

A Online Appendix: English education system

Ages and stages:

The English school curriculum is divided into blocks of years called “Key Stages” (KS). Children are assessed based on national standards at the end of each Key Stage. Key Stage 1 runs through ages 5 to 7 years, and Key Stage 2 through ages 7 to 11 years. Key Stage 2 is assessed at the end primary school at age 11 years. Children then transition to secondary school, where they must attend until age 16. Key Stage 3 runs from ages 11 to 14 years. During Key Stage 4, ages 14 to 16 years, children study a set of subjects in which they will sit the high-stakes General Certificate of Secondary Education examinations (GCSEs), or equivalents, at age 15/16 years. Pupils typically take between 5 and 10 subjects, and usually a passing grade in at least five of them—including Mathematics and English—is required to progress to many academic or other qualifications. A common school performance metric is the proportion of pupils achieving at least five GCSEs at grades A*-C (including English and Mathematics).

With satisfactory performance in GCSEs, pupils can study for A-levels, from age 16–18. A-Levels are the most common route to university entrance, and are taken either at the same secondary school, or some pupils move to another school or to a specialised “sixth form college” to take their A levels. Further Education (FE) colleges are an alternative to staying in an secondary school, and they provide vocational training or a mix of academic and vocational. Pupils typically enter Higher Education (University) from age 18.

School types:

The types of secondary schools in England in 2001 are described as follows. The first category are “state schools” which are controlled and funded by central government, and on which this study is based. All children in England between the ages of 5 and 16 are entitled to a free place at a state school. These schools follow the National Curriculum and are inspected by Ofsted (the government’s Office for Standards in Education, Children’s Services and Skills). About 93% of pupils attended such schools when our data was collected (97% unweighted). The sec-

ond category are “Independent Schools” (private schools) which about 7% of pupils attended (3% unweighted).

Within the category of state schools, there are community schools (67% of pupils in our data), voluntary-controlled (3%), voluntary-aided (11%), foundation schools (15%). *Community schools* are entirely run by the local council (Local Authority). *Foundation or Trust schools* are run by a local governing body. These schools were formerly called “Grant-maintained” schools. This was an initiative to allow more flexibility in provision of education where, by majority parental vote, schools could opt out of Local Authority control and be run by a governing body with more control over admissions and staffing. *Voluntary-aided schools* are typically religious or faith schools, which can admit pupils on religious affiliation grounds. *Voluntary-controlled schools* are almost all faith schools. They are a mix between community and voluntary-aided schools: similar to a community school, the local authority employs the staff and sets the entrance criteria, but the school land and buildings are owned by a charity, often a church, which also can appoint some members of the governing body.

Regardless of governance arrangement, all non-selective state schools have to comply with the school admissions code of practice which sets guidelines for fair admissions. The exception to this are grammar schools: about 3% of state schools are *Grammar schools*, a minority of academically selective state-funded schools. They select all or most of their pupils based on academic ability, assessed by the so-called “11 plus” exam. The 11-plus exam is distinct from the national Key Stage assessments that take place at the age of 11. The former are not nationally co-ordinated. Sometimes they are set and graded by individual grammar schools, and sometimes (usually in local authorities with several grammar schools) they are co-ordinated by the local authority.

B Online Appendix: Sample characteristics

Table WA.1: Regional distribution of sample

Region	N	Missed out	Distance to school (km)			Weighted N	Pr(Urban) Proportion
			Mean	Median	80th percentile		
North East	254	0.071	2.527	1.526	3.285	302.46	0.903
North West	703	0.086	2.986	1.877	3.670	702.85	0.885
Yorks & Humber	453	0.117	3.079	2.040	4.077	475.37	0.816
East Midlands	383	0.104	2.984	1.801	4.221	408.29	0.720
West Midlands	601	0.132	2.696	1.463	3.940	578.56	0.873
East of England	578	0.072	2.971	1.806	4.679	566.75	0.726
Greater London	796	0.170	2.534	1.796	3.852	589.01	1.000
South East	748	0.094	3.114	1.894	4.859	756.69	0.752
South West	415	0.048	3.401	1.837	5.615	496.23	0.651
Total	4931	0.101	2.938	1.788	4.200	4876.21	0.815

Notes: This table shows the proportion (mean) missing out on their first choice school (“missed out”); the mean, median and 80th percentile of distance from home to school in kilometres; the weighted sample size; the share of sample members living in an urban area, Pr(Urban); by region. The unweighted sample size underlying all of these statistics is denoted as N .

