Online Appendix for: Universalism and Political Representation: Evidence from the Field*

Benjamin Enke Raymond Fisman Luis Mota Freitas
Steven Sun

September 25, 2023

*Enke: Harvard University and NBER; enke@fas.harvard.edu. Fisman: Boston University; rfisman@bu.edu. Freitas: Northwestern University and Global Priorities Institute; luismotafreitas@u.northwestern.edu. Sun: Parthenon Capital.

A Additional Figures

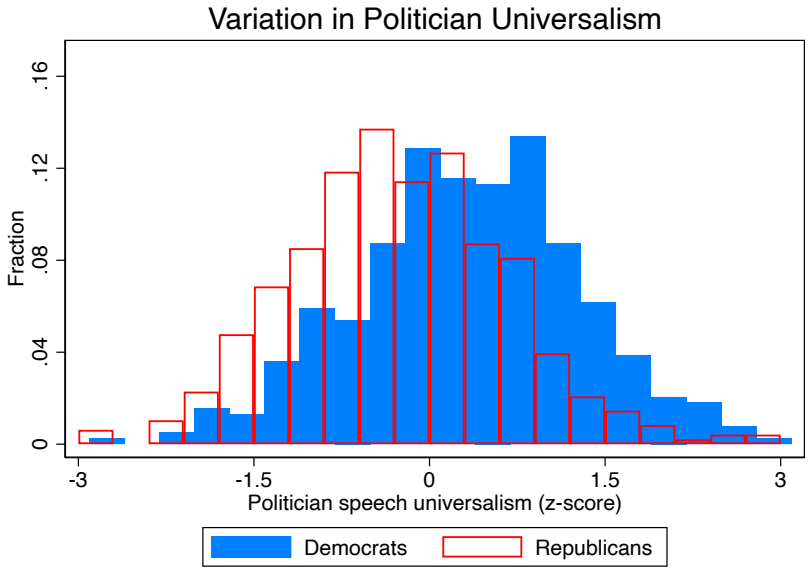

Figure A.1: Variation in politician universalism in speeches in the House of Representatives. Data are winsorized at $+/-3$ for readability.

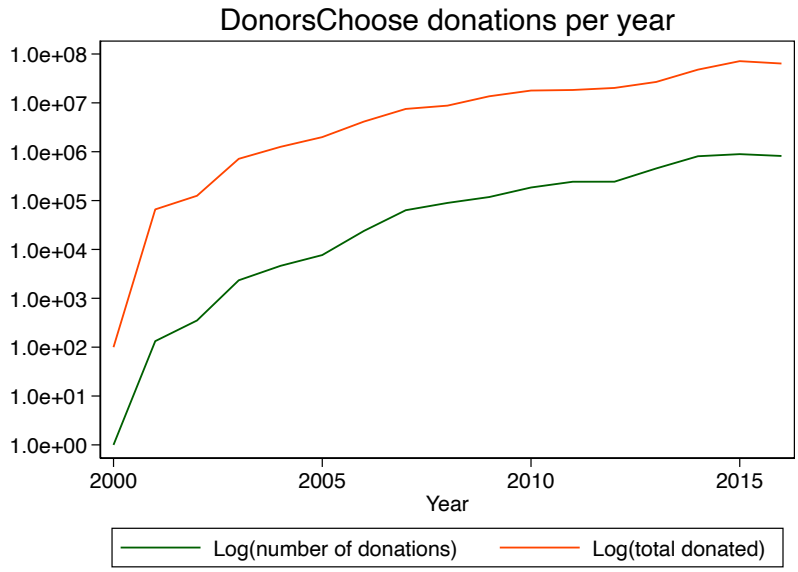

Figure A.2: Donations through the *DonorsChoose* platform per year.

