ONLINE APPENDIX

"High School Majors and Future Earnings"

Gordon B. Dahl, Dan-Olof Rooth, and Anders Stenberg

Appendix A. Robustness Checks

A.1 Alternative Definitions for Earnings

In this appendix we provide a variety of robustness checks, starting with an examination of alternative definitions for the earnings variable. Appendix Table A6 estimates whether the probability of being in our main sample, which excludes individuals with zero or low earnings between the ages of 37-39, jumps at the GPA cutoff. Six out of 30 estimates are statistically significant. The table also estimates whether having earnings above the threshold in all three years, versus at least one year, jumps at the cutoff. Seven out of 30 estimates are statistically significant. Both of these exercises are indicative of a small extensive labor market response to field of study. Therefore, we probe the robustness of our log earnings variable by using two alternative earnings measures which do not exclude any observations.

Our first alternative earnings measure uses earnings in levels as the outcome, including low earnings and zeros in any year (i.e., using the full population). The pattern of estimates appearing in Appendix Table A7 are similar to the baseline estimates using logs. The magnitudes are also roughly comparable, with the correlation between the two sets of estimates being 0.97.[28] As a second alternative, we use earnings rank. We calculate each individual's rank in the year-specific population earnings distribution for all individuals in Sweden between the ages of 16 to 64. Roughly the same number of estimates are statistically significant using this measure compared to our log earnings measure, with the correlation between the two sets of estimates being 0.95.[29]

A.2 Alternative RD Specifications

Alternative RD specifications appear in Appendix Table A8. For each robustness check, we report the correlation of the robustness estimates with the baseline estimates. We begin by exploring different parametric models for the RD regression, finding that the addition of quadratic terms in the running variable, a reduced bandwidth around the cutoff, or the

[28]For example, individuals choosing Engineering over Natural Science experience an earnings increase of $4,565 per year. Since the average earnings for this group is $54,668, this translates into an 8.4% increase in average earnings. This compares to a 6.4% increase when using log earnings. We weight the correlation, and the other correlations appearing in this section, by the inverse of the sum of the squared standard errors of the two estimates. While the baseline estimates and the levels estimates are both consistently estimated, they are measured with standard errors, and so the correlation coefficient could be biased.

[29]While not shown, we also explored 3 other modifications of our log earnings measure: (i) excluding publicly provided parental leave and sickness benefits from our earnings measure, (ii) adjusting the earnings threshold to account for inflation, but not wage growth, and (iii) using earnings between the ages of 39-41 instead of 37-39 (the oldest ages for which we observe occupation). All three of these modifications result in estimates which are similar to baseline.

addition of first-second choice specific intercept terms (i.e., 30 intercepts) does not appreciably change the estimates (although the standard errors sometimes increase for sparsely populated choice margins). Our next set of specification checks relax the parametric assumption of a two slope model. The sets of estimates from both the 12- and 60-slope RD models yield similar results compared to our baseline.[30] To see this visually, we plot the estimates for each of the first-second best combinations for the 12- and 60-slope models against the 2-slope model in Appendix Figure A8. The advantage of the 2-slope model, particularly relative to the 60-slope model, is that the estimates are substantially more precise for many of the combinations. Finally, we estimate our baseline model, but exclude the years 1982-84, when individuals were given a 0.5 GPA bonus for the first field on their ranking list and a 0.2 GPA bonus for the second field on their ranking list.[31,32]

A.3 Multiple Inference Adjustments

The findings are further robust to multiple inference adjustments using the False Discovery Rate (FDR) control (see, e.g., Anderson 2008). For the reduced form results in Table 4, 15 out of 17 estimates remain statistically significant, and for the fuzzy RD results in Table 5, 15 out of 17 estimates likewise remain significant (see Appendix Table A9).

A.4 Comparison to Kirkeboen et al. (2016)

Our paper leverages Kirkeboen et al.'s (2016) methodology (KLM) to account for second-best choices. We have more data and therefore the ability to implement a fuzzy RD design (allowing different slopes in the running variable of GPA on each side of the cutoff and using triangular weights) instead of their more parsimonious IV (including GPA as a single control variable). In Appendix Table A10, we explore what happens if we use their IV specification. When Engineering or Natural Science is a second choice, the estimates are all larger when using fuzzy RD. When Engineering is a second choice, all 4 estimates are statistically different at the 10% level and when Natural Science is a second choice, 2 out of 4 estimates are statistically different. For the remaining margins, the results are fairly similar.

Due to the nature of our data, our setting has several additional practical advantages compared to KLM's study of college major returns. First, individuals attend their local high

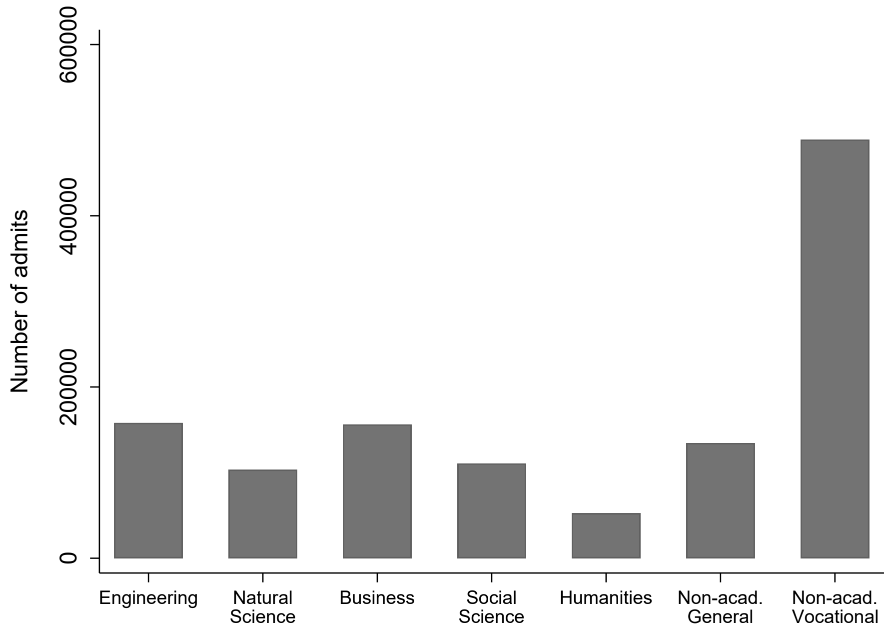
[30] Appendix Figure A7 illustrates why the 2 slope model yields similar estimates to the 12 slope model.

[31] This means that for these three years, the allocation mechanism was not strategy-proof. Instead, individuals could have been strategic about not putting their most preferred field first if they thought they wouldn't get in even with the GPA bonus.

[32] While not shown in the table, we also tried including a proxy measure of average class size (number of students divided by 30) as an additional regressor. This has virtually no effect on the estimates.

school and are therefore only choosing majors, so we do not have the confounding factor of institution choice. Second, we do not have to combine academic majors, whereas KLM needs to collapse college majors into 10 broad categories. Third, we observe earnings over two decades later, while KLM looks at earnings 8 years after admission (i.e., roughly 4 years after completing college).

