A. Empirical Appendix (Online Only)

Table A1: Application Score Components

Year	Total Score	Language Arts	Mathematics	Grades	Attendance	Leadership	Pers. Statemnt
2006							
Max Score (Weight)	100	20 (0.20)	20 (0.20)	40 (0.40)	20 (0.20)	0 (0.00)	0 (0.00)
2007							
Max Score (Weight)	100	21 (0.21)	21 (0.21)	48 (0.48)	10 (0.10)	0 (0.00)	0 (0.00)
2008							
Max Score (Weight)	100	21 (0.21)	21 (0.21)	48 (0.48)	10 (0.10)	0 (0.00)	0 (0.00)
2009							
Max Score (Weight)	106	21 (0.21)	21 (0.21)	48 (0.48)	10 (0.10)	3 (0.03)	3 (0.03)
2010							
Max Score (Weight)	106	21 (0.20)	21 (0.20)	48 (0.45)	10 (0.09)	3 (0.03)	3 (0.03)
2011							
Max Score (Weight)	120	21 (0.18)	21 (0.18)	48 (0.40)	10 (0.08)	10 (0.08)	10 (0.08)
2012							
Max Score (Weight)	120	21 (0.18)	21 (0.18)	48 (0.40)	10 (0.08)	10 (0.08)	10 (0.08)
2013							
Max Score (Weight)	120	21 (0.18)	21 (0.18)	48 (0.40)	10 (0.08)	10 (0.08)	10 (0.08)

Notes: Table presents overall admission score points and points associated with each component of the admission score for each application year in our sample. Numbers in parentheses represent the weight attached to each component when calculating the total application score.

Table A2: Labor Market Match Rate by Quarters Male Students

	Observed in Labor Market	
Quarters Count	No	Yes
1	53.57	46.43
2	42.05	57.95
3	40.32	59.68
4	40.80	59.20
5	42.42	57.58
6	38.22	61.78
7	37.26	62.74
8	38.89	61.11
9	40.58	59.42
10	37.86	62.14
11	37.26	62.74
12	37.80	62.20
13	40.06	59.94
14	37.38	62.62
15	36.89	63.11
16	37.22	62.78
17	39.50	60.50
18	37.79	62.21
19	37.83	62.17
20	38.25	61.75
21	41.37	58.63
22	39.19	60.81
23	39.23	60.77
24	39.43	60.57
25	41.67	58.33
Total	40.42	59.58

Notes: Table presents the fraction of the sample of male applicants observed in the labor market in a given quarter where quarters are enumerated starting in the first quarter of the calendar year five years after starting high school.

Table A3: Summary Statistics Above and Below Threshold

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Around	Below	Above		Around	Below	Above
		Threshold BW	Threshold BW	Threshold BW		Threshold BW	Threshold BW	Threshold BW
	Full Sample	10	10	10	Full Sample	10	10	10
Outcomes								
H.S. Graduation	0.84	0.83	0.76	0.87	0.83	0.82	0.77	0.85
Any College Enrollment	0.43	0.39	0.38	0.40	0.57	0.52	0.48	0.55
Semesters in College	1.98	1.66	1.54	1.73	2.79	2.37	2.13	2.51
10th Grade Reading Score	222	215	208	219	226	217	210	221
10th Grade Math Score	243	234	225	239	228	217	208	222
9th Grade Attendance	0.94	0.94	0.93	0.95	0.93	0.93	0.91	0.94
Total Earnings	70,685	71,069	61,393	76,742	50,837	51,273	49,154	52,562
Quarterly Earnings	5,581	5,500	4,810	5,905	4,055	3,984	3,836	4,074
Quarters with Earnings	10.91	11.22	11.00	11.35	11.05	11.41	11.30	11.48
Control Variables								
Asian	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Black	0.17	0.19	0.21	0.17	0.24	0.26	0.28	0.25
Hispanic	0.28	0.30	0.32	0.29	0.38	0.41	0.42	0.40
Free Lunch	0.59	0.64	0.69	0.61	0.74	0.79	0.82	0.77
English Learner	0.05	0.06	0.07	0.05	0.07	0.09	0.10	0.08

Notes: Table presents summary statistics for main outcomes used in analysis and control variables. Columns 1-4 report means for sample of male students. Columns 5-8 report means for sample of female students. Columns 1 and 5 report means for full sample of students that applied to a CTHSS school. Columns 2-4 and 6-7 report means for students around the admission threshold based on a 10 point bandwidth and above and below the threshold.

Table A4: First Stage Alternative Bandwidths Male Students

Outcome	(1) Bandwidth 6	(2) Bandwidth 8	(3) Bandwidth 10
Offer	0.523***	0.565***	0.583***
	(0.0298)	(0.0261)	(0.0229)
Controls	Yes	Yes	Yes
Observations	5,801	7,629	9,287

Notes: Table presents first-stage RD estimates for alternative bandwidths. Sample is limited to male students. All specifications include the full set of controls listed in Table 3 and CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A5: First Stage Estimates Labor Market (BW 10)

Outcome	Probability of Attending Male Students		Probability of Attending Female Students	
	(1)	(2)	(3)	(4)
Offer	0.597***	0.598***	0.550***	0.551***
	(0.0232)	(0.0234)	(0.0246)	(0.0253)
Controls	No	Yes	No	Yes
F	661	653	498	475
Observations	5,652	5,652	4,315	4,315

Notes: Table presents first-stage estimates based on sample of students in labor market. Estimates are for probability of attending a CTHSS school after receiving an offer where dependent variable is an indicator for attendance at a CTHSS school in 9th grade. Columns 1 and 2 are for sample of male students in labor market. Columns 3 and 4 are for sample of female students in labor market. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A6: School-by-Year Running Variable Density Test *p*-Value All Applicants Sample

School	2006	2007	2008	2009	2010	2011	2012	2013
				Application Year				
911	0.314	0.020	0.000	0.140	0.002	0.031	0.063	0.000
912	0.035	0.002	0.324	0.700	0.622	0.444	0.691	0.820
913	0.633	0.194	0.509	0.603	0.772	0.520	0.475	0.116
914	0.212	0.001	0.000	0.083	0.003	0.189	0.045	0.860
915	0.338	0.009	0.128	0.680	0.575	0.024	0.020	0.003
916	0.285	0.487	0.635	0.403	0.760	0.092	0.858	0.372
917	0.011	0.189	0.855	0.785	0.719	0.097	0.095	0.303
918	0.465	0.000	0.006	0.024	0.332	0.634	0.829	0.162
919	0.970	0.038	0.000	0.537	0.001	0.082	0.000	0.022
920	0.204	0.271	0.074	0.392	0.086	0.109	0.323	0.924
922	0.354	0.324	0.432	0.320	0.728	0.200	0.034	0.008
923	0.015	0.421	0.900	0.258	0.485	0.017	0.370	0.080
924	0.071	0.703	0.924	0.461	0.411	0.835	0.452	0.960
925	0.809	0.015	0.604	0.590	0.007	0.014	0.938	0.205
926	0.029	0.232	0.013	0.000	0.001	0.000	0.000	0.517
927	0.706	0.867	0.137	0.225	0.042	0.200	0.810	0.376

Notes: Table presents p-value associated with a McCrary test for manipulation for each school and application year. Shaded cells represent cases where we fail the McCrary test at the 10 percent significance level or below.

Table A7: Covariate Balancing Tests Male Students Alternative Bandwidths

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
						7th Grade Test	In Labor
Outcome	Asian	Black	Hispanic	Free Lunch	English Learner	Scores	Market Sample

				Bandwidth 6			
Offer	-0.00354	-0.0135	-0.00774	0.00435	0.0175	-1.681	-0.0234
	(0.00640)	(0.0168)	(0.0192)	(0.0231)	(0.0137)	(1.164)	(0.0246)
Observations	5,801	5,801	5,801	5,801	5,801	4,246	5,801

				Bandwidth 8			
Offer	-0.00298	-0.0186	-0.00654	-0.000910	0.0173	-1.443	-0.0147
	(0.00480)	(0.0134)	(0.0109)	(0.0153)	(0.0157)	(1.106)	(0.0224)
Observations	7,629	7,629	7,629	7,629	7,629	5,615	7,629

	Spending Per Pupil	Pupil Teacher Ratio	8th Grade Average Math Score	8th Grade Average Reading Score	Math % Proficient	Reading % Proficient
Outcome						

			Bandwidth 6			
Offer	93.51	0.0857	0.638	1.242	-0.0597	-0.162
	(208.6)	(0.0850)	(0.937)	(0.974)	(0.202)	(0.191)
Observations	682	5,765	2,032	2,032	5,575	5,575

			Bandwidth 8			
Offer	82.42	0.0521	0.518	0.938	-0.159	-0.222
	(218.6)	(0.0735)	(0.862)	(0.881)	(0.175)	(0.150)
Observations	905	7,582	2,717	2,717	7,337	7,337

Notes: Table presents balancing tests for sample of male students. Estimates are from a RD specification using local linear regression and the bandwidth listed at the top of each panel. Top 2 panels present balancing tests for individual-level covariates. Columns 1-4 of the bottom two panels present balancing tests for spending per-pupil, pupil-teacher ratio and 8th grade average test scores for sending middle schools. All specifications other than spending per-pupil include CTHSS school-by-year fixed effects and resident town fixed effects. Spending per pupil specification omits town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Balancing Tests Female Students

Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Individual-level Covariates			
	Asian	Black	Hispanic	Free Lunch	English Learner	7th Grade Test Scores	In Labor Market Sample
Offer	0.00258	-0.0155	0.00126	-0.0178	-0.00358	0.431	-0.00976
	(0.00401)	(0.0209)	(0.0211)	(0.0157)	(0.0187)	(1.749)	(0.0196)
Observations	7,629	7,629	7,629	7,629	7,629	5,684	7,629
				School / Town-level Covariates			
	Spending Per Pupil	Pupil Teacher Ratio	8th Grade Average Math Score	8th Grade Average Reading Score	Math % Proficient	Reading % Proficient	
Outcome							
Offer	0.113	-0.136	1.653*	2.065*	-0.0938	0.0122	
	(154.3)	(0.112)	(0.854)	(1.158)	(0.167)	(0.168)	
Observations	1,038	7,567	2,715	2,715	7,370	7,368	

Notes: Table presents balancing tests for sample of female students. Estimates are from a RD specification using local linear regression and a bandwidth of 10. Top panel presents balancing tests for individual-level covariates. Columns 1-4 of the bottom panel present balancing tests for spending per-pupil, pupil-teacher ratio and 8th grade average test scores for sending middle schools. Columns 5 and 6 present balancing tests for sending town % proficient in math and reading. Spending per-pupil is for sending middle schools in 2017, pupil teacher ratio is for 2006-2013, 8th grade average math and rading scores are from 2009 -2011, Math and Reading % Proficient are measured at the town level and are for 2006 - 2013. All specifications other than spending per-pupil include CTHSS school-by-year fixed effects and resident town fixed effects. Spending per pupil specification omits town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A9: Covariate Balancing Tests Labor Market Sample Male Students

Outcome	(1)	(2)	(3)	(4)	(5)	(6)
			Individual-level Covariates			
	Asian	Black	Hispanic	Free Lunch	English Learner	7th Grade Test Scores
Offer	0.00338	-0.00726	-0.00193	0.0203	0.0141	0.201
	(0.00472)	(0.0168)	(0.0225)	(0.0218)	(0.0222)	(1.210)
Observations	5,652	5,652	5,652	5,652	5,652	3,623
			School / Town-level Covariates			
	Spending Per Pupil	Pupil Teacher Ratio	8th Grade Average Math Score	8th Grade Average Reading Score	Math % Proficient	Reading % Proficient
Offer	-33.40	0.00402	1.159	1.224	0.106	0.0462
	(244.3)	(0.100)	(0.828)	(0.833)	(0.167)	(0.144)
Observations	793	5,612	2,602	2,602	5,444	5,444

