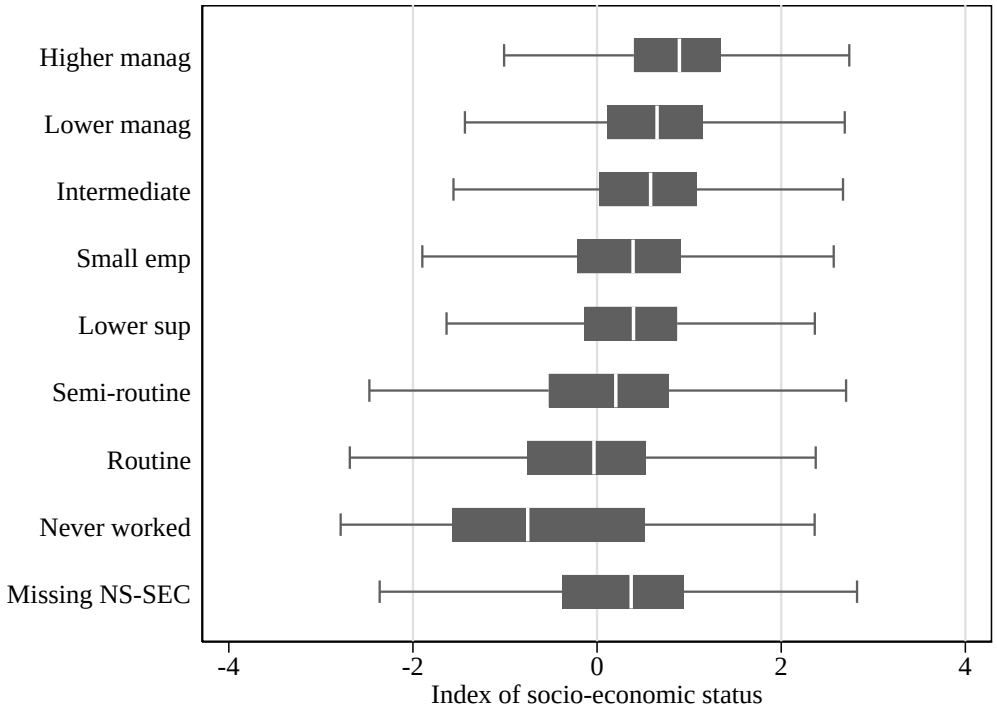
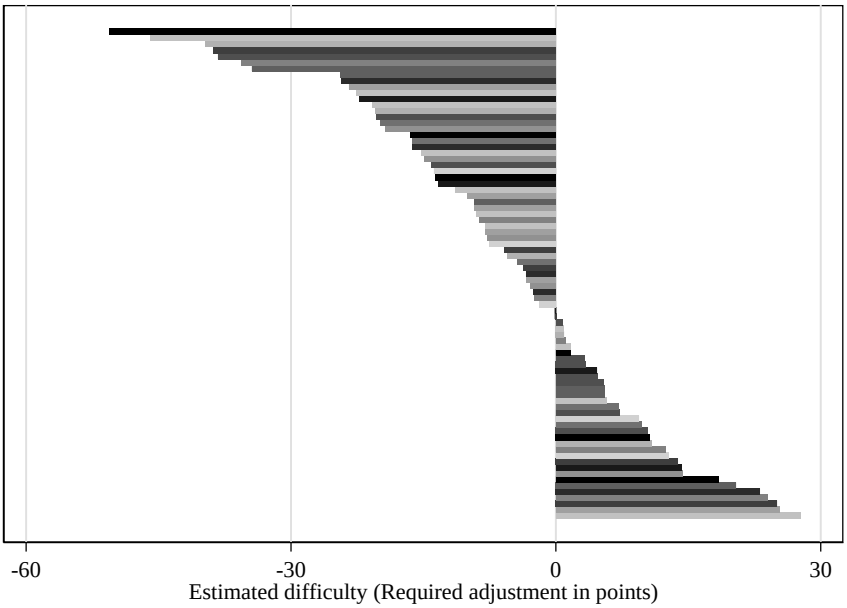
Appendix

Figure A1: Comparison of SES index with HESA NS-SEC



Source: NPD-HESA. n=138,969. Compares our SES index with the 'National Statistics-Socio-Economic Classification' (NS-SEC) which is available for around 80% of university attendees in our sample. The NS-SEC measure available in the HESA data is a fairly noisy categorical indicator of SES, since it relies on a mapping from the parental occupation which each student enters into their university application form, and has a relatively high level of non-response. Still, it is reassuring that our continuous measure of SES places the categories of the NS-SEC in a plausible ranking.

