Customs and Border Protection (CBP) Activities Mobilize Hispanic Voters

Cory Smith (MIT) ⓡ Donghee Jo (Northeastern), David Lazer (Northeastern)*

July 16, 2020

Abstract

Do activities by immigration enforcement agencies suppress or mobilize Hispanic voters? To answer this question, we exploit the sharp discontinuity in the legal authority of United States Customs and Border Protection (CBP) at the boundary of the 100-mile interior US border zone. We find that CBP activities increased Hispanic voter registration and turnout in the 2016 US general election by 1.8pp and 1.5pp, respectively. We suggest that the main mechanism is via personal experiences—observing or hearing about CBP activities—as opposed to elite-driven campaigns that are unlikely to be spatially discontinuous. We also estimate the electoral consequences of CBP activities through a simulation exercise and find a small increase in the Democratic Party's representation in the US Congress. *JEL Codes: D7, J15*.

*Circled "r" sign implies that the first author is determined by a coin flip between Donghee and Cory, a convention suggested in Ray and Robson (2018). We thank Daron Acemoglu, Sam Bazzi, Enrico Cantoni, Matt Lowe, and Vincent Pons for helpful comments. We are grateful for excellent research assistance by Hanyu Chwe. All errors are our own.

1 Introduction

In 2016, then-candidate Donald Trump brought the issue of immigration to the forefront of his campaign, with his famous slogan, “Build the wall.” In surveys conducted in 2016-2017, immigration was one of the most important (Pew Research Center, 2016) and polarizing (Pew Research Center, 2017) issues. Consistent with President Trump’s campaign promises, interior (“on-US-soil”) arrests by Immigration and Customs Enforcement (ICE) increased by 21% during the first year of his presidency (TRAC, 2018). The Hispanic population in the US has felt the impact of immigration enforcement most strongly. More than 90% of the people arrested by ICE within the US between November 2014 and May 2018 have been Latin Americans (TRAC, 2018).

Despite popular perceptions to the contrary, CBP agents are also frequently involved in interior operations, often far inside the border. Their legal grounds for such activities are especially strong within the 100-mile interior US border zone—the area within 100 miles of a US land or water border. It is only in this area that CBP agents may board vessels and vehicles to search for aliens, at which point they may question people under reasonable suspicions, even without probable cause. Based on this legal authority, CBP agents operate temporary and permanent checkpoints in the border zone and board buses and trains unannounced. At the interior boundary of this border zone (“the 100-mile boundary” hereafter), there is a sharp discontinuity in CBP’s legal authority. We exploit this geographic discontinuity to evaluate the impact of immigration enforcement on Hispanic voter turnout.

Theoretically, immigration enforcement activities have the potential to both suppress and mobilize Hispanic voters. On the one hand, CBP activities can increase the cost of registering and voting. On the other hand, voters’ direct and indirect experiences with CBP can motivate their political participation, especially when one of the main issues in the election is immigration enforcement.

In this paper, we empirically study the impact of interior immigration enforcement on Hispanic political participation in the 2016 US general election using RDD. We also suggest potential mechanisms of the political participation effect and simulate net electoral consequences.

The main finding of this paper is the mobilizing effect of CBP activities on Hispanic political participation. A discontinuous 1.8 percentage point increase in voter registration ($p < 0.01$) and a 1.5 percentage point increase in voter turnout ($p < 0.1$) were found among voting-eligible Hispanics residing just inside the 100-mile boundary compared with voting-eligible Hispanics residing just outside that area. This does not seem to be driven by population sorting or sample selection—no discontinuous jump was found on the total Hispanic voting-age population or their characteristics at the 100-mile boundary.

We also find a discontinuous 1.3pp vote margin swing ($p < 0.05$) toward the Democratic Party in the 2016 election, under the assumption that registered partisans vote for their party if they turn out to vote. Extrapolating this partisan effect to the entire border zone and to the last ten elections, a simulation exercise reveals that this vote swing amounts to a modest–but not trivial–flip from Republicans to Democrats in two congressional elections in 2010-2018 due to CBP's mobilizing effect compared with a counterfactual scenario where there is no CBP effect.

