Appendix A:
Geographic Regression Discontinuity

Data Sources

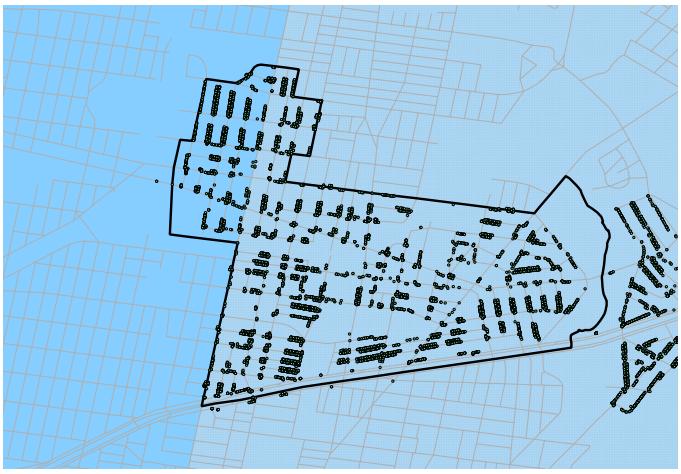
Data for this analysis come from three sources. First, the Pennsylvania Voter File, purchased from the Pennsylvania Secretary of State's Office, provides both the home address and voter history for all registered voters in Pennsylvania. This will be used to determine individual level voter turnout in the 2014 Pennsylvania Democratic Primary. Second, the U.S. Census Bureau provides geographic shapefile data for state, county, congressional district, and state legislative district boundaries. Third, the Data Science Toolkit provides geocoding services which allow me to determine the latitude and longitude coordinates of voters from their addresses. These data sources together will allow me to calculate the distance of individual voters from political borders – in this case the distance between a voter and the border of their congressional district. Finally, over 50 hours of elite interviews were conducted with the candidates, campaign staff, interest groups, and local activists during the campaign season to better understand the underlying dynamics of the race.

Compound Treatment Reduction

In geographic regression discontinuity models, researchers are often presented with situations in which more than one geographic 'treatment' affects the outcome of interest at the same time. In the primary in question, voters were exposed to primary campaigns for governor, congress, state senate, and state legislature. Election law in the United States is often administered at the county level, and exposure to campaign advertisements is often confined to particular media markets which are unique combinations of counties. In many of these instances the borders between these politically salient districts are the same. As Keele and Titiunik describe, this "poses a serious challenge if the researcher is interested in only one of those treatments since, absent any restrictions or assumptions, it will not be possible to separate the effect of the treatment of interest on the outcome from the effect of all other 'irrelevant' treatments" (2015).

In order to eliminate the issue of compound treatments, this project isolates areas that are within the same county, state senate, and state legislative district, that also contain portions both within and outside of the 13th Congressional district. Fortunately all of the 13th district falls within the Philadelphia media market, and is therefore held constant. By holding all other relevant political boundaries constant, we are essentially controlling for the effects of these alternative boundaries. I am able to isolate 16 regions along the border of the 13th district that meet these characteristics: six in Philadelphia and ten in Montgomery County. This provides us with 68,021 registered Democrats in Philadelphia, respectively, and 113,184 in Montgomery County. Figure A1 provides an example of a region within Philadelphia County that is within the 3rd State Senate district and the 179th State House district. The light blue area represents areas within this geographic subset that is within the 13th Congressional district, while the darker blue is within the 2nd Congressional district, which importantly did not have a contested Democratic primary. Each small dot represents a household with a registered voter.

Figure A1: Example of Compound Treatment Reduction



Measuring Geographic Distance

In order to measure the distance between each voter and the discontinuity (the congressional district boundary), I first geocoded the registered address of the voters that fell within these compound treatment reduction geographies. Less than 0.5% of voters were dropped because of incomplete addresses. Next, I converted the census provided shapefile of the district into a geometric polygon. From here I use the `geosphere` package in `R` to calculate the shortest geographic distance between each voter and the congressional district border. This package allows one to take into consideration the ellipsoidal shape of the earth when calculating geographic distance for greater accuracy.

Measuring Turnout

Voter turnout is measured as the percentage of registered voters recorded as voting either at the polling location or absentee. Because Pennsylvania is a closed primary state, turnout in the primary is measured only among registered Democrats.