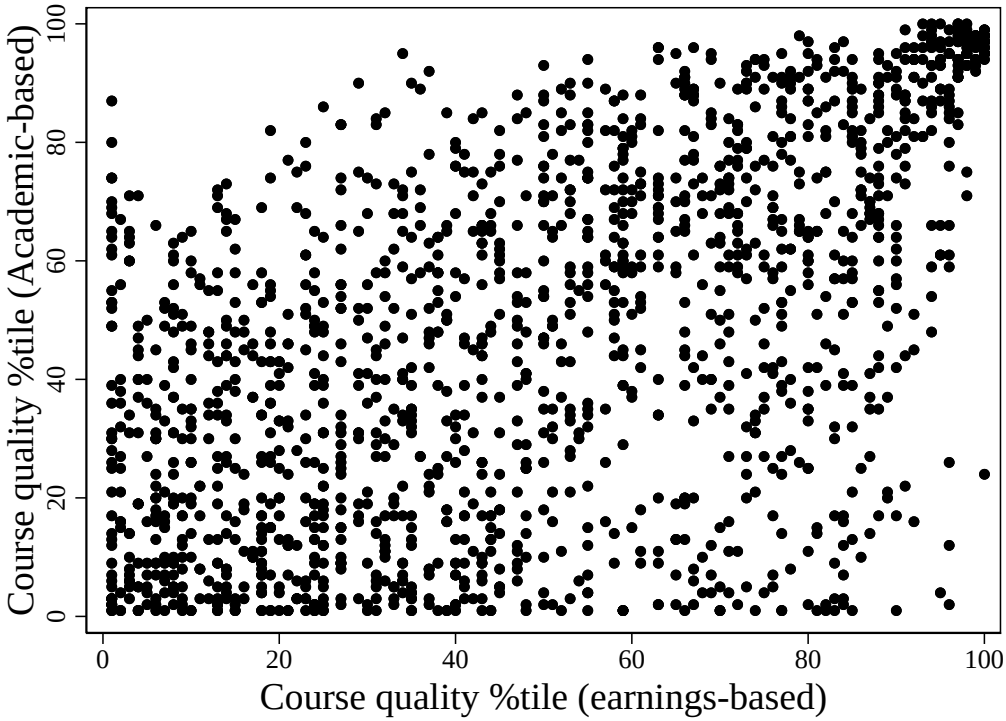
Figure A2: A-Level subject difficulty correction



Subject	Adjustment	Subject	Adjustment	Subject	Adjustment	Subject	Adjustment
Persian	-50.53	Dance	-15.23	PE	-3.39	IT	5.54
BTEC	-45.88	A&D: 3D design	-14.84	RE	-3.34	Ancient history	5.74
Bengali	-39.74	Drama	-14.09	Psychology (sci)	-2.90	Spanish	7.12
Urdu	-38.78	Fine Art	-13.82	Geography	-2.53	Accounting	7.25
Panjabi	-38.28	Art and Design	-13.64	Business studies and economics	-2.44	Mathematics (pure)	9.42
Turkish	-35.66	Portuguese	-13.26	English literature	-1.85	Greek	9.67
Polish	-34.43	Performing	-11.43	Psychology (soc)	-0.09	Computer studies	10.43
Film	-24.43	Dutch	-10.04	Home Economics	-0.02	Logic/philosophy	10.65
Communication	-24.31	D&T: Food	-9.27	Mathematics (statistics)	0.75	Mathematics	10.85
Russian	-23.44	Business	-9.20	Government and Politics	0.92	French	12.43
Modern Greek	-22.60	Development	-9.05	Other classical languages	0.93	Music	12.72
A&D: Photography	-22.21	D&T: Production	-8.68	Archaeology	1.08	German	13.83
A&D: Graphics	-20.78	Electronics	-7.99	Modern Hebrew	1.64	Environmental Science	14.24
A&D: Crit & Contextual studies	-20.43	D&T: Systems	-7.97	Italian	1.67	Latin	14.33
Media, film, and TV	-20.30	English	-7.74	Gujarati	3.21	Science	18.39
A&D: Textiles	-19.89	Chinese	-7.51	Music technology	3.40	Biology	20.41
Arabic	-19.35	Law	-5.85	Classical civilisation	4.67	Mathematics (further)	23.12
Vocational double award	-16.45	English language	-5.50	History	4.71	Chemistry	23.94
Sociology	-16.27	Geology	-4.35	History of art	5.42	Physics	25.05
Japanese	-16.26	Vocational A-Level	-3.70	Economics	5.51	Biology (human)	25.32
						Additional mathematics	27.74

