Supplementary Information for
Cory McCartan, Jacob Brown, and Kosuke Imai.
"Measuring and Modeling Neighborhoods."

Summary of the survey samples

Each of our surveys were approved by the Harvard Institutional Review Board (IRB-20-0938). Neighborhood survey respondents were recruited via e-mail using a list of email addresses attached to registered voter records. We randomly sampled a total of 1,514,612 potential respondents from this list within each city (533,333 in Miami, 476,605 in New York City, and 504,674 in Phoenix). We sent an email invitation to the sampled registered voters on non-holiday weekdays between December 21, 2020 and February 19, 2021. The invitation informed the potential respondent of the purpose of the survey, that they would be asked to draw their neighborhood on a map, and provided information on the researcher's affiliations and contact information. Of these e-mails, 38.5% failed to be delivered to the potential respondent's inbox, due to the email address either being invalid or the receiving email server rejecting the email. In total, 930,839 voters received survey invitations (329,624 in Miami, 275,449 in New York, and 325,766 in Phoenix). The Phoenix sample exhibited a higher response rate than the Miami and New York City sample, with a 0.5% response rate in Miami and New York and a 1.3% response rate in Phoenix. Although our survey tool is publicly available, we are not able to publicly release our survey data to protect the privacy of respondents.

Here we present summary statistics describing the demographics of the first survey sample (Table S1), the overall demographics of the cities in our sample (Table S2) and the demographics of the city council survey sample (Table S3).

Table S1 shows the overall summary statistics and those broken out by city for our first survey. The demographic comparison across cities also may inform the differences in response rates across cities. The higher response rate in Phoenix is likely due to higher quality of email lists in this city than in New York or Miami and due to an older sampling population being more likely to respond to the surveys.

Comparison of Table S2 to Tables S1 and S3 gives a sense of the representativeness of the sample relative to the adult population of the three metropolitan areas of the study. We find that our sample is more predominantly white, wealthier, educated, and more likely to be a homeowner than the population of each of the cities in our sample.

Table S1: Survey Sample Summary Statistics – Full Sample

	Miami (n = 473)		NYC (n = 450)		Phoenix (n = 1,585)		Pooled (n = 2,508)	
	Avg.	St. Dev.	Avg.	St. Dev.	Avg.	St. Dev.	Avg.	St. Dev.
Democrat	0.48	0.50	0.57	0.49	0.38	0.49	0.44	0.50
Republican	0.40	0.49	0.33	0.47	0.52	0.50	0.46	0.50
Vote Biden 2020	0.57	0.50	0.65	0.48	0.49	0.50	0.53	0.50
Age	61.78	14.16	60.13	13.37	64.11	12.98	62.89	13.41
Female	0.44	0.50	0.45	0.50	0.48	0.50	0.46	0.50
White	0.65	0.48	0.72	0.45	0.86	0.34	0.79	0.41
Income (1,000s)	104.91	49.35	117.30	48.35	110.88	48.03	110.91	48.59
College	0.67	0.47	0.68	0.47	0.62	0.48	0.64	0.48
Years Residence	17.47	53.24	23.70	58.59	19.11	58.29	19.57	57.30
Homeowner	0.86	0.35	0.74	0.44	0.90	0.30	0.86	0.34
Married	0.60	0.49	0.60	0.49	0.67	0.47	0.64	0.48
Children in Home	0.33	0.47	0.39	0.49	0.27	0.45	0.31	0.46

Table S2: Metropolitan Region Population Demographics. Proportions are out of adult census population in each metropolitan region, except for Democrat and Republican, which are out of total registered voters.

City	Miami	NYC	Phoenix
College	0.413	0.455	0.389
Homeowner	0.608	0.449	0.641
Median Income	$66,944	$85,267	$71,421
Registered	0.650	0.614	0.656
White	0.348	0.420	0.587
Black	0.197	0.181	0.046
Hispanic	0.416	0.263	0.295
Democrat	0.450	0.587	0.313
Republican	0.247	0.148	0.343

Table S3: Survey Sample Summary Statistics - City Council Survey

	Email		Meta		Pooled	
	Avg.	St. Dev.	Avg.	St. Dev.	Avg.	St. Dev.
Democrat	0.71	0.45	0.78	0.41	0.75	0.44
Republican	0.20	0.40	0.11	0.32	0.16	0.36
Eric Adams	0.34	0.47	0.17	0.37	0.26	0.44
Age	56.95	13.72	41.19	18.59	49.63	17.97
Female	0.49	0.50	0.31	0.46	0.41	0.49
White	0.64	0.48	0.77	0.42	0.70	0.46
Income	114,008	49,553	103,972	52,130	109,001	51,076
College	0.48	0.50	0.41	0.49	0.44	0.50
Years Residence	16.10	8.71	9.00	9.42	12.82	9.71
Homeowner	0.50	0.50	0.27	0.44	0.39	0.49
Married	0.50	0.50	0.32	0.47	0.42	0.49
Children in Home	0.37	0.48	0.19	0.39	0.28	0.45

For the city council survey, a total of 490,000 registered voters were sampled from voter lists and sent an email invitation on non-holiday weekdays between December 22, 2022 and January 25, 2023. Similar to the previous survey, the bounce rate for the email invitation was 43.3%.

Additional descriptive statistics of drawn neighborhoods

First survey

Table S4: Treatment Effect on Usable Neighborhoods

	Dependent variable:		
	Usable Neighborhood		
	Miami	New York City	Phoenix
Party Condition	0.034	0.081	0.035
	(0.052)	(0.057)	(0.036)
Party Placebo Condition	0.066	−0.037	−0.004
	(0.055)	(0.058)	(0.037)
Race Condition	−0.004	0.093	0.050
	(0.056)	(0.058)	(0.037)
Race Placebo Condition	0.081	0.026	0.025
	(0.054)	(0.058)	(0.036)
Age	−0.005***	−0.001	−0.004***
	(0.002)	(0.002)	(0.001)
College	0.021	0.053	0.045*
	(0.043)	(0.050)	(0.027)
Democrat	0.119*	0.006	0.053
	(0.064)	(0.077)	(0.047)
Female	0.011	0.007	0.011
	(0.035)	(0.037)	(0.024)
Homeowner	0.017	−0.080*	0.030
	(0.051)	(0.048)	(0.038)
Income	0.0004	0.002***	0.0004
	(0.0004)	(0.0004)	(0.0003)
Married	0.047	−0.063	0.006
	(0.039)	(0.043)	(0.027)
Republican	0.142**	−0.010	0.021
Vote Biden	0.110*	0.078	0.098**
	(0.062)	(0.063)	(0.039)
Years Residence	0.002	0.003	0.002
	(0.002)	(0.002)	(0.001)
Constant	0.260**	0.157	0.401***
	(0.114)	(0.134)	(0.083)
Observations	1,468	1,193	4,028
R^2	0.007	0.005	0.0004
Adjusted R^2	0.005	0.001	−0.001
Residual Std. Error	0.467 (df = 1463)	0.485 (df = 1188)	0.490 (df = 4023)
F Statistic	2.735** (df = 4; 1463)	1.366 (df = 4; 1188)	0.430 (df = 4; 4023)
Note:			*p<0.05; **p<0.01

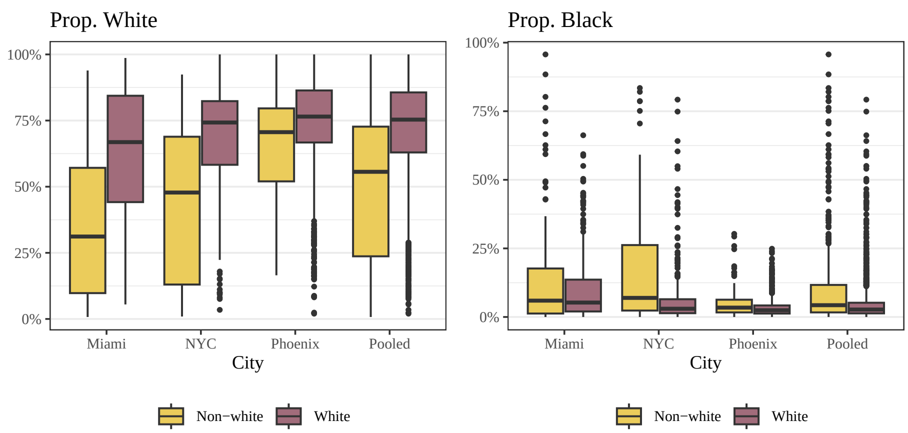

Figure S1: Racial demographics for respondent neighborhoods by respondent race (first survey)

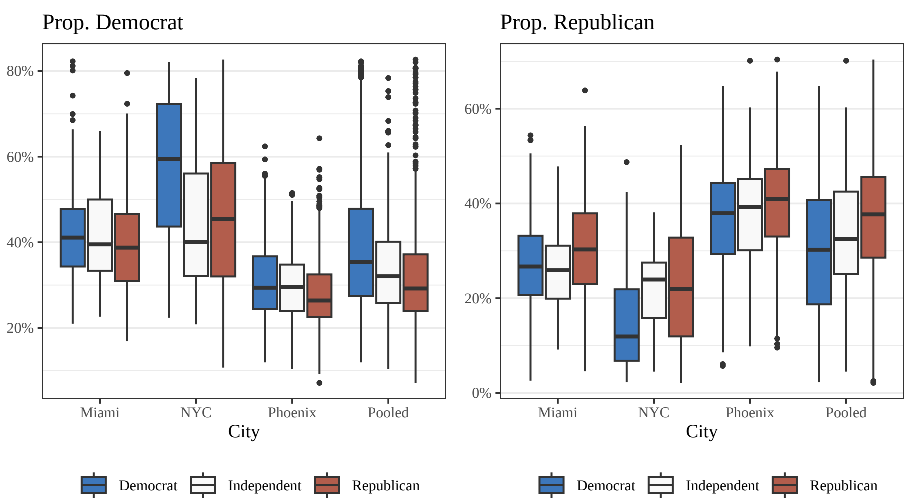

Figure S2: Party demographics for respondent neighborhoods by respondent party (first survey)

