Online Appendix

"The Immigrant Next Door"

Leonardo Bursztyn, Thomas Chaney, Tarek A. Hassan and Aakaash Rao

I. Additional Tables and Figures

Appendix Table A1—Summary statistics

	Obs.	Mean	Std. dev.	Median	Min	Max
Panel A: County-country-quarter level						
A.1: Ancestry, Charity 1 and 2						
2010 population from country f (thousands)	4,703,862	0.236	9.453	0.000	0.000	2,629.376
2010 IHS-transformed population from country f	4,703,862	1.188	2.073	0.007	0.000	15.475
2010 share of population from country f ($\times 100$)	4,703,862	0.124	1.356	0.000	0.000	72.765
A.2: Donations, Charity 1 and 2						
IHS-transformed number of donations to country f	4,703,862	0.019	0.182	0.000	0.000	7.71
A.3: Donations, Charity 2 only						
IHS-transformed dollar value of donations to country f	3,972,708	0.08	0.65	0.00	0.00	11.84
Panel B: County-quarter level						
B.1: Donations to Arab countries						
IHS-transformed number of donations	150,336	0.048	0.297	0.000	0.000	6.397
Panel C: Individual level						
C.1: Project Implicit						
Arab-Muslim IAT score	108,235	0.018	0.989	0.003	-4.208	4.39
Warmth toward Arab-Muslims	108,109	0.035	0.996	-0.315	-2.567	1.938
C.2: CCES						
Support for the Muslim Ban	57,195	0.529	0.499	1.000	0.000	1
Voted for Trump in 2016	98,205	0.463	0.499	0.000	0.000	1
C.3: Nationscape						
Favorability toward Muslims	189,273	-0.073	1.002	0.313	-1.668	1.304
Support for the Muslim Ban	58,750	0.309	0.462	0.000	0.000	1
Voted for Trump in 2016	171,944	0.534	0.499	1.000	0.000	1

Notes: The table presents summary statistics for all datasets used in the main analyses except the custom survey (summary statistics for which are presented in Appendix Table A4). Donations statistics are calculated from the pooled donations across Charity 1 and Charity 2.

Appendix Table A2—Top ten foreign countries by size of ancestral population

	Ancestry (thousands)	# counties	Peak arrival time
Mexico	22,903.85	3,136	1990-2000
Philippines	2,729.48	3,136	1990-2000
India	2,433.13	3,108	2000-2010
Japan	1,144.04	3,105	1990-2000
Haiti	868.67	2,596	1990-2000
Peru	662.80	3,125	2000-2010
Ecuador	606.75	3,121	2000-2010
Iran	419.04	2,882	1980-1990
Lebanon	371.66	3,047	1980-1990
Pakistan	371.52	2,844	1990-2000

Notes: Table A2 lists the top ten countries in our sample by size of ancestral population. For each country, Column 1 displays the size of the ancestral population (in thousands); Column 2 displays the number of counties with nonzero ancestral population; and Column 3 displays the decade in which the maximum number of immigrants from that country arrived in the U.S.

Appendix Table A3—Effect of ancestral presence on donations, separated by charity

	IHS(# donations)	Donations (dummy)	IHS($ donations)
Panel A: Charity 1			
IHS(Ancestry)	0.042	0.019	—
	(0.014)	(0.006)	—
First-stage F-statistic	53.79	53.79	—
Weak IV-robust p-value	< 0.01	< 0.01	
Dep. var. mean	0.009	0.007	—
Dep. var. sd	0.128	0.082	—
Observations	2,193,462	2,193,462	—
Panel B: Charity 2			
IHS(Ancestry)	0.068	0.033	0.203
	(0.030)	(0.015)	(0.091)
First-stage F-statistic	309.9	309.9	309.9
Weak IV-robust p-value	< 0.01	< 0.01	< 0.01
Dep. var. mean	0.013	0.010	0.051
Dep. var. sd	0.145	0.101	0.528
Observations	9,410,862	9,410,862	9,410,862
Foreign country × quarter FE	Yes	Yes	Yes
Distance controls	Yes	Yes	Yes
US county × quarter FE	Yes	Yes	Yes

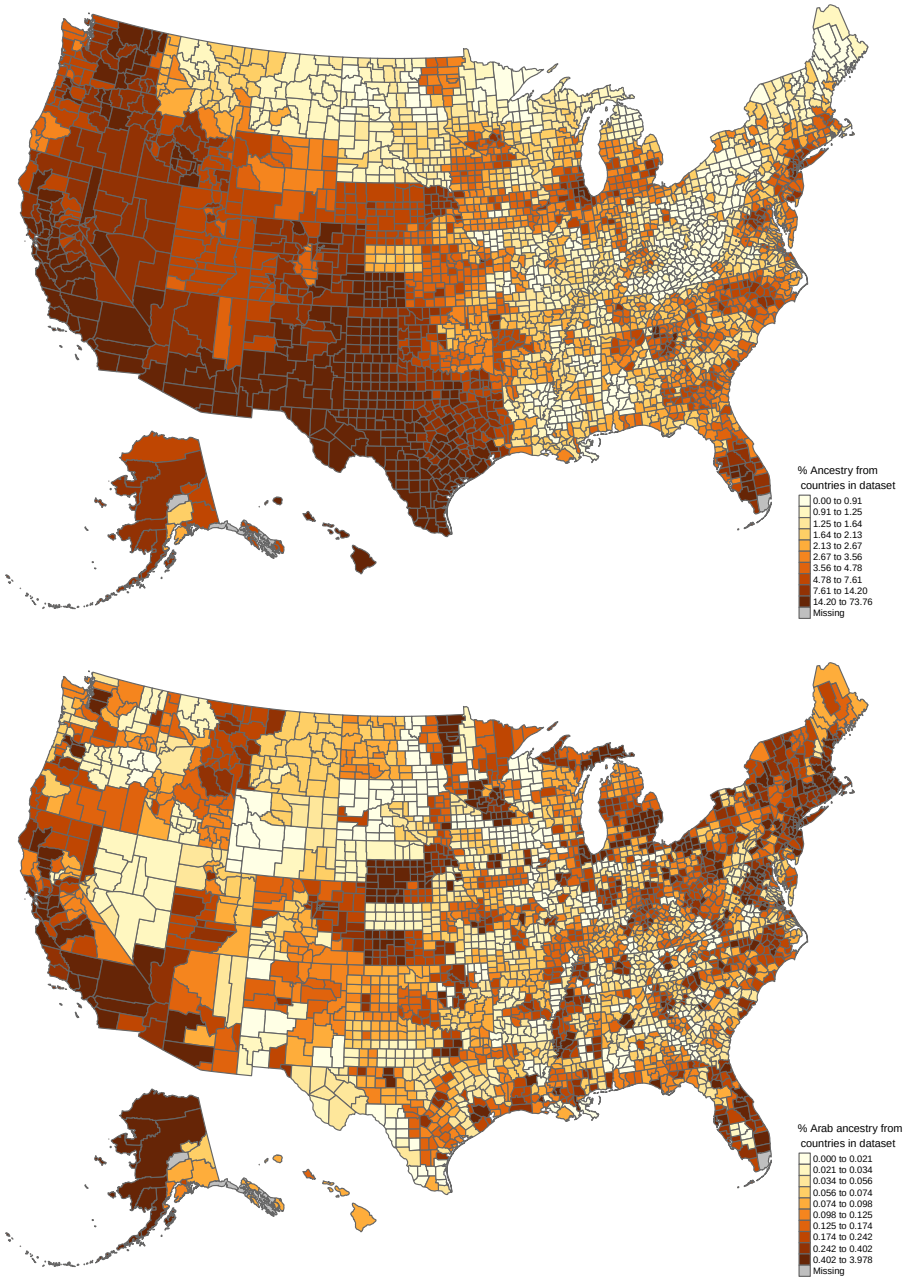
Notes: The table presents coefficient estimates from regressions at the county-country-quarter level. The dependent variable in Column 1 is the IHS-transformed number of donations from county to country in a quarter. The dependent variable in Column 2 is a dummy for the presence of at least one donation from county to country in a quarter. The dependent variable in Column 3 is the IHS-transformed total value of donations from county to country in a quarter. The main variable of interest is the IHS-transformed population with ancestry from country f in county d. In all columns, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. All specifications control for logged county-country distance and latitude difference. Standard errors are given in parentheses. Standard errors are clustered at the foreign country and domestic county levels.

Appendix Table A4—Survey Representativeness

	Survey mean	CCES mean
Age	52.416	50.344
Male	0.458	0.460
Hispanic	0.049	0.027
High school degree or higher	0.984	0.967
Family income		
under $20,000	0.071	0.121
$20,000 - 39,999	0.196	0.220
$40,000 - 59,999	0.197	0.197
$60,000 - 79,999	0.165	0.159
$80,000 - 99,999	0.109	0.100
$100,000 - 120,000	0.117	0.071
over $120,000	0.145	0.131
Census region		
Midwest	0.243	0.253
Northeast	0.168	0.199
South	0.383	0.349
West	0.206	0.200
Observations	5,063	115,930

