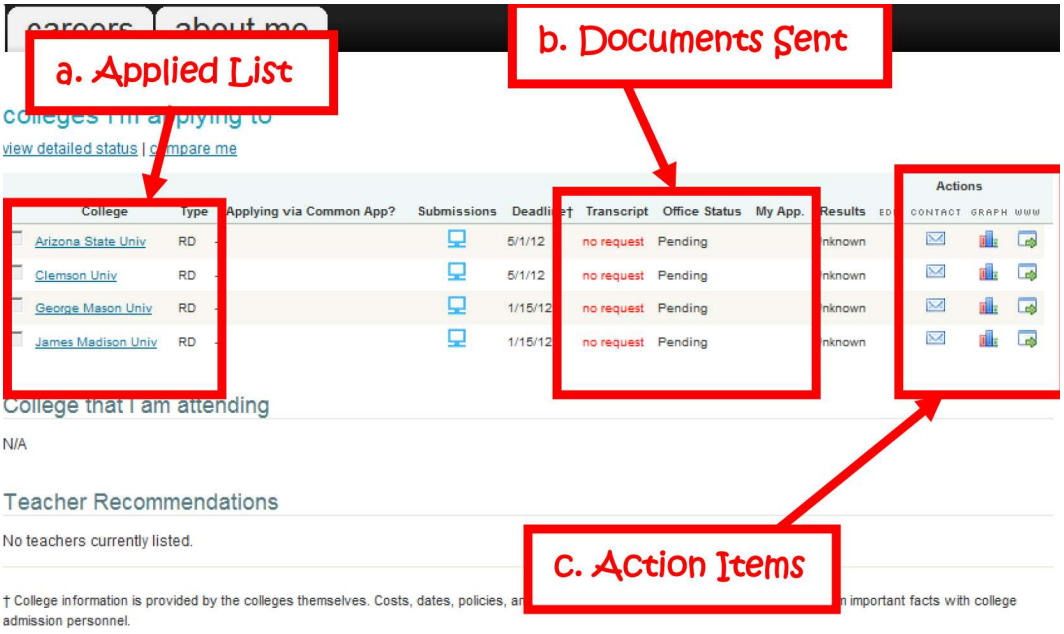
Appendix

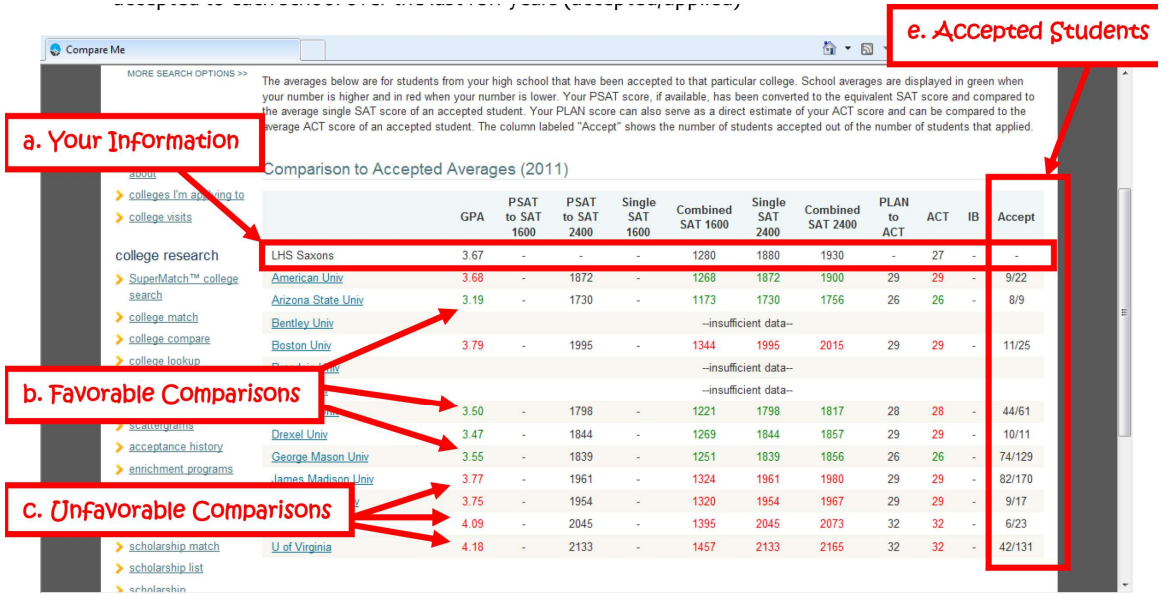
A Additional Tables and Figures

Figure A.1: Example of College Dashboard on Naviance



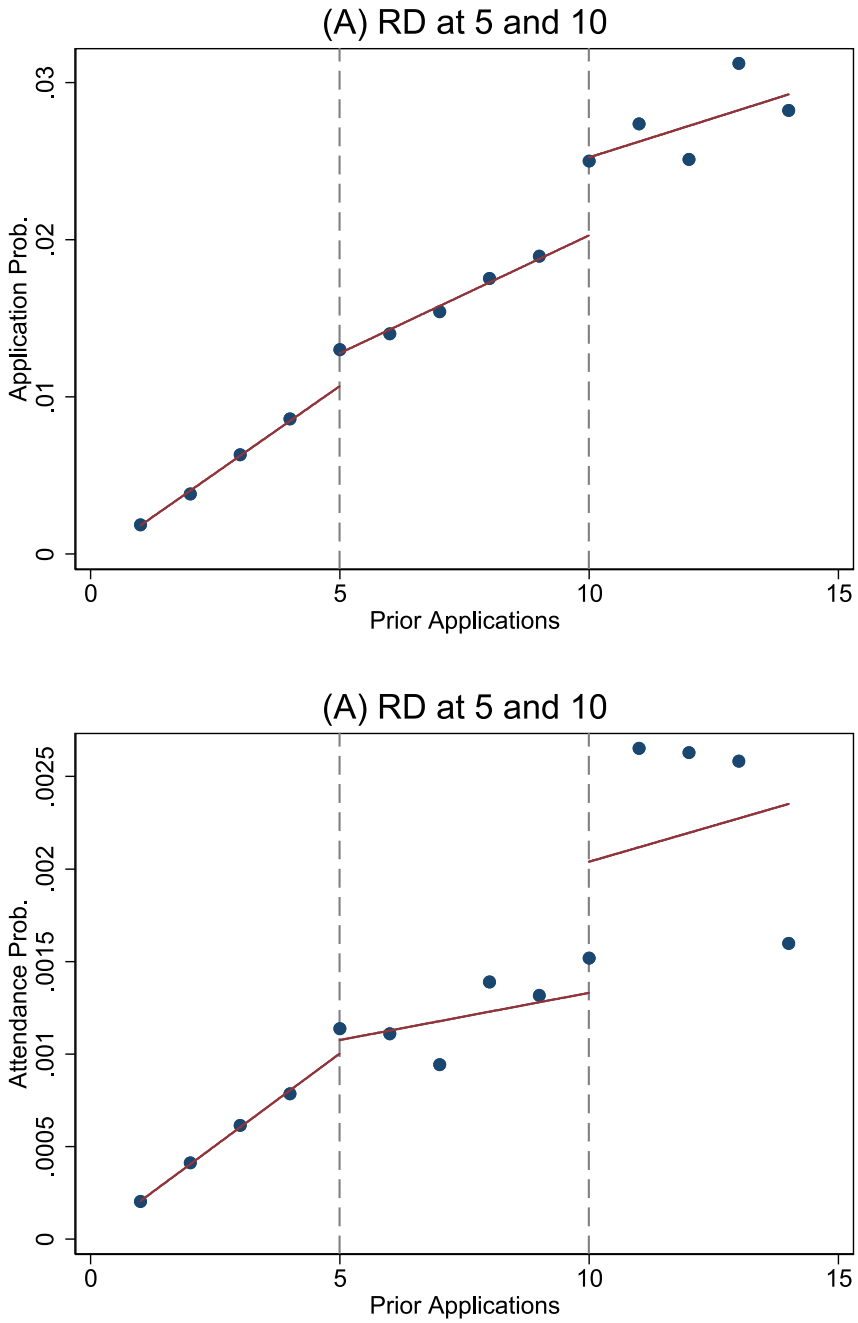
Note: This is an example of the college dashboard on Naviance. (This is not from the district studied). The red boxes were added as notes by the high school posting instructions on how to use Naviance. Source: Langley High School and Naviance.

Figure A.2: Example of College Comparisons on Naviance



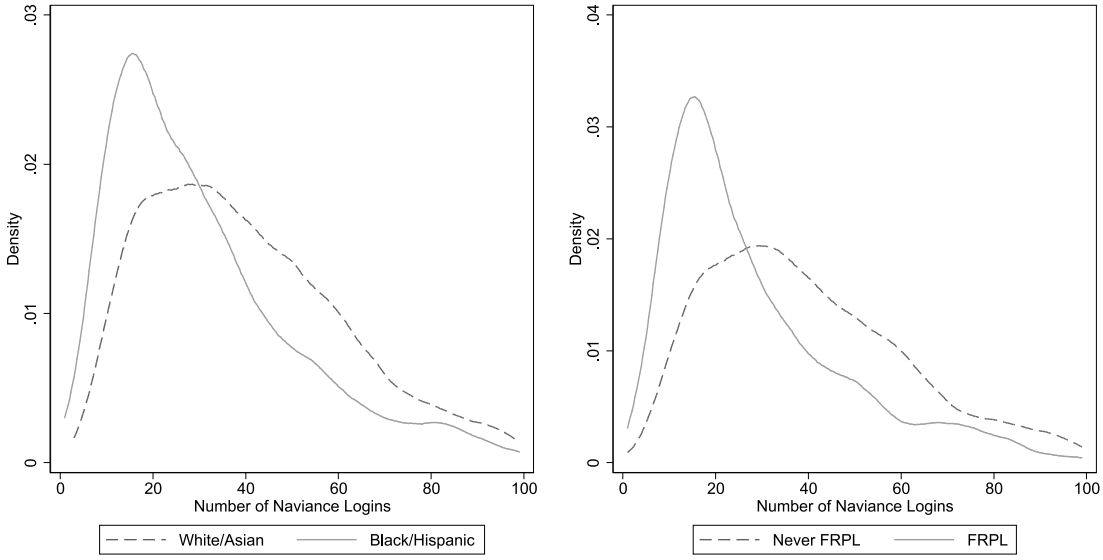
Note: This is an example of the how students can compare colleges on Naviance. (This is not from the district studied). The red boxes were added as notes by the high school posting instructions on how to use Naviance. Source: Langley High School and Naviance.