Notes: Table presents balancing tests for sample of male students observed in the labor market. Estimates are from a RD specification using local linear regression and a bandwidth of 10. Top panel presents balancing tests for individual-level covariates. Columns 1-4 of the bottom panel present balancing tests for spending per-pupil, pupil-teacher ratio and 8th grade average test scores for sending middle schools. Columns 5 and 6 present balancing tests for sending town % proficient in math and reading. Spending per-pupil is for sending middle schools in 2017, pupil teacher ratio is for 2006-2013, 8th grade average math and reading scores are from 2009 -2011, Math and Reading % Proficient are measured at the town level and are for 2006 - 2013. All specifications other than spending per-pupil include CTHSS school-by-year fixed effects and resident town fixed effects. Spending per pupil specification omits town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A10: Covariate Balancing Men Tests No Donut Hole Male Students

Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Individual-level Covariates						
	Asian	Black	Hispanic	Free Lunch	English Learner	7th Grade Test Scores	In Labor Market Sample
Offer	-0.00276	0.00231	-0.0111	-0.0103	0.0158	-0.958	-0.00538
	(0.00384)	(0.0126)	(0.00824)	(0.0159)	(0.00993)	(0.814)	(0.0174)
Observations	10,039	10,039	10,039	10,039	10,039	7,426	10,039
	School / Town-level Covariates						
	Spending Per Pupil	Pupil Teacher Ratio	8th Grade Average Math Score	8th Grade Average Reading Score	Math % Proficient	Reading % Proficient	
Offer	224.8	0.0365	0.470	0.624	-0.134	-0.148	
	(148.4)	(0.0831)	(0.813)	(0.861)	(0.129)	(0.122)	
Observations	1,189	9,966	3,630	3,630	9,646	9,646	

Notes: Table presents balancing tests for sample of male students. Estimates are from a RD specification using local linear regression and a bandwidth of 10. Top panel presents balancing tests for individual-level covariates. Columns 1-4 of the bottom panel present balancing tests for spending per-pupil, pupil-teacher ratio and 8th grade average test scores for sending middle schools. Columns 5 and 6 present balancing tests for sending town % proficient in math and reading. Spending per-pupil is for sending middle schools in 2017, pupil teacher ratio is for 2006-2013, 8th grade average math and reading scores are from 2009 -2011, Math and Reading % Proficient are measured at the town level and are for 2006 - 2013. All specifications other than spending per-pupil include CTHSS school-by-year fixed effects and resident town fixed effects. Spending per pupil specification omits town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A11: 2SLS Estimates (Bandwidth 10) Female Students

Outcome	(1) Grad	(2) Grad	(3) Sem Col	(4) Sem Col	(5) Sem Col 23+	(6) Sem Col 23+	(7) Any Coll 23+	(8) Any Coll 23+
Attend	0.0125	0.0123	-0.297	-0.289	-0.656	-0.672	-0.00915	-0.0119
	(0.0397)	(0.0397)	(0.295)	(0.293)	(0.500)	(0.480)	(0.0711)	(0.0693)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	7,627	7,627	7,629	7,629	4,044	4,044	4,044	4,044
Outcome	Total Earnings	Total Earnings	Quarterly Earnings	Quarterly Earnings	Quarters with Earnings	Quarters with Earnings	Quarterly Earnings 23+	Quarterly Earnings 23+
Attend	0.0249	0.0238	0.0405	0.0385	0.0147	0.0167	-0.000805	-0.00422
	(0.123)	(0.125)	(0.0754)	(0.0771)	(0.531)	(0.524)	(0.119)	(0.121)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4,315	4,315	4,315	4,315	4,315	4,315	2,676	2,676

Notes: Table presents 2SLS estimates for main outcomes based on sample of female students. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A12: 2SLS Estimates (Bandwidth 10) Combined Sample of Male & Female Students

Outcome	(1) Grad	(2) Grad	(3) Sem Col	(4) Sem Col	(5) Sem Col 23+	(6) Sem Col 23+	(7) Any Coll 23+	(8) Any Coll 23+
Attend	0.0647**	0.0642**	-0.427***	-0.397**	-0.392	-0.368	-0.0178	-0.0146
	(0.0248)	(0.0249)	(0.163)	(0.161)	(0.284)	(0.274)	(0.0389)	(0.0378)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	16,925	16,925	16,930	16,930	9,449	9,449	9,449	9,449

Outcome	Total Earnings	Total Earnings	Quarterly Earnings	Quarterly Earnings	Quarters with Earnings	Quarters with Earnings	Quarterly Earnings 23+	Quarterly Earnings 23+
Attend	0.248***	0.241***	0.195***	0.184***	0.595*	0.620**	0.180**	0.168**
	(0.0812)	(0.0814)	(0.0584)	(0.0581)	(0.302)	(0.299)	(0.0805)	(0.0791)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	9,981	9,981	9,981	9,981	9,981	9,981	6,319	6,319

Notes: Table presents 2SLS estimates for main outcomes based on the combined sample of male and female students. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A13: Fisher Exact Tests Male Students

Outcome	Graduation	Sem Col	Quarterly Earnings	Quarters with Earnings
	(1)	(2)	(3)	(4)
		p-values for Core Results		
Offer	0.002	0.016	0.000	0.025
		Fisher Exact Test *p*-values		
Offer	0.000	0.000	0.000	0.000
Observations	9,287	9,287	5,652	5,652

Notes: Table presents reduced-form RD *p*-values for estimated coefficient on offer. Top panel presents *p*-values associated with reduced form estimates. Bottom panel presents *p*-values associated with Fisher's exact test as proposed by Cattaneo et al. (2019).

Table A14: Falsification Test - First Stage (Bandwidth 10) Male Students

Outcome	(1) Cutoff -10	(2) Cutoff +10	(3) Cutoff +15	(4) Cutoff +20
A. Full Sample				
Offer	-0.00508	0.0173	0.00524	0.0220
	(0.00687)	(0.0195)	(0.0230)	(0.0188)
Controls	Yes	Yes	Yes	Yes
Observations	5,885	11,395	10,053	8,880
B. Labor Market Sample				
Offer	-0.000796	0.0111	0.0130	0.0311
	(0.00578)	(0.0245)	(0.0281)	(0.0269)
Controls	Yes	Yes	Yes	Yes
Observations	3,571	6,724	5,916	5,238

Notes: Table presents first-stage RD falsification tests based on pseudo cutoffs where we move the actual cutoff threshold: 1) down 10 points, 2) up 10 points, 3) up 15 points, and 4) up 20 points. Panel A is for the full set of male students that applied to a CTHSS school. Panel B is for sample of male students in labor market. All specifications include the full set of controls listed in Table 3 and CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A15: 2SLS Estimates (Bandwidth 10) Male Students No Donut Hole

Outcome	(1) Grad	(2) Grad	(3) Sem Col	(4) Sem Col	(5) Sem Col 23+	(6) Sem Col 23+	(7) Any Coll 23+	(8) Any Coll 23+
Attend	0.109***	0.108***	-0.536**	-0.538**	-0.518	-0.530	-0.0675	-0.0685
	(0.0328)	(0.0325)	(0.211)	(0.213)	(0.359)	(0.365)	(0.0419)	(0.0420)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	10,036	10,036	10,039	10,039	5,842	5,842	5,842	5,842

Outcome	Total Earnings	Total Earnings	Quarterly Earnings	Quarterly Earnings	Quarters with Earnings	Quarters with Earnings	Quarterly Earnings 23+	Quarterly Earnings 23+
Attend	0.431***	0.436***	0.318***	0.320***	0.914**	0.935**	0.340***	0.327***
	(0.0895)	(0.0900)	(0.0617)	(0.0599)	(0.451)	(0.460)	(0.0774)	(0.0696)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	6,142	6,142	6,142	6,142	6,142	6,142	3,915	3,915

Notes: Table presents 2SLS estimates for main outcomes based on sample of male students. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Columns with controls include full set of controls. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A16: 2SLS Estimates (Bandwidth 10) Male Students: Dropping School-Years that Fail Density Test for Manipulation

Outcome	(1) Grad	(2) Grad	(3) Sem Col	(4) Sem Col	(5) Sem Col 23+	(6) Sem Col 23+	(7) Any Coll 23+	(8) Any Coll 23+
Attend	0.130***	0.132***	-0.435	-0.456	0.0795	0.0765	-0.0225	-0.0227
	(0.0429)	(0.0417)	(0.339)	(0.344)	(0.450)	(0.453)	(0.0651)	(0.0664)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	5,351	5,351	5,354	5,354	3,334	3,334	3,334	3,334
Outcome	Total Earnings	Total Earnings	Quarterly Earnings	Quarterly Earnings	Quarters with Earnings	Quarters with Earnings	Quarterly Earnings 23+	Quarterly Earnings 23+
Attend	0.425***	0.436***	0.332***	0.339***	0.963	1.006*	0.427***	0.417***
	(0.129)	(0.122)	(0.0897)	(0.0832)	(0.585)	(0.578)	(0.136)	(0.125)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3,337	3,337	3,337	3,337	3,337	3,337	2,251	2,251

Notes: Table presents 2SLS estimates for main outcomes based on sample of male students. Sample is restricted by dropping any school-year observations that fail a density test for manipulation at a significance level of 0.10 or lower. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A17: Robustness Test of Core Reduced Form Results Dropping 2010 and Later Male Students

Outcome	(1) Grad	(2) Sem Col	(3) Quarterly Earnings	(4) Quarters With Earnings	(5) Quarterly Earnings 23+
	Main Reduced Form Results from Paper Sample of Male Students				
Offer	0.0580***	-0.278**	0.193***	0.681**	0.195***
	(0.0183)	(0.116)	(0.0371)	(0.305)	(0.0569)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	9,284	9,287	5,652	5,652	3,612
	Reduced Form Results Based on 2006-2009 Sample of Male Students				
Offer	0.0618**	-0.161	0.201***	1.038**	0.190***
	(0.0249)	(0.232)	(0.0405)	(0.425)	(0.0508)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	4,515	4,515	3,846	3,846	3,374

Notes: Table presents reduced-form RD estimates for main outcomes based on the sample of male students. Panel A replicates results presented in main paper for comparison purposes. Panel B limits the sample to application years 2006 - 2009. All specifications include the full set of controls, CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A18: 2SLS Estimates (Bandwidth 10) Male Students: Dropping Students that Applied to More than One School

Outcome	(1) Grad	(2) Grad	(3) Sem Col	(4) Sem Col	(5) Sem Col 23+	(6) Sem Col 23+	(7) Any Coll 23+	(8) Any Coll 23+
Attend	0.101***	0.0994**	-0.369	-0.390	-0.117	-0.140	-0.0195	-0.0230
	(0.0377)	(0.0382)	(0.274)	(0.295)	(0.393)	(0.411)	(0.0644)	(0.0665)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	5,864	5,864	5,865	5,865	3,825	3,825	3,825	3,825

Outcome	Total Earnings	Total Earnings	Quarterly Earnings	Quarterly Earnings	Quarters with Earnings	Quarters with Earnings	Quarterly Earnings 23+	Quarterly Earnings 23+
Attend	0.318***	0.328***	0.253***	0.261***	0.870	0.882	0.254**	0.247**
	(0.114)	(0.114)	(0.0765)	(0.0752)	(0.557)	(0.562)	(0.106)	(0.106)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3,849	3,849	3,849	3,849	3,849	3,849	2,616	2,616