Results

Table A1 provides the results of the geoRDD model for turnout in the primary. All models are estimated with the `rdrobust` package. The results present both the initial estimates and the estimates with fixed-effects for each compound reduction geography. The effective number of observations on each side of the threshold are also provided.

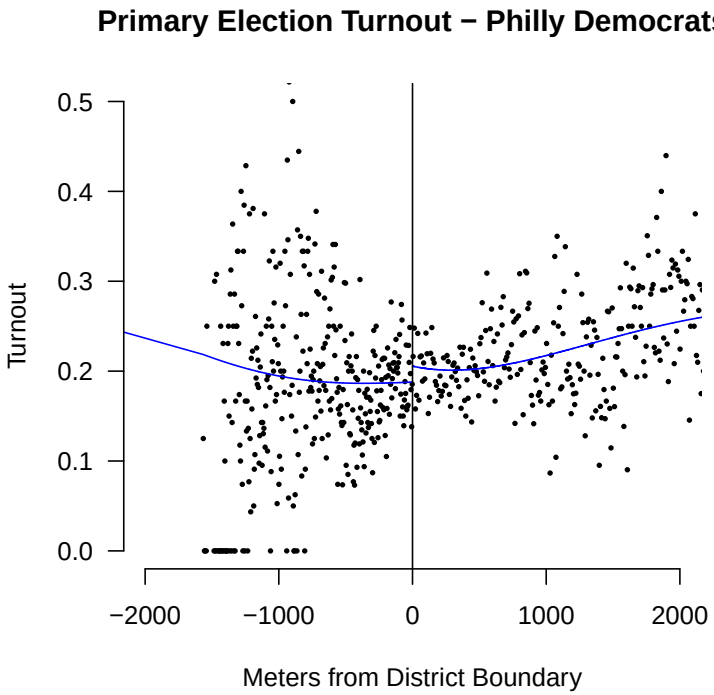
What is immediately apparent is the difference in outcomes between the counties. In Philadelphia there is roughly a 3% increase in turnout in the presence of a ground game operation. While 3% may seem modest given the race was eventually decided by an over 10 point margin, given that turnout in these sections was only 21%, that constitutes a 14%

change in turnout. No such effect is observed in Montgomery County, where no discernible ground game occurred. Figure A2 provides this trend graphically.

Table A1: Primary Election Turnout of Democrats

County	**Model**	**Estimate**	**Std. Err.**	**p-value**	**C.I.**	**F.E.**
Philly	Conventional	0.023*	0.010	0.026	0.003, 0.042	X
	Bias-Corrected	0.028**	0.010	0.006	0.008, 0.047	X
	Robust	0.028*	0.011	0.016	0.005, 0.050	X
	Conventional	0.032**	0.011	0.003	0.011, 0.053	✓
	Bias-Corrected	0.037***	0.011	0.001	0.016, 0.058	✓
	Robust	0.037**	0.012	0.002	0.014, 0.060	✓
Eff. Obs.	10,498 — 10,438					
MontCo	Conventional	0.001	0.008	0.900	-0.014, 0.016	X
	Bias-Corrected	0.005	0.008	0.563	-0.011, 0.020	X
	Robust	0.005	0.009	0.604	-0.013, 0.021	X
	Conventional	-0.003	0.013	0.797	-0.020, 0.022	✓
	Bias-Corrected	0.001	0.013	0.955	-0.025, 0.026	✓
	Robust	0.001	0.013	0.955	-0.025, 0.027	✓
Eff. Obs.	16,860 — 45,897					

Figure A2: Primary Election Turnout of Philadelphia Democrats



Appendix B:
Network Threshold Reduction

The below figure provides the existing network density for the three candidates competing in a 2014 open seat contest for Alabama's 6th district Republican Nomination. Moving left to right, the networks for Paul DeMarco, Chad Mathis, and Gary Palmer are presented with a threshold of 1, 2, and 3 donations. What is important to note is that as the threshold increases the networks become less dense, but the relative densities between the networks stays about the same. I did this exercise for 10 races in the 2014 cycle and found similar trends. However, in the regression analysis, a threshold of 1 donor is used consistently.

Figure A3: Network Density Threshold Sensitivity