City council survey

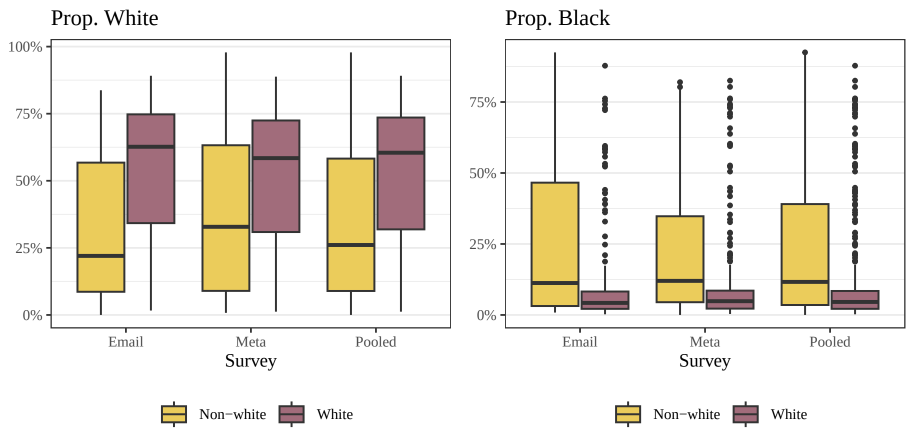

Figure S3: Racial demographics for respondent communities of interest by respondent race (city council survey)

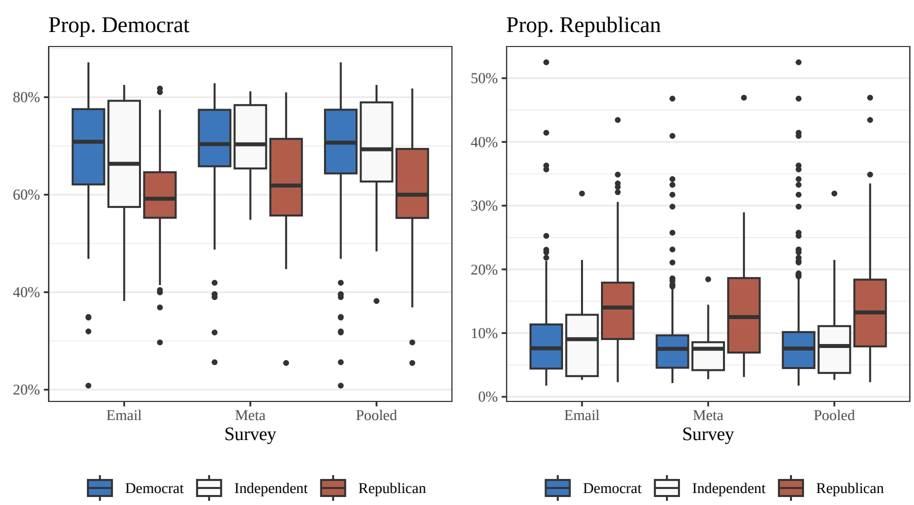

Figure S4: Party demographics for respondent communities of interest by respondent party (city council survey)

Computational details

As shown by Equation (1), the model can be expressed as a Bernoulli GLMM with complementary log-log link function. In this setup, every block that is included the neighborhood, as well as those blocks at the boundary of the neighborhood which could have been included ($C_i = 1$) but are not, becomes a separate GLMM observation. As a result, even moderate respondent sample sizes lead to many more block-level observations, although these block-level observations are of course dependent. In our data, for example, 309 respondents in the control group from Phoenix translate to 57,242 block-level observations.

The large number of block-level observations means that estimates will generally be precise for coefficients which vary at the block level and are shared across all respondents. At the same time, it can create computational efficiency problems for traditional Bayesian posterior sampling methods. Consequently, for our analysis we use a Normal approximation corrected by importance resampling, as implemented by the Stan modeling package Rstan. This approximation centers a Normal distribution at the posterior mode with covariance matrix the inverse of the curvature of the log posterior density at the mode. It subsequently samples 1,000 draws from this Gaussian approximation and performs importance resampling so that the draws better approximate the posterior.

Variable descriptions and model specifications

Table S5: Variable description and model specification.

Variable	Description	Interaction(s)	Baseline only?
Church	Block contains church		
Distance to church	Logarithm of distance to nearest church from block, in meters		
Park	Block contains park		
School	Block contains school	Children	
Distance to school	Logarithm of distance to nearest school from block, in meters	Children	
Children	Respondent has children at home	School, distance to school	*
Same block group	Block in same block group as respondent		
Same tract	Block in same tract as respondent		
Same road region	Block in same region bounded by major roads and railroads as respondent		
Population	Square root of the block population divided by 10,000		
Area	Square root of the area in square miles		
Fraction same race	Fraction of block of the same race as respondent (White, Black, Hispanic, Other)	Minority	*
Minority	Whether respondent is non-white, or Hispanic of any race	Fraction same race	*
Fraction same party	Fraction of block of the same party as respondent	Party	*
Party	Self-reported political party: Democratic, Republican, or independent	Fraction same party	*
Fraction same ownership	Fraction of block with same home ownership or rental status as respondent	Homeowner	*
Homeowner	Whether respondent owns their home	Fraction same ownership	*

Variable	Description	Interaction(s)	Baseline only?
Fraction same education	Fraction of block with same education as respondent	Education	*
Education	Respondent education; either “college” or “no college”	Fraction same education, Income	*
Income	Logarithm of median income of block group	Education	*
Age	Respondent age group: 0–40, 41–55, 56–65, 66-75, or 76+		*
Retired	Whether respondent is retired		*
Tenure	Square root of how long respondent has lived in current residence		*

Full and baseline model estimates

First survey

Tables S6 and S7 contain posterior summaries for all model coefficients on the original model scale. These models were fit using data from a random sample of 400 survey respondents from the control group, consisting of 78,771 individual block-level observations.

Table S6: Full model estimates.

Coefficient	City	Mean	Std. Dev.	Q5	Median	Q95
(Intercept)	Miami	-6.59	9.52	-22.30	-6.44	8.89
(Intercept)	NYC	-5.37	7.56	-17.36	-5.29	6.43
(Intercept)	Phoenix	-9.30	6.36	-19.82	-9.17	0.51
Church	Miami	-0.04	0.07	-0.16	-0.04	0.07
Church	NYC	0.11	0.05	0.03	0.11	0.19
Church	Phoenix	0.14	0.04	0.07	0.14	0.21
Distance	Miami	0.00	0.02	-0.03	0.00	0.03
Distance	Miami	0.13	0.03	0.08	0.13	0.18
Distance	NYC	0.12	0.02	0.08	0.12	0.15
Distance	NYC	0.01	0.03	-0.03	0.01	0.06
Distance	Phoenix	0.07	0.01	0.06	0.07	0.09
Distance	Phoenix	0.16	0.02	0.14	0.16	0.19
Park	Miami	-0.06	0.09	-0.21	-0.05	0.08
Park	NYC	0.15	0.05	0.08	0.15	0.23
Park	Phoenix	0.07	0.04	0.01	0.07	0.13
School	Miami	0.39	0.14	0.16	0.39	0.62
School	NYC	-0.15	0.12	-0.34	-0.15	0.06
School	Phoenix	0.15	0.07	0.04	0.16	0.27
Children	Miami	0.53	3.10	-4.48	0.57	5.39
Children	NYC	-0.77	2.79	-5.15	-0.78	3.73
Children	Phoenix	-1.09	2.05	-4.35	-1.14	2.39
Same block group	Miami	0.08	0.07	-0.04	0.08	0.19
Same block group	NYC	0.21	0.09	0.07	0.21	0.35
Same block group	Phoenix	0.08	0.03	0.03	0.08	0.13

Same tract	Miami	-0.17	0.09	-0.33	-0.17	-0.04
Same tract	NYC	0.06	0.12	-0.14	0.06	0.26
Same tract	Phoenix	-0.09	0.10	-0.27	-0.09	0.07
Same road region	Miami	-0.49	0.05	-0.57	-0.49	-0.41
Same road region	NYC	0.13	0.04	0.07	0.13	0.20
Same road region	Phoenix	-0.09	0.03	-0.14	-0.09	-0.04
Population	Miami	-0.22	0.26	-0.66	-0.22	0.21
Population	NYC	-1.21	0.20	-1.53	-1.21	-0.90
Population	Phoenix	-1.18	0.16	-1.43	-1.19	-0.89
Area	Miami	-0.37	0.08	-0.51	-0.37	-0.24
Area	NYC	-0.14	0.11	-0.32	-0.14	0.03
Area	Phoenix	0.00	0.04	-0.06	0.00	0.06
Fraction same race	Miami	-0.30	0.06	-0.39	-0.30	-0.21
Fraction same race	NYC	-0.12	0.05	-0.21	-0.12	-0.04
Fraction same race	Phoenix	-0.28	0.03	-0.33	-0.29	-0.24
Fraction same party	Miami	-0.26	0.09	-0.41	-0.26	-0.12
Fraction same party	NYC	-0.23	0.08	-0.35	-0.23	-0.10
Fraction same party	Phoenix	-0.29	0.07	-0.41	-0.29	-0.18
Fraction same ownership	Miami	-0.24	0.15	-0.48	-0.24	0.02
Fraction same ownership	NYC	-0.27	0.13	-0.49	-0.27	-0.07
Fraction same ownership	Phoenix	-0.28	0.12	-0.47	-0.28	-0.10
Fraction same education	Miami	0.25	0.15	0.00	0.26	0.48
Fraction same education	NYC	-0.53	0.10	-0.70	-0.53	-0.35
Fraction same education	Phoenix	-0.94	0.08	-1.07	-0.95	-0.82
Income	Miami	-0.12	0.05	-0.21	-0.12	-0.03
Income	NYC	-0.13	0.05	-0.21	-0.13	-0.06
Income	Phoenix	0.14	0.03	0.09	0.14	0.19
Age	Miami	0.01	0.16	-0.25	0.00	0.28
Age	NYC	0.01	0.13	-0.20	0.01	0.23
Age	Phoenix	-0.01	0.11	-0.17	-0.02	0.17
Education = no college	Miami	-2.20	3.20	-7.44	-2.23	3.13
Education = no college	NYC	0.06	2.83	-4.59	0.09	4.72
Education = no college	Phoenix	-0.33	2.26	-4.02	-0.28	3.31
Retired	Miami	0.50	3.10	-4.57	0.51	5.29
Retired	NYC	0.44	3.03	-4.71	0.51	5.28
Retired	Phoenix	0.11	2.44	-4.01	0.20	4.11
Tenure	Miami	-0.10	0.96	-1.71	-0.09	1.52
Tenure	NYC	-0.12	0.89	-1.55	-0.14	1.39
Tenure	Phoenix	0.24	0.77	-1.01	0.21	1.53
Party = ind	Miami	-0.52	4.32	-7.78	-0.57	6.71
Party = ind	NYC	0.41	4.73	-7.33	0.30	8.66
Party = ind	Phoenix	0.08	3.53	-5.81	0.15	5.90
Party = rep	Miami	-0.07	2.76	-4.57	-0.02	4.36
Party = rep	NYC	-0.27	2.60	-4.45	-0.30	4.09
Party = rep	Phoenix	-0.11	1.81	-3.08	-0.14	2.88
Minority	Miami	-0.46	3.14	-5.62	-0.47	4.53
Minority	NYC	0.03	2.69	-4.47	0.04	4.59
Minority	Phoenix	-0.15	2.57	-4.28	-0.16	4.00
Homeowner	Miami	-0.11	4.93	-8.35	0.02	7.57