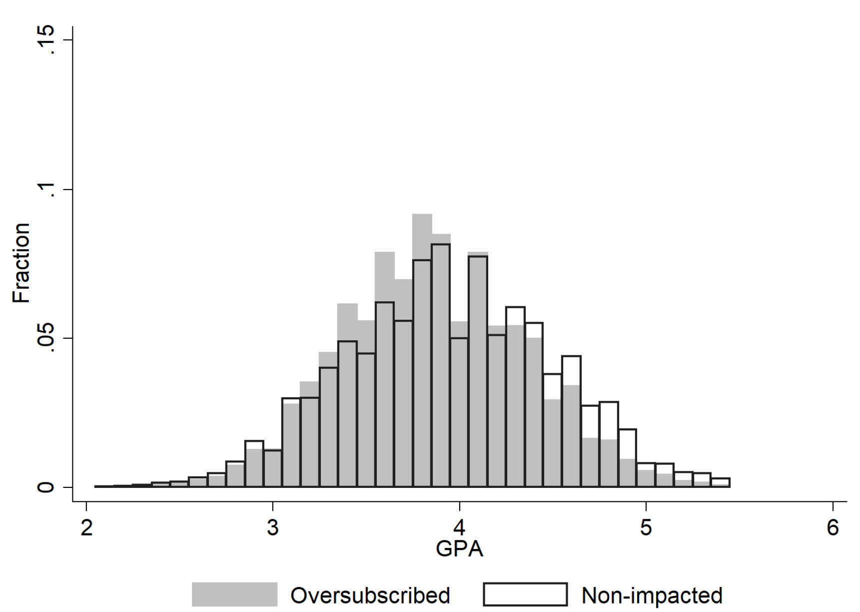
Figure A1. Number of students admitted to each major.



Notes: Admission to high school majors between 1977-1991. N=1,208,269.

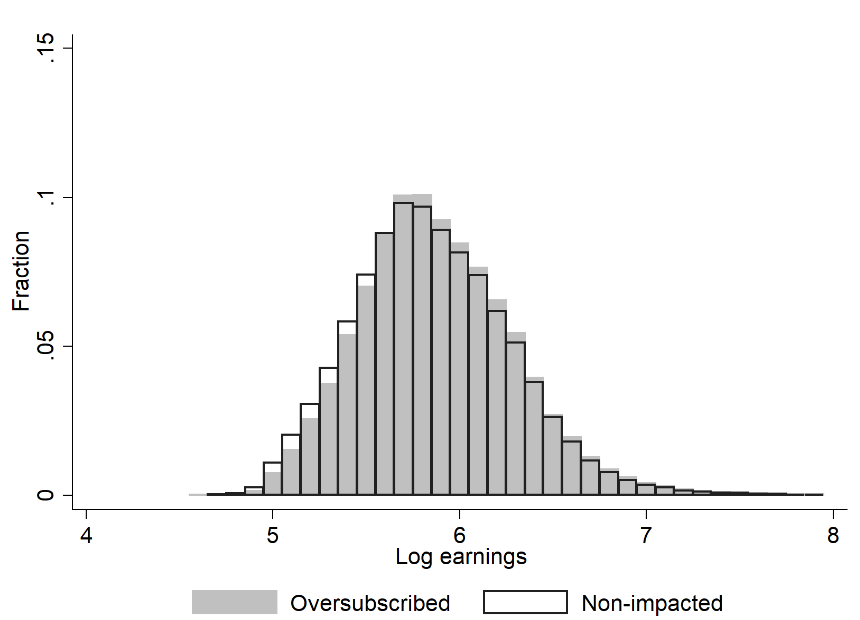
Figure A2. GPA and annual earnings in oversubscribed and non-impacted programs.

Panel A: Individual GPA



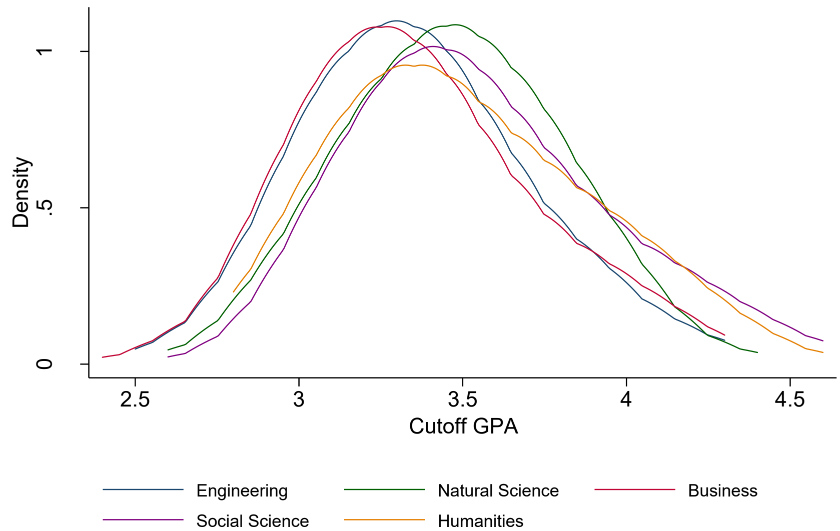
Notes: The grey bars plot the distribution of GPA for individuals in our baseline estimation sample of oversubscribed programs. The white bars plot the distribution of GPAs for individuals in non-impacted academic programs.

Panel B: Individual annual earnings at age 38



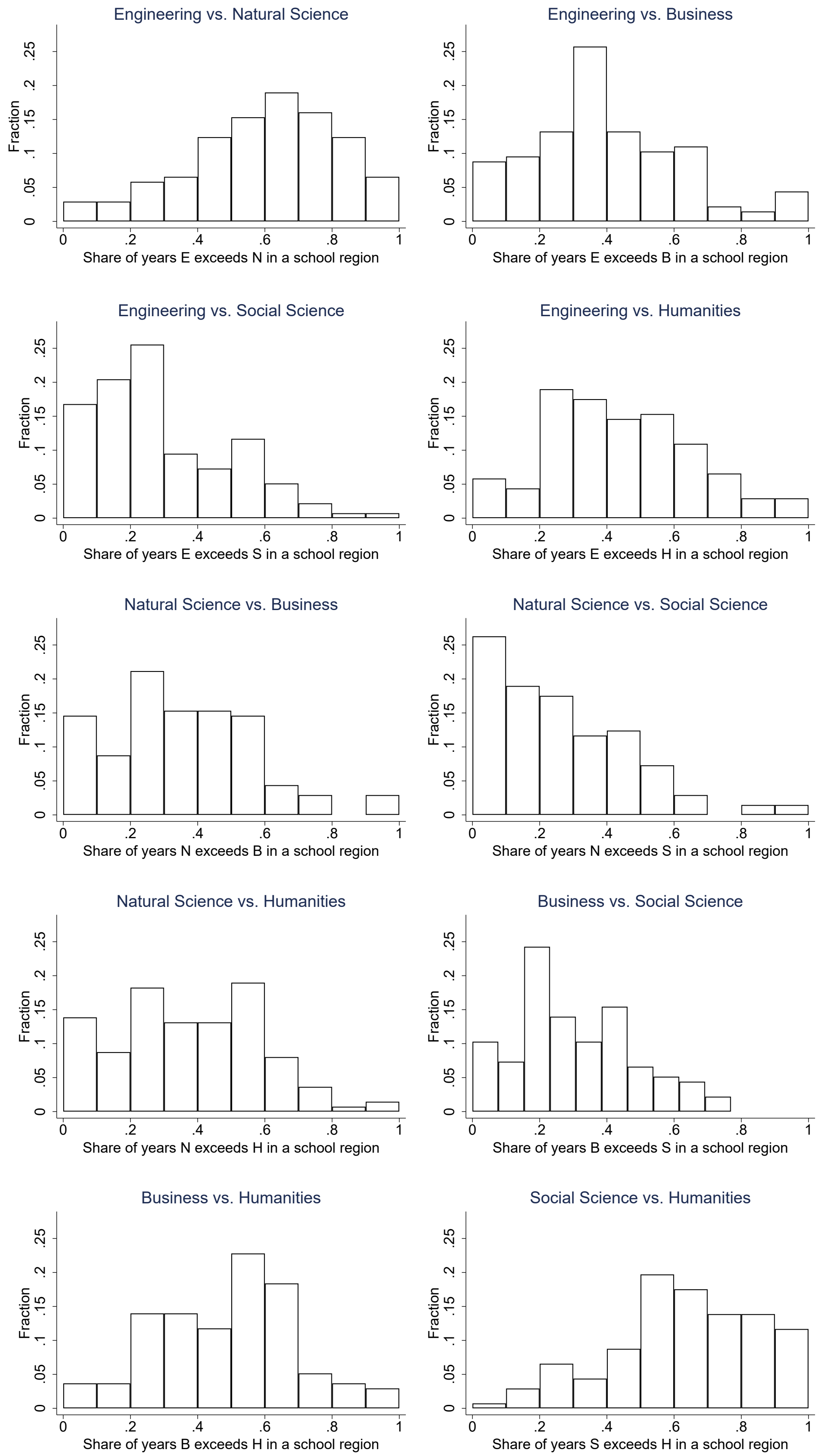
Notes: The grey bars plot the distribution of GPA for individuals in our baseline estimation sample of oversubscribed programs. The white bars plot the distribution of GPAs for individuals in non-impacted academic programs.