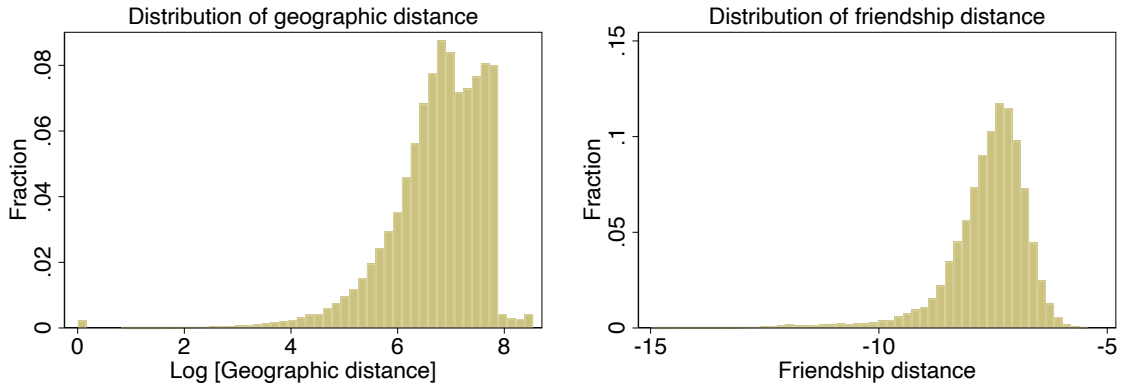

Figure A.3: Histogram of the distance between congressional districts. The left panel shows the distribution of log geographic distance, and the right hand panel shows the distribution of friendship distance.

B Additional Tables

Table B.1: Robustness checks, part 1

	Dependent variable:							
	Dem candidates vote share		Legislator DW-NOMINATE score			Speech universalism		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
District universalism (geographic distance)	10.5***	6.47***	-0.20***	-0.071***	-0.016*	0.22***	0.19***	0.15***
	(0.86)	(1.04)	(0.02)	(0.02)	(0.01)	(0.04)	(0.05)	(0.05)
Panel B								
District universalism (friendship distance)	9.96***	3.94***	-0.20***	-0.050**	-0.016**	0.21***	0.15**	0.16**
	(0.92)	(1.06)	(0.02)	(0.02)	(0.01)	(0.04)	(0.07)	(0.07)
Panel C								
District universalism (binarized geo. distance)	10.0***	5.92***	-0.19***	-0.077***	-0.025***	0.15**	0.15**	0.11
	(0.82)	(1.32)	(0.02)	(0.02)	(0.01)	(0.06)	(0.07)	(0.07)
Panel D								
District universalism (geo., excl. same state)	9.13***	5.92***	-0.17***	-0.086***	-0.014*	0.20***	0.15***	0.10*
	(0.95)	(1.01)	(0.02)	(0.02)	(0.01)	(0.05)	(0.05)	(0.05)
Congress FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Base controls	No	Yes	No	Yes	Yes	No	Yes	Yes
Democrat dummy	No	No	No	No	Yes	No	No	Yes
Observations	870	870	870	870	870	870	870	870

Notes. Columns (1)–(5) are OLS estimates and columns (6)–(8) WLS estimates, weighted by word count. Robust standard errors (clustered at the district level) in parentheses. Dependent variables are for the 113th and 114th Congresses. Each panel corresponds to a different regression. In panel A, the independent variable is district universalism computed based on geographic distance and in panel B it is computed based on friendship distance. In panel C, geographic distance is binarized based on a cutoff of 50 miles (80 km). Panel D restricts attention to geographic-based universalism using out-of-state donations. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.2: Robustness checks, part 2

	Dependent variable:							
	Dem candidates vote share		Legislator DW-NOMINATE score			Speech universalism		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
District universalism (high poverty donations sample)	11.1***	6.08***	-0.21***	-0.068***	-0.018**	0.22***	0.21***	0.19***
	(0.89)	(1.14)	(0.02)	(0.02)	(0.01)	(0.04)	(0.06)	(0.06)
Panel B								
District universalism (pre-114th Congress donations sample)	10.2***	4.49***	-0.20***	-0.041*	-0.011	0.25***	0.13**	0.11*
	(1.03)	(1.20)	(0.02)	(0.02)	(0.01)	(0.06)	(0.07)	(0.06)
Panel C								
District universalism (no matched donations sample)	10.7***	5.50***	-0.21***	-0.062***	-0.017**	0.22***	0.19***	0.17***
	(0.91)	(1.12)	(0.02)	(0.02)	(0.01)	(0.04)	(0.06)	(0.06)
Congress FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Base controls	No	Yes	No	Yes	Yes	No	Yes	Yes
Democrat dummy	No	No	No	No	Yes	No	No	Yes

Notes. Columns (1)–(5) are OLS estimates and columns (6)–(8) WLS estimates, weighted by word count. Robust standard errors (clustered at the district level) in parentheses. Dependent variables are for the 113th and 114th Congresses, except for Panel B. Each panel corresponds to a different regression. Panel A restricts attention high poverty schools. Panel B restricts attention to donations made before the 114th Congress, and dependent variables for the 114th Congress. Panel C excludes donations that have matching funds available. The number of observations is 870 for Panels A and C, and 435 for Panel B. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C Background on DonorsChoose

C.1 Visual layout and functionality of the DonorsChoose platform

We ensure our results are not artifacts of the layout or functionality of the DonorsChoose website. To do so, we examined all available screenshots of the platform's layout and functionality since its inception.

