Online Appendix:

Evidence-Based Decisions and Education Policymakers

Nozomi Nakajima *

November 2021

<Most recent version here>

*Ph.D. candidate, Harvard University (nnakajima@g.harvard.edu).

A Additional Figures and Tables

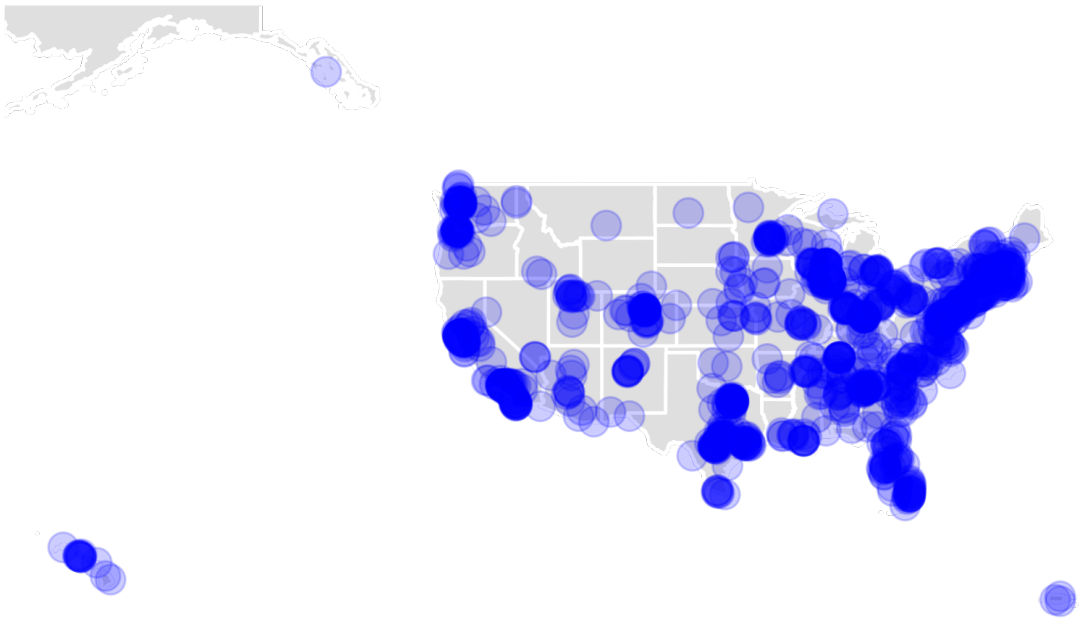


Figure A1: Geographic location of policymakers in study sample

Note: This figure shows the geographic location of the education agency that employs the policymakers in my study sample (N=2,079). The ZIP codes of the education agencies come from administrative records described in Section 3. The policymakers work in 49 states, Washington, D.C. and Puerto Rico.

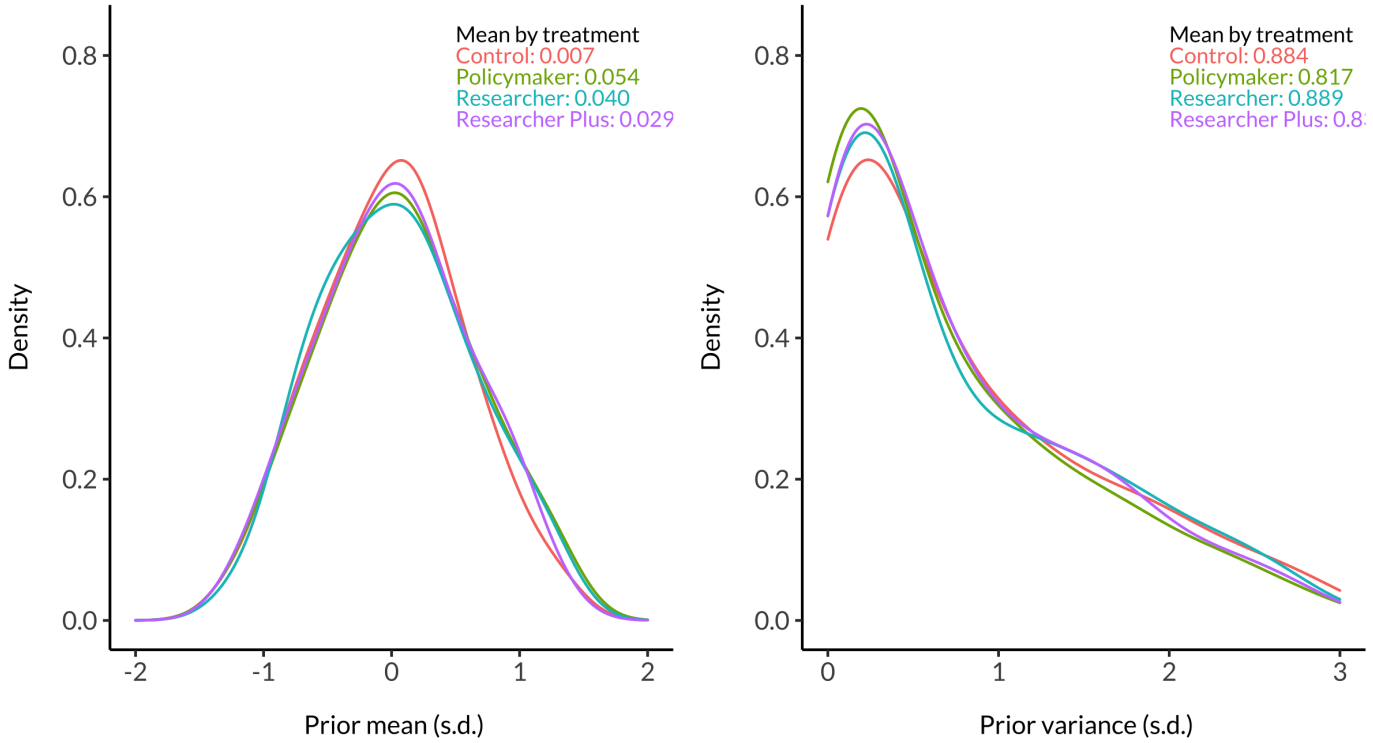


Figure A2: Prior mean and variance by treatment assignment in information experiment

Note: This figure presents the mean and variance of policymakers' predictions about the effect of urban charter schools on student achievement, grouped by information treatment. These responses were measured at the beginning of the survey. Policymakers reported support points and probabilities associated with each support point. Sample size is 2,079 policymakers.

