A Appendix A: Additional Tables and Figures

Table A1: NHPS school district characteristics

	2015	2017
A. Demographics		
Female	49.4%	49.3%
Asian	2.3%	2.2%
Black	41.9%	39.3%
Hispanic	40.2%	43.4%
Other	0.9%	1.4%
White	14.8%	13.9%
B. Student status		
Free/reduced meals eligible	58.7%	51.0%
English learners	14.2%	15.2%
Individualized education program	12.5%	13.5%

Notes: School district characteristics in academic year 2014-2015 ('2015') and 2016-2017 ('2017'). Reproduced from Connecticut Department of Education District Profile and Performance Reports, available online at edsight.ct.gov.

Table A2: Summary Statistics: Survey and Lottery Applications

School	Considered	In app.	1^{st} in app.	Most preferred	Revealed strategic	Mistaken strategic
Achievement First	64.8	9.1	0.6	3.8	3.5	1.1
Common Ground	55.7	16.7	4.9	4.2	1.7	0.7
Coop. Arts	63.6	22.5	10.9	19.2	11.5	2.5
Engineering & Sci	44.3	42.6	21.6	10.5	2.8	0.7
HS in the Community	54.5	20.1	2.4	3.5	1.7	0.7
Hill Regional	73.9	54.1	24.3	20.9	4.5	1.8
Hillhouse	75.0	6.6	0.5	1.7	1.2	1.2
Hyde	53.4	21.6	5.2	5.2	2.8	2.5
Metro Business	60.2	45.0	13.1	15.0	4.5	1.4
New Haven Academy	64.8	24.9	5.8	4.5	1.4	0.7
Riverside	39.8	3.0	1.2	0.0	0.0	0.0
Cross	66.7	40.9	15.9	13.0	3.5	1.8

	Probability		Marginal RSP		Capacity	
	2015	2017	2015	2017	2015	2017
Achievement First	93.5	14.1	(N,1)	(N,1)	28	2
Common Ground	20.9	5.6	(N,4)	(N,3)	56	70
Coop. Arts					147	155
Engineering & Sci	48.6	58.9	(N,1)	(N,1)	37	47
HS in the Community	61.7	94.2	(N,2)	(N,2)	69	68
Hill Regional	58.0	48.5	(N,1)	(N,1)	158	111
Hillhouse	100.0	88.3	(N,4)	(N,1)	17	12
Hyde	72.8	73.2	(N,2)	(N,2)	75	83
Metro Business	97.4	39.1	(N,1)	(N,1)	94	79
New Haven Academy	89.7	60.6	(N,2)	(N,1)	74	61
Riverside						
Cross	100.0	39.1	(N,4)	(N,1)	72	34

Notes: N=331 (87 in 2015, 244 in 2017) students in the survey who participated in the survey and matched to lottery data. Upper panel: All figures are percentages out of N. 'Considered' equals 1 when the respondent stated that he or she considered this school as a possible choice for his/her child and was only asked in 2015. 'In app.' displayes frequencies at which different schools appeared in lottery applications, while '1st in app' shows frequencies for first-ranked schools. 'Most preferred' referes to respondents' unconstrained first-choice school. 'Revealed strategic' show the rate at which respondents' unconstrained first-choice school was not ranked first on a lottery application. 'Mistaken strategic' tabulates the rate at which each school was played strategically but the RatEx odds of their first choice school were lower than the odds had they ranked their unconstrained first-choice first on the application. Lower panel: 'Probability' gives the odds, by year, that a student in the marginal round received a placement. 'Marginal RSP' is a pair describing the marginal report-specific priority. Y/N signify sibling priority (yes/no), while the number indicates the rank. Omitted for AF in 2017 (when no seats were available through the lottery) for Riverside (not assigned through main process) and for Coop. Arts (students may apply to different programs). 'Capacity' gives the number of seats available through assignment process at each school in each year. AF and Eng.& Sci. admit students through K-12 programs. See section 4.3 for details. Seat counts from Cross and Hillhouse are for non-neighborhood students.

Table A3: Demographic correlates of choice participation

	All surveyed			Choice participants	
	Participate	Place	Place MP	Place	Place MP
Black	-0.07	-0.05	-0.09	-0.04	-0.08
	(0.05)	(0.05)	(0.06)	(0.07)	(0.08)
White	-0.03	0.01	0.03	0.02	0.01
	(0.07)	(0.08)	(0.10)	(0.09)	(0.11)
Female	0.08	-0.02	-0.00	-0.05	-0.04
	(0.04)	(0.04)	(0.06)	(0.05)	(0.06)
Tract poverty rate	-0.17	-0.08	-0.52	-0.07	-0.45
	(0.19)	(0.21)	(0.25)	(0.25)	(0.28)
Dep. var. mean	0.807	0.581	0.369	0.720	0.434
N	358	358	312	289	265

Notes: Table displays regression results for regressions of a dummy indicating participation in the lottery ('Participate'), receiving placement at through the lottery ('Place') and placing at respondents' unconstrained first-choice school ('Place MP') on a demographic covariates. 'All surveyed' uses the full sample of survey respondents while 'Choice participants' conditions on those surveyed who participated in choice. Neighborhood school and year fixed effects included in all regressions (not shown). Robust standard errors in parentheses. See section 3.7 for details.

Table A4: Beliefs and application choices, conditional on preferences and first-listed schools

	(1) State 1st listed as MP	(2) State 1st listed as MP	(3) Placed	(4) Placed	(5) Placed
Subjective belief		-0.0321			0.165
		(0.0986)			(0.107)
RatEx	0.0455	0.0495	0.970	0.939	0.935
	(0.190)	(0.190)	(0.0426)	(0.146)	(0.147)
Placed	0.0603	0.0630			
	(0.0748)	(0.0755)			
Constant	0.644	0.657	0.00821	0.0680	-0.0197
	(0.105)	(0.115)	(0.0313)	(0.107)	(0.126)
Model test	0.421	0.692	0.218	0.611	0.311
N	186	186	2,101	186	186

Notes: Robust standard errors in parentheses. Specifications (1) and (2) include fixed effects for survey respondents' most-preferred schools. Specifications (3) through (5) include fixed effects for lottery applicants' first-listed schools. Model test displays p-values for a variety of statistical tests: (1) Placed=0 (2) Placed=0,Subjective belief=0 (3)-(4) RatEx=1, constant=0 (5) Subjective belief, RatEx=1, constant=0.

Table A5: Correlates of belief errors (cont.)

	D. Strategies		E. Participant characteristics		F. Recall	
	Optimism	Abs. Error	Optimism	Abs. Error	Optimism	Abs. Error
Hypothetical rank 2	41.6	10.6	41.7	10.5	41.7	11.4
	(1.4)	(1.9)	(1.4)	(1.8)	(1.3)	(1.7)
Have priority	-27.6	3.0	-24.0	5.7	-25.1	5.8
	(6.7)	(5.9)	(7.1)	(4.8)	(7.2)	(4.8)
Revealed strategic	3.7	-2.3				
	(5.0)	(2.6)				
Mistaken strategic	-5.4	4.7				
	(6.4)	(3.2)				
Mother			-3.8	-0.3		
			(4.6)	(2.3)		
Helped with application			-0.2	-0.1		
			(4.7)	(2.5)		
Correctly recall application					-0.1	-1.4
					(4.0)	(2.0)
N	941	941	1,045	1,045	1,155	1,155

Standard errors in parentheses. Errors clustered at the student level. Sample sizes change across panels due to covariate availability. All regressions include year fixed effects and exclude neighborhood schools from the sample. Correctly recall application is a dummy equal to one if a student both participated in the lottery and can correctly recall their first-listed school. See section 3.7 for additional description.

Table A6: Source of information and belief errors

A. Information sources	Mean	Optimism	Abs. Error
Visit fair	0.41	−3.4	−1.0
		(6.2)	(3.1)
Visit school	0.51	0.5	3.2
		(6.2)	(3.0)
Visit website	0.57	−6.9	−4.8
		(6.2)	(3.1)
Talk to teacher	0.54	0.6	−0.5
		(6.2)	(3.0)
Talk to counselor	0.47	−1.2	−2.1
		(6.2)	(3.2)
Talk to friend	0.42	1.6	−0.2
		(6.2)	(3.1)
Read catalog	0.65	−5.9	−4.8
		(6.7)	(3.1)
Read newspaper	0.25	5.3	−0.9
		(7.4)	(4.0)
Looked up capacity	0.24	1.7	−1.4
		(4.0)	(2.0)
Any admin. source	0.88	−12.8	−1.7
		(11.7)	(6.5)

B. Strategic play	Mean	Strategic	Mistaken strategic
Understand ranking penalty	0.23	3.9	3.8
		(6.7)	(5.7)
Understand priorities	0.12	−14.3	−6.6
		(8.4)	(6.8)
Understand both	0.04	1.5	2.0
		(15.3)	(13.0)

Panel A: Cells display independent variable means, regression coefficients, and standard errors from separate regressions of belief errors (column titles) against each information source, with controls for rank, priority, and year, when appropriate. Standard errors are clustered at the student level. Panel B: within-year bivariate regressions of indicators for strategic and mistaken strategic on indicators for understanding priorities, ranking penalty, and both mechanisms. Robust standard errors in parentheses. See section 3.7 for details.

Table A7: Probability of enrolling in a placed school: linear probability models

	Survey Only		Full Sample	
	(1)	(2)	(3)	(4)
Most-preferred	18.5	18.3		
	(6.4)	(6.2)		
Zoned to Cross	3.6	-0.9	-4.3	-3.7
	(6.4)	(6.7)	(2.5)	(2.6)
HS distance (km)			-0.7	0.8
			(0.4)	(0.7)
Dep. var. mean	70.05	70.05	69.02	69.02
N	207	207	1,388	1,388
School FE	No	Yes	No	Yes

Linear probability models of enrollment in a school in the year following the lottery, conditional on being placed in a school, on a survey-elicited dummy indicating whether a school is a student's most-preferred ('Most preferred'), as well as a control for students' default schools ('Zoned to Cross') and a year effect (not shown). Robust standard errors in parentheses.

Table A8: Median and 90% credible intervals for belief model covariances Σ^{η}.

		Low SES	
	η_0	η_{pri}	η_{round}
2015			
η_0	(64.22,71.18, 80.48)		
η_{pri}	(3.2,4.87, 6.97)	(6.03,6.72, 8.43)	
η_{round}	(-28.73,-25.53, -22.95)	(-2.52,-1.8, -1.16)	(8.22,9.2, 10.32)
2017			
η_0	(33.33,37.0, 41.82)		
η_{pri}	(7.15,11.14, 16.41)	(2.08,3.86, 6.94)	
η_{round}	(-1.46,-0.85, -0.41)	(-0.52,-0.28, -0.09)	(0.06,0.09, 0.13)
		High SES	
	η_0	η_{pri}	η_{round}
2015			
η_0	(69.23,88.74, 121.18)		
η_{pri}	(-13.15,-2.24, 10.57)	(0.9,4.33, 12.53)	
η_{round}	(-40.43,-30.05, -23.2)	(-3.46,0.71, 4.44)	(7.83,10.29, 13.54)
2017			
η_0	(36.4,42.23, 48.99)		
η_{pri}	(21.74,27.32, 33.35)	(12.91,17.86, 25.13)	
η_{round}	(-1.94,-1.19, -0.36)	(-1.31,-0.74, -0.22)	(0.06,0.12, 0.18)

Notes: Median and 90% credible intervals for belief error covariance terms. Panels split by SES and year. See Sections 5 and 6 for details.

Table A9: 90% credible intervals for preference shocks, Σ

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
2015												
Achievement First Amistad HS (1)	(223.17, 2070.74)											
Common Ground Charter (2)	(-84.33, 105.69)	(256.13, 2375.47)										
Coop. Arts and Humanities (3)	(-87.85, 107.02)	(-108.73, 93.22)	(267.52, 2540.02)									
Engineering & Science Univ. HS (4)	(-16.97, 225.89)	(-110.18, 100.22)	(-257.22, 7.14)	(269.89, 2459.05)								
High School in the Community (5)	(-74.0, 101.47)	(-89.92, 104.56)	(-109.94, 77.16)	(-76.28, 138.63)	(227.29, 2088.06)							
Hill Regional Career (6)	(-49.97, 144.17)	(-183.64, 42.38)	(-105.5, 69.46)	(-110.6, 81.98)	(-88.45, 85.13)	(260.56, 2395.17)						
Hillhouse (7)	(-91.52, 101.8)	(-71.88, 145.64)	(-82.96, 126.09)	(-119.84, 107.21)	(-91.51, 97.83)	(-137.56, 70.69)	(237.64, 2228.93)					
Hyde School (8)	(-37.71, 166.06)	(-110.88, 102.62)	(-257.11, 5.58)	(-38.44, 190.9)	(-37.58, 155.82)	(-8.2, 228.78)	(-71.18, 141.13)	(243.33, 2216.07)				
Metropolitan Business Academy (9)	(-72.09, 99.8)	(-150.58, 52.16)	(-123.58, 54.94)	(-2.91, 260.33)	(-25.34, 171.21)	(12.02, 294.31)	(-166.46, 46.48)	(-49.65, 125.44)	(224.88, 2154.97)			
New Haven Academy (10)	(-44.69, 149.79)	(-86.35, 108.51)	(-102.98, 69.08)	(-43.26, 156.29)	(-32.04, 157.15)	(-124.12, 50.28)	(-88.25, 97.3)	(-52.27, 136.68)	(-17.59, 190.36)	(227.91, 2135.45)		
Riverside Education Academy (11)	(-133.78, 84.36)	(-73.07, 176.65)	(-132.87, 101.71)	(-173.38, 82.08)	(-103.12, 104.48)	(-229.13, 28.43)	(-109.92, 119.12)	(-161.78, 68.46)	(-169.26, 42.02)	(-133.57, 75.64)	(267.03, 2482.9)	
Wilbur L. Cross High School (12)	(-132.43, 77.01)	(-140.86, 107.74)	(-86.55, 133.35)	(-138.92, 107.79)	(-114.45, 92.22)	(-141.47, 75.12)	(-117.44, 116.5)	(-199.1, 35.97)	(-171.03, 43.72)	(-168.33, 48.54)	(-84.14, 162.87)	(274.25, 2510.05)
2017												
Achievement First Amistad HS (1)	(141.38, 871.32)											
Common Ground Charter (2)	(-40.82, 55.89)	(159.79, 986.55)										
Coop. Arts and Humanities (3)	(-8.05, 99.07)	(-42.12, 53.74)	(164.43, 1008.67)									
Engineering & Science Univ. HS (4)	(-13.67, 94.5)	(-53.0, 45.0)	(-62.23, 31.72)	(164.29, 1022.42)								
High School in the Community (5)	(-37.68, 45.12)	(-32.1, 62.51)	(-47.6, 36.74)	(-40.05, 48.87)	(138.5, 859.39)							
Hill Regional Career (6)	(-36.11, 50.51)	(-87.33, 14.77)	(-54.0, 28.21)	(-94.16, 6.83)	(-29.63, 51.7)	(155.76, 987.91)						
Hillhouse (7)	(-42.12, 49.2)	(-61.77, 39.31)	(-44.48, 53.12)	(-58.3, 41.21)	(-51.0, 39.57)	(-35.36, 58.5)	(145.95, 900.31)					
Hyde School (8)	(-35.9, 48.92)	(-46.9, 47.15)	(-81.91, 13.87)	(-23.38, 69.59)	(-32.15, 50.17)	(-10.47, 90.31)	(-45.33, 44.95)	(140.99, 881.93)				
Metropolitan Business Academy (9)	(-49.22, 34.29)	(-62.38, 29.79)	(-68.82, 18.82)	(-64.73, 25.63)	(-16.06, 68.63)	(-0.12, 102.7)	(-48.4, 41.36)	(-11.71, 77.29)	(147.66, 919.11)			
New Haven Academy (10)	(-25.91, 63.52)	(-61.87, 32.67)	(-10.68, 82.59)	(-39.73, 51.93)	(-48.99, 35.88)	(-71.64, 14.12)	(-33.43, 60.77)	(-33.75, 51.34)	(-18.16, 64.7)	(149.05, 908.62)		
Riverside Education Academy (11)	(-54.23, 47.13)	(-66.54, 56.82)	(-66.95, 55.77)	(-64.69, 48.66)	(-56.93, 44.03)	(-71.84, 42.26)	(-55.46, 56.17)	(-56.42, 42.05)	(-65.66, 38.39)	(-46.72, 63.37)	(168.23, 1055.91)	
Wilbur L. Cross High School (12)	(-44.97, 40.5)	(-23.64, 69.84)	(-79.13, 15.7)	(-26.68, 67.76)	(-18.28, 75.13)	(-43.13, 44.06)	(-50.42, 43.2)	(-44.25, 33.71)	(-55.22, 29.68)	(-63.92, 22.11)	(-60.3, 41.74)	(143.9, 899.39)

Notes: 90% credible intervals for preference covariance matrix. See Sections 5 and 6 for details.