Figure A3. Distribution of GPA cutoffs by high school major.



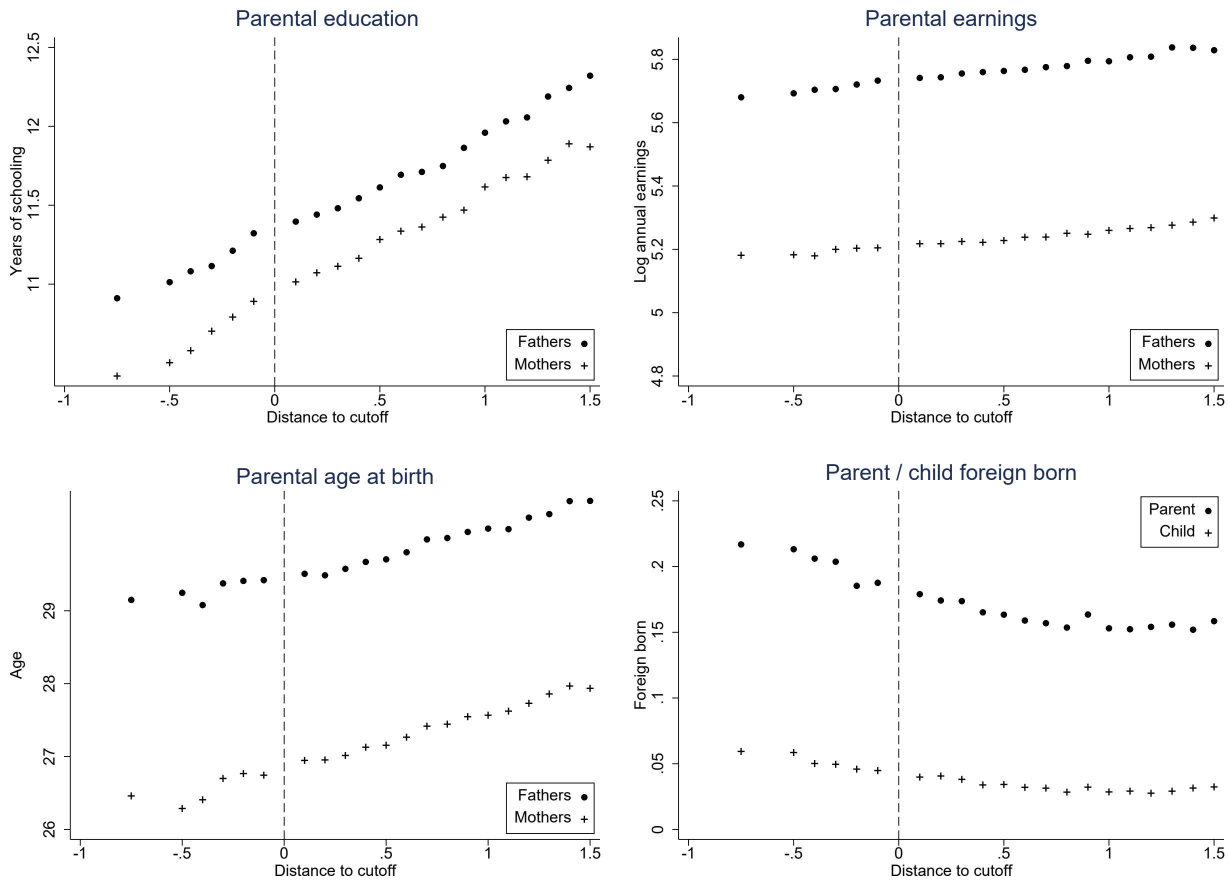
Notes: Kernel density estimates of GPA cutoffs by major, using an Epanechnikov kernel and a bandwidth of 0.2.

Figure A4. Within school region variation over time in relative major cutoffs.



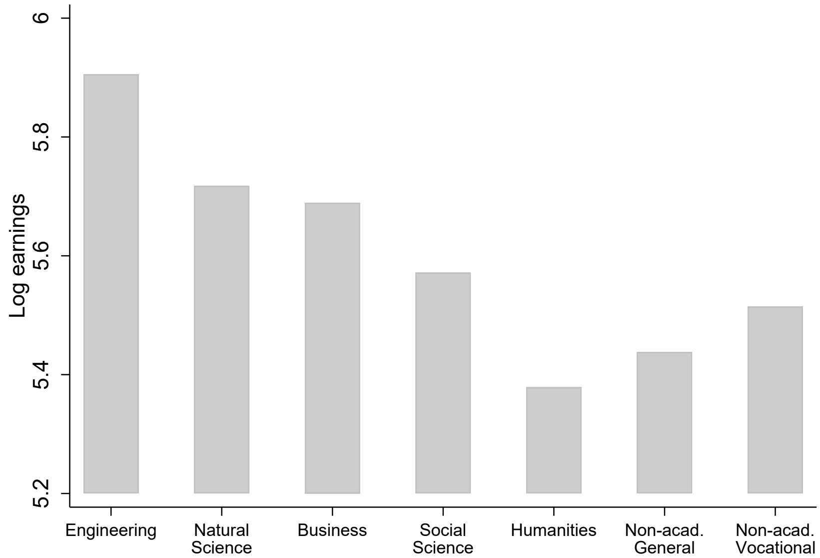
Notes: For each combination of majors, the graphs plot the distribution across school regions of the share of years a cutoff for one major exceeds another (conditional on at least one of the two majors being oversubscribed).

Figure A5. Smoothness of predetermined variables at the cutoff.



Notes: Each dot is the average for the relevant outcome in a .1 GPA bin, except for the leftmost dot which is a .5 bin due to sparsity. GPA is measured relative to a normalized cutoff of 0. Parent foreign born is a dummy for whether at least one parent is foreign born. See Appendix Table A4 and the text for further details.