Homeowner	NYC	-0.43	2.91	-5.14	-0.38	4.46
Homeowner	Phoenix	-0.48	3.64	-6.24	-0.42	5.40
School * children	Miami	-0.35	0.22	-0.70	-0.34	0.01
School * children	NYC	0.39	0.17	0.12	0.39	0.66
School * children	Phoenix	0.24	0.11	0.07	0.24	0.42
Children * distance	Miami	-0.03	0.05	-0.11	-0.03	0.05
Children * distance	NYC	0.07	0.04	0.01	0.07	0.15
Children * distance	Phoenix	0.12	0.03	0.08	0.12	0.16
Same tract * same road region	Miami	0.02	0.10	-0.15	0.02	0.19
Same tract * same road region	NYC	-0.24	0.13	-0.45	-0.23	-0.02
Same tract * same road region	Phoenix	-0.37	0.10	-0.53	-0.37	-0.20
Fraction same race * minority	Miami	0.25	0.09	0.10	0.25	0.40
Fraction same race * minority	NYC	0.27	0.12	0.07	0.27	0.46
Fraction same race * minority	Phoenix	-0.31	0.22	-0.66	-0.31	0.04
Fraction same party * party = ind	Miami	-0.12	0.20	-0.43	-0.12	0.20
Fraction same party * party = ind	NYC	-0.04	0.32	-0.57	-0.05	0.49
Fraction same party * party = ind	Phoenix	0.22	0.14	-0.02	0.22	0.45
Fraction same party * party = rep	Miami	-0.31	0.14	-0.54	-0.31	-0.07
Fraction same party * party = rep	NYC	-0.04	0.13	-0.25	-0.04	0.17
Fraction same party * party = rep	Phoenix	0.10	0.08	-0.02	0.10	0.23
Fraction same ownership * homeowner	Miami	0.52	0.18	0.22	0.53	0.83
Fraction same ownership * homeowner	NYC	0.42	0.16	0.15	0.42	0.67
Fraction same ownership * homeowner	Phoenix	0.26	0.13	0.05	0.26	0.47
Fraction same education * educ = no college	Miami	0.44	0.25	0.05	0.44	0.85
Fraction same education * educ = no college	NYC	1.41	0.21	1.05	1.41	1.77
Fraction same education * educ = no college	Phoenix	0.57	0.21	0.24	0.57	0.91
Income * education = no college	Miami	0.18	0.09	0.04	0.18	0.33
Income * education = no college	NYC	-0.10	0.09	-0.24	-0.10	0.04
Income * education = no college	Phoenix	0.05	0.08	-0.07	0.05	0.18
Alpha	Miami	1.42	0.05	1.34	1.42	1.50
Alpha	NYC	1.46	0.05	1.38	1.46	1.55
Alpha	Phoenix	1.50	0.03	1.46	1.50	1.54

Table S7: Baseline model estimates.

Coefficient	City	Mean	Std. Dev.	Q5	Median	Q95
(Intercept)	Miami	-7.72	1.12	-9.62	-7.71	-5.84
(Intercept)	NYC	-8.05	1.12	-9.91	-8.07	-6.27
(Intercept)	Phoenix	-8.82	0.89	-10.30	-8.83	-7.30
Church	Miami	-0.04	0.07	-0.15	-0.04	0.07
Church	NYC	0.13	0.05	0.05	0.13	0.21
Church	Phoenix	0.18	0.04	0.11	0.18	0.24
Distance	Miami	0.00	0.02	-0.04	0.00	0.03
Distance	Miami	0.11	0.02	0.07	0.11	0.15
Distance	NYC	0.12	0.02	0.09	0.12	0.16
Distance	NYC	0.05	0.02	0.01	0.05	0.08
Distance	Phoenix	0.09	0.01	0.07	0.09	0.11

Distance	Phoenix	0.20	0.01	0.18	0.20	0.23
Park	Miami	-0.02	0.09	-0.17	-0.02	0.12
Park	NYC	0.18	0.05	0.10	0.18	0.26
Park	Phoenix	0.09	0.04	0.03	0.08	0.15
School	Miami	0.26	0.11	0.08	0.26	0.43
School	NYC	0.05	0.09	-0.11	0.05	0.20
School	Phoenix	0.29	0.06	0.20	0.29	0.39
Same block group	Miami	0.07	0.07	-0.04	0.07	0.19
Same block group	NYC	0.17	0.09	0.03	0.17	0.32
Same block group	Phoenix	0.09	0.03	0.04	0.09	0.14
Same tract	Miami	-0.16	0.08	-0.29	-0.16	-0.01
Same tract	NYC	0.14	0.12	-0.06	0.14	0.34
Same tract	Phoenix	-0.06	0.10	-0.21	-0.06	0.11
Same road region	Miami	-0.51	0.05	-0.58	-0.51	-0.43
Same road region	NYC	0.16	0.04	0.10	0.16	0.22
Same road region	Phoenix	-0.13	0.03	-0.17	-0.13	-0.08
Population	Miami	-1.23	0.23	-1.60	-1.23	-0.84
Population	NYC	-1.74	0.18	-2.03	-1.74	-1.46
Population	Phoenix	-2.44	0.15	-2.68	-2.45	-2.17
Area	Miami	-0.30	0.08	-0.42	-0.30	-0.17
Area	NYC	-0.08	0.11	-0.26	-0.08	0.10
Area	Phoenix	0.02	0.04	-0.04	0.02	0.08
Same tract * same road region	Miami	0.00	0.09	-0.15	0.01	0.15
Same tract * same road region	NYC	-0.32	0.13	-0.54	-0.32	-0.11
Same tract * same road region	Phoenix	-0.43	0.10	-0.60	-0.43	-0.28
Alpha	Miami	1.42	0.05	1.34	1.42	1.49
Alpha	NYC	1.38	0.05	1.31	1.38	1.47
Alpha	Phoenix	1.51	0.03	1.46	1.51	1.55

City council survey

Tables S8 and S9 contain posterior summaries for all model coefficients on the original model scale. These models were fit using a training sample of 500 survey respondents (out of 627), consisting of 94,349 individual block-level observations.

Table S8: Full model estimates.

Coefficient	Mean	Std. Dev.	Q5	Median	Q95
(Intercept)	-6.03	3.24	-11.11	-5.97	-0.99
Church	0.10	0.02	0.07	0.10	0.12
Distance	0.06	0.01	0.04	0.06	0.07
Distance	0.05	0.01	0.04	0.05	0.07
Park	0.19	0.02	0.17	0.19	0.22
School	0.06	0.04	0.00	0.06	0.12
Children	0.03	1.71	-2.89	0.02	2.78
Same block group	0.25	0.06	0.14	0.25	0.35
Same tract	0.19	0.09	0.04	0.19	0.33
Same road region	-0.22	0.02	-0.25	-0.22	-0.19

Population	-1.18	0.07	-1.30	-1.18	-1.07
Area	-0.52	0.12	-0.71	-0.52	-0.32
Fraction same race	-0.32	0.03	-0.37	-0.32	-0.28
Fraction same party	-0.22	0.03	-0.26	-0.22	-0.18
Fraction same ownership	-0.11	0.04	-0.18	-0.11	-0.03
Fraction same education	-0.27	0.04	-0.34	-0.27	-0.20
Income	-0.11	0.02	-0.14	-0.11	-0.08
Age	0.01	0.06	-0.09	0.00	0.10
Education = no college	-3.58	1.96	-6.75	-3.60	-0.30
Retired	-0.14	2.25	-3.66	-0.20	3.63
Tenure	0.02	0.54	-0.85	0.03	0.87
Party = ind	-0.23	2.75	-4.73	-0.21	4.26
Party = rep	0.15	2.05	-3.30	0.14	3.50
Minority	0.10	1.66	-2.61	0.08	2.79
Homeownerother (please specify)	0.08	3.35	-5.40	0.13	5.49
Homeownerrenter	0.17	1.64	-2.56	0.14	2.76
School * children	-0.20	0.07	-0.32	-0.20	-0.08
Children * distance	-0.02	0.02	-0.06	-0.02	0.01
Same tract * same road region	-0.40	0.09	-0.55	-0.40	-0.25
Fraction same race * minority	0.16	0.06	0.06	0.15	0.25
Fraction same party * party = ind	0.12	0.13	-0.09	0.12	0.33
Fraction same party * party = rep	-0.53	0.15	-0.77	-0.53	-0.27
Fraction same ownership * homeownerother (please specify)	0.00	0.13	-0.22	0.01	0.22
Fraction same ownership * homeownerrenter	0.14	0.06	0.03	0.14	0.24
Fraction same education * educ = no college	0.79	0.11	0.62	0.79	0.97
Income * education = no college	0.27	0.04	0.21	0.27	0.34
Alpha	1.39	0.02	1.35	1.39	1.42

Table S9: Baseline model estimates.