Table WA.2: Comparison of key variables with population counterparts

	Male	White British	English First Language	KS4 points	5+ A*-C	FSM eligible
<i>Final sample</i>	0.51	0.88	0.94	361.61	0.57	0.12
<i>Population</i>	0.51	0.81	0.91	347.61	0.56	0.16

Notes: This Table reports the means (or proportions for binary variables), and samples sizes, of a set of demographic, socio-economic and attainment variables. The variables are selected based on the availability of variables in the population data available for this project. *Final LSYPE sample* reports these statistics for the final analytical sample used in the paper comprised of records with non-missing data on all outcomes and covariates (n=4,931). *Population* reports these statistics for the population counterparts: all 'on roll' pupils who were in Year 9 in the academic year 2003/04. The data source is the 2003/04 Pupil-Level Annual School Census (PLASC) linked to Key Stage attainment data. These population statistics were produced using statistical data from ONS via the Secure Research Service. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

C Online Appendix: Further results

Ethnicity sub-group analyses

Previous research has found negligible differences in the probability of gaining a place at a first-choice school by socio-economic status, but large differences by ethnicity. For instance, 90% of White British families were offered a place at their first-choice school, compared with 72% Asian and 62% of Black families applying for schools in England in the 2014/15 academic year (Burgess et al., 2019). As shown in Weldon (2018), this pattern is observed across the regions of the England—it is not, for instance, explained by clustering of black and ethnic minority families in the denser London school market. The reasons for this variation remain unclear, but one possible explanation is differences by ethnicity in engagement with the school choice process. White British families are more likely to list only one school: 41% of White British households list one school, compared to 17% of Asian households and 12% of Black households (Burgess et al., 2019). Ethnic minority families are more likely to list many schools, and select, on average, more ambitious (further away and better quality) schools as their first choice. This may lead to a higher chance of missing out based on the oversubscription criteria.

This motivates splitting the sample by ethnic origin of the child. Table WA.3 reports OLS and distance-matched results separately by ethnic group: White British, or Black, Asian or other minority ethnicity (BAME). The distance-matched results are less precise than OLS, as they are based on smaller sample sizes due to the restricted bandwidth, especially within subgroups. The only outcome for which the difference in effect size between subgroups is statistically significant across both specifications is mental ill-health, where the effects appear larger among White pupils. The effects on staying on in school after age 16 years, the minimum age of compulsory schooling at this time, are also consistently larger across OLS and the distance-matching results among White pupils, however the difference between subgroups is only statistically significant in the OLS results. This pattern could have several explanations, including differential school preferences by ethnicity, variation in residential sorting, or differential list lengths as discussed earlier. If White parents are not listing enough schools, when they do miss out they may be

allocated a school which they would not have chosen at all; While if BAME parents list more schools, in the way the system is designed to be used, and they miss out then the child has a higher chance of ending up in a fallback local school that is at least acceptable. However, without further data is difficult to be confident about the specific explanatory mechanism(s).

Table WA.3: Estimation results by ethnicity subgroup

	5+ A*-C	KS4 points	Stayed on	Attended uni	Mental ill-health	ln(income)	Fertility	Smoker
<i>OLS</i>								
BAME								
β	-0.02	-18.36	0.03	0.05	-0.15	-0.01	-0.00	0.02
(s.e.)	(0.04)	(12.82)	(0.04)	(0.04)	(0.26)	(0.01)	(0.03)	(0.04)
δ	-0.94	-3.73	1.88	2.90	-1.45	-2.86	-0.71	5.34
N	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337
White								
β	-0.04	-16.09**	-0.06**	0.01	0.74***	-0.01	0.06**	0.09***
(s.e.)	(0.03)	(7.07)	(0.03)	(0.02)	(0.24)	(0.01)	(0.03)	(0.03)
δ	8.98	8.57	10.81	-1.77	70.13	3.93	6.33	50.44
N	3,594	3,594	3,594	3,594	3,594	3,594	3,594	3,594
<i>p</i> -value	0.27	0.47	0.01	0.29	0.03	0.88	0.16	0.20
<i>Distance-matching</i>								
BAME								
ATT	-0.06	-19.80	-0.00	0.06	-0.13	-0.03**	0.06	0.06
(s.e.)	(0.05)	(14.79)	(0.04)	(0.06)	(0.32)	(0.02)	(0.04)	(0.04)
N	977	977	977	977	977	977	977	977
White								
ATT	-0.05	-12.56	-0.06*	0.02	1.04***	-0.02	0.05	0.10**
(s.e.)	(0.03)	(9.27)	(0.03)	(0.03)	(0.32)	(0.01)	(0.04)	(0.04)
N	1,625	1,625	1,625	1,625	1,625	1,625	1,625	1,625
<i>p</i> -value	0.75	0.68	0.31	0.56	0.01	0.43	0.82	0.53

