Appendix
for Online Publication

"Field of Study and Long-Term Mental Health"

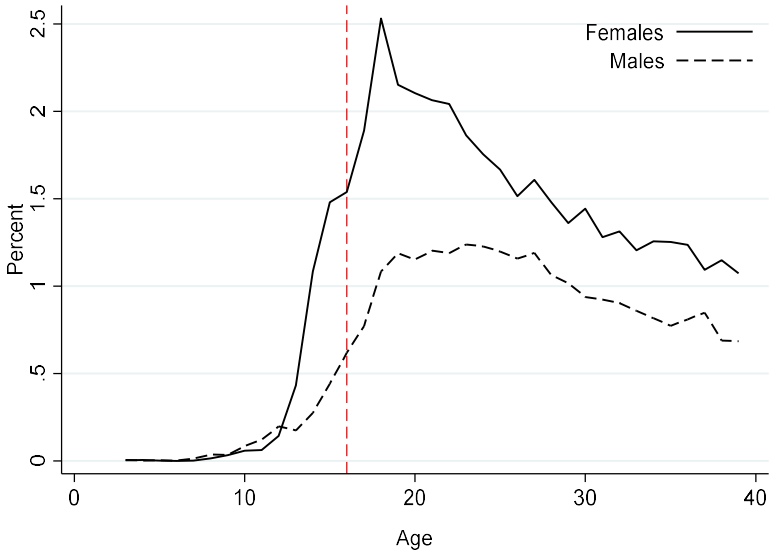
Anders Stenberg and Simona Tudor

Definition of the incidence of hospitalization due to mental disorders.

We restrict our definition of hospitalizations for mental disorders to International Classification of Diseases 10th Revision (ICD10) codes linked to environmental and lifestyle/socioeconomic factors rather than to biological factors. Specifically, in the definition of our hospitalizations for mental disorders we include individuals with one of the following ICD10 diagnostic codes (individuals often receive more than one diagnosis): F10–F19 mental and behavioral disorders due to psychoactive substance use; F30–F39 mood [affective] disorders; F40–F48 neurotic, stress-related and somatoform disorders; F50–F59 behavioral syndromes associated with physiological disturbances and physical factors. We exclude F00–F09 organic, including symptomatic, mental disorders; F20–F29 schizophrenia, schizotypal and delusional disorders; F60–F69 disorders of adult personality and behavior; F70–F79 mental retardation; F80–F89 disorders of psychological development; F90–F98 behavioral and emotional disorders with onset usually occurring in childhood and adolescence; F99–F99 unspecified mental disorders.

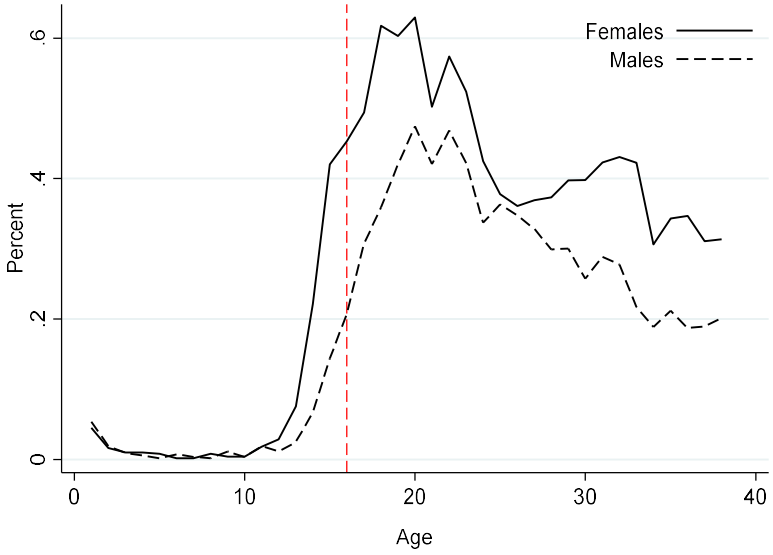
Figure A1. Total population first-time incidence of antidepressants and hospitalizations due to mental disorders.

Panel A: Incidence of antidepressants



Notes: Total population first-time incidence of prescribed antidepressant medications across age based on observations 2005 to 2019. Vertical line at age 16, the timing of high-school start. Incidence until age 14 is based on cohort born 2005, incidence age 15 on cohort born 2004 observed in 2019 conditioned on no incidence age 1-14, age 16 on cohort born 2003 observed in 2019 conditioned on no incidence age 2-15, and so forth, until age 39 based cohort born 1980 observed in 2019 conditioned on no incidence age 25-38.