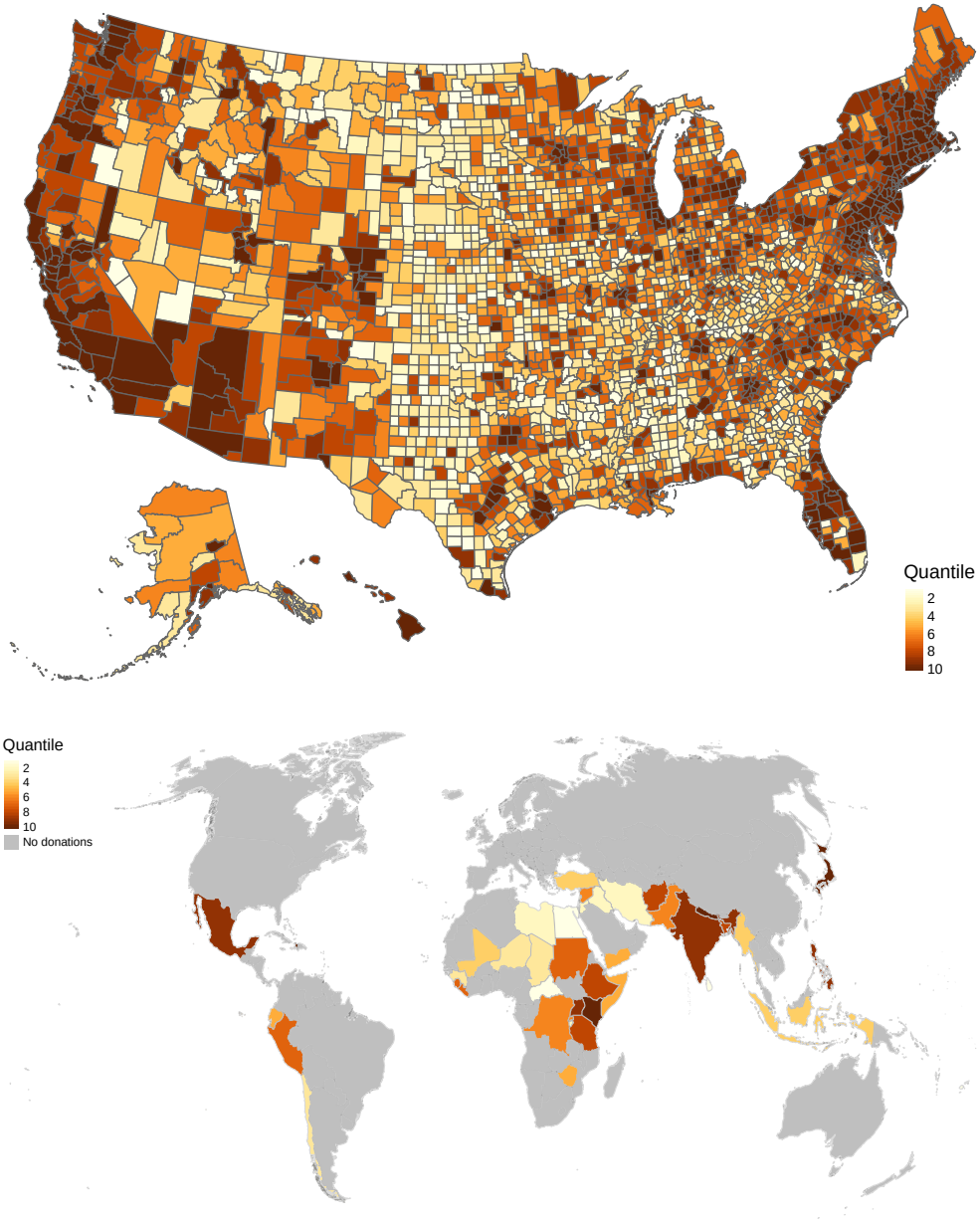
Notes: Column 1 presents means of respondent characteristics from our survey. Column 2 presents means of respondent characteristics from the 2016-2019 waves of the CCES.

Appendix Figure A1. Foreign ancestry share (top) and Arab ancestry share (bottom)



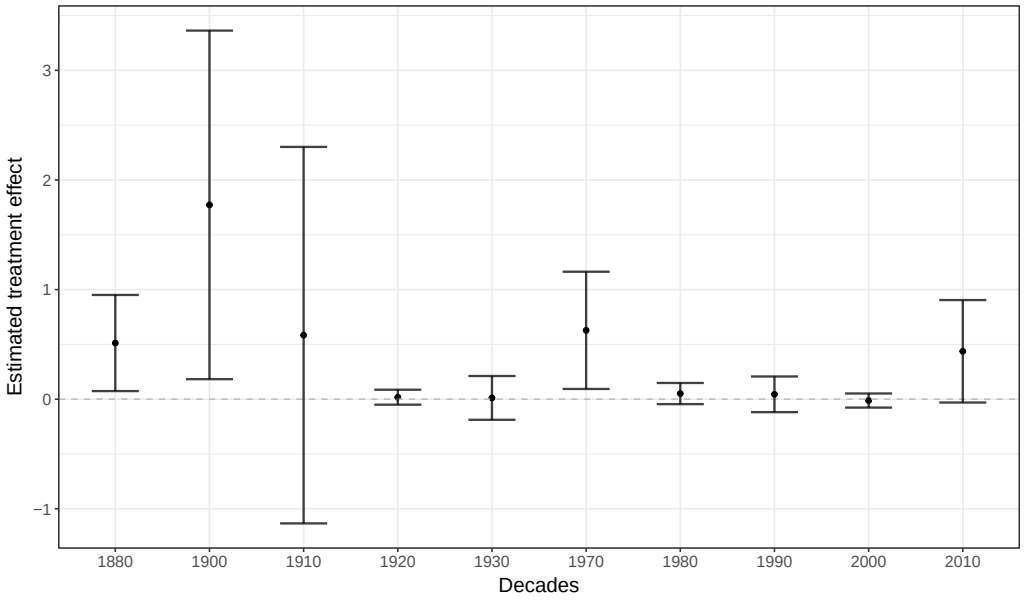
Notes: The top map plots the share of each county's population with ancestry from a country in our donations dataset. The bottom map plots the share of each county's population with ancestry from Arab countries in our donations dataset.

Appendix Figure A2. Donations by origin (top) and destination (bottom)



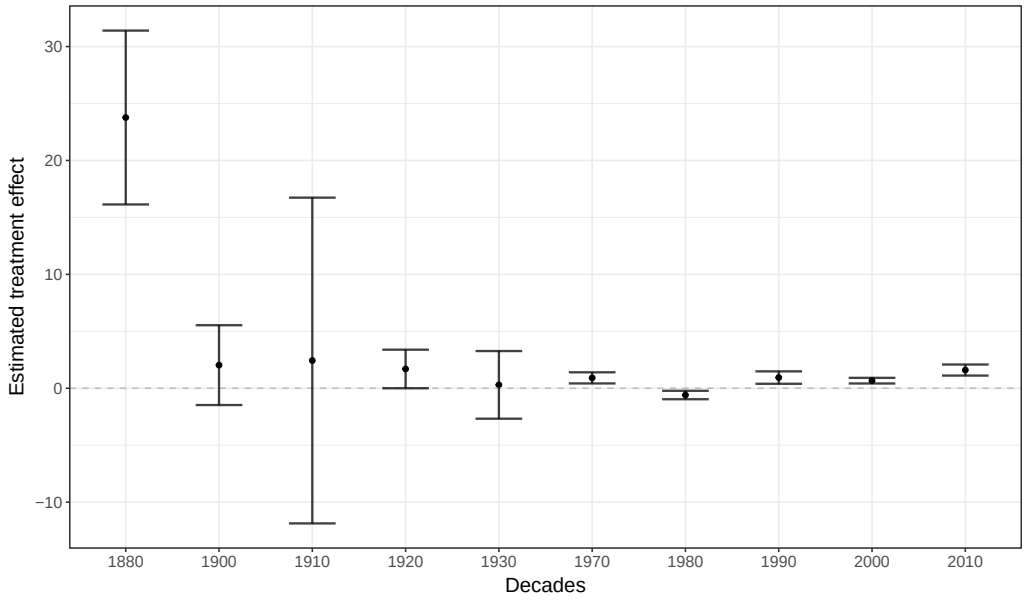
Notes: The top map plots the quantile of the number of donations in our dataset emanating from each domestic county. The bottom map plots the quantile of the number of donations in our dataset to each foreign country.

Appendix Figure A3. First-stage coefficients: all countries



Notes: Figure A3 presents coefficient estimates from regressions of IHS-transformed ancestry on the instruments in Equation (2). Following Burchardi, Chaney and Hassan (2019*b*), to facilitate the interpretation of coefficients as the marginal effect of migrations in that period, we sequentially orthogonalize each instrument with respect to the previous instruments. Error bars indicate 90% confidence intervals. Standard errors are clustered at the foreign country and domestic county levels.

Appendix Figure A4. First-stage coefficients: Arab-Muslim countries



Notes: Figure A4 presents coefficient estimates from regressions of IHS-transformed Arab-Muslim ancestry on the instruments in Equation (2). Following Burchardi, Chaney and Hassan (2019*b*), to facilitate the interpretation of coefficients as the marginal effect of migrations in that period, we sequentially orthogonalize each instrument with respect to the previous instruments. Error bars indicate 90% confidence intervals. Standard errors are clustered at the domestic county level.

Appendix Table A5—Stability of estimated effect of ancestral presence on donations, varying instruments

	(1)	(2)	(3)
	Eur. only pull	Excl. corr. origins	Excl. corr. dest.
IHS(Ancestry)	0.099	0.095	0.106
	(0.040)	(0.040)	(0.046)
First-stage F-statistic	133.3	160.0	202.0
Weak IV-robust p-value	< 0.01	< 0.01	0.43
Dep. var. mean	0.019	0.019	0.019
Dep. var. s.d.	0.182	0.182	0.182
Observations	4,703,862	4,703,862	4,703,862
Foreign country × quarter FE	Yes	Yes	Yes
Distance controls	Yes	Yes	Yes
US county × quarter FE	Yes	Yes	Yes

Notes: The table presents coefficient estimates from regressions at the county-country-quarter level. The dependent variable is the IHS-transformed number of donations from county to country in a quarter. The main variable of interest is the IHS-transformed population with ancestry from country f. In all columns, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. Column 1 uses an alternative construction of the instrument that calculates the pull factor based only on European emigrants; Column 2 uses an alternative construction of the instrument that excludes countries with correlated migrant flows; Column 3 uses an alternative construction of the instrument that excludes counties with correlated migrant flows. All specifications control for logged county-country distance and latitude difference. Standard errors are given in parentheses. Standard errors are clustered at the foreign country and domestic county levels.

Appendix Table A6—Effect of ancestral presence on donations: sensitivity to including principal components of interactions as instruments

	(1)	(2)	(3)	(4)	(5)	(6)
		IHS(# donations)			Donations (dummy)	IHS($ donations)
Panel A: IV, including principal components						
IHS(Ancestry)	0.139	0.132	0.132	0.107	0.047	0.329
	(0.028)	(0.032)	(0.033)	(0.043)	(0.021)	(0.137)
First-stage F-statistic	417.1	404.2	393.6	330.6	330.6	337.8
Weak IV-robust p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Panel B: IV, excluding principal components						
IHS(Ancestry)	0.137	0.130	0.130	0.114	0.052	0.354
	(0.025)	(0.028)	(0.029)	(0.044)	(0.023)	(0.138)
First-stage F-statistic	466.8	364.3	375.8	327.3	327.3	325.7
Weak IV-robust p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Observations	4,703,862	4,700,864	4,700,864	4,703,862	4,703,862	3,972,708
Foreign country × quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Distance controls	No	Yes	Yes	Yes	Yes	Yes
Demographic controls	No	Yes	Yes	—	—	—
US state × quarter FE	No	No	Yes	—	—	—
US county × quarter FE	No	No	No	Yes	Yes	Yes

Notes: The table presents coefficient estimates from regressions at the county-country-quarter level. The dependent variable in Columns 1–4 is the IHS-transformed number of donations from county to country in a quarter. The dependent variable in Column 5 is a dummy for the presence of at least one donation from county to country in a quarter. The dependent variable in Column 6 is the IHS-transformed total value of donations from county to country in a quarter (available only for Charity 2). The main variable of interest is the IHS-transformed population with ancestry from country f. In both panels, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ as excluded instruments. In Panel A, we include the first five principal components of the higher-order interactions of push and pull factors as additional excluded instruments. Columns 1–3 control for log 2010 population. Columns 2–6 include logged county-country distance and latitude difference. Columns 2 and 3 include the following county-level demographic controls (as of 2000): the shares of the population above 18, above 65, with a high school education, with a college education, below the poverty line, and living in a rural area; population density, the unemployment rate, and log income. Standard errors are given in parentheses. Standard errors are clustered at the foreign country and domestic county levels.