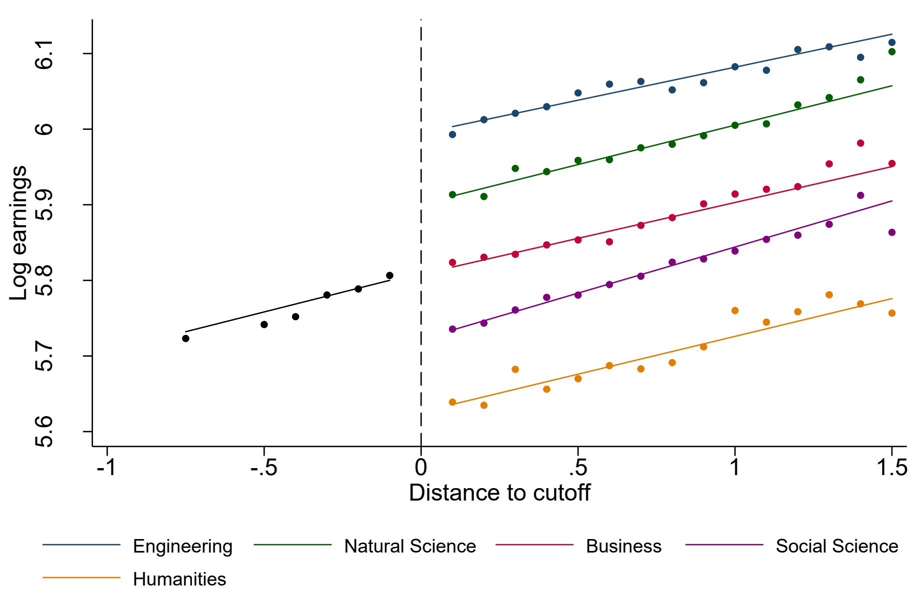
Figure A6. Log earnings for program completers with a GPA 3.4 or 3.5.



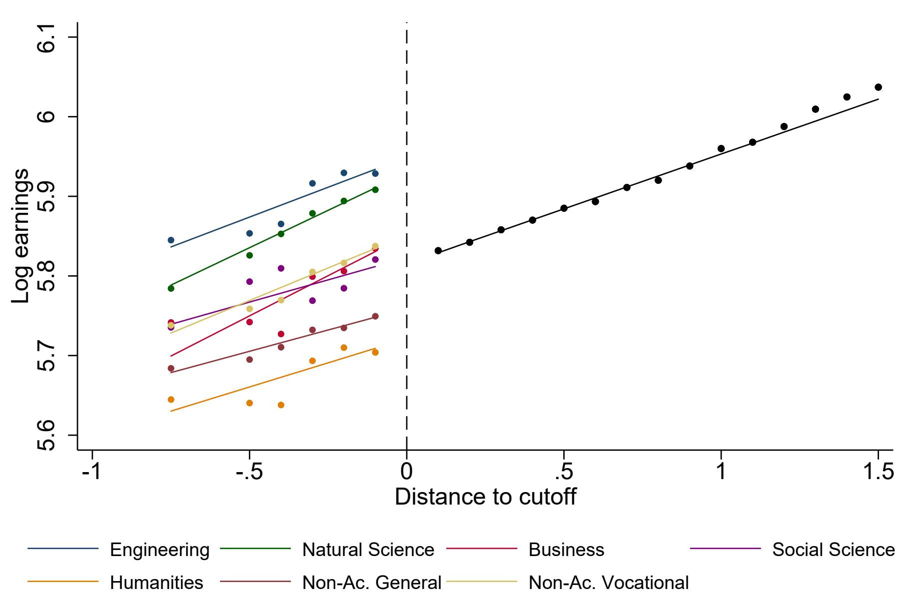
Notes: The figure uses the same sample as in Figure 1, but limited to students with a GPA of 3.4 or 3.5. Adult earnings measured between the ages of 37-39. N=90,988.

Figure A7. Comparison of 2 versus 12 slope models.

Panel A: Single slope below the cutoff, 5 separate slopes above the cutoff

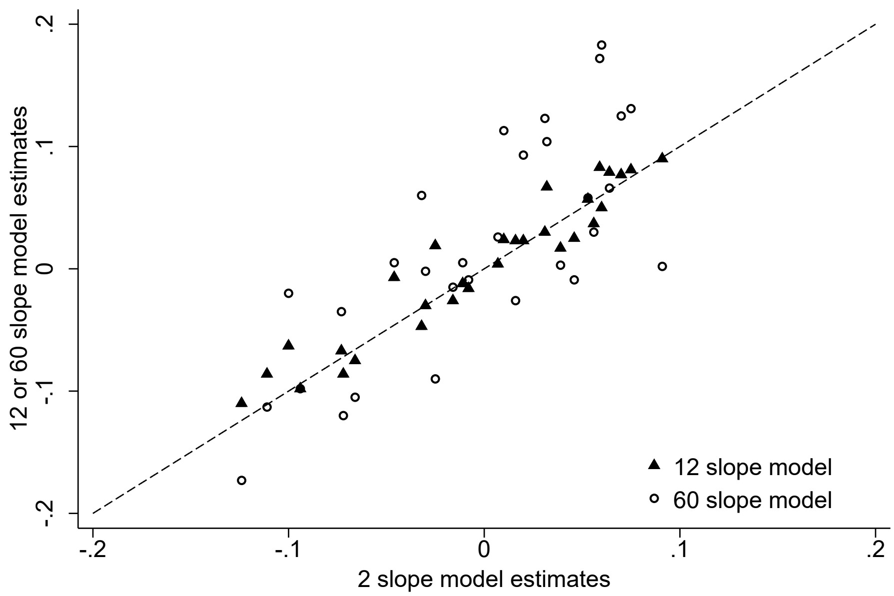


Panel B: 7 separate slopes below the cutoff, single slope above the cutoff



Notes: The top figure plots averages of log annual earnings in GPA bins, allowing for separate slopes for each of the five first-best choices to the right of the cutoff. While the graph makes clear the intercepts for the various first-best choices differ, the slopes are remarkably similar to one another. The bottom figure conducts a similar exercise, plotting the averages separately for each of the 7 next-best choices. Again, the intercepts for the various second-best choices differ, but not as much as they did for first-best choices in the top graph. And while the data are noisier to the left of the cutoff due to smaller sample sizes, the slopes are again similar to each other. The trend lines are RD estimates using the underlying data, no covariates, and triangular weights.

Figure A8. Comparison of fuzzy RD estimates using the 60 slope, 12 slope, and 2 slope models.



Notes: There are 30 estimates for each model, one for each first–second best choice combination. The dashed line is the 45° line.

Table A1. Number of observations by first-second choice combination.

	Second choice						
First choice	Engineering	Natural Science	Business	Social Science	Humanities	Non-acad. General	Non-acad. Vocational
Engineering	--	31,910	12,023	3,375	552	4,504	11,246
Natural Science	8,833	--	2,345	5,617	674	462	899
Business	7,656	6,687	--	29,850	8,135	18,254	13,559
Social Science	1,723	8,392	15,714	--	15,279	7,963	3,394
Humanities	413	566	2,305	7,202	--	2,233	1,269

Notes: Baseline sample of 233,034 individuals.

Table A2. Summary statistics for applicants with a first-choice academic program.

Variables	Oversubscribed programs	Share missing	Non-impacted programs	Share missing
Parent characteristics:				
Father age	29.74	.05	29.99	.07
Mother age	27.20	.02	27.33	.02
Father schooling	11.60	.05	11.29	.06
Mother schooling	11.23	.02	10.82	.02
Father earnings	5.76	.18	5.75	.20
Mother earnings	5.23	.25	5.20	.29
Foreign born parent	.16	0	.16	0
Child characteristics:				
Foreign born	.03	0	.03	0
Female	.51	0	.50	0
Age at application	15.99	0	15.99	0
GPA	3.86	0	3.94	0
Child outcomes:				
College degree	.45	0	.45	0
Log earnings	5.84	0	5.81	0
N	233,034		194,024	

Note: Years span 1977-1991. Parent and child characteristics are measured in the year of application (when the child is roughly 16 years old). Parent age refers to age at the time of the child's birth. Years of schooling inferred from highest education level. Earnings are measured between the ages of 37-39 and are converted to year 2016 US dollars using an exchange rate of 8.5 SEK to 1 USD. GPA is standardized to be mean 0 and variance 1 in the entire population, including those who do not apply to secondary school.