Figure A.3: Discontinuities at 5 and 10 Prior Applications



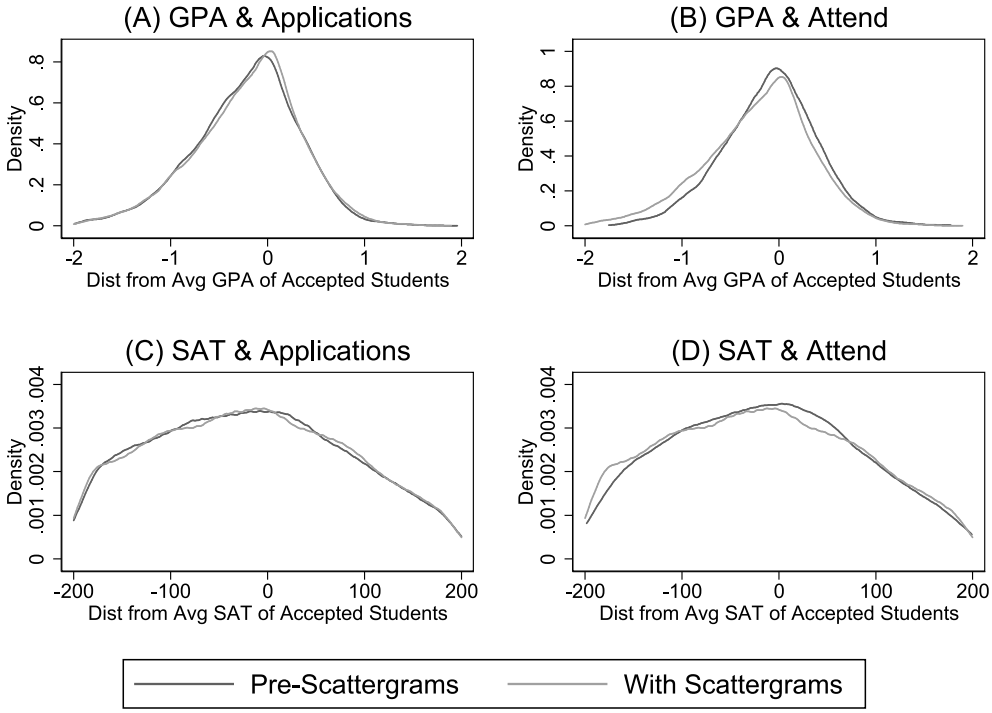
Note: The figures above show how the probability of a student applying to (A) or attending a college (B) changes based on the number of applications a college has received from the student's high school prior to the student's year of graduation (2015 or 2016). Each dot on the x-axis represents the exact number of applications sent from the high school. On the y-axis, the dot indicates the average fraction of students who applied to or attended each of the colleges with the associated number of prior applications. The fitted lines are from a local linear regression discontinuity model. The graph includes all student-college combinations for which at least one and fewer than fourteen applications from the student's high school were sent to the college since 2014 and prior to the student's graduation year. Scattergrams become visible when a college has received five or ten applications. High schools choose which visibility threshold to use but I could not determine which threshold applied to which high school.

Figure A.4: Density of Naviance Login Rates for the Class of 2017



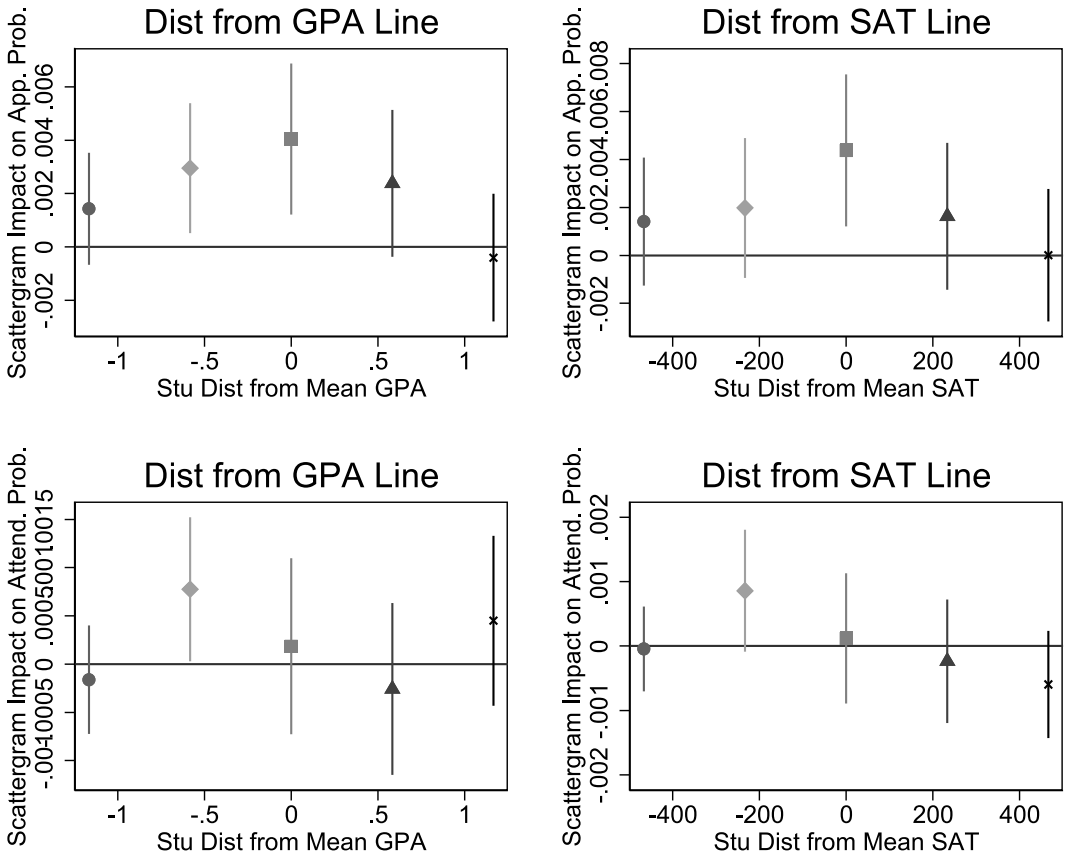
Note: The figures above indicate the densities for the number of times students logged onto Naviance. These login rates are only available for the class of 2017 but they include all logins since the fall of 2014. (These include all logins through the student's account and thus may capture parents or other individuals logging onto Naviance.) The panel on the left compares the login rates of Black and Hispanic students to those for white and Asian students. The panel on the right compares login rates for students who never received free or reduced-price lunch to students who received it at least one year while enrolled in the district.

Figure A.5: Application and Attendance Density by Distance from Mean GPA or SAT



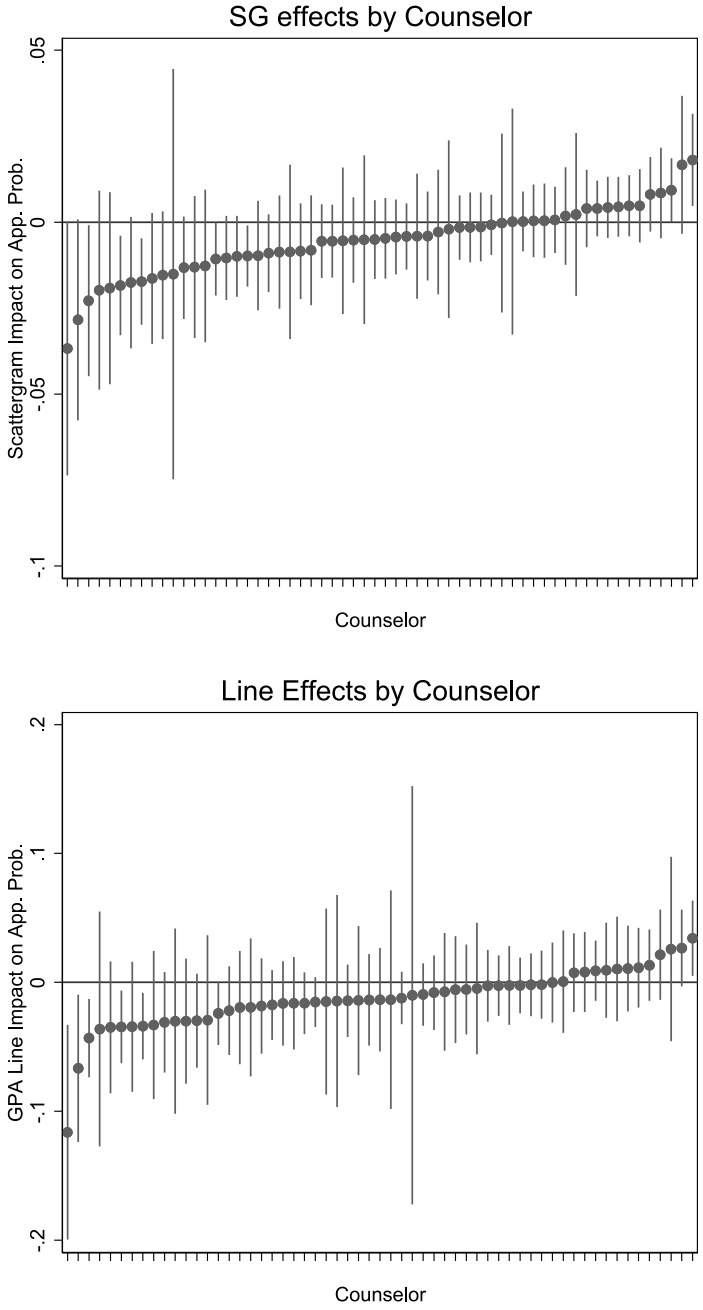
Note: The figures above show how the types of colleges to which students applied or attended shifted when scattergrams became available. In particular, they show the densities of applications (A and C) and attendance (B and D) as a function of the student's GPA or SAT distance from the average GPA or SAT, of all admitted students in the district, at the college to which they apply or attend. I use the district-wide averages because these may be a more accurate measure of the college's admissions criteria than the school averages, especially for colleges with only a few admitted students from a high school. This also enables me to calculate averages for colleges with only a few or no students admitted at some high schools. The figures are based on weighted GPAs and SAT scores on the 2400 scale. The "Pre-Scattergrams "line is based on the students graduating high school in 2014 and the "With Scattergrams" line is based on students graduating in 2015 and 2016. Students who graduated in 2014 could not see scattergrams.

Figure A.6: Magnitude of Scattergram Impact by Proximity to Typical Acceptee



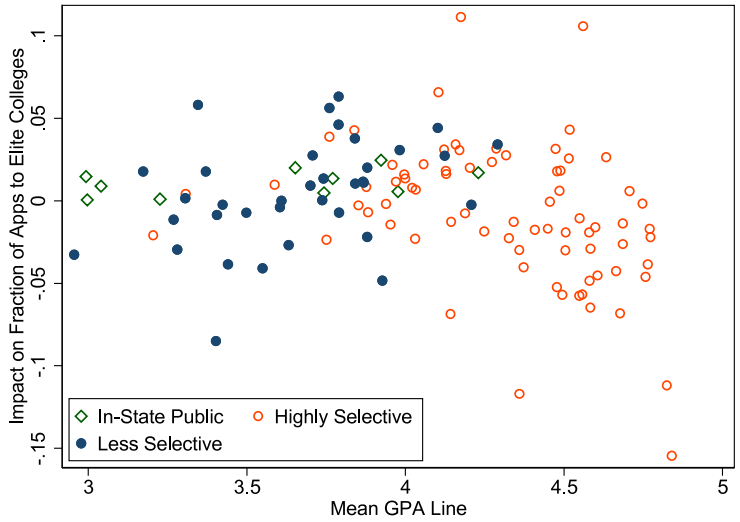
Note: The above figures show how the magnitudes of the discontinuities in application or attendance probabilities at the visibility thresholds vary based on how similar the students were to the typical acceptee's from their high school. I construct these by separately estimating discontinuities for students who are of varying distances from the typical acceptee's GPA and SAT. For the GPA, I bin students in .5 GPA intervals, starting with students who are within .25 GPA points of typical acceptee. For the SAT I use bins of 150. This is based on weighted GPAs and SAT scores on the 2400 scale. For students who could not see a scattergram, I calculate how far a student would have been from the typical acceptee line based on prior applications. The middle dot in panel (A) indicates, for students whose GPA was within .25 GPA points of the typical acceptee's, how much more likely they were to apply to a college if they could see its scattergram compared to similar students who could not see the scattergram. The bars indicate the standard errors of the discontinuity estimates (where standard errors are clustered by student). These estimates are based on regressions which include student and college by year fixed effects.

