Online Appendix for “Parental Education and the Responses to Higher SAT Scores”

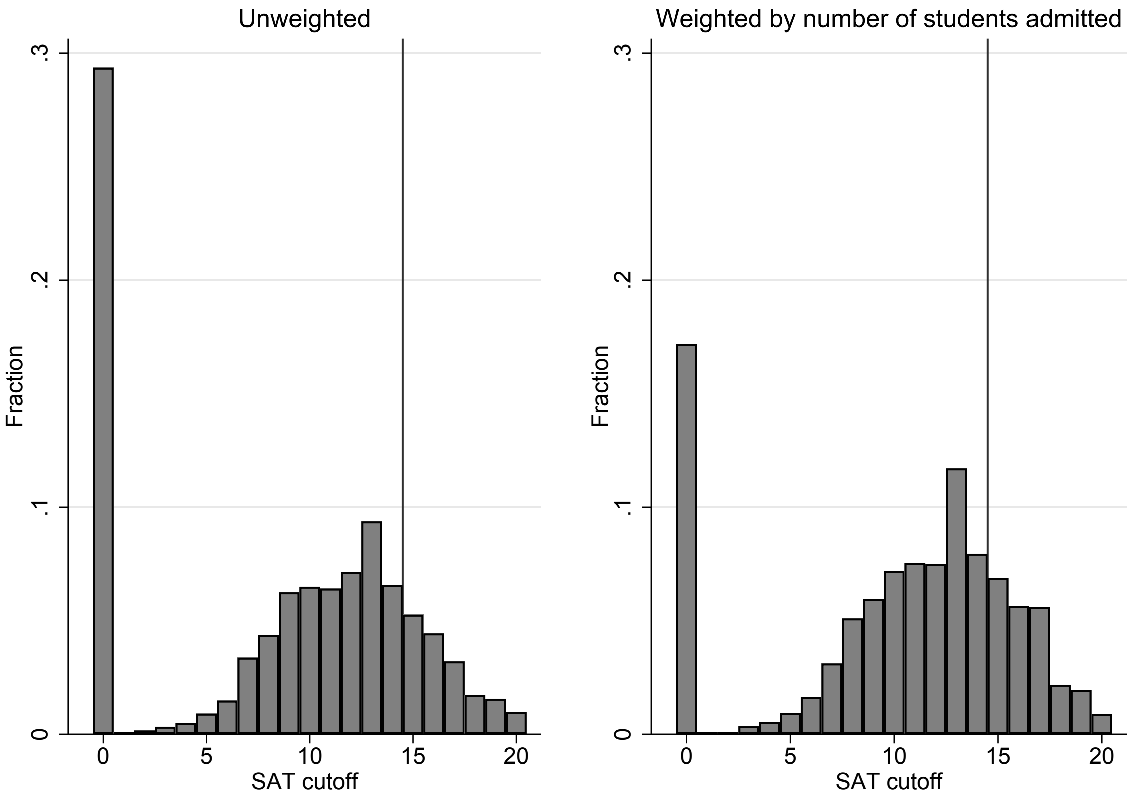
Georg Graetz (*Uppsala University*)

Björn Öckert (*IFAU*)

Oskar Nordström Skans (*Uppsala University*)