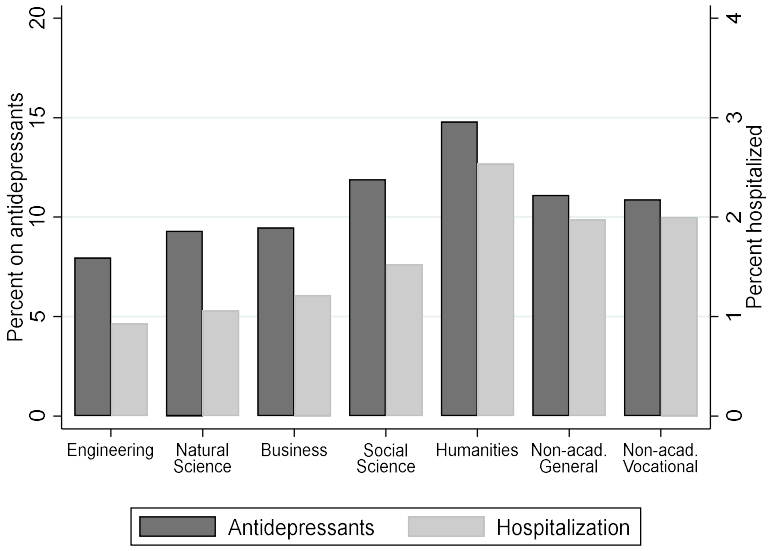
Panel B: Incidence of hospitalization due to mental disorders



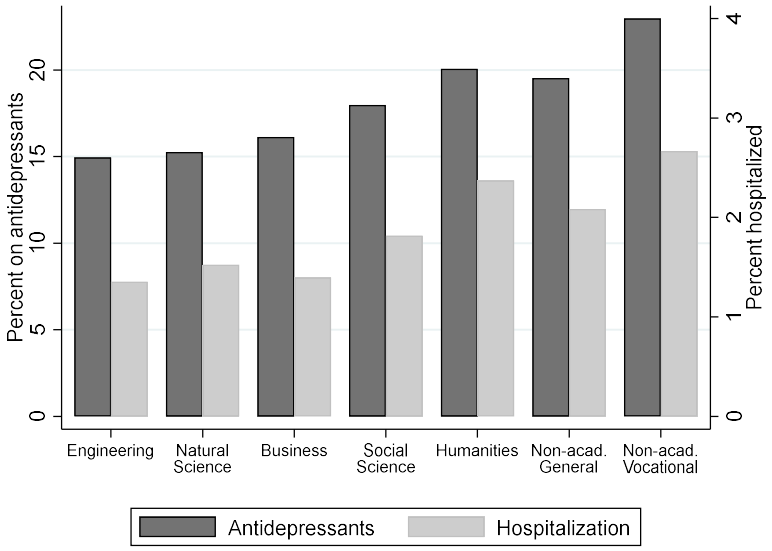
Notes: Total population first-time incidence of hospitalization due to mental disorders across age based on observations 1997 to 2018. Vertical line at age 16, the timing of high-school start. Incidence until age 21 is based on cohort born 1997, age 22 on cohort born 1996 and observed in 2018 conditioned on no incidence age 1-21, age 23 on cohort born 1995 and observed in 2018 conditioned on no incidence age 2-22, and so forth, until age 39 based cohort born 1979 and observed in 2018 conditioned on no incidence age 18-38.

Figure A2. Program completers and adult outcomes.

Panel A: Men - antidepressants and hospitalization (right y-axis)