Table A1: Policymaker preferences by model specifications

Dependent variable:	Forced choice		Percent weight
Model specification:	Linear probability	Logit	Linear regression
	(1)	(2)	(3)
Sample (ref: 500 students)			
2,000 students	0.109***	0.459***	5.137***
	(0.010)	(0.041)	(0.437)
15,000 students	0.164***	0.689***	7.983***
	(0.010)	(0.042)	(0.457)
Sites (ref: 1 site)			
10 sites	0.140***	0.592***	6.735***
	(0.009)	(0.039)	(0.419)
25 sites	0.142***	0.598***	6.919***
	(0.009)	(0.040)	(0.440)
Poverty (ref: +/- 45 pp)			
+/- 25 pp	0.052***	0.218***	2.150***
	(0.010)	(0.041)	(0.432)
+/- 5 pp	0.112***	0.474***	5.016***
	(0.010)	(0.043)	(0.446)
Race (ref: +/- 45 pp)			
+/- 25 pp	0.040***	0.169***	2.132***
	(0.009)	(0.040)	(0.405)
+/- 5 pp	0.031**	0.130**	2.204***
	(0.009)	(0.040)	(0.417)
Urban (ref: Different)			
Mix	0.126***	0.523***	5.832***
	(0.008)	(0.035)	(0.393)
Congruent	0.079***	0.328***	4.128***
	(0.009)	(0.037)	(0.426)
Design (ref: Observational)			
Experimental	0.014	0.060	0.174
	(0.007)	(0.031)	(0.339)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents the estimates of the effects of the randomly assigned study attributes on the probability of being selected to inform policymakers' decisions (forced choice) and the percent weight on policymakers' decisions (percent weight). Estimates are based on the regression specification in equation 1. Each column is the result of a different model specification. Robust standard errors are clustered at the respondent level. The reference category for each attribute is denoted in brackets.

Table A2: Robustness check of policymaker preference for research

	Panel A. Outcome: Forced choice									
	Task number		Order of study		Order of attributes		Order of questions		Placebo	
Attribute	F-statistic	p-value	F-statistic	p-value	F-statistic	p-value	F-statistic	p-value	Coef	(S.E.)
Design	0.49	0.743	0.131	0.718	1.867	0.172	1.186	0.276		
Sample	1.001	0.432	0.002	0.998	2.912	0.054	2.199	0.111		
Sites	0.839	0.568	0.08	0.923	0.394	0.674	0.766	0.465		
Poverty	1.66	0.103	0.597	0.551	1.223	0.294	0.207	0.813		
Race	1.426	0.18	1.167	0.311	0.224	0.799	1.53	0.216		
Urban	1.055	0.392	1.316	0.268	1.831	0.16	0.696	0.499		
Color of column									0.003	(0.007)
	Panel B. Outcome: Percent weight									
	Task number		Order of study		Order of attributes		Order of questions		Placebo	
Attribute	F-statistic	p-value	F-statistic	p-value	F-statistic	p-value	F-statistic	p-value	Coef	(S.E.)
Design	0.296	0.880	0.676	0.411	1.914	0.167	0.035	0.851		
Sample	1.599	0.119	0.087	0.917	1.805	0.165	1.054	0.349		
Sites	1.077	0.376	0.844	0.430	0.151	0.860	0.711	0.491		
Poverty	0.843	0.565	0.081	0.922	0.944	0.389	0.207	0.813		
Race	1.760	0.080	2.271	0.103	0.194	0.824	1.774	0.170		
Urban	0.663	0.725	0.965	0.381	1.175	0.309	0.875	0.417		
Color of column									0.294	(0.320)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table presents robustness checks for the effects of the randomly assigned study attributes on the probability that the study is selected to inform policymakers' decisions (Panel A) and the percent weight of the study on policymakers' decisions (Panel B). Each column is the result of a separate regression (N=20,790). Column 1 interacts study attributes with task number, which can range from task 1 to task 5. Column 2 interacts study attributes with the order of study, which can take on values of study A or study B. Column 3 interacts study attributes with the order of attributes, which can take on values from first to sixth. Column 4 interacts study attributes with the order of questions, which can either be that the forced choice question appeared first or the percent weight question appeared first. Columns 1-4 report the F-statistic and corresponding p-value for a test of whether the effect of the study attribute is equivalent across the interaction term. Column 5 reports the effect of the placebo (the color of the shaded column in the task table) on policymakers' responses. Robust standard errors clustered at the respondent level in parentheses.

Table A3: Balance of baseline characteristics of policymakers in analytic sample

	No information		Peer policymaker		Researcher		Researcher Plus		p-value of diff. in means		
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	(3) - (1)	(5) - (1)	(7) - (1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Female	0.792	0.406	0.811	0.392	0.779	0.415	0.795	0.404	0.462	0.605	0.931
White	0.619	0.486	0.622	0.485	0.639	0.481	0.631	0.483	0.913	0.511	0.675
Black	0.148	0.355	0.115	0.319	0.124	0.330	0.165	0.372	0.128	0.270	0.446
Hispanic	0.104	0.305	0.113	0.317	0.114	0.318	0.090	0.287	0.645	0.621	0.464
Asian	0.072	0.259	0.078	0.268	0.063	0.243	0.060	0.237	0.783	0.166	0.389
District Leader	0.447	0.498	0.456	0.499	0.404	0.491	0.420	0.494	0.414	0.992	0.450
District Admin	0.355	0.479	0.331	0.471	0.355	0.479	0.378	0.485	0.717	0.131	0.947
State Leader	0.060	0.238	0.068	0.252	0.069	0.254	0.065	0.247	0.613	0.561	0.723
Data Role	0.172	0.377	0.158	0.365	0.118	0.323	0.167	0.373	0.567	0.016	0.843
Urban	0.575	0.495	0.567	0.496	0.576	0.495	0.543	0.499	0.797	0.970	0.309
Rural	0.096	0.295	0.094	0.293	0.114	0.318	0.117	0.322	0.943	0.361	0.271
% FRPL students	52.279	28.083	51.721	29.156	52.402	28.618	50.267	28.197	0.760	0.946	0.254
% white students	42.443	28.201	41.637	28.325	39.726	28.006	42.119	28.060	0.654	0.128	0.854
Scientific reasoning scale	7.701	1.525	7.817	1.608	7.590	1.522	7.701	1.565	0.243	0.254	1.000
Absolute deviation of data and guess of FRPL students	9.775	8.745	9.991	9.043	10.432	9.551	10.083	9.422	0.704	0.259	0.588
Absolute deviation of data and guess of white students	9.675	9.247	10.430	9.776	9.981	9.129	9.954	9.531	0.213	0.600	0.636
Observations	501		487		493		521				

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table summarizes policymaker characteristics collected before the information treatment in the survey experiment for my analytic sample (N-2,002). Columns 1-8 report means and standard deviations for each treatment group. Columns 9-11 report p-values of the mean differences between treatment groups.