Notes: Table presents 2SLS estimates for main outcomes based on sample of male students. Sample is restricted to include students that only applied to one CTHSS school. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A19: 2SLS Estimates (Bandwidth 10) Male Students Including Special Education Students and Students that Applied to CTHSS in 10th Grade

Outcome	(1) Grad	(2) Grad	(3) Sem Col	(4) Sem Col	(5) Sem Col 23+	(6) Sem Col 23+	(7) Any Coll 23+	(8) Any Coll 23+
Attend	0.0730***	0.0705***	-0.306**	-0.329**	-0.194	-0.237	-0.0272	-0.0316
	(0.0261)	(0.0254)	(0.132)	(0.133)	(0.200)	(0.208)	(0.0291)	(0.0311)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	12,676	12,676	12,684	12,684	7,685	7,685	7,685	7,685
Outcome	Total Earnings	Total Earnings	Quarterly Earnings	Quarterly Earnings	Quarters with Earnings	Quarters with Earnings	Quarterly Earnings 23+	Quarterly Earnings 23+
Attend	0.415***	0.398***	0.296***	0.283***	1.041***	1.014***	0.248***	0.227***
	(0.0922)	(0.0905)	(0.0613)	(0.0594)	(0.368)	(0.370)	(0.0751)	(0.0720)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	7,731	7,731	7,731	7,731	7,731	7,731	5,199	5,199

Notes: Table presents 2SLS estimates for main outcomes based on sample of male students. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Sample includes special education students and students that applied to CTHSS in 9th grade. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A20: Reduced Form Effect Estimates by School by Year Male Students

Outcome	(1) Grad	(2) Sem Col	(4) Quarterly Earnings	(6) Quarters with Earnings
	Main Reduced Form Estimates from Paper			
Offer	0.0580***	-0.278**	0.193***	0.681**
	(0.0183)	(0.116)	(0.0371)	(0.305)
Observations	9,284	9,287	5,652	5,652
	Bertanha Weight Average			
Offer	0.0640***	-0.225	0.208***	0.822**
	(0.0247)	(0.155)	(0.054)	(0.372)
Observations	128	128	96	96
	Coefficient from Regression of Reduced Form Treatment Estimate on Cutoff Score			
Offer	-0.00387	-0.0125	-0.00274	-0.00908
	(0.00409)	(0.0224)	(0.00881)	(0.0635)
Observations	128	128	96	96

Notes: Table presents RD estimates for core outcomes based on sample of male students. Panel A replicates the results presented in Panel 3 of Table 6. Panel B presents estimates based on the methodology suggested by Bertanha (2020) where we estimate separate treatment effects for each school and year and then take a weighted average of the estimates where the weights correspond to the number of observations within the 10-point bandwidth for each school and year. The standard errors reported in Panel B represent bootstrapped standard errors based on 1,000 bootstrap estimates. Panel C presents estimates based on specifications where we regress the individual school and year treatment estimates on the cutoff score of each school in each year. Regressions are weighted by the number of observations within the 10-point bandwidth for each school and year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A21: 2SLS Quarterly Earnings Estimates Male Students

Outcome	(1)	(2)	(3)
	Quarterly Earnings		
	At Least 1 Quarter of Earnings	At least 2 Quarters of Earnings	At Least 4 Quarters of Earnings
Attend	0.323***	0.303***	0.309***
	(0.0640)	(0.0601)	(0.0594)
Controls	Yes	Yes	Yes
Observations	5,652	5,482	4,980

Notes: Table presents 2SLS estimates for quarterly earnings based on sample of male students. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. All specifications include the full set of controls listed in Table 3, CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A22: Quarterly Earnings Quantile Regression Estimates Male Students Reduced Form

Outcome	(1) 25th Percentile	(2) 50th Percentile	(3) 75th Percentile
		Quarterly Earnings	
Offer	0.219***	0.177***	0.177***
	(0.0623)	(0.0545)	(0.0494)
Controls	Yes	Yes	Yes
Observations	5,665	5,665	5,665

Notes: Table presents reduced form quantile regression estimates for quarterly earnings based on sample of male students. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. All specifications include the full set of controls listed in Table 3, CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A23: 2SLS Estimates Male Students: Educational Outcomes Labor Market Sample

Outcome	(1) Grad	(2) Grad	(3) Sem Col	(4) Sem Col	(5) Sem Col 23+	(6) Sem Col 23+	(7) Any Coll 23+	(8) Any Coll 23+
Attend	0.102***	0.103***	-0.369	-0.363	-0.0857	-0.0804	-0.00828	-0.00788
	(0.0320)	(0.0306)	(0.296)	(0.296)	(0.355)	(0.366)	(0.0437)	(0.0450)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	5,652	5,652	5,652	5,652	4,530	4,530	4,530	4,530

Notes: Table presents 2SLS estimates for main educational outcomes based on the sample of male students. Sample is restricted to students in the labor market sample. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A24: 2SLS Labor Market Estimates Male Students Including Students not Observed in 9th Grade

Outcome	(1) Total Earnings	(2) Total Earnings	(3) Quarterly Earnings	(4) Quarterly Earnings	(5) Quarters with Earnings	(6) Quarters with Earnings	(7) Quarterly Earnings 23+	(8) Quarterly Earnings 23+
Attend	0.443***	0.438***	0.326***	0.321***	1.161**	1.166**	0.347***	0.328***
	(0.0936)	(0.0912)	(0.0685)	(0.0650)	(0.508)	(0.510)	(0.102)	(0.0972)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	5,780	5,780	5,780	5,780	5,780	5,780	3,689	3,689

Notes: Table presents 2SLS estimates for main outcomes based on sample of male students. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Sample includes students for which we do not observe high school outcomes. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A25: Labor Market Participation Rate for Boundary versus Non-Boundary Towns Male Students

Towns		Observed in Labor Market	
		No	Yes
New York	Border	46.53	53.47
	Adjacent	41.56	58.44
	Interior	40.12	59.88
Total		40.45	59.55

		Observed in Labor Market	
		No	Yes
Massachusetts	Border	47.26	52.74
	Adjacent	38.92	61.08
	Interior	40.12	59.88
Total		40.16	59.84

		Observed in Labor Market	
		No	Yes
Rhode Island	Border	46.08	53.92
	Adjacent	39.17	60.83
	Interior	40.12	59.88
Total		40.29	59.71

Notes: Table presents the fraction of individuals ever observed in the labor market data by whether the town lies on a state border, is adjacent to a border town or adjacent to an adjacent town, or is an interior town. Panels 1, 2 and 3 present the results for the border between Connecticut and New York, Massachusetts and Rhode Island, respectively. Towns that are on the border of another state, but not on the border of the subject state for a given panel, are omited from the calcuations.

Table A26: 2SLS Estimates Male Students Dropping Students at State Boundary Male Students

Outcome	(1) Grad	(2) Grad	(3) Sem Col	(4) Sem Col	(5) Sem Col 23+	(6) Sem Col 23+	(7) Any Coll 23+	(8) Any Coll 23+
Attend	0.0933***	0.0932***	-0.568***	-0.564***	-0.419	-0.420	-0.0430	-0.0429
	(0.0347)	(0.0342)	(0.198)	(0.203)	(0.315)	(0.326)	(0.0428)	(0.0449)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	8,687	8,687	8,690	8,690	5,004	5,004	5,004	5,004

Outcome	Total Earnings	Total Earnings	Quarterly Earnings	Quarterly Earnings	Quarters with Earnings	Quarters with Earnings	Quarterly Earnings 23+	Quarterly Earnings 23+
Attend	0.458***	0.458***	0.327***	0.325***	1.304**	1.327**	0.333***	0.316***
	(0.101)	(0.100)	(0.0708)	(0.0675)	(0.510)	(0.518)	(0.105)	(0.0994)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	5,309	5,309	5,309	5,309	5,309	5,309	3,393	3,393

Notes: Table presents 2SLS estimates for main outcomes based on sample of male students and dropping students that reside in towns that directly border New York, Massachusetts or Rhode Island. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A27: Mechanism - 2SLS (BW 10) Female Students

Outcome	(1) 9th Grade Attendance	(2) 10th Grade Composite Test Scores	(3) 10th Grade Math Score	(4) 10th Grade Reading Score
Attend	0.0122*	0.0144	0.0352	-0.0293
	(0.00641)	(0.0721)	(0.0887)	(0.0723)
Controls	Yes	Yes	Yes	Yes
Industry Fixed Effects	No	No	No	No
Observations	7,629	4,876	4,914	4,938

Notes: Table presents 2SLS RD estimates for 9th grade days of attendance, standardized individual and composite 10th grade reading and math test scores. Sample is limited to female students. All estimates based on local linear regression and a 10 point bandwidth. All specifications include the full set of controls listed in Table 3 and CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 28: Mechanism Earnings Estimates

	(1)	(2)
H.S. Graduation	0.355***	0.338***
	(0.0516)	(0.0413)
9th Grade Attendance	0.477**	0.452**
	(0.207)	(0.191)
10th Grade Test Scores	0.0904***	0.0879***
	(0.0263)	(0.0137)
Observations	3,642	8,586
R-squared	0.097	0.103

Notes: Table presents estimates for the sample of male students based on regressions of the log of quarterly earnings at age 23 or older on an indicator for being a high school graduate, 9th grade attendance, 10th grade standardized math and reading scores, and the individual demographic controls listed in Table 3. All specifications also include CTHSS school-by-year fixed effects. Column 1 presents estimates based on our RD sample with a bandwidth of +/-10 points. Column 2 presents estimates based on the full sample of male students with earnings at age 23 or older. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A29: Counterfactual CTE 2SLS (BW 10) Male Students

Outcome	Graduation	Sem Col	Quarterly Earnings	Quarters with Earnings
	(1)	(2)	(3)	(4)
A. Spending Per Pupil (SPP)				
Attend	0.103***	-0.461**	0.335***	1.109**
	(0.0341)	(0.208)	(0.0727)	(0.443)
Attend*SPP	0.0196	-0.219	0.0112	-0.690
	(0.0335)	(0.198)	(0.0732)	(0.449)
Controls	Yes	Yes	Yes	Yes
Observations	9,162	9,165	5,564	5,564
B. Pupil Teacher Ratio (PTR)				
Attend	0.0999***	-0.489**	0.332***	1.263***
	(0.0333)	(0.209)	(0.0676)	(0.422)
Attend*PTR	-0.000527	0.0643	0.0101	0.423
	(0.0360)	(0.125)	(0.0579)	(0.430)
Controls	Yes	Yes	Yes	Yes
Observations	9,180	9,183	5,576	5,576
C. Average 10th Grade Math & English Scores				
Attend	0.107***	-0.435**	0.351***	1.203***
	(0.0336)	(0.205)	(0.0688)	(0.443)
Attend*Score	0.0225	-0.0116	0.0694	-0.530
	(0.0302)	(0.155)	(0.0508)	(0.334)
Controls	Yes	Yes	Yes	Yes
Observations	9,180	9,183	5,576	5,576

Notes: Table presents 2SLS RD estimates for main outcomes based on sample of male students. Panel A interacts the attend a CTE school indicator with average spending per pupil in a students' residential town high school(s). Panel B interacts the attend indicator with the pupil teacher ratio in a students' residential town high school while Panel C interacts the attend indicator with average standardized math and english 10th grade test scores in a students's residential town high school. All specifications include the full set of controls listed in Table 3, CTHSS school-by-year fixed effects and resident town fixed effects and interactions between either spending per pupil, pupil teacher ratio or average 10th grade math and reading scores and the running variable and the running variable interacted with offer. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A30: Heterogeneity Reduced Form (BW 10) Male Students