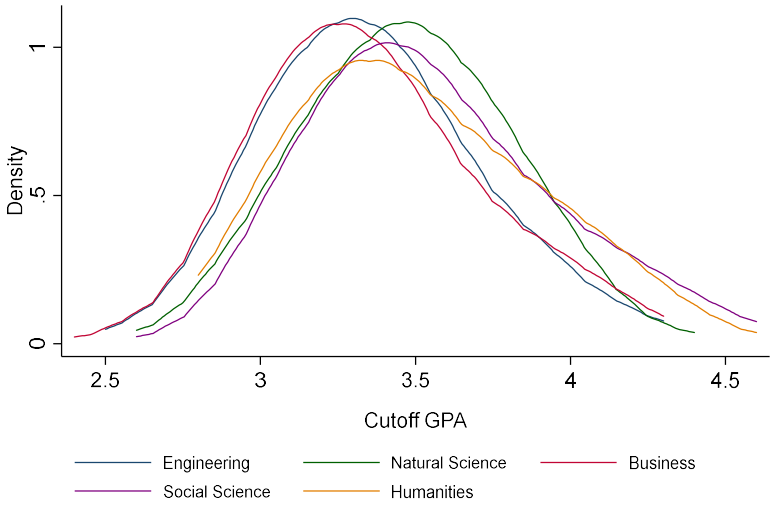


Panel B: Women - antidepressants and hospitalization (right y-axis)



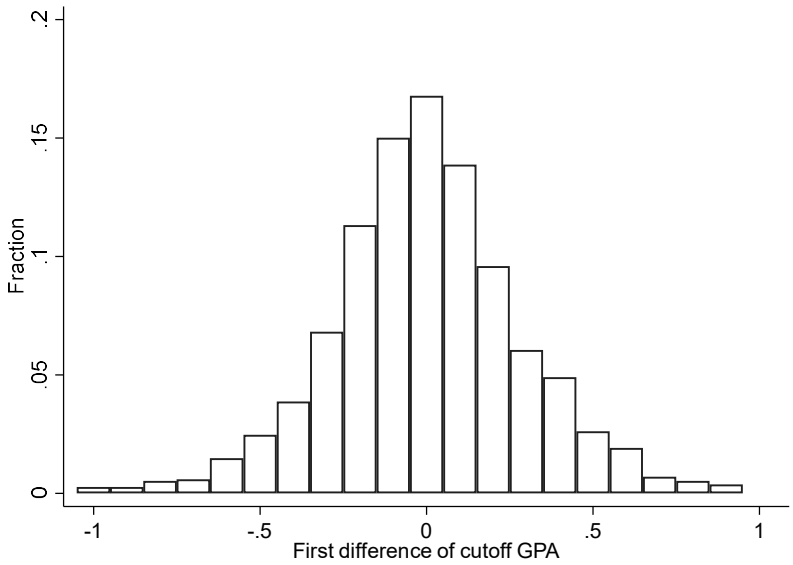
Notes: Sample of program completers who applied between 1977-1991. Incidence of prescribed antidepressants medications at age 40-45 and of hospitalizations due to mental disorders. N_{men}*= 553,824;* N_{women} *= 581,425,*

Figure A3. Distribution of GPA cutoffs by high school major



Notes: Kernel density estimates of GPA cutoffs by major for oversubscribed programs, applying an Epanechnikov kernel with bandwidth 0.2.

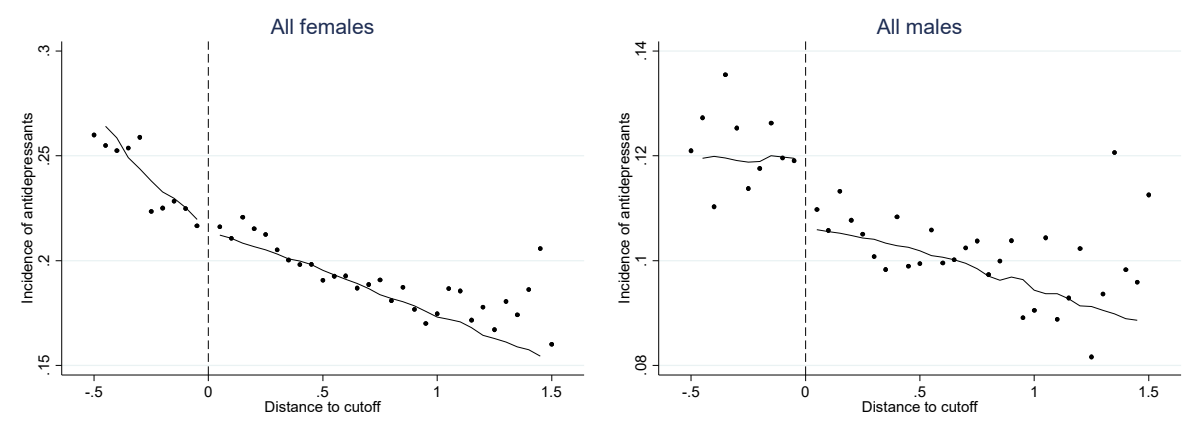
Figure A4. First-differenced cutoff GPA distribution