The spatial nature of the treatment effect suggests that the main mechanism is through the experience of self, family members, and close neighbors rather than through top-down campaigns, such as through the mass media, political rallies, or (not-spatially-assembled) online communities. Consistent with this interpretation, we do not find any evidence of discontinuous political gathering (rallies, protests, demonstrations, etc.) frequencies at the 100-mile boundary, a likely mechanism if the effect is mostly caused by elite-driven campaigns. Furthermore, the mobilization effects are driven by Hispanic voters in more densely populated areas, places where CBP activities are more likely to be visible to the general public. Finally, it appears that a significant portion of the turnout effect is driven by those who changed their minds from not registering to registering, ordinarily less engaged voters. Given that they are less likely to be regular consumers of political information, personal experiences may have played a major role in mobilizing them.

Related Literature. This paper contributes to a broad literature that studies how various political and judicial institutions affect political participation of minority and underprivileged voters. Institutional obstacles, such as inconvenient voter registration processes (Braconnier et al., 2017), misdemeanor convictions (White, 2019), and increased distance to polling locations (Cantoni, 2020) are reported to have a disproportionally negative effect on political participation by minority voters and underprivileged populations. Interestingly, Cantoni and Pons (2019) find a mobilizing effect from strict ID laws on Hispanic voters in the US, while generally finding a null effect on any other group. Focusing specifically on immigration enforcement, White (2016) finds a mobilizing effect from the Secure Communities Program on Latino voters. We provide reinforcing evidence from more granular (voter level) data in a different setting with a spatial identification strategy, which naturally restricts the space of potential mechanisms.

The remainder of the paper proceeds as follows. Section 2 describes the background, data, and empirical strategy. Section 3 provides the main results of this paper. Discussions on potential mechanisms can be found in Section 4, and Section 5 concludes.

2 Background, Data, and Empirical Strategy

2.1 Background

Section 287 of the Immigration and Nationality Act of 1952 (Pub.L. 82–414, 66 Stat. 163, enacted June 27, 1952) specifies "Any officer or employee of the Service [...] shall have power without warrant *within a reasonable distance* from any external boundary of the United States, to board and search for aliens any vessel within the territorial waters of the United States and any railway car, aircraft, conveyance, or vehicle" [*emphasis added*]. In 1953, the Department of Justice issued a rule (§ 287.1) redefining a reasonable distance to be "100 air miles from any external boundary of the United States."[1] Accordingly, CBP operates temporary and permanent interior checkpoints in this 100-mile border zone to stop and search vehicles. Also, it has been reported that CBP agents frequently board buses and trains and operate facial recognition drones in this border zone (Cato Institute, 2018).

Immigration enforcement within this 100-mile border zone is a significant policy issue affecting approximately two-thirds of the US population residing in this area (CityLab, 2018). In addition to the southern border area and beyond, all major cities on the East and West Coasts, such as New York, Boston, Seattle, San Francisco, and Los Angeles, are within this border zone.[2]

Almost all non-airport CBP duty locations are strategically placed inside the 100-mile boundary, and several of them are just inside the boundary as shown in Figure A1. This suggests that CBP is well aware of this rule and the officers are strategically operating within this constraint. Although the Department of Homeland Security does not publish data on CBP enforcement activities,[3] there are episodes suggesting that CBP agents recognize this discontinuity. For example, CBP officers who boarded a Greyhound bus outside the border zone soon retreated after a woman on the bus asserted, "You have no right to ask me for anything [...] We are not within 100 miles of a border so [these agents] have no legal right or jurisdiction here!" (ACLU, 2018). Thus, although we have been unable to locate data detailing the exact proportion of citizens familiar with the rule, at least some are familiar with its specifics. Regardless of the extent of knowledge about this discontinuity, however, Hispanic voters inside the border zone are treated differently compared with those outside, and this difference can create the treatment effect that this paper

[1] Air mile distance means distance "as the crow flies," instead of the shortest Euclidean distance through under the Earth surface.

[2] Several lawmakers find this number of miles to be excessive and have recently introduced bills (S. 2180 and H.R. 3853) to limit this zone to 25 miles from the southern border and exclude northern and coastal border zones, though these bills have not made much progress so far.

[3] To date, our effort to get this data through the Freedom of Information Act has been unsuccessful.

identifies.