Figure A.7: Effect Sizes by Counselor



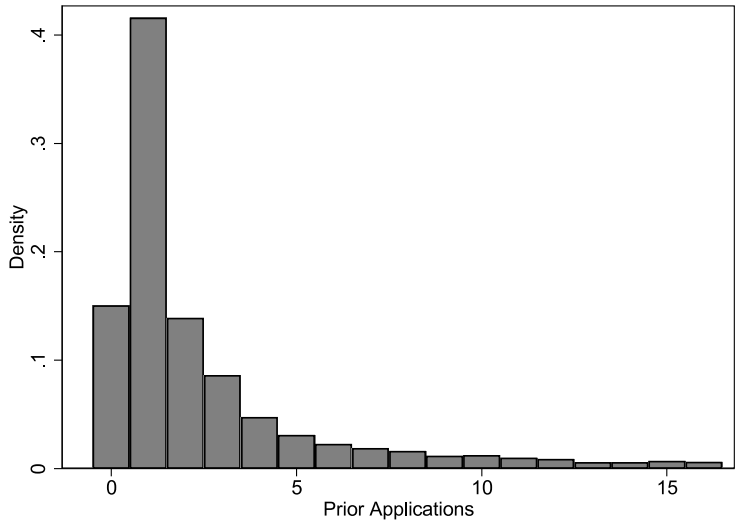
Notes:The figures above show the interactions between counselor effects and scattergram effects. Panel (A) shows how the effect of gaining access to a scattergram varies across counselors. Panel (B) shows how the impact of being below the GPA line varies across counselors. Counselors are assigned by last name so this variation may be due to differences in counseling practices rather than sorting into counselors.

Figure A.8: Impact of Individual Scattergrams on Elite College Applications



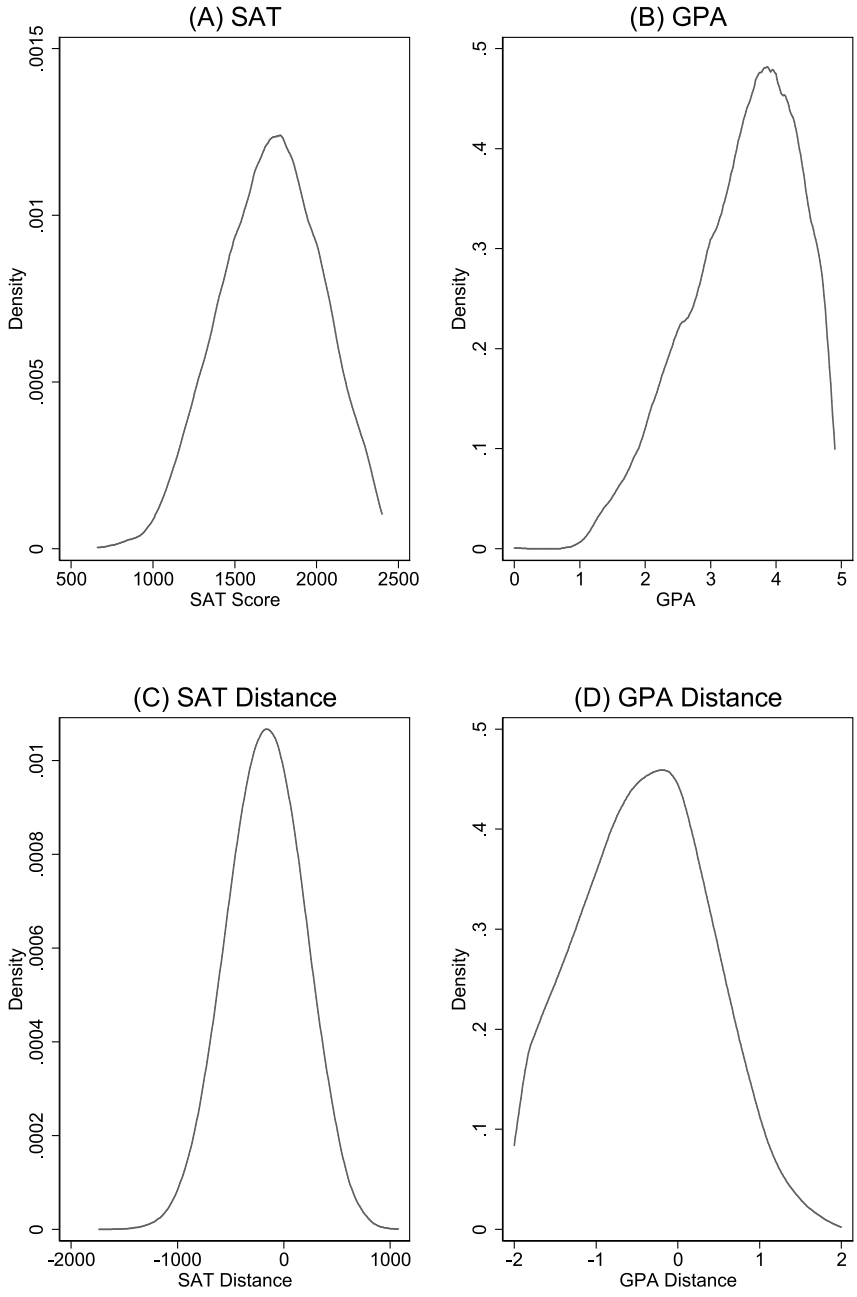
Note: The figure above plots the average impact of a college's typical acceptee GPA line on the fraction of elite colleges to which a student applies. Each dot represents the average impact of an individual college's line (across all the high schools). Elite colleges are the public and private colleges defined as "Elite" by Barron's *Profiles of American Colleges*. The x-axis represents the average location of the college's weighted GPA line, across all high schools in the district.

Figure A.9: Density of Prior Applications



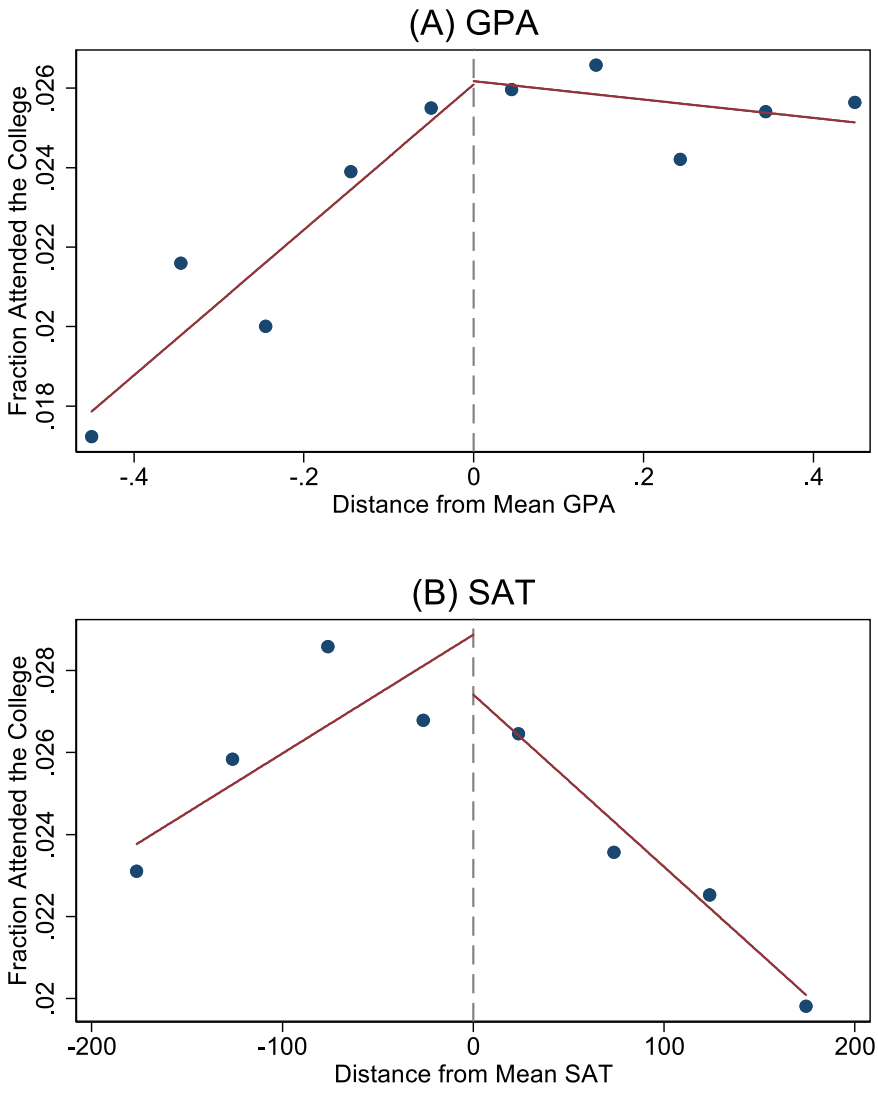
Note: The figure above depicts the density of prior applications received by colleges. Prior applications refers to the cumulative number of applications received by a college from a high school since 2014 but prior to the current year. For each high school, it includes the set of colleges which received an application from that high school between 2014 and 2016.

Figure A.10: Densities of SAT scores and GPAs



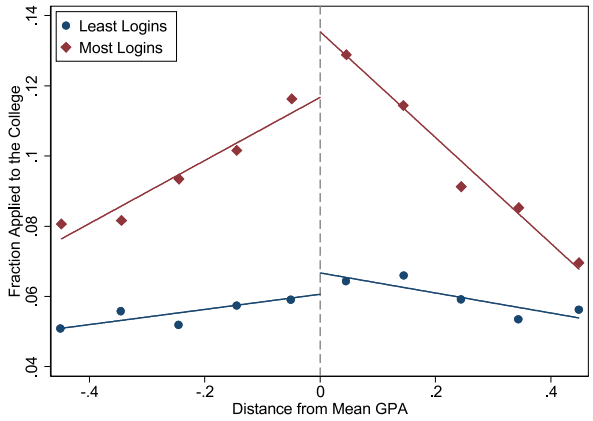
Note: The top row shows the densities for SAT scores (the old version on the 2400 scale) in panel (A) and weighted 11th grade GPAs in panel (B). The bottom row shows the densities for the distance of the student's SAT score (old 2400 version) and weighted GPA from the typical acceptee's SAT or GPA on the scattergram. Scores on the new version of the SAT have been converted to the old version equivalent score using the scale provided by the College Board. Maximum SAT scores are used in these figures. The figures are based on student-scattergram combinations since the same student has a different distance value for each scattergram. Thus, students may appear multiple times in each figure. There is no statistically distinguishable evidence of heaping on either side of the mean SAT or GPA lines.