Outcome	Graduation	Sem Col	Quarterly Earnings	Quarters with Earnings
	(1)	(2)	(3)	(4)
A. Free Lunch Status				
Offer	0.0559**	-0.297	0.241***	0.982**
	(0.0241)	(0.200)	(0.0658)	(0.495)
Offer*Free Lunch	0.00494	0.0249	-0.0693	-0.492
	(0.0473)	(0.369)	(0.0779)	(0.594)
Controls	No	No	No	No
Observations	9,284	9,287	5,652	5,652
B. Race/Ethnicity				
Offer	0.0703***	-0.316*	0.222***	0.393
	(0.0242)	(0.172)	(0.0587)	(0.392)
Offer*Black/Hispanic	-0.0244	0.0835	-0.0639	0.493
	(0.0363)	(0.226)	(0.0697)	(0.542)
Controls	No	No	No	No
Observations	9,284	9,287	5,652	5,652
C. Central City				
Offer	0.0596**	-0.352**	0.222***	0.347
	(0.0265)	(0.154)	(0.0523)	(0.293)
Offer*Central City	-0.00330	0.181	-0.0721	0.952**
	(0.0389)	(0.222)	(0.0582)	(0.425)
Controls	No	No	No	No
Observations	9,284	9,287	5,652	5,652
D. School-by-Year Treatment Effects				
Free Lunch	0.0884	0.472	-0.0556	0.850
	(0.149)	(0.813)	(0.266)	(1.916)
Black/Hispanic	0.0468	0.0103	0.0172	1.306
	(0.112)	(0.613)	(0.202)	(1.449)
Central City	-0.0387	0.0177	-0.140	0.801
	(0.0932)	(0.509)	(0.174)	(1.254)
Controls	No	No	No	No
Observations	128	128	96	96

Notes: Table presents reduced form RD estimates for main outcomes based on sample of male students. Panel A interacts the offer to attend a CTE school indicator with the indicator for free lunch eligible students. Panel B interacts the offer indicator with an indicator for whether a student is Black or Hispanic. Panel C interacts the offer indicator with an indicator for whether a student resides in one of Connecticut's five central cities. All estimates based on local linear regression and a 10-point bandwidth. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects and interactions between the relevant characteristic (free lunch status, ethnicity and central city) and the running variable and the running variable interacted with offer. Robust standard errors, clustered at the school-by-year and town levels in parentheses. Panel D presents estimates based on specifications where we regress the individual school and year treatment estimates on the share of free lunch students, black or Hispanic students and students from one of Connecticut's 5 central cities. The shares are constructed based on the attributes of students within a 10-point bandwidth around the cutoff score for each school and year. Regressions are weighted by the number of observations within the 10-point bandwidth for each school and year.*** p<0.01, ** p<0.05, * p<0.1

Figure A1: Individual School & Year RD Point Estimates versus Cutoffs:

H.S. Graduation
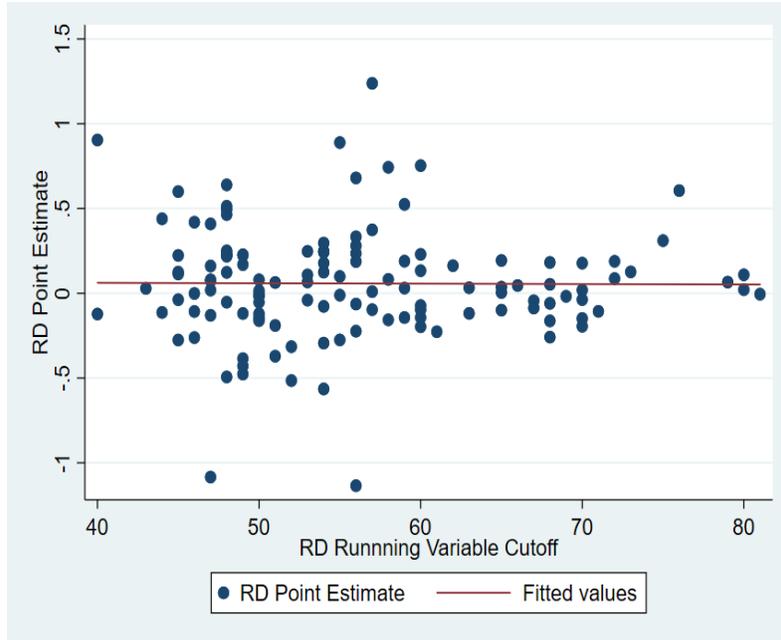
Log Quarterly Earnings
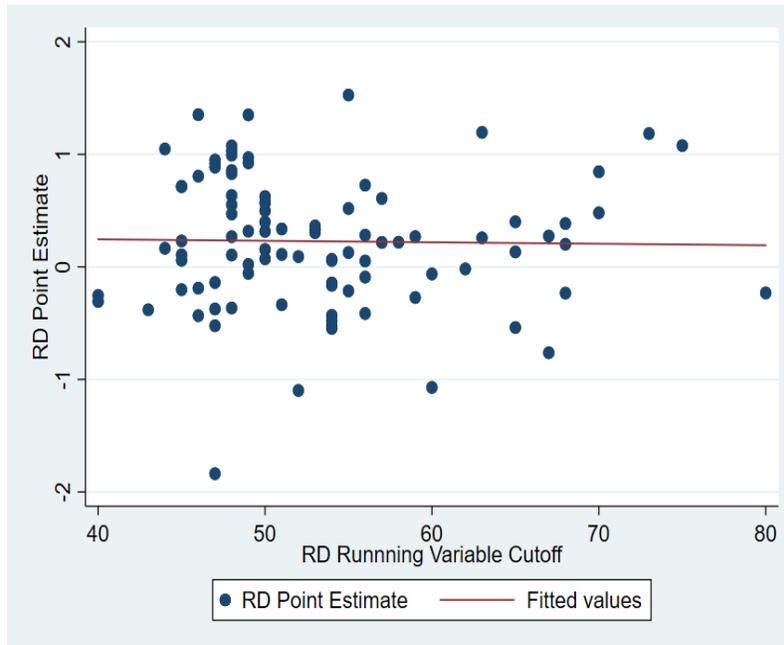
B. Methodological Appendix (Online Only)

As discussed in the paper, the Connecticut Technical High School System (CTHSS) admits students based on a discrete score variable that runs between zero and a maximum of between 100 and 120 depending upon the year considered. This score is the sum of individual components based on standardized state level mathematics and language arts test scores, grade point average, attendance, and for 2009 or later, points assigned based on extracurricular activities and a written statement. The weights assigned to each component vary from year to year, and the exact weights are shown in Appendix Table A1.

Two issues arise, however, that complicate our analysis relative to a simple fuzzy regression discontinuity. First, the individual schools never establish a formal cut-off score, and in fact we can find many examples of individuals receiving acceptance letters who have scores below the scores of individuals who did not receive acceptance letters for the same school and application year. Therefore, there are two sources of noise associated with the discontinuity: deviations of school administrators from a high to low score admission/acceptance rule and differences in take-up among accepted applicants; and accordingly the true location of the discontinuity is unknown. The second issue that arises is that the individual score components are both correlated with each other and discrete in nature leading to raw score distributions that have heaps or mass points at specific locations in the distribution, making it impossible to directly identify manipulation at the threshold or boundary using standard McCrary tests.

B1. Empirical Identification of Cut-offs

While receipt of an acceptance letter is not perfectly determined by a student's application score, the mass of the acceptances for each school and year appear concentrated above apparent score thresholds, and this concentration is especially true for the initial batch of acceptance letters sent out by the schools each year. Therefore, we follow Porter and Yu (2015) and select the threshold for each year and school from the empirical distribution of applications.

Porter and Yu (2015) recommend estimating the location of the threshold by selecting the threshold that maximizes the size of the discontinuity, or specifically:

$$\widehat{X^*} = argmax_{X^*} \hat{\alpha}(X^*) \tag{B1}$$

where the treatment (T) is defined by

$$T = \begin{cases} 1 \ if \ y^* \geq 0 \\ 0 \ if \ y^* < 0 \end{cases}, \ T^* = \begin{cases} f_1(X) + \varepsilon & if \ X < X^* \\ f_2(X) + \alpha + \varepsilon & if \ X \geq X^* \end{cases}, \ f_1(X^*) = f_2(X^*) \tag{B2}$$

where $f_1(X)$ and $f_2(X)$ are continuous and differentiable functions. The inclusion of different functions on either side of the threshold allows for differential processes for non-compliance. In practice, we estimate equation (A2) using a linear probability model and specifying $f_1(X)$ and $f_2(X)$ as linear functions of X.

Porter and Yu (2015) estimate $\widehat{X^*}$ using the discontinuity in the outcome in order to develop second stage specification tests for the existence of treatment effects at that unknown discontinuity. However, Porter and Yu's specification tests require both continuous outcomes and a smooth distribution of the population over scores X. Therefore, we take a different approach exploiting fuzzy regression discontinuity approaches. We select the thresholds applying equation (B1) to the actual school decision to send an acceptance letter to an applicant since that is the decision process that creates the discontinuity. Then, with the estimated thresholds in hand, we estimate a two stage least squares model of student outcomes (y) that incorporates the additional noise created by student take-up so that the first stage is whether we see the applicant in the technical high school the year after they applied (A). Specifically, the fuzzy RD takes the following form:

$$y = \beta A + g_1(X) + d(\widehat{X^*} \leq X)g_2(X) + \varepsilon_2 \tag{B3}$$

where A is instrumented by

$$A = \tilde{\alpha}d(\widehat{X^*} \leq X) + h_1(X) + d(\widehat{X^*} > X)h_2(X) + \varepsilon_3 \tag{B4}$$

and $d(\widehat{X^*} \leq X)$ is an indicator function that takes the value one when the condition is satisfied, and both $g_1(\widehat{X^*}) = g_2(\widehat{X^*})$ and $h_1(\widehat{X^*}) = h_2(\widehat{X^*})$.

The disadvantage of this approach is that our analysis is conditional on the estimated discontinuity and the estimated discontinuity may not represent the true discontinuity. However, in a 2SLS context, we only need to establish the power of the instrument and the validity of the exclusion restriction. The issue of whether a discontinuity exists can be determined by examining the power of the instrument. If the estimated threshold has significant power to explain student attendance A then the discontinuity exists, and the second stage estimates and standard errors from 2SLS will be consistent as long as the exclusion restriction is valid. Determining the power of the instrument, however, is not straightforward since the cutoff has been selected using the same data that is then used to estimate the first stage equation. Naturally, treatment (receiving an acceptance letter) affects attendance and so the F-test may be biased upwards providing misleading evidence on instrument power because we selected $\widehat{X^*}$ to

maximize the discontinuity for T which is strongly correlated with A. We follow Card, Mas and Rothstein (2008) and address this concern using a hold-out sample.[1]

B1.1 Selection of Cutoffs using Hold-out Samples

In our case, we will divide the applicants in each school and year into equal sized analysis and hold-out samples so that we can use the analysis sample to select the thresholds and use the hold-out sample to estimate the first stage for attendance and examine the power of the instrument. Specifically, we do the following:

1) For each school and year, divide all applicants into deciles, assign each applicant a random number, place applicants in the hold-out sample if they are above the median on the random number and in the analysis sample if below, split median applicants half in hold-out/half in analysis samples, and in the case of an odd number of median applicants assign the last median applicant randomly to either the hold-out or analysis sample.

2) Using our preferred bandwidth of 10 points and linear running variables, estimate equation B1 for each school and year starting with a candidate cutoff score at 10 so that the bottom of the 20-point band is at a score of zero using the analysis sample.