Table A10: Distance-Metric Welfare: Benchmark and Counterfactuals, 2015

	Mean welfare			Welfare differences		
	Baseline	RatEx	DA	RatEx − Baseline	DA − Baseline	No Survey DA − Baseline
A1. Posterior distribution of mean distance-metric welfare						
Mean	9.942	14.985	13.882	5.043	3.941	−1.653
Median	8.878	13.242	12.328	4.463	3.507	−1.047
95% CI	[5.353, 23.409]	[8.083, 35.345]	[7.439, 32.552]	[2.661, 11.588]	[2.100, 8.908]	[−5.455, −0.507]
A2. High-SES mean minus low-SES mean						
Mean	−2.552	−3.740	−3.643	−1.188	−1.091	1.127
Median	−2.308	−3.342	−3.279	−1.044	−0.954	0.656
95% CI	[−5.381, −1.001]	[−8.056, −1.722]	[−7.713, −1.679]	[−2.555, −0.324]	[−2.425, −0.276]	[0.225, 3.882]

	Truthful	Strategic	Drops	Stops
B. DA-4 - baseline under different strategy types				
Mean	4.018	4.239	4.018	4.016
Median	3.586	3.782	3.586	3.593
95% CI	[2.140, 8.970]	[2.263, 9.701]	[2.140, 8.968]	[2.140, 8.949]

	0%	25%	50%	75%	100%
C. Share submitting baseline application under DA-4					
Mean	4.018	2.982	1.947	0.916	−0.142
Median	3.586	2.656	1.745	0.805	−0.115
95% CI	[2.140, 8.970]	[1.625, 6.632]	[1.033, 4.260]	[0.451, 2.198]	[−0.430, 0.043]

	Switch to DA		Keep baseline mechanism	
	School and priority	School	School and priority	School
D. Eliminate specific error components under DA-4 and baseline				
Mean	3.165	3.173	3.256	3.261
Median	2.825	2.817	2.909	2.910
95% CI	[1.723, 7.036]	[1.728, 7.053]	[1.778, 7.273]	[1.790, 7.291]

Notes: This table describes the posterior distribution of mean welfare in the baseline case and under policy counterfactuals for 2015 households. Welfare is measured using miles traveled as the numeraire good. Panels A1 and A2: 'Baseline' is baseline mechanism given observed beliefs. 'RatEx' is the baseline mechanism under rational expectations beliefs. 'DA' is the strategy-proof deferred acceptance mechanism. 'RatEx-baseline' and 'DA-baseline' columns compare welfare differences under the listed mechanisms. 'No survey DA-base' column compares welfare under the sophisticated DA and baseline mechanisms using model estimates based on rational expectations beliefs. Panel A2 displays differences in each of these objects between high-SES and low-SES households. Panel B: difference between DA welfare and baseline welfare under 'drop' and 'stop' DA play (columns 1-4) and sophisticated truncated DA-4. See text for details. Panel C: Welfare gain from switch to DA from baseline by share of households continuing to submit 'baseline' applications. See text for details. Panel D: Welfare change from switch to DA from baseline under strategic truncated DA with school- and school by priority-specific errors (columns 1+2), and welfare change from switching to only school- and school by priority-specific errors while keeping the baseline mechanism. See text for details.

Table A11: Distance-Metric Welfare: Benchmark and Counterfactuals, 2017

	Mean welfare			Welfare differences		
	Baseline	RatEx	DA	RatEx − Baseline	DA − Baseline	No Survey DA − Baseline
A1. Posterior distribution of mean distance-metric welfare						
Mean	18.899	22.811	22.809	3.912	3.910	−1.950
Median	17.479	21.271	21.074	3.585	3.601	−1.466
95% CI	[11.945, 33.064]	[14.251, 39.515]	[14.310, 39.742]	[2.185, 7.269]	[2.351, 6.987]	[−6.363, −0.778]
A2. High-SES mean minus low-SES mean						
Mean	−3.222	−3.795	−3.988	−0.573	−0.767	0.162
Median	−3.133	−3.623	−3.806	−0.530	−0.711	0.123
95% CI	[−6.054, −1.293]	[−7.084, −1.408]	[−7.363, −1.544]	[−1.704, 0.323]	[−1.936, 0.234]	[−0.453, 1.109]

	Truthful	Strategic	Drops	Stops
B. DA-4 - baseline under different strategy types				
Mean	2.891	3.052	2.868	2.889
Median	2.637	2.812	2.631	2.636
95% CI	[1.829, 5.065]	[1.794, 5.507]	[1.826, 5.035]	[1.827, 5.061]

	0%	25%	50%	75%	100%
C. Share submitting baseline application under DA-4					
Mean	2.891	2.421	1.946	1.446	0.938
Median	2.637	2.230	1.794	1.334	0.860
95% CI	[1.829, 5.065]	[1.507, 4.223]	[1.156, 3.445]	[0.856, 2.522]	[0.545, 1.696]

	Switch to DA		Keep baseline mechanism	
	School and priority	School	School and priority	School
D. Eliminate specific error components under DA-4 and baseline				
Mean	1.451	1.447	0.218	0.228
Median	1.325	1.325	0.200	0.214
95% CI	[0.859, 2.606]	[0.852, 2.611]	[−0.007, 0.545]	[0.007, 0.541]

Notes: This table describes the posterior distribution of mean welfare in the baseline case and under policy counterfactuals for 2017 households. Welfare is measured using miles traveled as the numeraire good. Panels A1 and A2: 'Baseline' is baseline mechanism given observed beliefs. 'RatEx' is the baseline mechanism under rational expectations beliefs. 'DA' is the strategy-proof deferred acceptance mechanism. 'RatEx-baseline' and 'DA-baseline' columns compare welfare differences under the listed mechanisms. 'No survey DA-base' column compares welfare under the sophisticated DA and baseline mechanisms using model estimates based on rational expectations beliefs. Panel A2 displays differences in each of these objects between high-SES and low-SES households. Panel B: difference between DA welfare and baseline welfare under 'drop' and 'stop' DA play (columns 1-4) and sophisticated truncated DA-4. See text for details. Panel C: Welfare gain from switch to DA from baseline by share of households continuing to submit 'baseline' applications. See text for details. Panel D: Welfare change from switch to DA from baseline under strategic truncated DA with school- and school by priority-specific errors (columns 1+2), and welfare change from switching to only school- and school by priority-specific errors while keeping the baseline mechanism. See text for details.

Table A12: Distance-Metric Welfare: Benchmark and Counterfactuals, Accurate Recall Only

	Mean welfare			Welfare differences		
	Baseline	RatEx	DA	RatEx − Baseline	DA − Baseline	No Survey DA − Baseline
A1. Posterior distribution of mean distance-metric welfare						
Mean	14.062	18.262	17.801	4.200	3.739	−1.801
Median	13.548	17.076	17.021	3.629	3.321	−1.211
95% CI	[5.481, 29.306]	[8.508, 37.679]	[7.876, 36.696]	[2.078, 9.137]	[2.143, 7.542]	[−6.165, −0.542]
A2. High-SES mean minus low-SES mean						
Mean	−2.814	−3.591	−3.601	−0.777	−0.788	0.644
Median	−2.503	−3.183	−3.231	−0.719	−0.725	0.410
95% CI	[−6.593, −1.011]	[−8.045, −1.420]	[−7.990, −1.414]	[−2.217, 0.305]	[−2.193, 0.126]	[−0.366, 2.228]

	Truthful	Strategic	Drops	Stops
B. DA-4 - baseline under different strategy types				
Mean	3.498	3.505	3.481	3.499
Median	3.071	3.051	3.052	3.073
95% CI	[1.883, 7.381]	[1.780, 7.727]	[1.875, 7.381]	[1.882, 7.445]

	0%	25%	50%	75%	100%
C. Share submitting baseline application under DA-4					
Mean	3.498	2.749	1.988	1.211	0.424
Median	3.071	2.406	1.777	1.135	0.274
95% CI	[1.883, 7.381]	[1.484, 5.699]	[1.112, 3.939]	[0.512, 2.593]	[−0.315, 1.647]

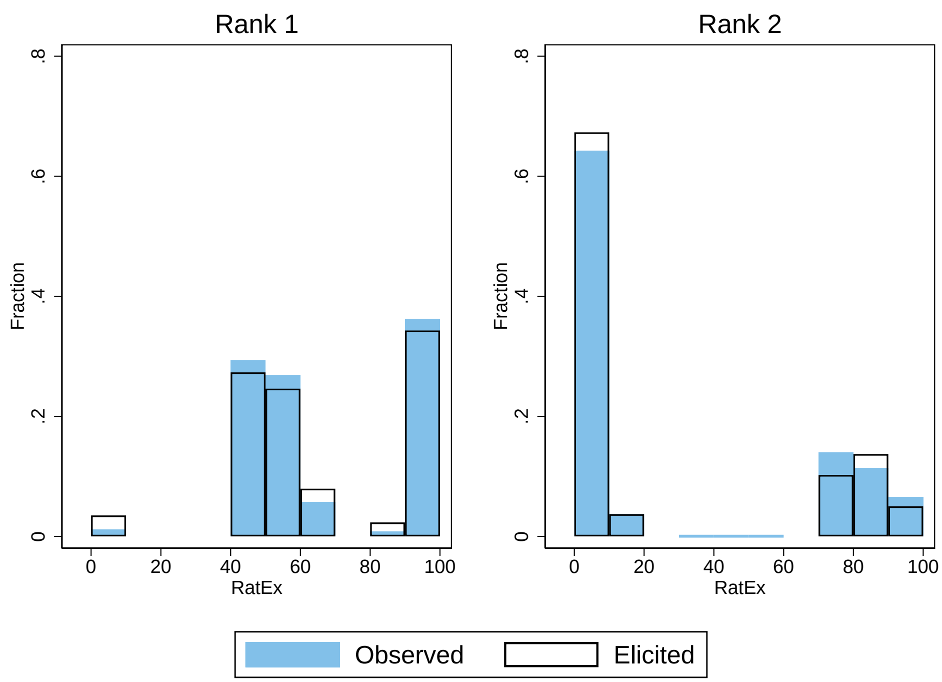
	Switch to DA		Keep baseline mechanism	
	School and priority	School	School and priority	School
D. Eliminate specific error components under DA-4 and baseline				
Mean	2.324	2.327	1.767	1.775
Median	1.945	1.945	1.368	1.368
95% CI	[0.927, 5.433]	[0.923, 5.469]	[0.018, 5.633]	[0.025, 5.603]

Notes: This table describes the posterior distribution of mean welfare in the baseline case and under policy counterfactuals. We restrict the survey data used for belief and preference estimation to the subset of respondents with correct recall of the submitted application. Welfare is measured using miles traveled as the numeraire good. Panels A1 and A2: 'Baseline' is baseline mechanism given observed beliefs. 'RatEx' is the baseline mechanism under rational expectations beliefs. 'DA' is the strategy-proof deferred acceptance mechanism. 'RatEx-baseline' and 'DA-baseline' columns compare welfare differences under the listed mechanisms. 'No survey DA-base' column compares welfare under the sophisticated DA and baseline mechanisms using model estimates based on rational expectations beliefs. Note that this is the same as reported in Table 7 because the survey is not used. Panel A2 displays differences in each of these objects between high-SES and low-SES households. Panel B: difference between DA welfare and baseline welfare under 'drop' and 'stop' DA play (columns 1-4) and sophisticated truncated DA-4. See text for details. Panel C: Welfare gain from switch to DA from baseline by share of households continuing to submit 'baseline' applications. See text for details. Panel D: Welfare change from switch to DA from baseline under strategic truncated DA with school- and school by priority-specific errors (columns 1+2), and welfare change from switching to only school- and school by priority-specific errors while keeping the baseline mechanism. See text for details.

Figure A1: Share of Students within each School Zone

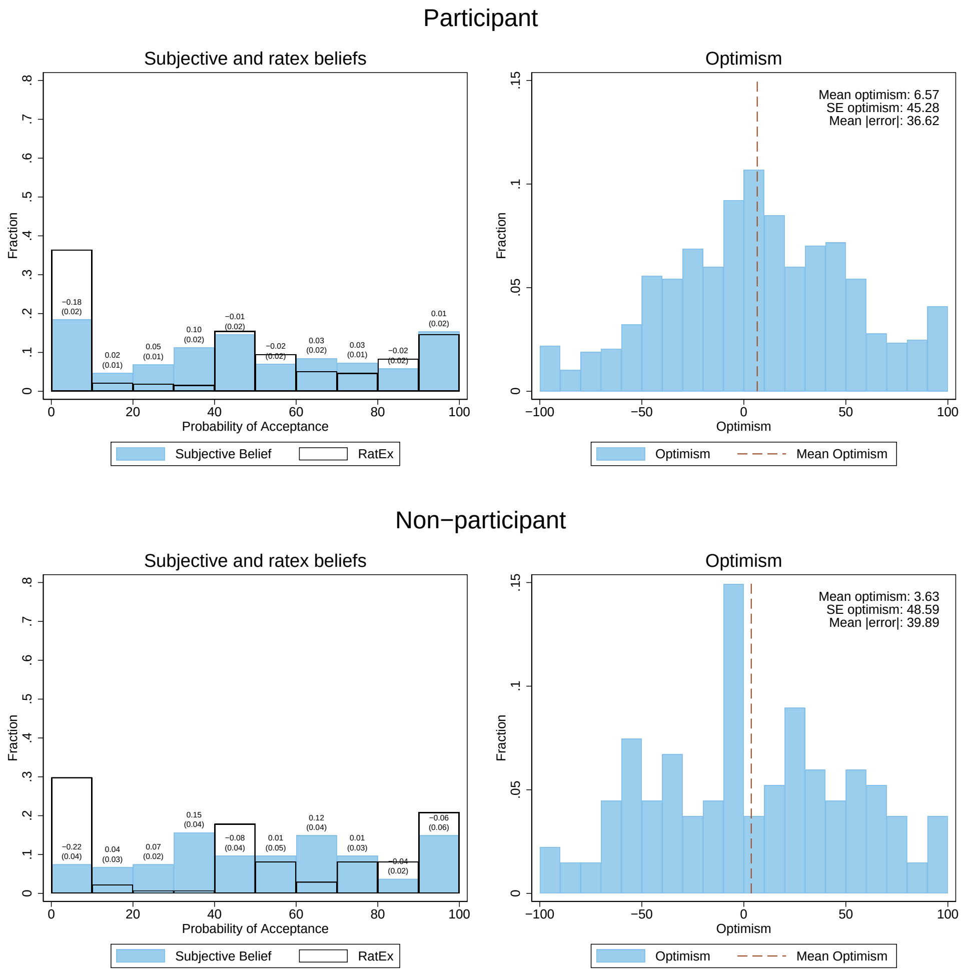
This figure displays the geographic distribution of sample universe and surveyed population. Size of circles reflect shares of population and surveyed individuals, respectively. Each point represents the physical centroid (as opposed to a within-tract population weighted centroid) of a census tract in the city of New Haven. Census tracts incorporate non-habitable features of the landscape and their centroids may lie in uninhabited areas.

Figure A2: Ratex admissions probabilities of actual and hypothetical applications



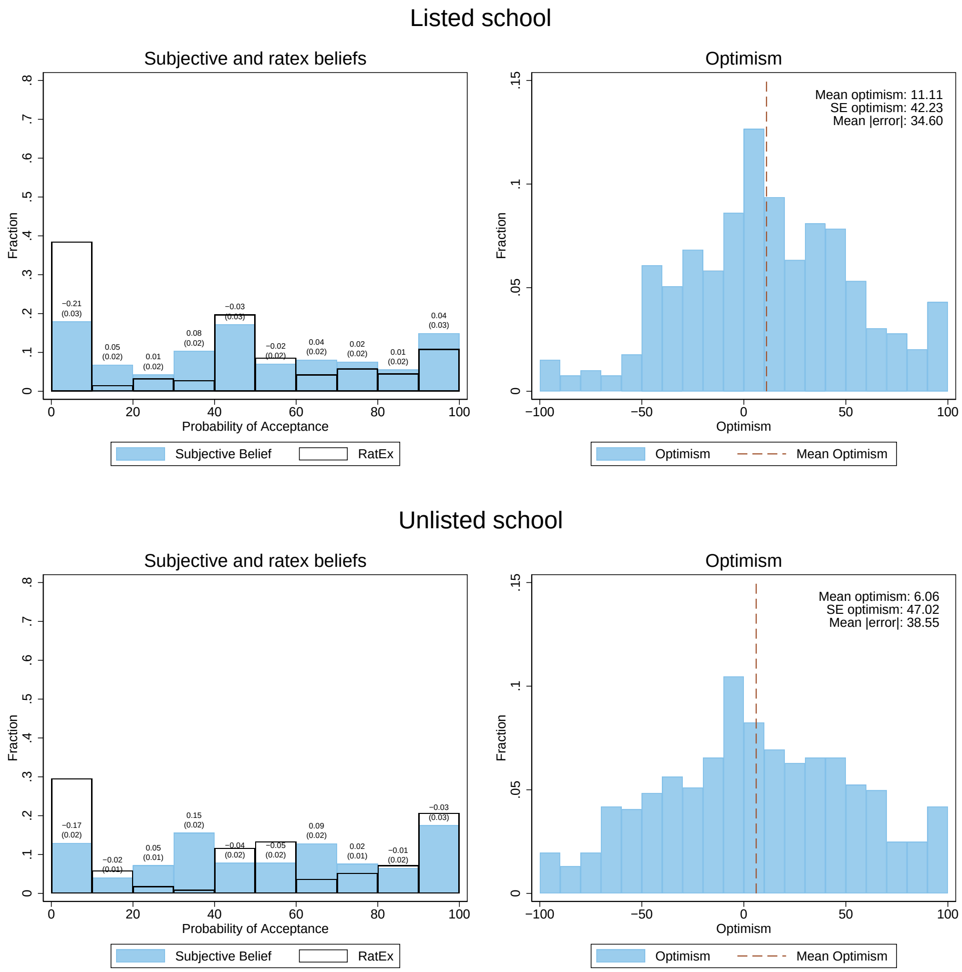
Notes: $N = 3,129$ for observed RatEx (1,567 for rank 1 schools, 1,562 for rank 2 schools), $N = 975$ for hypothetical RatEx (516 in rank 1 applications; 459 rank 2 applications). Sample of schools for which RatEx are tabulated is all schools except neighborhood schools and Co-Op Arts. For observed applications, the sample is the entire universe of lottery participants while the sample for elicited applications is hypothetical application-ranks in surveyed sample. Bins have width 0.10.