Appendix Table A7—Effect of ancestral presence on donations dropping recent periods from the instrument

	IHS(# donations)					
Includes decades until:	2010	2000	1990	1980	1970	1930
IHS(Ancestry)	0.114	0.107	0.106	0.103	0.101	0.065
	(0.044)	(0.050)	(0.050)	(0.051)	(0.051)	(0.088)
First-stage F-statistic	327.3	247.1	282.0	320.4	374.6	253.8
Weak IV-robust p-value	< 0.01	< 0.01	0.06	< 0.01	< 0.01	< 0.01
Observations	4,703,862	4,703,862	4,703,862	4,703,862	4,703,862	4,703,862
Foreign country × quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Distance controls	Yes	Yes	Yes	Yes	Yes	Yes
US county × quarter FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table presents coefficient estimates from regressions at the county-country-quarter level. The dependent variable is the IHS-transformed number of donations from county to country in a quarter. The main variable of interest is the IHS-transformed population with ancestry from country f. In the first column, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ as excluded instruments. Columns 2–6 incrementally drop the last decade of the instrument; i.e., Column 2 includes $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2000}$, Column 3 includes $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,1990}$, and so on. All columns control for logged county-country distance and latitude difference as well as foreign country × quarter and domestic county × quarter fixed effects. Standard errors are given in parentheses. Standard errors are clustered at the foreign country and domestic county levels.

Appendix Table A8—Stability of estimated effect of ancestral presence on donations, varying population

	(1)	(2)	(3)	(4)
	European donors	Other continents	Other countries	No country restriction
IHS(Ancestry)	0.107	0.110	0.116	0.157
	(0.043)	(0.045)	(0.048)	(0.077)
First-stage F-statistic	330.6	330.6	330.6	330.6
Weak IV-robust p-value	< 0.01	0.19	< 0.01	< 0.01
Dep. var. mean	0.019	0.021	0.022	0.024
Dep. var. s.d.	0.182	0.192	0.200	0.209
Observations	4,703,862	4,703,862	4,703,862	4,703,862
Foreign country × quarter FE	Yes	Yes	Yes	Yes
Distance controls	Yes	Yes	Yes	Yes
US county × quarter FE	Yes	Yes	Yes	Yes

Notes: The table presents coefficient estimates from regressions at the county-country-quarter level. Column 1 limits the sample to European donors; Column 2 additionally limits the sample to donors whose name is matched to a country on a different continent than the receiving country; Column 3 additionally limits the sample to donors whose name is matched to a country different than the receiving country; Column 4 presents the results for all donors with no limitation of the sample. The dependent variable is the IHS-transformed number of donations from county to country in a quarter. The main variable of interest is the IHS-transformed population with ancestry from country f. In all columns, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. All specifications control for logged county-country distance and latitude difference. Standard errors are given in parentheses. Standard errors are clustered at the foreign country and domestic county levels.

Appendix Table A9—Effect of ancestral presence on donations: Bounding misclassification error

	IHS(# donations)						
	All countries				Pooled Arab		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A:	*Baseline specification*						
IHS(Ancestry)	0.139	0.132	0.132	0.107	0.388	0.371	0.397
	(0.028)	(0.032)	(0.033)	(0.043)	(0.048)	(0.057)	(0.059)
Panel B:	*Lower bound accounting for potential misclassification*						
IHS(Ancestry)	0.120	0.110	0.109	0.077	0.378	0.358	0.385
	(0.019)	(0.021)	(0.022)	(0.027)	(0.047)	(0.055)	(0.058)
First-stage F-statistic	417.1	404.2	393.6	330.6	466.4	361.9	317.5
Observations	4,703,862	4,700,864	4,700,864	4,703,862	150,336	150,336	150,336
Foreign country × quarter FE	Yes	Yes	Yes	Yes	—	—	—
Distance controls	No	Yes	Yes	Yes	No	Yes	Yes
Demographic controls	No	Yes	Yes	—	No	Yes	Yes
US state × quarter FE	No	No	Yes	—	No	No	Yes
US county × quarter FE	No	No	No	Yes	No	—	—
Quarter FE	—	—	—	—	Yes	Yes	—

Notes: The table presents coefficient estimates from regressions at the county-country-quarter level. Panel A replicates our main specifications in Table 1 Columns 1–4 and Table 5 Columns 1–3. Panel B presents the lower bound after accounting for potential misclassification of surnames resulting from NamSor's algorithm. More particularly, we approximate the donations made by people who could be misclassified into the European surname group and subtract the number from our total number of donations. The dependent variable is the IHS-transformed number of such donations from county to country in a quarter in Columns 1–4 and the IHS-transformed number of such donations from the county to Arab League countries in a quarter in Columns 5–7. The main variable of interest is the IHS-transformed population with ancestry from country f. In all columns, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. Columns 1–3 and 5–7 control for log 2010 population. Columns 2–4 and 6–7 include logged county-country distance and latitude difference. Columns 2, 3, 6, and 7 include the following county-level demographic controls (as of 2000): the shares of the population above 18, above 65, with a high school education, with a college education, below the poverty line, and living in a rural area; population density, the unemployment rate, and log income. Standard errors are given in parentheses. Standard errors are clustered at the foreign country and domestic county levels.

Appendix Table A10—Effect of ancestral presence on donations, excluding different countries and Census regions

	IHS(# donations)			
Panel A: Excluding different countries				
Countries excluded:	Muslim-majority	Arab	Latin American	non-Arab African
IHS(Ancestry)	0.078	0.082	0.069	0.263
	(0.040)	(0.036)	(0.036)	(0.148)
First-stage F-statistic	397.7	349.7	722.1	157.6
Weak IV-robust p-value	< 0.01	< 0.01	< 0.01	< 0.01
Dep. var. mean	0.026	0.022	0.018	0.019
Dep. var. s.d.	0.218	0.198	0.176	0.183
Observations	2,479,020	3,605,562	4,195,506	3,025,032
Panel B: Excluding different census regions				
Census region excluded:	Northeast	South	Midwest	West
IHS(Ancestry)	0.113	0.111	0.111	0.089
	(0.042)	(0.046)	(0.047)	(0.039)
First-stage F-statistic	397.7	349.7	722.1	157.6
Weak IV-robust p-value	< 0.01	< 0.01	< 0.01	< 0.01
Dep. var. mean	0.015	0.025	0.024	0.016
Dep. var. s.d.	0.162	0.208	0.204	0.165
Observations	4,378,579	2,570,785	3,122,417	4,039,805
Foreign country × quarter FE	Yes	Yes	Yes	Yes
Distance controls	Yes	Yes	Yes	Yes
US county × quarter FE	Yes	Yes	Yes	Yes

Notes: The table presents coefficient estimates from regressions at the county-country-quarter level. The main variable of interest is the IHS-transformed population with ancestry from country f in county d. In all columns, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. In Panel A, Column 1 excludes Muslim-majority countries; Column 2 excludes Arab League countries; Column 3 excludes Latin American countries; and Column 4 excludes African countries which are not members of the Arab League. In Panel B, Column 1 excludes domestic counties in the Northeast; Column 2 excludes domestic counties in the South; Column 3 excludes domestic counties in the Midwest; and Column 4 excludes domestic counties in the West. All specifications control for logged county-country distance and latitude difference. Standard errors are given in parentheses. Standard errors are clustered at the foreign country and domestic county levels.

Appendix Table A11—Effect of ancestral presence on donations, different choices of clustering

	(1) All countries	(2) Arab countries (pooled)
	IHS(# donations)	
IHS(Ancestry)	0.107	0.397
Robust SE	(0.004)	(0.017)
Clustering: Foreign country	(0.044)	—
Clustering: Domestic county	(0.009)	(0.059)
Clustering: Domestic state	(0.012)	(0.084)
2-way clustering: Country/county	(0.043)	—
2-way clustering: Country/state	(0.042)	—
Dep. var. mean	0.019	0.048
Dep. var. sd	0.182	0.297
Observations	4,703,862	150,336
Distance controls	Yes	Yes
Foreign country × quarter FE	Yes	No
US county × quarter FE	Yes	No
Demographic controls	—	Yes
US state × quarter FE	—	Yes

Notes: The table presents coefficient estimates from regressions at the county-country-quarter level. We present standard errors associated with different choices of clustering. In Column 2, only donations to Arab League countries are included. In Column 1, the dependent variable is the IHS-transformed number of donations from county to country in a quarter. In Column 2, the dependent variable is the IHS-transformed number of donations from the county to Arab League countries in a quarter. The main variable of interest in Column 1 is the IHS-transformed population with ancestry from country f, while it is the IHS-transformed population with ancestry from Arab countries in Column 2. In both columns, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. All specifications control for logged county-country distance and latitude difference. Column 2 additionally includes the following county-level demographic controls (as of 2000): the shares of the population above 18, above 65, with a high school education, with a college education, below the poverty line, and living in a rural area; population density, the unemployment rate, log income, and log 2010 population.