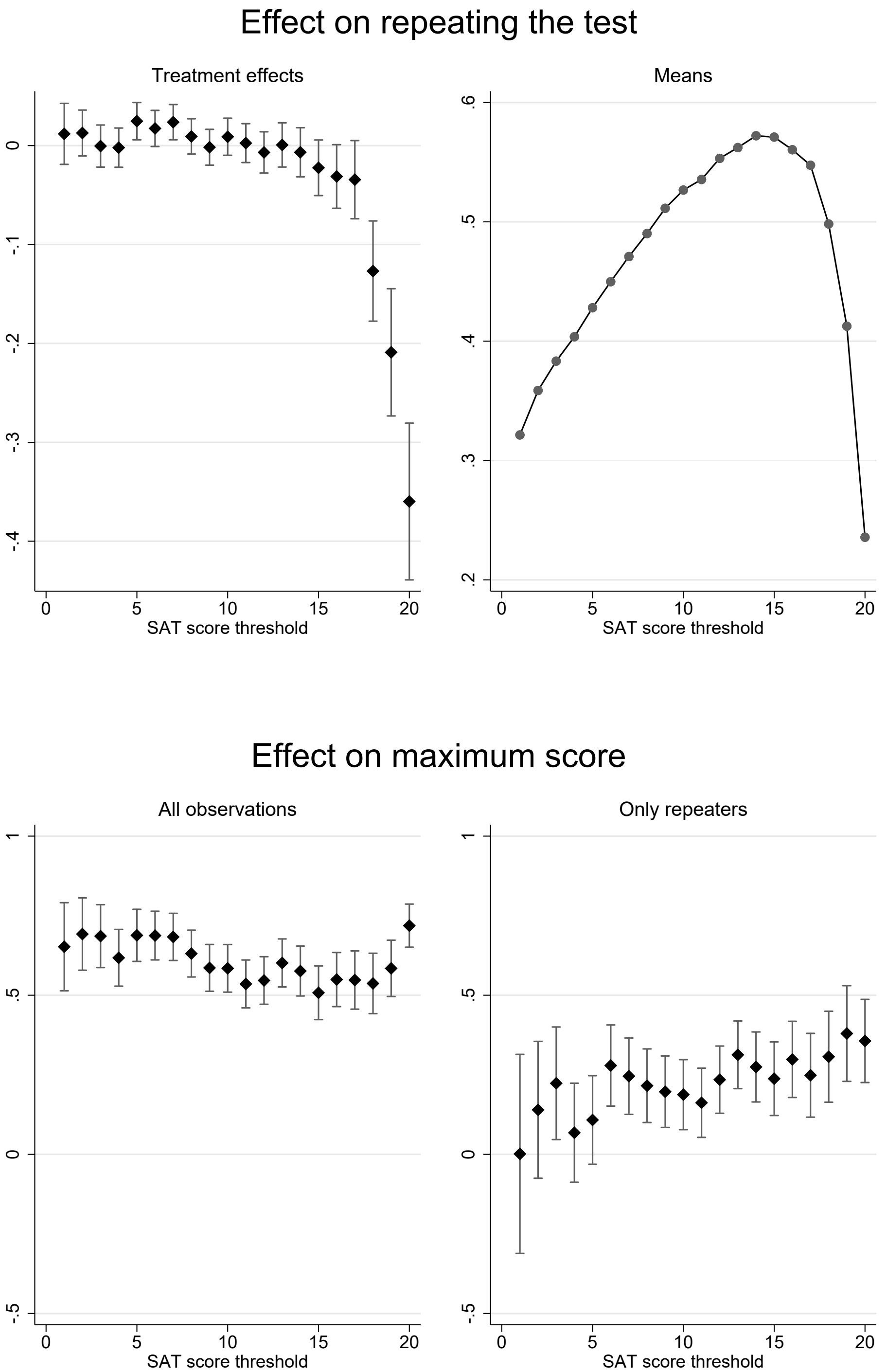
December 27, 2022

A Appendix figures and tables



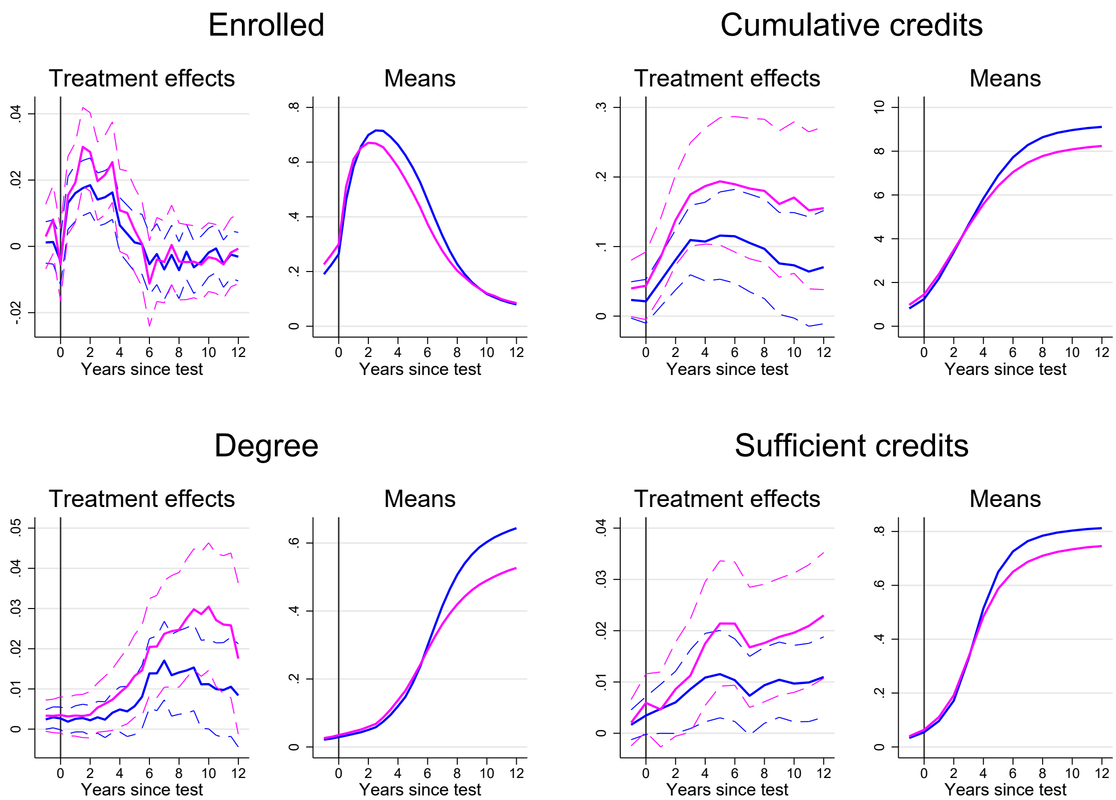
Notes: A cutoff score of zero indicates that the program was undersubscribed. Most of our results focus on the thresholds to the right of the vertical lines.

Figure A1: Distribution of SAT cutoffs across programs from 1999 fall admissions



Notes: The estimated effects of crossing the indicated threshold on repeating the test and on the maximum life-time score are displayed together with 95-percent confidence intervals. Effects are estimated by RD, implemented via local linear regressions with a bandwidth of three raw marks on each side. Only first tests are included. The top right panel shows the means of repeating by normalized score.

Figure A2: The effects of a higher score on the probability of repeating and on the maximum life-time score



Notes: In the panels titled 'Treatment effects', point estimates (solid) and cluster-robust 95-percent confidence intervals (dashed) from estimating equation (2) are plotted against years since taking the test. In the panels titled 'Means', means of the dependent variable within the estimation sample are plotted against years since taking the test. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Results from the full and locally low GPA samples are shown in blue and magenta, respectively.

Figure A3: The effects of a higher score on enrollment, credits, and graduation

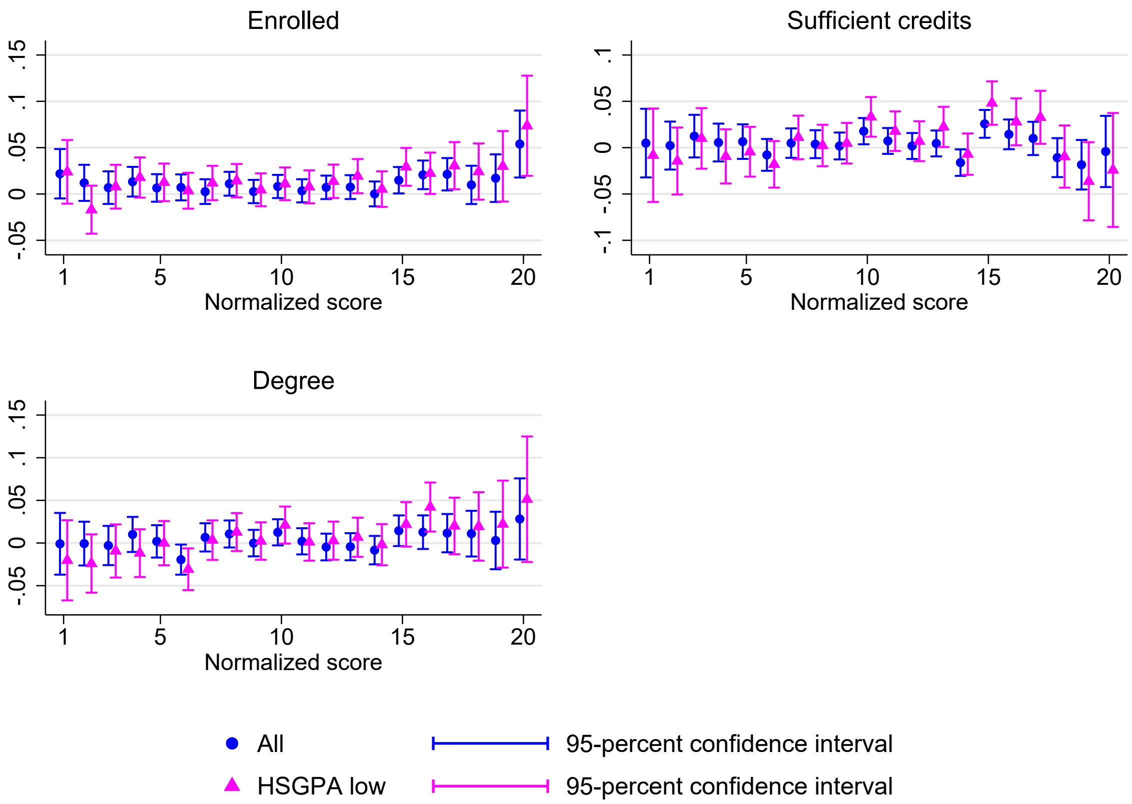
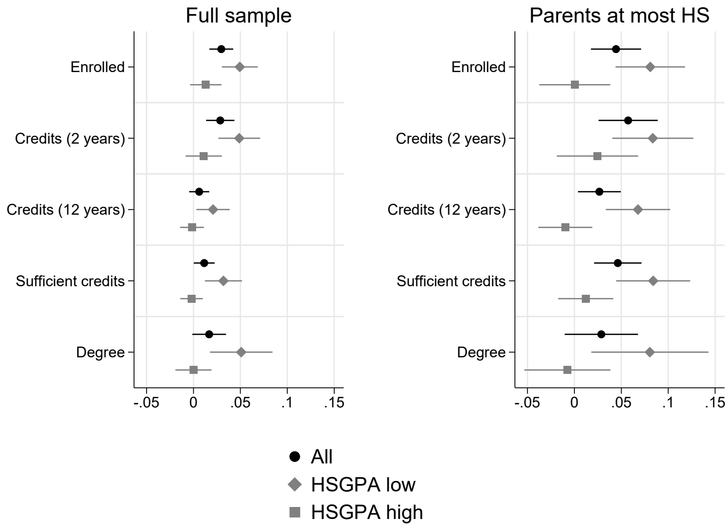
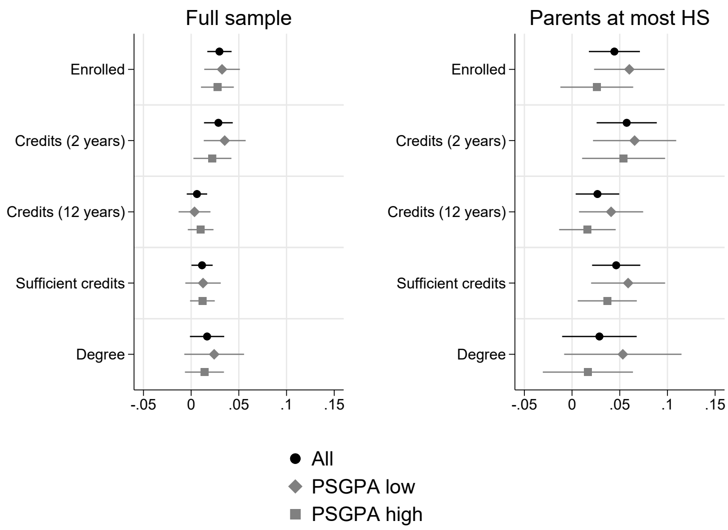