Throughout the relevant time period, it is *not* the case that projects are sorted by closest proximity to each donor on the website. Instead, for a significant portion of our sample period, the default sort for projects on the platform was by urgency, which DonorsChoose defines as a combination of the lowest cost to complete, highest economic need, and fewest days left to expiration of the project.

The website's layout also does not vary across space. That is, to the best of our knowledge, at any given time all donors observe the same platform layout regardless of location, and given the default sort, they observe exactly the same projects when they first arrive at the platform. Below, we present a screenshot of the DonorsChoose platform as accessed in June 2019. (The reader may note that one of the projects includes matching funds from Google.org. Approximately 10 percent of listings include such matches; our results are virtually identical if we omit these listings; see Panel C of Table B.2.)

Throughout our sample period, the options available to filter and sort projects were constant. Most importantly, the ability to search through and filter projects based on location was and continues to be a salient (usually the highest) option available on the screen. This feature makes a donor's selection of a project based on geography particularly straightforward, and potentially enhances the case for our claim that geographic distance is a relevant metric employed by donors in selecting projects.

C.2 Additional Notes on Methodology

Data Cleaning. Our raw data consist of 6,211,940 individual donations made between March 2000 and October 2016. Beginning in 2007, donations are made to projects in all states in the United States plus the District of Columbia.

In addition to dropping observations with missing geographic or donation data, we exclude donations in which either the donor or the recipient school is located outside of the 50 states and the District of Columbia.

Aggregation to Congressional District level. ZIP codes provided in the DonorsChoose data were used to map donors and projects to their respective congressional districts. Note that for reasons of anonymity, donor ZIP codes were truncated at the first three