Appendix Table A12—Effect of ancestral presence on donations, percent functional form

	(1)	(2)	(3)	(4)
	# of donations, per capita			
Percent country ancestry	0.016	0.017	0.018	0.019
	(0.009)	(0.007)	(0.007)	(0.007)
First-stage F-statistic	240.4	246.7	274.7	281.4
Weak IV-robust p-value	< 0.01	< 0.01	< 0.01	< 0.01
Dep. var. mean	0.019	0.019	0.019	0.019
Dep. var. sd	0.463	0.463	0.463	0.463
Observations	4,703,862	4,700,864	4,700,864	4,703,862
Foreign country × quarter FE	Yes	Yes	Yes	Yes
Distance controls	No	Yes	Yes	Yes
Demographic controls	No	Yes	Yes	—
US state × quarter FE	No	No	Yes	—
US county × quarter FE	No	No	No	Yes

Notes: The table presents coefficient estimates from regressions at the county-country-quarter level. The dependent variable is the number of donations per capita from county to country in a given quarter. The main variable of interest is the percentage of the population with ancestry from country f. In all columns, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. Columns 1–3 control for log 2010 population. Columns 2–4 include logged county-country distance and latitude difference. Columns 2 and 3 include the following county-level demographic controls (as of 2000): the shares of the population above 18, above 65, with a high school education, with a college education, below the poverty line, and living in a rural area; population density, the unemployment rate, and log income. Standard errors are given in parentheses. Standard errors are clustered at the foreign country and domestic county levels.

Appendix Table A13—Effect of ancestral presence on donations by ancestry: aggregating to continents

	(1)	(2)	(3)	(4)
	IHS(# donations)			
Panel A: Effect of continent ancestry				
IHS(Ancestry), Continent exc. country	0.053	0.116	0.102	0.105
	(0.015)	(0.026)	(0.027)	(0.026)
First-stage F-statistic	4992.1	3685.9	3051.2	2727.1
Weak IV-robust p-value	< 0.01	< 0.01	< 0.01	0.04
Panel B: Effect of continent and country ancestry				
IHS(Ancestry)	0.056	0.067	0.061	0.060
	(0.021)	(0.013)	(0.013)	(0.013)
IHS(Ancestry), Continent exc. country	−0.007	0.015	0.009	0.011
	(0.024)	(0.024)	(0.023)	(0.023)
F-stat IHS(Ancestry)	48,254.97	36,527.74	1,375.46	180.37
F-stat IHS(Ancestry), Continent exc. country	65,011.80	426,747.72	36,843.66	8,445.44
Observations	4,703,862	4,703,862	4,700,864	4,700,864
Continent × quarter FE	No	Yes	Yes	Yes
Distance controls	No	No	Yes	Yes
Demographic controls	No	No	Yes	Yes
US state × quarter FE	No	No	No	Yes

Notes: The table presents coefficient estimates from regressions at the county-country-quarter level. The dependent variable is the IHS-transformed number of donations from county to country in a quarter. The main variables of interest are the IHS-transformed population with ancestry from country f and the IHS-transformed population with ancestry from continent c, excluding country f. In all columns, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. To instrument the 2010 IHS-transformed population with ancestry aggregated to the continent level c excluding foreign country f, we modify the push factor, $I^s_{f,-r(d)}$, in Equation (2) as the total number of migrants arriving from continent c in period s, excluding those from country f and those who settle in d's region, i.e. $I^s_{c(-f),-r(d)}$. Columns 1–4 control for log 2010 population. Columns 3 and 4 include logged county-country distance and latitude difference and the following county-level demographic controls (as of 2000): the shares of the population above 18, above 65, with a high school education, with a college education, below the poverty line, and living in a rural area; population density, the unemployment rate, and log income. The table reports Sanderson-Windmeijer conditional first-stage F-statistics. Standard errors are given in parentheses. Standard errors are clustered at the foreign country and domestic county levels.

Appendix Table A14—Effect of presence of Arab ancestry on auxiliary measures of prejudice and social norms

	(1) Standards	(2) Disapproval	(3) Beliefs (1)	(4) Beliefs (2)
Panel A: IV				
IHS(Arab ancestry)	0.046	0.010	0.086	0.091
	(0.026)	(0.030)	(0.037)	(0.042)
AP F-statistic	9.960	9.934	9.949	9.963
Weak IV-robust p-value	< 0.01	0.95	< 0.01	< 0.01
Panel B: OLS				
IHS(Arab ancestry)	0.023	0.016	0.035	0.033
	(0.005)	(0.005)	(0.007)	(0.007)
Observations	107,425	107,350	107,707	107,816
State FE	Yes	Yes	Yes	Yes
Individual-level demographics	Yes	Yes	Yes	Yes

Notes: The table presents coefficient estimates from regressions at the individual level. The dependent variables represent agreement with different statements about prejudice and social norms; all outcomes are scaled to mean zero and standard deviation one such that higher values indicate less prejudice. "Standards" refers to the statement "Because of today's standards I try to appear nonprejudiced toward Arab Muslims" (Column 1); "Disapproval" refers to the statement "I attempt to appear nonprejudiced toward Arab Muslims in order to avoid disapproval from others" (Column 2); "Beliefs (1)" refers to the statement "I am personally motivated by my beliefs to be nonprejudiced toward Arab Muslims" (Column 3); and "Beliefs (2)" refers to the statement "Because of my personal values, I believe that using stereotypes about Arab Muslims is wrong" (Column 4). Only respondents who self-reported their reason for taking the Project Implicit test as "Assigned for work," "Assigned for school," or "Assigned for discussion group" are included. The main variable of interest is the 2010 IHS-transformed population with ancestry from Arab League countries. In Panel A, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. All specifications control for log 2010 population. Individual demographics include age, male, age squared, and age × male. Standard errors are given in parentheses. Standard errors are clustered at the county level.

Appendix Table A15—Effect of presence of Arab ancestry on attitudes toward Arab-Muslims, forced and unforced respondents

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV	(6) IV	(7) IV
Panel A:	*Score on Arab-Muslim IAT (std., higher score = less prejudiced)*						
IHS(Arab ancestry)	0.011	0.062	0.057	0.063	0.062	0.044	0.054
	(0.005)	(0.017)	(0.019)	(0.022)	(0.026)	(0.017)	(0.021)
IHS(non-Euro ancestry)					−0.023		
					(0.017)		
Avg. race IAT score						0.377	
						(0.048)	
2012 Rep. vote share							−0.127
							(0.044)
AP F-statistic	—	13.98	11.11	6.645	6.653	6.812	6.154
Weak IV-robust p-value	—	0.16	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Observations	228,411	228,411	225,768	225,768	225,768	225,768	225,768
Panel B:	*Warmth toward Arab-Muslims (std., higher score = more favorable)*						
IHS(Arab ancestry)	0.036	0.132	0.128	0.108	0.111	0.077	0.089
	(0.007)	(0.023)	(0.030)	(0.031)	(0.033)	(0.025)	(0.030)
IHS(non-Euro ancestry)					−0.046		
					(0.021)		
Avg. race IAT score						0.604	
						(0.060)	
2012 Rep. vote share							−0.259
							(0.059)
AP F-statistic	—	14.00	11.15	6.588	6.614	6.770	6.095
Weak IV-robust p-value	—	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Observations	228,902	228,902	226,301	226,301	226,301	226,300	226,301
State FE	No	No	Yes	Yes	Yes	Yes	Yes
Individual-level demographics	No	No	Yes	Yes	Yes	Yes	Yes
County-level demographics	No	No	No	Yes	Yes	Yes	Yes

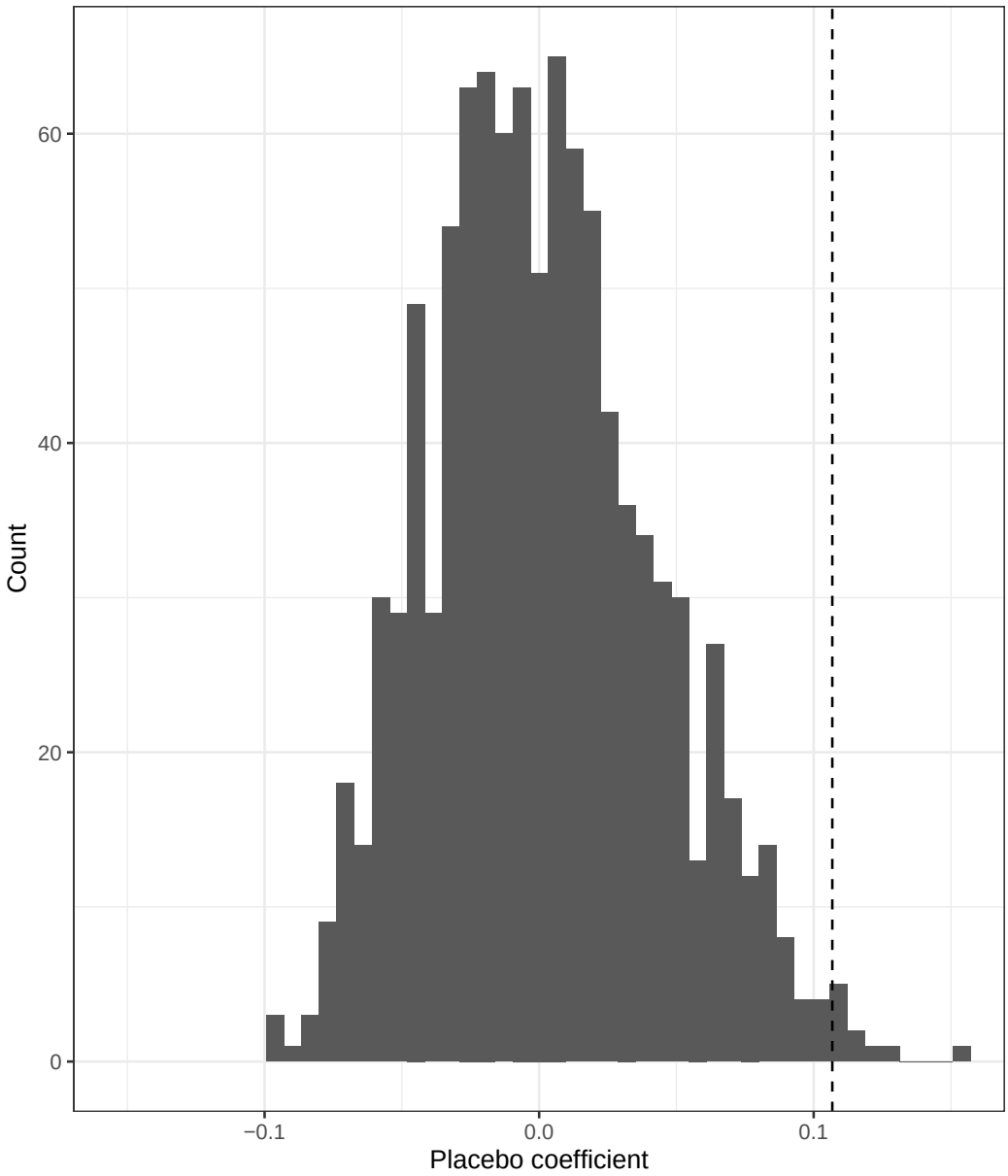
Notes: The table presents coefficient estimates from regressions at the individual level. The dependent variable in Panel A is the score on the Arab-Muslim IAT (from Project Implicit); the dependent variable in Panel B is the stated warmth toward Arab-Muslims (also from Project Implicit). Both measures are scaled to take mean zero and standard deviation one. The main variable of interest is the 2010 IHS-transformed population with ancestry from Arab League countries. We include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. All specifications control for log 2010 population. Individual demographics include age, male, age squared, and age × male. County-level demographic controls are as of 2000 and include the shares of the population above 18, above 65, with a high school education, with a college education, below the poverty line, and living in a rural area; population density, the unemployment rate, and log income. Standard errors are clustered at the county level.