Table A4: Balance of baseline characteristics of policymakers that started surveys

	No information		Peer policymaker		Researcher		Researcher Plus		p-value of diff. in means		
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	(3) - (1)	(5) - (1)	(7) - (1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Attrited (1 = Yes)	0.0649	0.247	0.0737	0.261	0.0722	0.259	0.0843	0.278	0.578	0.641	0.209
Female	0.778	0.416	0.816	0.388	0.769	0.422	0.799	0.401	0.123	0.708	0.386
White	0.629	0.484	0.613	0.487	0.637	0.481	0.642	0.48	0.596	0.774	0.631
Black	0.144	0.352	0.12	0.325	0.121	0.326	0.164	0.37	0.232	0.255	0.362
Hispanic	0.103	0.304	0.11	0.314	0.116	0.32	0.086	0.281	0.676	0.494	0.335
Asian	0.068	0.253	0.077	0.267	0.069	0.253	0.056	0.229	0.793	0.212	0.965
District Leader	0.441	0.497	0.449	0.498	0.404	0.491	0.44	0.497	0.493	0.968	0.747
District Admin	0.357	0.479	0.337	0.473	0.356	0.479	0.366	0.482	0.816	0.181	0.597
State Leader	0.061	0.24	0.068	0.252	0.07	0.256	0.064	0.245	0.644	0.54	0.844
Data Role	0.173	0.379	0.157	0.364	0.121	0.326	0.162	0.369	0.464	0.015	0.616
Urban	0.566	0.496	0.571	0.495	0.569	0.496	0.55	0.498	0.864	0.925	0.586
Rural	0.094	0.292	0.098	0.297	0.108	0.311	0.116	0.321	0.826	0.42	0.211
% FRPL students	52.085	27.995	52.227	29.169	52.765	28.592	50.567	27.908	0.935	0.689	0.358
% white students	42.4	28.088	41.059	28.361	39.72	28.29	42.285	27.992	0.432	0.114	0.945
Scientific reasoning scale	7.721	1.521	7.796	1.596	7.621	1.538	7.693	1.563	0.427	0.278	0.762
Absolute deviation of data and guess of FRPL students	9.974	8.895	10.117	9.262	10.223	9.377	9.993	9.404	0.795	0.65	0.972
Absolute deviation of data and guess of white students	9.723	9.473	10.504	9.807	10.099	9.216	9.968	9.584	0.18	0.504	0.664
Observations	555		543		554		593				

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table summarizes policymaker characteristics collected before the information treatment in the survey experiment for the full sample (i.e., anyone who started the survey). Columns 1-8 report means and standard deviations for each treatment group. Columns 9-11 report p-values of the mean differences between treatment groups. Attrition is defined as respondents who did not complete the main survey and/or the follow-up survey.