Figure A4: The effects of a higher score on enrollment, credits, and graduation, by threshold



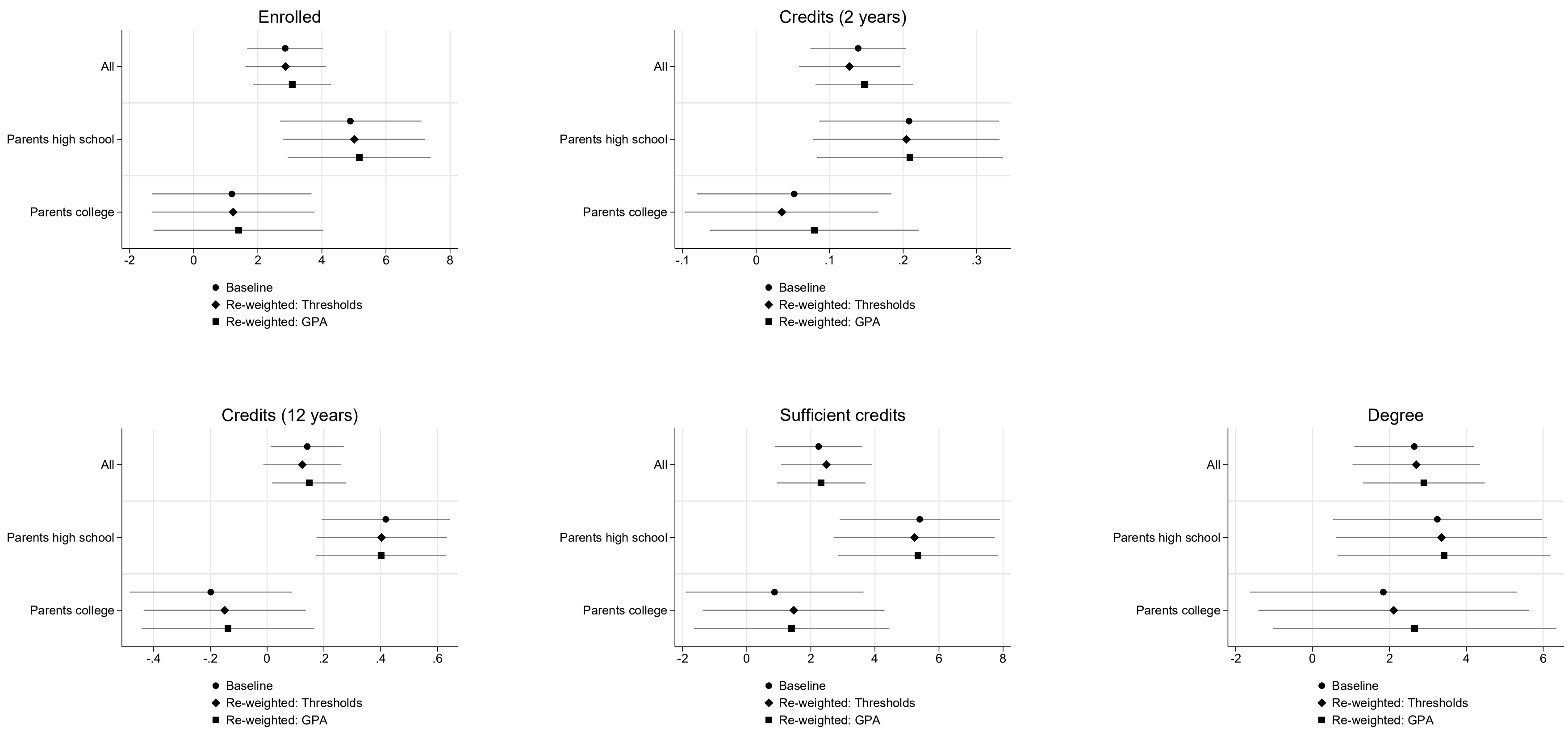
(a) High school GPA relative to local median



(b) Compulsory school GPA relative to local median

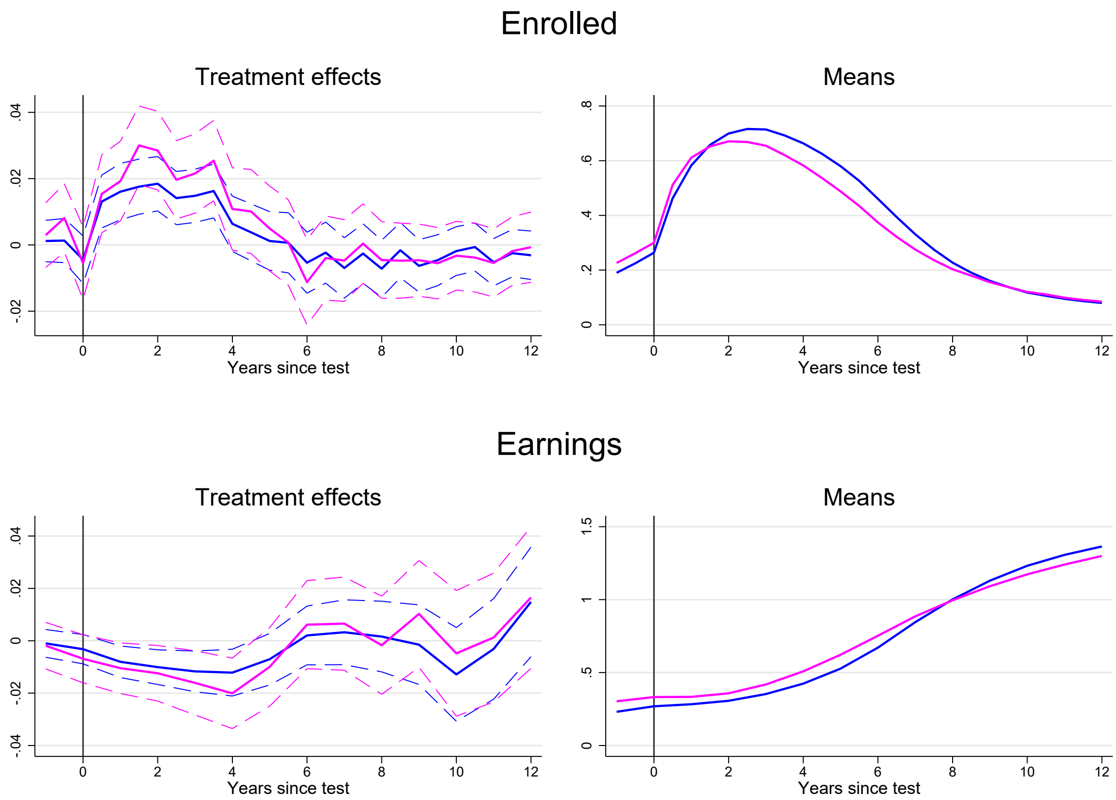
Notes: Results are shown from specifications as in Table 2, but with different sample splits as indicated. Point estimates and confidence intervals have been re-scaled by outcome means to facilitate a graphical comparison.