Appendix Table A16—Effect of presence of Arab ancestry on attitudes toward Arab-Muslims and political preferences, representative sample

	(1) Favorability	(2) Trump	(3) Muslim Ban
Panel A: IV			
IHS(Arab ancestry)	0.114	−0.061	−0.076
	(0.029)	(0.018)	(0.018)
AP F-statistic	10.32	10.59	10.61
Weak IV-robust p-value	< 0.01	< 0.01	< 0.01
Panel B: OLS			
IHS(Arab ancestry)	0.034	−0.015	−0.014
	(0.005)	(0.003)	(0.003)
Observations	189,273	171,944	58,750
State FE	Yes	Yes	Yes
Individual-level demographics	Yes	Yes	Yes

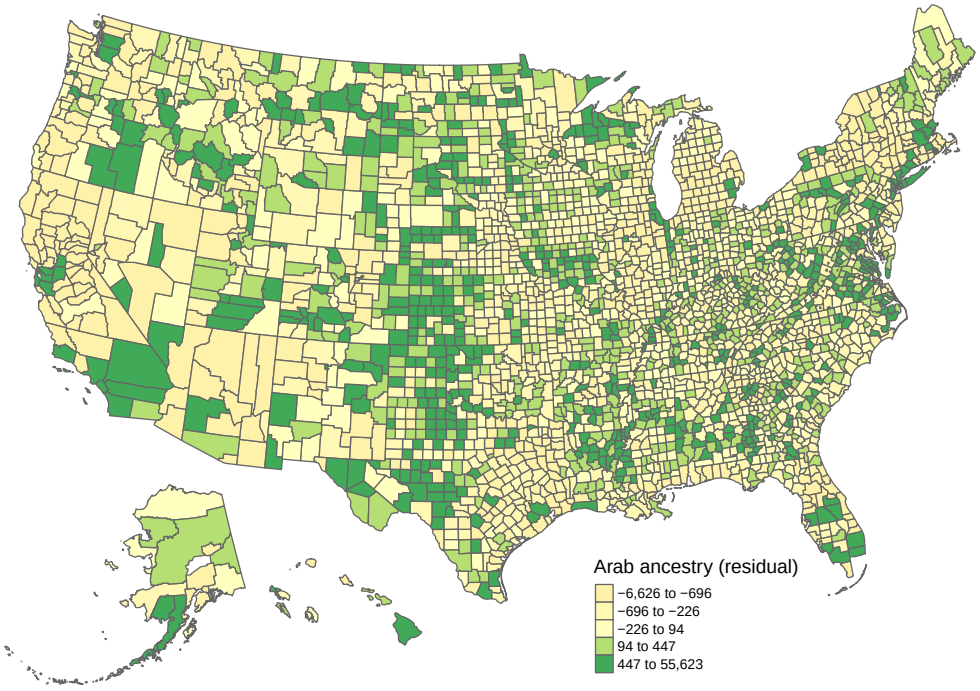
Notes: The table presents coefficient estimates from regressions at the individual level. The dependent variable in Column 1 is the stated favorability toward Muslims; the dependent variable in Column 2 is self-reported Trump votership; and the dependent variable in Column 3 is stated support for the Muslim Ban. The data is from Nationscape. The main variable of interest is the 2010 IHS-transformed population with ancestry from Arab League countries. In Panel A, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. All specifications control for log 2010 population. Individual demographics include age, male, age squared, and age × male. Standard errors are given in parentheses. Standard errors are clustered at county level.

Appendix Figure A5. Effect of ancestral presence on donations, permutation test



Notes: Figure A5 presents the results of a permutation test in which we permute ancestry and the excluded instruments, such that our regression estimates an average of the effect of the presence of one ancestral group on donations toward another country. The dotted line is placed at the true coefficient estimate. We include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. We control for logged county-country distance and latitude difference as well as foreign country × quarter and domestic county × quarter fixed effects.

Appendix Figure A6. Residualized predicted values of Arab-Muslim ancestry

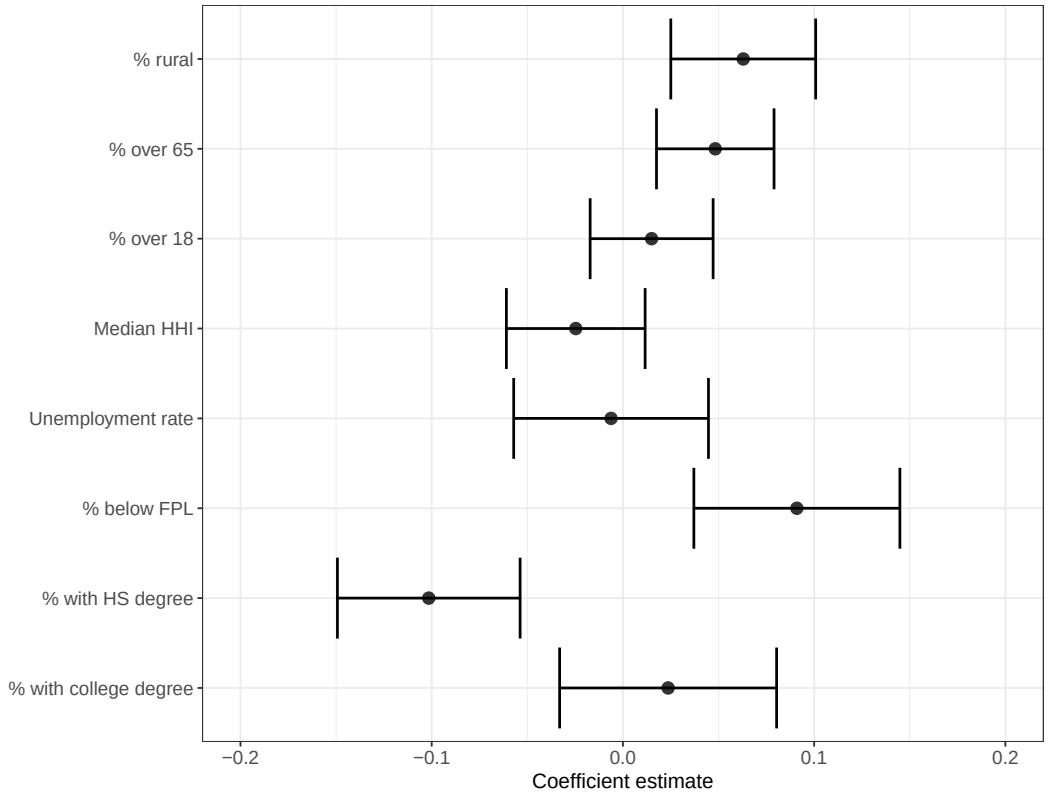


Notes: Figure A6 maps the residualized values of predicted Arab-Muslim ancestry, where we use $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as instruments and residualize by state fixed effects, log population, and the following county-level demographic controls: the shares of the population with a high school education, with a college education, and population density as of 2000.

Appendix Figure A7. Arab-ancestry population across counties

Notes: Figure A7 plots the IHS-transformed 2010 population of each US county against the IHS-transformed 2010 Arab-ancestry population of that county.

Appendix Figure A8. Balance test of Arab-Muslim instruments



Notes: Figure A8 presents coefficient estimates from regressions of a number of demographic characteristics (scaled to take mean zero and standard deviation one) on the predicted values of IHS-transformed Arab-Muslim ancestry (scaled similarly). We include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. All regressions control for log 2010 population and include state fixed effects. Error bars indicate 95% confidence intervals. Standard errors are clustered at the state level.

Appendix Table A17—Effect of presence of Arab ancestry on attitudes toward different groups

	(1)	(2)	(3)	(4)	(5)	(6)
	Arab-Muslim		Asian		Black	
	IAT	Warmth	IAT	Warmth	IAT	Warmth
Panel A:	*Unweighted*					
IHS(Arab ancestry)	0.075	0.136	0.036	0.035	0.019	0.029
	(0.027)	(0.033)	(0.029)	(0.028)	(0.017)	(0.015)
AP F-statistic	9.907	9.950	10.66	11.57	10.17	10.21
Weak IV-robust p-value	< 0.01	< 0.01	0.12	< 0.01	< 0.01	< 0.01
Observations	107,110	106,999	74,871	34,926	1,128,738	1,128,106
Panel B:	*Reweighted to match Arab-Muslim test-takers on observables*					
IHS(Arab ancestry)	0.075	0.136	0.034	0.032	0.021	0.033
	(0.027)	(0.033)	(0.030)	(0.030)	(0.017)	(0.014)
AP F-statistic	9.907	9.950	10.70	11.65	10.10	10.14
Weak IV-robust p-value	< 0.01	< 0.01	0.03	0.06	< 0.01	< 0.01
Observations	107,110	106,999	69,899	32,626	1,042,155	1,042,092
Panel C:	*Limiting to counties in Arab-Muslim data, unweighted*					
IHS(Arab ancestry)	0.075	0.136	0.036	0.036	0.017	0.029
	(0.027)	(0.033)	(0.028)	(0.027)	(0.016)	(0.014)
AP F-statistic	9.907	9.950	10.87	12.09	10.45	10.49
Weak IV-robust p-value	< 0.01	< 0.01	0.13	< 0.01	< 0.01	< 0.01
Observations	107,110	106,999	74,667	34,824	1,124,106	1,123,490
Panel D:	*Limiting to counties in Arab-Muslim data, reweighted*					
IHS(Arab ancestry)	0.075	0.136	0.034	0.033	0.019	0.032
	(0.027)	(0.033)	(0.029)	(0.030)	(0.017)	(0.014)
AP F-statistic	9.907	9.950	10.91	12.21	10.38	10.42
Weak IV-robust p-value	< 0.01	< 0.01	0.03	0.05	< 0.01	< 0.01
Observations	107,110	106,999	69,703	32,527	1,037,977	1,037,927
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual demographics	Yes	Yes	Yes	Yes	Yes	Yes