B Survey Details

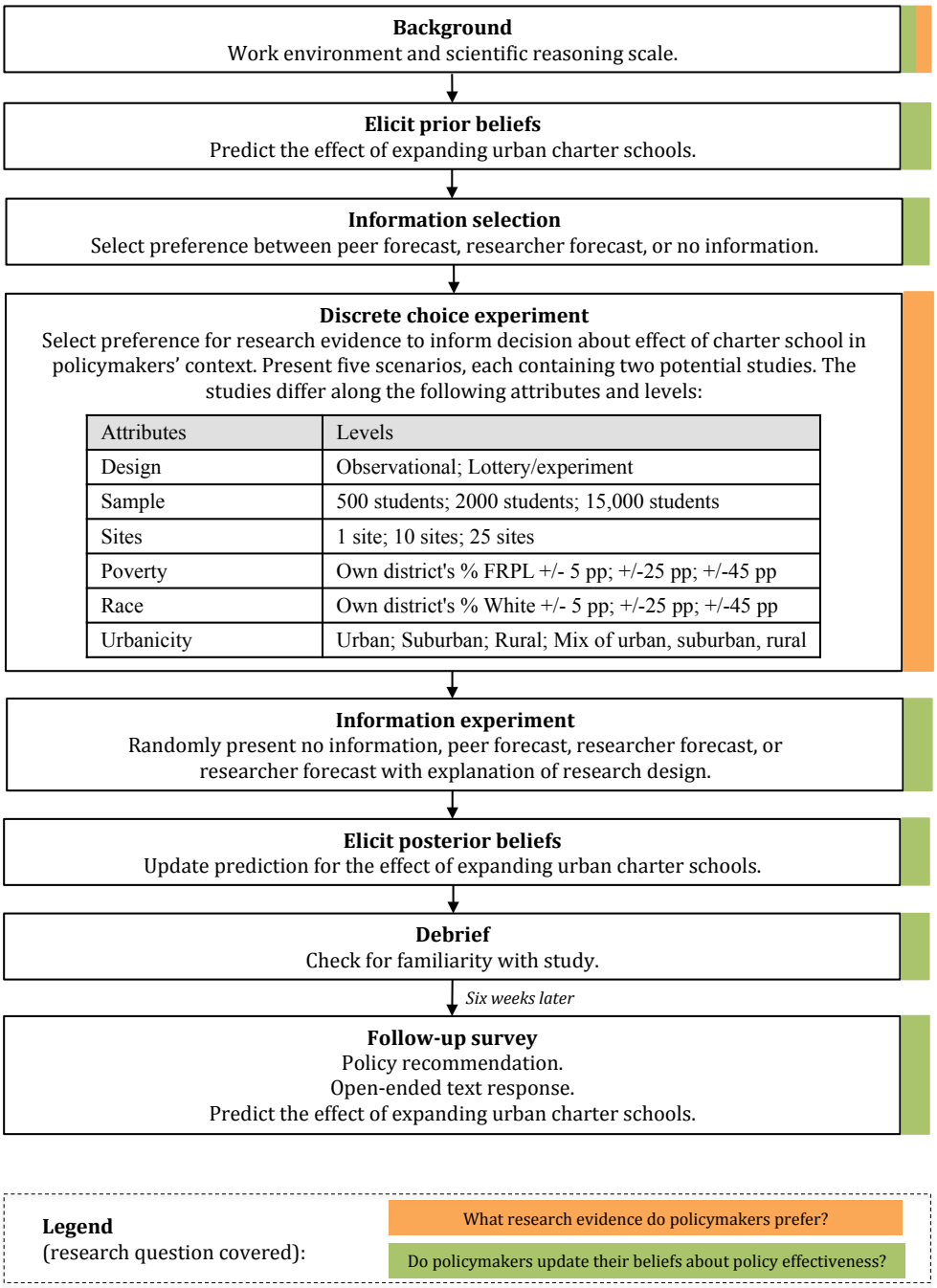
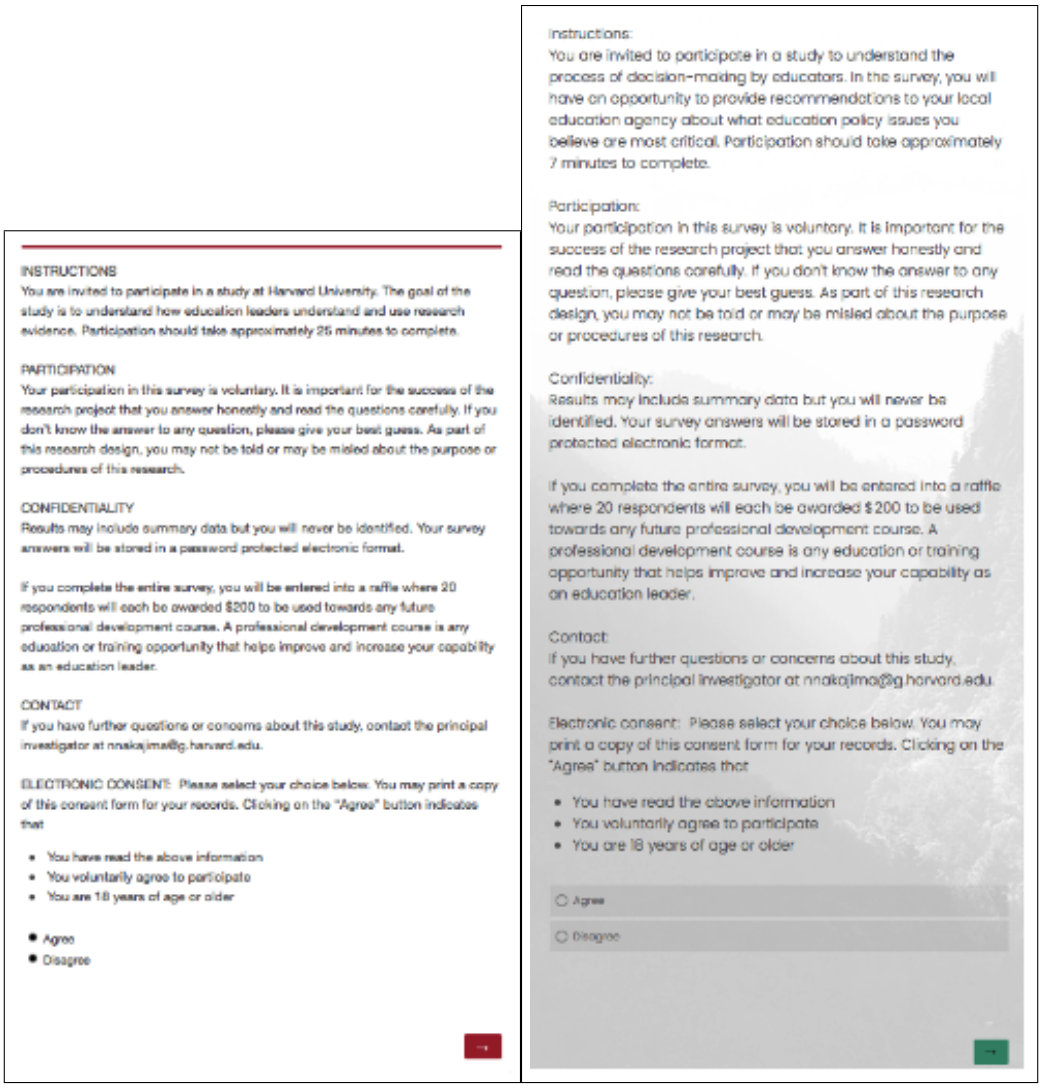


Figure B1: Overview of the survey design

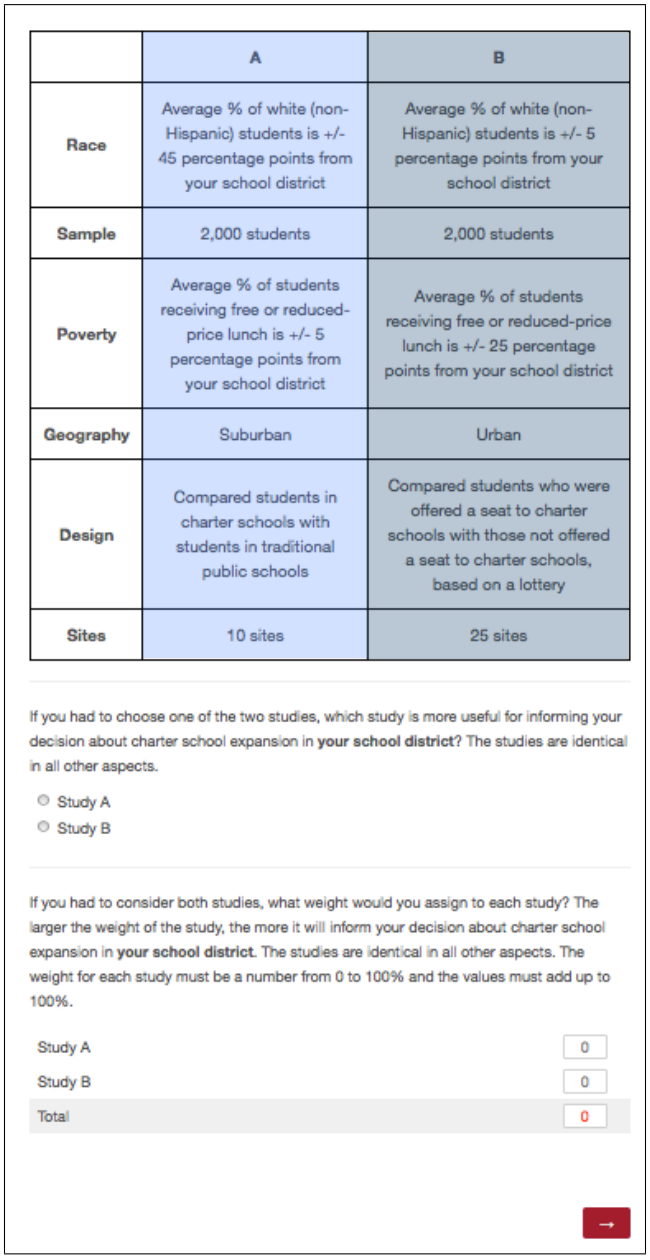
Note: This figure shows the flow of the survey. Each part of the survey contains questionnaire/items relevant to one of the two research questions (see legend). An exception is the background section, which contains questionnaire items used to address both research questions.