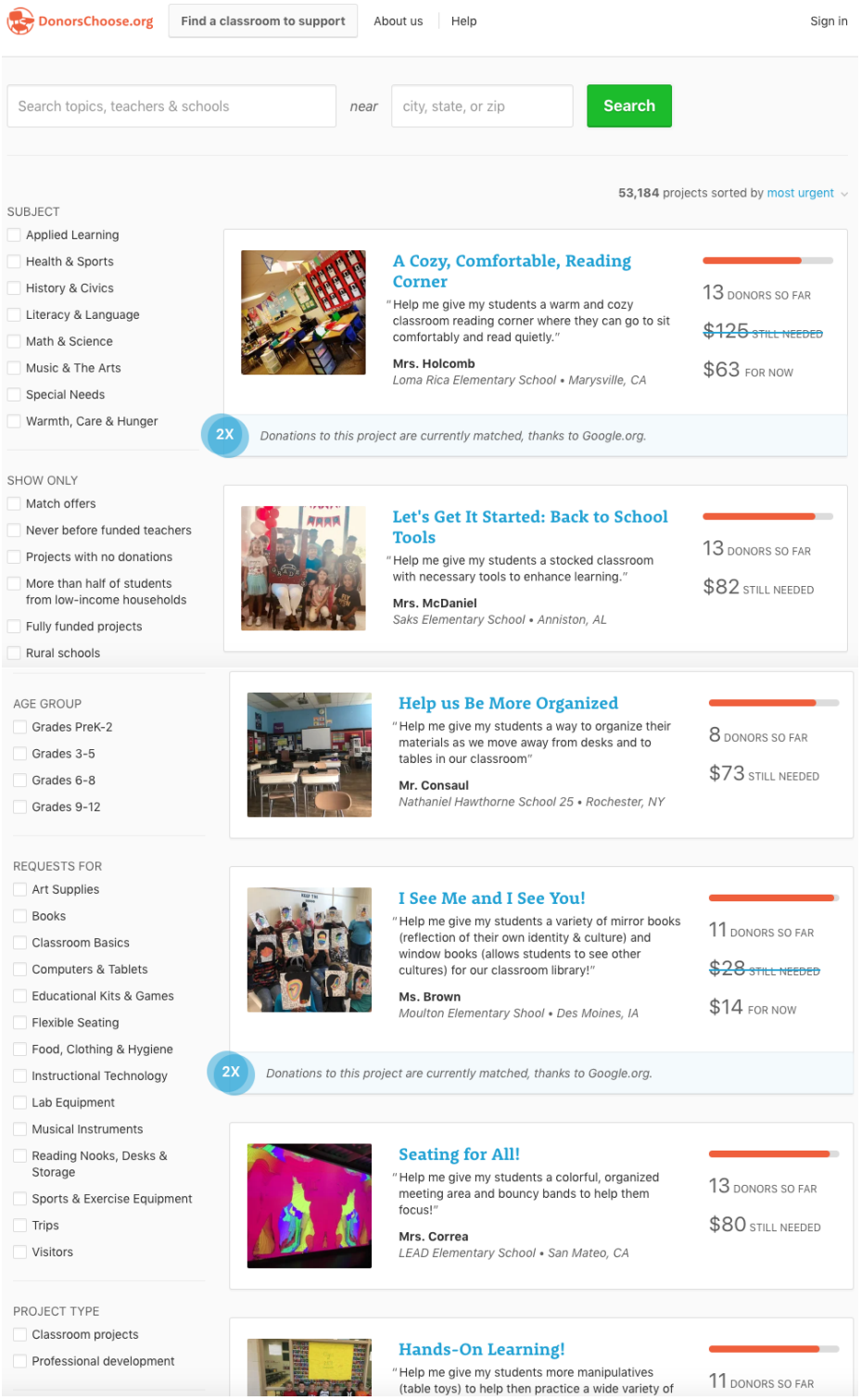

Figure C.1: Screenshot of DonorsChoose platform in June 2019. Note the ability to search for projects near any given geographical location at the top of the page, the options available to the donor with which to filter projects, and the "Double Your Impact" promotion applied to the topmost project presented. Additional options available with which to filter projects included the project's target age group, request type (e.g., art supplies, books, classroom basics, etc.), project type (classroom projects or professional development), and buckets for amount needed ($50 and under, $100 and under, etc.).

digits, which added a layer of uncertainty to congressional district (CD) mappings, beyond the usual fuzziness of ZIP-to-CD mappings. Thus, through data provided by the United States Census Bureau, every donation was first mapped to the area formed by all possible *full* ZIP codes corresponding to the truncated ZIP code from DonorsChoose, and then in turn to a given CD based on all congressional districts overlap with that area. Because this mapping is not one-to-one, when aggregating donations to relevant source CDs, all observations were weighted by the degree of a fuzzy match to relevant CDs. For example, if based on the provided ZIP code a donation could have originated from either MA-2 or MA-3, this donation would appear twice in our merged data once all donations were mapped to donor congressional districts. In turn, each of these two observations would then be weighted by the share of the 3-digit ZIP code area population in each of these congressional districts when aggregating donation statistics by pairs of donor and recipient CDs. An analogous aggregation procedure was used for other variables in our analysis which were not provided at the CD level.

C.2.1 Social Distance Data

Data on the social connectedness and the "relative probability of friendship" between pairs of counties in the United States was obtained from Facebook. The construction of these data is covered in Bailey et al. (2018). The Social Connectedness Index (SCI) reflects the aggregate number of Facebook friendship links within or between counties. The "relative probability of friendship" normalizes for county populations by dividing the SCI by the product of the number of Facebook users in each of the two counties.

We aggregate this "relative probability of friendship" data to the congressional district level by using the aggregation procedure described in Bailey et al. (2021). Since mappings from county to congressional district are not one-to-one, the aggregation from county to this geographic level accounts for the possibility of a fuzzy match, by weighting observations by the share of the county population in each possible congressional district that a given county could map to.

This aggregation from county-pair SCIs and relative probabilities of friendship forms our measure of "friendship distance.'. Specifically, we define the social distance between a donor in geographic entity i and a recipient in a geographic entity j of the same level as $-\ln(\text{rel. prob. of friendship}_{i,j})$.

