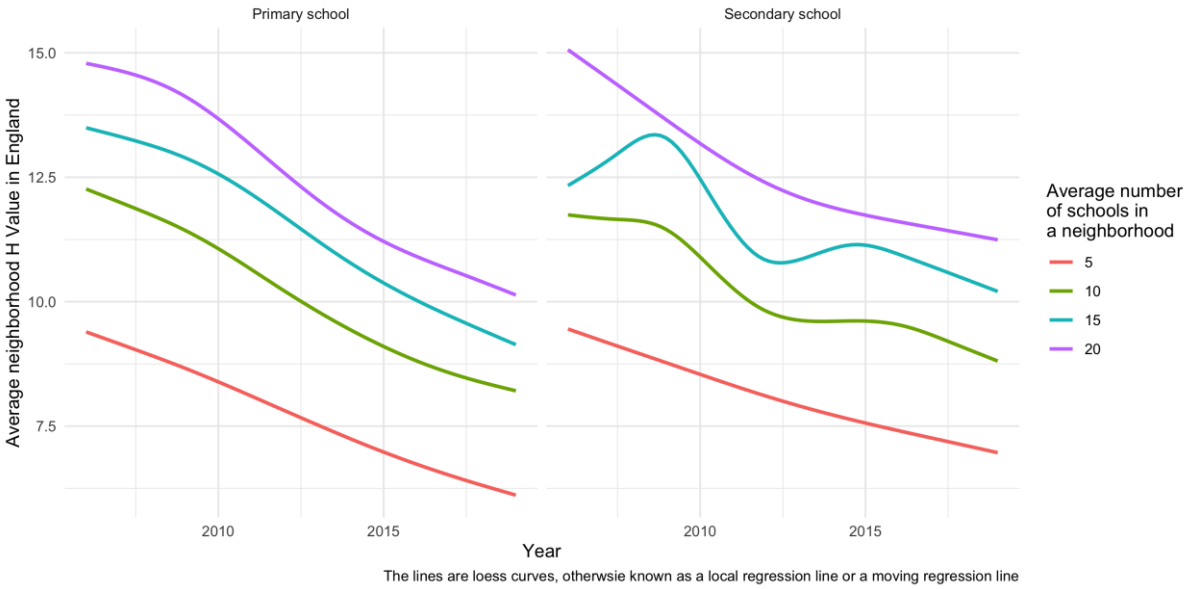
The following subsection outlines a suggested approach to measuring ethnic school segregation across various geographic scales. The study calculated *H* using four sets of pseudo-neighbourhoods, where pseudo-neighbourhoods contained an average of either 5, 10, 15 and 20 schools. The box plot below shows the distributions of *H* values for pseudo-neighbourhoods of different sizes for 2019.

Figure 87: The distribution of H values for pseudo-neighbourhoods with an average of 5, 10, 15 and 20 schools



This approach is particularly helpful when measuring ethnic school segregation across time as it becomes possible to investigate whether changes are consistent between pseudo-neighbourhoods of different sizes. As shown in the line graph below, the average level of ethnic school segregation in the UK fell consistently for pseudo-neighbourhoods of all four sizes, suggesting that the major findings of this thesis were not scale dependent.

Figure 88: Plotting the average level of ethnic school segregation within pseudo-neighbourhoods with an average of 5, 10, 15 and 20 schools between 2006-2019



This technique also allows us to compare the relative change for pseudo-neighbourhoods of different sizes. For example, the average *H* value calculated using primary school pseudo-neighbourhoods which contained an average of 5 schools fell by a greater proportion than pseudo-neighbourhoods which contained an average of 20 schools. This can improve our understanding of whether changes in segregation differs at different geographic scales of analysis.