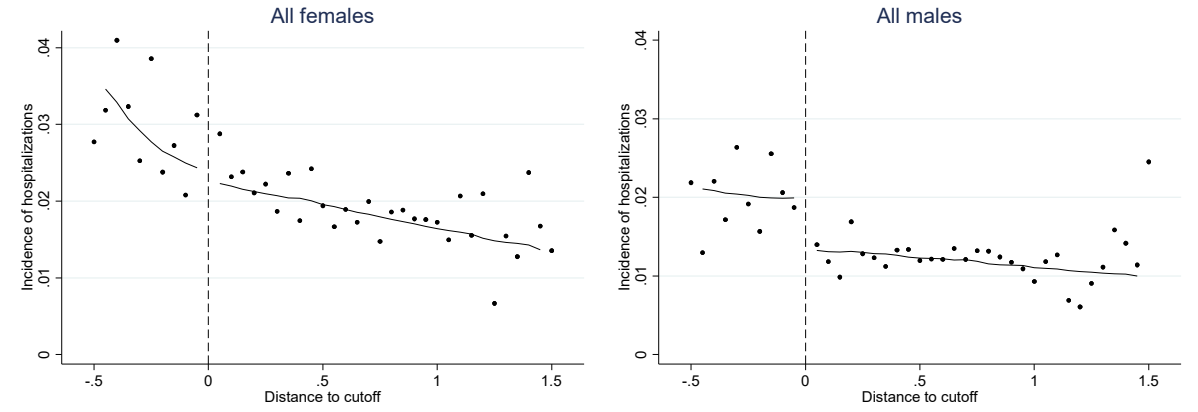
Notes: Current minus lagged cutoff GPA, where the sample is limited to majors which are competitive two years in a row in a school region.

Figure A5. Gender specific incidence of indicators of mental health in adult age by distance to cutoff.

Panel A: Incidence of antidepressants at age 40-45



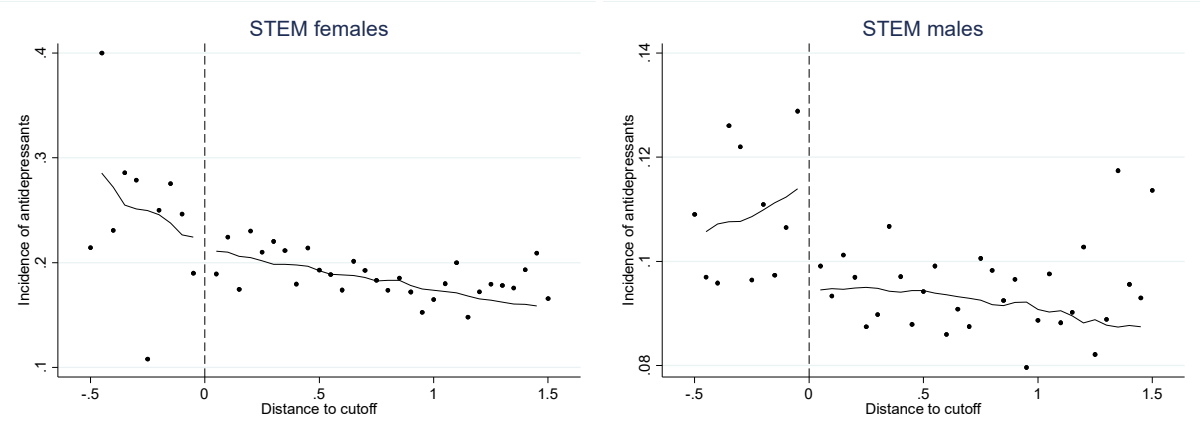
Panel B: Incidence of hospitalization at age 36-44