Coefficient	Mean	Std. Dev.	Q5	Median	Q95
(Intercept)	-7.13	1.48	-9.57	-7.11	-4.64
Church	0.09	0.02	0.07	0.09	0.12
Distance	0.07	0.01	0.05	0.07	0.08
Distance	0.04	0.01	0.03	0.04	0.06
Park	0.20	0.02	0.17	0.20	0.22
School	0.00	0.03	-0.05	-0.01	0.05
Same block group	0.22	0.06	0.12	0.22	0.32
Same tract	0.23	0.09	0.09	0.23	0.38
Same road region	-0.27	0.02	-0.29	-0.27	-0.23
Population	-1.79	0.07	-1.90	-1.79	-1.68
Area	-0.24	0.10	-0.41	-0.24	-0.07
Same tract * same road region	-0.45	0.09	-0.60	-0.45	-0.30
Alpha	1.39	0.02	1.36	1.39	1.43

Aggregate-level prediction

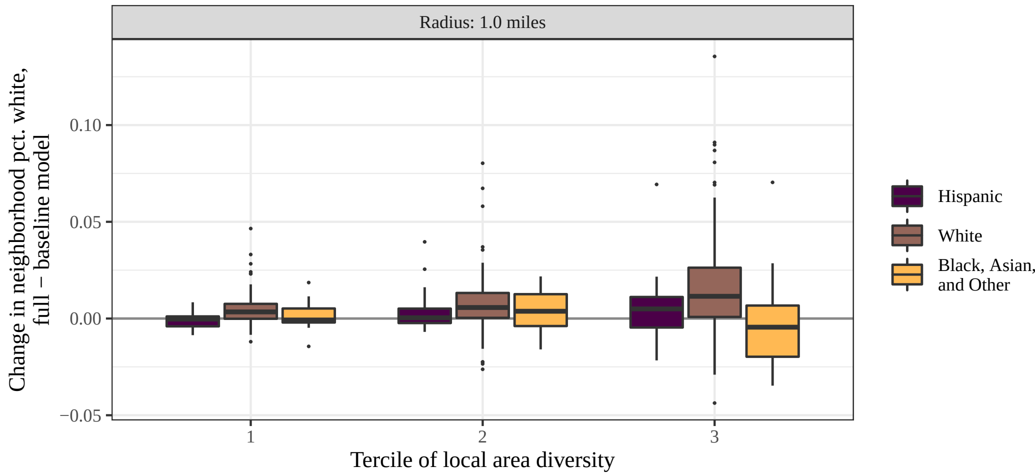

Figure S5: For each respondent, we compare the posterior mean of the fraction white in their predicted neighborhoods between the full and baseline models. Positive values indicate that using demographic information (full model) leads to a predicted neighborhood that is more white, on average. These differences in neighborhood fraction white are broken out by tercile of local area diversity and by the race of the respondent. Section S5 of the SI contains the full results tables for the full and baseline models.

Here we demonstrate how, in aggregate, the modeled relationship between co-racial and co-partisan demographics and census block inclusion produces predicted neighborhoods with different racial and partisan makeups across voters of different races and parties. Figure S5 presents boxplots showing the median and interquartile ranges of the change in proportion White in predicted neighborhoods, comparing the baseline model to the full model. The full model considers demographic information, while the baseline model does not, so the difference between the two is evidence of how much more homogeneous subjective neighborhoods become when demographics are considered.

This comparison is plotted separately by tercile of local racial diversity, to illustrate that, when voters live in areas where it is plausible to include or exclude out-group neighbors, they tend to do so. We measure local racial diversity as the standard deviation in the White percentage of each block within a fixed radius of a voter's residence. Since respondents must include or exclude whole blocks, measuring diversity according to block-level statistics is appropriate. Higher standard deviations indicate more block-level variation in racial composition and thus more opportunity to exclude out-group neighbors. Figure S6 shows the variation in local diversity across respondents.

Three quarters of White respondents' predicted neighborhoods contain greater proportions of White residents under the full model compared to the baseline model, with the disparity increasing as neighborhood diversity increases. We further see evidence, in the most mixed neighborhoods, of predicted neighborhoods for Black and Asian respondents with lower numbers of White residents, further evidence of the impact of demographics on subjective racial neighborhoods. Figure S5 shows the results for a one-mile radius, but the results are robust to different specifications.

In Figure S7, we plot the same comparison for partisan demographics, showing the boxplots of the difference in proportion Democratic between the full and baseline models across terciles of partisan diversity. Boxplots are shown separately by the self-reported partisan identification of the voter, ranging from strong Republican to strong Democrat.

On average, predicted neighborhoods for Democrats are slightly more Democratic in the full model compared to the baseline model, although the interquartile ranges overlap with zero across diversity terciles. Predicted neighborhoods for Republicans, on the other hand, are noticeably less Democratic in the full model compared to the baseline model, and the largest disparity is seen for strong Republicans, where the gap reaches 0.37

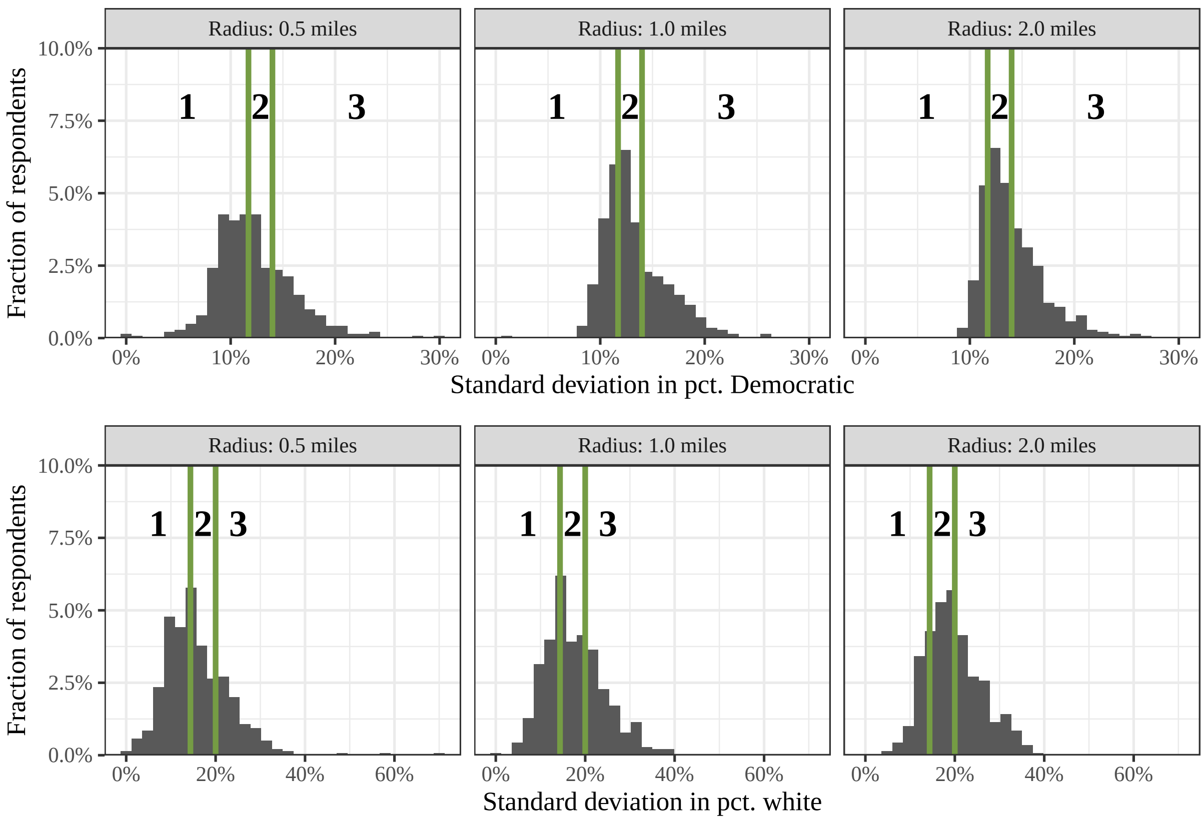

Figure S6: Local partisan and racial diversity for each respondent, as measured by the block-level standard deviation of the respective variables for blocks within a 0.5, 1.0, and 2.0-mile radius. Terciles are indicated by vertical lines and bold labels.

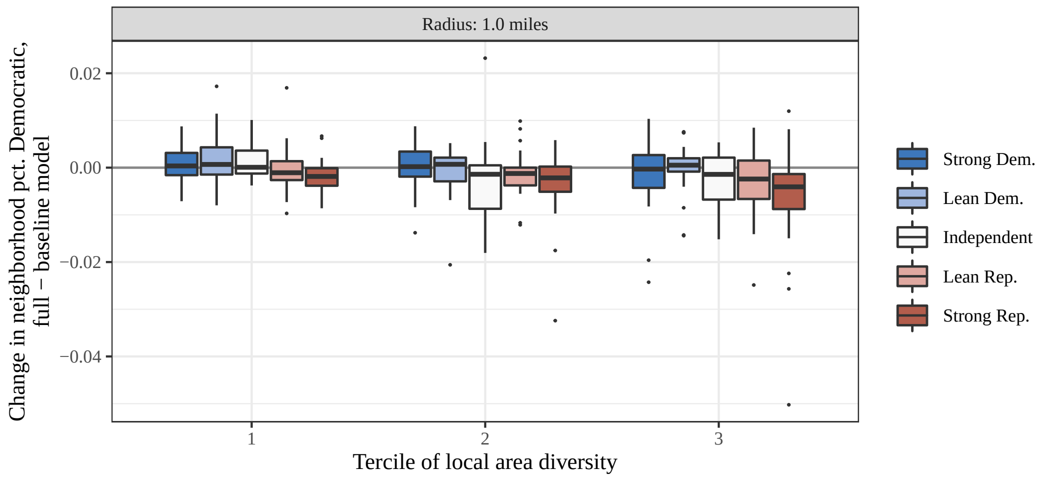

Figure S7: For each respondent, we compare the posterior mean of the fraction of Democrats in their predicted neighborhoods between the full and baseline models. Positive values indicate that using demographic information (full model) leads to a predicted neighborhood that is more Democratic, on average. These differences in neighborhood fraction Democratic are broken out by tercile of local area diversity and by the party identification of the respondent. Section S5 of the SI contains the full results tables for the full and baseline models.

percentage points in the most politically diverse areas. Similar to the racial comparison, the degree of difference is increasing with partisan diversity, and thus the potential to draw neighborhoods more differentiated by partisanship.