(a) Main survey **(b)** Follow-up survey

Figure B2: Instructions

Note: This figure shows screenshots of the instruction page at the beginning of the main survey (a) and follow-up survey (b). To obfuscate the connection between the main survey and the follow-up survey, the surveys use different layouts and fonts.



	A	B
Race	Average % of white (non-Hispanic) students is +/- 45 percentage points from your school district	Average % of white (non-Hispanic) students is +/- 5 percentage points from your school district
Sample	2,000 students	2,000 students
Poverty	Average % of students receiving free or reduced-price lunch is +/- 5 percentage points from your school district	Average % of students receiving free or reduced-price lunch is +/- 25 percentage points from your school district
Geography	Suburban	Urban
Design	Compared students in charter schools with students in traditional public schools	Compared students who were offered a seat to charter schools with those not offered a seat to charter schools, based on a lottery
Sites	10 sites	25 sites

Figure B3: Example task in discrete choice experiment

Note: This figure shows a screenshot of an example task in the discrete choice experiment. The full set of possible attribute values are described in Section 2.2 of the main manuscript.

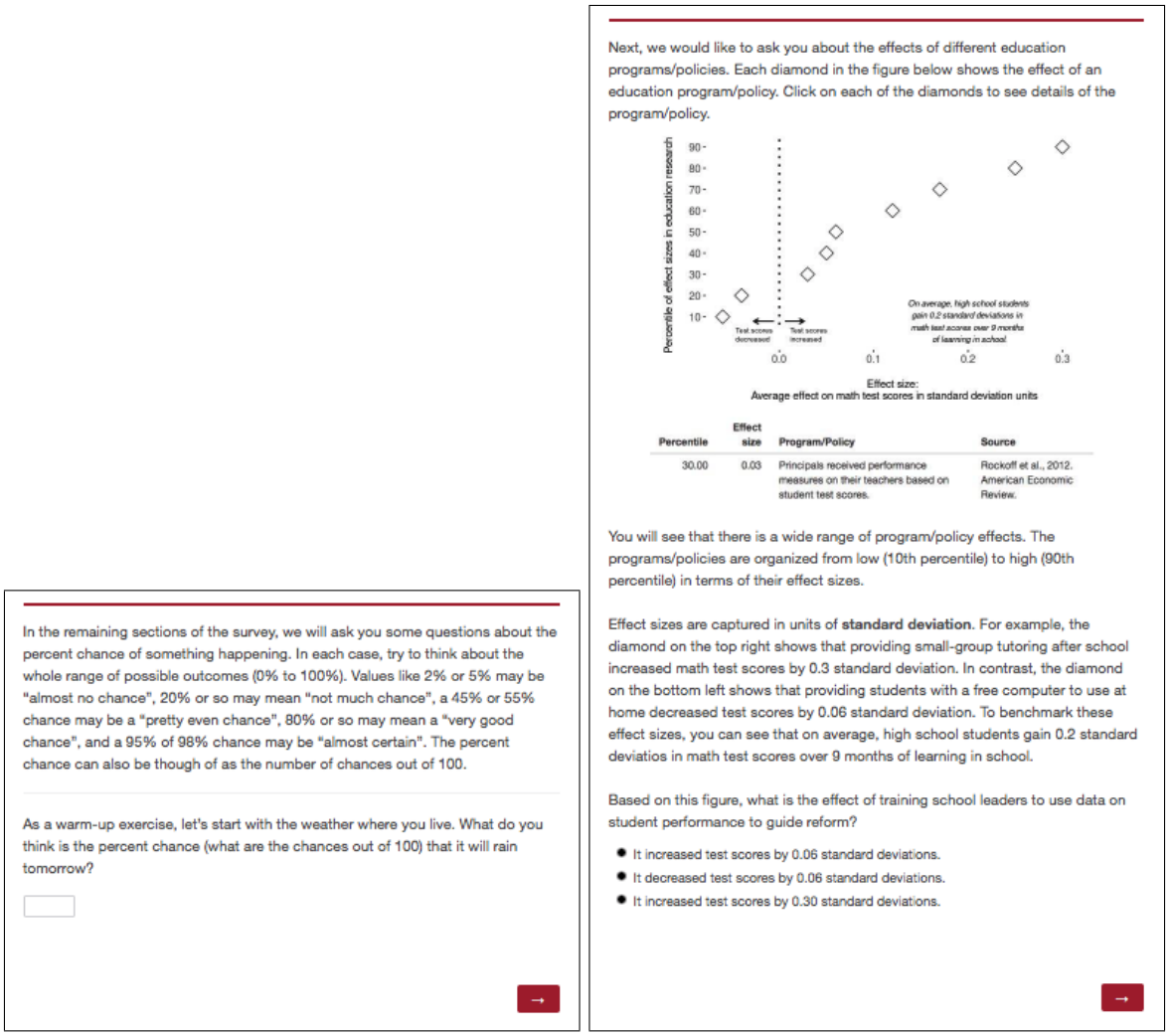


Figure B4: Warm-up exercises

Note: This figure shows screenshots of the warm-up exercises in the information experiment. The first question is modeled after the opening questions used in the New York Federal Reserve Board's Survey of Economic Expectations and the University of Michigan Survey of Consumers. It is intended to familiarize respondents with answering probability questions. The second question is an interactive data visualization that familiarizes respondents with interpreting effect sizes and understanding the range of effect sizes that have been reported in education research. The interactive data visualization was created based on the range of effect sizes and interventions described in Kraft (2020). Readers can interact with the visualization here: https://edpolicyeffectsize.shinyapps.io/effectsize shinyapp/