Figure A.11: Attendance Probability by Distance from Typical Acceptee Lines



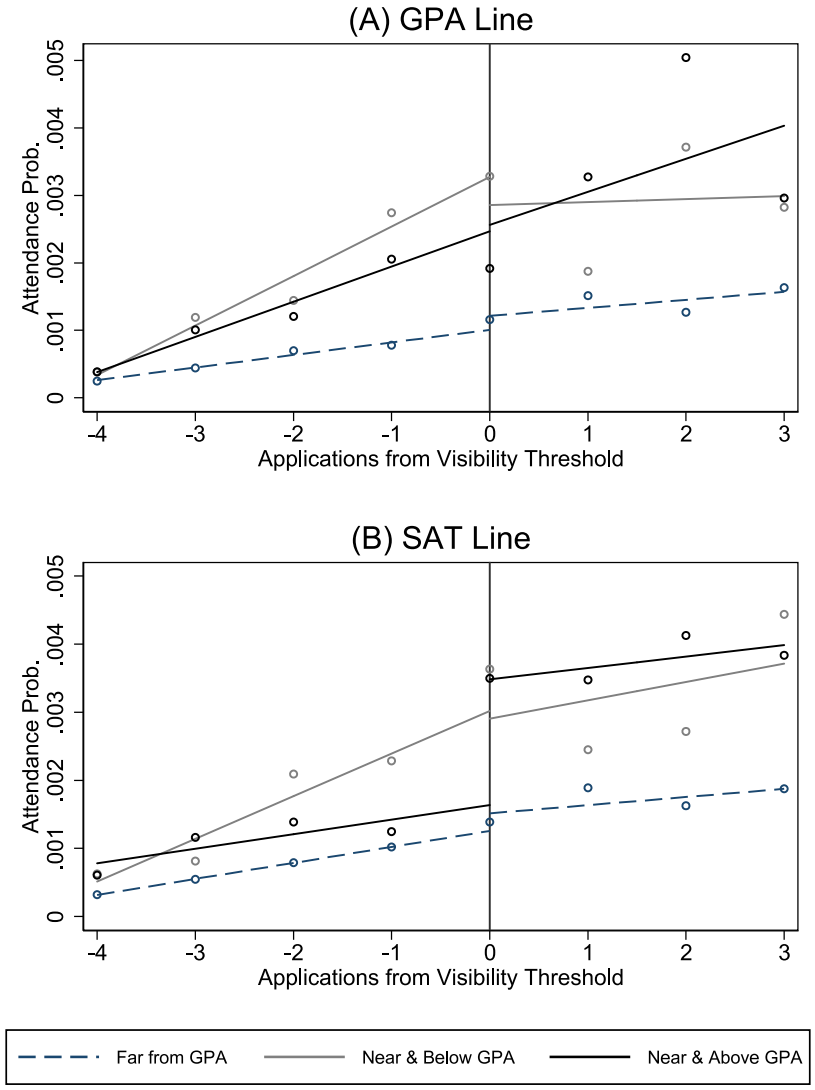
Note: The figures above show how attendance rates varied based on a student's position on a scattergram relative to the typical acceptee's GPA (in panel A) and SAT (in panel B). For panel (A), Weighted GPAs from 11th grade are used to determine the distance from the mean weighted GPA line depicted on the scattergram when the student is in 12th grade. The data are binned in intervals of 0.1 from the threshold at zero. For panel (B), students' maximum SAT scores on the old 2400 scale are used. The data are binned in intervals of 50 from the threshold at zero. The fitted lines come from a local linear regression discontinuity model with a bandwidth of 0.5 for panel (A) and a bandwidth of 200 for panel (B). The y-axis represents the fraction of students in each bin who attended the college (in 2015 or 2016). A bin includes multiple scattergrams (and colleges) and it may include the same students multiple times (but for different scattergrams). Observations are student-scattergram combinations (where scattergrams are based on colleges which received at least 10 prior applications).

Figure A.12: Application Probability by Distance from Mean GPA and Naviance Logins



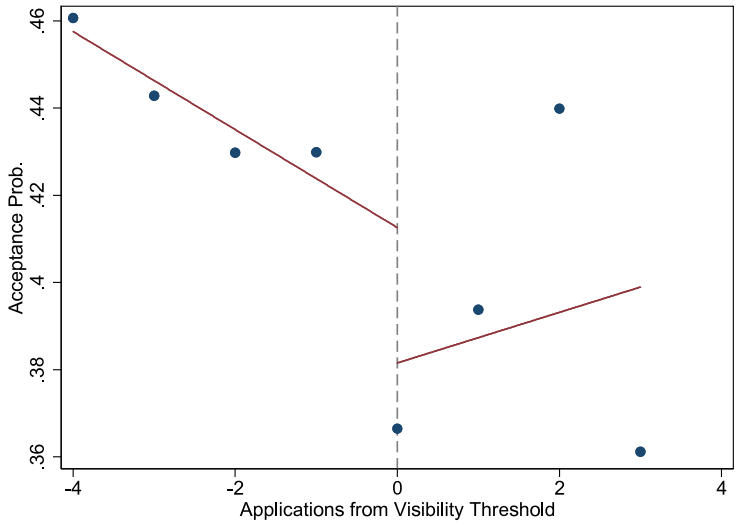
Note: The figure above shows how application rates varied for the class of 2017 based on a student's position on a scattergram relative to the typical acceptee's GPA. Login data are only available for students who graduated from the district in 2017. The red line indicates students who logged onto Naviance the most (top 50%) since 2014. The blue line is based on students who logged on the least (bottom 50%). Logins count any time the student's account is used. Weighted GPAs from 11th grade are used to determine the distance from the mean weighted GPA line depicted on the scattergram when the student is in 12th grade. The data are binned in intervals of 0.1 from the threshold at zero. The y-axis represents the fraction of students in each bin who applied to the college. The fitted lines come from a local linear regression discontinuity model with a bandwidth of 0.5. Observations are student-scattergram combinations (where scattergrams are based on colleges which received at least 10 prior applications).

Figure A.13: Impact of Scattergram Visibility on Attendance by Proximity to Typical Acceptee



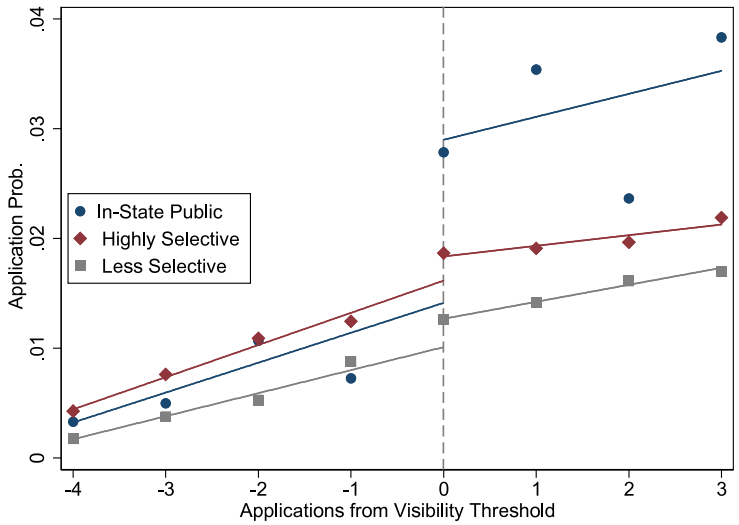
Note: The figures above show how the probability of a student attending a college changes when a college crosses a scattergram visibility threshold, and how this varies based on the proximity of the student to the typical acceptee lines. Panel (A) is based on the weighted GPA lines and near is defined as within .1 GPA points. Panel (B) is based on the SAT 2400 scale and near is defined as within 50 SAT points. I computed hypothetical typical acceptee lines for colleges which did not meet the cutoff for a scattergram based on the prior applications and used these to compute near, far, above and below, for student-college combinations to the left of the RD threshold. Students to the left of the RD threshold would not have seen these lines. Observations are student-college-threshold combinations. I used distances to both thresholds (five and ten) where relevant. The X-axis shows how far a college was, in terms of applications, from each of these minimum applicant thresholds (in 2015 and 2016). The dots on the y-axis represents the fraction of students who attended a college with previous applications x distance from the threshold.

Figure A.14: Impact of Scattergram Visibility on Acceptance (Conditional on Applying)



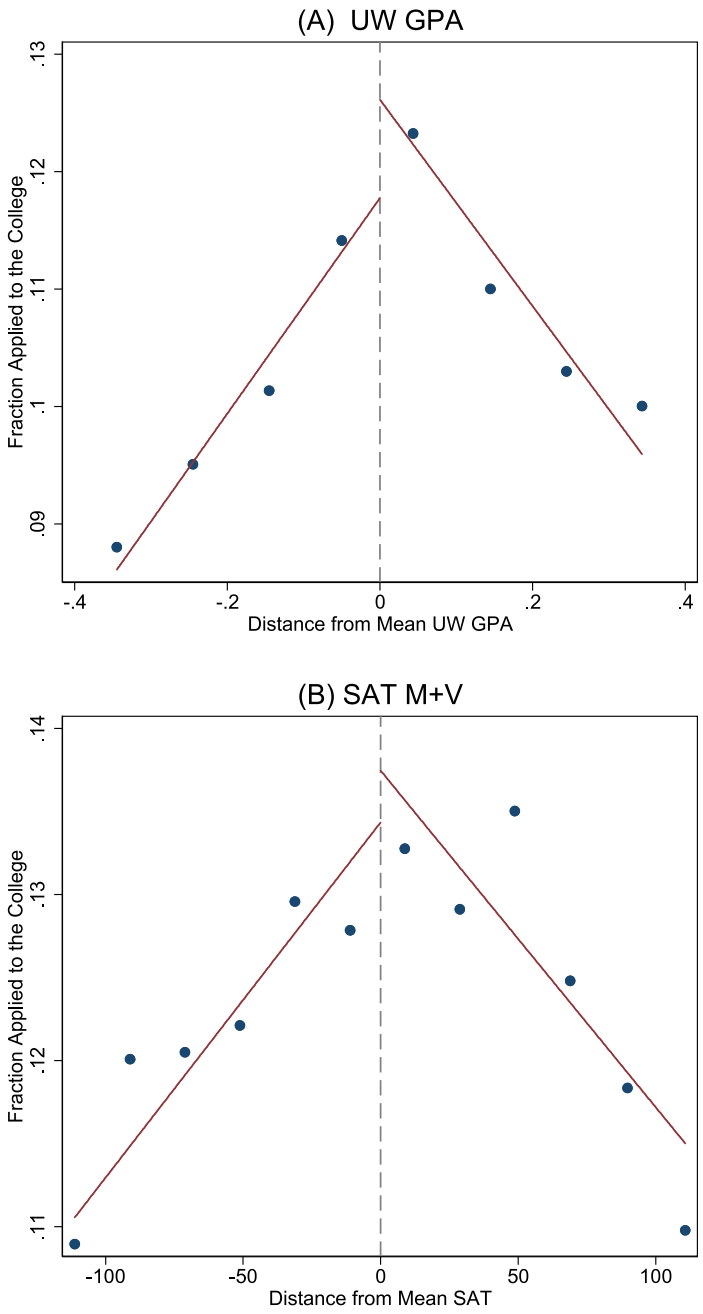
Note: This figure shows how the probability of a student being accepted to a college, conditional on applying, changes when a college crosses a scattergram visibility threshold. A college's scattergram becomes visible to students after it receives five or ten applications from the student's high school. (I do not know which threshold each high school uses.) The X-axis shows how far a college was, in terms of applications, from each of these minimum applicant thresholds (in 2015 and 2016). Since I use both thresholds, college-high school combinations with 5 to 8 applications in the previous year are included twice in this graph for the same student. Observations are student-college-threshold combinations. The dots on the y-axis represents the fraction of students who were accepted to the college, conditional on applying. The fitted lines are from a local linear regression discontinuity model with a bandwidth of 4 applications.