3) Re-estimate these models incrementing the candidate cutoff score by 1 each time and ending 10 points away from the maximum score so that the top of the band at the maximum.

4) Select the cutoff for each school and year as the cutoff that provides the maximum estimate of α.[2]

5) Center the scores for each school and year by subtracting the cutoff and pool all years and schools.

6) Estimate equation (B4) for school attendance and calculate the F-statistic associated with the indicator $d(\widehat{X^*} \leq X)$ using the hold-out sample.

We conduct this hold-out simulation four times, and the resulting F-statistics always fall between 456 and 628 for the full sample and between 458 and 674 for the donut hole sample. We also estimate the first stage models separately for each school and year. The means of the estimated discontinuities over all schools and years range between 0.525 and 0.540 for the four simulations very close to our full sample first stage estimate of 0.582. The fraction of thresholds that are significant at the 10 percent level ranges

---

[1] Card, Mas and Rothstein (2008) calculate tipping points associated with racial segregation for each metropolitan area using a subset of census tracts, and then analyzed the effects of being above the tipping point using the remaining tracts.

[2] Note that we verify that the cutoffs are never anywhere near the edge of the score range considered.

between 0.795 and 0.843. Further, the average threshold size across years for any school or simulation never falls below 0.329, and the fraction of significant thresholds for any school in a simulation never falls below 0.571. However, while the empirically selected thresholds provide a very strong instrument for explaining school attendance across the sample of CTHSS schools, the magnitude of the estimated threshold varies considerably and systematically across the simulations. The correlation between the average threshold for each school across the simulations is very high ranging between 0.77 and 0.91. We observe less evidence of a systematic difference in thresholds across years with the correlation between the average threshold for each year ranging between -0.10 and 0.56. Appendix Table B1 presents the correlations of the average thresholds by school across the four simulations in panel 1, and by year in panel 2. Appendix Table B2 presents the full set of correlations for one of the simulations as an example.

B2. Examining the Empirical Distribution of Scores

We next conduct a series of analyses first intended to demonstrate that the lumpy nature of the raw distribution arises naturally from the data as opposed to manipulation around the threshold (sections B2.1 and B2.2) and then second to demonstrate how features in the centered score distribution could arise naturally from the application process (section B2.3).

B2.1 The Effect of Dropping Students at Threshold

While the underlying test scores, grade point averages and attendance data are relatively continuous, the system for assigning points divides students into more aggregated bins leading to mass points in the empirical distribution of application scores. The raw distribution of applicant admission scores are shown in the left hand side of Figure B1 separately for 2006 to 2008, 2009-2010 and 2011 to 2014. This division of the data is chosen because prior to 2008 the score did not include points from the written statement and extracurricular activities, and in 2011 those points increased from a total of 6 points to a total of 20 points. The distribution contains significant mass points as well as holes in the distribution, which would typically raise concerns about manipulation at the cutoff. However, the right-hand side of Figure B1 shows the same distributions after dropping students for each school and application year whose score was exactly at the cutoff and dropping those students has at most modest effects on the smoothness of the raw score distribution.

As shown in Figure B1, the elimination of applications who have the same score as the admissions letter threshold selected for their school and year does little to change the irregular nature of the distribution. In order assess how much of the change in the distribution we can explain, we calculate the square root of the mean squared error deviation between the full empirical distribution and the distribution minus the applications at the cutoff for each specific application school and year. These

calculations are conducted separately for the periods 2006-2008, 2009-2010 and 2011-2013 as shown in Figure B1. We then conduct a simple simulation that involves the following steps:

1) Select the cutoffs chosen by all schools in a given year as a population of relevant cutoffs for that year in order to take into account the fact that the environment changes from year to year likely changing the location of any mass points.
2) For each school and year, select one cutoff randomly from the population of cutoffs for that year omitting the school's own cutoff for that year.
3) Pool the data across schools and relevant years to obtain a simulated distribution.
4) Delete all applications from the pool for that school and year at the simulated cutoff, and calculate the root mean squared error from the full distribution.
5) Repeat steps 2 through 4 and average the resulting root mean squared errors.

While the simulation cannot completely explain the deviations caused by dropping applications at the cutoff, the simulations explain much of the deviation especially for 2006-2008. Specifically, the mean root squared error from comparing the full raw distribution to the distributions after dropping applications at one of the other school's cutoffs baseline explains 71 percent of the effect of dropping applications at the school's actual cutoff for the 2006-2008 timeframe. This exercise explains 62 and 60 percent of the effect for 2009-2010 and 2011-2013 timeframes, respectively.

B2.2 Simulation of Raw Test Score Distribution

We have argued that the discrete nature of the individual components in the total score and the natural correlation between these components leads to the large number of mass points and holes in the empirical application score distribution. To demonstrate that these components can generate the patterns we observe, we next conduct the following exercise.

1) We measure the correlation between the four or six components in each year.
2) We use these correlations and draw simulated data for each of these components from a multivariate normal distribution in order to match the exact number of applications at each school and year.
3) We then map these components through the normal CDF to obtain a probability, and then assign discrete scores to each component based on the empirical frequency of each discrete score for that school and year.

4) We add the components, and the total resulting scores form the simulated distribution for each year.

The bottom panels of Figures B2-B9 present the results of this exercise for each year between 2006 and 2013. As mentioned above, the left side panel of Figures B2-B9 shows the raw distribution, and the right side panel of each figure shows a simulation of the distribution. While we cannot perfectly replicate the mass points and holes in the true distribution (likely because we have to assume a form for the unknown, underlying latent distribution), these simulations consistently generate a substantial number of sizable mass points and holes in the simulated data. Note that in 2009 we do not observe the sub-scores separately for extracurricular activities and the written statement and so conduct the simulation in 2009 using just the original four components.

B2.3 Simulation of Centered Test Score Distribution

Another unusual feature arises once we center the application score distributions by school and year. The left-hand side of Figures B10-B17 show the centered application score distribution by school and year and there is a substantial cliff to the left of the cutoff location in most cases. These cliffs could be consistent with substantial manipulation, placing students to the right of the cutoff with some noise so that an unexplained mass of students is just to the right of the cutoff. However, it is also possible that the cliffs in the centered score distribution arise naturally from the mass points and holes in the raw distribution. If a school is trying to issue a specific number of acceptances and typically works down the list from high to low scores, then the stopping point for this process is more likely to land on a mass point as opposed to a hole because the mass points contain many students and any of those students could help fill the school's capacity.

To examine whether we can replicate the cliffs we observe in the centered application score distributions, we conduct another simulation. For each school, we capture the number of students admitted each year, and view these numbers as an empirical distribution of the number of students the school might admit. For each school, we also calculate the fraction of students over all years who received an admission letter whose test score falls above the threshold, and the fraction of students over all years who are above the threshold, but did not receive an admission letter as a share of the number of students above the threshold. We then use these ratios to scale the potential numbers of students admitted up, based on some students not being admitted even though they were above the threshold, and down based on students being admitted whose scores were below the threshold. We then conduct the following exercise:

1) For each year and school, we select a scaled number of admitted students from the empirical distribution over all years for that school.

2) We then count down the distribution from high to low test scores for the year and school until we have admitted the scaled number of students. The score of the last student admitted is the simulated threshold cutoff test score.

3) We then use the simulated cutoffs for every year and school to center the test score distribution on the simulated cutoff and then pool the centered distributions over all schools.

The result of this exercise is shown on the right-hand side of Figures B10-B17. Specifically, as noted above, for each year, the left-hand side of Figures B10-B17 show the actual empirical distribution of the re-centered application score and the right-hand side shows the simulated distribution based on the process described above. The resulting simulated distributions are relatively similar to the empirical distributions for the years of 2006-2009, but starting in 2010, and especially in 2011, the simulated centered distributions become much smoother while we continue to see mass points and cliffs in the empirical distribution. While pooling centered distributions should in principle smooth over mass points, we have no explanation for the time pattern of this phenomenon. However, given our inability to generate data that looks similar to the actual centered distribution after 2009, we rerun all our analyses dropping the years between 2010 and 2014. These results are shown in Appendix Table A17 and as noted in the paper, with the exception of college attendance, our core results are robust to this sample restriction.

Table B1
Correlations between Estimated Mean Threshold Sizes

Correlations by School over Sample of Years

Simulation #'s	1	2	3
2	0.871916		
3	0.773716	0.819969	
4	0.840485	0.912581	0.89444

Correlations by Year over Sample of Schools

Simulation #'s	1	2	3
2	0.136102		
3	0.557161	0.470346	
4	0.402145	0.051942	-0.10463

Table B2
Estimated Thresholds for one Simulated Set of Hold-out Samples

Application School/Year	2006	2007	2008	2009	2010	2011	2012	2013	Mean Threshold	Significant at 0.10
911	0.871	0.266	0.881	0.586	0.921	0.831	0.903	0.704	0.745	1.000
912	-0.053	0.411	0.432	0.499	0.450	0.426	0.558	0.652	0.422	0.750
913	0.333	0.493	0.529	1.116	0.374	1.021	-0.299	0.924	0.561	0.750
914	0.506	0.610	0.652	0.271	0.648	0.726	0.818	0.695	0.616	0.875
915	0.563	0.658	0.520	0.602	0.337	0.858	0.647	0.425	0.576	1.000
916	0.674	0.646	0.563	-0.001	0.172	0.228	0.073	0.605	0.370	0.625
917	0.706	0.693	0.100	0.791	0.805	0.581	NA	0.596	0.610	0.750
918	0.592	0.836	0.506	0.697	0.515	0.013	-0.729	0.266	0.337	0.750
919	0.375	0.745	0.821	0.795	0.594	0.582	0.556	0.461	0.616	1.000
920	0.747	0.297	0.341	0.138	0.726	0.461	0.109	0.475	0.412	0.750
922	0.763	0.756	0.359	0.901	-0.094	0.734	0.620	0.152	0.524	0.750
923	0.652	0.606	0.322	0.117	0.810	0.291	0.670	0.436	0.488	0.875
924	0.790	0.474	0.576	-0.692	0.728	0.559	0.560	-0.362	0.329	0.875
925	0.436	0.281	0.528	0.447	0.802	0.865	0.870	0.747	0.622	0.875
926	0.777	0.487	0.665	0.476	0.723	0.599	0.544	0.348	0.577	1.000
927	1.077	0.560	0.681	0.002	0.587	0.582	0.875	0.465	0.604	0.750
Mean Threshold	0.613	0.551	0.530	0.422	0.569	0.585	0.452	0.474	0.525	
Fraction Significant at 0.10	0.938	0.875	0.938	0.688	0.813	0.938	0.733	0.813		0.843

Figure B1: Uconditional Distribution of Running Variable

2006-2008

Actual Distribution   Dropping Observations at the Cut-off
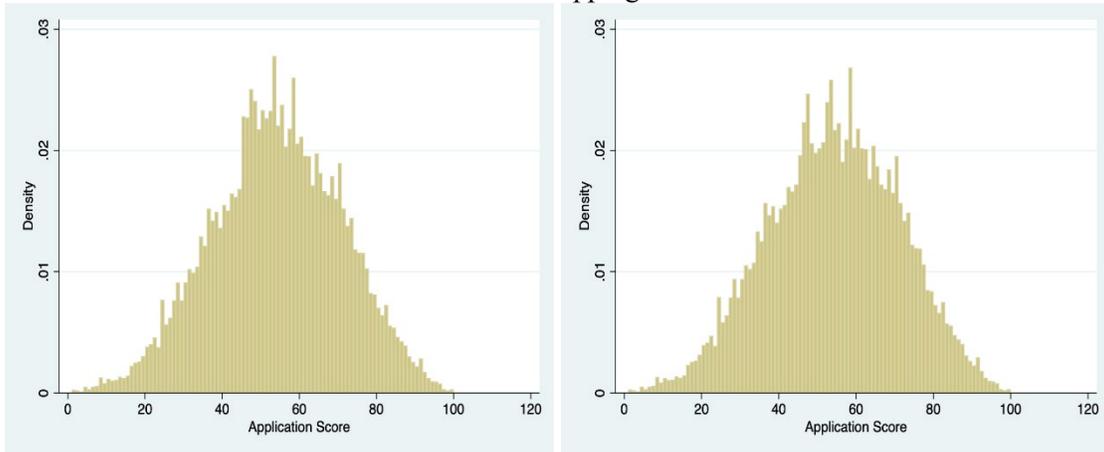
2009-2010