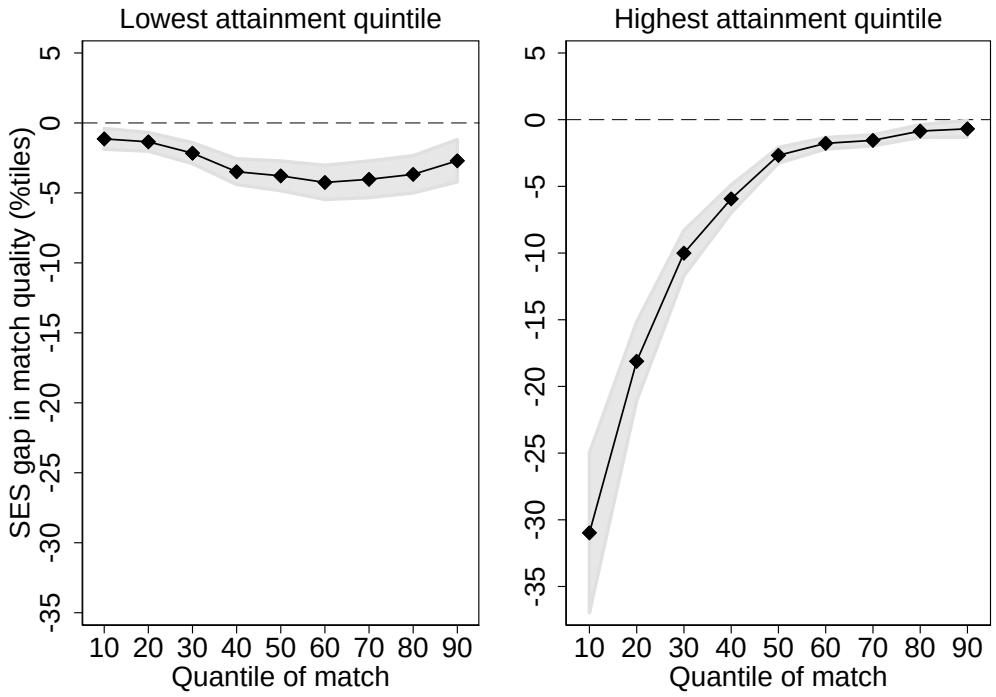
Source: NPD-HESA. n=138,969. Units are QCA (Qualifications and Curriculum Authority) points, where one A-Level grade is 30 points. Adjustment method discussed in Technical Appendix.

Figure A3: Academic- and earnings-based measures of match



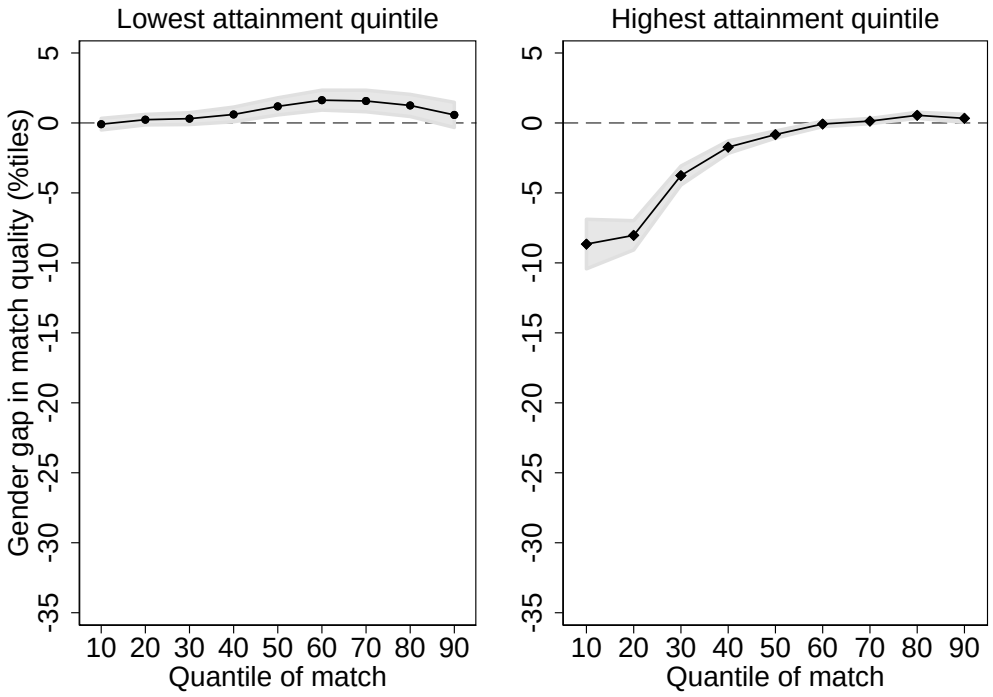
Source: NPD-HESA. n=1,722. Notes: Each point represents a university degree, plotted against our two quality measures: Median graduate earnings percentile; Median student entry qualification percentile.