Figure A3: Ratex beliefs, subjective beliefs, and optimism by choice participation



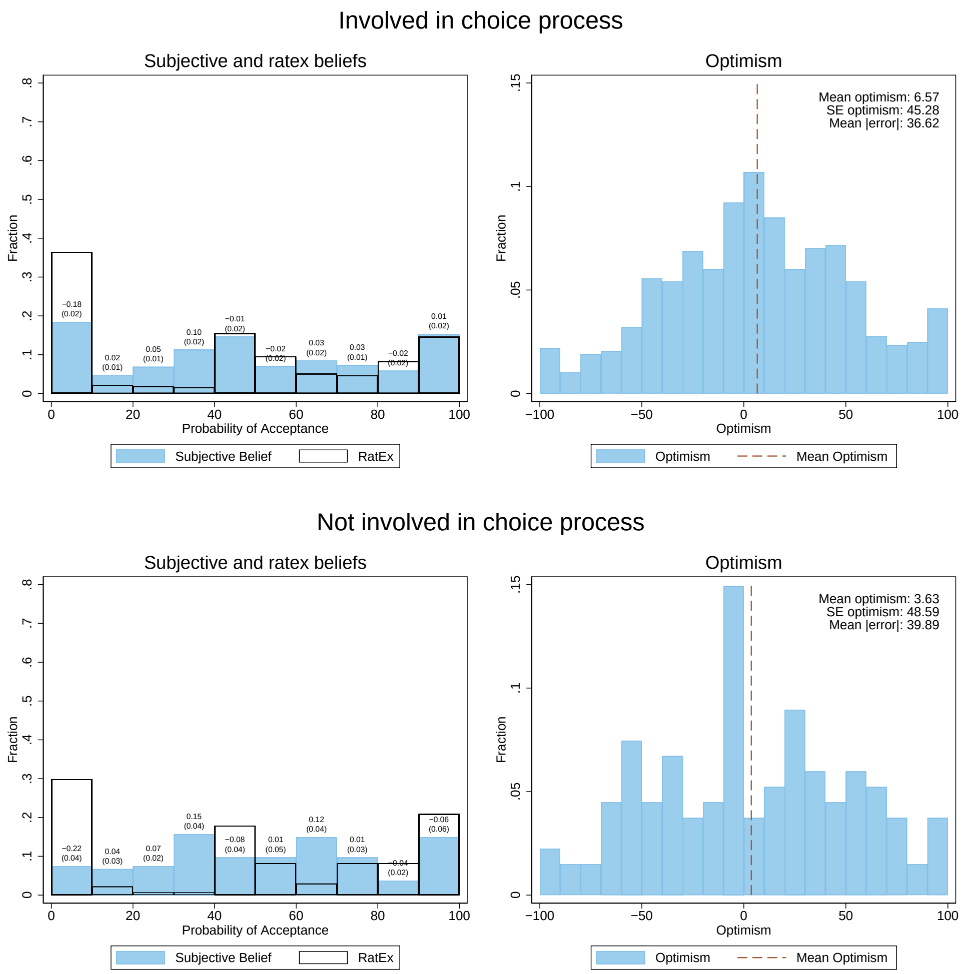
Upper graph: households participating in school choice. Lower graph: non-participants. Left panel: distribution of subjective and rational expectations assignment probabilities. Text reports gap in fraction of subjective reports and and RatEx values in the bin, with standard errors clustered at the respondent level in parentheses below. Right panel: distribution of optimism. Bars show shares of population within bins of width 10. Red line indicates mean of the distribution. In both panels, beliefs for second-ranked options are conditional on non-admission to the first-ranked choice.

Figure A4: Ratex beliefs, subjective beliefs, and optimism by whether school was listed



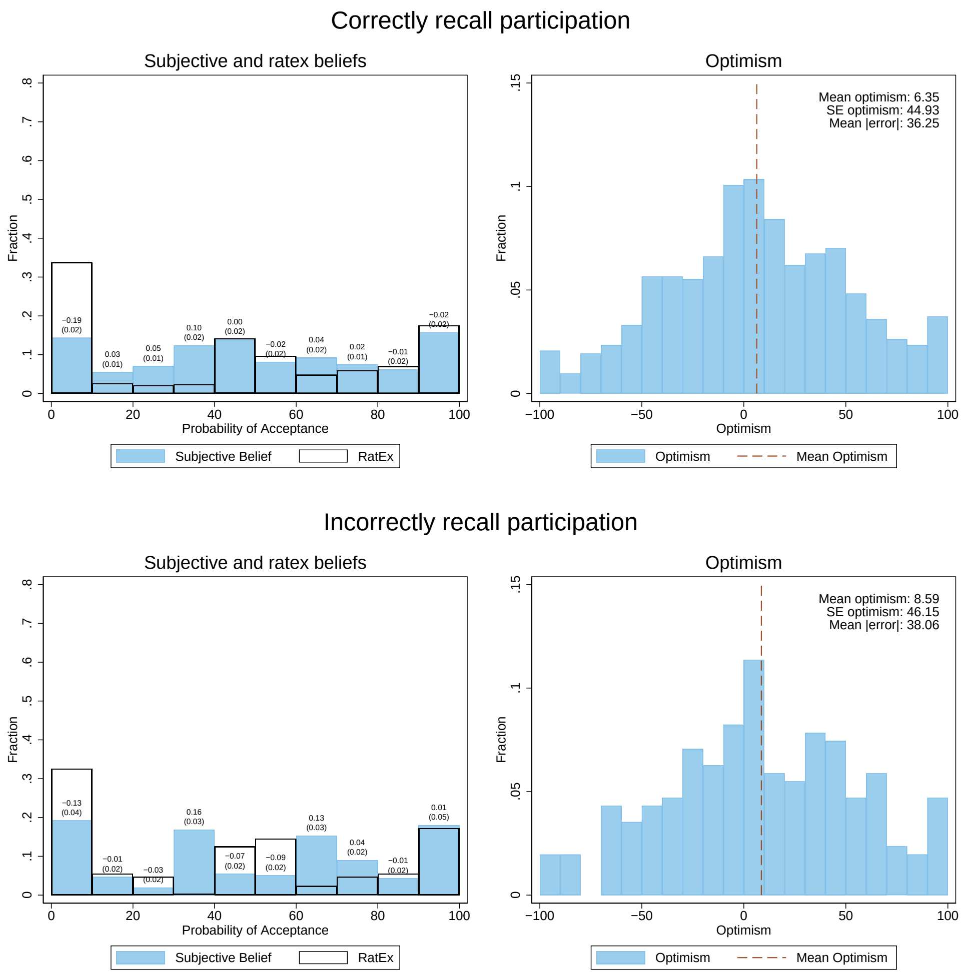
Upper graph: beliefs for schools listed on a household's school choice application. Lower graph: beliefs for unlisted schools. Left panel: distribution of subjective and rational expectations assignment probabilities. Text reports gap in fraction of subjective reports and and RatEx values in the bin, with standard errors clustered at the respondent level in parentheses below. Right panel: distribution of optimism. Bars show shares of population within bins of width 10. Red line indicates mean of the distribution. In both panels, beliefs for second-ranked options are conditional on non-admission to the first-ranked choice.

Figure A5: Ratex beliefs, subjective beliefs, and optimism by whether respondent was involved in school choice process



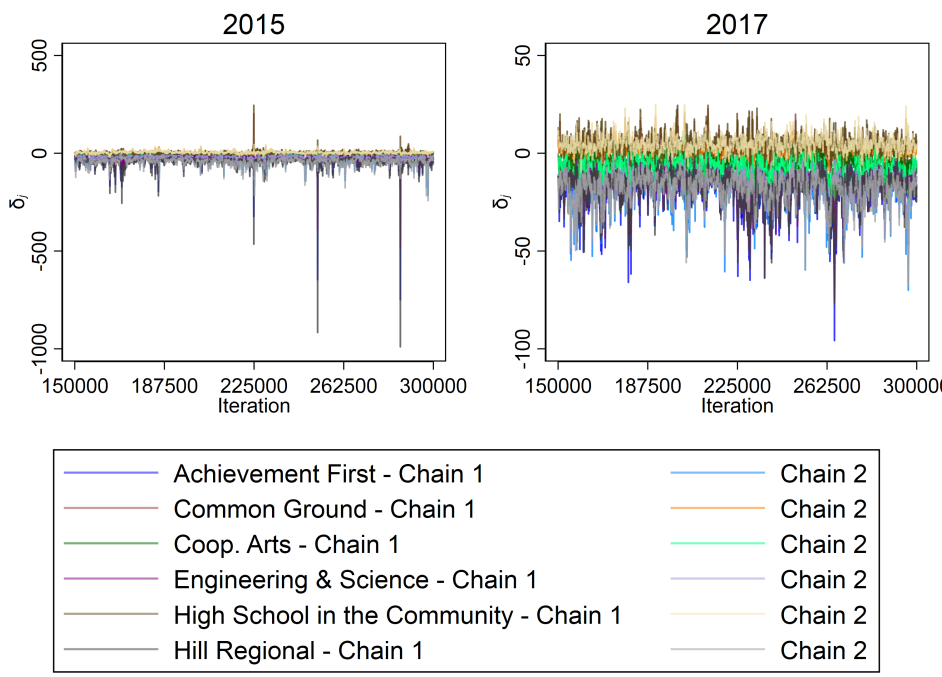
Upper graph: beliefs for respondents reporting personal involvement in school choice process. Lower graph: respondents not reporting personal involvement. Left panel: distribution of subjective and rational expectations assignment probabilities. Text reports gap in fraction of subjective reports and and RatEx values in the bin, with standard errors clustered at the respondent level in parentheses below. Right panel: distribution of optimism. Bars show shares of population within bins of width 10. Red line indicates mean of the distribution. In both panels, beliefs for second-ranked options are conditional on non-admission to the first-ranked choice.

Figure A6: Ratex beliefs, subjective beliefs, and optimism by whether respondent correctly recalled the submitted application



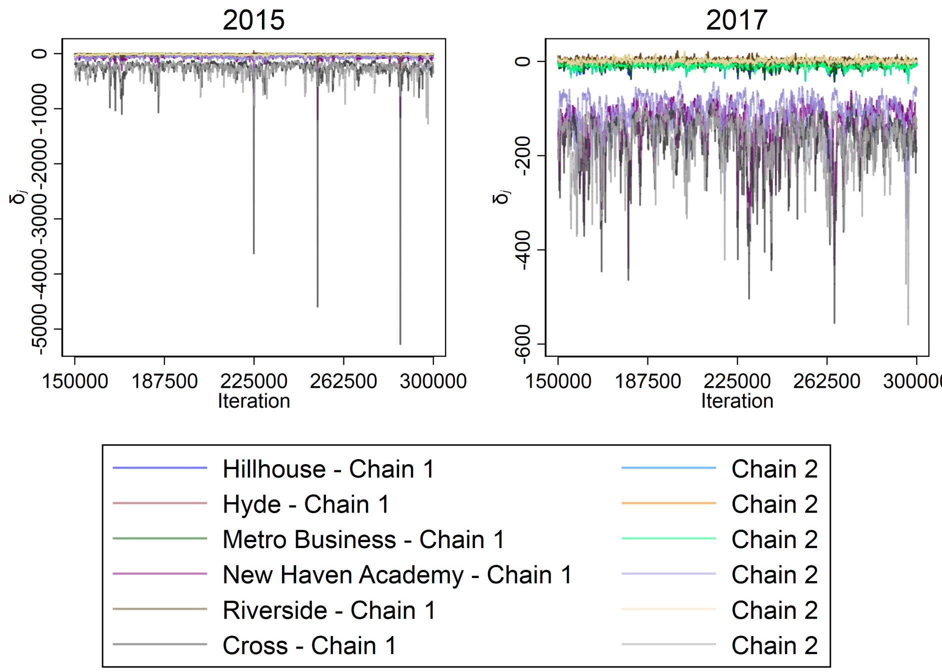
Upper graph: beliefs for respondents with correct recall of school choice application. Lower graph: respondents with incorrect recall. Left panel: distribution of subjective and rational expectations assignment probabilities. Text reports gap in fraction of subjective reports and and RatEx values in the bin, with standard errors clustered at the respondent level in parentheses below. Right panel: distribution of optimism. Bars show shares of population within bins of width 10. Red line indicates mean of the distribution. In both panels, beliefs for second-ranked options are conditional on non-admission to the first-ranked choice.

Figure A7: Trace plots: δ_j



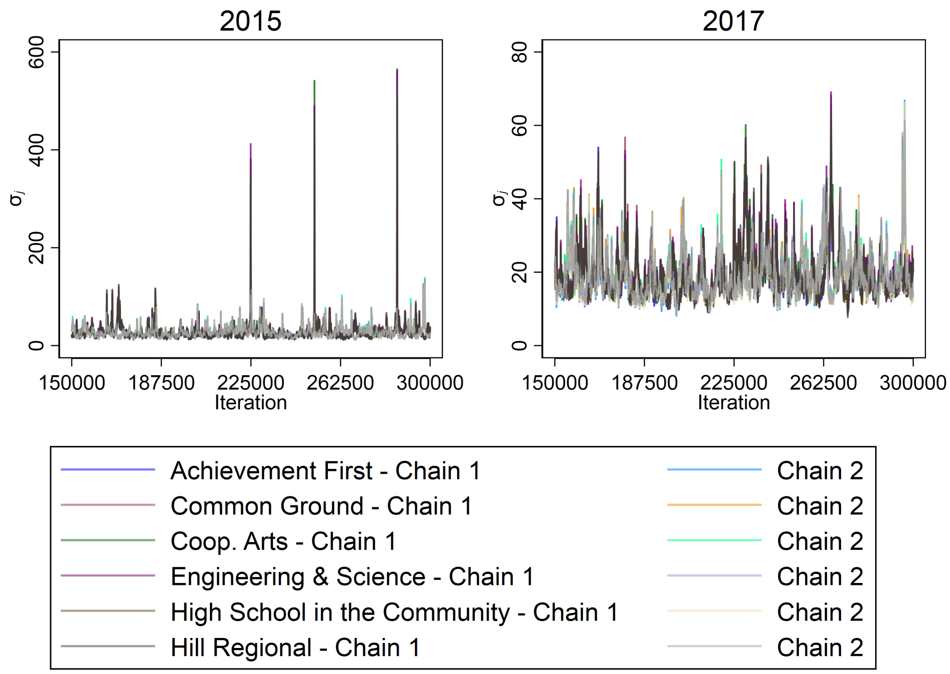
Trace plots for estimates of school-specific preference shifters δ_j split by year. For 2015, the Gelman-Rubin PSRF convergence statistics are, in order: 1.000, 1.010, 1.001, 1.002, 1.005, 1.000. For 2017, the Gelman-Rubin PSRF convergence statistics are, in order: 1.008, 1.003, 1.008, 1.009, 1.002, 1.012. See Section 5 for estimation details.

Figure A8: Trace plots: δ_j (con't)



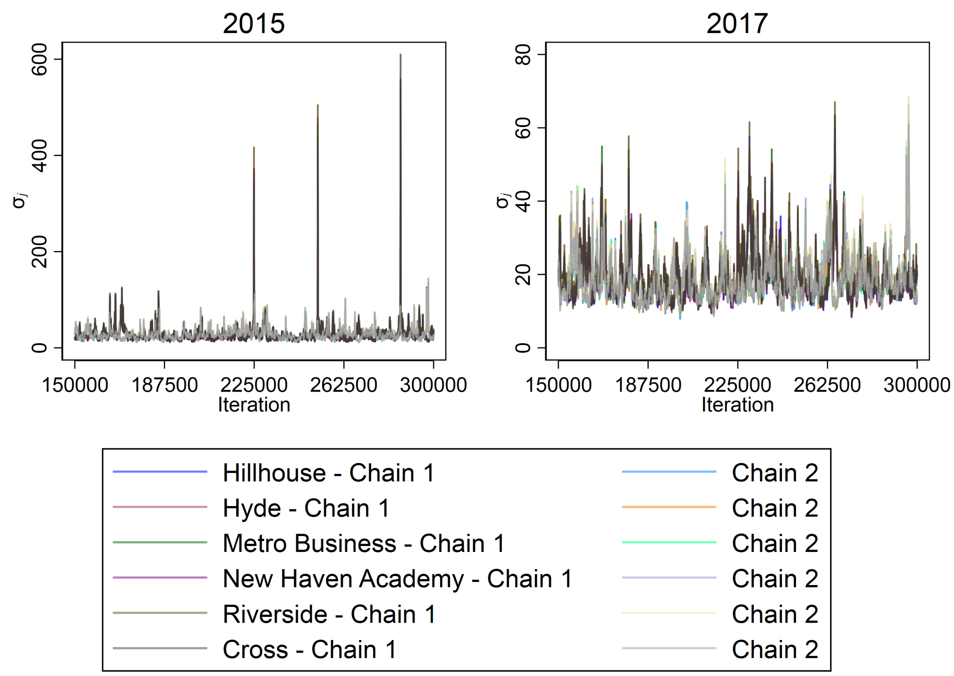
Trace plots for estimates of additional school-specific preference shifters δ_j split by year. For 2015, the Gelman-Rubin PSRF convergence statistics are, in order: 1.001, 1.003, 1.001, 1.000, 1.000, 1.000. For 2017, the Gelman-Rubin PSRF convergence statistics are, in order: 1.008, 1.005, 1.007, 1.045, 1.006, 1.010. See Section 5 for estimation details.