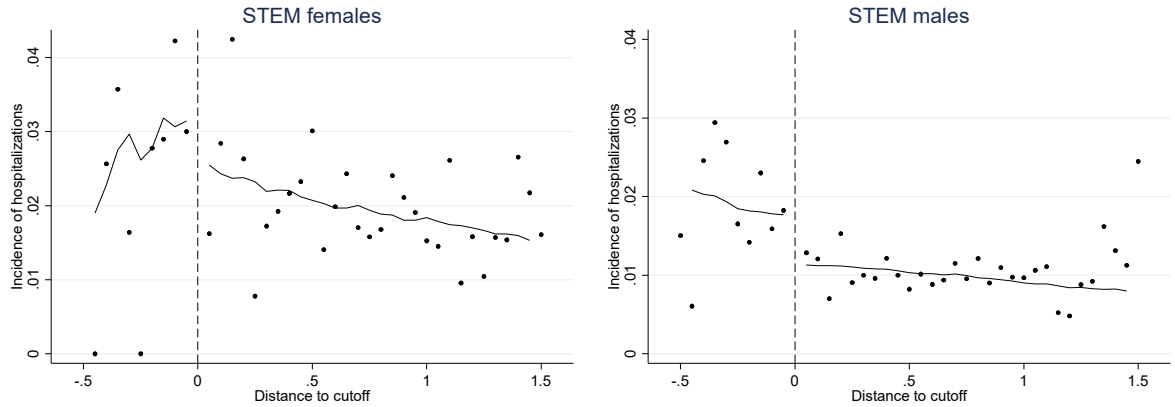
Notes: Each observation is the average incidence of antidepressants at age 40-45 (panel A) or hospitalization due to mental illness (panel B) in GPA bins measured relative to a cutoff normalized to zero. The regressions are estimated within a window of -0.5 to 1.5, applying triangular weights and using linear functions of GPA which are allowed to vary to the left and to the right of the cutoff for each field of study category (antidepressants, columns 1 and 2), with squared terms to the left of the cutoff when estimating the impact of hospitalization (columns 3 and 4). Covariates included are fixed effects for year and school region, dummies for preferred program category as well as for next-best alternative. The number of observations is for females 128,388 and for males 118,686.

Figure A6. Sample with STEM as first choice. Gender specific incidence of indicators of mental health in adult age by distance to cutoff.