Figure A4: SES gaps in severity of academic-based match



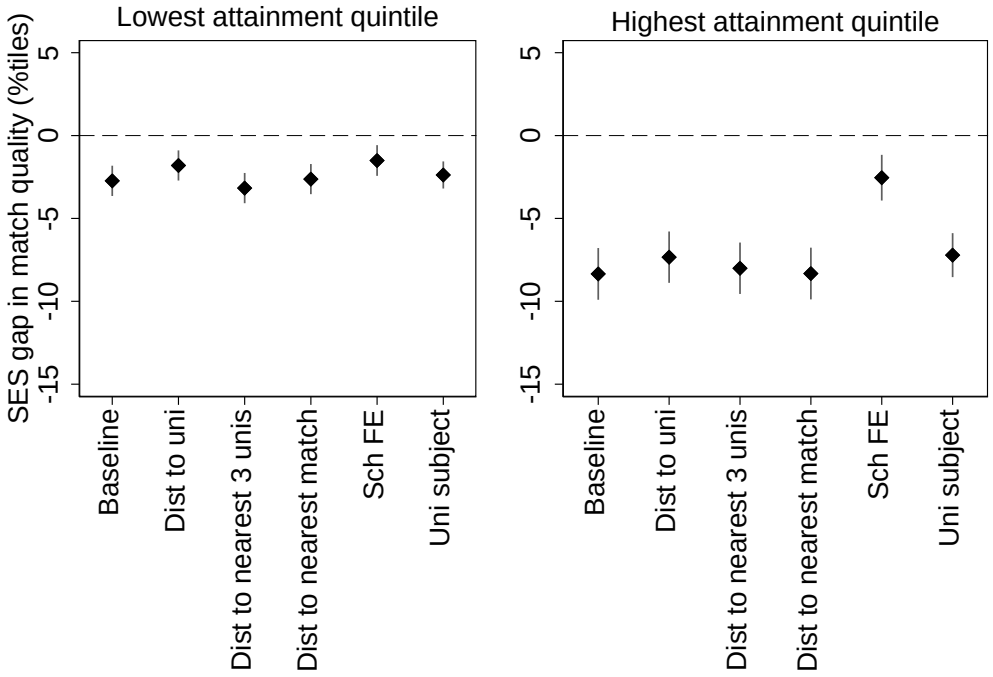
Source: NPD-HESA. n=138,969. Notes: Each point represents the SES match gap between groups 1 and 5 from specification 2, estimated for each decile of the match distribution. Controls are dummies for ethnicity, English as an Additional Language, Special Educational Needs, and gap year before college, and cubics in age 11 and age 16 exam results. We present the 95 percent confidence intervals, with standard errors clustered at the secondary school level.