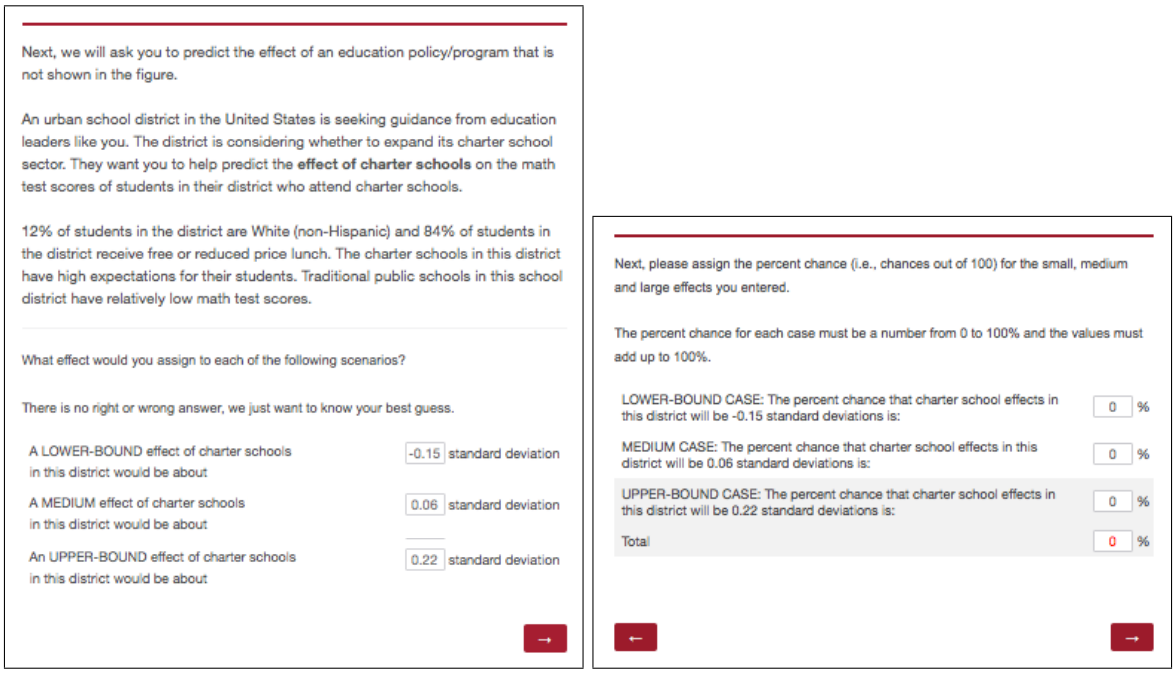
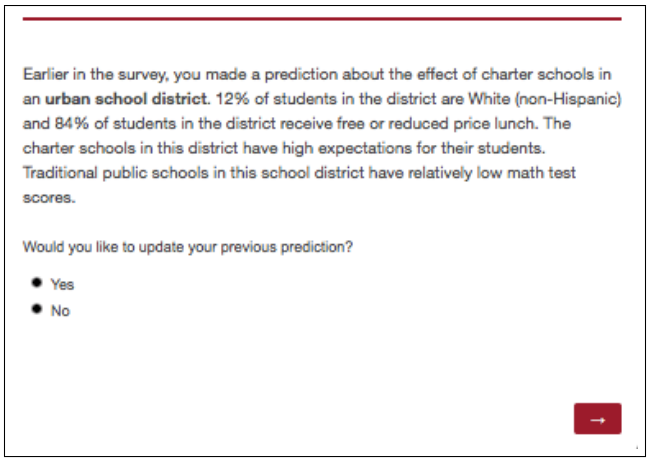
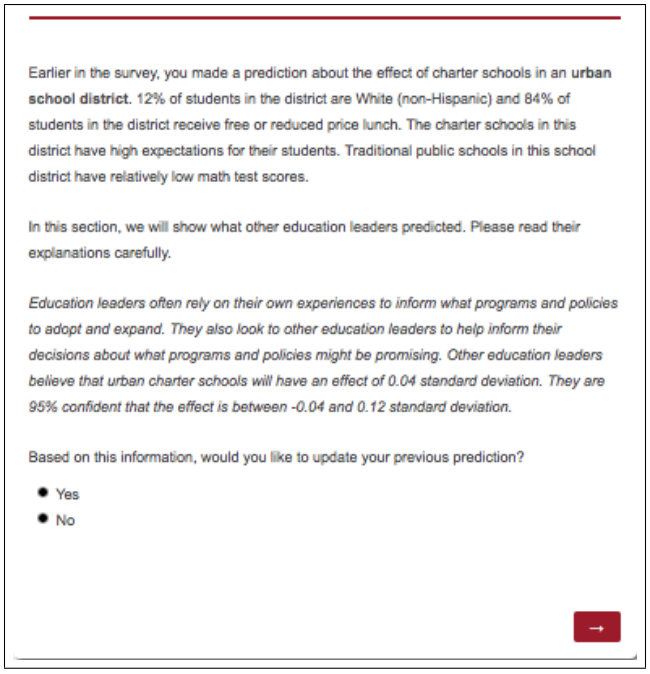