2.2 Data

Individual level voter data comes from TargetSmart, a private company specializing in voter data collection. The company collates information about 191 million voters primarily from state voter roll data. The information available in the dataset includes age, gender, race, address with geo-codes, congressional district, party, registration records, and voting records of each voter. It additionally collects information on 58 million "potential" voters using their proprietary methods. This potential voter data enables analysis—albeit imperfectly—on the general population of eligible voters, not just registered voters, which is an endogenously selected subset.[4]

Approximately two-thirds (20 million out of 32 million) of the Hispanic voting-age population in the US are accounted for in the TargetSmart data.[5] Since TargetSmart data is primarily collected for election consulting, almost all under-age populations are excluded. Further, anyone who was ineligible to vote in the 2016 election is excluded from the analyses in this paper. Table 1 provides summary statistics. TargetSmart's data shows that the Hispanic population living near the 100-mile boundary (Column 2) has very similar characteristics to the general Hispanic population in the US (Column 1). The sample used in Column 2—the Hispanic population living within 3 miles inside the 100-mile boundary—is the primary sample we use throughout our RDD analysis in this paper.

In addition to TargetSmart data, we use 2010 US census data to further demonstrate that there is no Hispanic population discontinuity at the 100-mile boundary. For simulation exercises for electoral consequences, we use MIT Election Lab data (MIT Election Data and Science Lab, 2019). For political gathering frequencies, we use a dataset by Crowd Counting Consortium (2020), where we extract and geo-code 8,814 gatherings held during February 1, 2017 - January 31, 2018, the earliest publicly available dates.

2.3 Empirical Strategy

Our empirical strategy is a single-dimensional RDD, where the running variable is the directed shortest distance from the 100-mile boundary (< 0 if inside the 100-mile border zone). The main

[4] To the best of our knowledge, the information collection about potential voters is conducted online and, therefore, is unlikely to be affected by the spatial discontinuity. This is supported by falisification tests showing no discontinuity on demographic characteristics at the 100-mile boundary (see Columns 4-7 of Table 2 and (d)-(g) of Figure 1).

[5] Classification as Hispanic comes from the voter files; TargetSmart then updates the classification using their commercial sources.

specification is as follows.

$$y_i = \alpha_c + \beta \cdot Treat_i + f(distance_i) + X_i + \varepsilon_i \quad (1)$$

The dependent variable of interest is y_i. α_c is county fixed effect. $Treat_i$ is an indicator of whether i is located within the 100-mile border zone, and $f(distance_i)$ is a function of RD running variable. We use linear approximation with different slopes below and above the cutoff. X_i includes age, race, gender, latitude, and longitude. A rectangular kernel with 3-mile bandwidth is used. Since we conduct RD analyses over space, we report spatial standard errors (Conley, 2010) allowing correlation in ε_i between any two people residing within 100km from each other as our preferred specification by default.[6]

The histogram of the running variable, directed distance from the 100-mile boundary, is shown in Figure A2. Although the histogram passes the eye-test, the McCrary test (McCrary, 2008) is also conducted by binning the Hispanic voting-age population within each county by distance from the 100-mile boundary and running an RD regression with county fixed effect as in Equation (1). As shown in Table A1, we fail to reject the null hypothesis of no density jump from this test using a range of bin sizes and bandwidths.[7]

3 Main Results

The first two columns of Table 2 provide the main results of this paper from specification (1). Geographic and demographic controls in Columns 1-3 in the table consist of age, gender, latitude, longitude, and county fixed effects. Political participation by Hispanics just inside the 100-mile boundary is greater than participation by those just outside. This implies that CBP activities have a mobilizing effect on Hispanic voters—registration is increase by 1.8pp (Column 1), and turnout is increased by 1.5pp (Column 2). Based on this, increased motivation for political change or expressive voting appears to outweigh the increase in various costs of political participation induced by CBP activities. This is in stark contrast to precise null effects found for whites shown in Table A2. The null results in Table A2 rule out the mobilization effects of CBP activities

[6] Alternative specifications with clustered standard errors are reported in the Appendix. Also in the Appendix is a figure showing the results' sensitivity to the bandwidth choice. Bandwidth for our main specification is chosen to give enough effective sample size without biasing the estimates. We are unaware of a bandwidth selection technique with spatial standard errors. The cluster-robust (clustered in congressional districts) MSE-minimizing bandwidths calculated for residualized outcome variables are 2-4 miles depending on the outcome variable, which is similar to our bandwidth of choice.