Notes: β : OLS coefficient on treatment variable; ATT: average treatment effect on the treated; N: observations; δ : size of proportional selection on unobservables to drive ATT to zero; estimates in brackets are robust standard errors clustered by school. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All analyses are weighted using the survey weights. *p*-value is the *p*-value from a test of equality of coefficients between groups. BAME; child's reported ethnicity is Black, Asian and Minority Ethnicity. White; child's reported ethnicity is White British. Each specification includes the following covariates: gender; month-of-birth; older siblings; whether English is the main language spoken at home; Free School Meal eligibility; whether the child has special educational needs; Key Stage 2 scores; average Key stage 2 scores at the child's primary school; housing tenure; family income; parental occupational social class; parental education; parental religion; Principal Components Analysis summary of information used to choose schools; presence of local Grammar school; distance to closest private school; distance-weighted share of good-, very good- or excellent-rated schools within the 80th percentile of home-to-school distances travelled in that child's Local Authority; Local Authority fixed effects.

Table WA.4: Estimation results by ethnicity subgroup - school attributes in controls

	5+ A*-C	KS4 points	Stayed on	Attended uni	Mental ill-health	ln(income)	Fertility	Smoker
<i>OLS</i>								
BAME								
β	-0.01	-17.23	0.02	0.06	-0.18	-0.02	-0.01	0.03
<i>s.e.</i>	(0.04)	(12.89)	(0.04)	(0.04)	(0.27)	(0.01)	(0.03)	(0.04)
δ	-0.49	-3.80	1.48	4.15	-1.56	-2.90	-1.49	12.52
N	1,337	1,337	1,337	1,337	1,337	1,337	1,337	1,337
White								
β	-0.02	-5.97	-0.04*	0.04*	0.75***	-0.01	0.04	0.08**
<i>s.e.</i>	(0.03)	(6.88)	(0.03)	(0.02)	(0.24)	(0.01)	(0.03)	(0.03)
δ	1.61	1.17	3.57	-2.51	52.57	2.02	2.02	12.57
N	3,594	3,594	3,594	3,594	3,594	3,594	3,594	3,594
<i>p</i> -value	0.35	0.60	0.02	0.42	0.03	0.95	0.23	0.23
<i>Distance-matching</i>								
BAME								
ATT	-0.06	-23.58	-0.04	0.08	-0.38	-0.04**	0.03	0.07*
<i>s.e.</i>	(0.05)	(15.74)	(0.05)	(0.06)	(0.35)	(0.02)	(0.04)	(0.04)
N	977	977	977	977	977	977	977	977
White								
ATT	-0.04	-8.96	-0.05	0.04	1.03***	-0.01	0.04	0.09**
<i>s.e.</i>	(0.03)	(8.59)	(0.03)	(0.03)	(0.32)	(0.01)	(0.04)	(0.04)
N	1,625	1,625	1,625	1,625	1,625	1,625	1,625	1,625
<i>p</i> -value	0.72	0.41	0.81	0.49	0.00	0.22	0.98	0.74

Notes: β : OLS coefficient on treatment variable; ATT: average treatment effect on the treated; N: observations; δ : size of proportional selection on unobservables to drive ATT to zero; estimates in brackets are robust standard errors clustered by school. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All analyses are weighted using the survey weights. *p*-value is the *p*-value from a test of equality of coefficients between groups. BAME; child's reported ethnicity is Black, Asian and Minority Ethnicity. White; child's reported ethnicity is White British. Each specification includes the following covariates: gender; month-of-birth; older siblings; whether English is the main language spoken at home; Free School Meal (FSM) eligibility; whether the child has special educational needs (SEN); Key Stage 2 scores; average Key stage 2 scores at the child's primary school; housing tenure; family income; parental occupational social class; parental education; parental religion; Principal Components Analysis summary of information used to choose schools; presence of local Grammar school; distance to closest private school; distance-weighted share of good-, very good- or excellent-rated schools within the 80th percentile of home-to-school distances travelled in that child's Local Authority; Local Authority fixed effects. School-level controls: % gaining 5+ A*-C grades in 2004; KS2-KS4 value-added; OFSTED rating of good and above?; % with English as first language; % of pupils White British; FSM band of the school; % with SEN; school type.