Actual Distribution   Dropping Observations at the Cut-off
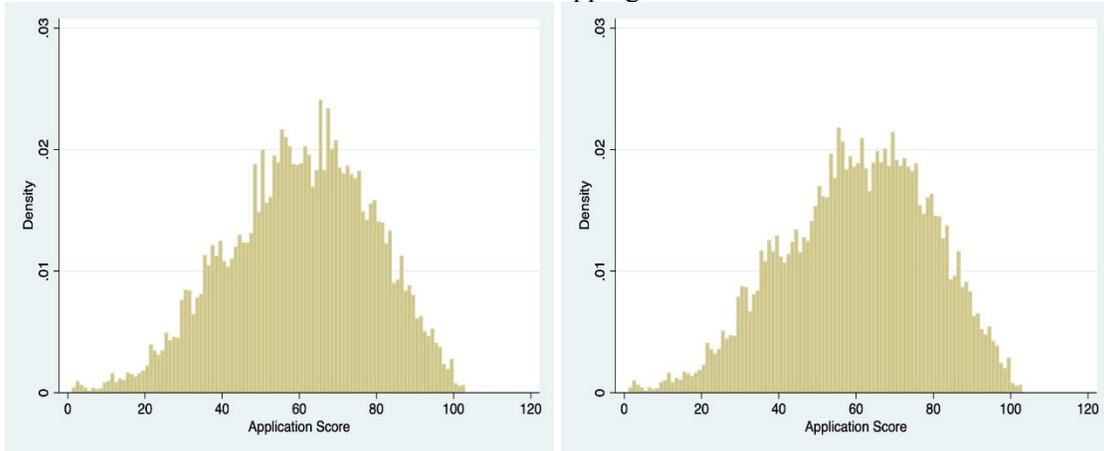
2011-2013

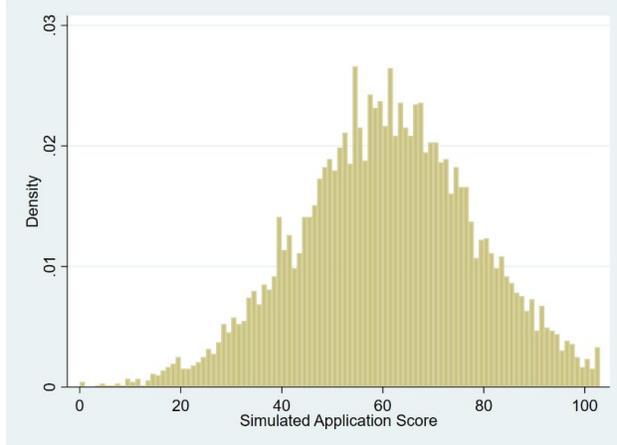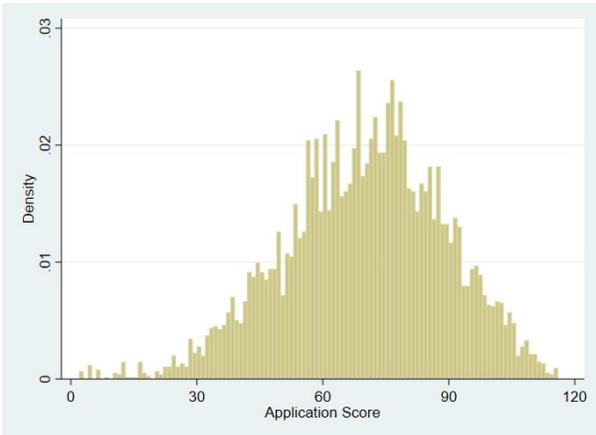Actual Distribution   Dropping Observations at the Cut-off
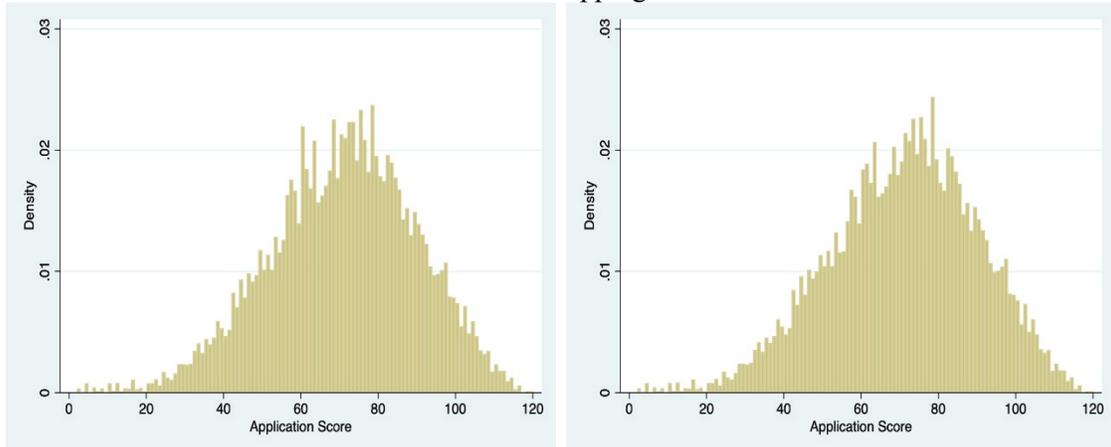
Figure B2
Application Score Distributions: Application Year 2006
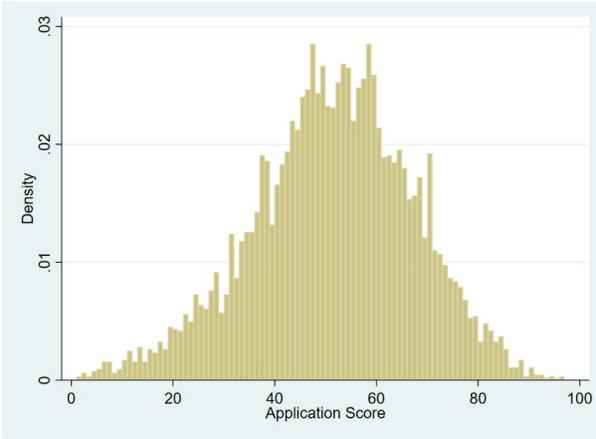
Full Sample
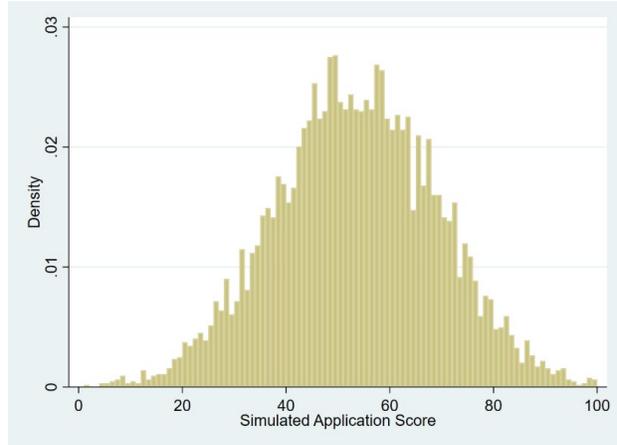
Simulated Sample