Figure A5: Other margins of heterogeneity: GPAs



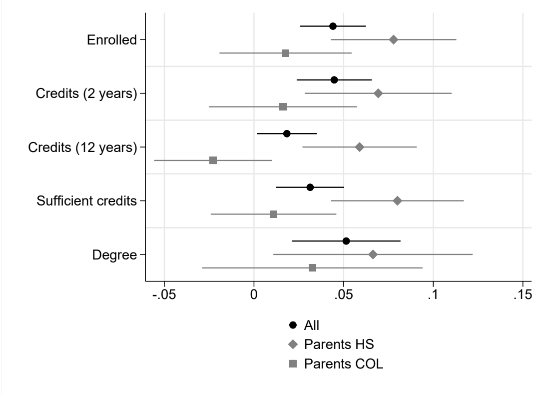
Notes: Results are shown from specifications as in Table 2. "Baseline" exactly replicates results from that table. "Re-weighted: Threshold" refers to a re-weighted sample in which the distribution of thresholds within each parental education group is the same as that in the full sample. "Re-weighted: GPA" refers to a re-weighted sample in which the distribution of the high school GPA at each threshold is nearly the same within each parental education group as in the full sample, using local deciles of the high school GPA as a basis for re-weighting.

Figure A6: Results from re-weighted specifications

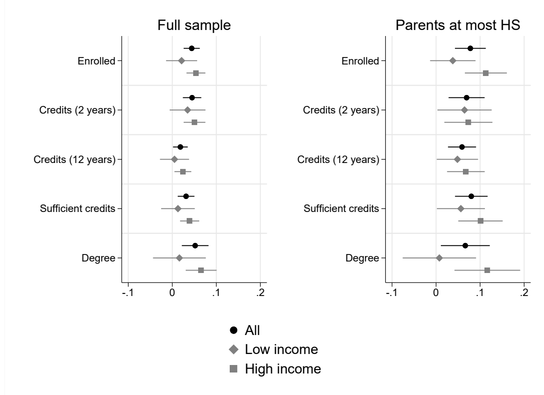


Notes: In the panels titled 'Treatment effects', point estimates (solid) and cluster-robust 95-percent confidence intervals (dashed) from estimating equation (2) are plotted against years since taking the test. In the panels titled 'Means', means of the dependent variable within the estimation sample are plotted against years since taking the test. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Results from the full and locally low GPA samples are shown in blue and magenta, respectively. Earnings are normalized by the amount one would earn while receiving the 10th-percentile wage and working full time for a whole year.

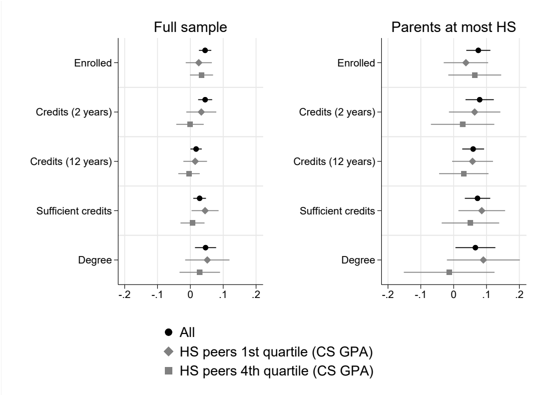
Figure A7: The effects of a higher score on enrollment and earnings