Figure A.15: Impact of Scattergram Visibility on Application Probability by College Type



Note: This figure compares the fraction of students who attended a college with the distance of the student's weighted 11th grade GPA from the typical acceptee line she could see and the type of college. Observations are student-college combinations, and the college in this pair must have received at least ten previous applications from the student's high school to be included in this graph. The data are binned in intervals of 0.1 from the threshold at zero. The fitted lines come from a local linear regression discontinuity model with a bandwidth of 0.5. Colleges are broken into highly selective and less selective categories based on Barron's selectivity ratings. The in-state public colleges are excluded from the selectivity groups so that each student-college combination appears at most once in this figure.

Figure A.16: Application Probabilities by Distance from Other Lines



Note: The figures above show how attendance rates varied based on a student's position on a scattergram relative to the typical acceptee's GPA (in panel A) and SAT (in panel B). For panel (A), unweighted GPAs from 11th grade are used to determine the distance from the mean unweighted GPA line depicted on the scattergram when the student is in 12th grade. The data are binned in intervals of 0.1 from the threshold at zero. For panel (B), students' maximum SAT scores on the old 1600 scale are used. The data are binned in intervals of 20 from the threshold at zero. The fitted lines come from a local linear regression discontinuity model with a bandwidth of 0.5 for panel (A) and a bandwidth of 100 for panel (B). The y-axis represents the fraction of students in each bin who attended the college (in 2015 or 2016). A bin includes multiple scattergrams (and colleges) and it may include the same students multiple times (but for different scattergrams). Observations are student-scattergram combinations (where scattergrams are based on colleges which received at least 10 prior applications).

Table A.1: Additional Summary Statistics

		Free/Reduced Lunch		White or	Black or	Weighted By Scattergram Obs.	
	All	Never	Yes	Asian	Hispanic	GPA BW	SAT BW
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(A) Demographics							
White	0.49	0.58	0.16	0.75	0.00	0.57	0.57
Asian	0.17	0.17	0.16	0.25	0.00	0.22	0.22
Black	0.20	0.14	0.45	0.00	0.72	0.11	0.11
Hispanic	0.08	0.05	0.18	0.00	0.28	0.05	0.05
Free/Reduced Lunch	0.21	0.00	1.00	0.10	0.47	0.10	0.09
(B) Academics							
GPA (11th gr. weighted)	3.41	3.58	2.79	3.64	2.88	3.97	3.89
SAT(M+V+W)	1689	1740	1444	1765	1477	1821	1836
Attend 4-yr Coll	0.60	0.67	0.32	0.68	0.42	0.81	0.82
Attend 2-yr Coll	0.24	0.20	0.38	0.21	0.32	0.11	0.11
Persist in Coll	0.78	0.83	0.60	0.84	0.65	0.90	0.91
(C) Applications							
Number of Apps	5.15	5.50	3.89	5.41	4.60	6.51	6.70
Num Reach Apps	1.53	1.40	2.10	1.23	2.34	1.40	1.26
Num Match Apps	2.31	2.53	1.38	2.55	1.73	3.01	3.25
Num Safety Apps	1.29	1.48	0.50	1.55	0.63	1.84	1.91
Highly Selective	3.98	4.30	2.63	4.33	3.07	5.23	5.25
Acceptances	2.51	2.78	1.53	2.77	1.92	3.36	3.43
(D) Attendance							
Reach	0.19	0.17	0.31	0.15	0.33	0.26	0.09
Match	0.54	0.54	0.54	0.55	0.51	0.43	0.63
Safety	0.27	0.29	0.16	0.30	0.16	0.31	0.29
Highly Selective	0.56	0.60	0.32	0.62	0.37	0.71	0.69
(E) Scattergrams							
Total	47	50	38	51	39	59	59
In GPA Bandwidth	18	20	8	21	10	32	30
In SAT Bandwidth	11	13	5	13	6	19	22
Relevant	21	23	10	25	12	36	35
N	7,647	6,004	1,643	5,005	2,156	134,023	85,451

Note: Column 1 contains the full sample of students. They all appear in the scattergram introduction regressions. The number of times they appear depends on the number of colleges which received an application from their high school between 2014 and 2017 and how many of these colleges fell within the bandwidth. Column (2) contains all students who never received free or reduced-price lunch while enrolled in the district, while column (3) contains students who received it. Students who indicate two or more races or report a race that is not white, Black, Asian, or Hispanic are excluded from columns (4) and (5). In columns (6) and (7) there is one observation for each student-scattergram combination for which the student the student is near the GPA line (column 6) or SAT line (column 7). I define near to the GPA line as students' whose weighted GPAs are within .5 GPA points of the typical acceptee's weighted GPA. I define near to the SAT line as students' whose SAT scores (on the 2400 scale) are within 150 points of the typical acceptee's SAT score. Free/reduced lunch is an indicator for students who ever received free or reduced-price lunch while enrolled in the district. Students who indicate two or more races are excluded from the race categories in Panel (A). GPA refers to 11th grade weighted GPA and SAT refers to the maximum SAT on the old 2400 scale. New SAT scores have been converted to the old 2400 scale. Scattergrams refers to the minimum number of scattergrams to which a student had access based on her graduation year and high school. It is the number of colleges with at least 10 prior applicants. If a college was using the minimum of five applicants, more scattergrams would have been visible. Attend 4-yr college is an indicator for whether the student attended a four-year college within six months of graduating high school. Attend 2-yr is similarly defined but for two-year colleges. Reach schools are colleges at which the student's maximum SAT score is below the 25th percentile of accepted students' SATs, as reported to IPEDS in 2015 by the college. Match schools are colleges at which the student's maximum SAT score is within the interquartile range of accepted students' SATs. Safety schools are colleges at which the student's maximum SAT score is above the 75th percentile of accepted students' SATs. Selectivity ratings are based on Barron's 2009 selectivity index. Where this is missing, selectivity rankings from IPEDS in 2002 are used.

Table A.2: Impact with One Observation per Student-College and Randomly Selected Threshold

		Near GPA		Near SAT		Near Both	Near
	All	.5	.1	150	50	.5 & 150	Neither
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(A) Applied							
Visible	0.0031***	0.0045***	0.0061**	0.0049***	0.0061***	0.0069***	0.0025***
	(0.0004)	(0.0010)	(0.0024)	(0.0013)	(0.0022)	(0.0016)	(0.0004)
CCM	0.0136	0.0228	0.0271	0.0246	0.0264	0.0272	0.0083
(B) Attended							
Visible	0.0002	0.0005	-0.0005	0.0002	0.0005	0.0003	0.0001
	(0.0001)	(0.0003)	(0.0007)	(0.0004)	(0.0007)	(0.0005)	(0.0001)
CCM	0.0011	0.0023	0.0030	0.0027	0.0030	0.0033	0.0004
N	2,394,811	608,959	120,755	394,676	140,159	247,144	1,638,275

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). All regressions include fixed effects for each student and college by year. Observations are all student-college combinations for which the college was within four applications of the threshold at five or ten and the college received at least one application from the student's high school between 2014 and 2016. When a student-college combination is in the bandwidth for two thresholds, I randomly select one threshold observation to keep. This happens for colleges with five to eight prior applications since they fall in the bandwidth for both thresholds. W GPA refers to weighted GPAs, which are on a five point scale, and these SAT scores are on the (old) 2400 scale. CCM refers to the mean application or attendance probability predicted at a college at the threshold if its scattergram had not been made visible.

Table A.3: Bandwidth and Fixed Effects Comparisons for Scattergram Impacts

	Main	SEs	Fixed Effects			Bandwidths			Triangular
	FE: Stu, CollxYr BW: 4	Kolesar & Rothe	None	Coll HS Yr	Coll Student	0-20	1-14	+/- 3	Kernel
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(A) Applied									
Visible	0.0027***	0.0027*	0.0029***	0.0028***	0.0028***	0.0019***	0.0018***	0.0027***	0.0031 ***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0004)	(0.0005)	(0.0006)
Visible & Near Lines	0.0056***	0.0056*	0.0080***	0.0068***	0.0058***	0.0022	0.0040**	0.0071***	0.0089***
	(0.0017)	(0.0014)	(0.0015)	(0.0016)	(0.0016)	(0.0014)	(0.0016)	(0.0020)	(0.0019)
(B) Attended									
Visible	0.0001	0.0002	0.0001	0.0002	0.0002	0.0002**	0.0001	0.0001	0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Visible & Near Lines	0.0001	0.0010**	0.0010**	0.0005	0.0005	0.0003	0.0004	0.0002	0.0011*
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0006)	(0.0006)
N	2,565,375	2,565,375	2,565,375	2,565,375	2,565,375	4,513,998	2,660,037	1,236,352	4,521,645
N Near Lines	272,995	272,995	273,097	273,096	272,998	294,843	287,923	166,597	294,945

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). Observations are student-college-threshold combinations for all colleges which received at least one application from the student's high school between 2014 and 2016. Near lines is defined as within .5 GPA points of the weighted GPA line and 150 SAT points of the SAT M+V+W line. All regressions in columns (1)-(5) are based on a bandwidth 4 applications. All regressions in columns (1)-(2) and (6)-(8) include student and college by year fixed effects. For columns (6) - (8), a college is in the bandwidth (x) if the number of applications it received in the prior years is in the noted range. Column (9) contains the result from a triangular kernel specification with a bandwidth of 4.

Table A.4: Impact of Scattergrams by Proximity to Other Typical Acceptee Lines

		Near UW GPA		Near SAT M+V		Near Both	Near
	All	.5	.1	150	50	.5 & 150	Neither
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(A) Applied							
Visible	0.0027***	0.0035***	0.0055***	0.0025**	0.0024	0.0030**	0.0024***
	(0.0004)	(0.0008)	(0.0020)	(0.0011)	(0.0018)	(0.0013)	(0.0005)
CCM	0.0137	0.0193	0.0249	0.0240	0.0268	0.0261	0.0080
(B) Attended							
Visible	0.0001	0.0002	0.0004	0.0003	-0.0001	0.0005	0.0001
	(0.0001)	(0.0003)	(0.0006)	(0.0003)	(0.0006)	(0.0004)	(0.0001)
CCM	0.0011	0.0020	0.0024	0.0024	0.0029	0.0027	0.0003
N	2,565,375	908,296	191,989	623,977	227,209	436,252	1,469,312

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). All regressions include fixed effects for each student and college by year. Observations are all student-college-threshold combinations for which the college was within four applications of the threshold at five or ten and the college received at least one application from the student's high school between 2014 and 2016. Student-college combinations are included twice for colleges with five to eight prior applications since they fall in the bandwidth for both thresholds. UW GPA refers to unweighted GPAs, which are on a four point scale, and these SAT scores are on the (old) 1600 scale. CCM refers to the mean application or attendance probability predicted at a college at the threshold if its scattergram had not been made visible.