Figure A9: Trace plots: preference shocks $\sqrt{\Sigma_{(j,j)}}$



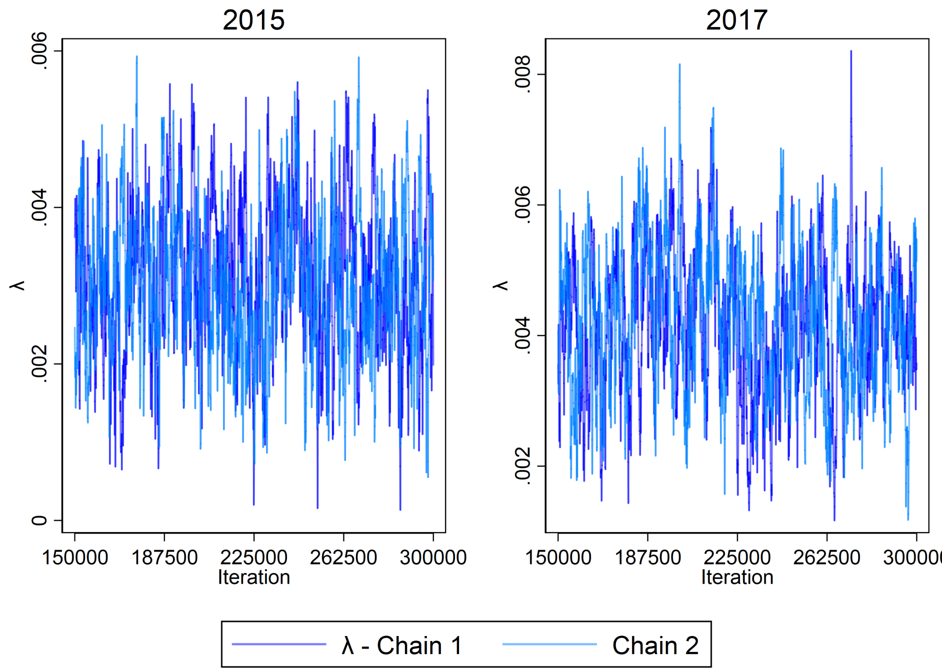
Trace plots for preference variance matrix terms $\sqrt{\Sigma_{j,j}}$ split by year. For 2015, the Gelman-Rubin PSRF convergence statistics are, in order: 1.000, 1.000, 1.000, 1.000, 1.000, 1.000. For 2017, the Gelman-Rubin PSRF convergence statistics are, in order: 1.010, 1.009, 1.011, 1.011, 1.009, 1.011. See Section 5 for estimation details.

Figure A10: Trace plots: preference shocks $\sqrt{\Sigma_{(j,j)}}$ (con't)



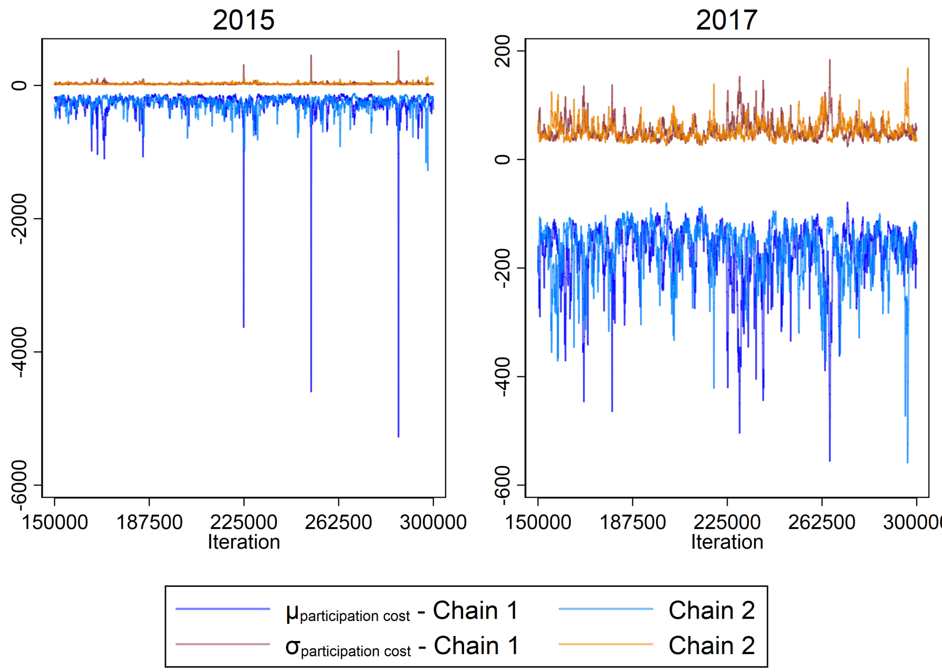
Trace plots for additional preference variance matrix terms $\sqrt{\Sigma_{j,j}}$ split by year. For 2015, the Gelman-Rubin PSRF convergence statistics are, in order: 1.000, 1.000, 1.000, 1.000, 1.000, 1.000. For 2017, the Gelman-Rubin PSRF convergence statistics are, in order: 1.011, 1.010, 1.009, 1.009, 1.012, 1.010. See Section 5 for estimation details.

Figure A11: Trace plots: enrollment shock parameter λ



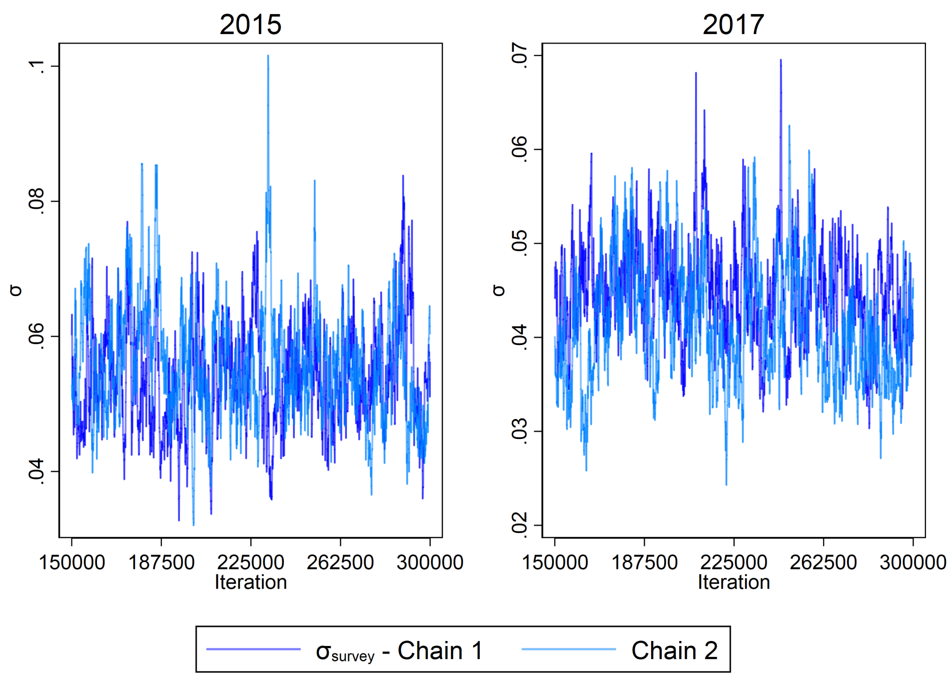
Trace plots for the enrollment shock scale parameter for the distribution of ϵ^e_{ij}, $\frac{1}{\lambda}$. The Gelman-Rubin PSRF convergence statistic for λ is 1.001 in 2015 and 1.013 in 2017. See section 5 for estimation details. See Section 4 for parameter definitions.

Figure A12: Trace plots: placement cost



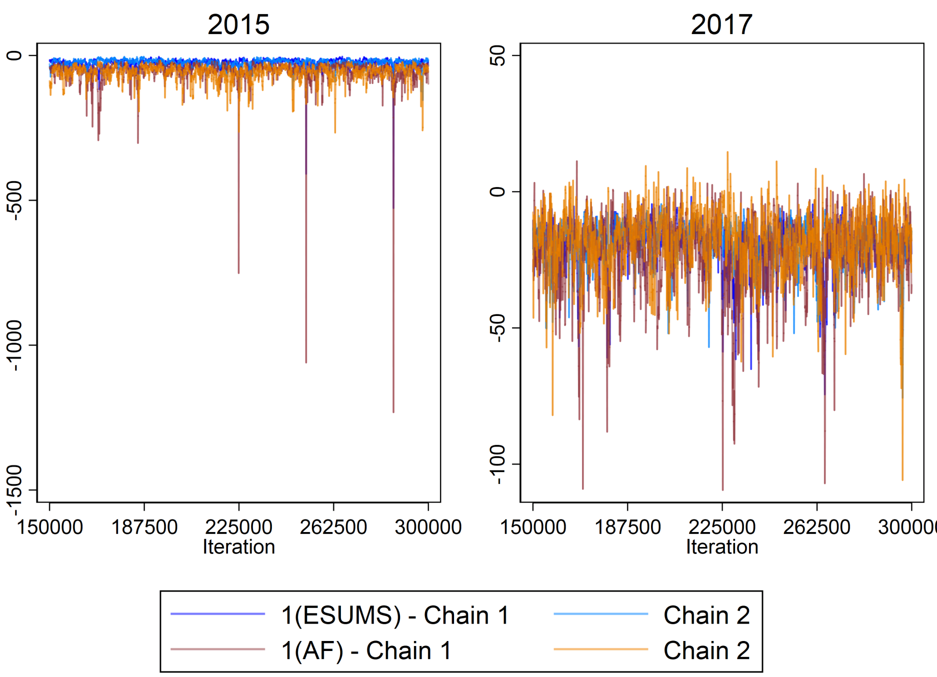
Trace plots for placement cost parameters split by year. The Gelman-Rubin PSRF convergence statistic for μ_b is 1.000 in 2015 and 1.010 in 2017. The Gelman-Rubin PSRF convergence statistic for σ_b is 1.000 in 2015 and 1.010 in 2017. See section 5 for estimation details. See Section 4 for parameter definitions.

Figure A13: Trace plots: measurement error σ_{survey}



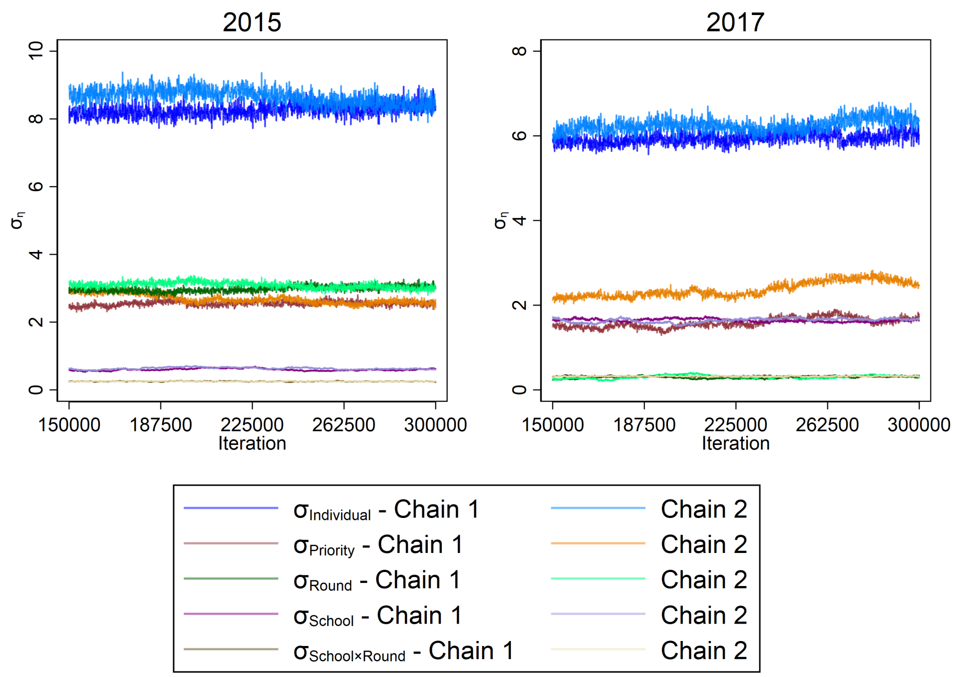
Trace plots for survey preference measurement error variance parameter split by year. The Gelman-Rubin PSRF convergence statistic for σ_{survey} is 1.010 in 2015 and 1.042 in 2017. See section 5 for estimation details. See Section 4 for parameter definitions.

Figure A14: Trace plots: other utility parameters



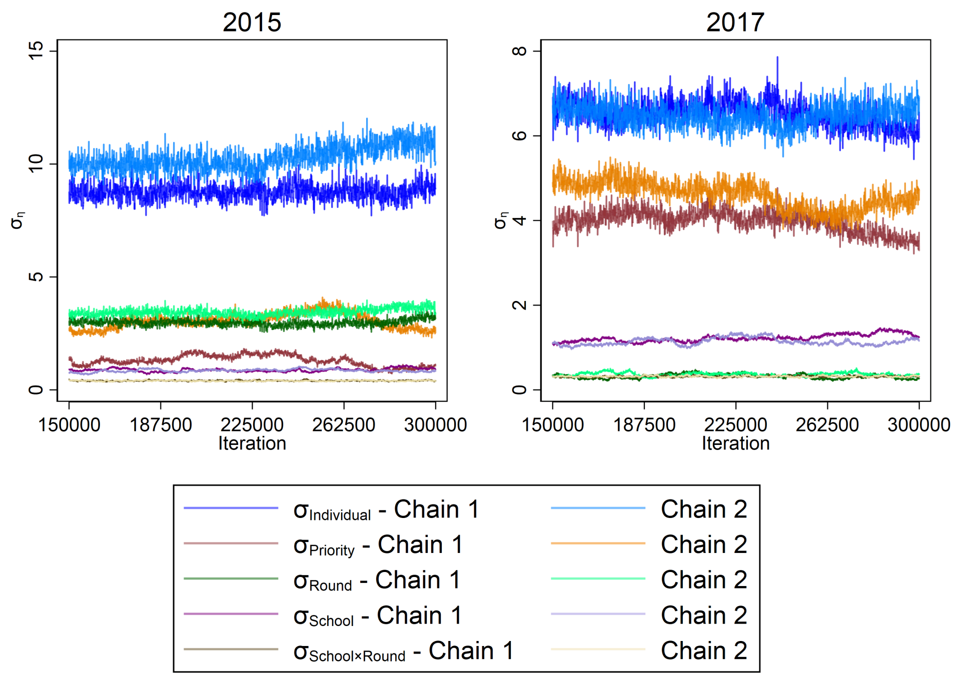
Trace plots for miscellaneous utility parameters split by year. The Gelman-Rubin PSRF convergence statistic for ESUMS is 1.000 in 2015 and 1.011 in 2017. The Gelman-Rubin PSRF convergence statistic for AF is 0.999 in 2015 and 1.015 in 2017. See section 5 for estimation details. See 4 for parameter definitions.

Figure A15: Trace plots: belief variances, low SES



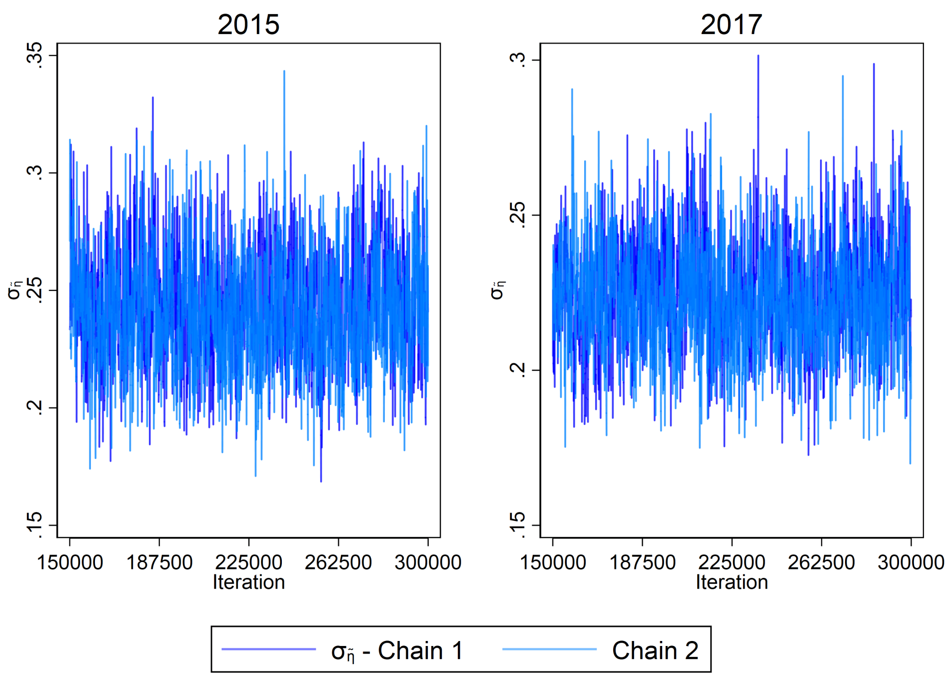
Trace plots for estimates of belief variance parameters for low SES households, split by year (left and right panels, respectively). The Gelman-Rubin PSRF convergence statistics for σ_{η_0} (low SES), $\sigma_{\eta_{pri}}$ (low SES), $\sigma_{\eta_{round}}$ (low SES), σ_{η_j} (low SES), $\sigma_{\eta_j r}$ (low SES), are, in order: 1.476, 1.339, 1.389, 1.142, and 1.015 in 2015 and 1.424, 3.890, 1.021, 1.269, and 1.004 in 2017. See Section 5 for estimation details.