Table A3. Comparison of major cutoffs across years within the same school region.

	Fraction of years with a higher cutoff		
Major combinations	1st major	2nd major	No difference
Engineering vs. Natural Science	.37	.25	.38
Engineering vs. Business	.28	.42	.30
Engineering vs. Social Science	.21	.53	.27
Engineering vs. Humanities	.31	.38	.31
Natural Science vs. Business	.24	.46	.30
Natural Science vs. Social Science	.18	.51	.31
Natural Science vs. Humanities	.24	.38	.39
Business vs. Social Science	.24	.48	.28
Business vs. Humanities	.37	.32	.31
Social Science vs. Humanities	.47	.21	.32

Notes: The table reports the average fraction of years with a higher cutoff for one major compared to another within the same school region. If both majors have a cutoff in a given year in the same school region, we compare the two to determine which is higher. If one major has a cutoff, but the other does not, we record the major with the cutoff as having a higher cutoff. "No difference" can either reflect that both majors have cutoffs which are equal or that neither major was oversubscribed.

Table A4. Balancing tests for pre-determined characteristics.

	Years schooling Father	Years schooling mother	Log earnings father	Log earnings mother	Age at birth father	Age at birth mother	Foreign born parent	Child foreign born
	-.084	-.039	-.008	-.005	-.169	.013	-.004	-.003
	(.059)	(.054)	(.010)	(.009)	(.138)	(.118)	(.009)	(.005)
N	249,424	259,434	214,308	196,991	249,174	258,879	263,856	263,856

Notes: Each column is an estimate from a separate RD regression which uses the 2 slope model; linear functions of the running variable of normalized GPA; a window of -1.0 to 1.5; triangular weights; fixed effects for year, school region, and program. There is a common jump for all first-second best choice combinations. Standard errors in parentheses.

** p<.10, ** p<.05, *** p<.01*

Table A5. Heterogeneity by age, gender, and parental education.

	Baseline (age 37-39)	Age 27-29	Males	Females	Low parental education	High parental education
E vs. N	.064***	.029**	.052***	.010	.061***	.066***
	(.017)	(.014)	(.019)	(.023)	(.018)	(.017)
E vs. B	.007	.001	.002	-.036	-.002	.013
	(.018)	(.014)	(.020)	(.023)	(.019)	(.019)
E vs. S	.059**	-.039*	.052*	.023	.068**	.053**
	(.025)	(.021)	(.027)	(.033)	(.028)	(.026)
E vs. H	.070*	.003	.064	.023	.065	.074*
	(.039)	(.031)	(.045)	(.055)	(.055)	(.042)
E vs. G	.010	-.027**	.001	-.020	.007	.015
	(.017)	(.013)	(.020)	(.039)	(.019)	(.020)
E vs. V	.020	.016	.007	-.015	.022	.019
	(.015)	(.012)	(.017)	(.030)	(.016)	(.016)
N vs. E	.039	-.011	.014	.041	.013	.050**
	(.025)	(.020)	(.027)	(.029)	(.028)	(.025)
N vs. B	.056**	.055**	.060*	.035	.027	.068**
	(.028)	(.023)	(.033)	(.031)	(.036)	(.029)
N vs. S	.075***	.023	.048	.075**	.055*	.084***
	(.028)	(.023)	(.032)	(.030)	(.031)	(.028)
N vs. H	.060	.010	-.012	.071*	.108**	.036
	(.037)	(.030)	(.063)	(.038)	(.045)	(.041)
N vs. G	.031	-.050	.106	-.065	.096	-.063
	(.052)	(.042)	(.074)	(.061)	(.061)	(.072)
N vs. V	-.032	-.051	-.073	-.024	-.010	-.040
	(.040)	(.033)	(.055)	(.044)	(.050)	(.044)
B vs. E	.046**	.014	.045*	.020	.020	.065***
	(.021)	(.017)	(.024)	(.024)	(.022)	(.022)
B vs. N	.091***	.071***	.102***	.066***	.068***	.105***
	(.017)	(.014)	(.021)	(.019)	(.019)	(.018)
B vs. S	.053***	.025*	.076***	.035*	.040**	.061***
	(.016)	(.014)	(.019)	(.018)	(.017)	(.017)
B vs. H	-.008	.016	.003	-.015	-.007	-.007
	(.018)	(.014)	(.026)	(.019)	(.018)	(.019)
B vs. G	-.011	-.010	.021	-.033***	-.017	-.003
	(.010)	(.008)	(.015)	(.012)	(.011)	(.011)
B vs. V	-.016	-.022**	-.008	-.027**	-.021*	-.007
	(.011)	(.009)	(.016)	(.014)	(.012)	(.012)
S vs. E	-.072***	-.069***	-.097***	-.069**	-.054*	-.078***
	(.026)	(.022)	(.033)	(.029)	(.031)	(.028)
S vs. N	.016	.022	-.001	.012	.016	.017
	(.018)	(.015)	(.024)	(.019)	(.020)	(.019)
S vs. B	-.066***	-.013	-.085***	-.060***	-.071***	-.062***
	(.014)	(.011)	(.019)	(.015)	(.015)	(.014)
S vs. H	-.030*	-.000	-.087***	-.026	-.020	-.035**
	(.017)	(.014)	(.024)	(.018)	(.018)	(.018)
S vs. G	-.073***	-.058***	-.099***	-.067***	-.067***	-.078***
	(.013)	(.011)	(.021)	(.015)	(.015)	(.015)
S vs. V	-.094***	-.063***	-.207***	-.075***	-.070***	-.111***
	(.016)	(.014)	(.030)	(.018)	(.019)	(.018)
H vs. E	.032	-.076	-.054	.019	.011	.057
	(.141)	(.118)	(.558)	(.115)	(.164)	(.179)
H vs. N	-.025	-.019	-.260**	.022	-.020	-.029
	(.039)	(.032)	(.104)	(.038)	(.055)	(.044)
H vs. B	-.124***	-.054***	-.196***	-.113***	-.125***	-.125***
	(.021)	(.017)	(.043)	(.022)	(.023)	(.023)
H vs. S	-.046**	-.047***	-.163***	-.032	-.035	-.056***
	(.021)	(.018)	(.031)	(.022)	(.022)	(.021)
H vs. G	-.100***	-.073***	-.149*	-.089***	-.052	-.147***
	(.028)	(.023)	(.077)	(.029)	(.036)	(.030)
H vs. V	-.111***	-.091***	-.403***	-.091***	-.137***	-.085**
	(.031)	(.026)	(.149)	(.031)	(.038)	(.036)
Corr.	.83		.77		.91	
N	233,034	220,360	233,034		232,882	

Notes: See notes to Table 4 and text for details. Standard errors in parentheses. N males = 113,893, females = 119,195, low-skilled parents = 95,825 and high-skilled parents = 136,654. * $p<.10$, ** $p<.05$, *** $p<.01$

Table A6. Probability of above threshold earnings and earnings in all three years.