Table A.5: Scattergram Impacts on Applications and Attendance by Distance to GPA and SAT

	W GPA				SAT M+V+W		
	All	BW 1	BW .5	BW .1	BW 300	BW 150	BW 50
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(A) Apply							
Visible	0.0027***	0.0034***	0.0040***	0.0060**	0.0032***	0.0038***	0.0048**
	(0.0004)	(0.0007)	(0.0010)	(0.0025)	(0.0009)	(0.0013)	(0.0023)
Visible & Above	0.0021**	0.0024**	0.0019	0.0070**	0.0032**	0.0035*	0.0038
	(0.0009)	(0.0011)	(0.0015)	(0.0035)	(0.0013)	(0.0018)	(0.0032)
Visible & Below	0.0022***	0.0035***	0.0053***	0.0052	0.0026**	0.0037**	0.0071**
	(0.0005)	(0.0009)	(0.0014)	(0.0037)	(0.0013)	(0.0018)	(0.0033)
CCM	0.0158	0.0183	0.0237	0.0268	0.0213	0.0242	0.0262
(B) Attend							
Visible	0.0001	0.0003	0.0004	-0.0005	0.0003	0.0000	0.0003
	(0.0001)	(0.0002)	(0.0003)	(0.0007)	(0.0003)	(0.0004)	(0.0007)
Visible & Above	0.0000	0.0000	-0.0003	-0.0007	-0.0002	-0.0007	0.0010
	(0.0003)	(0.0003)	(0.0005)	(0.0010)	(0.0004)	(0.0006)	(0.0009)
Visible & Below	0.0001	0.0005*	0.0011**	-0.0007	0.0007*	0.0007	-0.0004
	(0.0001)	(0.0003)	(0.0004)	(0.0012)	(0.0004)	(0.0006)	(0.0012)
CCM	0.0019	0.0022	0.0031	0.0030	0.0025	0.0033	0.0023
N	2,565,375	1,167,906	666,731	132,319	800,508	432,073	153,384
N Above	671,907	525,023	320,169	68,653	380,179	214,080	80,384
N Below	1,186,729	642,597	346,223	62,910	420,202	217,793	72,589

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). All regressions include fixed effects for each student and college by year. Observations are all student-college-threshold combinations for which the college was within four applications of the threshold at five or ten and the college received at least one application from the student's high school between 2014 and 2016. Student-college combinations are included twice for colleges with five to eight prior applications since they fall in the bandwidth for both thresholds. CCM refers to the mean application or attendance probability predicted at a college at the threshold if its scattergram had not been made visible. Near is defined as within .5 GPA points or 150 SAT points (based on weighted GPAs and SAT scores on the (old) 2400 scale.

Table A.6: Average Treatment Effects of Scattergram Visibility on Attendance

		Free/Reduced Lunch		White or	Black or	In-St. Public	Other Colleges	
	All	Never	Ever	Asian	Hispanic	Colleges	High Sel.	Less Sel.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(A) All Students								
Visible	0.0007***	0.0007***	0.0002*	0.0005***	0.0006***	0.0042***	0.0004***	0.0008***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0009)	(0.0001)	(0.0001)
CCM	0.0007	0.0008	0.0004	0.0007	0.0006	0.0016	0.0007	0.0007
N	8,914,720	7,018,780	1,895,940	5,844,978	2,503,108	300,304	2,939,362	5,062,656
(B) Near Lines								
Visible	0.0011***	0.0008***	0.0026**	0.0006**	0.0027***	-0.0004	0.0009**	0.0012***
	(0.0003)	(0.0003)	(0.0010)	(0.0003)	(0.0008)	(0.0067)	(0.0004)	(0.0004)
CCM	0.0018	0.0018	0.0011	0.0017	0.0020	0.0039	0.0019	0.0016
N	583,508	520,768	62,740	451,196	98,194	27,742	233,076	303,676

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). All regressions include fixed effects for each student and college by year. Observations are all student-college-threshold combinations. Panel (A) includes LATE estimates based on the RDD for colleges within four applications of the threshold at five or ten. Panel (B) includes all colleges and relies on college by year and student fixed effects for identification. CCM refers to the mean attendance probability predicted at a college at the threshold if its scattergram had not been made visible. Near is defined as within .5 GPA points or 150 SAT points. This is based on weighted GPA points and SAT points on the old 2400 scale. Column (2) is based on students who never received free or reduced-price lunch from the district. Column (3) contains all students who received it at least once while enrolled in the district. Students who indicate two or more races or report a race that is not white, Black, Asian, or Hispanic are excluded from columns (4) and (5). The in-state public colleges are excluded from the highly and less selective college categories in columns (7) and (8). Selectivity ratings are based on Barron's 2009 selectivity index. Where this is missing, selectivity rankings from IPEDS in 2002 are used.

Table A.7: Results using Alternative Definitions of Typical Acceptee Lines

	GPA		SAT			Both GPA (Wtd) & SAT(M+V+W)		
	Weighted	Unweighted	M+V+W (2400)	M+V (1600)	Any Line	Below Both	Below at Least One	Below Just One
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(A) Applied								
Below Line	-0.0107***	-0.0103***	0.0040	-0.0051	-0.0049	0.0022	0.0029	-0.0092***
	(0.0035)	(0.0034)	(0.0038)	(0.0037)	(0.0041)	(0.0042)	(0.0027)	(0.0027)
Dist Above	-0.1034***	-0.1641***	-0.0002***	-0.0004***				
	(0.0106)	(0.0157)	(0.0000)	(0.0000)				
Dist Below	-0.0846***	-0.0349**	-0.0001***	-0.0000				
	(0.0098)	(0.0156)	(0.0000)	(0.0000)				
N	123,429	131,271	101,188	98,970	71,342	71,342	71,342	192,382
(B) Attended								
Below Line	0.0003	-0.0005	0.0016	-0.0029	-0.0015	0.0016	0.0012	0.0007
	(0.0016)	(0.0014)	(0.0018)	(0.0019)	(0.0018)	(0.0018)	(0.0011)	(0.0011)
Dist Above	-0.0091*	-0.0140**	-0.0001***	-0.0001***				
	(0.0048)	(0.0056)	(0.0000)	(0.0000)				
Dist Below	-0.0246***	-0.0129**	-0.0000**	0.0000				
	(0.0042)	(0.0052)	(0.0000)	(0.0000)				
N	123,429	131,271	101,188	98,970	71,342	71,342	71,342	192,382

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). College by year and high school fixed effects are included. The optimal bandwidths are calculated as described in Calonico, Cattaneo and Titiunik (2014). New SAT scores have been converted to the old scale. All columns include controls for 11th grade GPA, maximum SAT score, gender, special education and dummy variables for race and ever receiving free or reduced-price lunch. Column (5) compares students who are below all the lines (weighted, unweighted, SAT M+V and SAT M+V+W) to students who are above at least one line. Column (6) compares students below both the weighted GPA and SAT line to students who are above at least one line. Column (7) compares students who are below the weighted GPA, SAT M+V+W line or both, to students who are above both lines. Column (8) compares students who are below the weighted GPA or SAT M+V+W line (but not both), to students who are above both lines. N refers to the number of student-scattergram combinations on which the regression is based.

Table A.8: Results for Alternative Specifications Around GPA Line

	Main	GPA Bandwidth		Donut RD		Alternate Specifications		Other Fixed Effects		
		0.4	0.6	+/- .05	+/- .1	Quadr Dist	Triangular Kernel	None No Controls	Student	Scattergram
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(A) Applied										
Below GPA	-0.011***	-0.008**	-0.010***	-0.012***	-0.016***	-0.005	-0.012***	-0.013***	-0.006*	-0.012***
	(0.004)	(0.004)	(0.003)	(0.004)	(0.005)	(0.005)	(0.004))	(0.004)	(0.004)	(0.003)
Dist Above GPA	-0.103***	-0.094***	-0.109***	-0.102***	-0.108***	-0.098***		-0.096***	-0.119***	-0.086***
	(0.011)	(0.014)	(0.009)	(0.012)	(0.014)	(0.035)		(0.010)	(0.010)	(0.009)
Dist Below GPA	-0.085***	-0.088***	-0.093***	-0.081***	-0.079***	-0.143***		-0.080***	-0.089***	-0.086***
	(0.010)	(0.013)	(0.008)	(0.011)	(0.012)	(0.035)		(0.009)	(0.009)	(0.009)
(B) Attended										
Below GPA	0.000	-0.008**	-0.010***	0.001	-0.001	0.000	0.000	0.000	0.001	0.000
	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Dist Above GPA	-0.009*	-0.094***	-0.109***	-0.007	-0.008	-0.026		-0.001	0.002	-0.015***
	(0.005)	(0.014)	(0.009)	(0.006)	(0.007)	(0.017)		(0.005)	(0.005)	(0.004)
Dist Below GPA	-0.025***	-0.088***	-0.093***	-0.025***	-0.023***	-0.040**		-0.019***	-0.043***	-0.016***
	(0.004)	(0.013)	(0.008)	(0.005)	(0.005)	(0.016)		(0.004)	(0.004)	(0.004)
N	123,429	100,729	144,229	111,233	97,703	123,429	345,548	131,704	131,437	131,704

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). The main regression, is in column (1). It is based on a bandwidth of .5 GPA points, college by year and high school fixed effects, as well as controls. Controls include 11th grade GPA, maximum SAT score, gender, special education, and dummy variables for race and ever receiving free or reduced-price lunch. All estimates are for weighted GPAs in 11th grade. N refers to the number of student-scattergram combinations on which the regression is based. The specifications in columns (1-7) include college by year and high school fixed effects and controls for student characteristics. The donut RD columns (4 and 5) exclude student observations in which the student is within .05 or .1 GPA points of the GPA line. A quadratic term is added for GPA distance in column (6). Control variables are excluded from columns (7)-(10).

Table A.9: Results for Alternative Specifications Around SAT Line

	Main	SAT Bandwidth		Donut RD		Alternate Specifications		Other Fixed Effects		
		0.4	0.6	+/- 10	+/- 20	Quadr Dist	Triangular Kernel	None No Controls	Student	Scattergram
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(A) Applied										
Below SAT	0.0040	0.0050	0.0033	0.0043	0.0002	0.0071	-0.006	0.0042	0.0050	0.0045
	(0.0038)	(0.0052)	(0.0036)	(0.0047)	(0.0055)	(0.0058)	(0.005)	(0.0041)	(0.0041)	(0.0038)
Dist Above SAT	-0.0002***	-0.0001**	-0.0002***	-0.0002***	-0.0002***	-0.0001		-0.0001***	-0.0002***	-0.0001***
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)		(0.0000)	(0.0000)	(0.0000)
Dist Below SAT	-0.0001***	-0.0001	-0.0001***	-0.0001***	-0.0001**	-0.0001		-0.0002***	-0.0002***	-0.0002***
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)		(0.0000)	(0.0000)	(0.0000)
(B) Attended										
Below SAT	0.0016	-0.0000	0.0025	0.0018	0.0007	0.0005	-0.002	0.0019	0.0027	0.0019
	(0.0018)	(0.0024)	(0.0018)	(0.0023)	(0.0026)	(0.0027)	(0.002)	(0.0020)	(0.0020)	(0.0018)
Dist Above SAT	-0.0001***	-0.0001*	-0.0000***	-0.0000**	-0.0000**	-0.0000		-0.0000***	-0.0000***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)		(0.0000)	(0.0000)	(0.0000)
Dist Below SAT	-0.0000**	0.0000	-0.0000***	-0.0000	-0.0000	-0.0000		-0.0000**	-0.0001***	-0.0000*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)		(0.0000)	(0.0000)	(0.0000)
N	101,188	57,700	111,554	80,245	74,474	101,188	291,959	101,188	101,004	101,188

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). The main regression, is in column (1). It is based on a bandwidth of 150, college by year and high school fixed effects, as well as controls. Controls include 11th grade GPA, maximum SAT score, gender, special education, and dummy variables for race and ever receiving free or reduced-price lunch. All estimates are for SAT scores on the 2400 scale. New scores have been converted to the old scale. N refers to the number of student-scattergram combinations on which the regression is based. The specifications in columns (1-7) include college by year and high school fixed effects and controls for student characteristics. The donut RD columns (4 and 5) exclude student observations in which the student is within 10 or 20 points of the SAT line. A quadratic term is added for SAT distance in column (6). Control variables are excluded from columns (7)-(10).