[7] McCrary (2008) suggests the bandwidth size be half of the bandwidth used for the main results—1.5 miles in our case. This corresponds to Columns 1, 4, and 7 of Table A1.

on all races, and this is consistent with our background knowledge that the main target of CBP activities has been Hispanics.

Residual RD plots of these outcome variables are shown in Figure 1. Both of our main results, registration and turnout, show clear discontinuity at the boundary. These results are robust to (i) a range of bandwidths (Figure A3), (ii) using cluster-robust errors instead of spatial (Table A3), and (iii) excluding the top 10 most-populated precincts inside and outside the border zone, respectively, which excludes approximately 10% of the total observations (Table A4).

Although turnout per se matters in the democratic process, it is useful to examine any potential partisan implications of the mobilization effect found above. We estimate a plausible partisanship effect by re-running our main regression on a variable equal to 1 if the voter turned out *and* is a registered Democrat, and equal to -1 if the voter turned out *and* is a registered Republican. The result is reported in Column 3 of Table 2—the Democratic Party vote share margin compared to the Republican Party increased by 1.3pp at the discontinuity. To contextualize the effect, we consider US House of Representatives elections in relevant sample states from 2010 to 2018. Sampling the differential partisan turnout effect according to the t-distribution of the resulting estimate, on average the Democratic Party gains two house seats across these five midterm elections relative to the counterfactual scenario without the 100-mile border enforcement.[8]

4 Discussions

No Population Sorting. It is theoretically possible that the result is driven by population sorting, where Hispanic voters, who are likely to be voting selectively, migrate into the border zone. However, as can be seen in Columns 4-8 of Table 2 and (d)-(g) of Figure 1,[9] we detect no difference in any demographic characteristic reported in the voter files: age, gender, race, party registration, or partisan leaning as computed by TargetSmart. Furthermore, the 2010 census data provides reinforcing evidence that there was no Hispanic population discontinuity at the boundary—see Table 3.[10] This lack of evidence for population sorting also can be considered as a falsification test for the RDD.

[8]Note that this exercise essentially precludes any effect on many congressional districts with already unequal support to one of the parties. Changes in vote margins without election result changes can still have political consequences.

[9]Panel (h) of Figure 1 shows a sign of potential imbalance, but we fail to reject no discontinuity in our regression specification (Column 8 of Table 2). The result from regression is also consistent with the results from the 2020 census data (Table 3).

[10]We use centroid of each census block for bandwidth kernel calculations. Census blocks that straddle the 100-mile boundary are dropped from the analysis.

Mechanism. Through what mechanism do CBP activities mobilize Hispanic voters? One possibility is via top-down campaigning by political activists and the mass media. Another is motivation induced by direct and indirect experience of self, family, friends, and neighbors. The evidence favors an explanation that it is not just top-down persuasion and that individual experiences may play a major role.

First, the spatial nature of our main causal evidence favors the individual experience mechanism. It is implausible that political activism and mass media coverage have a sharp discontinuity at the 100-mile boundary, whereas personal experience with CBP is likely to have spatial discontinuity.

Second, we do not find evidence of discontinuous political gathering frequencies, a potential indicator of elite-driven activism, at the 100-mile boundary (Table 4). As described in Section 2.2, this publicly-available data is from Crowd Counting Consortium (2020). For this exercise, we first create bin$\times$county level observations, where binning was conducted using the directed distance from the 100-mile boundary. We then count event frequencies in each bin$\times$county. We try (i) several combinations of bin sizes and bandwidths, (ii) event frequency counts as well as its inverse hyperbolic sine transformation, (iii) event sizes, and (iv) filtering events using prominent immigration-related keywords; we find consistent null effects.