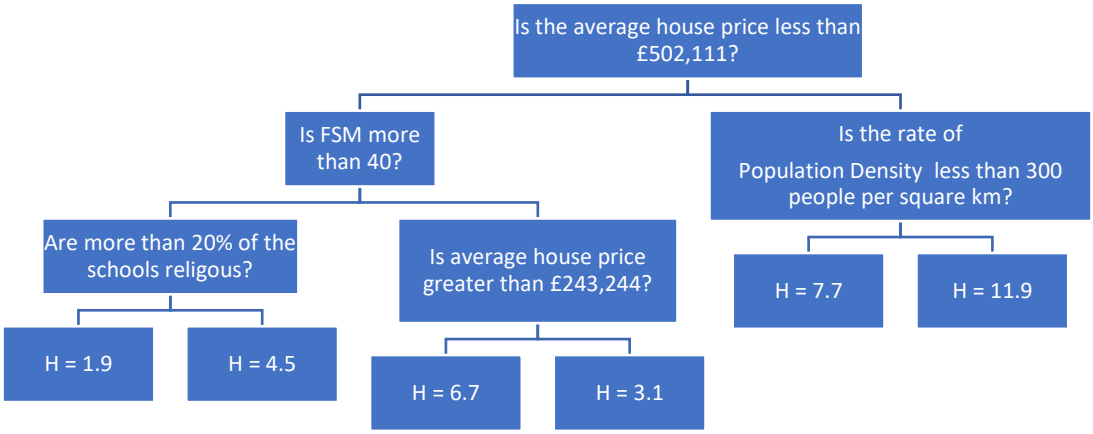
11.28 Random Forest Models

Based on the findings of the pilot study, meta regression analysis was initially deemed unlikely to explain a substantial proportion of the variation in *H*. Consequently, the thesis employed Random Forest models to provide an estimation for the combined predictive power of the study's predictor variables on *H* and to explore which variables had the greatest importance in estimating a pseudo-neighbourhood's *H* value. While these findings were ultimately not reported in the main body of the text, the results could help guide future research with possible variable and model selection. This section will explain how random forest models work before presenting the findings of the analysis.

11.28.1 Explaining random forest models

Random Decision Forests is a tree-based modelling technique that is highly regarded for its predictive power (Casella, Fienberg and Olkin, 2015). Tree based models sort data into increasingly small groupings using a series of binary decision rules that minimalize the within-group variation in the outcome variable.

Figure 89: A fictitious tree-based model to predict the level of segregation based on the level of FSM, house price and ethnic diversity (E).

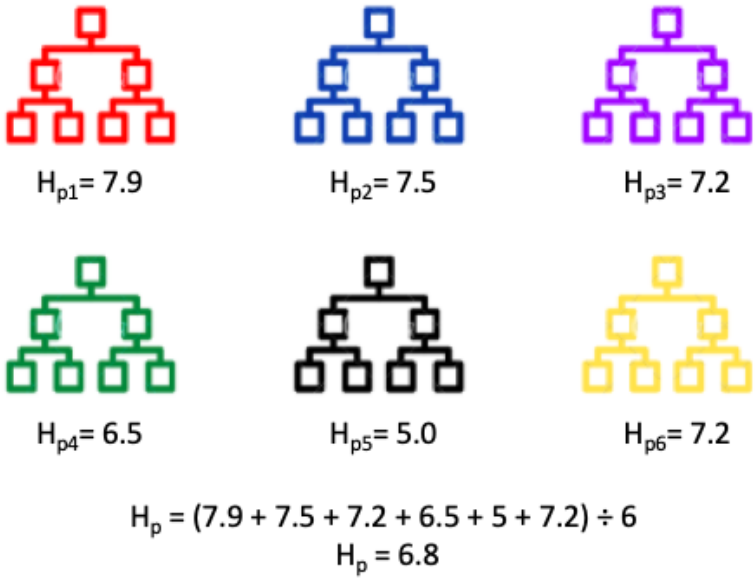


Unfortunately, a single tree model normally yields a poor prediction as they tend to over fit the data. As tree models naturally add decision rules until each data point is sorted into a homogenous group, later decision rules split data based on noise in the sample data rather than predictive features of the wider population. Consequently, most tree-based models include a method(s) of removing some of the later decision rules, a process known as "pruning" the tree.

Random Forests offer one of the best techniques for creating tree-based models which avoids overfitting the data. Proposed by Prof Ho in 1995, random forests create a series of uncorrelated tree-based models using a random selection of predictor variables and a bootstrapped sample of the data (Kam Ho, 1995). Each of the uncorrelated trees is then used to produce a prediction for each data point. While each individual prediction is poor, the combined predictions of multiple trees often results in "dramatic improvements in

predictive accuracy"(Casella, Fienberg and Olkin, 2015). See Figure 90 for a diagrammatic representation of this process.

Figure 90: Diagram to show how random forest models work.