	Above threshold earnings	Earnings in all three years
E vs. N	.024*	-.006
	(.013)	(.016)
E vs. B	.024*	-.018
	(.013)	(.017)
E vs. S	.031	.041
	(.020)	(.025)
E vs. H	-.014	-.019
	(.030)	(.039)
E vs. G	.029**	-.020
	(.012)	(.016)
E vs. V	.016	.010
	(.011)	(.014)
N vs. E	.013	.035
	(.019)	(.024)
N vs. B	.006	.042
	(.022)	(.027)
N vs. S	-.010	.070**
	(.022)	(.030)
N vs. H	-.020	.055
	(.031)	(.038)
N vs. G	.009	.004
	(.040)	(.052)
N vs. V	.030	.015
	(.033)	(.041)
B vs. E	.003	.036*
	(.015)	(.020)
B vs. N	-.003	.029*
	(.013)	(.016)
B vs. S	-.012	.065***
	(.013)	(.017)
B vs. H	-.036**	.022
	(.016)	(.020)
B vs. G	.009	.002
	(.008)	(.010)
B vs. V	-.005	.019*
	(.009)	(.011)
S vs. E	-.008	.023
	(.018)	(.025)
S vs. N	.015	.030
	(.014)	(.018)
S vs. B	.003	.009
	(.011)	(.014)
S vs. H	-.031**	.023
	(.016)	(.020)
S vs. G	.024**	-.007
	(.011)	(.015)
S vs. V	.007	.015
	(.013)	(.018)
H vs. E	-.054	.333**
	(.100)	(.145)
H vs. N	-.014	.047
	(.032)	(.042)
H vs. B	.008	.006
	(.019)	(.025)
H vs. S	.006	.065**
	(.020)	(.025)
H vs. G	.036	-.032
	(.027)	(.035)
H vs. V	.024	.036
	(.029)	(.038)
N	250,522	233,034

Notes: Sample in first column includes all individuals with earnings, including zeros and below-threshold earnings. Earnings in all three years is defined as having above-threshold earnings in all three years (when aged 37-39). Standard errors in parentheses.
** p<.10, ** p<.05, *** p<.01*

Table A7. Robustness to alternative earnings measures.

	Baseline	Earnings in levels	Earnings rank
E vs. N	.064***	4.565***	.032***
	(.017)	(1.095)	(.012)
E vs. B	.007	.182	.005
	(.018)	(1.279)	(.012)
E vs. S	.059**	3.340**	.030*
	(.025)	(1.634)	(.018)
E vs. H	.070*	4.379*	.043
	(.039)	(2.417)	(.028)
E vs. G	.010	-1.361	.006
	(.017)	(1.122)	(.012)
E vs. V	.020	.272	.029***
	(.015)	(1.023)	(.010)
N vs. E	.039	3.432**	.016
	(.025)	(1.680)	(.017)
N vs. B	.056**	3.006	.037*
	(.028)	(2.055)	(.019)
N vs. S	.075***	4.934***	.053***
	(.028)	(1.800)	(.020)
N vs. H	.060	2.267	.056**
	(.037)	(2.222)	(.026)
N vs. G	.031	-.064	.021
	(.052)	(3.064)	(.037)
N vs. V	-.032	-3.060	-.009
	(.040)	(2.476)	(.029)
B vs. E	.046**	5.435***	.021
	(.021)	(1.564)	(.014)
B vs. N	.091***	7.786***	.045***
	(.017)	(1.223)	(.012)
B vs. S	.053***	4.073***	.043***
	(.016)	(1.074)	(.011)
B vs. H	-.008	.696	.020
	(.018)	(1.029)	(.013)
B vs. G	-.011	-.773	-.004
	(.010)	(.681)	(.007)
B vs. V	-.016	-.636	.004
	(.011)	(.845)	(.008)
S vs. E	-.072***	-2.283	-.030*
	(.026)	(1.775)	(.018)
S vs. N	.016	.993	.004
	(.018)	(1.202)	(.013)
S vs. B	-.066***	-4.568***	-.029***
	(.014)	(1.095)	(.010)
S vs. H	-.030*	-1.188	.006
	(.017)	(1.001)	(.013)
S vs. G	-.073***	-4.995***	-.048***
	(.013)	(.825)	(.010)
S vs. V	-.094***	-5.195***	-.045***
	(.016)	(1.079)	(.012)
H vs. E	.032	1.299	.085
	(.141)	(8.461)	(.097)
H vs. N	-.025	-.493	-.000
	(.039)	(2.474)	(.028)
H vs. B	-.124***	-8.127***	-.055***
	(.021)	(1.269)	(.016)
H vs. S	-.046**	-3.026**	-.009
	(.021)	(1.268)	(.016)
H vs. G	-.100***	-6.873***	-.066***
	(.028)	(1.579)	(.021)
H vs. V	-.111***	-6.517***	-.053**
	(.031)	(1.728)	(.023)
N	233,034	250,522	250,522

Notes: See notes to Table 4. Column 1 is the baseline results reported in Table 5, while the other two specifications include zero and below-threshold earnings. Column 2 uses earnings in levels measured in real terms relative to 2016, converted to U.S. dollars using an exchange rate of 8.50 Swedish crowns per dollar. Column 3 uses earnings rank. Standard errors in parentheses. E, N, B, S, H, G, N stand for Engineering, Natural Science, Business, Social Science, Humanities, General non-academic, and Vocational non-academic, respectively.
** p<.10, ** p<.05, *** p<.01*

Table A8. Specification checks.

Margin	Baseline	Quadratic	Smaller bandwidth	1st-2nd intercepts	12 slopes	60 slopes	Excluding 1982-84
E vs. N	.064***	.071***	.065***	.072***	.079**	.066*	.060***
	(.017)	(.024)	(.021)	(.019)	(.032)	(.036)	(.020)
E vs. B	.007	.014	.013	.026	.004	.026	.001
	(.018)	(.024)	(.022)	(.032)	(.028)	(.039)	(.020)
E vs. S	.059**	.067**	.083***	.080	.083**	.172***	.048*
	(.025)	(.032)	(.031)	(.059)	(.041)	(.060)	(.028)
E vs. H	.070*	.077*	.066	.231	.077	.125	.065
	(.039)	(.043)	(.051)	(.206)	(.050)	(.082)	(.046)
E vs. G	.010	.017	.029	-.024	.024	.113***	.020
	(.017)	(.024)	(.022)	(.026)	(.020)	(.038)	(.020)
E vs. V	.020	.027	.036*	.025	.023	.093***	.017
	(.015)	(.022)	(.019)	(.017)	(.019)	(.030)	(.016)
N vs. E	.039	.045	.0294	.013	.017	.003	.047*
	(.025)	(.030)	(.031)	(.033)	(.047)	(.059)	(.027)
N vs. B	.056**	.062*	.055	.043	.037	.030	.044
	(.028)	(.032)	(.035)	(.063)	(.040)	(.067)	(.030)
N vs. S	.075***	.082**	.088**	.080	.081*	.131*	.063**
	(.028)	(.032)	(.036)	(.051)	(.047)	(.069)	(.030)
N vs. H	.060	.067*	.102*	.246**	.050	.183*	.061
	(.037)	(.040)	(.056)	(.105)	(.050)	(.106)	(.040)
N vs. G	.031	.038	.054	.071	.030	.123	-.006
	(.052)	(.055)	(.068)	(.084)	(.058)	(.127)	(.056)
N vs. V	-.032	-.025	.001	-.020	-.047	.060	-.043
	(.040)	(.044)	(.058)	(.058)	(.048)	(.122)	(.043)
B vs. E	.046**	.052**	.037	.067**	.025	-.009	.049**
	(.021)	(.026)	(.025)	(.028)	(.036)	(.044)	(.023)
B vs. N	.091***	.097***	.067***	.075**	.090***	.002	.084***
	(.017)	(.023)	(.021)	(.030)	(.029)	(.038)	(.020)
B vs. S	.053***	.059***	.052***	.062***	.057*	.058	.045**
	(.016)	(.022)	(.019)	(.019)	(.031)	(.035)	(.018)
B vs. H	-.008	-.001	-.007	-.021	-.016	-.009	-.008
	(.018)	(.023)	(.021)	(.028)	(.029)	(.037)	(.020)
B vs. G	-.011	-.006	-.009	-.014	-.012	.005	-.010
	(.010)	(.016)	(.012)	(.011)	(.013)	(.017)	(.012)
B vs. V	-.016	-.010	-.013	-.013	-.026*	-.015	-.018
	(.011)	(.017)	(.014)	(.012)	(.015)	(.020)	(.012)

Table A8. Specification checks, continued.