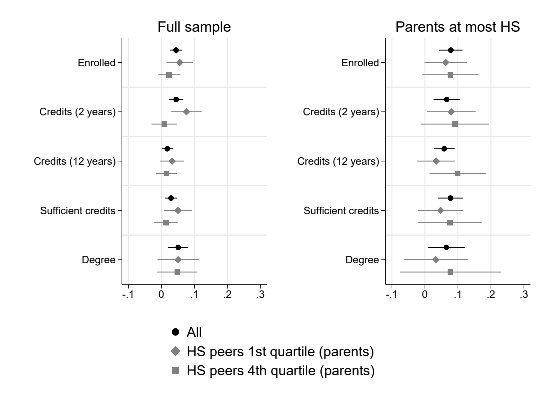
(a) Parental education



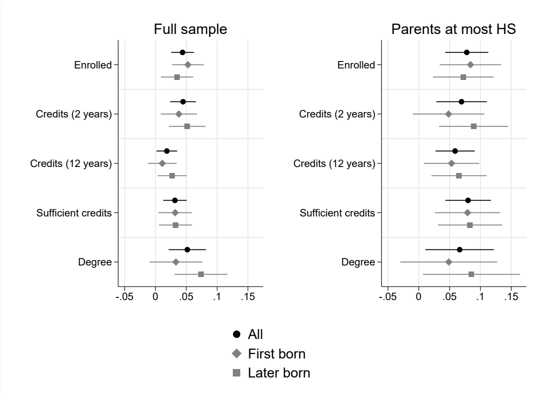
(b) Parental income



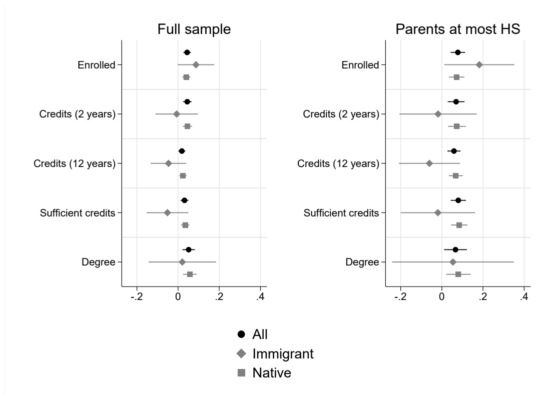
(c) High school peers: CS GPA



(d) High school peers: Parental education



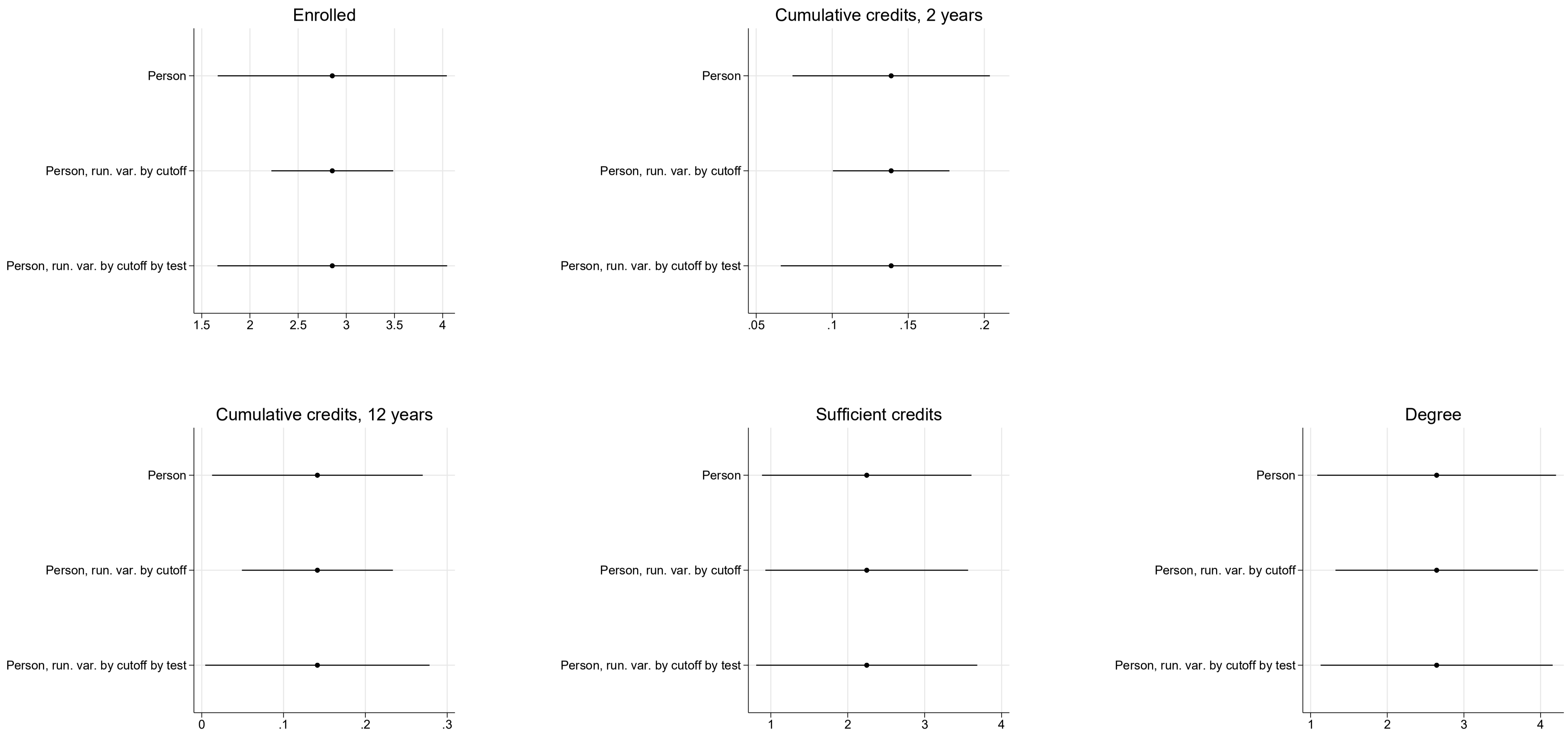
(e) Birth order



(f) Immigrant status

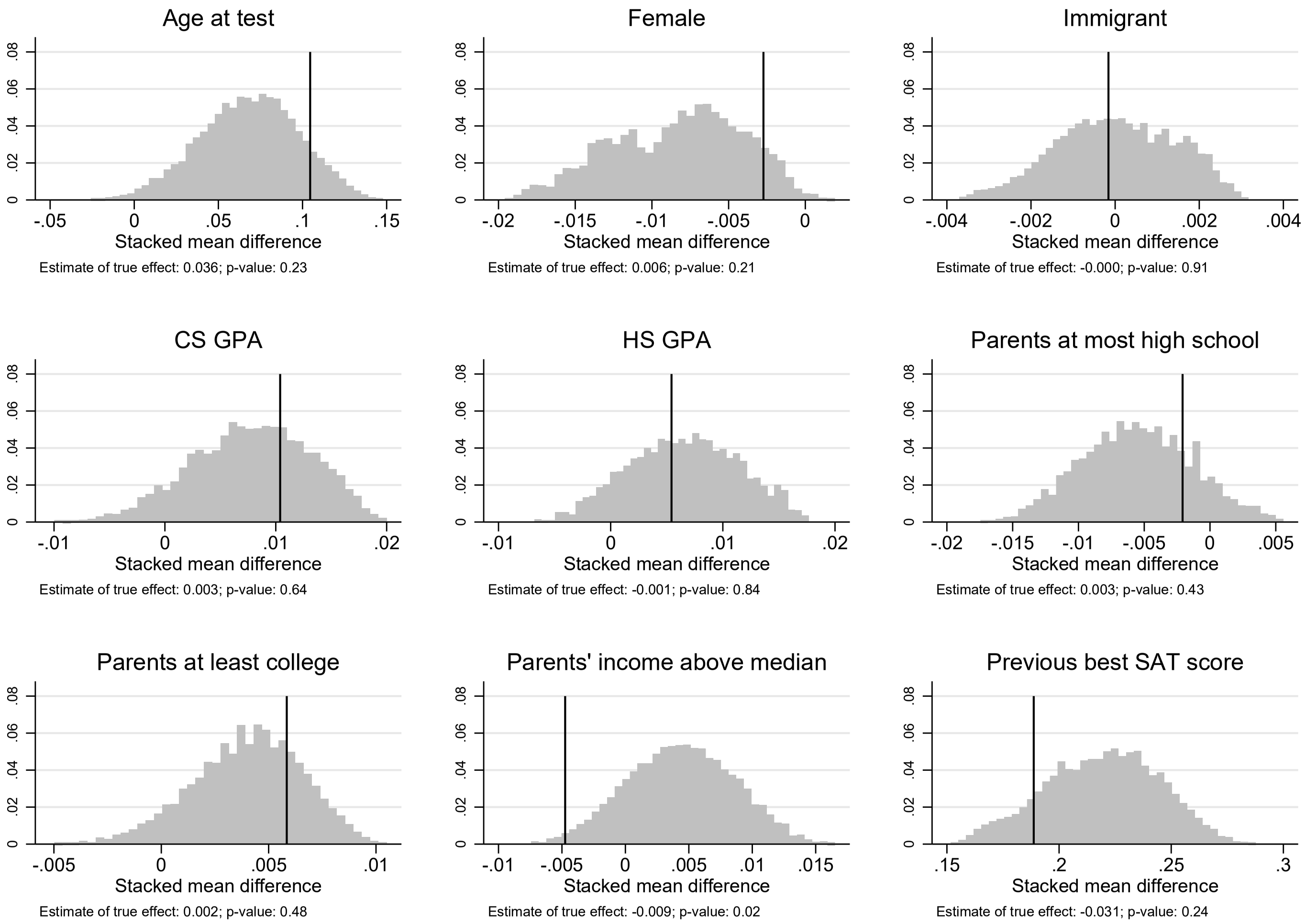
Notes: Results are shown from specifications as in Table 2, but with different sample splits as indicated. Point estimates and confidence intervals have been re-scaled by outcome means to facilitate a graphical comparison.