Table A.10: Impact of Mean Lines on Applications by Year and Logins

	2015-2016	2015-2017	2015	2016	2017		
					All	Many Logins	Few Logins
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(A) GPA Line on Apps							
Below GPA	-0.011***	-0.013***	-0.014**	-0.008*	-0.016***	-0.017***	-0.008*
	(0.004)	(0.003)	(0.006)	(0.004)	(0.004)	(0.005)	(0.005)
Dist Above GPA	-0.103***	-0.112***	-0.123***	-0.094***	-0.125***	-0.148***	-0.052***
	(0.011)	(0.008)	(0.018)	(0.012)	(0.011)	(0.013)	(0.014)
Dist Below GPA	-0.085***	-0.074***	-0.093***	-0.082***	-0.051***	-0.065***	-0.022**
	(0.010)	(0.008)	(0.017)	(0.011)	(0.009)	(0.013)	(0.011)
N	123,429	216,655	45,624	84,059	11,2872	77,792	39,406
(B) SAT Line on Apps							
Below SAT	0.0040	-0.0021	-0.0025	0.0092*	-0.0106***	-0.0102**	-0.0099*
	(0.0038)	(0.0025)	(0.0068)	(0.0051)	(0.0036)	(0.0046)	(0.0056)
Dist Above SAT	-0.0002***	-0.0003***	-0.0002***	-0.0001***	-0.0004***	-0.0005***	-0.0003***
	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Dist Below SAT	-0.0001***	-0.0001***	-0.0000	-0.0002***	-0.0000	-0.0001	0.0000
	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0001)
N	101,188	222,325	36,949	52,259	90,598	65,364	26,950

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). College by year and high school fixed effects are included, as well as controls for 11th grade GPA, maximum SAT score, gender, special education, and dummy variables for race and ever receiving free or reduced-price lunch. The optimal bandwidths, as described in Calonico, Cattaneo and Titiunik (2014), are calculated for each regression. All estimates are for weighted GPAs and SAT scores on the old 2400 scale. New scores have been converted to the old scale. The outcome is applying to the college associated with the scattergram treating the student. N refers to the number of student-scattergram combinations on which the regression is based. Column (1) shows the main results which are based on students who graduated in 2015 and 2016. Login records are only available for students who graduated in 2017. Column (6) is based on students who were in the top 50% in terms of Naviance logins. Column (7) is based on students in the bottom 50%. Students who logged onto Naviance more were in the bandwidth for more scattergrams, which is why the N in column (6) is much larger than the N in column (7).

Table A.11: Cumulative Impact of Scattergrams for Minority and Low-Income Students

	Applications				Acceptances	Attend College				
	Num.	Reach	Match	Safety		Reach	Match	Safety	Four-yr	Persist
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total SGs	-0.007	-0.026**	0.039***	-0.015**	0.010	-0.004**	0.008***	-0.004***	-0.000	0.005
	(0.019)	(0.013)	(0.010)	(0.006)	(0.010)	(0.002)	(0.002)	(0.001)	(0.002)	(0.004)
Reach SGs	0.002	0.143***	-0.003	-0.130***	0.023	0.023***	-0.009	-0.020***	-0.006	-0.000
	(0.080)	(0.054)	(0.043)	(0.026)	(0.044)	(0.008)	(0.010)	(0.006)	(0.009)	(0.015)
Match SGs	-0.015	-0.036**	0.056***	-0.029***	0.015	-0.004	0.010***	-0.006***	0.001	0.010*
	(0.027)	(0.018)	(0.014)	(0.009)	(0.015)	(0.003)	(0.003)	(0.002)	(0.003)	(0.006)
Safety SGs	-0.019	-0.113***	0.056**	0.042***	-0.001	-0.022***	0.018***	-0.001	-0.004	-0.005
	(0.048)	(0.032)	(0.026)	(0.016)	(0.026)	(0.004)	(0.006)	(0.004)	(0.005)	(0.010)
In-St Public SGs	0.053	-0.049	0.163***	-0.045	0.030	-0.014	0.043***	-0.010	0.023**	0.005
	(0.103)	(0.070)	(0.055)	(0.034)	(0.056)	(0.010)	(0.012)	(0.008)	(0.012)	(0.017)
N	1,409	1,409	1,409	1,409	1,409	1,409	1,409	1,409	1,409	623

Note: Heteroskedasticity robust standard errors are in parentheses. (*p<.10 **p<.05 *** p<.01). High school and year fixed effects are included. I control for academic achievement using fixed effects for 50 point intervals of maximum SAT scores, and .1 point intervals of students' weighted 11th grade GPAs. Controls include demographic indicators for race (white, asian, black or hispanic), free-or-reduced price lunch, special education, and gender. There is one observation per student. Persistence refers to persistence into a second year of college. These data are only available for students who graduated high school in 2015. Reach schools are colleges at which the student's maximum SAT score is below the 25th percentile of accepted students' SATs, as reported in IPEDS in 2015. Match schools are colleges at which the student's maximum SAT score is within the interquartile range of accepted students' SATs. Safety schools are colleges at which the student's maximum SAT score is above the 75th percentile of accepted students' SATs. Acceptances are self-reported but I corrected the self reports if a student attended a college where an acceptance decision was not reported. I assume a student must have been accepted to a college if she attends the college.

Table A.12: Scattergram Impacts by Year

	2015	2016	2017	2015-2016	2015-2017
	(1)	(2)	(3)	(4)	(5)
(A) Applied					
Visible	0.0037***	0.0026***	0.0009***	0.0031***	0.0022***
	(0.0008)	(0.0005)	(0.0003)	(0.0004)	(0.0003)
Visible & Near Lines	0.0068**	0.0067***	0.0018*	0.0067***	0.0041***
	(0.0031)	(0.0019)	(0.0010)	(0.0017)	(0.0010)
(B) Attended					
Visible	0.0002	0.0003**	0.0003**		
	(0.0002)	(0.0001)	(0.0001)		
Visible & Near Lines	0.0005	0.0006	0.0006		
	(0.0010)	(0.0006)	(0.0005)		
N	1,021,258	1,427,266	1,603,743	2,448,524	4,052,267
N Near Lines	105,559	149,987	246,606	255,546	502,152

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). All regressions include fixed effects for each student and college by year. Observations are all student-college-threshold combinations for which the college received at least one application from the student's high school between 2014 and 2017. Near is defined as within .5 GPA points or 150 SAT points. This is based on weighted GPA points and SAT points on the old 2400 scale. All regressions are based on a bandwidth of 4 applications.

Table A.13: Balance Table for Colleges with and without Scattergrams

	Private (1)	Out-of-State Public (2)	In-State Public (3)	Selectivity Tier (4)	Net Price (5)	Enrollment (6)
Visible	0.0309	-0.0068	-0.0529	3.4982	-245.2943	-2.2939**
	(0.0357)	(0.0100)	(0.0369)	(7.5569)	(620.4679)	(1.0809)
N	7,956	7,956	7,485	7,485	7,638	7,485

Note: Standard errors clustered by high school and year (combinations) are in parentheses. (*p<.10 **p<.05 *** p<.01). All colleges within four applications of the thresholds at five or ten are included. Selectivity Tier refers to Barron's rankings in 2009. Net Price and Enrollment numbers come from Ipeds in 2015.

Table A.14: Balance Table for Students Above and Below Typical Acceptee Lines

	Below GPA (1)	Below SAT (2)
White/Asian	-0.003	0.001
	(0.004)	(0.004)
Black/Hispanic	0.005	-0.001
	(0.004)	(0.004)
Female	0.004	0.002
	(0.006)	(0.006)
Free or Reduced-Price Lunch	-0.002	0.005
	(0.003)	(0.003)
Special Ed	0.004**	0.000
	(0.002)	(0.001)
N	131,704	101,188

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). Estimates are from a regression of the indicator for being below the typical acceptee line (for a particular scattergram) on the demographic variable and the distance of one's GPA or SAT from the line. High school and college by year fixed effects are included. The bandwidths are .5 GPA points and 150 SAT points. All estimates are for weighted GPAs and SAT scores on the 2400 scale. New SAT scores have been converted to the old scale. N refers to the number of student-scattergram combinations on which the regression is based.

Table A.15: Impact of Mean SAT and GPA on Attendance

	All	White or Asian	Black or Hispanic	In-St. Public Colleges	Other Colleges	
					High Sel.	Less Sel.
	(1)	(2)	(3)	(4)	(5)	(6)
(A) GPA						
Below GPA	0.000	-0.001	0.001	0.007	-0.000	0.002
	(0.002)	(0.002)	(0.004)	(0.009)	(0.002)	(0.002)
Dist Above GPA	-0.011*	-0.016**	-0.019**	-0.010	-0.001	-0.001
	(0.006)	(0.006)	(0.009)	(0.026)	(0.009)	(0.003)
Dist Below GPA	-0.025***	-0.022***	-0.024***	-0.111***	-0.009	-0.006
	(0.005)	(0.006)	(0.007)	(0.029)	(0.007)	(0.004)
CCM	0.026	0.026	0.028	0.113	0.010	0.005
N	110,013	85,875	26,522	19,081	43,719	39,620
(B) SAT						
Below SAT	0.0013	0.0031*	0.0083*	0.0088	-0.0013	0.0037*
	(0.0019)	(0.0018)	(0.0048)	(0.0095)	(0.0016)	(0.0019)
Dist Above SAT	-0.0001***	-0.0001***	-0.0000	-0.0002**	-0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)
Dist Below SAT	-0.0000*	-0.0001***	-0.0000	-0.0001	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)
CCM	0.025	0.027	0.019	0.111	0.007	0.005
N	97,226	105,567	16,012	15,082	49,691	29,159

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). College by year and high school fixed effects are included, as well as controls for 11th grade GPA, maximum SAT score, gender, special education, and dummy variables for race and ever receiving free or reduced-price lunch. Optimal bandwidths for each regression are calculated as described in Calonico, Cattaneo and Titiunik (2014). All estimates are for weighted GPAs, which are on a five point scale, in panel (A) and SAT scores on the 2400 scale in panel (B). New scores have been converted to old ones where relevant. The outcome is attending the college with the relevant scattergram. N refers to the number of student-scattergram combinations on which the regression is based. Students who indicate two or more races or report a race that is not white, Black, Asian, or Hispanic are excluded from columns (2) and (3). The in-state public colleges are excluded from the highly and less selective college categories in columns (5) and (6). Selectivity ratings are based on Barron's 2009 selectivity index. Where this is missing, selectivity rankings from IPEDS in 2002 are used. CCM refers to the mean attendance probability for students with GPAs or SATs just above the typical acceptee's.