11.28.2 Applying random forest models to this study

The pilot study created 100 random Forrest models using the R library, randomForest (Liaw and Wiener, 2002). Like with the meta regression analysis, each model was fitted to a different set of pseudo-neighbourhood boundaries with the model outputs stored and then averaged across all models. Each model used the same independent variables used in the meta-regression models.[215]

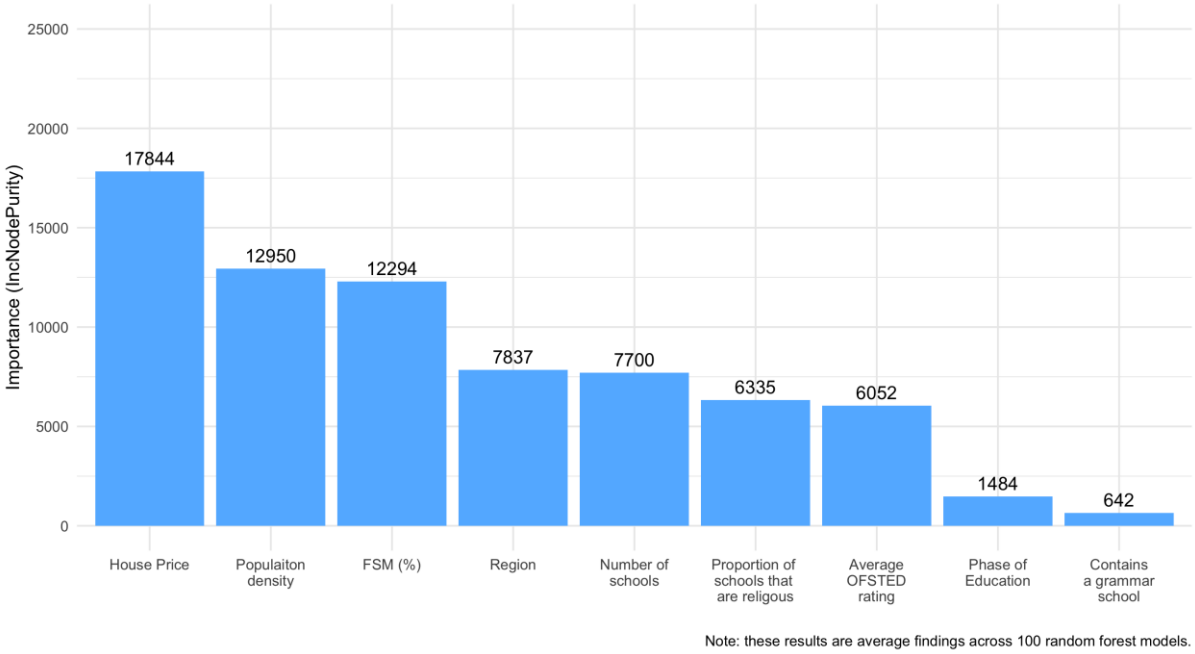
[215] These are FSM, the average price paid for a terrace house in the neighbourhood, the neighbourhood's ethnic diversity (*E*) and the percentage of students in the neighbourhood who are White British.

Two outputs were extracted from the Random Forest models. The first was the variation in *H* which was explained by all predictor variables. The second output was a metric called IncNodePurity, which evaluates the importance of each variable by calculating the "total decrease in node impurities from splitting on [each] variable, averaged over all trees [within a single Random Forrest model]." (Liaw and Wiener, 2002). While IncNodePurity is hard to directly interpret, it provides a useful comparative tool to measure the relative importance of each variable. Both figures will help future researchers select the most appropriate variables to explain patterns of ethnic school segregation in England. Unfortunately, Random Forest models are a "black-box" model which offer little insight into how these variables relate to one another. It is therefore left to future research to untangle the precise nature of these relationships.

11.28.3 Findings from the random forest models

Across 100 random forest models, around 34.6% of the variation in *H* was explained using all independent variables included in this study. As shown in the bar chart below, certain variables provided a greater reduction in the prediction error of the model than others. Out of the nine variables, a neighbourhood's average cost of housing was the most predictive. This was followed by the pseudo neighbourhood's rate of FSM, its region, its population density, the number of schools, the proportion of schools that were religious, the average OFSTED rating, the phase of education and whether the neighbourhood contained a grammar school.

Figure 91: Comparing the predictive power of each independent variable on a pseudo-neighbourhood's H value.



Comparing the results from the random forest models to the coefficients of the meta-regression analysis provide a useful insight into which variables may be important for predicting ethnic school segregation but may have non-linear, or more complex, associations to *H*. For example, the results from the random forest models suggest that a pseudo-neighbourhood's level of FSM and its population density may be more important for predicting ethnic school segregation than the results from the meta-regression models suggests. In contrast, the proportion of a schools in a pseudo-neighbourhood that are religious is one of the most important variables found by the meta regression models but is less important in reducing predictive error in the random forest models. This provides

partial evidence to suggest that the linear regression models *may* have given it undue importance.