Figure A16: Trace plots: belief variances, high SES



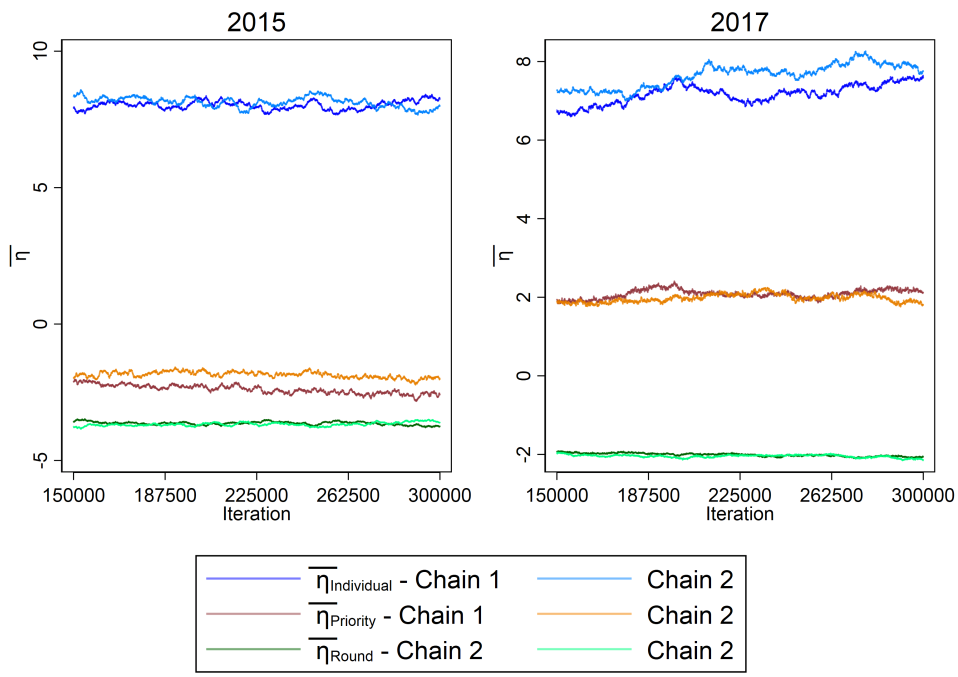
Trace plots for estimates of belief variance parameters for high SES households, split by year (left and right panels, respectively). The Gelman-Rubin PSRF convergence statistics for σ_{η_0} (high SES), $\sigma_{\eta_{pri}}$ (high SES), $\sigma_{\eta_{round}}$ (high SES), σ_{η_j} (high SES), $\sigma_{\eta_{jr}}$ (high SES), are, in order: 1.987, 2.974, 1.714, 1.039, and 1.004 in 2015 and 1.021, 1.697, 1.089, 1.214, and 1.014 in 2017. See Section 5 for estimation details.

Figure A17: Trace plots: measurement error $\sigma_{\tilde{\eta}}$



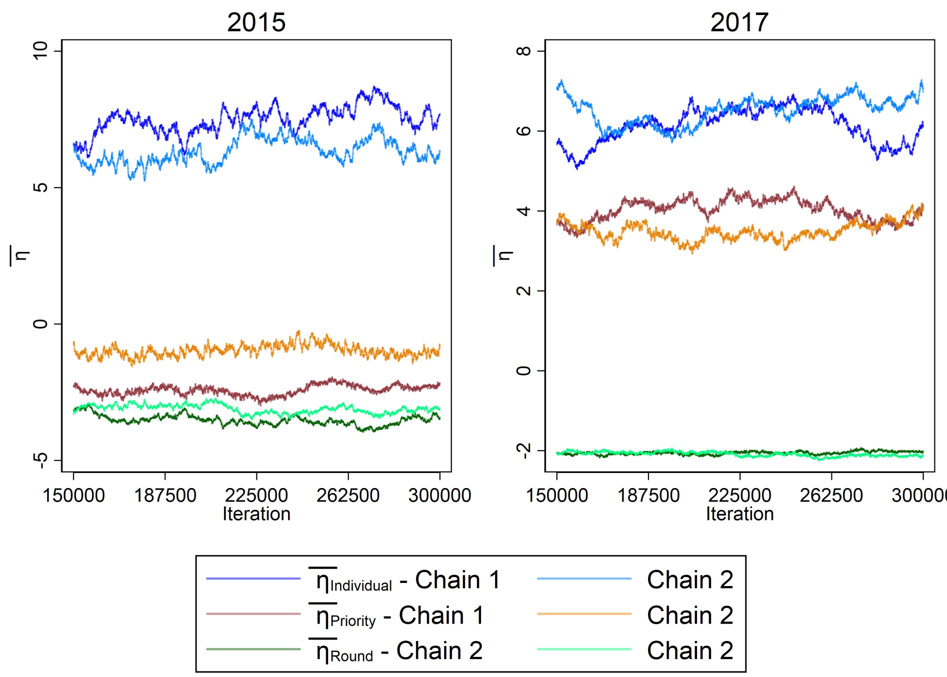
Trace plots for survey belief measurement error variance parameter split by year. The Gelman-Rubin PSRF convergence statistic for $\sigma_{\tilde{\eta}}$ is 1.003 in 2015 and 1.006 in 2017. See section 5 for estimation details. See Section 4 for parameter definitions.

Figure A18: Trace plots: belief means, low SES



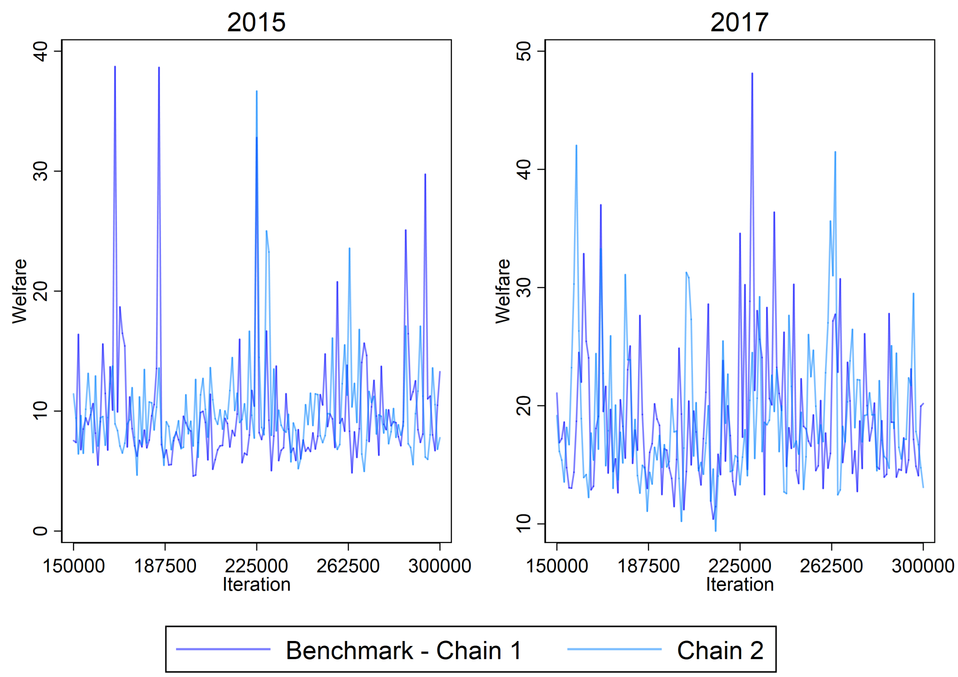
Trace plots for estimates of belief mean parameters for low SES households, split by year (left and right panels, respectively). The Gelman-Rubin PSRF convergence statistics for $\overline{\eta}_0$ (low SES), $\overline{\eta}_{priority}$ (low SES), $\overline{\eta}_{round}$ (low SES), 1.053, 2.873, and 1.056 in 2015 and 1.541, 1.128, and 1.329 in 2017. See Section 5 for estimation details.

Figure A19: Trace plots: belief means, high SES



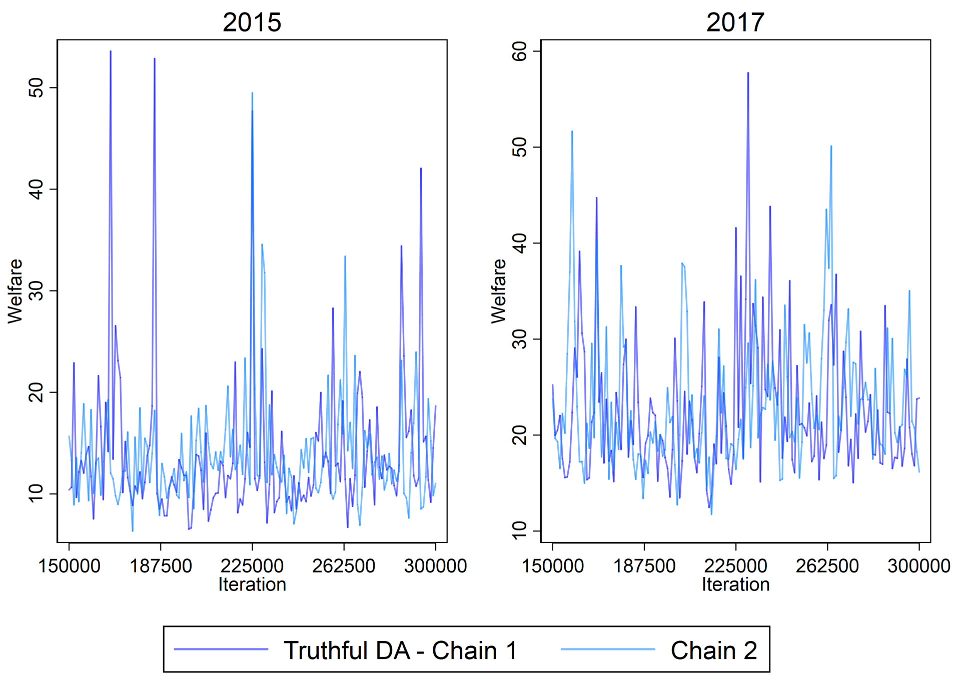
Trace plots for estimates of belief mean parameters for high SES households, split by year (left and right panels, respectively). The Gelman-Rubin PSRF convergence statistics for $\overline{\eta}_0$ (high SES), $\overline{\eta}_{priority}$ (high SES), $\overline{\eta}_{round}$ (high SES), 1.690, 3.637, and 1.660 in 2015 and 1.183, 1.403, and 1.360 in 2017. See Section 5 for estimation details.

Figure A20: Trace plots: welfare in the baseline model



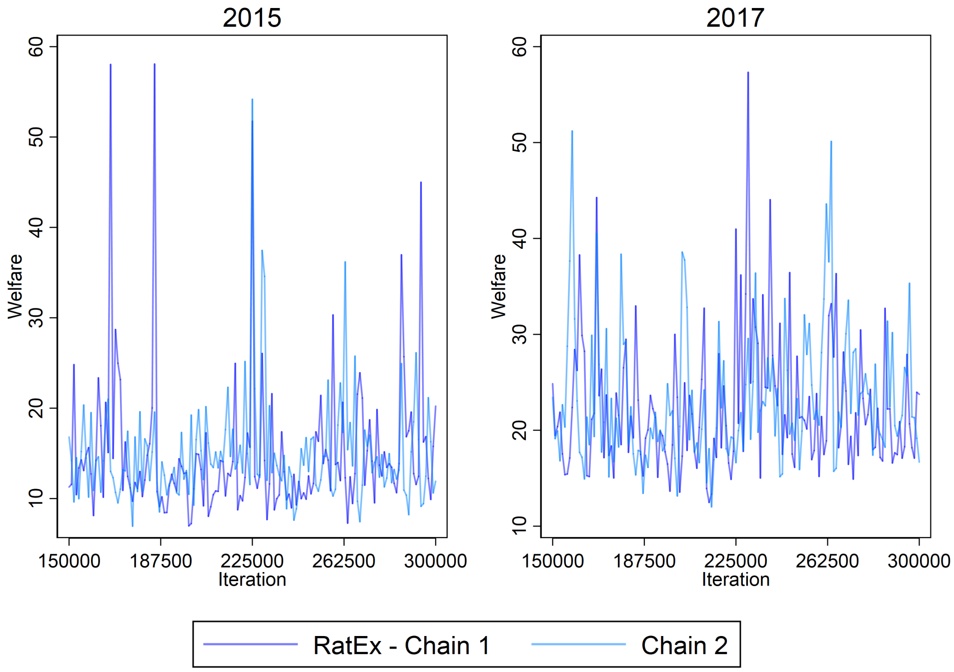
Trace plots for estimates of welfare in the baseline model. The Gelman-Rubin PSRF convergence statistic is 0.996 in 2015 and 1.004 in 2017. See section 5 for estimation details.

Figure A21: Trace plots: welfare in the 'DA' model



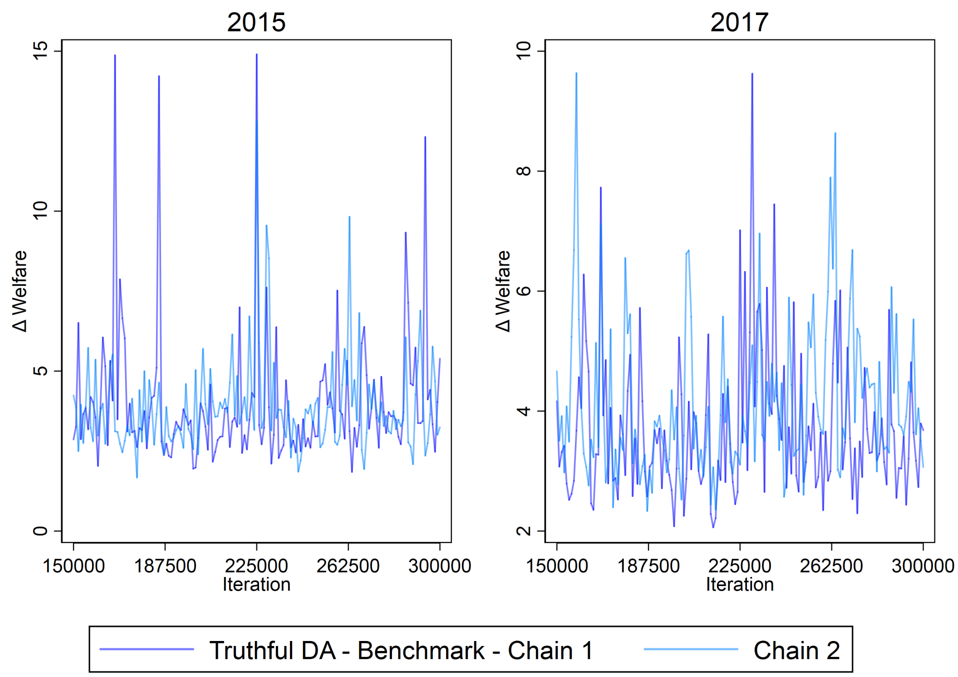
Trace plots for estimates of welfare in the 'DA' model. The Gelman-Rubin PSRF convergence statistic is 0.996 in 2015 and 1.005 in 2017. See section 5 for estimation details.