Third, we find that the mobilization effect is especially pronounced in densely populated areas. As shown in Table 5, there are heterogeneous treatment effects with respect to population density for both registration (panel A) and turnout (panel B).[11] Hispanic voters in above-median population density areas have 2.3pp and 2.2pp treatment effects on registration and turnout, whereas those in below-median population density areas have 0.54pp and -1.2pp treatment effects, lower than above-median's, although the difference is statistically significant only for the turnout heterogeneity ($p < 0.05$).

Finally, a back-of-the-envolpe calculation reveals that approximately two-thirds of the turnout effect is driven by registration. Under the assumption that all registered voters turn out at the same rate (i.e., no further CBP mobilization effect conditional on registration) of 55%,[12] we should see $1.8\text{pp} \times 0.55 = 0.99\text{pp}$ turnout effect, which is approximately two-thirds of the turnout effect (1.5pp) that we find in Column 2 of 2. This implies that CBP activities had a significant impact on marginal Hispanics who are undecided about registering. These marginal voters are likely to be politically less engaged, less informed, and not easily reachable through top-down

[11] Population density is defined as number of Hispanic voters in a 0.1 latitude $\times$ 0.1 longitude grid.

[12] This number comes from the fact that among registered Hispanics who reside within 3 miles from the 100-mile boundary, 55% voted.

approaches, which gives support to the mechanism through individual experiences.

5 Conclusion

In this paper, we provide empirical evidence that CBP activities mobilize Hispanic voters. CBP is granted enhanced powers within an arbitrarily-defined boundary—100 miles from the US border. We use an RDD to show that Hispanic voters within this area register and turn out at a higher rate. We do not find evidence of population sorting, and there is supporting evidence for the mechanism of direct and indirect experience with CBP driving this mobilization effect.

As in any paper using an RDD strategy, the treatment effect found in this paper is local, and therefore extrapolating the result to other areas of the US should be done with caution. However, it is definitely encouraging for external validity to see similar underlying characteristics between Hispanics living near our cutoff boundary and national Hispanic voters (Table 1). Our results thus speak to general trends in the United States whose electoral environment is increasingly polarized on the topic of immigration policy. Immigration enforcement can affect elections not only through its symbolic power, but through its direct effects on voters.

References

ACLU, "One Woman Who Knew Her Rights Forced Border Patrol Off a Greyhound Bus," June 2018. [Written by Mitra Ebadolahi online; posted 13-June-2018].

Braconnier, Céline, Jean-Yves Dormagen, and Vincent Pons, "Voter registration costs and disenfranchisement: experimental evidence from France," *American Political Science Review*, 2017, *111* (3), 584–604.

Cantoni, Enrico, "A precinct too far: Turnout and voting costs," *American Economic Journal: Applied Economics*, 2020, *12* (1), 61–85.

_ **and Vincent Pons**, "Strict ID Laws Don't Stop Voters: Evidence from a US Nationwide Panel, 2008–2016," Technical Report, National Bureau of Economic Research 2019.

Cato Institute, "Walling Off Liberty: How Strict Immigration Enforcement Threatens Privacy and Local Policing," November 2018. [Written by Matthew Feeney online; posted 1-November-2018].

CityLab, "Inside the Massive U.S. 'Border Zone'," May 2018. [Written by Tanvi Misra online; posted 14-May-2018].

Conley, Timothy G, "Spatial econometrics," *Microeconometrics*, 2010, pp. 303–313.

Crowd Counting Consortium, "Crowd Data," 2020.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of econometrics*, 2008, *142* (2), 698–714.

MIT Election Data and Science Lab, "U.S. House 1976-2018," 2019.

Pew Research Center, "2016 Campaign: Strong Interest, Widespread Dissatisfaction," 2016.

_ , "The Partisan Divide on Political Values Grows Even Wider," 2017.

Ray, Debraj and Arthur Robson, "Certified random: A new order for coauthorship," *American Economic Review*, 2018, *108* (2), 489–520.

TRAC, "Immigration and Customs Enforcement Arrests," 2018. Data retrieved from https://trac.syr.edu/phptools/immigration/arrest/.

White, Ariel, "When threat mobilizes: Immigration enforcement and Latino voter turnout," *Political Behavior*, 2016, *38* (2), 355–382.

_ , "Misdemeanor Disenfranchisement? The demobilizing effects of brief jail spells on potential voters," *American Political Science Review*, 2019, *113* (2), 311–324.