D Most common eMFD words in Congressional speeches data

Table D.1: 20 most frequent eMFD words in congressional speeches

Ranking	Term	Frequency
1	people	114479
2	time	99601
3	president	99435
4	speaker	86166
5	going	64816
6	work	61930
7	states	60363
8	country	58531
9	want	57966
10	act	56262
11	senator	53195
12	know	51902
13	support	51841
14	house	51201
15	need	50814
16	state	50044
17	committee	48681
18	new	48471
19	government	48054
20	think	47174

E Data sources

- DonorsChoose donations and projects: DonorsChoose (2016)[1]
- IRS individual donations: Internal Revenue Service (2022)
- IRS charity received donations: Urban Institute, National Center for Charitable Statistics (2023)
- Geographic location: U.S. Census Bureau (2022a)
- Relationship between geographic areas: U.S. Census Bureau (2022b)

[1]These datasets have been removed from the original source since our access date, but a snapshot of the download page at the time is available here: `https://web.archive.org/web/20181022125924/https://research.donorschoose.org/t/download-opendata/33`. We provide the files in our replication package.

- Social Connectedness Index: Data for Good at Meta (2021)
- House representative vote share: MIT Election Data and Science Lab (2017)
- President vote share: MIT Election Data and Science Lab (2018)
- House representative ideology: Lewis et al. (2023)
- Congressional Speeches: Gentzkow et al. (2018)
- eMFD: Hopp et al. (2021)
- Stop words: Porter and Boulton (2006)
- Population, median household income, share of population with college degree, White ethnic share: U.S. Census Bureau (2022d)
- Congressional District boundary location: U.S. Census Bureau (2022c)
- Coastline: Natural Earth (2018)

References

Bailey, Michael, Abhinav Gupta, Sebastian Hillenbrand, Theresa Kuchler, Robert Richmond, and Johannes Stroebel, "International Trade and Social Connectedness," *Journal of International Economics*, March 2021, *129*, 103418.

_ , Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong, "Social Connectedness: Measurement, Determinants, and Effects," *Journal of Economic Perspectives*, 2018, *32* (3), 259–280.

Data for Good at Meta, "Facebook Social Connectedness Index," 2021. https://data.humdata.org/dataset/social-connectedness-index (accessed Nov 16, 2022).

DonorsChoose, "DonorsChoose Project and Donation Data," 2016. https://research.donorschoose.org/t/download-opendata/33 (accessed Sep 5, 2018).

Gentzkow, Matthew, Jesse M. Shapiro, and Matt Taddy, "Congressional Record for the 43rd-114th Congresses: Parsed Speeches and Phrase Counts," 2018. https://data.stanford.edu/congress_text (accessed Dec 3, 2022).

Hopp, Frederic R, Jacob T Fisher, Devin Cornell, Richard Huskey, and René Weber, "The extended Moral Foundations Dictionary (eMFD): Development and applications of a crowd-sourced approach to extracting moral intuitions from text," *Behavior research methods*, 2021, *53*, 232–246.

Internal Revenue Service, "Individual Income Tax Statistics - ZIP Code Data," 2022. https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-statistics-zip-code-data-soi (accessed Aug 19, 2023).

Lewis, Jeffrey B., Keith Poole, Howard Rosenthal, Adam Boche, Aaron Rudkin, and Luke Sonnet, "Voteview: Congressional Roll-Call Votes Database," 2023. https://voteview.com/ (accessed Jan 30, 2023).

MIT Election Data and Science Lab, "U.S. House 1976-2022," 2017. https://doi.org/10.7910/DVN/IG0UN2 (accessed Feb 2, 2023).

_ , "County Presidential Election Returns 2000-2020," 2018. https://doi.org/10.7910/DVN/VOQCHQ (accessed Aug 7, 2023).

Natural Earth, "Coastline," 2018. https://www.naturalearthdata.com/downloads/10m-physical-vectors/10m-coastline/ (accessed May 20, 2018).

Porter, Martin and Richard Boulton, "Stop words," 2006. http://snowball.tartarus.org/algorithms/english/stop.txt (accessed Dec 5, 2022).

Urban Institute, National Center for Charitable Statistics, "Core File for Public Charities," 2023. http://nccs-data.urban.org (accessed Aug 20, 2023).

U.S. Census Bureau, "Gazetteer files," 2022. https://www.census.gov/geographies/reference-files/time-series/geo/gazetteer-files.2017.html (accessed Oct 31, 2018).

_ , "Relationship files," 2022. https://www.census.gov/geographies/reference-files/time-series/geo/relationship-files.2010.html (accessed Nov 8, 2022).

_ , "TIGER/Line Shapefiles," 2022. https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2013.html (accessed Feb 13, 2023).

_ , "US Census Bureau Data," 2022. https://data.census.gov/ (accessed Nov 23, 2022).