Figure A22: Trace plots: welfare in the 'RatEx' model



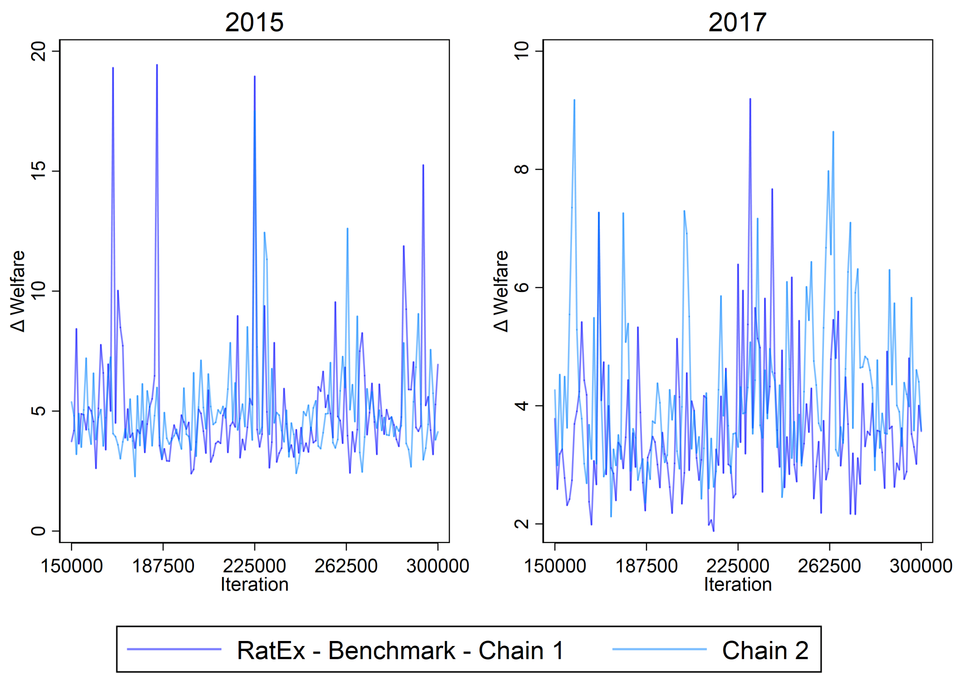
Trace plots for estimates of welfare in the 'RatEx' model. The Gelman-Rubin PSRF convergence statistic is 0.996 in 2015 and 1.007 in 2017. See section 5 for estimation details.

Figure A23: Trace plots: Δ welfare 'DA'−baseline



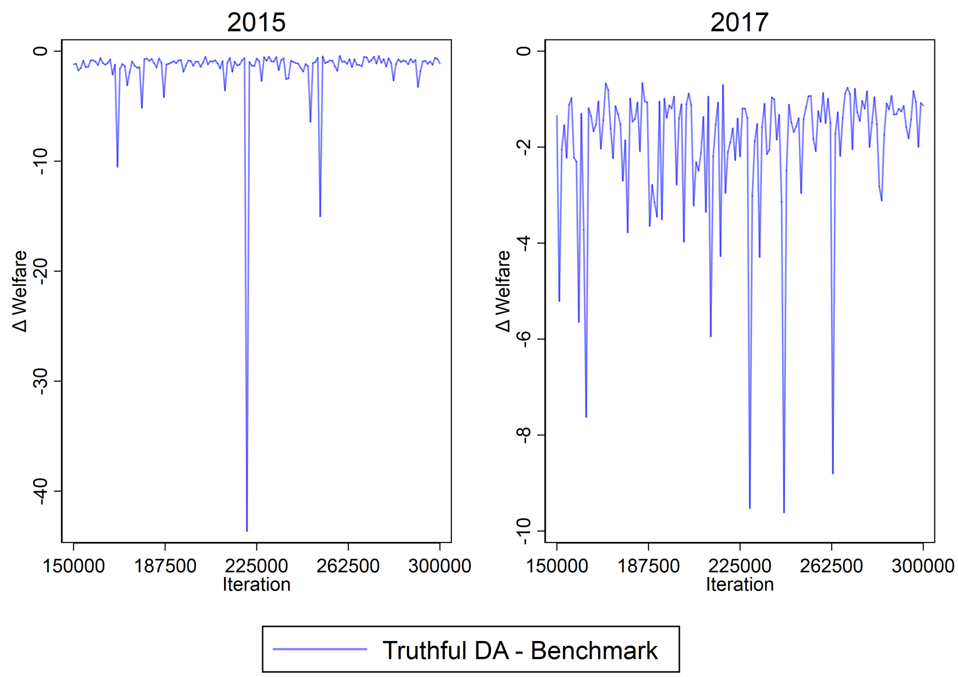
Trace plots for estimates of Δ welfare, 'DA'−baseline. The Gelman-Rubin PSRF convergence statistic is 0.996 in 2015 and 1.007 in 2017. See section 5 for estimation details.

Figure A24: Trace plots: Δ welfare ‘RatEx’−baseline



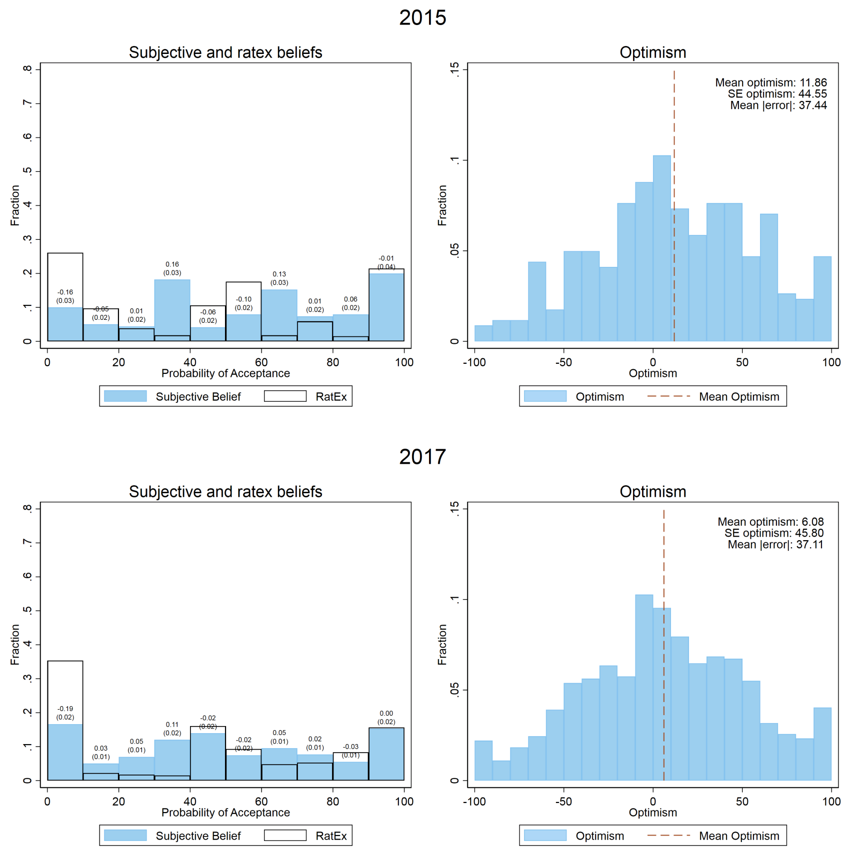
Trace plots for estimates of Δ welfare, ‘RatEx’−baseline. The Gelman-Rubin PSRF convergence statistic is 0.996 in 2015 and 1.040 in 2017. See section 5 for estimation details.

Figure A25: Trace plots: Δ welfare, 'No survey' 'DA'−baseline



Trace plots for estimates of Δ welfare, 'No survey' 'DA'−baseline. The Gelman-Rubin PSRF convergence statistic is 0.996 in 2015 and 1.007 in 2017. See section 5 for estimation details.

Figure A26: Ratex beliefs, subjective beliefs, and optimism by survey year



Upper graph: 2015. Lower graph: 2017. Left panel: distribution of subjective and rational expectations assignment probabilities. Text reports gap in fraction of subjective reports and and RatEx values in the bin, with standard errors clustered at the respondent level in parentheses below. Right panel: distribution of optimism. Bars show shares of population within bins of width 10. Red line indicates mean of the distribution. In both panels, beliefs for second-ranked options are conditional on non-admission to the first-ranked choice.

B Appendix B: Estimation details

B.1 Constraints implied by optimal behavior

We can write the constraints implied by reported preferences and the optimality of enrollment decisions in matrix form as follows:

$$\Lambda'_{i,(shock)} \begin{pmatrix} u_i \\ \epsilon_i^{survey} \\ \epsilon_i^e \end{pmatrix} \geq 0.$$

If i reported first and second choices j_1 and j_2, respectively, then the first column of $\Lambda_{i,(shock)}$ contains 1's in the j_1th and $(J + j_1)$th places, and -1 in the j_2th and $(J + j_2)$th places.[18] The next $J - 1$ columns similarly require

$$u_{i,j_2} + \epsilon_{i,j_2}^{survey} > u_{i,j} + \epsilon_{i,j}^{survey} \quad \text{for } j \neq j_1, j_2.$$

If i was placed in school j and enrolled there, then the final column of $\Lambda_{i,(shock)}$ contains 1 in the jth place and -1 in the $2J + 1$th place. If i was placed in j but did not enroll there, the final column contains -1 in the jth place and 1 in the final place.

When i receives a placement $j > 0$, then $\epsilon_i^e \equiv \epsilon_{ij}^e - \epsilon_{i0}^e$ denotes the difference between his shock for school j and his shock for his outside option. Because both of these shocks are distributed T1EV, their difference ϵ_i^e has a logistic distribution.

B.2 Starting values

We first construct feasible belief shifts $shift_{ijr}$ for all i, j, and r. Where the survey provides no constraints, we start at $shift_{ijr} = 0$, i.e. at the rational-expectations value. We pick points interior to the relevant intervals when households report beliefs. Given these values of $shift_{ijr}$, we set initial measurement error $\tilde{\eta}_{ijr}$=0 for all i, j, r.

Next, given the feasible beliefs, we use linear programming techniques to construct strictly feasible utilities $u_i \in \mathbb{R}^J$ and placement payoff terms $b_i \in \mathbb{R}$. A utility vector u_i and benefit b_i are (strictly) feasible if the observed report a_i is optimal conditional on the beliefs p_i, that is if $\Gamma_i (v_i + b_i)' > 0$, where Γ_i is the matrix of constraints induced by the optimality of the observed report, given the cutoff distribution and shift terms $shift_{ijr}$ for all schools $j \in J$ and rounds

[18] If i reported a first but not a second choice, we similarly construct $\Lambda_{i,(shock)}$ using the resulting inequalities.

$r \in \{1,2,3,4\}$. We allow the set of possible reports to include an empty list, which we interpret as nonparticipation.

Finally, we use linear programming again to pick strictly feasible enrollment-time shocks ϵ_i^e and measurement errors ϵ_i^{survey}.

To obtain a starting value, we choose $\lambda = 1$.

We now describe the prior distributions and the MCMC procedure that we use to estimate the model parameters.

B.3 Prior distributions

We begin with prior distributions over the preference parameters and belief parameters. We place priors directly on β, Σ, μ_b, σ_b, and σ_{survey} as well as on the belief parameters separately by SES category. In order to minimize the priors' influence on our estimates, we choose the following diffuse priors:

$$
\begin{aligned}
\lambda &\sim Gamma(2, .5) \\
(\beta_{-dist}, \mu_b)|\lambda &\sim N(0, 100 * I/\lambda) \\
\Sigma|\lambda &\sim IW(100, I/\lambda) \\
{\sigma^2}_{survey}, {\sigma^2}_b|\lambda &\sim InverseGamma(1, \lambda^{-2}) \text{ iid} \\
\overline{\eta} &\sim N(0, 100 * I) \\
\Sigma^\eta &\sim IW(4, I) \\
\sigma^2_{\eta_{school}}, \sigma^2_{\eta_{school \times round}}, \sigma^2_{\tilde{\eta}} &\sim InverseGamma(1, 1) \text{ iid}
\end{aligned}
$$

Here, β_{-dist} denotes the coefficients on all preference shifters other than *distance*. Other than the stated dependence on λ, we assume that the priors are independent.

Let $\tilde{\beta} = \beta\lambda$. Because $\beta_{distance}$ is normalized to -1, we have $\tilde{\beta}_{distance} = -\lambda$. Similarly, define $\tilde{\Sigma} = \Sigma\lambda$, $\tilde{\sigma}_b = \sigma_b\lambda$, and $\tilde{\sigma}_{survey} = \sigma_{survey}\lambda$. Let $\tilde{\mu}_b = \mu_b\lambda$. We then have $\tilde{\beta}_{-dist} \sim N(0, 100 * I), \Sigma \sim IW(100, I), \tilde{\sigma}^2_{survey} \sim InverseGamma(1, 1)$, and $\tilde{\sigma}^2_b \sim InverseGamma(1, 1)$, independently.

B.4 MCMC iteration

Let u_i denote the vector $\{u_{ij}\}_{j \in J}$. Similarly, let ϵ_i denote the vector of preference measurement errors, η_i the random coefficients in beliefs, and $shift_i$ the matrix of shift terms for household i.

Let $u = \{u_i\}_{i \in I}$ denote the matrix of utilities of all households.

For each i, let $\tilde{u}_i = u_i \lambda$. Let $\tilde{b}_i = b_i \lambda$. Let $\tilde{\epsilon}_i^e = \epsilon_i^e \lambda$. We augment the data with $\tilde{u}_i, \tilde{b}_i, \tilde{\epsilon}_i^e, \tilde{\epsilon}_i^{survey}, \eta_i, \tilde{\eta}_i$, and $shift_i$ for each household i.

We iterate through the following steps, which consist of sampling from the conditional posterior distributions of utilities, utility shocks, beliefs, belief measurement error, application costs, and model parameters:

1. Draw λ from its posterior distribution conditional on the data, augmented data, and parameters.
2. Draw mean-utility parameters $\tilde{\beta}^{(s+1)}$ and mean benefit $\tilde{\mu}_b^{(s+1)}$ from the distribution of $\tilde{\beta}|\tilde{u}^{(s)}, \tilde{\Sigma}^{(s)}$ and $\tilde{\mu}_b|\tilde{b}^{(s)}, \tilde{\sigma}_b^{(s)}$
3. Draw variance of benefit term $(\tilde{\sigma}_b^2)^{(s+1)}$ from the distribution of $\tilde{\sigma}_c^2|\tilde{\mu}_b^{(s+1)}, \tilde{b}^{(s)}$.
4. Draw variance of shocks to reported preferences $\tilde{\sigma}^2_{survey}$ from the distribution of $\tilde{\sigma}^2_{survey}|\tilde{\epsilon}^{survey}$.
5. Draw covariance matrix $\tilde{\Sigma}^{(s+1)}$ from the distribution of $\tilde{\Sigma}|\tilde{\beta}^{(s+1)}, \tilde{u}^{(s)}$.
6. Draw the parameters of the belief distribution from their posterior conditional on $shift_i$ and belief random effects η_i^0, $\eta_i^{priority}$, η_i^{round}, and $\{\eta_{ij}\}_{j \in J}$ for all i. Draw belief measurement error variance $\sigma^2_{\tilde{\eta}}$ from its posterior distribution given $\tilde{\eta}$.
7. For each individual in the dataset:
   (a) Draw utility $\tilde{u}_i^{(s+1)}$ from the posterior distribution of $\tilde{u}_i$ given $\tilde{\beta}, \tilde{\Sigma}$, i's decision to accept or decline his placement (if offered one), and constraints implied by the optimality of i's report.
   (b) Draw $\tilde{b}_i^{(s+1)}$ from the posterior distribution of $\tilde{b}_i$ given $v_i(\tilde{u}_i^{(s+1)})$ and constraints implied by the optimality of i's report.
   (c) Draw shock realizations $\tilde{\epsilon}_i^{survey}$ and $\tilde{\epsilon}_i^e$ from their posterior distributions given $\tilde{u}_i$ and the household's decisions.
   (d) Draw belief random effects η_i^0, $\eta_i^{priority}$, η_i^{round}, and $\{\eta_{ij}\}_{j \in J}$ from their posterior distribution given $shift_i$, $\overline{\eta}$, Σ_η, $\sigma^2_{\eta_{school \times round}}$, and $\sigma^2_{\eta_{school}}$.
   (e) Draw belief measurement error $\tilde{\eta}_i$ from its posterior given $shift_i$, belief random effects, and the constraints imposed by the elicited belief measures.
   (f) Draw $shift_i$ from its posterior distribution conditional on $\tilde{\eta}_i, \eta_i^0$, $\eta_i^{priority}$, η_i^{round}, $\{\eta_{ij}\}_{j \in J}$, v_i, b_i, and the constraints imposed by the survey.

B.5 Updating λ

Under the data augmentation strategy outlined above, λ enters the likelihood only via the transformed coefficient on distance, $\tilde{\beta}_{distance} = -\lambda$. Each time we update λ, we use a sequence of 10 Metropolis-Hastings steps with symmmetric normal proposal densities with variance 0.01.[19] Observe that $\lambda\epsilon_{ij} = \tilde{u}_{ij} - x'_{ij}\tilde{\beta}_{-dist} + distance_{ij}\lambda$. The likelihood of λ conditional on $\tilde{\Sigma}$, $\tilde{u}$, and observables (distance, x) is therefore given by

$$\phi(\tilde{u}_i - x'_i\tilde{\beta}_{-dist} + distance_i\lambda; \mathbf{0}, \tilde{\Sigma}),$$

where $\phi(v; m, \Sigma)$ is the density of a multivariate normal distribution $MVN(m, \Sigma)$ evaluated at v.

B.6 Updating utilities

In order to update utilities, for each individual we iterate through the various schools, updating the terms $\tilde{u}_{ij}$ sequentially. Because $\tilde{u}_i$ is jointly normal, the distribution of $\tilde{u}_{ij}|\tilde{u}_{i,-j}, \beta, \Sigma$ is normal with known mean and variance.