Figure B5: Belief elicitation task

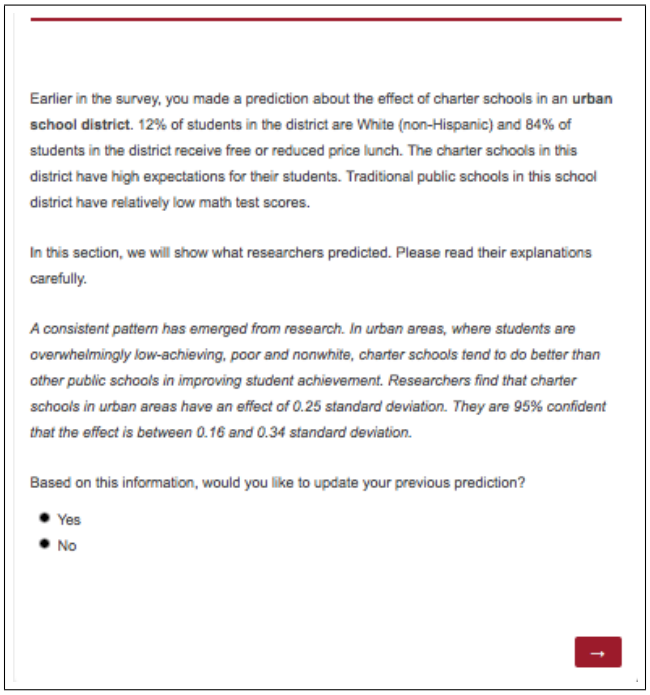
Note: This figures shows screenshots of the belief elicitation task in the information experiment. Modeled after Altig et al. (2020), respondents are asked to freely select three support points and then assign probabilities to each. The example above shows a respondent who entered the support points of -0.15, 0.06, and 0.22 standard deviation. These entries are automatically updated so that they are displayed when assigning probabilities.



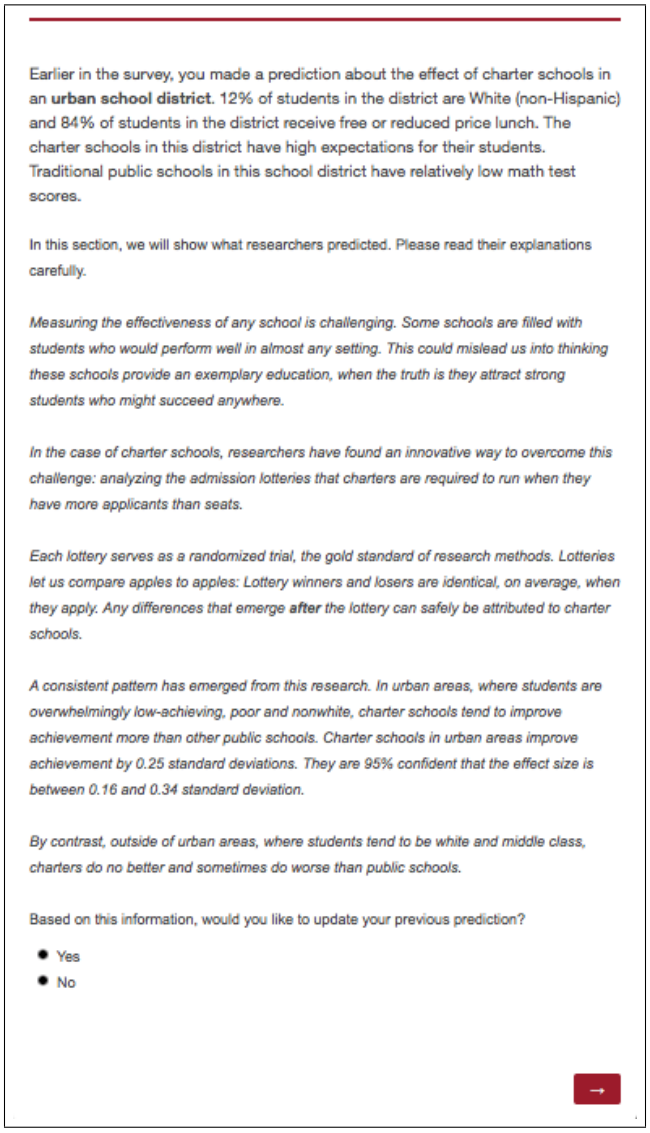
(a) Control group



(b) Peer policymaker treatment group



(c) Researcher treatment group



(d) Researcher plus treatment group

Figure B6: Information experiment

Note: This figures shows screenshots of the information intervention. The four conditions (control, policymaker, researcher, and researcher-plus) were randomly assigned such that each respondent only saw one of the four screenshots (a)-(d). Details of the information experiment are described in Section 2.2. of the main manuscript

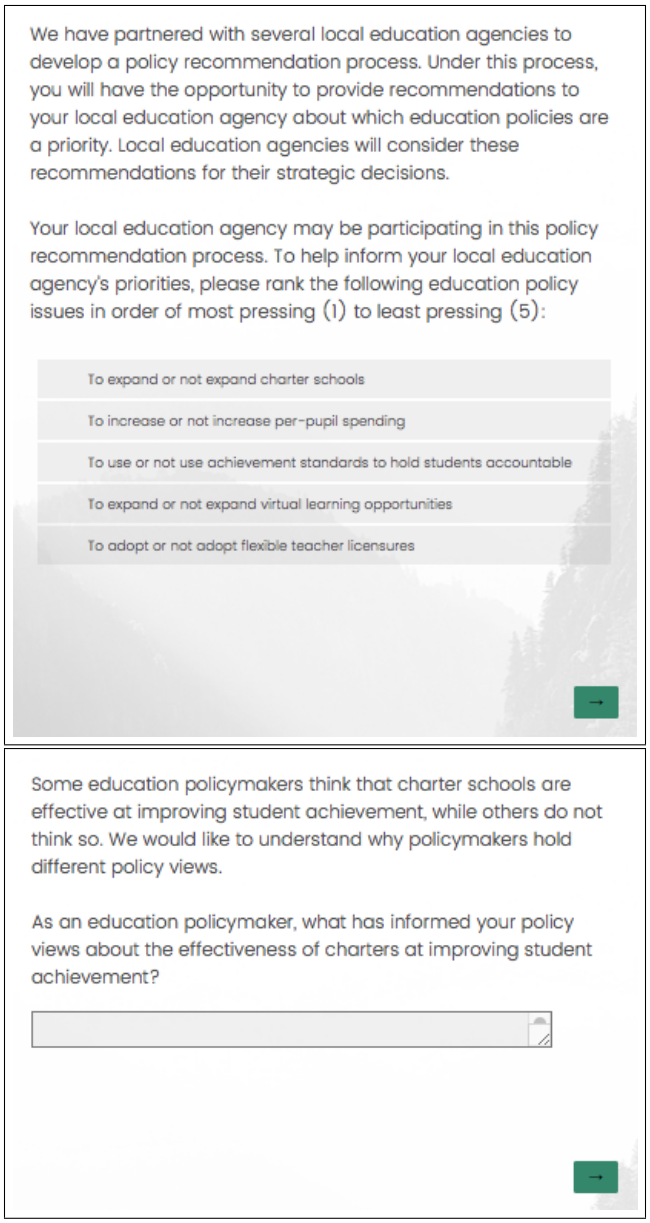


Figure B7: Follow-up survey

Note: This figures shows screenshots of the policy recommendation process and open-ended text in the follow-up survey.