Panel A: Incidence of antidepressants at age 40-45

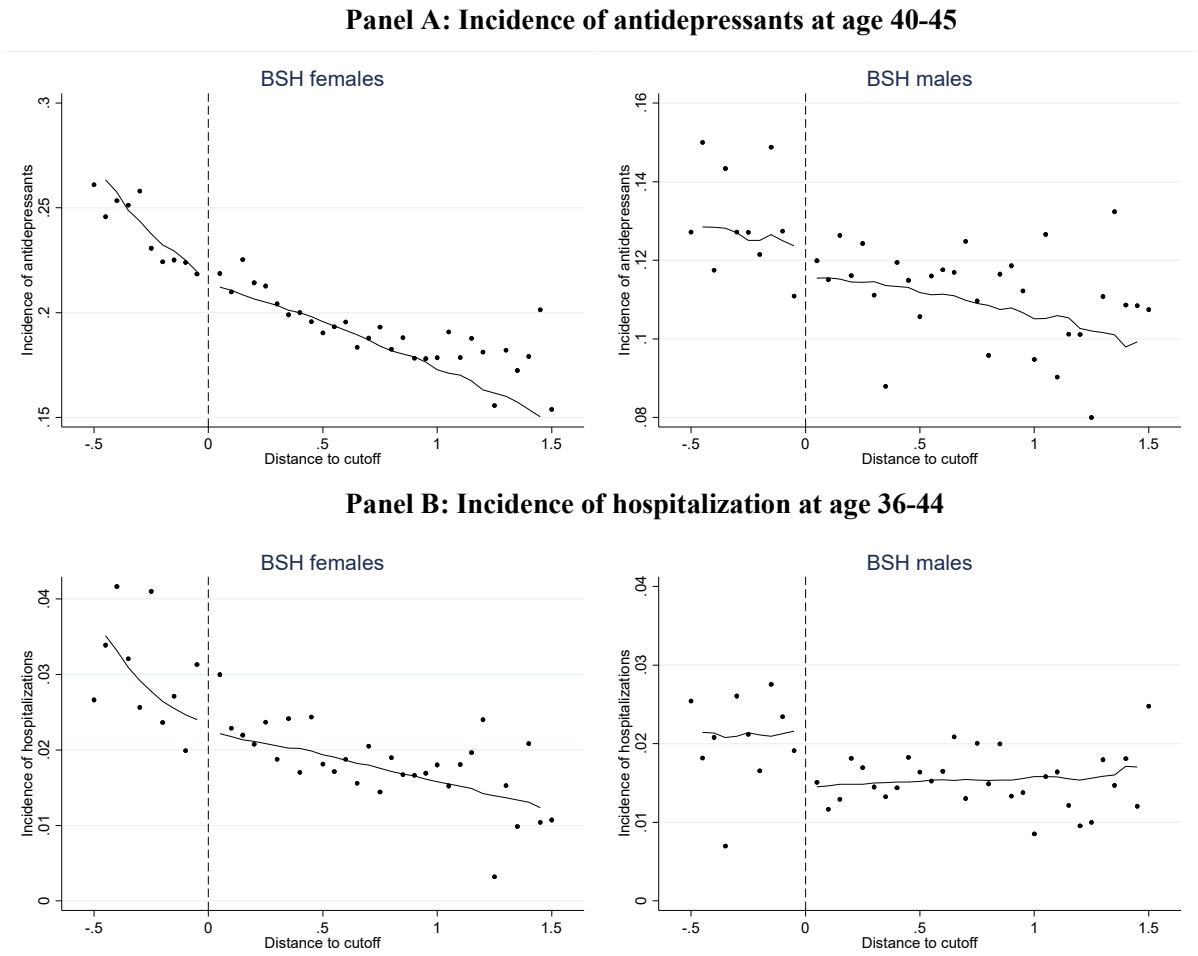


Panel B: Incidence of hospitalization at age 36-44



Notes: Each observation is the average incidence of antidepressants at age 40-45 (panel A) or hospitalization due to mental illness (panel B) in GPA bins measured relative to a cutoff normalized to zero. The regressions are estimated within a window of -0.5 to 1.5, applying triangular weights and using linear functions of GPA which are allowed to vary to the left and to the right of the cutoff for each field of study category (antidepressants, columns 1 and 2), with squared terms to the left of the cutoff when estimating the impact of hospitalization (columns 3 and 4). Covariates included are fixed effects for year and school region, dummies for preferred program category as well as for next-best alternative. The number of observations is for females in STEM 21,286, for males in STEM 65,071.

Figure A7. Sample with BSH as first choice. Gender specific incidence of indicators of mental health in adult age by distance to cutoff.



Notes: Each observation is the average incidence of antidepressants at age 40-45 (panel A) or hospitalization due to mental illness (panel B) in GPA bins measured relative to a cutoff normalized to zero. The regressions are estimated within a window of -0.5 to 1.5, applying triangular weights and using linear functions of GPA which are allowed to vary to the left and to the right of the cutoff for each field of study category (antidepressants, columns 1 and 2), with squared terms to the left of the cutoff when estimating the impact of hospitalization (columns 3 and 4). Covariates included are fixed effects for year and school region, dummies for preferred program category as well as for next-best alternative. The number of observations is for females in BSH 107,102 and for males in BSH 53,615.

Table A1. Program specific summary statistics of main analysis sample.

	First choice:				
	Engineering	Natural science	Business	Social science	Humanities
Women	0.17	0.49	0.60	0.72	0.86
Age when applying	16.00	15.98	16.00	15.99	15.99
GPA	3.75	4.04	3.65	3.80	3.76
GPA adjusted	3.88	4.11	3.75	3.94	3.90
Foreign born	0.03	0.04	0.03	0.03	0.05
Foreing born parent	0.16	0.17	0.16	0.16	0.21
Father earnings	5.76	5.86	5.74	5.77	5.72
Mother earnings	5.22	5.31	5.21	5.26	5.23
Father schooling	11.57	12.78	11.20	11.90	11.48
Mother schooling	11.17	12.34	10.83	11.63	11.12
Father age	29.71	29.84	29.69	29.89	29.99
Mother age	27.15	27.46	27.11	27.34	27.30
2nd choice					
Engineering	--	0.47	0.09	0.03	0.03
Natural science	0.50	--	0.08	0.16	0.04
Business	0.19	0.12	--	0.30	0.17
Social science	0.05	0.30	0.35	--	0.52
Humanities	0.01	0.04	0.10	0.30	--
Non-ac. General	0.07	0.02	0.22	0.15	0.16
Non-ac. Vocational	0.17	0.05	0.16	0.06	0.09
Observations	66,306	20,051	89,194	55,959	15,564

Notes: Oversubscribed programs are defined by major, year, and school region. Parent characteristics are measured in the year of application (the child's 16th year since birth).

Table A2. Comparison of major cutoffs across years within the same school region.

Major combinations	Fraction of years with a higher cutoff		
	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 A3. Balancing tests for pre-determined characteristics.