Margin	Baseline	Quadratic	Smaller bandwidth	1st-2nd intercepts	12 slopes	60 slopes	Excluding 1982-84
S vs. E	-.072***	-.065**	-.081**	-.069*	-.086**	-.120**	-.068**
	(.026)	(.031)	(.032)	(.041)	(.042)	(.057)	(.029)
S vs. N	.016	.022	.009	.018	.023	-.026	.007
	(.018)	(.024)	(.023)	(.035)	(.031)	(.040)	(.021)
S vs. B	-.066***	-.060***	-.071***	-.073***	-.075***	-.105***	-.075***
	(.014)	(.019)	(.017)	(.016)	(.024)	(.027)	(.016)
S vs. H	-.030*	-.024	-.021	-.031	-.030	-.002	-.026
	(.017)	(.024)	(.021)	(.020)	(.030)	(.034)	(.019)
S vs. G	-.073***	-.068***	-.065***	-.063***	-.067***	-.035	-.073***
	(.013)	(.020)	(.016)	(.015)	(.016)	(.025)	(.015)
S vs. V	-.094***	-.088***	-.095***	-.105***	-.098***	-.098***	-.096***
	(.016)	(.022)	(.020)	(.022)	(.020)	(.033)	(.018)
H vs. E	.032	.055	.116	1.329	.067	.104	.027
	(.141)	(.152)	(.210)	(3.550)	(.196)	(.412)	(.157)
H vs. N	-.025	-.018	-.024	-.039	.019	-.090	-.044
	(.039)	(.044)	(.054)	(.076)	(.052)	(.093)	(.045)
H vs. B	-.124***	-.118***	-.132***	-.108***	-.110***	-.173***	-.127***
	(.021)	(.026)	(.025)	(.031)	(.030)	(.047)	(.024)
H vs. S	-.046**	-.039	-.031	-.086**	-.007	.005	-.055**
	(.021)	(.027)	(.025)	(.035)	(.037)	(.048)	(.024)
H vs. G	-.100***	-.092***	-.079**	-.082*	-.063*	-.020	-.098***
	(.028)	(.034)	(.034)	(.044)	(.033)	(.067)	(.032)
H vs. V	-.111***	-.103***	-.103***	-.142***	-.086**	-.113	-.106***
	(.031)	(.037)	(.039)	(.051)	(.037)	(.073)	(.033)
Corr. w/ baseline	1.00	.99	.98	.95	.97	.76	.99
N	233,034	233,034	169,403	233,034	233,034	233,034	186,796

*Notes: See notes to Table 4. The baseline estimates correspond to those reported in Table 5. Column 2 adds in quadratic terms in the running variable, column 3 reduces the bandwidth to + or - .75, and column 4 includes first–second choice specific intercept terms. The next two columns use the 12 slope model (one slope for each of the 5 first choices and the 7 second choices) and the 60 slope model (separate slopes to the left and right of the cutoff for each first–second choice combination). Column 6 excludes the years 1982-84; these years added GPA bonuses for the first and second choices on an individual's ranking list. Standard errors in parentheses. * p<.10, ** p<.05, *** p<.01*

Table A9. Multiple inference correction.

	Second choice						
First choice	Engineering	Natural Science	Business	Social Science	Humanities	Non-acad. General	Non-acad. Vocational
			Panel A: q-values after FDR control for Table 4				
Engineering	--	<.001	.699	.073	.184	.625	.107
Natural Science	.177	--	.020	.011	.134	.625	.386
Business	.018	<.001	--	<.001	.482	.261	.177
Social Science	.001	.482	<.001	--	.172	<.001	<.001
Humanities	.747	.482	<.001	.017	--	.007	.007
			Panel B: q-values after FDR control for Table 5				
Engineering	--	<.001	.708	.046	.134	.616	.246
Natural Science	.195	--	.096	.019	.169	.616	.528
Business	.068	<.001	--	.004	.708	.355	.244
Social Science	.019	.514	<.001	--	.146	<.001	<.001
Humanities	.818	.612	<.001	.068	--	.001	.001

Notes: The table reports multiple inference corrected q-values (False Discovery Rate control) using the qqvalue package in Stata (method: simes).

Table A10. Comparison to KLM.

	Baseline	KLM IV
E vs. N	.065***	.027*
	(.017)	(.016)
E vs. B	.007	-.001
	(.018)	(.017)
E vs. S	.059**	.052**
	(.025)	(.024)
E vs. H	.070*	.068*
	(.039)	(.039)
E vs. G	.010	.015
	(.017)	(.016)
E vs. V	.020	.030**
	(.015)	(.014)
N vs. E	.039	-.042
	(.025)	(.027)
N vs. B	.056**	.054**
	(.028)	(.027)
N vs. S	.075***	.078***
	(.028)	(.027)
N vs. H	.060	.054
	(.037)	(.034)
N vs. G	.031	.066
	(.052)	(.046)
N vs. V	-.032	.004
	(.040)	(.035)
B vs. E	.046**	-.009
	(.021)	(.022)
B vs. N	.091***	.062***
	(.017)	(.016)
B vs. S	.053***	.056***
	(.016)	(.016)
B vs. H	-.008	.001
	(.018)	(.018)
B vs. G	-.011	-.008
	(.010)	(.009)
B vs. V	-.016	-.002
	(.011)	(.010)
S vs. E	-.072***	-.143***
	(.026)	(.028)
S vs. N	.016	-.033*
	(.018)	(.017)
S vs. B	-.066***	-.073***
	(.014)	(.014)
S vs. H	-.030*	-.030*
	(.017)	(.017)
S vs. G	-.073***	-.074***
	(.013)	(.012)
S vs. V	-.094***	-.073***
	(.016)	(.015)
H vs. E	.033	-.343***
	(.140)	(.131)
H vs. N	-.025	-.092***
	(.039)	(.035)
H vs. B	-.124***	-.128***
	(.021)	(.020)
H vs. S	-.046**	-.055***
	(.021)	(.020)
H vs. G	-.100***	-.090***
	(.028)	(.026)
H vs. V	-.111***	-.077***
	(.031)	(.029)
Corr. w/ baseline	1.00	.89
N	233,034	233,034

Notes: See notes to Table 4 and text in Section 4.6 for details. Standard errors in parentheses.
** p<.10, ** p<.05, *** p<.01*

Table A11. Comparison to OLS estimates.