Table A.16: Bandwidth Comparisons for Subgroup Results

		Free/Reduced Lunch		White or	Black or	In-St. Public	Other Colleges	
	All	Never	Ever	Asian	Hispanic	Colleges	High Sel.	Less Sel.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(A) BW: +/- 4								
Visible	0.003***	0.003***	0.002*	0.003***	0.003***	0.013***	0.003***	0.002***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)
N	2,394,811	1,895,046	499,765	1,581,564	663,378	58,730	892,047	1,444,034
(B) BW: +/- 3								
Visible	0.003***	0.003***	0.001	0.003***	0.002	0.007**	0.003***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)
N	1,122,462	890,641	231,821	745,508	307,209	34,742	518,596	569,124
(C) BW: 1-14								
Visible	0.003***	0.003***	0.001*	0.003***	0.002***	0.012***	0.003***	0.002***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)
N	2,448,524	1,938,859	509,665	1,618,331	676,992	59,569	928,389	1,460,566
(D) BW: 0 - 20								
Visible	0.003***	0.003***	0.002**	0.003***	0.002***	0.012***	0.002***	0.003***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.002)	(0.001)	(0.001)
N	4,343,434	3,413,033	930,401	2,838,708	1,227,753	103,730	1,410,797	2,828,907

Note: Heteroskedasticity robust standard errors clustered by student are in parentheses. (*p<.10 **p<.05 *** p<.01). All regressions include fixed effects for each student and college by year. Observations are all student-college-threshold combinations for which the college received at least one application from the student's high school between 2014 and 2016. Near is defined as within .5 GPA points or 150 SAT points. This is based on weighted GPA points and SAT points on the old 2400 scale. Panel (A) contains the main results. Panels (B), (C), and (D) contain alternate bandwidths of prior applications. Column (2) is based on students who never received free or reduced-price lunch from the district. Column (3) contains all students who received it at least once while enrolled in the district. Students who indicate two or more races or report a race that is not white, Black, Asian, or Hispanic are excluded from columns (4) and (5). The in-state public colleges are excluded from the highly and less selective college categories. Selectivity ratings are based on Barron's 2009 selectivity index. Where this is missing, selectivity rankings from IPEDS in 2002 are used.

B Additional Details on Naviance and the District's Implementation

Naviance is a popular technology used by schools and counselors to help with the college and career counseling process. It is owned by Hobsons and has versions available for all students in grades K-12. I focus on the features used by the high school students and center my research on the scattergrams. It is important to note that this is just one of the many features available in Naviance, though it is perhaps the most novel feature of the platform.

More than forty percent of students in the U.S. use Naviance and information on some of the districts using it, its implementation and features are available on Naviance's website `https://www.naviance.com/`. Naviance provides a similar implementation package to all schools and districts that purchase it, though there are some options for customizing implementation based on the district's needs. The district I studied believed the guidance they received from Naviance to be fairly typical. This guidance included a set of materials and trainings to help them get started as well as a liason whom they could contact with questions. Additional materials and tutorial videos can be found on Naviance's website. Naviance also provides ongoing professional development, and summaries of updates to Naviance's platform, as well as tips, are periodically emailed to counselors.

Naviance is sold directly to schools, not students, and is intended for use with a school counselor. Thus, as in most districts, counselors in this district were responsible for implementing Naviance. Counselors received some training from Naviance and then they were responsible for providing guidance to students and parents around how to use the platform. The district counseling office also provided some guidance to the schools and students. For instance, the district suggested that schools set up time during the school day for students to register with Naviance and explore its main features, such as the quizzes, scattergrams, and how to save colleges.

Counselors set up information sessions for parents and students, and logged on with students during school hours. They also provided students some specific suggestions around how to use the platform. Some counselors encouraged students to start by taking the quizzes or looking at previously popular colleges.

The district office mentioned some concerns about the typical acceptee scores and how they

were being used. They wanted counselors to help students interpret these, but it was unclear how much this message permeated through the schools. The district did not explicitly encourage or discourage use of the typical acceptee scores in students' application choices. It appears that different counselors provided different suggestions about how to use Naviance. This is consistent with the lack of a district-wide application strategy. (The district was excited about this project because they wanted to understand how counselors and students used Naviance and how it impacted the advice counselors provided students.)

Counselors and students each have their own accounts. Counselors can use their acccounts to send recommendations and nudges to students. They may also look at scattergrams and other student data to help them determine which colleges to recommend to students. Students use their accounts to access Naviance's features, which include exploring and saving colleges. Some districts allow parents to create their own accounts but in many districts, including the one I study, parents do not have their own accounts. They often used their student's account so it is unclear how much the usage rates are driven by students or parents.

Districts have some control over which information students can see on the scattergrams. During the time I study, schools could set the minimum number of applications needed to make a scattergram visible. These minimums can either be set by the district or by individual schools. The threshold a school selects applies to all colleges but it can be changed at any time. The district I studied informed me that they used the minimums of five and ten and I sat down with a district administrator to examine how they could change the settings. This setting no longer exists in the version of Naviance the district currently uses, so I cannot see what rule the schools are currently using. I also lack power to use the data to detect which cutoffs each high school used.

Districts and schools can also choose whether to limit the number of cohorts that populate the scattergram or use all available data. If the data are not limited, the number of available scattergrams and data points will continue to grow over time. During the period I studied, the district did not limit the cohorts for which data were visible.

This district adds data on a graduating class in the June that the cohort graduates so that students have updated information when searching for colleges over the summer. During their

12th grade, the class of 2015 could access data on the class of 2014. The class of 2016 had access to two sets of scattergrams - one during the 2015 school year and another during 2016 school year. Students logged onto Naviance more during 12th grade than 11th grade, and most application decisions are made in 12th grade, so I focus on the 12th grade scattergrams. Students exploring scattergrams during the fall of their 12th grade will see their 11th grade GPA populated on the scattergram and in the college dashboard.

C Details on the Data and Accuracy of Self-Reported Admissions

Application data are based on requests for students' transcripts to be sent to colleges. This measure may inaccurately count too many colleges in a student's application portfolio if a student decides not to complete the application after requesting a transcript. Transcripts cost a few dollars to send, so it is unlikely that this is happening in many cases. It is possible that this approach misses some applications if students send unofficial transcripts. This is probably not a big concern because most colleges request an official transcript and the transcript request in Naviance triggers a request for a counselor recommendation, which is also necessary at many colleges. In 2015-2016, only 10 students did not submit a transcript request for a four-year college they attended (as indicated by the NSC records) within 6 months of graduating high school.

C.1 Self-Reported Admissions Data

Admissions decisions are self-reported in the senior survey. If the self-reports are inaccurate, or if many students do not report their expereinces, the admissions information students see may be biased. Approximately 90% of students respond to the senior survey, however, many students appear to under-report rejections. For this reason, the district treats non-responses as rejections. Fifteen percent of applications are reported to end in rejections and 53% in admissions. Under-reporting of admissions appears to be less of a problem. 10% of students who apply to at least one college report no admissions. This could be because they were admitted nowhere or because they did not respond to the survey. 9% of students who applied to five or more colleges report

no acceptances. This is probably driven by non-response since rather than not being accepted anywhere.

At least 3% of students under-report acceptances. When they complete the survey, 3% of students report plans to attend a college but do not report an acceptance at that college. NSC records indicate that 13% of students attend a college where they did not report an acceptance. Some of the latter discrepancy could be driven by students getting off the waitlist over the summer. In addition, 69% of students' self-reported attendance plans match the NSC records. (I do not use the attendance self-reports in my study.) Districts could use NSC records to update the accuracy of the self-reported data they put into Naviance. This may not be difficult if the district purchases the Naviance feature which links NSC records to student records in Naviance.

Over-reporting of acceptances may also be an issue. 32% of students report acceptances everywhere they apply, but this is largely driven by students who only apply to a few colleges. 5.5% of students who apply to five or more colleges report being accepted everywhere they apply. This may be true or some of it could be driven by students quickly or carelessly responding to the survey.

While missing admissions data may bias the accuracy of the admissions information students see, it will not bias the estimates of the treatment I am studying. Admissions is not one of my main outcomes, and when I use it as an outcome, I correct the admissions self-reports with the attendance self-reports and NSC records. I assume that if a student attends a college she must also have been admitted to it.

C.2 Definitions

I define *safety* colleges as those where the student's SAT is above the 75th percentile of all accepted students, as reported to IPEDS in 2015. Reach colleges are defined as those where the student's SAT is below the 25th percentile. Match colleges are defined as those where a student's SAT score is whithin the inter-quartile range. Elite colleges are the public and private institutions defined as elite by Barron's *Profiles of American Colleges*.

In section 5 I briefly mention medium popular colleges. Medium popular colleges are defined

as those where where 5 to 20% of the high school students applied in 2014. Application rates are constant for the most popular colleges, and the least popular colleges experience an increase in applications, likely due to increased awareness from the scattergrams. This seems reasonable if we expect students to always apply to the local public colleges but to substitute across out of state or private colleges based on how much they know about them and their admissions probability. This is consistent with scattergrams broadening the set of schools to which students apply.

I also briefly examine a college's competitors in section 5. College X's closest competitor is defined as the college Y which is most popular among students who applied to college X, relative to its average popularity in the sample.

C.3 A Note on Recreating Scattergrams

Schools have several options for the orientation of the scattergrams, so one needs to confirm the decision rules used by the district or school before using the raw data to reconstruct scattergrams outside of Naviance. Some of these decision rules include the minimum number of applications needed to make it visible, the GPA scale and the number of cohorts with visible data. I visually confirmed that my identification of the typical acceptee lines matched what students observed in 2017 for twenty colleges at one high school. I focus on the version of the scattergram students could see in the fall of their 12th grade. At this time, students could see how their 11th grade GPA compared to the average admit's.

D Additional Robustness Checks

I find no evidence of manipulation of the running variables in any of the regression discontinuity designs I use. Figure A.9 shows no spike in the density of observations with exactly five or ten prior applications. In addition, Table A.13 shows that the colleges just above and below these thresholds are similar in terms of their sector, location, selectivity, price and enrollment.

I find no evidence that students manipulate their SATs to be just above the typical acceptee's (Figure A.10), despite evidence that, in other settings, students retake the SAT until their score is

equal to or above minimum admissions thresholds (Goodman, Hurwitz, & Smith, 2017; Goodman, Gurantz, & Smith, forthcoming). The density of SAT scores and the distance of SAT scores from the lines are smooth, as per McCrary (2008). GPAs are typically more difficult to manipulate because they are a combination of many grades over a multi-year period and they are calculated to two decimal points. Manipulation does not appear to be a problem for weighted GPAs (Figure A.10). The density of weighted GPAs, and their distance from the typical acceptee's weighted GPA, are also smooth as per McCrary (2008). Since students view the relevant line after their 11th grade GPA is fixed they cannot manipulate it. While some could view a college's scattergram in 11th grade, the location of the line could have been different so this should not lead to heaping around the 12th grade line. Table A.14 shows no significant differences in observable characteristics for students just above and below the typical acceptee's GPA or SAT.