The restriction $\Gamma'_i(\tilde{v}_i+\tilde{b}_i) \geq 0$ implies that $\tilde{v}_{ij}$ must belong to a (known) interval whose endpoints depend on $\tilde{v}_{i,-j}$ and $\tilde{b}_i$.[20] Recall that $\tilde{v}_{ij} = \log(1 + \exp(\tilde{u}_{ij}))$ is a monotone transformation of $\tilde{u}_{ij}$. Therefore, conditional on the optimality of the report and the current values of other variables and parameters, updating $\tilde{u}_{ij}$ consists of drawing from a truncated normal distribution.

B.7 Updating preference shocks

We draw shock realizations $\tilde{\epsilon}_i^{survey}$ and $\tilde{\epsilon}_i^{e}$ from their posterior distributions given $\tilde{u}_i$ and the household's decisions. The procedure is analogous to drawing utilities subject to linear constraints represented by a matrix Γ. Here we have: $\Lambda'_{i,(shock)}\begin{pmatrix} \tilde{u}_i \\ \tilde{\epsilon}_i^{survey} \\ \tilde{\epsilon}_i^{e} \end{pmatrix} \geq 0$.

B.8 Updating beliefs

The remaining steps are standard Gibbs-sampler steps, with the exception of the updates to belief shift terms $shift_{ijr}$ and belief measurement error $\tilde{\eta}_i$.

[19] In principle, one step per iteration would suffice. We find that more steps lead to faster convergence.

[20] Similarly, $\tilde{b}_i$ must belong to an interval with known endpoints that depend on $\tilde{v}_i$.

To update each of these parameters we take a sequence of Metropolis-Hastings steps with normal proposal densities. We tune the variance of the proposal density so that roughly a third of the draws are accepted. In particular, we take a sequence of 5 Metropolis-Hastings steps within each update of $\tilde{\eta}_i$ or $shift_i$. A single Metropolis-Hastings step to update $shift_i$ is as follows.

$shift_i$ can be represented as a $(J \times R)$ matrix. We draw a $(J \times R)$ matrix of iid normal shocks, $\Delta(shift_{ijr}) \sim N(0, \sigma_{proposal})$, and construct a new proposal $shift'_i = shift_i + \Delta(shift_i)$. We then compute the likelihood ratio $a = \frac{\ell(shift_i + \Delta(shift_i))}{\ell(shift_i)}$, where

$$\ell(shift_i) = \prod_{j,r} \phi(shift_{ijr} - \eta_i^0 - \eta_i^{priority} * priority_{ij} - \eta_i^{round} * r - \eta_{ij}; \sigma^2_{\eta_{\text{school}\times\text{round}}})$$

where $\phi(x;\sigma)$ is the density of a normal distribution with mean zero and variance σ^2 evaluated at x. To understand this expression, observe that

$$\eta_{ijr} = shift_{ijr} - \eta_i^0 - \eta_i^{priority} * priority_{ij} - \eta_i^{round} * r - \eta_{ij}.$$

$shift_i + \Delta(shift_i)$ is consistent with the survey iff

$$shift_{ijr} + \Delta(shift_{ijr}) + \tilde{\eta}_{ijr} \in I_{ijr}$$

where I_{ijr} is the reported interval.

If $a > 1$ and the proposal is consistent with the survey and with the observed report, the proposal is accepted and we set

$$shift_i := shift_i + \Delta(shift_i).$$

If $a < 1$ and the proposal is consistent with the survey and observed report, we accept it with probability a. We reject the proposal with probability 1 if it violates the constraints imposed by the survey or causes the observed report to become non-optimal.

Once $shift_i$ is updated, we recalculate Γ_i accordingly.

The update to belief measurement error $\tilde{\eta}_i$ similarly consists of sequence of 5 Metropolis-Hastings steps. A key distinction is that we update each element of belief measurement error $\tilde{\eta}_{ijr}$ separately. We keep track of measurement error only for schools and rounds at which we elicited beliefs. For these schools and rounds, in each MH step we draw a proposal $\tilde{\eta}_{ijr} + \Delta_{ijr}$ where $\Delta \sim N(0, \sigma^2_{\underbrace{}_{proposal}})$ and accept with the appropriate MH acceptance probability.

B.9 Convergence properties

Trace plots for parameter estimates are reported in Online Appendix Figures A7 through A19. The trace plots show that mean and variance parameters from the preference model are precisely estimated with potential scale reduction factors (PSRFs) close to one in every case. Belief model estimates also show evidence of convergence. The notable exception is for the $\sigma_{\eta_{pri}}$ and $\overline{\eta}_{priority}$ parameters. These parameters affect beliefs for the relatively small share of households with sibling priority, and their estimation relies on data from the smaller group of surveyed households who were asked about schools at which they have a sibling. Any non-convergence in belief parameter estimates that may exist does not lead to convergence issues for our core estimates of counterfactual welfare effects. As shown in Figures A20 through A25, the estimates of welfare levels and differences reported in the next section all have PSRFs of almost exactly one.

C Appendix C: Alternate model

This section describes our alternative specification, which treats as exogenous students' enrollment decisions following the choice process. This specification is close to that of Agarwal and Somaini (2018), but integrates belief and preference data from our survey.

C.1 Model

Our alternative model consists of three stages. First, applicants learn their preferences over schools and costs of applying to schools. Second, they choose whether to participate in the school choice process and, if they participate, what report to submit. Third, the lottery runs and participants receive placements. Utility is realized as a function of students' placements.

Students $i \in I$ have underlying preferences over schools $j \in J$ according to:

$$u_{ij} = \delta_j + X_{ij}\beta + \epsilon_{ij},$$

where X_{ij} are observed school and student characteristics. The errors ϵ_i are distributed according to

$$\epsilon_i \sim MVN(0, \Sigma),$$

iid across households, where Σ is unrestricted. X_i consists of the same observables as in our main specification: distance, a full set of school dummies, a low-SES indicator, distance to the zoned school, and identity of the zoned school.

Household i chooses an application portfolio a to solve

$$\max_a \left(\sum_j \tilde{p}_{ija} u_{ij} \right).$$

Subjective beliefs $\tilde{p}_{ija}$ are modeled as in our main specification.

C.2 Estimation

As before, we normalize $\beta_{dist} = -1$. We use the same priors as our main specification, the same number of draws, and the same burn-in period. Our estimation procedure is modified as follows. There is no matriculation-time shock ϵ_i^e or accept/decline cost b_i, so we do not track these variables.

Let u_i denote the vector $\{u_{ij}\}_{j \in J}$. Similarly, let ϵ_i denote the vector of preference measurement errors, η_i the random coefficients in beliefs, and $shift_i$ the matrix of shift terms for household i. Let $u = \{u_i\}_{i \in I}$ denote the matrix of utilities of all households.

1. Draw mean-utility parameters $\beta^{(s+1)}$ from the distribution of $\beta | u^{(s)}, \Sigma^{(s)}$.

2. Draw variance of shocks to reported preferences σ^2_{survey} from the distribution of $\sigma^2_{survey} | \epsilon^{survey}$.

3. Draw covariance matrix $\Sigma^{(s+1)}$ from the distribution of $\Sigma | \beta^{(s+1)}, u^{(s)}$.

4. Draw the parameters of the belief distribution from their posterior conditional on $shift$ and belief random effects η_i^0, $\eta_i^{priority}$, η_i^{round}, and $\{\eta_{ij}\}_{j \in J}$ for all i. Draw belief measurement error variance $\sigma^2_{\tilde{\eta}}$ from its posterior distribution given $\tilde{\eta}$.

5. For each individual in the dataset:

   (a) Draw utility $u_i^{(s+1)}$ from the posterior distribution of u_i given β, Σ and constraints implied by the optimality of i's report.

   (b) Draw shock realizations ϵ_i^{survey} from their posterior distributions given u_i and the household's decisions.

   (c) Draw belief random effects η_i^0, $\eta_i^{priority}$, η_i^{round}, and $\{\eta_{ij}\}_{j \in J}$ from their posterior distribution given $shift_i$, $\overline{\eta}$, Σ_η, $\sigma^2_{\eta_{school \times round}}$, $\sigma^2_{\eta_{school}}$, and measurement error $\tilde{\eta}_i$.

   (d) Draw belief measurement error $\tilde{\eta}_i$ from its posterior given $shift_i$, belief random effects, and the constraints imposed by the elicited belief measures.

   (e) Draw $shift_i$ from its posterior distribution conditional on η_i^0, $\eta_i^{priority}$, η_i^{round}, $\{\eta_{ij}\}_{j \in J}$, u_i, and the constraints imposed by the survey and optimality of i's report.

C.3 Results

In this section we provide the analogues of Table 6, Table 7, and Figures 5, 6, and C3. Our findings are qualitiatively identical to our main findings in terms of welfare ordering of counterfactuals. Quantitiatvely, welfare gains from the switch to DA are similar in percentage terms to those reported in Table 6. Welfare levels relative to the outside option are lower across all specifications. See section 6.3 for additional discussion.

Table C1: Distance-Metric Welfare: Benchmark and Counterfactuals

	Mean welfare			Welfare differences		
	Baseline	RatEx	DA	RatEx − Baseline	DA − Baseline	No Survey DA − Baseline
A1. Mean distance metric						
Mean	2.840	3.741	3.749	0.901	0.909	−0.198
Median	2.762	3.629	3.651	0.939	0.900	−0.198
95% CI	[1.449, 4.256]	[2.437, 5.183]	[2.347, 5.257]	[0.640, 1.116]	[0.784, 1.113]	[−0.284, −0.125]
A2. SES gap						
Mean	−0.600	−0.819	−0.819	−0.219	−0.219	0.081
Median	−0.576	−0.811	−0.808	−0.225	−0.218	0.075
95% CI	[−1.074, −0.199]	[−1.235, −0.350]	[−1.237, −0.312]	[−0.416, −0.001]	[−0.412, −0.026]	[0.010, 0.168]

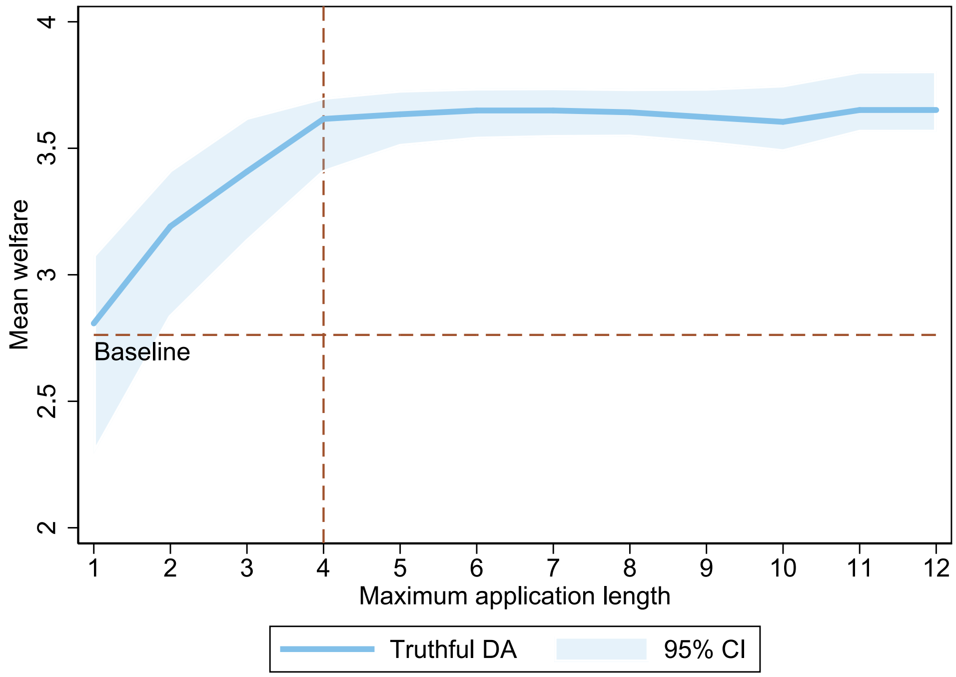
	Truthful	Strategic	Drops	Stops
B. Mistakes under DA				
Mean	0.792	0.838	0.819	0.791
Median	0.795	0.855	0.814	0.795
95% CI	[0.619, 0.964]	[0.667, 1.022]	[0.679, 0.966]	[0.618, 0.963]

	0%	25%	50%	75%	100%
C. Share submitting baseline application under DA-4					
Mean	0.792	0.634	0.464	0.288	0.111
Median	0.795	0.632	0.463	0.294	0.099
95% CI	[0.619, 0.964]	[0.539, 0.747]	[0.382, 0.540]	[0.184, 0.389]	[−0.008, 0.240]

	Switch to DA		Keep baseline mechanism	
Quantile	School and priority	School	School and priority	School
D. Error components				
Mean	0.557	0.559	0.355	0.359
Median	0.549	0.551	0.356	0.360
95% CI	[0.343, 0.810]	[0.343, 0.814]	[0.028, 0.731]	[0.034, 0.730]

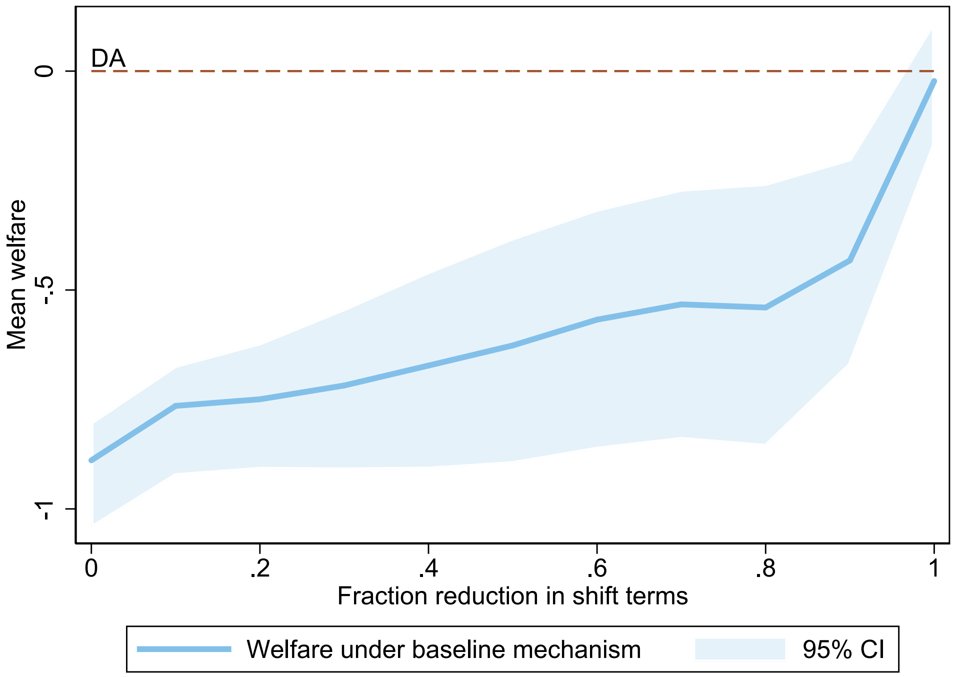
Notes: This table describes the posterior distribution of mean welfare in the baseline case and under policy counterfactuals. Welfare is measured using miles traveled as the numeraire good. Panels A1 and A2: 'Baseline' is baseline (New Haven or Boston) mechanism given observed beliefs. 'RatEx' is the baseline mechanism under rational expectations beliefs. 'DA' is the strategy-proof deferred acceptance mechanism. 'RatEx-baseline' and 'DA-baseline' columns compare welfare differences under the listed mechanisms. 'No survey DA-base' column compares welfare under the DA and baseline mechanisms using model estimates based on rational expectations beliefs. Panel A2 displays differences in each of these objects between high-SES and low-SES households. Panel B: difference between DA welfare and baseline welfare under 'drop' and 'stop' DA play (columns 1-4) and sophisticated truncated DA-4. See text for details. Panel C: Welfare gain from switch from baseline to truncated DA-4 by share of households continuing to submit 'baseline' applications. See text for details. Panel D: Welfare change from switch from baseline to strategic truncated DA with school- and school by priority-specific errors (columns 1+2), and welfare change from switching to only school- and school by priority-specific errors while keeping the baseline mechanism. See text for details. Calculations use alternative model.

Figure C1: Welfare under naive DA by list length



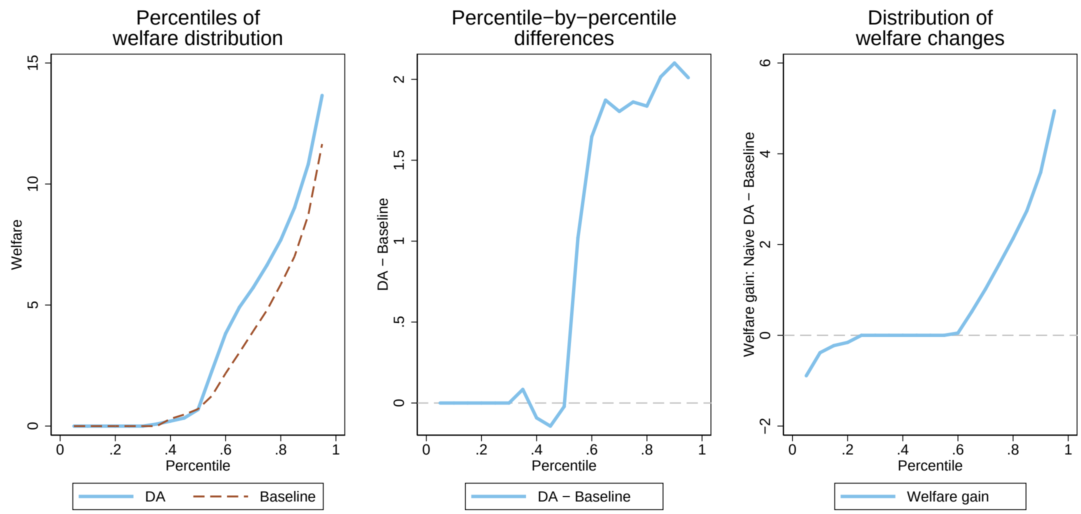
Notes: median of posterior mean welfare distribution (vertical axis) under truthful DA policy counterfactual by application length (horizontal axis). 'Baseline' line is median of posterior mean welfare under the baseline mechanism and observed beliefs with an application length of four. Calculations use alternative model.