	Years of schooling father	Years of schooling mother	Log earnings father	Log earnings mother	Age at birth father	Age at birth mother	Foreign born parent	Child foreign born
	-.0511	-.0278	-.0045	-.0047	-.0908	-.0354	.0010	.0004
	(.0313)	(.0286)	(.0052)	(.0045)	(.0725)	(.0621)	(.0046)	(.0024)
N	234,050	243,120	201,514	184,914	233,845	242,653	247,068	247,074

Notes: Each column is an estimate from a separate RD regression, where the outcome is a linear function of the running variable (normalized GPA) within a window of -0.5 to 1.5, using triangular weights; fixed effects for year, school region, and program, and a common slope on each side of the cutoff. Standard errors in parentheses.
** p<.10, ** p<.05, *** p<.01*

Table A4. Estimates separate for each combination of first and second choices.

	Antidepressants	Hospitalizations	% with 2nd choice	Earnings 37-39
Eng vs. Nat sci	-.0131	-.0068**	50	.065***
	(.0090)	(.0030)		(.017)
Eng vs. Bus	-.0345***	-.0099***	19	.007
	(.0093)	(.0032)		(.018)
Eng vs. Soc sci	-.0108	-.0057	5	.059**
	(.0109)	(.0038)		(.025)
Eng vs. Hum	-.0077	.0003	1	.070*
	(.0193)	(.0086)		(.039)
Eng vs. Gen non-ac	-.0122	-.0035	7	.010
	(.0094)	(.0032)		(.017)
Eng vs. Voc non-ac	-.0189**	-.0051*	17	.020
	(.0086)	(.0027)		(.015)
Nat sci vs. Eng	-.0033	-.0066**	47	.039
	(.0098)	(.0033)		(.025)
Nat sci vs. Bus	-.0240**	-.0008	12	.056**
	(.0120)	(.0048)		(.028)
Nat sci vs. Soc sci	-.0086	-.0030	30	.075***
	(.0106)	(.0038)		(.028)
Nat sci vs. Hum	.0128	.0010	4	.060
	(.0190)	(.0074)		(.037)
Nat sci vs. Gen non-ac	.0025	-.0094**	2	.031
	(.0217)	(.0043)		(.052)
Nat sci vs. Voc non-ac	.0199	.0104	5	-.032
	(.0175)	(.0072)		(.040)
Bus vs. Eng	-.0212**	-.0093***	9	.046**
	(.0085)	(.0031)		(.021)
Bus vs. Nat sci	-.0155*	-.0048	8	.091***
	(.0088)	(.0033)		(.017)
Bus vs. Soc sci	-.0181***	-.0080***	35	.053***
	(.0068)	(.0024)		(.016)
Bus vs. Hum	-.0042	-.0077***	10	-.008
	(.0079)	(.0027)		(.018)
Bus vs. Gen non-ac	-.0031	-.0023	22	-.011
	(.0062)	(.0020)		(.010)
Bus vs. Voc non-ac	-.0069	-.0049**	16	-.016
	(.0067)	(.0021)		(.011)
Soc sci vs. Eng	.0030	-.0039	3	-.072***
	(.0125)	(.0047)		(.026)
Soc sci vs. Nat sci	-.0028	-.0050	16	.016
	(.0088)	(.0033)		(.018)
Soc sci vs. Bus	-.0019	-.0033	30	-.066***
	(.0072)	(.0026)		(.014)
Soc sci vs. Hum	.0118	-.0025	30	-.030*
	(.0073)	(.0026)		(.017)
Soc sci vs. Gen non-ac	.0123	-.0020	15	-.073***
	(.0076)	(.0025)		(.013)
Soc sci vs. Voc non-ac	.0068	.0015	6	-.094***
	(.0099)	(.0037)		(.016)
Hum vs. Eng	.0135	-.0087	3	.033
	(.0228)	(.0070)		(.140)
Hum vs. Nat sci	.0266	-.0014	4	-.025
	(.0207)	(.0072)		(.039)
Hum vs. Bus	.0290**	.0040	17	-.124***
	(.0114)	(.0045)		(.021)
Hum vs. Soc sci	.0344***	.0029	52	-.046**
	(.0083)	(.0031)		(.021)
Hum vs. Gen non-ac	.0385***	.0070	16	-.100***
	(.0126)	(.0048)		(.028)
Hum vs. Voc non-ac	.0761***	.0051	9	-.111***
	(.0163)	(.0058)		(.031)
N	247,074	247,074		233,034

Notes: See notes to Table 4 and text for details. Earnings estimates from Dahl et al. (2023). Standard errors in parentheses.