Figure A5: Gender gaps in severity of academic-based match



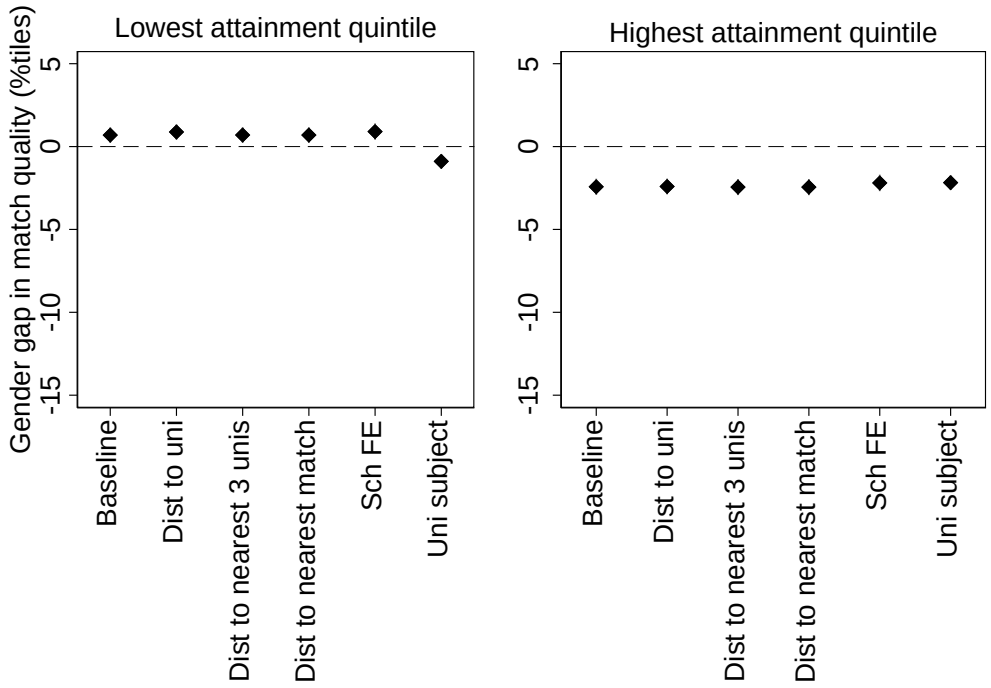
Source: NPD-HESA. n=138,969. Notes: Each point represents the gender match gap between groups 1 and 5 from specification 2, estimated for each decile of the match distribution. Controls are dummies for ethnicity, English as an Additional Language, Special Educational Needs, and gap year before college, and cubics in age 11 and age 16 exam results. We present the 95 percent confidence intervals, with standard errors clustered at the secondary school level.

Figure A6: SES gaps in academic-based match, conditional on geography, subject, and schools



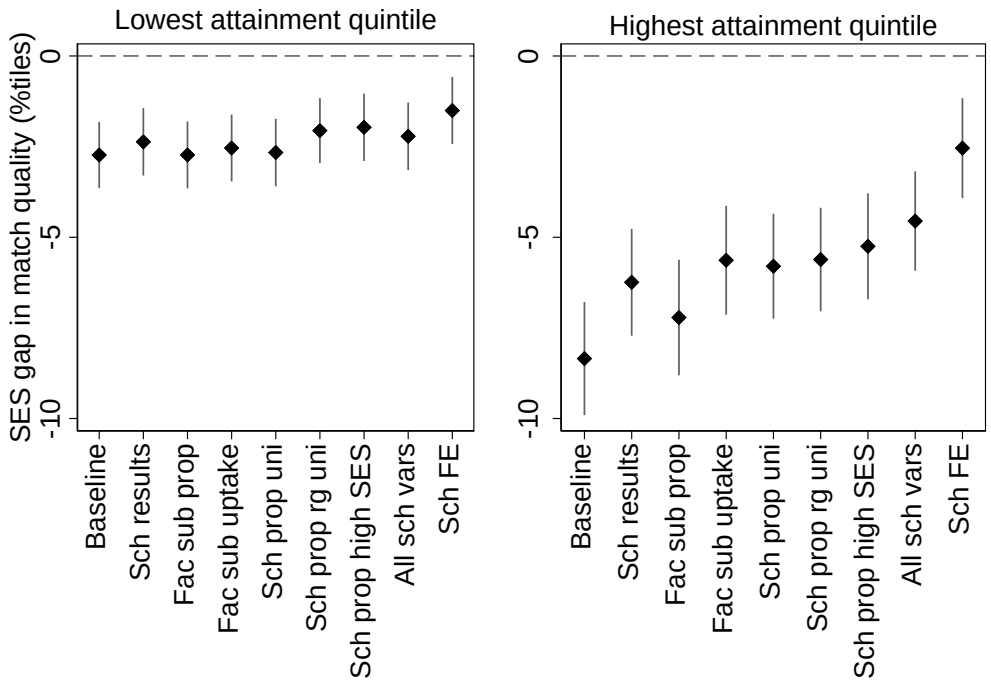
Source: NPD-HESA. n=138,969. Notes: Each point represents the SES match gap between groups 1 and 5 from specification 3, estimated for the top and bottom quintiles of the achievement distribution. The baseline controls are dummies for ethnicity, English as an Additional Language, Special Educational Needs, and gap year before college, and cubics in age 11 and age 16 exam results. We present the 95 percent confidence intervals, with standard errors clustered at the secondary school level.