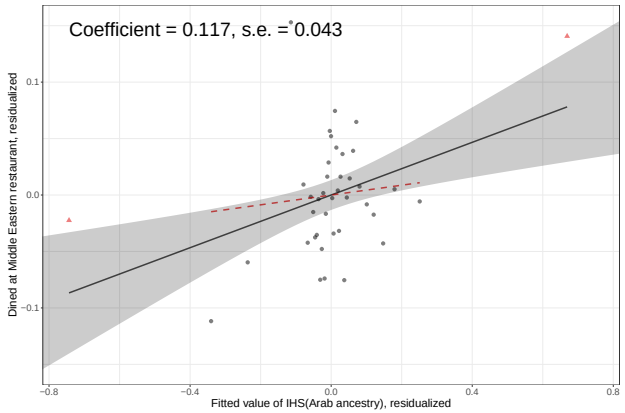
Notes: The table presents coefficient estimates from regressions at the individual level. The dependent variables in Columns 1, 3, and 5 are IAT scores toward Arab-Muslims, Asians, and Black Americans, respectively. The dependent variables in Columns 2, 4, and 6 are stated warmth toward Arab-Muslims, Asians, and Black Americans, respectively. Panel A weights all observations equally. We conduct a t-test to test the null that effects on attitudes toward Asians and Blacks are equal to the effects on Arab-Muslims. The resulting p-value is 0.31 for the comparison of Columns 1 and 3; 0.076 for Columns 1 and 5; 0.02 for Columns 2 and 4; and 0.003 for Columns 2 and 6. Columns 3–6 of Panel B reweight observations to match the sample of Columns 1–2 on age, gender, education, and Hispanic status; Columns 3–6 of Panel C limit the sample to counties with at least one Arab-Muslim IAT; and Columns 3–6 of Panel D first limit the sample to counties with at least one Arab-Muslim IAT, then reweight observations to match the sample of Columns 1–2 on age, gender, education, and Hispanic status. All measures are scaled to take mean zero and standard deviation one. Only respondents who self-reported their reason for taking the Project Implicit test as "Assigned for work," "Assigned for school," or "Assigned for discussion group" are included. The main variable of interest is the 2010 IHS-transformed population with ancestry from Arab League countries. We include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. All specifications control for log 2010 population. Standard errors are given in parentheses. Standard errors are clustered at the county level.

Appendix Table A18—Effect of presence of Arab ancestry on political preferences, individual Romney control

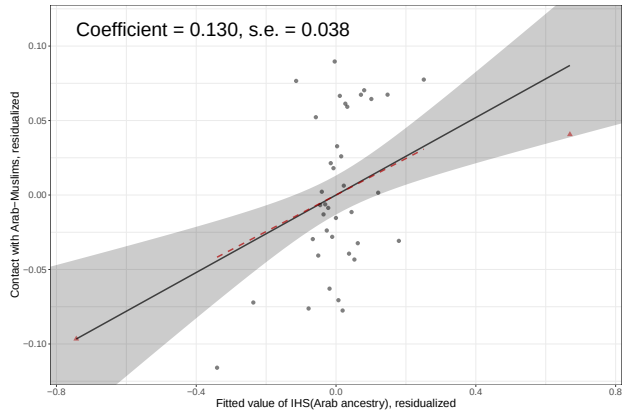
	(1) OLS	(2) IV	(3) IV	(4) IV
	Voted for Trump in 2016			
IHS(Arab ancestry)	−0.012	−0.052	−0.061	−0.035
	(0.003)	(0.014)	(0.021)	(0.019)
Voted for Romney in 2012	0.739	0.736	0.725	0.724
	(0.005)	(0.005)	(0.005)	(0.005)
AP F-statistic	—	17.28	9.486	5.306
Weak IV-robust p-value	—	< 0.01	< 0.01	< 0.01
Observations	32,748	32,748	32,748	32,748
State FE	No	No	Yes	Yes
Individual-level demographics	No	No	Yes	Yes
County-level demographics	No	No	No	Yes

Notes: The table presents coefficient estimates from regressions at the individual level. The dependent variable is self-reported Trump votership. The data is from the CCES. The main variable of interest is the 2010 IHS-transformed population with ancestry from Arab League countries. We include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. All specifications control for log 2010 population. Individual demographics include age, male, age squared, and age × male. County-level demographic controls are as of 2000 and include the shares of the population above 18, above 65, with a high school education, with a college education, below the poverty line, and living in a rural area; population density, the unemployment rate, and log income. Standard errors are given in parentheses. Standard errors are clustered at the county level.

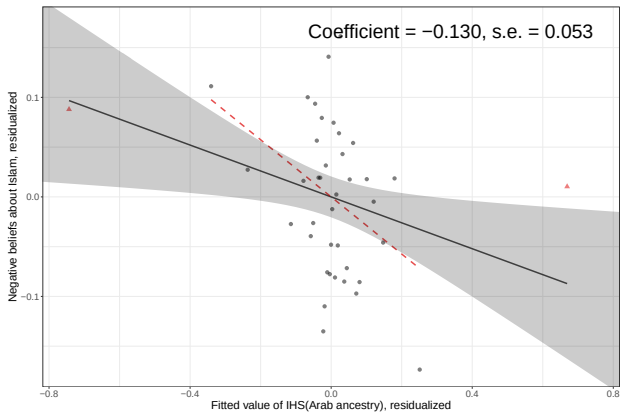
Appendix Figure A9. Binned scatter plots of contact and knowledge



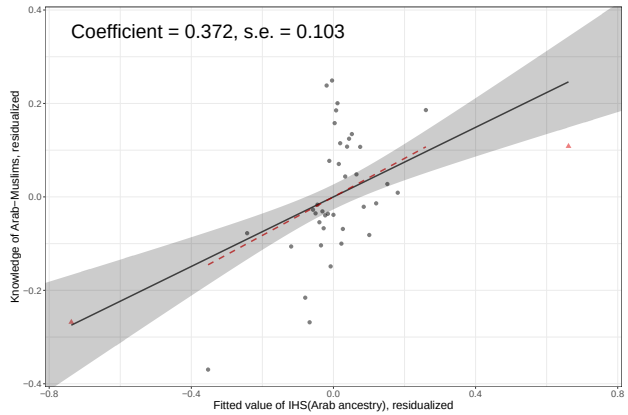
(a) Visited Middle Eastern restaurant



(b) Contact with Arab-Muslims



(c) Negative beliefs about Islam



(d) Knowledge about Islam

Notes: Figure A9 presents binned scatter plots displaying the relationship between the fitted values of IHS(Arab ancestry) and four outcomes: an indicator taking value one if the respondent reports ever visiting a Middle Eastern restaurant, an indicator taking value one if the respondent personally knows an Arab-Muslim friend, neighbor, or colleague; a measure of the respondent's negative beliefs about Islam; and an index measuring respondents' knowledge of Islam. The main variable of interest is the 2010 IHS-transformed population with ancestry from Arab League countries. We include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. We residualize outcomes and instruments by the controls used in Columns 1–5 of Table 8. Red triangles are used to indicate the top and bottom 2.5% of the data by fitted values; the red dotted line indicates the regression fit after dropping observations in the top and bottom 2.5% of fitted values. Standard errors are clustered at the county level. 95% confidence intervals are reported.

Appendix Table A19—Robustness across different definitions of Muslim ancestry

	(1) IAT	(2) Warmth	(3) Muslim Ban	(4) Trump vote	(5) Contact	(6) Knowledge
Panel A: Arab-Muslim ancestry						
IHS(Ancestry)	0.075	0.136	−0.077	−0.075	0.130	0.372
	(0.027)	(0.033)	(0.024)	(0.027)	(0.038)	(0.103)
AP F-statistic	9.907	9.950	9.650	9.779	8.300	7.903
Weak IV-robust p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Panel B: Ancestry from Muslim Ban countries						
IHS(Ancestry)	0.061	0.124	−0.089	−0.070	0.095	0.128
	(0.027)	(0.034)	(0.030)	(0.037)	(0.035)	(0.079)
AP F-statistic	3.808	3.794	7.462	7.026	17.63	17.39
Weak IV-robust p-value	< 0.01	< 0.01	< 0.01	< 0.01	0.02	< 0.01
Panel C: Ancestry from Muslim-majority countries						
IHS(Ancestry)	0.076	0.131	−0.086	−0.074	0.130	0.324
	(0.028)	(0.038)	(0.029)	(0.034)	(0.029)	(0.065)
AP F-statistic	6.526	6.592	5.575	5.790	5.785	5.945
Weak IV-robust p-value	< 0.01	< 0.01	< 0.01	0.01	0.02	< 0.01
Dep. var. mean	0.018	0.035	0.529	0.463	0.397	0.000
Dep. var. sd	0.989	0.996	0.499	0.499	0.489	1.000
Observations	107,110	106,999	57,195	98,205	5,051	4,757
State FE	Yes	Yes	Yes	Yes	No	No
Individual demographics	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table presents coefficient estimates from regressions at the individual level. The dependent variables in Columns 1–6 are scores on the Arab-Muslim IAT, stated warmth toward Arab-Muslims, stated support for the Muslim Ban, self-reported Trump voting, an indicator for whether the respondent has an Arab-Muslim friend, workplace acquaintance, or neighbor, and a normalized index of knowledge about Arab-Muslims. The dependent variables in Columns 1–2 are drawn from Project Implicit; the dependent variables in Columns 3–4 are drawn from the CCES; and the dependent variables in Columns 5–6 are drawn from our survey. In Columns 1 and 2, only respondents who self-reported their reason for taking the Project Implicit test as "Assigned for work," "Assigned for school," or "Assigned for discussion group" are included. The main variable of interest in Panel A is the 2010 IHS-transformed population with ancestry from Arab League countries; the main variable of interest in Panel B is the 2010 IHS-transformed population with ancestry from countries affected by Executive Order 13769 ("Muslim ban"); and the main variable of interest in Panel C is the 2010 IHS-transformed population with ancestry from Muslim-majority countries. We include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. All specifications control for log 2010 population. Individual demographics include age, male, age squared, and age × male. Standard errors are given in parentheses. Standard errors are clustered at the county level.