Figure A8: Other margins of heterogeneity: Demographics and peers



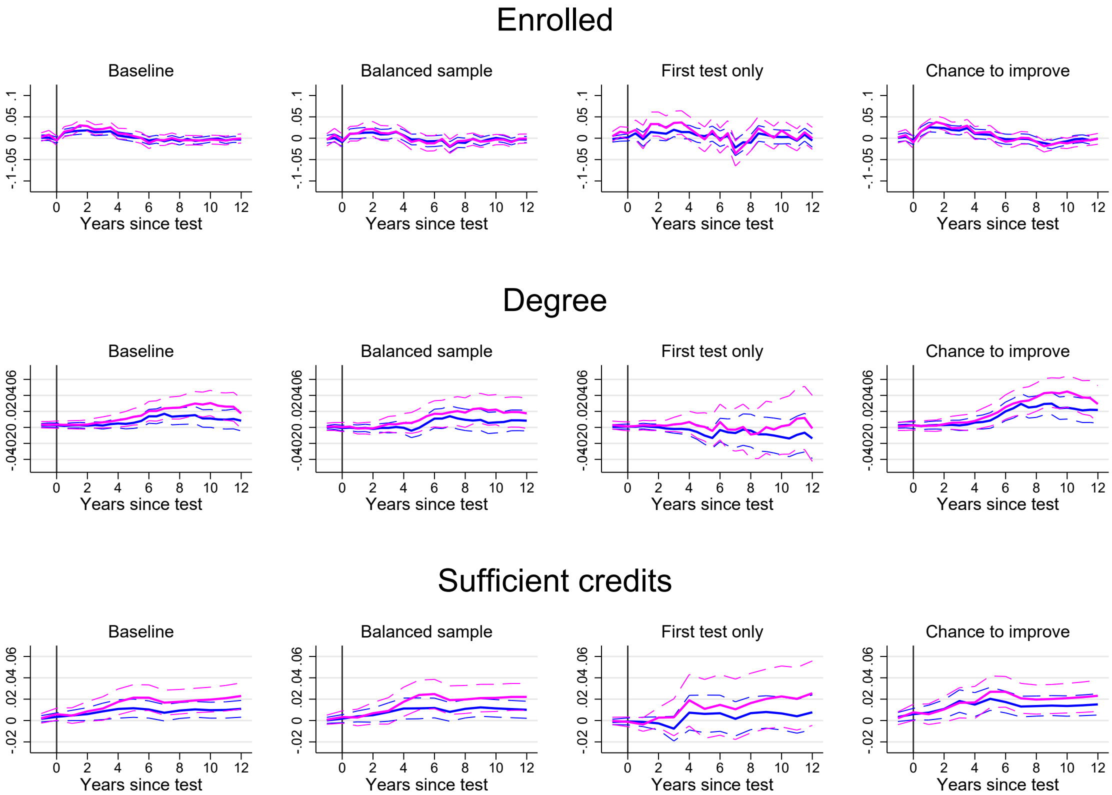
Notes: Point estimates and 95-percent confidence intervals are shown from specifications as in Table 2, varying the level of clustering. "Person" is the baseline specification exactly as in Table 2. The other two specifications allow for correlated errors along an additional dimension, as indicated.

Figure A9: Alternative clustering choices



Notes: Vertical lack lines represent the difference in the indicated outcomes between students just above and just below a threshold, averaged across thresholds 15-20. The histograms show the distribution of such stacked differences for all possible combinations across thresholds 15-20 of one-point increments in the raw score (one per threshold) that do *not* cross a threshold. "Estimate of true effect" refers to the differences between the values indicated by the black lines and the mean of the distributions represented by the histograms. *p*-values are calculated as the fraction of placebo estimates that are at least as far from their mean as the black lines.

Figure A10: Permutation inference: Pre-determined outcomes



Notes: Point estimates (solid) and cluster-robust 95-percent confidence intervals (dashed) from estimating equation (2) are plotted against years since taking the test. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Results from the full and locally low GPA samples are shown in blue and magenta, respectively.

Figure A11: The effects of a higher score on enrollment, graduation, and credits—robustness checks

Table A1: SAT and GPA cutoffs from 1999 fall admissions

College	Program	SAT cutoff	GPA cutoff
Linköping University	Medicine	20	2.4
Karolinska Institute	Medicine	20	2.4
Linköping University	Industrial Engineering	20	2.3
Uppsala University	Medicine	20	2.3
University of Gothenburg	Medicine	20	2.3
Lund University	Medicine	20	2.3
Karolinska Institute	Speech-Language Pathology	20	2.3
University of Gothenburg	European Studies	20	2.2
University of Gothenburg	Psychology	20	2.2
Swedish University of Agricultural Sciences	Veterinary Medicine	19	2.2
Stockholm School of Economics	Economics & Business	19	2.2
Royal Institute of Technology	Industrial Engineering	19	2.0
Royal Institute of Technology	Architecture	19	1.9
Royal Institute of Technology	Engineering Physics	18	2.0
Royal Institute of Technology	Computer Engineering	18	1.6
Stockholm University	Information Systems	18	1.4
University of Gothenburg	Information Systems	18	1.4
University of Gothenburg	Law	17	1.8
Stockholm University	Economics & Business	17	1.6
Stockholm University	Law	17	1.5
Chalmers University of Technology	Computer Engineering	17	1.4
Uppsala University	Economics & Business	16	1.7
Lund University	Law	16	1.6
Lund University	Economics & Business	16	1.5
Royal Institute of Technology	Electrical Engineering	16	1.1
Uppsala University	Law	15	1.5
University of Gothenburg	Social Services	15	1.4
Linköping University	Economics & Business	15	1.2
Chalmers University of Technology	Electrical Engineering	15	1.0
Stockholm University	Multimedia Pedagogy	14	1.2
Umeå University	Economics & Business	14	1.0
Lund University	Electrical Engineering	14	0.8
Royal Institute of Technology	Mechanical Engineering	14	0.7

Notes: 'SAT cutoff' is the lowest SAT score among the students admitted via the SAT quota, and 'GPA cutoff' is the lowest high school GPA (standardized by the mean and standard deviation of the 1999 high school cohort) among the students admitted via the GPA quota. For values of the SAT cutoff below 20, only the four largest programs (in terms of total number of students admitted) are listed for each cutoff, to save space. The data on cutoffs were obtained from `https://www.uhr.se/studier-och-antagning/Antagningsstatistik/Tidigare-terminer/Antagningsstatistik-urval-2-ar-2010-och-tidigare/` on 16 Feb 2018.

Table A2: The effects of receiving the top SAT score on enrollment and graduation

	Two years after test		Twelve years after test		
	(1)	(2)	(3)	(4)	(5)
	Enrolled	Cumulative credits	Cumulative credits	Sufficient credits	Degree
A. All					
Above cutoff	7.57	0.068	-0.24	-2.58	5.02
	(2.73)	(0.19)	(0.34)	(3.13)	(3.80)
Mean of dep. var.	69.9	4.7	9.3	78.7	60.4
Obs. (total—person-exam—person)	4,708—4,708—4,204		3,063—3,063—2,805		
B. Parents at most high school					
Above cutoff	12.8	-0.14	0.66	7.25	21.8
	(6.18)	(0.45)	(0.72)	(6.97)	(7.95)
Mean of dep. var.	66.5	4.9	8.8	76.3	55.6
Obs. (total—person-exam—person)	1,063—1,063—978		715—715—666		
C. Parents at least college					
Above cutoff	7.90	0.66	-0.59	-8.02	-1.63
	(4.84)	(0.32)	(0.62)	(5.00)	(6.80)
Mean of dep. var.	74.0	4.6	10.0	83.4	68.3
Obs. (total—person-exam—person)	1,368—1,368—1,197		860—860—770		
p-value	0.53	0.14	0.19	0.07	0.02

Notes: Results from the same specifications as in Table 2 are shown, except that only the highest threshold is included in the sample. See also the notes to Table 2.