Figure A7: Gender gaps in academic-based match, conditional on geography, subject, and schools



Source: NPD-HESA. n=138,969. Notes: Each point represents the gender match gap between groups 1 and 5 from specification 3, estimated for the top and bottom quintiles of the achievement distribution. The baseline controls are dummies for ethnicity, English as an Additional Language, Special Educational Needs, and gap year before college, and cubics in age 11 and age 16 exam results. We present the 95 percent confidence intervals, with standard errors clustered at the secondary school level.

Figure A8: SES gaps in academic-based match, conditional on school characteristics



Source: NPD-HESA. n=138,969. Notes: Each point represents the SES match gap between groups 1 and 5 from specification 3, estimated for the top and bottom quintiles of the achievement distribution. The baseline uses our main specification. We present the 95 percent confidence intervals, with standard errors clustered at the secondary school level. “Sch results” is the mean age 18 exam score at the school, “Fac sub prop” is facilitating subjects as a proportion of A-level subjects available, “Fac sub uptake” is facilitating subjects as a proportion of all A-levels taken, “Sch prop uni” is proportion of school students going to university, “Sch prop rg uni” is proportion of school students going to a Russell Group university, “Sch prop high SES” is the proportion of school students in the top SES quintile.

Table A1: Sum of total overmatch and undermatch

	Total Overmatch	Total Undermatch	Prop. to reallocate for balance
Overall Academic Match	**1,171,280**	**-1,176,518**	**0.11%**
Overall Earnings Match	**1,574,769**	**-1,620,956**	**0.72%**
Figure 3 Gender Academic Match			
Women	634,155	-710,391	-2.84%
Men	537,125	-466,127	3.54%
Figure 3 Gender Earnings Match			
Women	702,692	-1,087,532	-10.75%
Men	872,077	-533,424	12.05%

Source: NPD-HESA, HMRC n=138,969. Match is defined as the percentile difference in quality between the student and the course they enrolled on. Student quality is defined by their age 18 difficulty adjusted exam performance. Course quality is defined by two measures. Academic Quality defined by the course's median student's difficulty adjusted exam performance. Potential Earnings Quality defined by the course's median student's earnings five years after graduation. Columns 1 and 2 present the total over(under)match as the product of the number of individuals and their extent of over(under)match. Column 3 presents the proportion of mismatch required to balance overmatch and undermatch.

Table A2: SES and Gender Conditional Match Gaps across Additional Alternative Specifications