For both race and partisanship, the changes in neighborhood composition as a result of factoring in demographics are small in magnitude. This reflects, we believe, the overwhelming influence of residential segregation and sorting. Voters' preferences for homogeneous neighborhoods are already reflected in their choice of residence; all that we measure here is the marginal predictive effect of this preference on their subjective definition of neighborhood, given that residence.

Predictive performance compared to ZCTAs

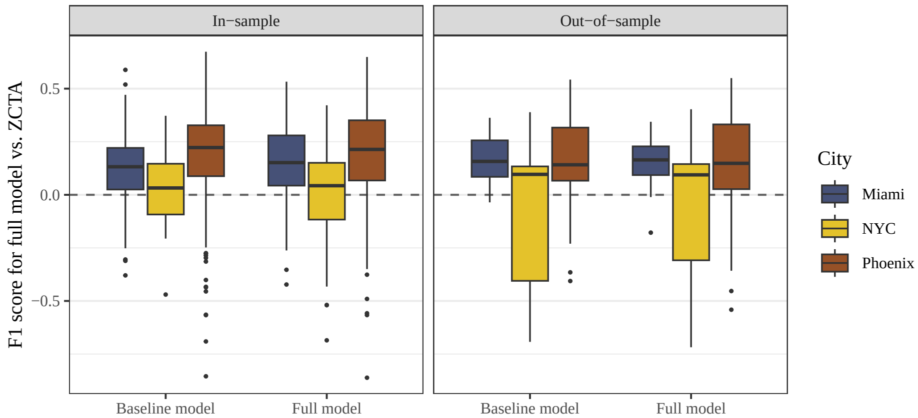

Figure S8: Posterior median of the difference in F1 scores between a neighborhood predicted by the model prediction and a census ZCTA. The boxplot shows the variation in this median difference accros the respondents included in the model fitting (left plot) and excluded from the model fitting (right plot). Section S5 of the SI contains the full results tables for the full and baseline models.

Additional model fits

Model fit with urban-suburban indicator

We fit the full model specification again, but include an indicator for whether a block belongs to the primary city in each metro area. This indicator is also interacted with the same-tract indicator, and the fraction-same-race variable.

These additional model coefficients are summarized in Figure S9. Results are mixed for the direct effect (which affects the size of the neighborhood) and same-tract interaction. In all 3 cities, the same-race preference is slightly stronger in the urban area compared to suburban areas. The main same-race effect (corresponding to the coefficient for suburban voters) remains positive for both white and minority voters. Thus while there is some evidence of urban-suburban heterogeneity, directionally the results are consistent with the overall findings.

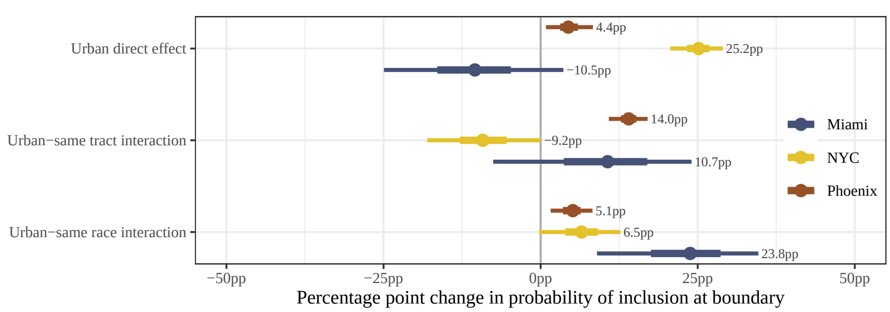

Figure S9: Full model estimates for urban indicator variable and interactions. The full coefficient estimates for this model are available in CSV format as part of the APSR Dataverse files.

Model fit with home prices

We fit the full model specification again, but include the (log of the) median home value in each block group. The home value covariate is also interacted with the same-race and same-party variables, and with the respondent's educational group (the indicator for not having attended college).

These additional model coefficients are summarized in Figure S10. Both the direct effect and interaction terms are relatively small. In New York and Miami, there is some evidence that non-college respondents are more likely to include a region if it has higher home prices. Overall the main conclusions about party and race in-group preference remain unchanged.

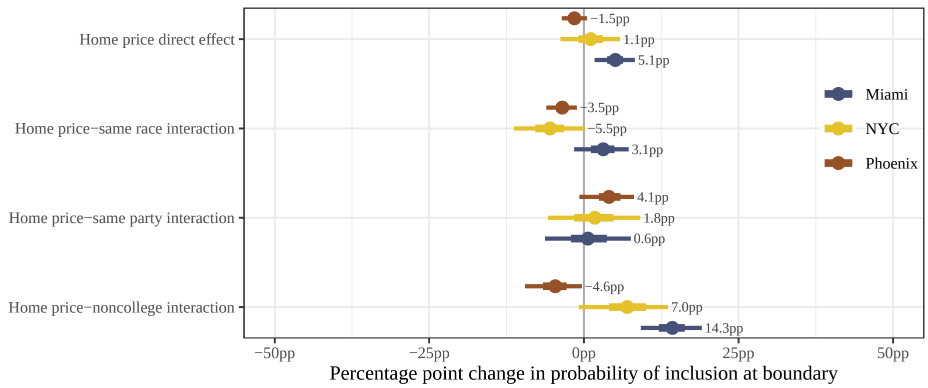

Figure S10: Full model estimates with home prices control. The full coefficient estimates for this model are available in CSV format as part of the APSR Dataverse files.

Model fit with community centers and other cultural institutions

We fit the full model specification again, this time only on the sample from New York City, where we acquired data on other types of cultural institutions besides churches and schools. These include: youth centers, senior centers, libraries, community centers, and other cultural institutions. We include an indicator for the presence of any of these buildings in a census block. We estimate the full model specification with this indicator as an additional covariate. Figure S11 reports the results, demonstrating the robustness of the original variables to the inclusion of this covariates. We further see that community centers and other cultural institutions exert a similar effect as churches on inclusion in subjective neighborhoods, with the present of any of these making it less likely a place is included.

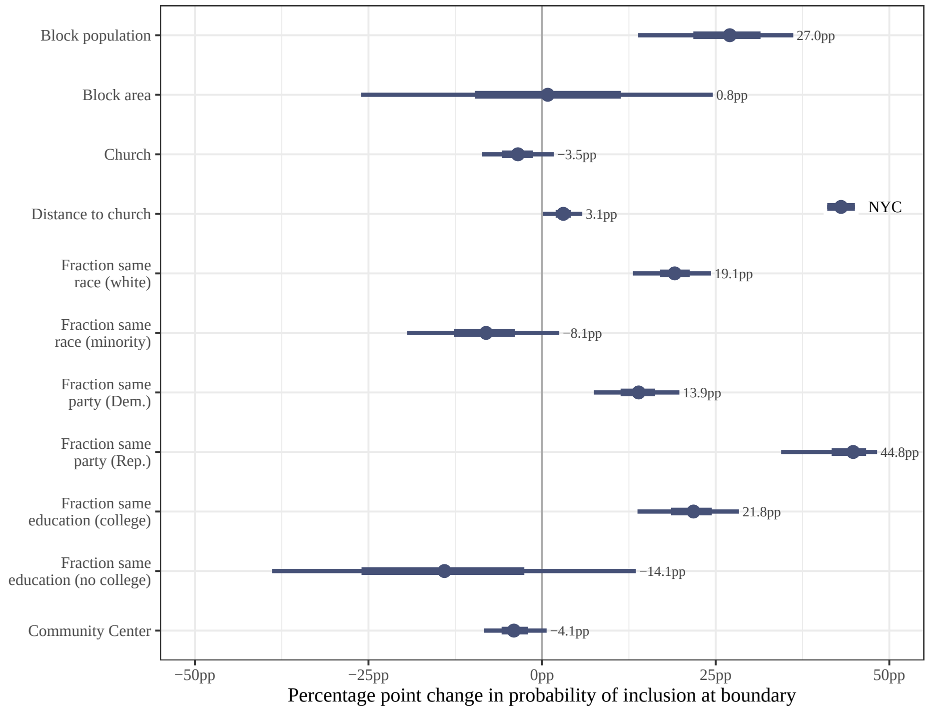

Figure S11: Full model estimates for community center indicator variable. The full coefficient estimates for this model are available in CSV format as part of the APSR Dataverse files.

Model fit with turnout and party interactions

We fit the full model specification again, but include interactions between party and fraction same race in each census block. We do this for the both surveys, the subjective neighborhoods survey and the city council survey. Figure S12 and S13 report the same race coefficients for Democrats and Republicans. We find that, among whites, that while respondents of both parties prefer census blocks with more same race residents, this preference is stronger for Republicans than for Democrats.