D Online Appendix: Robustness checks

Table WA.5: OLS results restricted to the sample with within-school variation

	5+ A*-C	KS4 points	Stayed on	Attended uni	Mental ill-health	ln(income)	Fertility	Smoker
<i>A: OLS</i>								
β	-0.04	-14.10**	-0.05**	0.04	0.67***	-0.01	0.03	0.09***
(s.e.)	(0.02)	(6.63)	(0.02)	(0.02)	(0.21)	(0.01)	(0.03)	(0.03)
δ	-3.08	-4.96	-4.57	4.37	56.05	1.27	158.4	-17.00
N	2,607	2,607	2,607	2,607	2,607	2,607	2,607	2,607
<i>B: OLS with school attributes in controls</i>								
β	-0.02	-8.24	-0.04*	0.05**	0.70***	-0.01	0.02	0.08***
(s.e.)	(0.02)	(6.50)	(0.02)	(0.02)	(0.21)	(0.01)	(0.03)	(0.03)
δ	-2.93	-7.92	-5.26	8.08	134.8	0.99	6.61	-95.8
N	2,607	2,607	2,607	2,607	2,607	2,607	2,607	2,607

Notes: The analyses in this Table are restricted to pupils in schools which have within-school variation in the treatment. β : OLS coefficient on treatment variable; N: observations; δ : size of proportional selection on unobservables to drive ATT to zero; estimates in brackets are robust standard errors clustered by school. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All analyses are weighted using the survey weights. Each specification includes the following covariates: gender; month-of-birth; older siblings; whether English is the main language spoken at home; Free School Meal (FSM) eligibility; whether the child has special educational needs (SEN); Key Stage 2 scores; average Key stage 2 scores at the child's primary school; housing tenure; family income; parental occupational social class; parental education; parental religion; Principal Components Analysis summary of information used to choose schools; presence of local Grammar school; distance to closest private school; distance-weighted share of good-, very good- or excellent-rated schools within the 80th percentile of home-to-school distances travelled in that child's Local Authority; Local Authority fixed effects. Panel B additionally includes the following school-level controls: % gaining 5+ A*-C grades in 2004; KS2-KS4 value-added; OFSTED rating of good and above?; % with English as first language; % of pupils White British; FSM band of the school; % with SEN; school type.

Table WA.6: Distance-matched effects of missing out on first-choice school - optimal bandwidth

VARIABLES	5+ A*-C	KS4 points	Stayed on	Attended uni	Mental ill-health	ln(income)	Fertility	Smoker
ATT	-0.03	-13.27*	-0.04	0.03	0.66***	-0.02**	0.04	0.07***
(s.e.)	(0.02)	(7.52)	(0.02)	(0.03)	(0.22)	(0.01)	(0.03)	(0.03)
Y_0	0.58	363.44	0.50	0.36	2.20	9.61	0.27	0.18
N	3389	3389	3389	3389	3389	3389	3389	3389
Bandwidth	6.918	6.918	6.918	6.918	6.918	6.918	6.918	6.918

Notes: ATT: average treatment effect on the treated; Y_0 : model-adjusted outcome mean for control group; Bandwidth is the radius in kilometres from which control units are selected, selected using a data-driven bandwidth selection procedure. N: unweighted cell count for all statistics in this table; Estimates in brackets are the standard errors.

Table WA.7: Distance-matched effects of missing out on first-choice school with school-attributes in controls

VARIABLES	5+ A*-C	KS4 points	Stayed on	Attended uni	Mental ill-health	ln(income)	Fertility	Smoker
ATT	-0.01	-8.52	-0.00	0.04	0.69***	-0.02	0.05	0.07**
(s.e.)	(0.03)	(11.63)	(0.03)	(0.03)	(0.23)	(0.02)	(0.03)	(0.03)
Y_0	0.56	358.85	0.47	0.35	2.20	9.60	0.26	0.19
N	3039	3039	3039	3039	3039	3039	3039	3039

Notes: ATT: average treatment effect on the treated; Y_0 : model-adjusted outcome mean for control group; N: unweighted cell count for all statistics in this table; Estimates in brackets are the standard errors. Each specification includes the following covariates: gender; month-of-birth; older siblings; whether English is the main language spoken at home; Free School Meal (FSM) eligibility; whether the child has special educational needs (SEN); Key Stage 2 scores; average Key stage 2 scores at the child's primary school; housing tenure; family income; parental occupational social class; parental education; parental religion; Principal Components Analysis summary of information used to choose schools; presence of local Grammar school; distance to closest private school; distance-weighted share of good-, very good- or excellent-rated schools within the 80th percentile of home-to-school distances travelled in that child's Local Authority; Local Authority fixed effects. School-level controls: % gaining 5+ A*-C grades in 2004; KS2-KS4 value-added; OFSTED rating of good and above?; % with English as first language; % of pupils White British; FSM band of the school; % with SEN; school type.