C Designing the Discrete Choice Experiment

A key design consideration for the discrete choice experiment is the realism of the studies encountered by the respondents. As noted in Section 2.2, there is a large repository of charter school effectiveness studies with considerable variation along the six attributes studied in this paper. The table below summarizes the various charter school studies along some of the attributes that were varied in this paper. The levels of each attribute in the discrete choice experiment were designed based on the range of plausible values based on the table below.

Study	**Research design**	**Sample size**	**No. of schools**	**Setting**	**Urbanicity**	**% FRPL**	**% White**
Abdulkadiroğlu, A., Angrist, J.D., Dynarski, S.M., Kane, T.J. and Pathak, P.A., 2011. "Accountability and flexibility in public schools: Evidence from Boston's charters and pilots".The Quarterly Journal of Economics, 126(2), pp.699-748.	Lottery	2,837	8 schools	Boston, MA	Urban	71.4	16.1
Cohodes, S., Setren, E.M., Walters, C.R., Angrist, J.D. and Pathak, P.A., 2013. "Charter school demand and effectiveness: A Boston update". Massachusetts Dept. of Elementary and Secondary Education.	Lottery	15,311	13 schools	Boston, MA	Urban	72.1	18.1
Angrist, J.D., Pathak, P.A. and Walters, C.R., 2013. "Explaining charter school effectiveness". American Economic Journal: Applied Economics, 5(4), pp.1-27.	Lottery	9,141	26 schools	MA	Urban, Suburban, Rural	40.6	63.9
Angrist, J.D., Dynarski, S.M., Kane, T.J., Pathak, P.A. and Walters, C.R., 2012. "Who benefits from KIPP?". Journal of Policy Analysis & Management, 31(4), pp.837-860.	Lottery	856	1 school	Lynn, MA	Urban	82.5	16.8

Abdulkadiroğlu, A., Angrist, J.D., Hull, P.D. and Pathak, P.A., 2016. "Charters without lotteries: Testing takeovers in New Orleans and Boston". American Economic Review, 106(7), pp.1878-1920.	Observational	2,205	11 schools	Boston, New Orleans	Urban	80.4-91.2	1.9-13.5
Gleason, P., Clark, M., Tuttle, C.C. and Dwoyer, E., 2010. "The Evaluation of Charter School Impacts: Final Report". National Center for Education Evaluation and Regional Assistance, NCEE 2010-4029.	Lottery	2,330	36 schools	15 states	Urban, Suburban, Rural	33	56
Tuttle, C.C., Gleason, P., Knechtel, V., Nichols-Barrer, I., Booker, K., Chojnacki, G., Coen, T. and Goble, L. 2015. "Understanding the Effect of KIPP as It Scales: Vol. I: Impacts on Achievement and Other Outcomes". Mathematica Policy Research, Inc.	Lottery	1,988	24 schools	KIPP elementary & middle schools in 10 states	Urban, Suburban, Rural	81	1.2
Tuttle, C.C., Gill, B., Gleason, P., Knechtel, V., Nichols-Barrer, I. and Resch, A., 2013. "KIPP Middle Schools: Impacts on Achievement and Other Outcomes. Final Report". Mathematica Policy Research, Inc.	Lottery	977	13 school	KIPP middle schools in 6 states	Urban, Suburban, Rural	79.7	1.4
Furgeson, J., Gill, B., Haimson, J., Killewald, A., McCullough, M., Nichols-Barrer, I., Verbitsky-Savitz, N., Teh, B.R., Bowen, M., Demeritt, A. and Hill, P., 2012. "Charter-school management organizations: Diverse strategies and diverse student impacts." Mathematica Policy Research, Inc.	Observational	1,630	16 schools	14 states	Urban, Suburban, Rural	76	7
Hoxby, C.M., Murarka, S. and Kang, J., 2009. "How New York City's charter schools affect achievement". New York City Charter Schools Evaluation Project, pp.1-85.	Lottery	14,903	42 schools	NYC	Urban	92	14.3

Dobbie, W. and Fryer Jr, R.G., 2013. "Getting beneath the veil of effective schools: Evidence from New York City". American Economic Journal: Applied Economics, 5(4), pp.28-60.	Lottery	31,779	29 schools	NYC	Urban	87	2
Dobbie, W. and Fryer Jr, R.G., 2015. "The medium-term impacts of high-achieving charter schools". Journal of Political Economy, 123(5), pp.985-1037.	Lottery	541	1 school	NYC	Urban	87.8	1.8
Dobbie, W. and Fryer Jr, R.G., 2011. "Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone". American Economic Journal: Applied Economics, 3(3), pp.158-87.	Lottery	842	1 school	NYC	Urban	68	2.4
Hoxby, C.M. and Rockoff, J.E., 2005. "The impact of charter schools on student achievement". Working Paper.	Lottery	2,664	3 schools	Chicago	Urban	80	14
Hastings, J.S., Neilson, C.A. and Zimmerman, S.D., 2012. "The effect of school choice on intrinsic motivation and academic outcomes". NBER Working Paper (No. w18324).	Lottery	3,377	4 schools	Unknown	Urban	94.5	2.5
Curto, V.E. and Fryer Jr, R.G., 2014. "The potential of urban boarding schools for the poor: Evidence from SEED". Journal of Labor Economics, 32(1), pp.65-93.	Lottery	303	1 school	DC	Urban	67.8	16.3
Center for Research on Education Outcomes, 2009. "Multiple choice: Charter school performance in 16 states". Technical Report.	Observational	1,733,758	2403 schools	16 states	Urban, Suburban, Rural	20-90	1.5-97
Center for Research on Education Outcomes, 2013. "National Charter School Study". Technical Report.	Observational	5,000,000	5068 schools	27 states	Urban, Suburban, Rural	54	35

Center for Research on Education Outcomes, 2015. "Urban Charter School Study". Technical Report.	Observational	1,018,510	-	22 states	Urban	11-93	-

Note: This table summarizes charter school studies along the six attributes used in the discrete choice experiment. Dashed line indicates cells that have missing information (i.e., source could not be confirmed). Some cells under % FRPL and % white report a range of values because the paper provided breakdowns by schools/states.