Table A3: The effects of a higher SAT score on enrollment abroad, on upgrading, downgrading, and on dropping out

	(1) Enrolled	(2) Enrolled, Sweden or abroad	(3) Enrolled abroad	(4) Upgrade	(5) Downgrade	(6) Dropout
A. All						
Above cutoff	2.85	2.78	-0.074	0.45	0.85	-0.72
	(0.61)	(0.60)	(0.14)	(0.30)	(0.31)	(0.42)
Mean of dep. var.	64.8	66.0	1.2	2.4	2.5	13.6
Obs. (total—person-exam—person)		120,540—89,688—55,332			120,540—89,688—55,332	
B. Parents at most high school						
Above cutoff	4.89	4.83	-0.056	0.95	0.59	-2.13
	(1.12)	(1.11)	(0.23)	(0.49)	(0.54)	(0.82)
Mean of dep. var.	62.8	63.8	0.9	2.2	2.7	15.9
Obs. (total—person-exam—person)		36,827—27,358—18,452			36,827—27,358—18,452	
C. Parents at least college						
Above cutoff	1.19	0.87	-0.31	-0.023	0.52	-0.40
	(1.27)	(1.25)	(0.34)	(0.73)	(0.70)	(0.82)
Mean of dep. var.	67.5	69.1	1.6	2.9	2.5	10.2
Obs. (total—person-exam—person)		25,312—18,772—10,168			25,312—18,772—10,168	
p-value	0.03	0.02	0.54	0.27	0.94	0.14

Notes: Results from estimating (2) are reported. The unit of observation is person-test-threshold. Samples are restricted to individuals whose GPA is below the local median. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Coefficients, standard errors, and means have been multiplied by 100. Standard errors, clustered by individual, in parentheses.

B Descriptive evidence on SAT takers

Here, we briefly discuss a set of descriptive patterns regarding SAT participation and performance as a general background. The Swedish SAT is not compulsory and it is possible to apply for college using only the high school GPA. Nevertheless, within the most recent cohorts in our data about half of high school graduates take the test at least once before age 30 and a quarter participate more than once. Students typically take the test for the first time in the fall of the year they graduate from high school. It is, however, not uncommon to take the test later on, even for the first time, as is reflected by the age distribution of test takers. See Figure B1, which illustrates incidence, frequency, and timing of test taking.

We report predictors of SAT participation in Table B1. As expected, raw correlations between ability measures (compulsory school GPA and high school GPA) and the probability of ever taking the SAT are positive. Notably, however, the high school GPA has two potentially counteracting effects on the incentives to take the test. It is correlated with ability, which should increase the usefulness of the test, but it is also an alternative 'application currency' which should reduce the incentives to take the test (conditional on ability). Along these lines, the impact of high school GPA is found to be *negative* when we control for compulsory school GPA (which only measures ability, not application currency). Thus, somewhat simplified, the results suggest that the typical test takers are able students without very attractive alternative admission opportunities. Notably, compulsory school GPA remains a very strong predictor of SAT participation even conditional on high school GPA, gender, immigration status, and parental education: a one-standard-deviation increase in the compulsory school GPA is associated with a 20 percentage point increase in the probability of taking the SAT.

The relationships between grades and the probability of repeating the test are similar to those for incidence. Compulsory school GPA (that is, ability) is positively associated with repeating, whereas high school GPA is negatively associated once we condition on compulsory school GPA. See Table B2.

Females are more likely to take the test, but less likely to repeat. Natives and individuals with high-educated parents are more likely to ever take the test and to repeat. These patterns remain (although a bit muted) after controlling for grades. However, after controlling for the SAT score in the first attempt, the association between female and repeating turns positive (see also the discussion of the *causal* effect of a higher score on repeating in Section 3).

Turning to the predictors of performance on the SAT, we find large differences related to ability indicators and demographic background. Females and foreign born score worse on the SAT by about 50 and 80 percent of a standard deviation, respectively. As expected, GPAs are positive predictors of SAT performance, and when controlling for the GPAs, as well as parental education, the coefficient on female becomes even more negative. First-time takers and older test takers perform worse, although the association with age is positive after controlling for GPA. Controlling for individual fixed effects, first-time takers still do worse, and older test takers do better, both of which likely reflect learning. See Table B3.

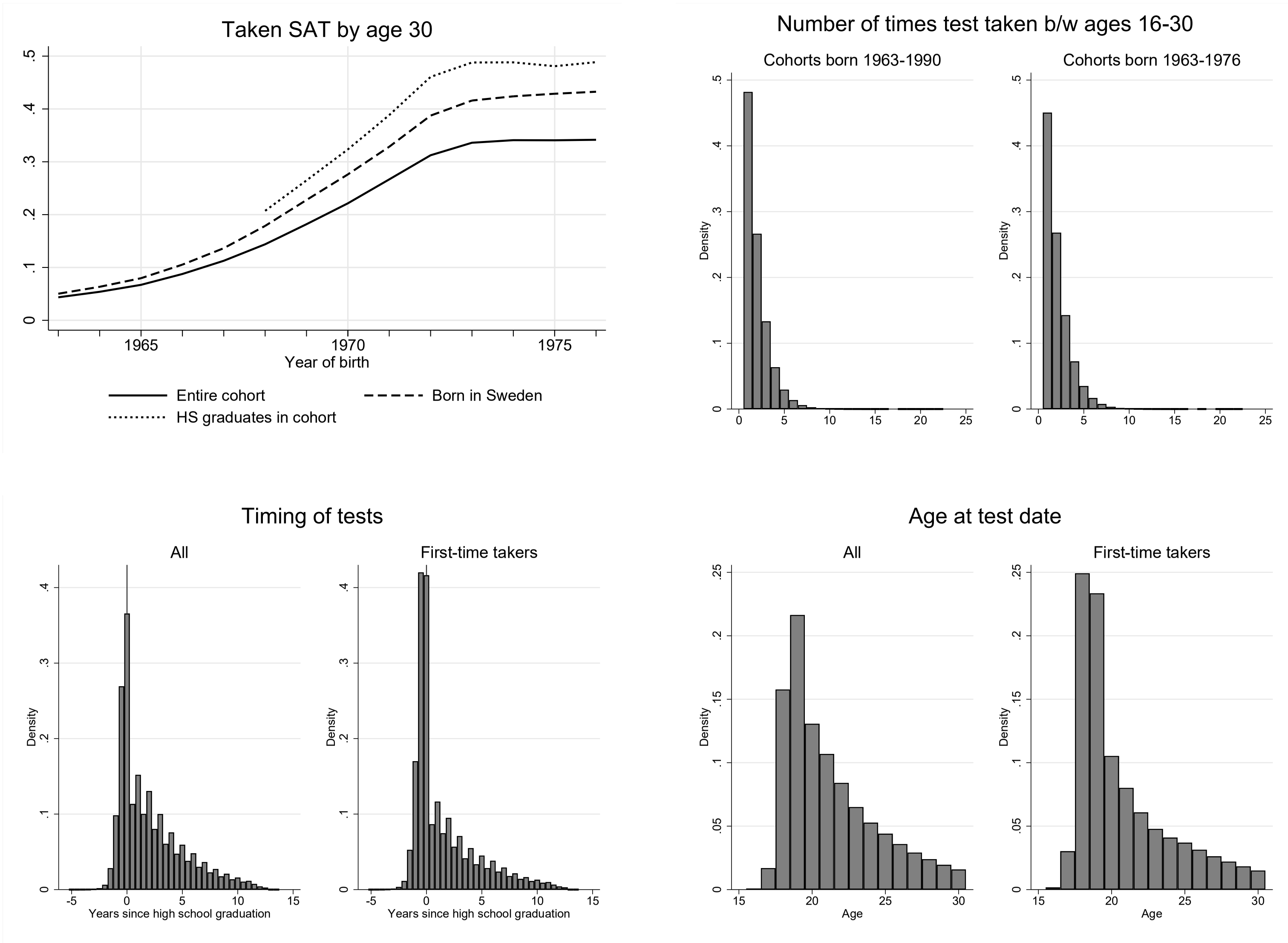


Figure B1: The incidence, frequency, and timing of taking the SAT

Table B1: Predictors of ever taking the SAT

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	0.037	0.061	0.063	-0.0096	0.020	-0.0084	-0.00032
	(0.00038)	(0.00071)	(0.00070)	(0.00068)	(0.00071)	(0.00068)	(0.00068)
Foreign born	-0.21	-0.053	-0.076	-0.0079	-0.014	-0.010	-0.028
	(0.00036)	(0.0016)	(0.0016)	(0.0015)	(0.0016)	(0.0015)	(0.0015)
Parents at most high school			-0.16				-0.082
			(0.00074)				(0.00073)
Both parents hold college degree			0.15				0.067
			(0.0013)				(0.0013)
CSGPA (z-score)				0.20		0.21	0.20
				(0.00037)		(0.00054)	(0.00055)
HSGPA (z-score)					0.11	-0.014	-0.020
					(0.00034)	(0.00047)	(0.00047)
Restricted sample			✓	✓	✓	✓	✓
R-squared	0.13	0.14	0.18	0.27	0.19	0.27	0.28
Obs	3,865,021	1,539,231	1,539,231	1,539,231	1,539,231	1,539,231	1,539,231