Figure C2: Mean welfare by reduction in scale of shift term



Notes: median of posterior distribution of differences in mean welfare between baseline and DA (vertical axis) by fraction reduction in $shift_{ijr}$ terms (horizontal axis). Calculations use alternative model.

Figure C3: Percentiles of the welfare distribution



Notes: Left panel: posterior mean welfare by centile of welfare distribution under baseline and strategy-proof DA. Middle panel: centile-by-centile differences in welfare between DA and baseline policies. Right panel: percentiles of welfare gain distribution from switch to strategy-proof DA from baseline. Calculations use alternative model.

D Appendix D: Back of the envelope calculation

We have shown that the welfare effects of changes in choice mechanism and informational environment represent large shares of mean utility relative to students' outside options. To place welfare effects in broader context, we conduct a back of the envelope calculation that maps distance-metric utility to travel time, and travel time to dollars. There were 18,947 students enrolled in NHPS grades Kindergarten through 12 in the 2014-2015 academic year. All were assigned to schools through the placement process or following a decision not to participate. There are 180 school days in the year, and each student must travel both to and from school, for an estimated 6.8 million trips per year. From Table 7, students receive per-trip welfare gains equivalent to 3.9 fewer miles traveled per trip from a switch to the DA mechanism, for a total welfare gain of 27 million fewer miles per year. Using Google Maps walk- and drive-time measures and assuming that students who live within one mile of a school choose to walk, we compute average hours per mile of travel time to the enrolled school as 0.21, for a total time gain of 5.7 million hours. Valuing students' time at $10 per hour, the total dollar value of the welfare gain from the switch is roughly $57 million, or 70% of the $82 million NHPS spent on teachers in 2014-2015 (NHPS, 2014). These are large effects for a change that is close to costless. For a benchmark, the well-known Project STAR experiment reduced class size by about 30%, from 22 students per class to 15 (Krueger, 1999; Chetty et al., 2011).[21]

[21] Our one-mile walk zone threshold is conservative relative to state guidelines for high school students; see Lohman (2014). We are also conservative in several other dimensions. Drive-times are based on car travel; buses are slower. Students in cars and younger walking students are often accompanied by adults, whose welfare we do not include in our calculation. Our $10 per hour valuation of time is based on the minimum wage in Connecticut, which was $10.10 in January 2017. For the average student, the present value of an hour of school attendance is likely higher. Finally, we do not include Pre-K students even though many Pre-K students also use the choice process.

E Appendix E: Fieldwork overview

The survey was implemented in 2015 and then again in 2017. The two surveys were similar in scope and objectives. We present the details of each survey below.

E.1 2015 Survey Procedures

E.1.1 Data

Administrative student-level data was procured in coordination with the New Haven Public Schools (NHPS). The data contained information for approximately 20,000 students present in the NHPS' enrollment records, and included student race, gender, school lunch status, test scores, and other information. Similar to the city of New Haven's resident population, the NHPS has a majority-minority student body, where nearly 60% of students are eligible for free lunch and more than 80% of students are black or Hispanic.

E.1.2 Sample Selection

The survey universe was sampled from the population of enrolled students in the New Haven Public Schools. The students were observed in enrollment administrative records. Only households with students that applied for either Kindergarten or 9th grade were selected for the survey. $1,589$ households with children applying to Kindergarten were selected, while $1,423$ households with children applying to 9th grade were selected.

E.1.3 Survey Implementation

Survey Overview The survey asked the parents or guardians of past school choice applicants questions about:

- Their knowledge of the administrative aspects of the school placement process.
- Their own involvement in both the school placement, and school choice process.
- How they obtained information about the process.
- Their preferences regarding school attributes

The survey was programmed using SurveyCTO and loaded on Samsung Galaxy Tab 7s tablets. The survey was also tested in small focus groups on three occasions during the two months prior to the field work.

Survey Team The team of surveyors was composed by ten active members who were recruited using online advertisement and Yale University's physical bulletin boards. All the surveyors received a two-day training that prepared them for the use of the tablet and regulation regarding interacting with human subjects. Almost half of the surveyors were bilingual English, and Spanish speakers which was useful given that a significant proportion of the population in New Haven is Hispanic.

Surveyor Training The two day training covered the following topics:

- Day 1: Introduction regarding data confidentiality and safety. Logistics procedures were discussed.
- Day 2: Practical training of the instrument in a random neighborhood where we tested their skills to approach the families and their accuracy while using the instrument.
- CITI Certificate: All surveyors had to complete an online course for IRB purposes where they learned about dealing with human research subjects, and confidential information.

Outreach Parent's participation was voluntary, and there was no compensation (neither monetary, nor non-monetary) for their participation.

- In partnership with the NHPS, the district contacted the households via phone-calls to announce their participation in the project.
- When the surveyors visited each house, they announced the project and handed in a business card (See Figure E2) with the study's contact information. Parents or guardians who agreed to participate signed an informed consent form.
- In case of finding no one at home, a door hanger (See Figure E4) was left with contact information.
- Surveyors also had the chance to reschedule the interview if the respondent had time issues at the moment.

Administration Survey personnel followed a pre-defined protocol while out on the field:

- Surveyors wore branded t-shirts and IDs identifying them as part of the survey team.
- If the surveyor was attempting a door-to-door survey:
- Surveyors approached selected households and introduced themselves. They asked if the person answering is the parent or guardian of an NHPS student eligible for the kindergarten school choice.

- If the parent or guardian was present, the surveyor went through the remainder of the introduction, and then through the consent script. The script identified the surveyor as a member of the team and briefly describes the project. Respondents also received a business card containing the survey team's contact information.

- Parents who agreed to continue were administered the survey with the surveyor, knowing from the consent form that they are free to interrupt their participation at any time.

E.2 2017 Survey Procedures

E.2.1 Data

For the 2017 fieldwork, administrative data was procured in the same fashion as the 2015 process. The NHPS gave the project access to student enrollment records (26,780 actively enrolled), which include race, gender, English-language learner (ELL) status, and special education status. Similar to the 2015 process, more than 80% of the student body is either black or Hispanic, while about 15% of the actively enrolled students were in an ELL program.

E.2.2 Sample Selection

The survey universe corresponds to the population of enrolled students in the NHPS. From this sample universe the following conditions were applied to select the sample:

- Keep only 8th grade students.
- Keep only New Haven residents.
- Keep only students with current enrollment status.

The sample universe consisted of $1,589$ students.

E.2.3 Survey Implementation

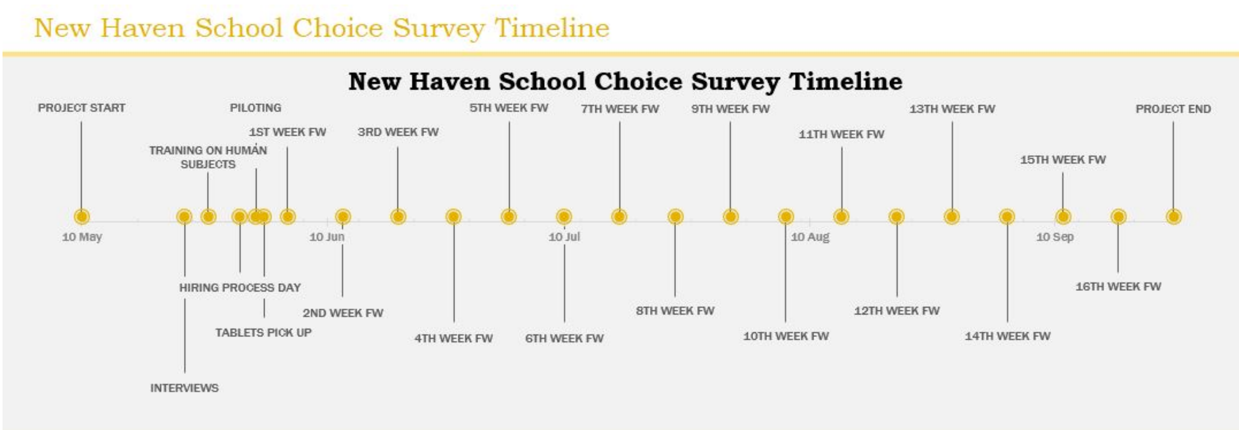


Figure E1: 2017 Fieldwork Timeline

Survey Overview The survey asked the parents or guardians of school choice eligible kindergarten children questions about:

- Their knowledge of the administrative aspects of the school placement process.
- The things parents are most interested in when choosing a school.
- What their perceptions are regarding the process.
- Their knowledge about the New Haven Promise program for college funding.

Two surveys were programmed using SurveyCTO and loaded onto tablets, which the surveyors used to administer the surveys. The first survey was administered to parents, while the second one was administered to students, under the parent's consent. Both surveys were piloted, and tested in two instances: first on a group of surveyors surveying each other, and secondly, on a small field pilot. Both surveys were administered during the household visit, given the parent's consent. The parent's survey took approximately 25 minutes to complete while the student's survey took approximately 15 minutes. Surveys were administered from early June 2017 until late September 2017. However, due to a SurveyCTO coding error, beliefs questions were not asked to households who did not participate in the NHPS' School Choice process (62 households). These households were resurveyed from late December 2017 to late January 2018, resulting in 20 re-surveyed households.

Survey Team Surveyors were recruited via open job calls posted on both physical, and digital university job boards, and online job websites. Additionally, local universities' social sciences departments were contacted so that a notice about the position was sent to their respective mailing lists.

The goal was to build a team of six surveyors that are representative of the NHPS' student population, and organize them into three teams of two. Three Spanish speaking surveyors were hired, and matched with non-Spanish speaking surveyors who were also representative of the student population. Each surveyor worked a total of 21 hours per week, which amounts to three 7-hour work days. Surveyor remuneration was as follows: The hourly rate for surveyors was $12, plus a bonus of $20 per completed survey. The requirements and details of the position included:

- Age: 21 years and above.
- Language requirements: Good Communication skills and ability to clearly read and write in English. Spoken and written Spanish is a plus.

- Additional requirements: A responsible, reliable dependable worker, who preferably lives in or is familiar with New Haven.
- Mandatory training session: Firstly, surveyors will attend a one-day compulsory, on-site training session. Secondly, surveyors will have to complete both the CITI Training's 'National Bureau of Economic Research (NBER) Responsible Conduct of Research (RCR) for Social & Behavioral', and the National Institutes of Health's (NIH) 'Protecting Human Research Participants' (PHRP) online trainings.
- Work Schedule: 21 hours per week.

Surveyor Training Training consisted of a two-part program led by the fieldwork coordinators. The first part consisted of two training sessions: The first was an on-site training, with the purpose of going over surveying and data collection best practices, while the second one was the completion of both the National Bureau of Economic Research (NBER) CITI Training for RCR and the National Institute of Health's (NIH) Protecting Human Research Participants (PHRP) online training courses. The second part consisted of a guided field exercise designed to reinforce concepts learned during the first training session. All coordinators and surveyors were compelled to attend, and complete all parts of the training program in order to be qualified for the data collection process.

Outreach The main criteria for recruitment was that families have children who are eligible for participating in the 9th grade School Choice program ran by the NHPS during the 2017-2018 school year.

Outreach was implemented over two dimensions:

- **Phone contact**

  The fieldwork coordinators established a call center at office space used by the project's team. This call center would emit calls attempting to recruit survey participants using the confidential contact information that was shared by the NHPS - which was done after surveyors were administered the required training. The initial phone call attempted to inform potential survey participants about the study, while also attempting to schedule a home visit. Depending on the outcome of the phone call, contacts were categorized into different priority groups, i.e. the participant's phone is out of service, the call went straight to voicemail, or the call rang but there's no voicemail set up. If the call was not successful, but leaving a voicemail was an option, a voicemail was left. It included a contact phone number, and an email address so potential participants wishing to schedule an appointment or decline their participation could do so at their own convenience. All households were called three times, except for when on one of these calls participation was declined.

- **Door to door**

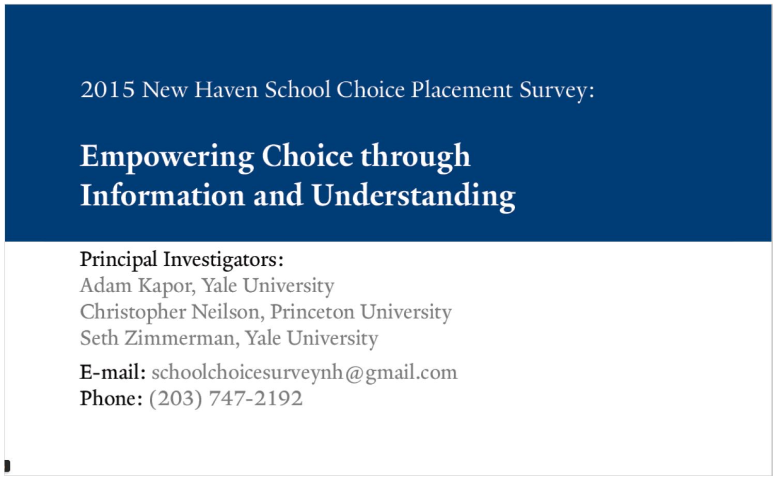
  Parallel to the call center, surveyors were in the field visiting households. Households were visited mostly during the work week, with a pair of surveyors working throughout the weekends. Door-to-door surveying was also mostly done in the afternoon and into the evening as an attempt to coincide with the work schedule of potential participants. However, surveyors were mostly using their time going to appointments scheduled over the phone, and only did door-to-door attempts in between scheduled appointments. Whenever surveyors were visiting households, business cards were used (See Figure E3) when potential participants were not home. All households were visited three times, except for when one of these attempts resulted in a decline.

Potential participants were also informed that participation in the study was completely voluntary, and without compensation.

Administration Survey personnel followed a pre-defined protocol while out on the field:

- Surveyors wore branded t-shirts and IDs identifying them as part of the survey team.
- If the surveyor was fulfilling a previous phone scheduled appointment:
  - Surveyors approached the scheduled household during the specified time and introduced themselves. They asked if the person answering is the parent or guardian of an NHPS student eligible for the kindergarten school choice.
- If the surveyor was attempting a door-to-door survey:
  - Surveyors approached selected households and introduced themselves. They asked if the person answering is the parent or guardian of an NHPS student eligible for the kindergarten school choice.
- If the parent or guardian was present, the surveyor went through the remainder of the introduction, and then through the consent script. The script identified the surveyor as a member of the team and briefly describes the project. Respondents also received a business card containing the survey team's contact information.
- Parents who agreed to continue were administered the survey with the surveyor, knowing from the consent form that they are free to interrupt their participation at any time.

E.2.4 Field Materials Used



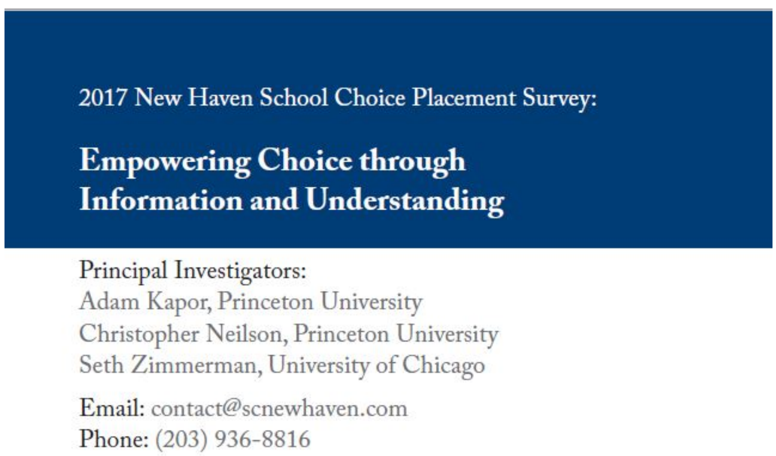
2015 New Haven School Choice Placement Survey:

Empowering Choice through Information and Understanding

Principal Investigators:
Adam Kapor, Yale University
Christopher Neilson, Princeton University
Seth Zimmerman, Yale University

E-mail: schoolchoicesurveynh@gmail.com
Phone: (203) 747-2192

Figure E2: Business card used during 2015 process



2017 New Haven School Choice Placement Survey:

Empowering Choice through Information and Understanding

Principal Investigators:
Adam Kapor, Princeton University
Christopher Neilson, Princeton University
Seth Zimmerman, University of Chicago

Email: contact@scnewhaven.com
Phone: (203) 936-8816

Figure E3: Business card used during 2017 process

Figure E4: Door hanger used during fieldwork