	Panel A: Academic-based match											
	Baseline		No diff adjustment		KS2-based		Degree size weights		631 subject level		Combined hons	
	Lowest	Highest	Lowest	Highest	Lowest	Highest	Lowest	Highest	Lowest	Highest	Lowest	Highest
Achievement quintile	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SES quintile (ref 5th)												
1	-2.73	-8.35	-2.09	-8.90	-1.65	-3.59	-2.83	-7.93	-2.32	-7.34	-2.71	-7.78
	(0.47)	(0.80)	(0.44)	(0.84)	(0.71)	(0.93)	(0.48)	(0.76)	(0.42)	(0.77)	(0.46)	(0.79)
2	-2.57	-5.02	-1.89	-5.45	-1.90	-1.66	-2.64	-4.73	-2.26	-4.37	-2.25	-5.00
	(0.39)	(0.51)	(0.37)	(0.54)	(0.59)	(0.61)	(0.39)	(0.48)	(0.36)	(0.48)	(0.38)	(0.50)
3	-1.19	-3.48	-0.55	-3.77	-1.86	-2.60	-1.24	-3.27	-1.24	-3.03	-1.16	-3.33
	(0.36)	(0.35)	(0.36)	(0.38)	(0.54)	(0.48)	(0.36)	(0.33)	(0.35)	(0.34)	(0.36)	(0.35)
4	-1.13	-1.99	-0.98	-2.28	-1.17	-1.52	-1.16	-1.85	-0.91	-1.79	-1.02	-1.92
	(0.33)	(0.27)	(0.33)	(0.30)	(0.52)	(0.42)	(0.33)	(0.26)	(0.32)	(0.26)	(0.32)	(0.27)
Women	0.69	-2.43	1.49	-3.31	0.81	-3.82	0.72	-2.26	0.98	-2.51	0.89	-2.26
	(0.24)	(0.25)	(0.22)	(0.27)	(0.38)	(0.36)	(0.24)	(0.24)	(0.23)	(0.24)	(0.24)	(0.24)
Clusters	2135	2005	2135	2005	2121	1983	2135	2005	2135	2005	2135	2005
n	27794	27786	27794	27786	26554	26580	27794	27786	27794	27786	27794	27786
	Panel B: Earnings-based match											
Achievement quintile	Lowest	Highest	Lowest	Highest	Lowest	Highest	Lowest	Highest				
SES quintile (ref 5th)												
1	-5.95	-7.92	-5.47	-7.97	-6.66	-0.42	-5.79	-7.76				
	(0.57)	(0.88)	(0.58)	(0.95)	(0.76)	(1.12)	(0.56)	(0.86)				
2	-3.05	-4.91	-2.64	-4.87	-3.72	0.19	-2.98	-4.84				
	(0.49)	(0.57)	(0.51)	(0.61)	(0.70)	(0.71)	(0.48)	(0.56)				
3	-1.78	-4.29	-1.45	-4.18	-2.48	-1.74	-1.73	-4.21				
	(0.45)	(0.45)	(0.46)	(0.47)	(0.61)	(0.58)	(0.44)	(0.44)				
4	-0.73	-2.24	-0.55	-2.27	-0.64	-0.94	-0.70	-2.21				
	(0.43)	(0.34)	(0.45)	(0.36)	(0.60)	(0.47)	(0.42)	(0.33)				
Women	-7.48	-8.06	-8.30	-10.07	-2.38	-6.94	-7.32	-7.94				
	(0.32)	(0.32)	(0.33)	(0.34)	(0.42)	(0.43)	(0.31)	(0.32)				
Clusters	2135	2005	2135	2005	2121	1983	2135	2005				
n	27794	27786	27794	27786	26554	26580	27794	27786				

Source: NPD-HESA. N=138.969. All specifications control for ethnicity, English as an Additional Language, Special Educational Needs, and gap year before college, and cubics in age 11 and age 16 exam results. These prior achievement controls are omitted for the KS2-based regressions.

Table A3: SES and Gender Conditional Match Gaps Using Within Subject Measures

	Academic Match				Earnings Match			
	Main Measure (No Preferences)		Within-subject (Subject Preference)		Main Measure (No Preferences)		Within-subject (Subject Preference)	
Achievement Quintile	Lowest (1)	Highest (2)	Lowest (3)	Highest (4)	Lowest (5)	Highest (6)	Lowest (7)	Highest (8)
SES quintile								
1	-2.73	-8.35	-3.54	-9.29	-5.95	-7.92	-8.00	-10.52
	(0.47)	(0.8)	(0.54)	(0.92)	(0.57)	(0.88)	(0.70)	(0.96)
2	-2.57	-5.02	-3.21	-5.38	-3.05	-4.91	-3.01	-5.74
	(0.39)	(0.51)	(0.48)	(0.59)	(0.49)	(0.57)	(0.57)	(0.64)
3	-1.19	-3.48	-1.63	-3.91	-1.78	-4.29	-1.61	-4.33
	(0.36)	(0.35)	(0.43)	(0.45)	(0.45)	(0.45)	(0.49)	(0.50)
4	-1.13	-1.99	-1.91	-2.19	-0.73	-2.24	-1.02	-2.50
	(0.33)	(0.27)	(0.40)	(0.37)	(0.43)	(0.34)	(0.46)	(0.40)
Women	0.69	-2.43	-0.36	-3.34	-7.48	-8.06	0.68	-2.65
	(0.24)	(0.25)	(0.27)	(0.34)	(0.32)	(0.32)	(0.31)	(0.37)
Clusters	2135	2005	2135	2005	2135	2005	2135	2005
N	27794	27786	27794	27786	27794	27786	27794	27786
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: NPD-HESA, HMRC n=138,969. Notes: All specifications control for ethnicity, English as an Additional Language, Special Educational Needs, and gap year before college, and cubics in age 11 and age 16 exam results. SES baseline is category 5 (highest quintile) while the other quintiles are included in the model but not reported here. Match is defined by the difference between student and course quality percentiles. Student quality defined by their percentile within in age 18 difficulty adjusted exam performance distribution for their undergraduate subject choice. Course quality is defined by two measures. Academic Quality defined by their percentile in the courses' median student's prior achievement distribution among students studying that undergraduate subject. Potential Earnings Quality defined by their percentile in the courses' median student earnings five years after graduation distribution among students studying that undergraduate subject. All quality measures are weighted by student enrolment, from which the deciles are obtained.