	Fuzzy RD Baseline	OLS (w/o GPA)	OLS (w/ GPA)
E vs. N	.065***	.015***	.041***
	(.017)	(.004)	(.004)
E vs. B	.007	.076***	.047***
	(.018)	(.002)	(.002)
E vs. S	.059**	.135***	.130***
	(.025)	(.003)	(.003)
E vs. H	.070*	.195***	.172***
	(.039)	(.004)	(.004)
E vs. G	.010	.222***	.125***
	(.017)	(.004)	(.004)
E vs. V	.020	.230***	.137***
	(.015)	(.003)	(.004)
N vs. E	.039	-.015***	-.041***
	(.025)	(.004)	(.004)
N vs. B	.056**	.061***	.006
	(.028)	(.003)	(.004)
N vs. S	.075***	.120***	.090***
	(.028)	(.004)	(.004)
N vs. H	.060	.180***	.131***
	(.037)	(.004)	(.004)
N vs. G	.031	.207***	.084***
	(.052)	(.004)	(.005)
N vs. V	-.032	.215***	.096***
	(.040)	(.004)	(.005)
B vs. E	.046**	-.076***	-.047***
	(.021)	(.002)	(.002)
B vs. N	.091***	-.061***	-.006
	(.017)	(.003)	(.004)
B vs. S	.053***	.059***	.084***
	(.016)	(.002)	(.002)
B vs. H	-.008	.119***	.125***
	(.018)	(.003)	(.003)
B vs. G	-.011	.146***	.079***
	(.010)	(.003)	(.004)
B vs. V	-.016	.154***	.091***
	(.011)	(.003)	(.003)
S vs. E	-.072***	-.135***	-.130***
	(.026)	(.003)	(.003)
S vs. N	.016	-.120***	-.090***
	(.018)	(.004)	(.004)
S vs. B	-.066***	-.059***	-.084***
	(.014)	(.002)	(.002)
S vs. H	-.030*	.060***	.041***
	(.017)	(.003)	(.003)
S vs. G	-.073***	.087***	-.005
	(.013)	(.004)	(.004)
S vs. V	-.094***	.095***	.007*
	(.016)	(.003)	(.004)
H vs. E	.033	-.195***	-.172***
	(.140)	(.004)	(.004)
H vs. N	-.025	-.180***	-.131***
	(.039)	(.004)	(.004)
H vs. B	-.124***	-.119***	-.125***
	(.021)	(.003)	(.003)
H vs. S	-.046**	-.060***	-.041***
	(.021)	(.003)	(.003)
H vs. G	-.100***	.027***	-.046***
	(.028)	(.004)	(.004)
H vs. V	-.111***	.035***	-.034***
	(.031)	(.004)	(.004)
Corr. w/ baseline	1.00	.13	.44
N	233,034	233,034	233,034

Notes: See notes to Table 4 and text in Section 4.6 for details. Standard errors in parentheses.
** p<.10, ** p<.05, *** p<.01*

Table A12. Mediation analysis.

	Baseline	Controls for years of schooling	Share explained	Controls for college major	Share explained	Controls for occupation	Share explained	All controls	Share explained
Panel A: Significant baseline estimates									
E vs. N	.064***	.048***	0.25	.052***	0.19	.023	0.64	.013	0.80
	(.017)	(.017)		(.017)		(.014)		(.015)	
E vs. S	.059**	.040	0.32	.065***	-0.10	.034	0.42	.030	0.49
	(.025)	(.025)		(.024)		(.021)		(.021)	
E vs. H	.070*	.055	0.21	.078**	-0.11	.016	0.77	.019	0.73
	(.039)	(.038)		(.037)		(.032)		(.031)	
N vs. B	.056**	.033	0.41	.047*	0.16	.039*	0.30	.031	0.45
	(.028)	(.028)		(.027)		(.023)		(.023)	
N vs. S	.075***	.047*	0.37	.072***	0.04	.057**	0.24	.054**	0.28
	(.028)	(.028)		(.027)		(.023)		(.023)	
B vs. E	.046**	.044**	0.04	.014	0.70	.019	0.59	.007	0.85
	(.021)	(.021)		(.021)		(.018)		(.017)	
B vs. N	.091***	.090***	0.01	.054***	0.41	.034**	0.63	.022	0.76
	(.017)	(.017)		(.017)		(.014)		(.014)	
B vs. S	.053***	.052***	0.02	.038**	0.28	.020	0.62	.015	0.72
	(.016)	(.016)		(.016)		(.014)		(.013)	
S vs. E	-.072***	-.081***	-0.13	-.065**	0.10	-.027	0.63	-.031	0.57
	(.026)	(.026)		(.025)		(.022)		(.022)	
S vs. B	-.066***	-.070***	-0.06	-.040***	0.39	-.026**	0.61	-.024**	0.64
	(.014)	(.014)		(.014)		(.012)		(.012)	
S vs. H	-.030*	-.035**	-0.17	-.002	0.93	-.026*	0.13	-.017	0.43
	(.017)	(.017)		(.017)		(.014)		(.014)	
S vs. G	-.073***	-.088***	-0.21	-.043***	0.41	-.022**	0.70	-.022**	0.70
	(.013)	(.013)		(.013)		(.011)		(.011)	
S vs. V	-.094***	-.112***	-0.19	-.038**	0.60	-.036***	0.62	-.031**	0.67
	(.016)	(.016)		(.016)		(.013)		(.013)	
H vs. B	-.124***	-.122***	0.02	-.072***	0.42	-.063***	0.49	-.055***	0.56
	(.021)	(.021)		(.020)		(.017)		(.017)	
H vs. S	-.046**	-.053**	-0.15	-.007	0.85	-.020	0.57	-.012	0.74
	(.021)	(.021)		(.020)		(.017)		(.017)	
H vs. G	-.100***	-.109***	-0.09	-.060**	0.40	-.052**	0.48	-.048**	0.52
	(.028)	(.028)		(.027)		(.023)		(.023)	
H vs. V	-.111***	-.124***	-0.12	-.051*	0.54	-.065**	0.41	-.057**	0.49
	(.031)	(.031)		(.030)		(.026)		(.026)	
Share explained >.50			0		5		10		12

Table A12. Mediation analysis, continued.

	Baseline	Controls for years of schooling	Share explained	Controls for college major	Share explained	Controls for occupation	Share explained	All controls	Share explained
				Panel B: Insignificant baseline estimates					
E vs. B	.007	-.005	-	.013	-	-.002	-	-.010	-
	(.018)	(.018)		(.018)		(.015)		(.015)	
E vs. G	.010	-.005	-	.039**	-	-.004	-	-.003	-
	(.017)	(.017)		(.017)		(.015)		(.015)	
E vs. V	.020	-.003	-	.038***	-	-.003	-	-.006	-
	(.015)	(.015)		(.015)		(.013)		(.013)	
N vs. E	.039	.009	-	.014	-	.030	-	.019	-
	(.025)	(.025)		(.025)		(.021)		(.021)	
N vs. H	.060	.036	-	.063*	-	.024	-	.027	-
	(.037)	(.036)		(.034)		(.030)		(.029)	
N vs. G	.031	.012	-	.059	-	.026	-	.030	-
	(.052)	(.051)		(.050)		(.043)		(.043)	
N vs. V	-.032	-.068*	-	.005	-	-.011	-	-.010	-
	(.040)	(.039)		(.038)		(.033)		(.032)	
B vs. H	-.008	.003	-	-.001	-	-.026*	-	-.019	-
	(.018)	(.017)		(.017)		(.014)		(.014)	
B vs. G	-.011	-.009	-	-.005	-	-.020**	-	-.018**	-
	(.010)	(.010)		(.010)		(.008)		(.008)	
B vs. V	-.016	-.017	-	-.009	-	-.025***	-	-.024***	-
	(.011)	(.011)		(.011)		(.009)		(.009)	
S vs. N	.016	.005	-	.018	-	.009	-	.007	-
	(.018)	(.018)		(.018)		(.015)		(.015)	
H vs. E	.032	.020	-	-.064	-	-.026	-	-.043	-
	(.141)	(.139)		(.135)		(.113)		(.112)	
H vs. N	-.025	-.034	-	.003	-	-.011	-	-.006	-
	(.039)	(.038)		(.036)		(.032)		(.031)	
N	233,034	233,034		232,462		233,034		232,462	

Notes: "Share explained" is defined as 1 – [baseline estimate / (baseline estimate - mediation estimate)]. Panel A reports the estimates and the share explained for baseline estimates which are statistically different from zero. Share explained is not reported in panel B for the insignificant estimates, as it provides little insight for small and noisy estimates. Standard errors in parentheses.
** p<.10, ** p<.05, *** p<.01*