We also fit the full model to the city council survey including an interaction for whether or not the respondent reported that they voted in the 2021 mayoral election. We do not fit the equivalent model to the subjective neighborhood survey because self-reported 2020 turnout was too high (>95%) in the sample to estimate this interaction in every city. Figure S14 reports the results, showing that voters who voted in the 2021 mayoral election also gave greater preference to racial similarity when drawing their communities of interest. We see a similar disparity for voting and preferences for partisan homophily among Republicans but not for Democrats, where Democrats who do not vote are slightly more influenced by party demographics when defining their communities of interest.

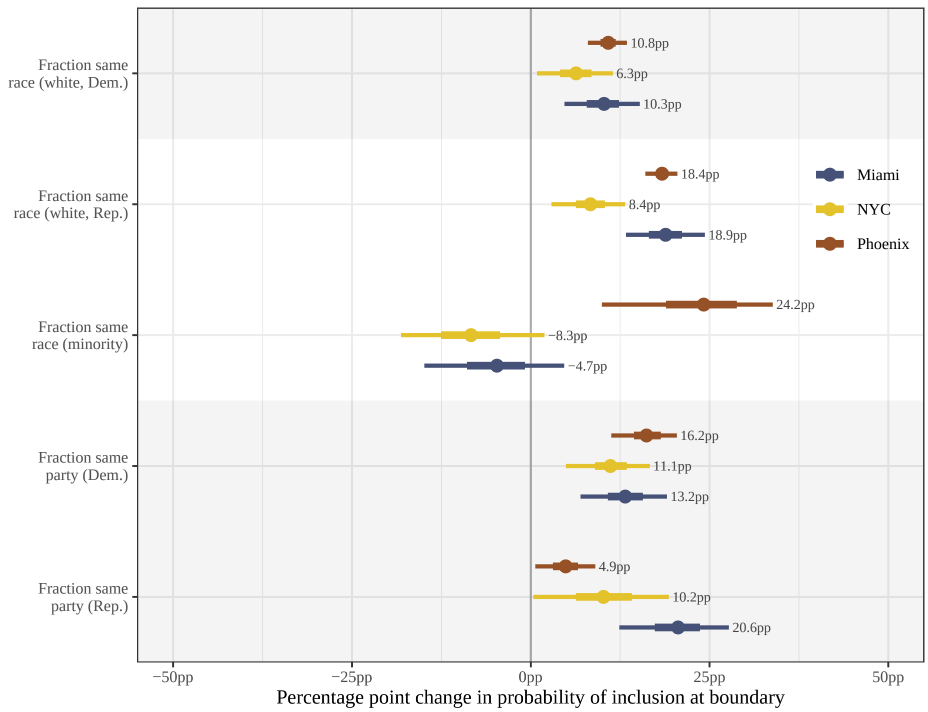

Figure S12: Subjective neighborhood survey: Selected full model coefficient posteriors, scaled to show the percentage point change in probability of a block's inclusion for a baseline probability of 50%. Plotted are 90% and 50% credible intervals, with posterior medians displayed to the right of each interval. The full coefficient estimates for this model are available in CSV format as part of the APSR Dataverse files.

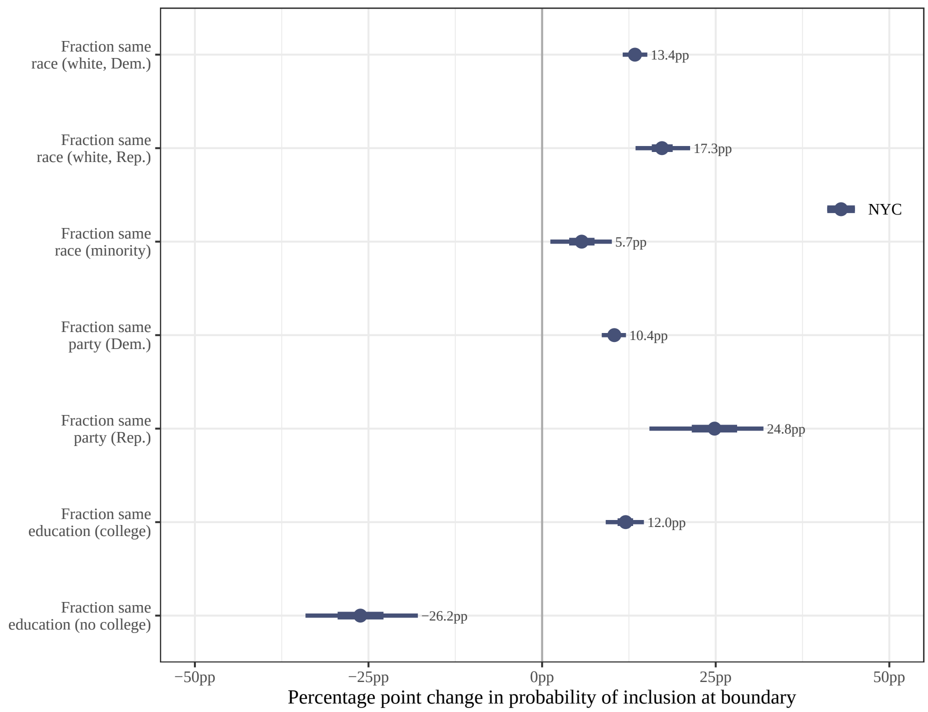

Figure S13: City council survey: Selected full model coefficient posteriors, scaled to show the percentage point change in probability of a block's inclusion for a baseline probability of 50%. Plotted are 90% and 50% credible intervals, with posterior medians displayed to the right of each interval. The full coefficient estimates for this model are available in CSV format as part of the APSR Dataverse files.

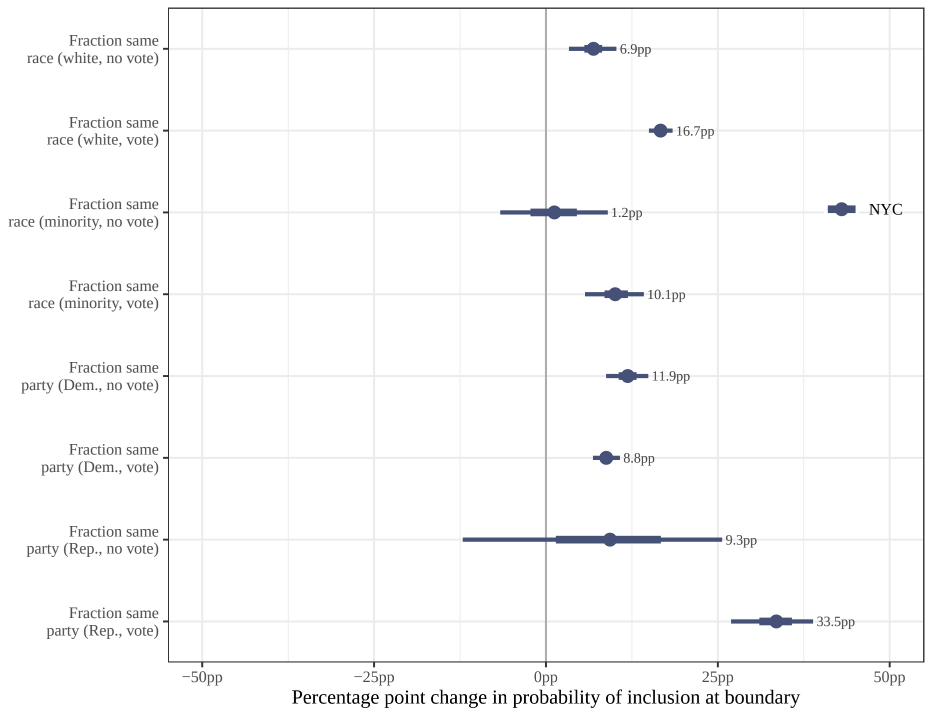

Figure S14: City council survey: Selected full model coefficient posteriors, scaled to show the percentage point change in probability of a block's inclusion for a baseline probability of 50%. Plotted are 90% and 50% credible intervals, with posterior medians displayed to the right of each interval. The full coefficient estimates for this model are available in CSV format as part of the APSR Dataverse files.

Model fit on Email versus Meta survey

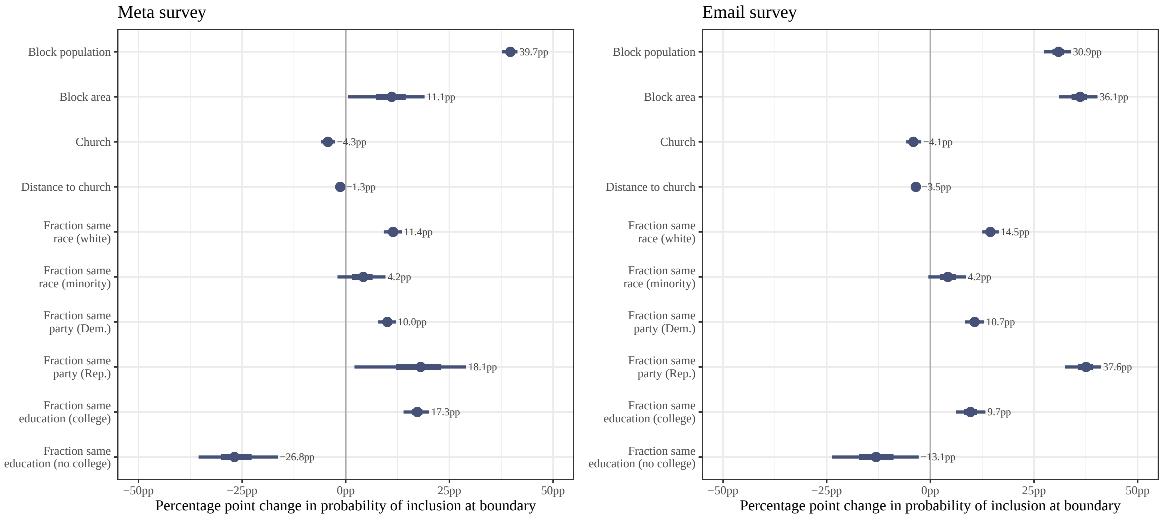

Figure S15: City council survey: Selected full model coefficient posteriors, scaled to show the percentage point change in probability of a block's inclusion for a baseline probability of 50%. Plotted are 90% and 50% credible intervals, with posterior medians displayed to the right of each interval. The full coefficient estimates for this model are available in CSV format as part of the APSR Dataverse files.