Technical Appendix

A.1 Difficulty Adjustments

Our approach to accounting for varying difficulty in age 18 examinations across different subjects follows Kelly (1976) and Coe et al. (2008). We calculate difficulty scores for all subjects based on the full set of examination results for the population of students taking age 18 examinations in 2008. To calculate the difficulty of each subject, we subtract each participating student's score in that subject from her average score across all other subjects, and sum the result across all participating students. This gives us a difficulty correction factor for each subject. As Kelly (1976) and Coe et al. (2008) note, these initial difficulty correction factors are likely to underestimate the variation in difficulty across subjects, since students who take 'hard' subjects tend to combine them with other 'hard' subjects, and those who take 'easy' subjects tend to combine them with other 'easy' subjects. We therefore 'correct' each student's score using the initial difficulty correction factors and repeat the process. We do this ten times. With each repetition, the difficulty correction get smaller, and after ten times, the required adjustments have effectively shrunk to zero.

Figure A2 and the accompanying table show the total difficulty correction factor applied to each subject. The units are "QCA points", and 30 points represents one A-Level grade. The 'easiest' subject, Persian, with a total difficulty correction of -51, is therefore 2.6 grades easier than the most difficult subject, Additional Mathematics, which has a total difficulty correction of 28. Intuitively, this means that students who take Persian tend to score higher in that subject than they do in others, while those who take Additional Mathematics tend to score lower.

NB: In the case of Persian and other minority languages, it may be that many students who take the subject already have some understanding of the language, which decreases the perceived difficulty of the subject using our method (see Ofqual, 2017). Only a small numbers of students take these minority language A-Levels – none of A-Level Persian, Bengali, Urdu, Panjabi, Turkish, Polish, or Russian, is taken by more than 100 students in our estimation sample.

A.2 Alternate Match Metrics

Appendix Table A2 presents a number of alternative specifications to supplement further the robustness checks discussed in Section 4. As discussed, we may be concerned that different students choose to study difference A-Level subjects, which are not of equal value. To attempt to reduce the effect of this on our estimated SES and gender gaps, we adjust our total points by a subject difficulty rating. Our first additional robustness specification in columns 3 and 4 presents the SES and gender gradients without this adjustment, to show that it is making little difference to our estimates.

A further alternative method of dealing with the potential endogeneity of A-Level subject choice is to alternative measures of student achievement to rank students and degrees (for academic-based match). We discuss the use of compulsory exams at age 16 in Section 4. Here we extend this even further to rank students based on their qualifications at age 11 (Key Stage 2). At age 11 all students in England take the same three exams in English, maths and science, which completely removes the issue of student choice. We sum the scores across subjects for each student and then calculate their national percentile rank. The degree ranks are in turn calculated on the basis of the median students using these measures. The results are similar for this very early ranking of student achievement, with a slight reduction in the SES gradient for high-attainers.

Finally, columns 11 and 12 consider an alternative way to specific students on combined honours degrees. About 10% of our sample are doing these types of degrees, where each subject studied falls into more than one group of our 23 subject classification. In our baseline results we assign these students their highest weighted subject, or if weighting is equal they are given the first subject listed. In these results, we assign them according to their highest 2 weighted subjects, so there are 117 different categories including single and combined honours. Note it is not possible to carry out this test for our earnings-based measure because we do not observe later earnings for these combined honours degrees. Our results are again very similar using this alternative specification.