Appendix Table A20—Effect of presence of Arab ancestry, percent functional form

	(1) Donations	(2) IAT	(3) Warmth	(4) Muslim Ban	(5) Trump vote	(6) Contact	(7) Knowledge
Percent Arab ancestry	0.123	0.073	0.177	−0.130	−0.151	0.257	0.588
	(0.059)	(0.026)	(0.052)	(0.049)	(0.051)	(0.103)	(0.225)
AP F-statistic	498.55	10.77	10.79	10.92	10.28	13.28	12.53
Weak IV-robust p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Dep. var. mean	0.043	0.018	0.035	0.529	0.463	0.397	0.000
Dep. var. sd	0.527	0.989	0.996	0.499	0.499	0.489	1.000
Observations	150,336	107,110	106,999	57,195	98,205	5,051	4,757
State FE	–	Yes	Yes	Yes	Yes	No	No
State × quarter FE	Yes	No	No	No	No	No	No
Distance controls	Yes	No	No	No	No	No	No
County-level demographics	Yes	No	No	No	No	No	No
Individual demographics	No	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table presents coefficient estimates from regressions at the county-quarter (Column 1) and individual (Columns 2–7) levels. The dependent variable in Column 1 is the number of donations per capita from the county to Arab League countries in a quarter. The dependent variables in Columns 2–7 are scores on the Arab-Muslim IAT, stated warmth toward Arab-Muslims, stated support for the Muslim Ban, self-reported Trump voting, an indicator for whether the respondent has an Arab-Muslim friend, workplace acquaintance, or neighbor, and a normalized index of knowledge about Arab-Muslims. The dependent variables in Columns 2–3 are drawn from Project Implicit; the dependent variables in Columns 4–5 are drawn from the CCES; and the dependent variables in Columns 6–7 are drawn from our survey. Only respondents who self-reported their reason for taking the Project Implicit test as "Assigned for work," "Assigned for school," or "Assigned for discussion group" are included. The main variable of interest is the percentage of the population with ancestry from Arab countries. In all columns, we include $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,2010}$ and the first five principal components of the higher-order interactions of push and pull factors as excluded instruments. All specifications control for log 2010 population. Individual demographics include age, male, age squared, and age × male. Standard errors are given in parentheses. Standard errors are clustered at the county level.

Appendix Table A21—Effect of presence of Arab ancestry on white flight

	(1)	(2)
Panel A:	*Selective white flight index*	
IHS(Ancestry)	0.025	0.017
	(0.003)	(0.002)
First-stage F-statistic	63.72	50.63
Weak IV-robust p-value	< 0.01	< 0.01
Dep. var. mean	0.062	0.062
Dep. var. s.d.	0.029	0.029
Observations	9,415	9,415
Panel B:	*Selective white flight index, by subgroup*	
IHS(Ancestry) × Married	−0.002	−0.002
	(0.0003)	(0.0003)
IHS(Ancestry) × Female	0.0004	0.0004
	(0.0001)	(0.0001)
IHS(Ancestry) × College	0.001	0.001
	(0.0004)	(0.0004)
IHS(Ancestry) × Age	0.001	0.001
	(0.0003)	(0.0003)
IHS(Ancestry) × Income	0.002	0.002
	(0.0005)	(0.0005)
Year FE	Yes	Yes
US state FE	No	Yes

Notes: The table presents coefficient estimates from regressions at the country-county-decade level. The dependent variable is the selective white flight index, defined in Section II.D; in Panel A, the index is computed from the full sample, whereas in Panel B, two separate indices are computed for each dimension of heterogeneity (one for each subgroup). The endogenous variable in Panel A is the IHS-transformed population with ancestry from Arab League countries. Each row of Panel B presents a separate regression of the selective white flight index for a given subgroup on an indicator for the subgroup, IHS-transformed population with ancestry from Arab League countries, and the interaction of the indicator and IHS-transformed ancestral population. The excluded instruments in Panel A are $\{I^t_{f,-r(d)}(I^t_{-c(f),d}/I^t_{-c(f)})\}_{t=1880,\ldots,1980}$ and the first five principal components of the higher-order interactions; in Panel B, we additionally include as instruments the interaction of each instrument with the subgroup indicator. Standard errors are given in parentheses. Standard errors are clustered at the foreign country and domestic county levels.

II. Data Appendix

A. *Details on the construction of migration and ethnicity data*

County residence is defined at the level of historic counties, and at the level of historic county groups or PUMAs starting in 1970. Whenever necessary, we use contemporaneous population weights to transition data from the historic county group or PUMA to historic county, and then area weights to transition data from the historic county to 1990 counties. Stated ancestry often corresponds to foreign countries in their 1990 borders (e.g. "Syrian"), though not always. In cases with ambiguous correspondence (e.g. "Kurdish"), we construct transition matrices that map into 1990 national boundaries using approximate population weights when feasible and approximate area weights otherwise.

Calculation of post-1880 flow of immigrants

For each census wave after 1880, we count the number of individuals in each historic US domestic county d who were born in historic foreign country f (as identified by birthplace variable "bpld" in the raw data) that had immigrated to the United States since the last census wave that contains the immigration variable (not always 10 years earlier). Then we transform these data

- from the non-1990 foreign-country ("bpld") level to the 1990 foreign-country level using bpld-to-country transition matrices.
- from the US-county group/puma level to the US-county level using group/puma-to-county transition matrices.
- from the non-1990 US-county level to the 1990 US-county level using county-to-county transition matrices.
- from the post-1990 US-county level to the 1990 US county level. Based on the information from `https://www.census.gov/programs-surveys/geography/technical-documentation/county-changes.html`, a new county is either created from part of ONE 1990 county or assigned a new FIPS code after 1990, so we manually change that county's FIPS code to what it was in 1990. A few counties' boundaries have been changed after 1990 but that only involved a tiny change in population, so we ignore these differences.

Calculation of pre-1880 stock of immigrants

The initial 1880 Census did not report the immigration date. Thus, for the year 1880, we calculate for each historic US county d the number of individuals who were born in a historic foreign country f (no matter when they immigrated). We add to those calculations

the number of individuals in county d who were born in the United States, but whose parents were born in historic foreign country f. (If the parents were born in different countries, we count the person as half a person from the mother's place of birth, and half a person from the father's place of birth). Then we transform these data

- from the pre-1880 foreign-country ("bpld") level to the 1990 foreign-country level using the pre-1880 country-to-country transition matrix.
- from the pre-1880 US-county level to the 1990 US-county level using the pre-1880 county-to-county transition matrix.

Calculation of stock of ancestry (1980, 1990, 2000, and 2010)

For the years 1980, 1990, 2000, and 2010, we calculate for each US county group the number of individuals who state as primary ancestry ("ancestr1" variable) some nationality/area. We transform the data

- from the ancestry-answer ("ancestr1") level to the 1990 foreign-country level using ancestry-to-country transition matrices.
- from the US-county group/puma level to the US county-level using group/puma-to-county transition matrices.
- from the non-1990 US-county level to the 1990 US-county level using county-to-county transition matrices.
- from the post-1990 US-county to the 1990 US-county level. Based on the information from `https://www.census.gov/programs-surveys/geography/technical-documentation/county-changes.html`, a new county is either created from part of ONE 1990 county or assigned a new FIPS code after 1990, so we manually change that county's FIPS code to what it was in 1990. A few counties' boundaries have been changed after 1990 but that only involved a tiny change in population, so we ignore the difference.

Correcting mis-classification of donors due to intermarriages

- From "ancestry1d", for each respondent i we construct a dummy variable $ancestry_i(f)$ equal to 1 if respondent i has ancestry from f.
- From "ancestry1d", "sex", "sploc", we compute for each respondent i a dummy variable, $wife_i(f)$, equal to 1 if respondent i is female, of ancestry from country f, and married to a spouse of European ancestry.
- From "ancestry1d" and "ancestry2d", we compute for each respondent i a dummy variable, $child_i(f)$ equal to 1 if respondent i claims both ancestry from f and from

Europe (i.e. was conceived from the union between parents or more distant ancestors from f and Europe).

- From "ancestry1d", "sex", "sploc", we compute the share p_f of marriages between a spouse with ancestry from f and a spouse with European ancestry that are of the form European wife, Non-European spouse.

- Using "perwt" (sample weights) and the above constructed variables, we define for each country f (or for the group of Arab countries), the following share of respondents with ancestry from f who may potentially have a European surname (either a woman adopting her spouse's name upon her marriage, or children emanating from such marriages),

$$s_f = \frac{\sum_{i=1}^{N} perwt_i \cdot \max\left(wife_i(f), child_i(f) \times (1 - p_f)\right)}{\sum_{j=1}^{N} perwt_i \cdot ancestry_i(f)}.$$

- We remove, from the donations originating from county d and going to country f made by donors we identify as having European ancestry, donations where the donor may have ancestry from f but have been mis-classified as having European ancestry,

$$\#\text{Donations Corrected}^t_{d,f} = \#\text{Donations}^t_{d,f} - \frac{s_f}{1 - s_f} \times \#\text{Donations from } d\text{-ancestry donor}^t_{d,f},$$

where '#Donations from d-ancestry donor$^t_{d,f}$' are all donations made from d to f at time t by donors whom we classify as having ancestry from f.

B. *Details on other demographic data*

We source county-level population and population density from IPUMS. Our data on average age, racial composition, average household income, and educational attainment is drawn from the 2018 round of the American Community Survey (U.S. Census Bureau, 2018*a*,*b*). Our county-level measures of poverty are provided by the US Census Bureau under the 2018 Small Area Income and Poverty Estimates (SAIPE; U.S. Census Bureau (2018*c*)) programs. Our data on unemployment is from the US Bureau of Labor Statistics' 2019 Local Area Unemployment Statistics (LAUS; Bureau of Labor Statistics (2018)). Additional data on the age distribution and rurality comes from the 2000 Census (U.S. Census Bureau, 2000*a*,*b*,*c*, 2010; Manson et al., 2016).