Table 1: Summary statistics

	Hispanic population in the US	Hispanics in US within 3 miles from 100-mile boundary
	(1)	(2)
Age	43.3	42.1
	(17.6)	(17.4)
Gender = Female	.516	.508
	(.5)	(.5)
Registered Democrat	.301	.245
	(.459)	(.43)
Registered Republican	.081	.0693
	(.273)	(.254)
Registered (2016)	.796	.798
	(.403)	(.401)
Voted (2016)	.482	.443
	(.5)	(.497)
Observations	20387457	404885

Notes: Standard errors are reported in the parentheses.

Table 2: Discontinuity at the 100-mile boundary

	Main outcomes: political participation			Baseline demographic characteristics				
	Registered (1)	Voted (2)	Vote Swing to Democrat (3)	Age (4)	Male (5)	Democrat (6)	Partisan Score (7)	Hispanic (8)
Border Zone Treatment	.018***	.015*	.013**	-.02	.0018	.0083	.55	-.0036
	(.0058)	(.0082)	(.0061)	(.41)	(.0019)	(.0081)	(1.3)	(.014)
County FE	Y	Y	Y	Y	Y	Y	Y	Y
Demographic controls	Y	Y	Y	N	N	N	N	N
Observations	395179	395179	395179	404862	404862	404862	404862	3216638
Sample mean	.8	.44	.12	44	.44	.18	77	.13
Sample s.d.	.4	.5	.44	17	.5	.53	26	.33

Notes: Each column in this table originates from a separate RDD regression with 3-mile bandwidth and rectangular kernel. The running variable (directed distance from 100-mile boundary) is included as a linear control, separately below and above the cutoff. We also control for latitude, longitude, age, and gender. For Columns 4-8, age and gender controls are excluded. Columns 1-7 include only Hispanic voting-age population, whereas Column 8 has all races.In parentheses below coefficients, spatial standard errors are shown with 100km cutoff. In Column 8 standard errors are clustered by county due to limitations in computational capacity. Note that this is also the most conservative standard error scheme for our purpose here (most likely to give rejection among all standard error schemes we consider). *** p<0.01, ** p<0.05, * p<0.1.

Table 3: No evidence of population sorting: 2010 census block

	Sample: 2010 census block					
	(asinh) Population	(asinh) Hispanic Population	Hispanic (Percentage)	White (Percentage)	Black (Percentage)	Asian (Percentage)
	(1)	(2)	(3)	(4)	(5)	(6)
Border Zone Treatment	-.081	-.079	.37	-.11	-.11	-.094
	(.072)	(.098)	(1.7)	(1.5)	(.76)	(.22)
County FE	Y	Y	Y	Y	Y	Y
Observations	146311	146311	85860	85860	85860	85860
Sample mean	2.2	1.7	13	75	7.6	1.7
Sample s.d.	2.1	3.2	26	33	20	6.7

Notes: Each column in this table originates from a separate RDD regression with 3-mile bandwidth and rectangular kernel. The running variable (directed distance from 100-mile boundary) is included as a linear control, separately below and above the cutoff. We also control for latitude, longitude, and log(area). In parentheses below coefficients, spatial standard errors are shown with 100km cutoff. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: No discontinuity in protest frequency

	Bin size: 0.05 miles			0.1 miles			0.2 miles		
	BW:1.5	BW:3	BW:4.5	BW:1.5	BW:3	BW:4.5	BW:1.5	BW:3	BW:4.5
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Border Zone	-.0033	-.0046	-.0059	-.006	-.0094	-.011	-.017	-.02	-.021
	(.0033)	(.0035)	(.005)	(.0063)	(.007)	(.0096)	(.013)	(.013)	(.016)
County FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	6480	12960	19440	3240	6480	9720	1512	3240	4752
Sample mean	.0056	.006	.0049	.011	.012	.0096	.02	.021	.018
Sample s.d.	.092	.092	.082	.13	.13	.12	.19	.19	.17