Survey representativeness

To assess the sensitivity of our main effects to any unrepresentativeness of the survey, we fit the full model specification, this time adding interactions between race and party homophily variables and respondent race, homeowner status, retirement status, and college education. Figure S16 summarizes these interaction coefficients. While the effects can vary by these variables, in almost every case we still observed an overall positive effect of partisan and racial homophily on inclusion (i.e. even negative interaction coefficients are not large enough to switch the sign of the overall effect). This is remarkable given the relatively small sample sizes for some demographics, and the high amount of individual heterogeneity observed for other aspects of the drawn neighborhoods.

To assess the sensitivity of our main effects to any unrepresentativeness of the survey, we fit the full model specification, this time adding interactions between race and party homophily variables and respondent race, homeowner status, retirement status, and college education. Figure S16 summarizes these interaction coefficients. While the effects can vary by these variables, in almost every case we still observed an overall positive effect of partisan and racial homophily on inclusion (i.e. even negative interaction coefficients are not large enough to switch the sign of the overall effect). This is remarkable given the relatively small sample sizes for some demographics, and the high amount of individual heterogeneity observed for other aspects of the drawn neighborhoods.

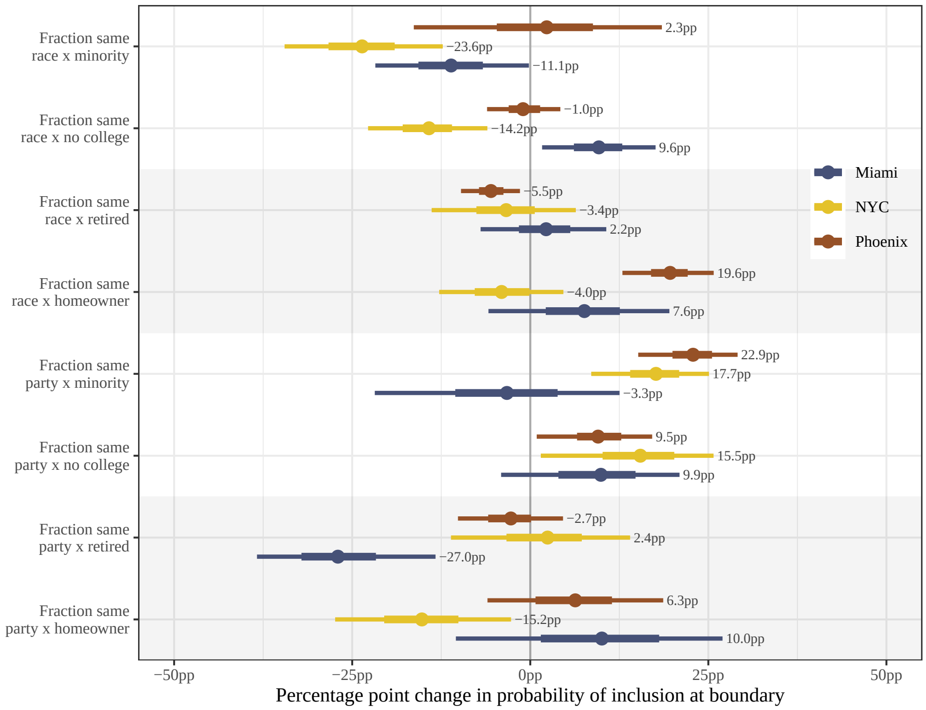

Figure S16: Full model estimates for interactions with race, homeownership, retirement, and college. The full coefficient estimates for this model are available in CSV format as part of the APSR Dataverse files.

Time spent drawing maps

First survey

Figure S17 shows that respondents which drew valid neighborhoods spent more time on average with the drawing tool, as would be expected. The results show that respondents who drew usable neighborhoods on average spent 4.17 minutes on the map (median 2.54 minutes). The subsequent table reports associational measures between various covariates and the time spent drawing the map, which is measured in seconds, binned into three categories, and reported in the table columns. Time spent on the map is not well correlated with individual characteristics or neighborhood features.

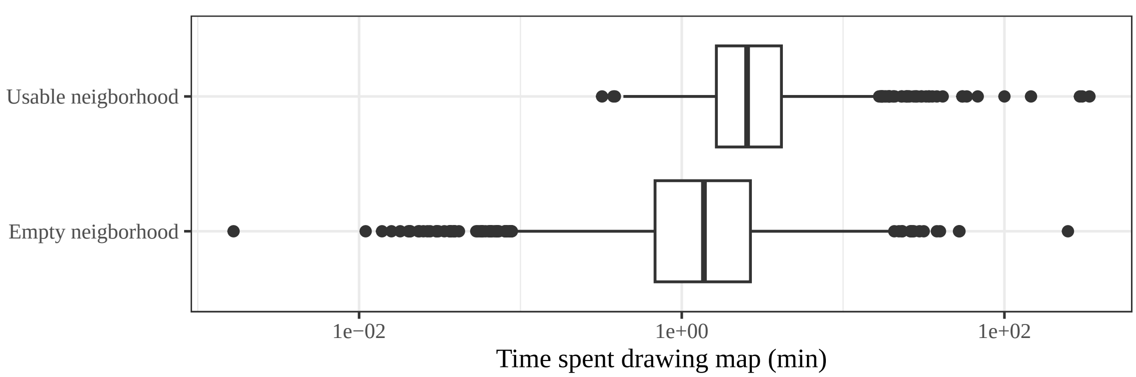

Figure S17: From the subjective neighborhoods survey. Distributions plotted separately for valid and empty neighborhoods.

Table S10: Descriptive statistics by total time spent drawing neighborhood (minutes)

	(0,1] (N=1328)	(1,2] (N=1556)	(2,Inf] (N=2732)	p value
usable				< 0.001[1]
FALSE	1133 (85.3%)	833 (53.5%)	1068 (39.1%)	
TRUE	195 (14.7%)	723 (46.5%)	1664 (60.9%)	
group				0.004[1]
C	293 (22.1%)	308 (19.8%)	518 (19.0%)	
P	237 (17.8%)	307 (19.7%)	611 (22.4%)	
PH	287 (21.6%)	277 (17.8%)	490 (17.9%)	
R	240 (18.1%)	316 (20.3%)	558 (20.4%)	
RH	271 (20.4%)	348 (22.4%)	555 (20.3%)	
city				0.052[1]
miami	322 (24.2%)	342 (22.0%)	549 (20.1%)	
new-york	241 (18.1%)	266 (17.1%)	466 (17.1%)	
phoenix	765 (57.6%)	948 (60.9%)	1717 (62.8%)	
party				< 0.001[1]
dem_strong	298 (22.4%)	423 (27.2%)	763 (27.9%)	
dem_lean	194 (14.6%)	308 (19.8%)	561 (20.5%)	
independent	153 (11.5%)	162 (10.4%)	209 (7.7%)	
rep_lean	288 (21.7%)	305 (19.6%)	547 (20.0%)	
rep_strong	395 (29.7%)	358 (23.0%)	652 (23.9%)	
age				< 0.001[2]
Mean (SD)	59.950 (13.861)	62.715 (13.057)	64.484 (13.028)	
Range	18.000 - 105.000	21.000 - 105.000	18.000 - 105.000	
gender				0.178[1]
N-Miss	28	16	14	
female	611 (47.0%)	674 (43.8%)	1255 (46.2%)	
male	689 (53.0%)	866 (56.2%)	1463 (53.8%)	
education				< 0.001[1]
N-Miss	34	23	34	
no_hs	11 (0.9%)	1 (0.1%)	2 (0.1%)	
some_hs	5 (0.4%)	16 (1.0%)	16 (0.6%)	
hs	85 (6.6%)	109 (7.1%)	183 (6.8%)	
some_coll	258 (19.9%)	240 (15.7%)	478 (17.7%)	
grad_2yr	130 (10.0%)	164 (10.7%)	269 (10.0%)	
grad_4yr	427 (33.0%)	512 (33.4%)	892 (33.1%)	
postgrad	378 (29.2%)	491 (32.0%)	858 (31.8%)	
retired				< 0.001[1]
No	834 (62.8%)	824 (53.0%)	1320 (48.3%)	
Yes	494 (37.2%)	732 (47.0%)	1412 (51.7%)	
race				< 0.001[1]
N-Miss	83	42	64	
aapi	34 (2.7%)	34 (2.2%)	57 (2.1%)	
black	63 (5.1%)	49 (3.2%)	95 (3.6%)	
hisp	190 (15.3%)	177 (11.7%)	276 (10.3%)	
indig	12 (1.0%)	14 (0.9%)	9 (0.3%)	
multi	32 (2.6%)	26 (1.7%)	52 (1.9%)	
white	914 (73.4%)	1214 (80.2%)	2179 (81.7%)	
homeowner				0.052[1]
No	204 (15.4%)	214 (13.8%)	335 (12.3%)	
Yes	1124 (84.6%)	1342 (86.2%)	2397 (87.7%)	

City council survey

Figure S18 shows that respondents which drew valid communities of interest spent more time on average with the drawing tool. The following table reports correlations between various covariates and the time spent drawing the map, which is measured in seconds, binned into three categories, and reported in the table columns.