Table A5. Incidence of antidepressamts at age 45-50 and at age 44-45. Reduced form and IV estimates.

Panel A: All	Reduced form	IV-enrolled	IV-completed	Mean
Antidepressants 45-50	-.0058	-.0105	-.0142	[.2023]
	(.0048)	(.0087)	(.0117)	
N	247,074	247,074	247,074	
Antidepressants 44-45	-.0022	-.0039	-.0053	[.1164]
	(.0036)	(.0065)	(.0088)	
	247,074	247,074	247,074	
Panel B: STEM	Reduced form	IV-enrolled	IV-completed	Mean
Antidepressants 45-50	-.0171**	-.0368**	-.0515**	[.1485]
	(.0087)	(.0187)	(.0261)	
N	86,357	86,357	86,357	
Antidepressants 44-45	-.0093	-.0200	-.0279	[.0799]
	(.0061)	(.0132)	(.0184)	
N	86,357	86,357	86,357	
Panel C: BSH				
Antidepressants 45-50	-.0023	-.0039	-.0052	[.2210]
	(.0058)	(.0098)	(.0132)	
N	160,717	160,717	160,717	
Antidepressants 44-45	.0001	.0002	.0003	[.1291]
	(.0044)	(.0074)	(.0100)	
N	160,717	160,717	160,717	

Notes: See note to table 4. The outcome variable is the incidence of prescribed antidepressants at age 45-50 or at age 44-45. For the age windows 40–45 and 45–50, data need to be slightly adjusted for the youngest or the oldest cohorts. Using antidepressants at age 40–45 as outcome, we apply for the oldest cohorts age 41–46 if born in 1964, age 42–47 for the cohort born in 1963, age 43–48 for the cohort born in 1962, and age 44–49 for the cohort born in 1961. Conversely, when we use age 45–50 we instead need to adjust the outcome of the youngest cohorts, age 44–49 for the cohort born in 1971, 43–48 for the cohort born in 1972, 42–47 for the cohort born in 1973, 41–46 for the cohort born in 1974 and 40–45 for the cohort born in 1975. The mean of the dependent variable is calculated for the sample where an individual's GPA is within plus or minus 0.2 GPA points of the admission cutoff. Standard errors in parentheses.

** p<0.10, ** p<0.05, *** p<0.01*

Table A6. Incidence of hospitalization at age 36-40 and at age 40-44. Reduced form and IV estimates.

Panel A: All	Reduced form	IV-enrolled	IV-completed	Mean
Hospitalizations 36-40	-.0022**	-.0034**	-.0047**	[.0114]
	(.0010)	(.0016)	(.0023)	
N	247,074	247,074	247,074	
Hospitalizations 40-44	-.0035***	-.0055***	-.0076***	[.0129]
	(.0011)	(.0017)	(.0024)	
	247,074	247,074	247,074	
Panel B: STEM	Reduced form	IV-enrolled	IV-completed	Mean
Hospitalizations 36-40	-.0033*	-.0058*	-.0086*	[.0088]
	(.0019)	(.0034)	(.0050)	
N	86,357	86,357	86,357	
Hospitalizations 40-44	-.0051**	-.0092**	-.0136**	[.0110]
	(.0020)	(.0036)	(.0054)	
N	86,357	86,357	86,357	
Panel C: BSH				
Hospitalizations 36-40	-.0017	-.0026	-.0035	[.0123]
	(.0013)	(.0019)	(.0026)	
N	160,717	160,717	160,717	
Hospitalizations 40-44	-.0029**	-.0044**	-.0059**	[.0136]
	(.0013)	(.0020)	(.0027)	
N	160,717	160,717	160,717	

Notes: See notes to table 4. The outcome variable is the incidence of hospitalizations related to mental disorders at age 36-40 (panel A) and age 40-44 (panel B). IV-enrolled uses as a first stage whether the individual enrolled in their first-best major, as a function of whether their GPA exceeded the admissions cutoff. IV-completed uses as first stage whether the individual completed their first-best major. The mean of the dependent variable is calculated for the sample where an individual's GPA is within plus or minus 0.2 GPA points of the admission cutoff. Standard errors in parentheses.
** p<0.10, ** p<0.05, *** p<0.01*