Notes: Results from OLS regressions are shown. The dependent variable is an indicator for ever having taken the test. All regressions control for year-of-birth fixed effects. Samples include all individuals born between 1963 and 1990. The restricted sample consists of individuals whose compulsory school GPA, high school GPA, and parents' education are observed. Robust standard errors in parentheses.

Table B2: Predictors of repeating the SAT

	(1)	(2)	(3)	(4)	(5)
Female	-0.037	-0.032	-0.039	-0.030	0.0082
	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)
Foreign born	-0.061	-0.064	-0.066	-0.072	-0.030
	(0.0032)	(0.0032)	(0.0032)	(0.0032)	(0.0032)
Parents at most high school		-0.052		-0.053	-0.040
		(0.0014)		(0.0014)	(0.0014)
Both parents hold college degree		0.060		0.068	0.061
		(0.0020)		(0.0020)	(0.0020)
CSGPA (z-score)			0.071	0.064	0.021
			(0.0012)	(0.0012)	(0.0013)
HSGPA (z-score)			-0.058	-0.064	-0.081
			(0.00096)	(0.00096)	(0.00098)
Score fixed effects					✓
R-squared	0.13	0.14	0.14	0.14	0.17
Obs	541,309	541,309	541,309	541,309	541,309

Notes: Results from OLS regressions are shown. The dependent variable is an indicator for repeating the test at least once within five years of the first test. The sample includes only first tests. All regressions control for date-of-test and cohort dummies. Robust standard errors in parentheses.

Table B3: Predictors of performance in the SAT

	(1)	(2)	(3)	(4)	(5)
Female	-1.69	-2.27	-1.60	-2.27	
	(0.013)	(0.011)	(0.013)	(0.011)	
Foreign born	-2.38	-1.80	-2.26	-1.68	
	(0.039)	(0.032)	(0.036)	(0.030)	
Parents at most high school		-0.81		-0.74	
		(0.012)		(0.012)	
Both parents hold college degree		0.97		0.86	
		(0.017)		(0.016)	
CSGPA (*z*-score)		2.07		2.21	
		(0.011)		(0.011)	
HSGPA (*z*-score)		1.12		1.34	
		(0.0092)		(0.0088)	
First time			-2.72	-1.81	-0.68
			(0.0081)	(0.0072)	(0.0067)
Age at test			-0.064	0.36	0.57
			(0.0024)	(0.0021)	(0.0025)
Individual fixed effects					✓
R-squared	0.05	0.32	0.13	0.41	0.94
Obs	1,102,932	1,102,932	1,102,932	1,102,932	1,102,932

Notes: Results from OLS regressions are shown. The dependent variable is the normalized SAT score ranging from 0 to 20 in integer steps. One step corresponds to about one third of a standard deviation. All regressions control for cohort dummies. Standard errors, clustered by individual, in parentheses.

C Extensions to the theoretical model

C.1 Repeating

A salient feature of our empirical setting is that students may repeat the SAT. The choice of repeating can easily be incorporated into our framework, which leads to the predictions that individuals are more likely to repeat the test if they have a higher probability of succeeding at the next attempt; if they are more patient; if they face lower costs of preparing for the test; and if they enjoy a higher return to a higher-quality college program over their current option (working or being enrolled in a lower-quality program). Parental education may correlate negatively with discount rates and costs of preparing for the SAT. High-SES students will then be more likely to repeat the test, and will exhibit lower intention-to-treat effects of higher scores for long-run educational attainment.

To formally establish these results, start by supposing that individuals have infinite lives and let the life-time utility of the outside option (not repeating) be denoted by V^O. Consider repeating as an option for individuals who work as well as those who are already enrolled, so that $V^O \in \{V^J, V_0^C\}$. Let the discount factor be denoted by β. Suppose that utility flows are constant across periods, except when paying a waiting cost or when switching states. The flow value of the outside option during a single period of waiting can then be written as $(1-\beta)V^O$.

Consider an individual who has not succeeded at the SAT yet (in the sense of achieving the higher score). Let us assume that it is possible to repeat the SAT once, at the beginning of the next period; and that a new enrollment decision can be made immediately after taking the test. Further, we assume that repeating means that a cost κ is incurred, reflecting any study that must be undertaken in preparation of the test. The probability of a higher score at the next test is p. Thus, individuals decide to retake if and only if

$$(1-\beta)V^O - \kappa + \beta \left\{ pV_1^C + (1-p)V^O \right\} > V^O,$$

which can be rewritten, also making concrete the two possible outside options, as

$$p\beta\left(V_1^C - V^J\right) > \kappa, \qquad p\beta\left(V_1^C - V_0^C\right) > \kappa. \tag{C1}$$

The comparative statics predictions stated at the start of this section follow directly from equation (C1).

Regarding heterogeneity, note that if low-SES individuals are of Type II, then they could actually be *more* likely to repeat the test, since $V_{II}^C(q_1) - V^J > V_I^C(q_1) - V_I^C(q_0)$. Differences in study costs and discount rates may counteract this, however. If low-SES individuals were of Type III, then based on valuations alone they would be less likely to repeat the test than high-SES individuals since $V_{III}^C(q_1) - V^J < V_I^C(q_1) - V_I^C(q_0)$.

C.2 Differences to admissions discontinuities

The model highlights the contrast between our empirical setting and admission cutoffs used by Kirkeboen et al. [2016] and others. These studies rely on a centralized admission system which can be assumed to elicit a truthful reporting of students' preferences across college programs, and which allocates program slots among applicants based on unpredictable grade cutoffs. In such a setting, the instrument will be a dummy variable for being above the cutoff for a given program c. The complying population for this instrument are individuals who enroll in program c if the instrument is switched on, and in their next-best program if it is switched off (in practice marginal students with cross-field preferences). This setting provides the researcher with variation in program or field of study in isolation. This is crucial for estimating the economic returns to field of study, and for informing policies that change the number of slots in given programs or fields.

In contrast, studying variation in enrollment opportunities arising from SAT-discontinuities allows us to unravel a large set of adjustments in terms of educational choices and outcomes. For example, our strategy allows us to identify effects on college participation that arise if expanded opportunities really do increase the likelihood of participating in 'always-attainable' programs as discussed above. This is not possible with admissions discontinuities since the option of participating in the lower-ranked program is removed when accepted at a higher-ranked program within centralised admissions systems. In addition, our strategy allows us to detect adjustments that arise due to increased returns to information-gathering before students rank their choices, a type of insight that cannot be derived at the later (admissions) stage when the ranking has already been submitted.