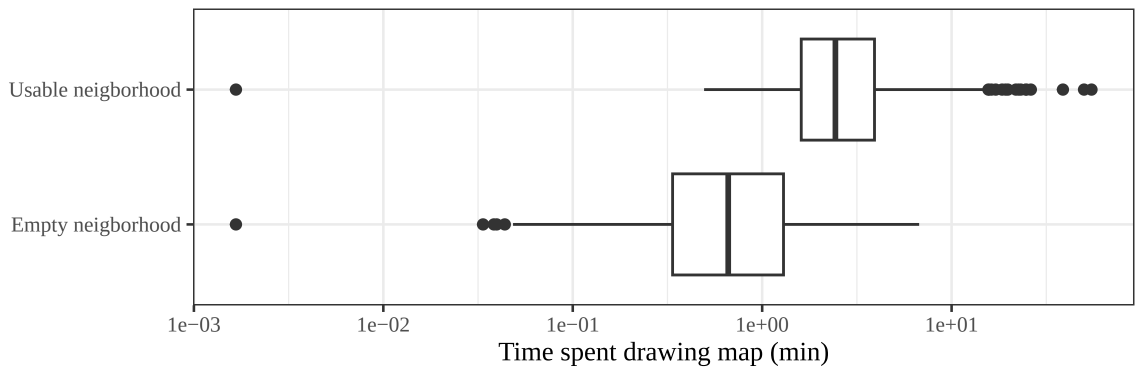

Figure S18: From city council survey. Distributions plotted separately for valid and empty neighborhoods.

Table S11: Descriptive statistics by total time spent drawing neighborhood (minutes)

	(0,1] (N=326)	(1,2] (N=393)	(2,Inf] (N=659)	p value
usable				< 0.001[1]
FALSE	270 (82.8%)	90 (22.9%)	50 (7.6%)	
TRUE	56 (17.2%)	303 (77.1%)	609 (92.4%)	
group				0.759[1]
N-Miss	84	50	37	
C	52 (21.5%)	65 (19.0%)	122 (19.6%)	
P	49 (20.2%)	73 (21.3%)	124 (19.9%)	
PH	48 (19.8%)	71 (20.7%)	117 (18.8%)	
R	39 (16.1%)	69 (20.1%)	137 (22.0%)	
RH	54 (22.3%)	65 (19.0%)	122 (19.6%)	
city				< 0.001[2]
new-york	326 (100.0%)	393 (100.0%)	659 (100.0%)	
party				< 0.001[1]
N-Miss	66	37	22	
dem_lean	164 (63.1%)	275 (77.2%)	495 (77.7%)	
independent	41 (15.8%)	30 (8.4%)	51 (8.0%)	
rep_lean	55 (21.2%)	51 (14.3%)	91 (14.3%)	
age				0.155[3]
N-Miss	77	44	29	
Mean (SD)	50.040 (17.841)	47.585 (17.655)	50.605 (18.152)	
Range	18.000 - 105.000	18.000 - 89.000	18.000 - 105.000	
gender				0.155[1]
N-Miss	84	56	47	
female	115 (47.5%)	127 (37.7%)	246 (40.2%)	
male	127 (52.5%)	210 (62.3%)	366 (59.8%)	
education				0.002[1]
N-Miss	85	47	35	
no_hs	2 (0.8%)	1 (0.3%)	0 (0.0%)	
some_hs	8 (3.3%)	4 (1.2%)	7 (1.1%)	

	(0,1] (N=326)	(1,2] (N=393)	(2,Inf] (N=659)	p value
hs	22 (9.1%)	26 (7.5%)	28 (4.5%)	
some_coll	36 (14.9%)	29 (8.4%)	80 (12.8%)	
grad_2yr	21 (8.7%)	17 (4.9%)	28 (4.5%)	
grad_4yr	70 (29.0%)	128 (37.0%)	216 (34.6%)	
postgrad	82 (34.0%)	141 (40.8%)	265 (42.5%)	
retired				0.119[1]
N-Miss	79	45	30	
No	184 (74.5%)	288 (82.8%)	479 (76.2%)	
Yes	63 (25.5%)	60 (17.2%)	150 (23.8%)	
race				< 0.001[1]
N-Miss	109	65	62	
aapi	13 (6.0%)	12 (3.7%)	20 (3.4%)	
black	37 (17.1%)	23 (7.0%)	41 (6.9%)	
hisp	45 (20.7%)	35 (10.7%)	69 (11.6%)	
indig	2 (0.9%)	2 (0.6%)	4 (0.7%)	
multi	5 (2.3%)	13 (4.0%)	23 (3.9%)	
white	115 (53.0%)	243 (74.1%)	440 (73.7%)	
homeowner				0.506[1]
N-Miss	84	48	34	
Homeowner	97 (40.1%)	133 (38.6%)	247 (39.5%)	
Other (please specify)	6 (2.5%)	14 (4.1%)	37 (5.9%)	
Renter	139 (57.4%)	198 (57.4%)	341 (54.6%)	

Support for new housing construction analysis

We also collect respondent attitudes on the construction of new housing in their neighborhoods and test how opposition to new housing intensifies is related preferences. We do so by creating an indicator variable for whether respondents support a ban on the construction of new housing in their neighborhood and interacting it with the fraction same race and fraction same party in each census block. We further interact the homeowner variable with these terms to see how this varies across homeowners and renters. We then report the influence of race and party demographics for respondents who do and do not support a housing ban, separately for homeowners and renters. Figure S19 reports these coefficients, showing that the influence of homophily by race or party does not substantially vary by whether the respondent supports new housing in their neighborhood.

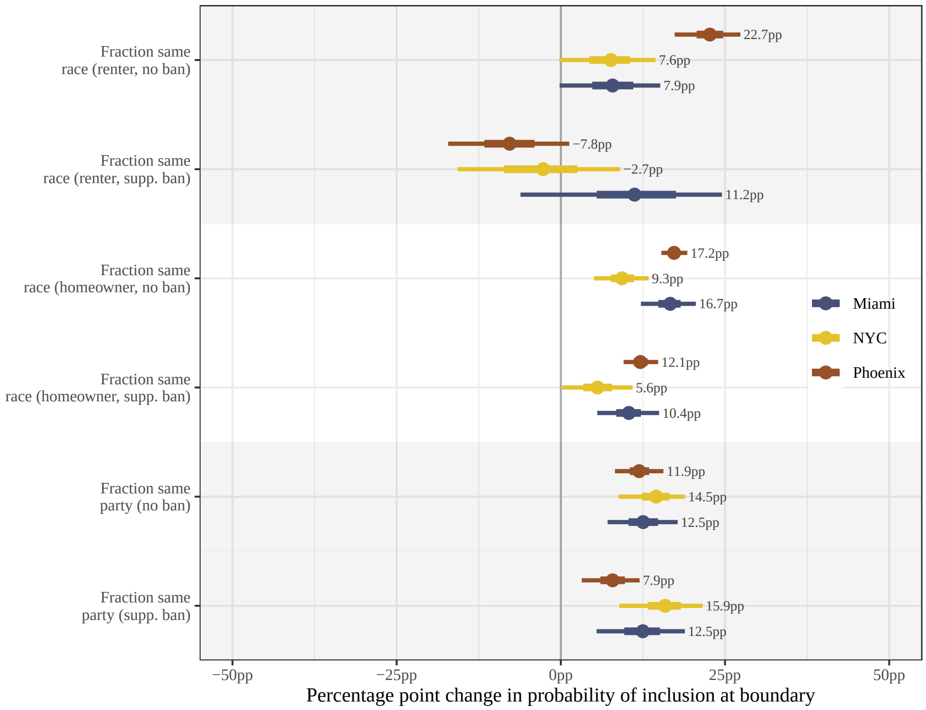

Figure S19: Selected full model coefficient posteriors, scaled to show the percentage point change in probability of a block's inclusion for a baseline probability of 50%. Plotted are 90% and 50% credible intervals, with posterior medians displayed to the right of each interval.

Neighborhood trust analysis

Next, we conduct a similar exercise for the measures of whether respondents express trust in their neighbors. Figure S20 reports the effects of fraction same race and party separately for respondents who are above the median level of expressed neighbor trust (median calculated from sample of respondents in each city). Again, we find that the effects on race and party demographics are generally consistent across these comparisons.

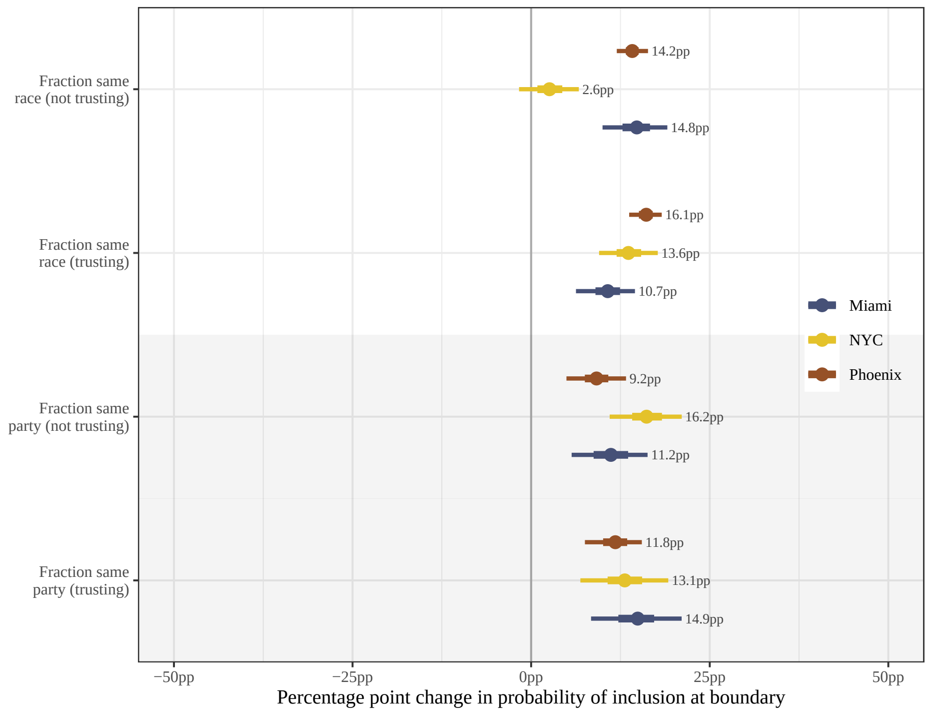

Figure S20: Selected full model coefficient posteriors, scaled to show the percentage point change in probability of a block's inclusion for a baseline probability of 50%. Plotted are 90% and 50% credible intervals, with posterior medians displayed to the right of each interval.