C. *Details on geographic data*

We use data on countries, counties and postal codes to geocode the donations dataset (Duncalfe, 2020; U.S. Department of Agriculture, 2012; Simple Maps, 2017; Goodtables.io, 2020). We additionally use data on postal codes and regions to geolocate survey data (Din and Wilson, 2020). To account for changing spatial units over time we use a set of transition

Appendix Table B1—Description of each IPUMS wave

Wave	Description
1880	We use the 10% sample with oversamples; the sample is weighted, so we use the provided person weights to get to a representative sample; we use the region identifiers statefip and county.
1900	We use the 5% sample; the sample is weighted, so we use the provided person weights to get to a representative sample; we use the region identifiers statefip and county.
1910	We use the 1% sample; the sample is unweighted; we use the region identifiers statefip and county.
1920	We use the 1% sample; the sample is weighted, so we use the provided person weights to get to a representative sample; we use the region identifiers statefip and county.
1930	We use the 5% sample; the sample is weighted, so we use the provided person weights to get to a representative sample; we use the region identifiers statefip and county.
1970	We use the 1% Form 1 Metro sample; the sample is unweighted; we use the region identifiers statefip and cntygp97 (county group 1970); note that only four states can be completely identified because metropolitan areas that straddle state boundaries are not assigned to states; identifies every metropolitan area of 250,000 or more.
1980	We use the 5% State sample; the sample is unweighted; we use the region identifiers statefip and cntygp98 (county group 1980); the sample identifies all states, larger metropolitan areas, and most counties over 100,000 population.
1990	We use the 5% State sample; the sample is weighted, so we use the provided person weights to get to a representative sample; we use the region identifiers statefip and puma; the sample identifies all states, and within states, most counties or parts of counties with 100,000 or more population.
2000	We use the 5% Census sample; the sample is weighted, so we use the provided person weights to get to a representative sample; we use region identifiers statefip and puma; the sample identifies all states, and within states, most counties or parts of counties with 100,000 or more population.
2010	We use the American Community Service (ACS) 5-Year sample; the sample is weighted, so we use the provided person weights to get to a representative sample; we use region identifiers statefip and puma, which contain at least 100,000 persons; the 2006-2010 data contains all households and persons from the 1% ACS samples for 2006, 2007, 2008, 2009 and 2010, identifiable by year.

matrices. The produce maps we use shapefiles from International Union of Architects (Belgiu, 2015) and Walker (2021). To group states into regions we use definitions from U.S. Census Bureau (2021).

We compute the distance between foreign country f and a US county d, $Distance_{f,d}$, as the great circle distance between the county and country centroids, measured in kilometers. The latitude difference between a foreign country f and a US county d, $LatitudeDifference_{f,d}$, is the absolute difference between the latitudes of the two, measured in degrees.[32] References to distance as a control include both distance and latitude difference.

[32]Geo-coordinates for counties and countries are sourced from `www.geonames.org` and `www.cepii.fr` respectively, with a county's latitude and longitude as the average of that of all postal codes within the county, and a country's latitude and longitude as that of the largest city within the country.

III. NamSor Classification

A. *Validation*

We are not aware of any published attempts to validate NamSor's algorithm matching names to countries of origin, though research examining the accuracy of NamSor's gender-matching algorithm (Van Buskirk, Clauset and Larremore, 2022) and NamSor's Census designation (Asian, Black, non-Latino, Hispanic Latino, white non-Latino) algorithm (Krishnan, Singer and Zhang, 2021) has found these algorithms to be highly accurate.

We conduct an additional validation using a random 250,000 person sample from the North Carolina Voter Registration Data (Sood, 2020), which contains registrants' first and last names alongside self-reported ethnicity (Asian, African American, American Indian or Alaskan Native, Two or More Races, Other, Native Hawaiian or Pacific Islander, Undesignated, and White). Given that we use this classification exercise to exclude donors with ancestry from non-European countries, we are primarily concerned with classification errors of the type: Reports Asian/Native Hawaiian or Pacific Islander/Other, Classified as European. We find that this error occurs for fewer than one percent (2,322 of 250,000) of cases, suggesting that any bias induced by erroneously including these donors is negligible.

B. *Data Privacy*

Privacy for individual microdata was maintained at all stages of the data process, with no organization receiving more information than necessary. A 3-way Non Disclosure Agreement was signed by relevant parties to ensure that the following data privacy procedure was adhered to:

1) The charitable organization sends the research team the donation data, stripped of identifying information including names and addresses, with each donation containing a unique anonymized identifier (ID)

2) The charitable organization sends the third party NamSor a list containing *only* the ID of the donations and the name associated with each donation

3) Based on these names, NamSor determines the most likely origin country of the name

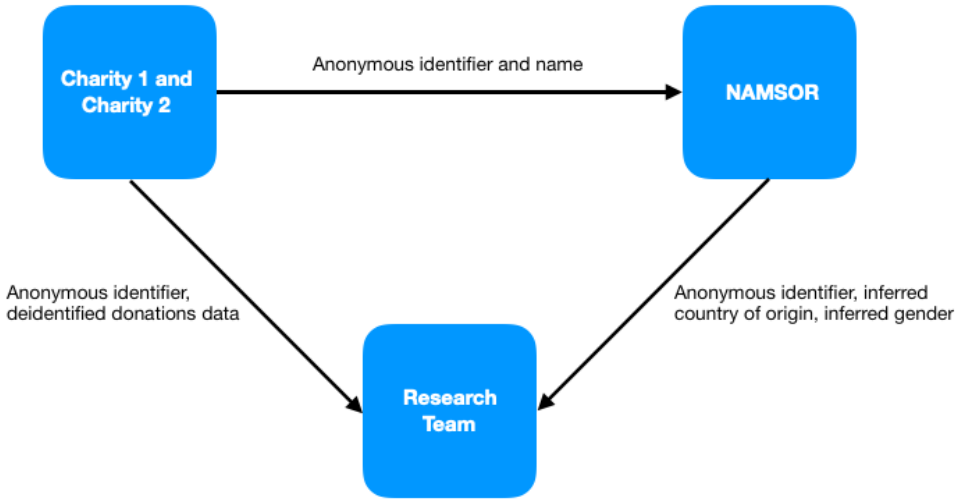
4) NamSor sends the research team a list containing *only* the ID of the donations and the origin country associated with each donation

5) The research team uses the donation ID to match up the donation data from the charitable organization and the origin country data from NamSor

A summary of the process is displayed below in Appendix Figure C1.

In this way, the organizations only receive the information that they need, and no more. The charitable organization does not receive NamSor data regarding origin countries for donor

names, NamSor does not receive any variables regarding donations except for the donor's name, and the research team does not receive any personally identifying information for any donation. Finally, data was shared using a number of secured Dropbox folders only shared with the intended recipients of the data.



Appendix Figure C1. Data Flow for Privacy

IV. Contact Survey Questionnaire

Demographics

Please indicate your gender.

- ○ Male
- ○ Female
- ○ Other/prefer not to answer

In what year were you born?

Were you born in the US?

- ○ Yes
- ○ No

What was your family's gross household income in 2019 in US dollars?

Do you have any children?

- ○ Yes
- ○ No

How many people are in your household?

Which of the following best describes your race or ethnicity?

- African American/Black
- Asian/Asian American
- Caucasian/White
- Native American, Inuit or Aleut
- Native Hawaiian/Pacific Islander
- Other

Are you of Hispanic, Latino, or Spanish origin?

- Yes
- No

Are you of Arab or Middle Eastern origin?

- Yes
- No

Which category best describes the highest level of education you have completed?

- 12th grade or less, but no high school diploma
- Graduated high school or equivalent
- Some college, no degree
- Associate degree
- Bachelor's degree
- Post-graduate degree

Are you married or in a long-term domestic partnership?

- ○ Yes
- ○ No

In general, how would you describe your physical health?

- ○ Excellent
- ○ Very good
- ○ Good
- ○ Only fair
- ○ Poor

What is your present religion, if any?

▾

County

What is the FIPS code of your current county of residence? If you are unsure, here is one way to look up your FIPS code:

1. Enter your address into https://www.whatcountyamiin.com/ to find your county name
2. Use your state name and the county name to look up the FIPS code on this page: https://www.nrcs.usda.gov/wps/portal/nrcs/detail/ma/home/?cid=nrcs143_01369

Your FIPS code will be a 5-digit number, possibly starting with 0. **Please note that your FIPS code is not your ZIP code!**

Please ensure that your FIPS code is correct. If it does not match your device location, we may be forced to terminate your survey.

For how many years have you lived in this county?

- Just moved in the last year
- 1-5 years
- 5-10 years
- 10-20 years
- 20-30 years
- 30+ years

Politics

In politics, as of today, do you consider yourself a Republican, a Democrat, or an Independent?

- Republican
- Democrat
- Independent

In politics, as of today, do you lean towards the Republican Party or lean towards the Democratic Party?

- The Republican Party
- The Democratic Party
- Do not lean toward either party

In politics, as of today, would you call yourself a strong Democrat or not a very strong Democrat?

- Strong
- Not very strong

In politics, as of today, would you call yourself a strong Republican or not a very strong Republican?

- ○ Strong
- ○ Not very strong

Who did you vote for in the 2012 Presidential election?

- ○ Mitt Romney
- ○ Barack Obama
- ○ Other
- ○ I did not vote

Who did you vote for in the 2016 Presidential election?

- ○ Donald Trump
- ○ Hillary Clinton
- ○ Other
- ○ I did not vote

Who did you vote for in the 2020 Presidential election?

- ○ Donald Trump
- ○ Joe Biden
- ○ Other
- ○ I did not vote

So far as you and your family are concerned, how worried are you about your current financial situations?

- ○ Extremely worried
- ○ Very worried

- ○ Moderately worried
- ○ A little worried
- ○ Not at all worried

Which of the following networks do you watch at least once a week? If you watch multiple networks, please choose the one you watch most often.

- ○ Fox News
- ○ CNN
- ○ MSNBC
- ○ None of the above

Contact

We would now like to ask about your close friends and family members, neighbors, workplace acquaintances, and others with whom you regularly interact (i.e. speak with at least once a month).

For each of the groups below, please check the box if a member of that group is among each group.

	Close friends and family members	Neighbors	Workplace acquaintances	Others with whom I regularly interact	Service or hospitality workers
African-Americans	☐	☐	☐	☐	☐
Arabs and/or Muslims	☐	☐	☐	☐	☐

Knowledge

We'd now like to ask you some questions about various religions.

What is Ramadan?

- ○ Hindu festival of lights
- ○ Jewish prayer for the dead
- ○ An Islamic holy month
- ○ Festival celebrating Buddha's birth

Which text is most closely associated with Hinduism?

- ○ Tao Te Ching
- ○ Vedas
- ○ Quran
- ○ Mahayana sutras

Which of the following are among the Five Pillars of Islam?

(You can select multiple options.)

- ☐ Fasting (sawm)
- ☐ Profession of faith (shahada)
- ☐ Charity to community members in need (zakat)
- ☐ Maintaining physical and mental health (sahi)
- ☐ Holy war against non-believers (jihad)
- ☐ Pilgrimage (hajj)
- ☐ Subservience of women and children to men (alnisa)

What percentage of the US population is Muslim? Please write your answer as a number, with 0 meaning that none of the US population is Muslim and 100 meaning that the entire US population is Muslim.

Restaurant

Have you ever eaten at a Middle Eastern restaurant? (For example, Iranian/Persian, Turkish, Egyptian, or Afghani restaurants)

- ○ Yes
- ○ No

End

Thank you for participating in our survey!