Figure B3
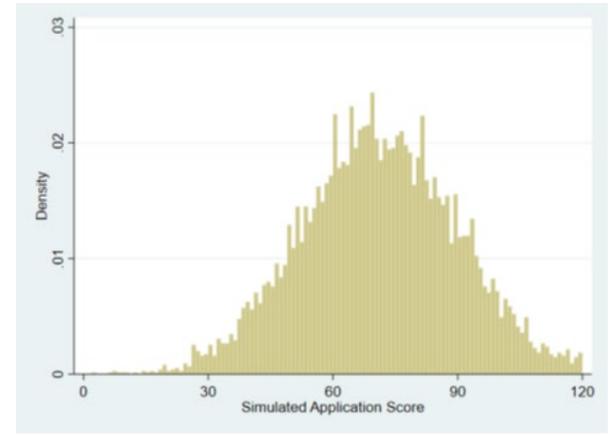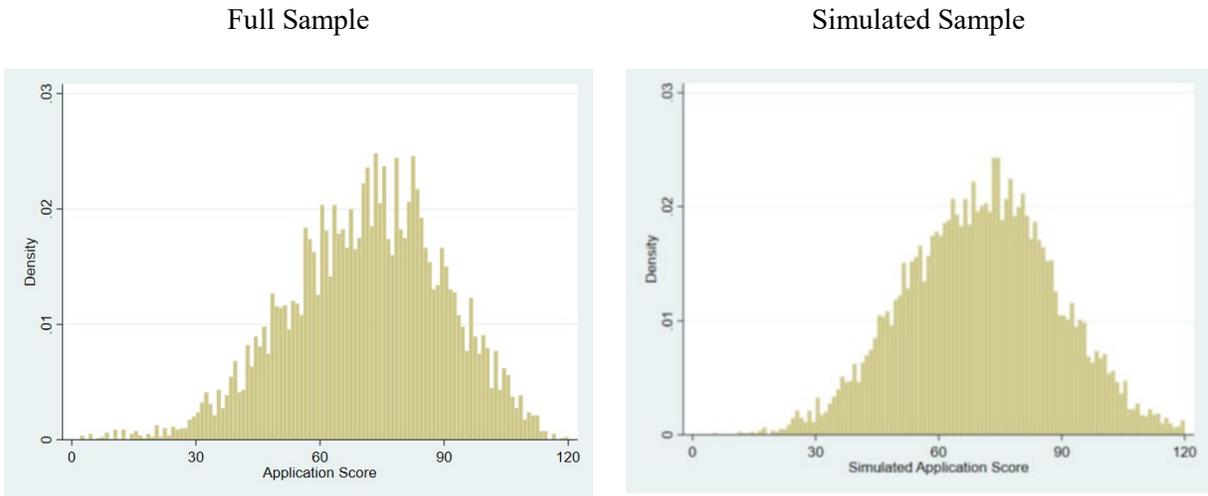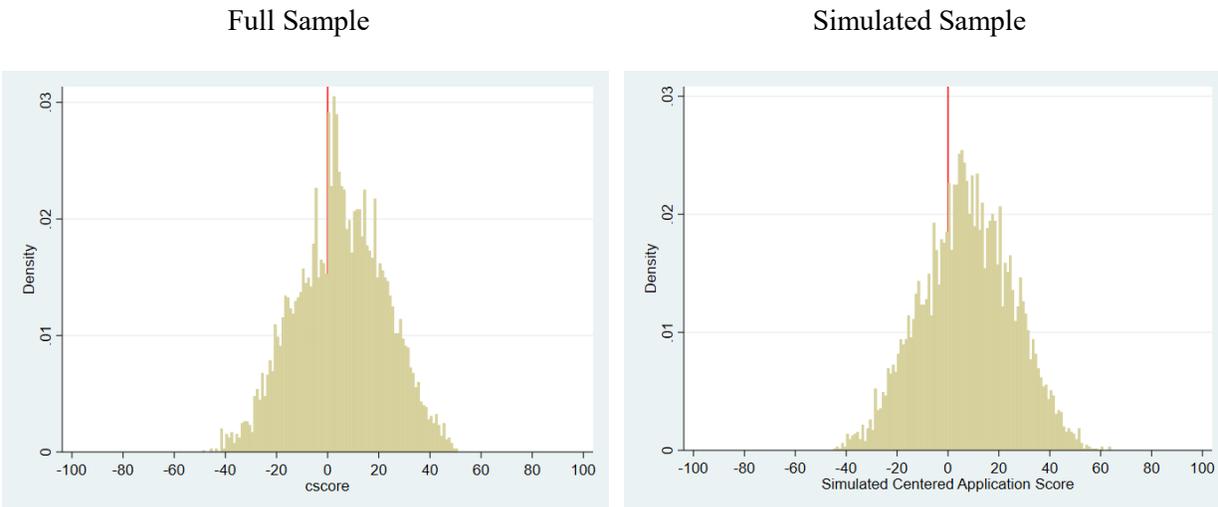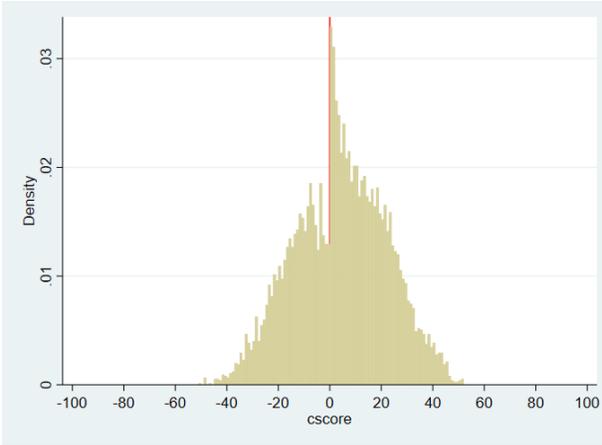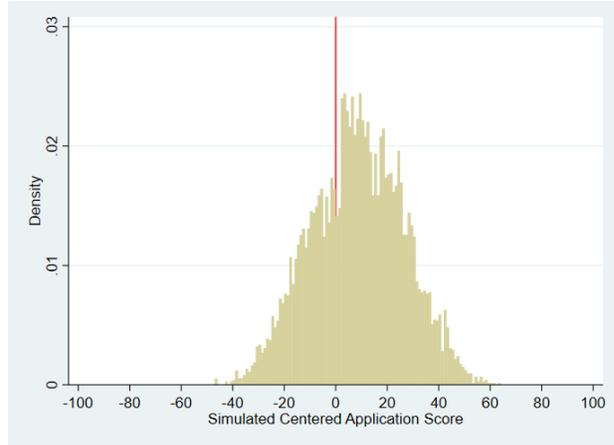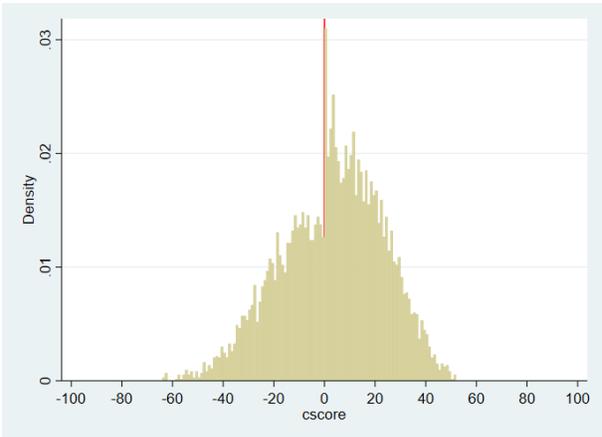Application Score Distributions: Application Year 2007
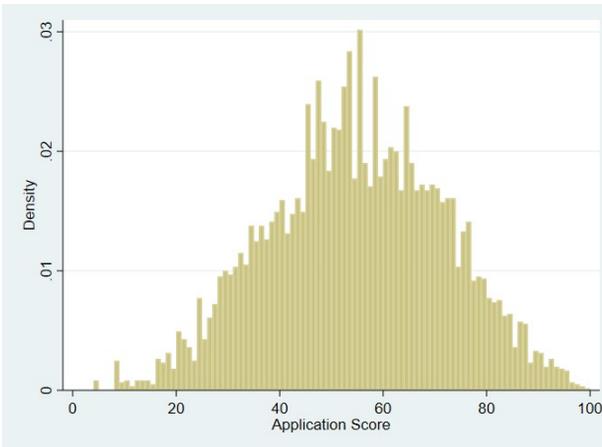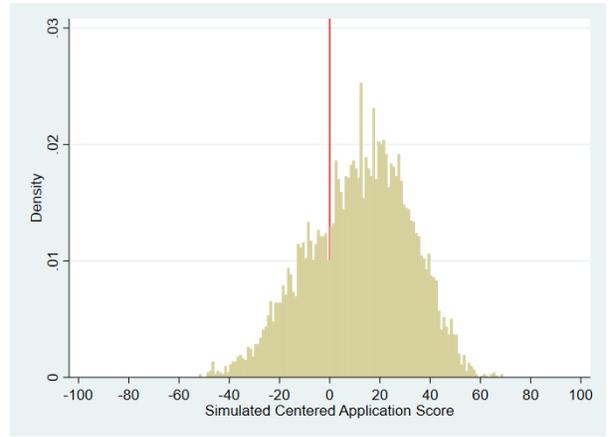
Full Sample
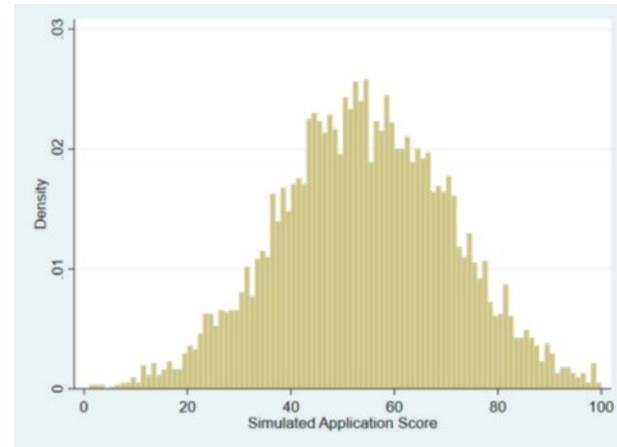
Simulated Sample

Figure B4
Application Score Distributions: Application Year 2008
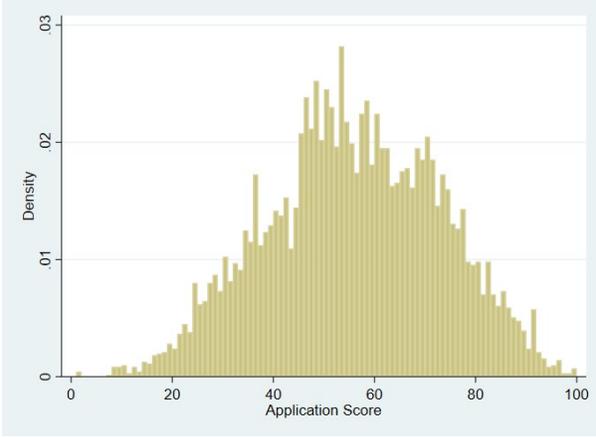
Full Sample
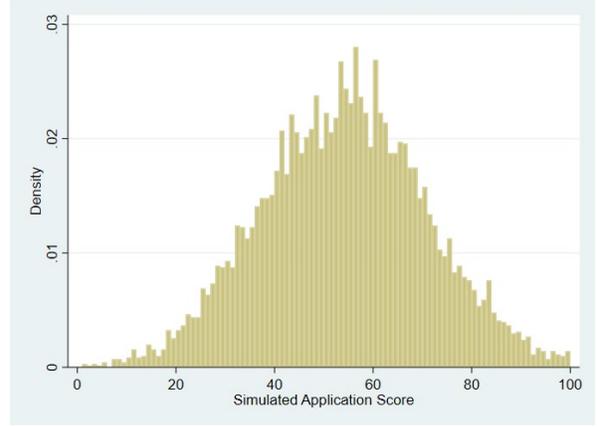
Simulated Sample

Figure B5
Application Score Distributions: Application Year 2009
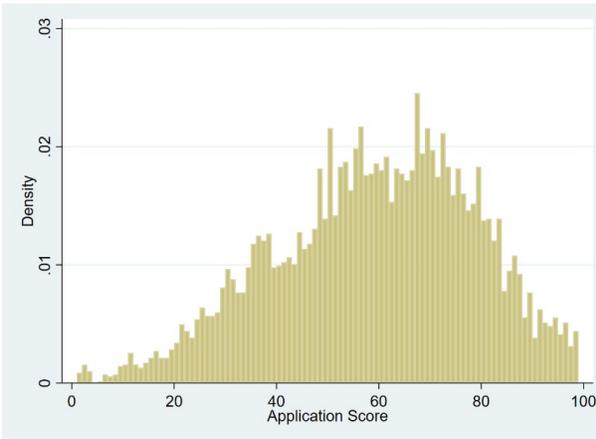
Full Sample
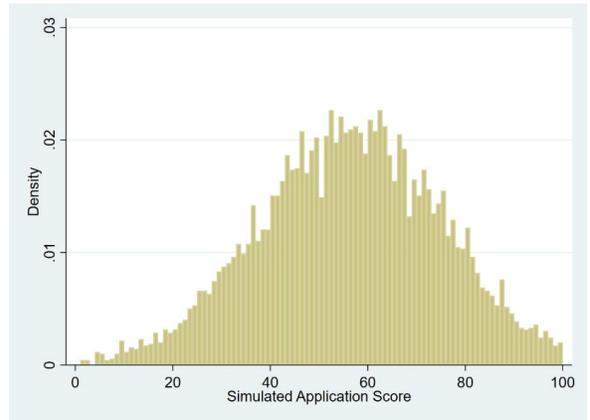
Simulated Sample

Figure B6
Application Score Distributions: Application Year 2010
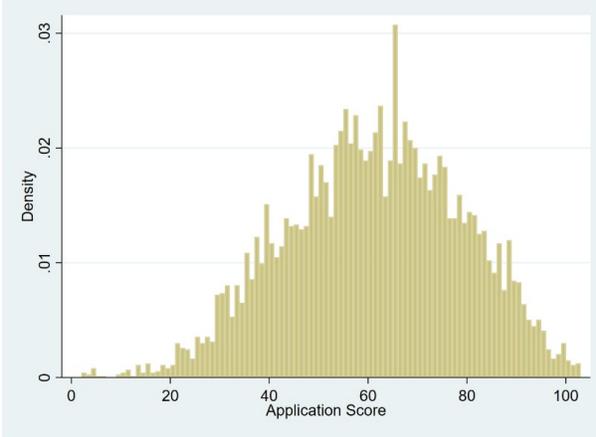
Full Sample

Simulated Sample

Figure B7
Application Score Distributions: Application Year 2011

Full Sample

Simulated Sample

Figure B8
Application Score Distributions: Application Year 2012

Full Sample   Simulated Sample

Figure B9
Application Score Distributions: Application Year 2013

Full Sample   Simulated Sample

44
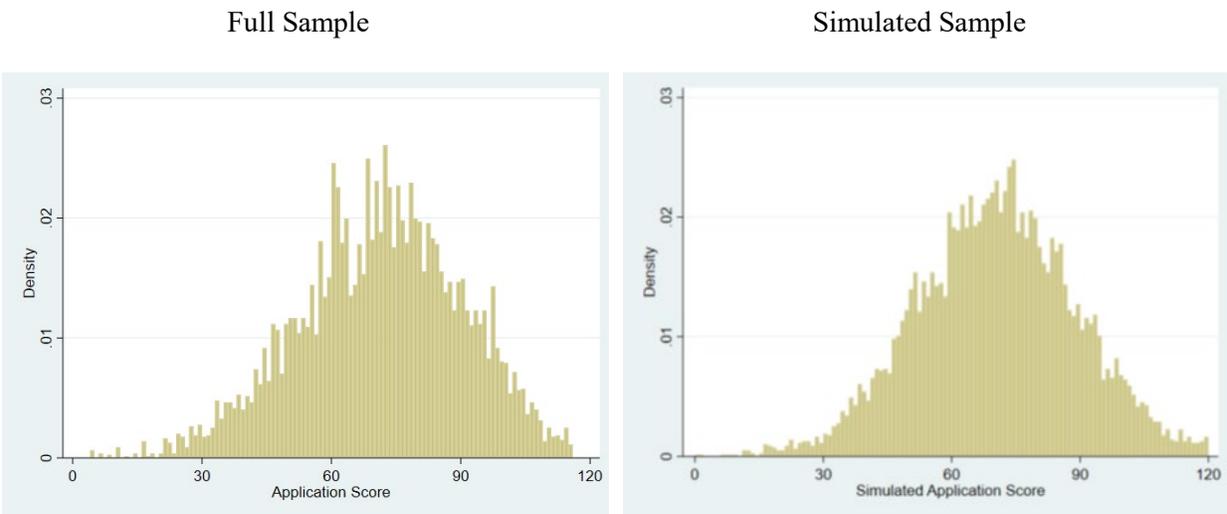
Figure B10
Centered Score Distributions: Application Year 2006

Full Sample
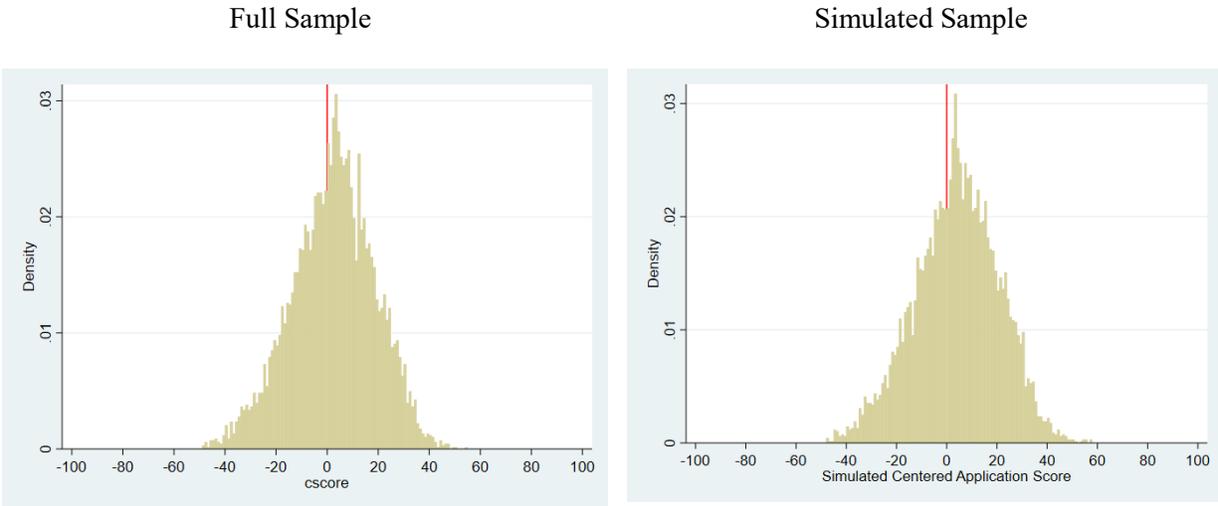
Simulated Sample

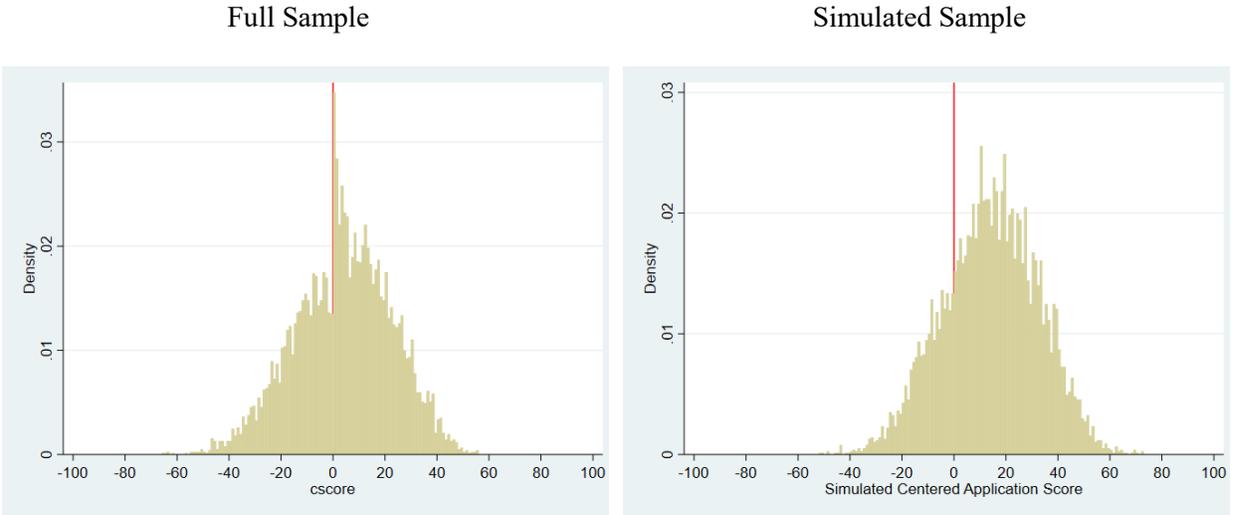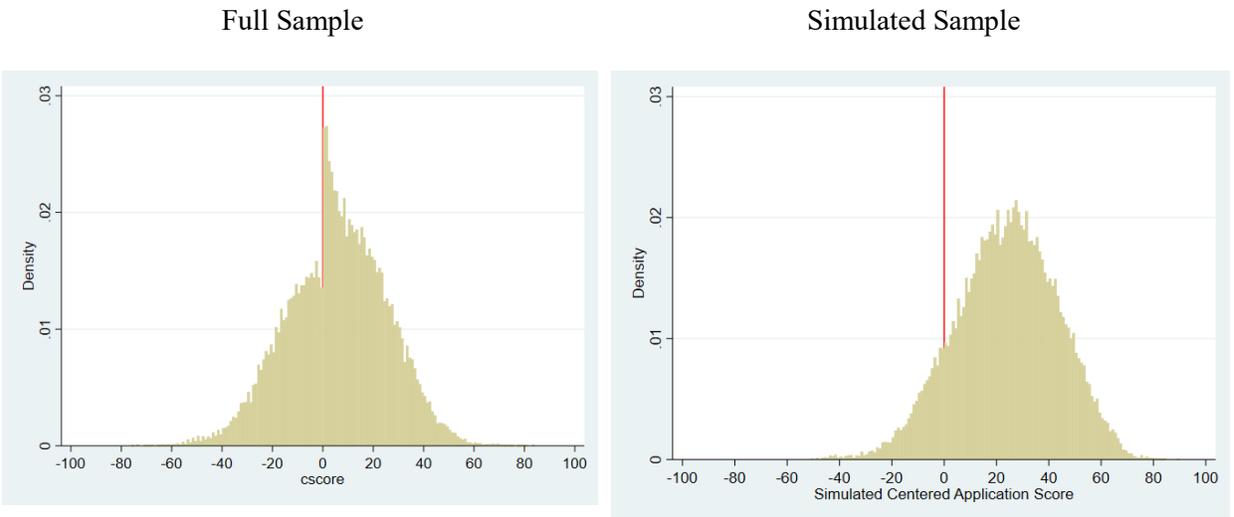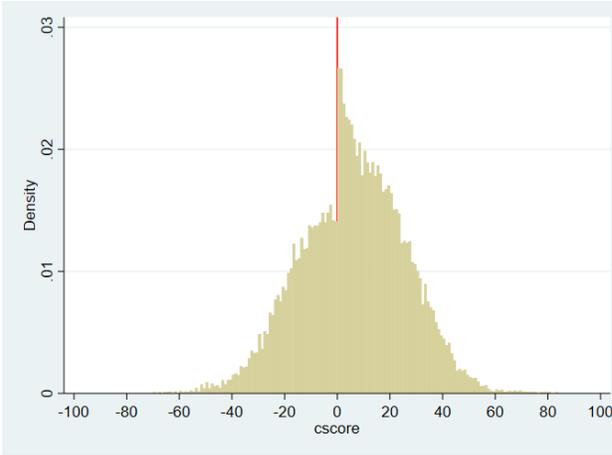Figure B11
Centered Score Distributions: Application Year 2007
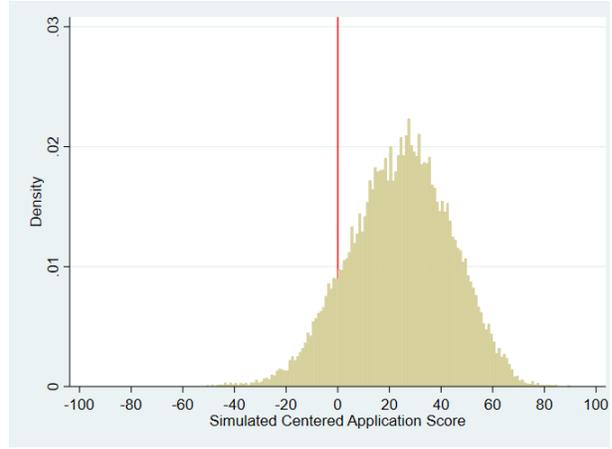
Full Sample

Simulated Sample

Figure B12
Centered Score Distributions: Application Year 2008

Full Sample

Simulated Sample

Figure B13
Centered Score Distributions: Application Year 2009

Full Sample

Simulated Sample

Figure B14
Centered Score Distributions: Application Year 2010

Full Sample

Simulated Sample

Figure B15
Centered Score Distributions: Application Year 2011

Full Sample

Simulated Sample

Figure B16
Centered Score Distributions: Application Year 2012
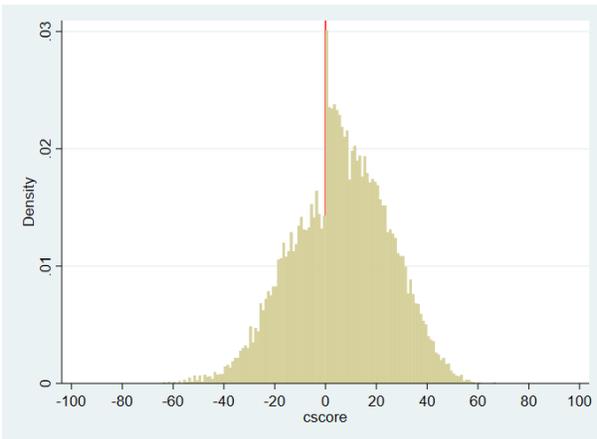
Full Sample

Simulated Sample

Figure B17
Centered Score Distributions: Application Year 2013

Full Sample

Simulated Sample