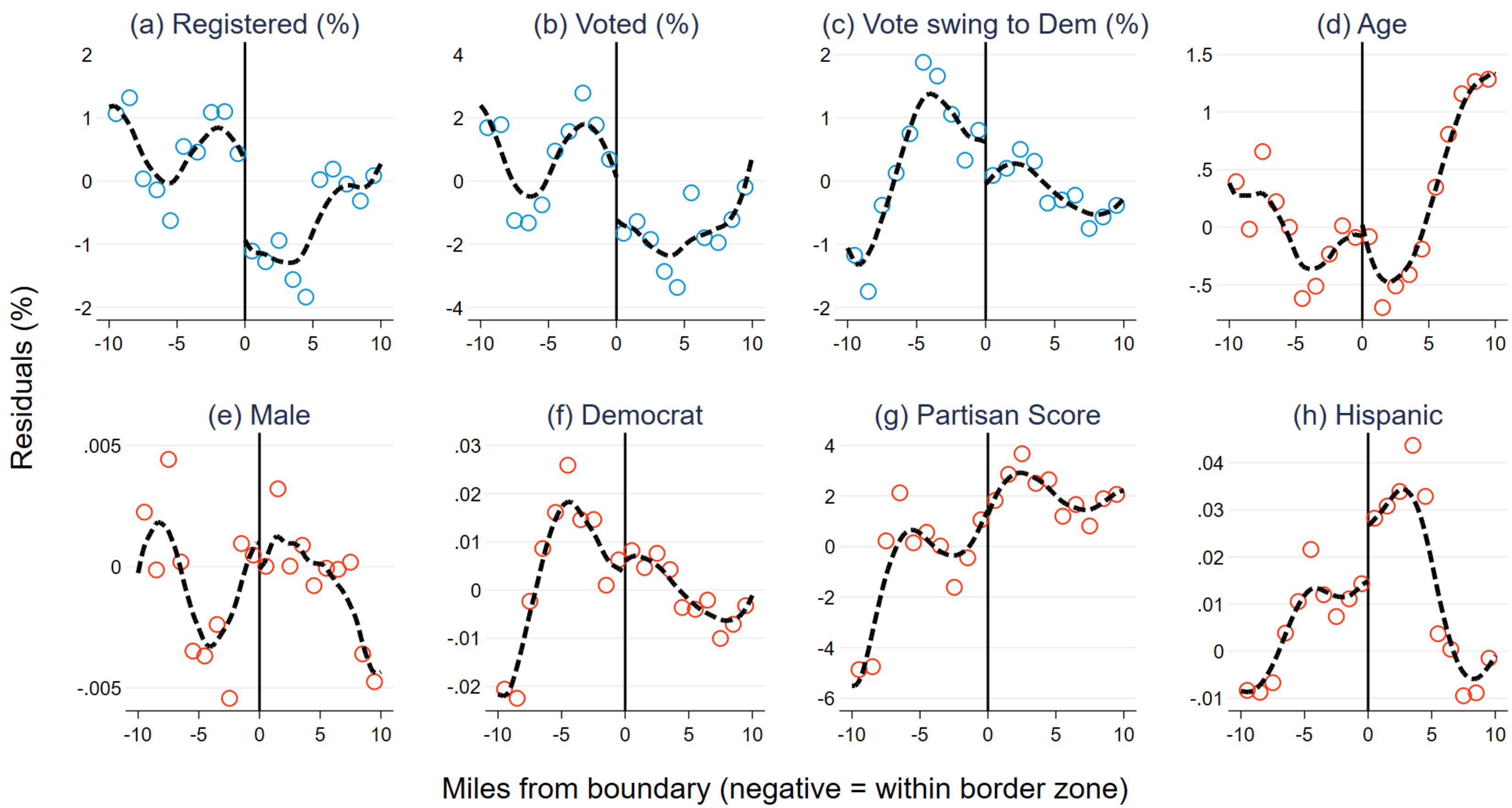
Notes: Each column in this table originates from a separate RDD regression with the specified bandwidth and rectangular kernel. We count the number of protests by distance bins (bin size specified above) within each county and run bin × county level regression. The dependent variable is inverse hyperbolic sine transformation of number of protests in each bin. The results are similar if we use (i) number of protests instead of asinh, (ii) protest size instead of frequency, or (iii) immigration-related events only. Standard errors (in parentheses below coefficients) are clustered by county. We report the results from county clustering to be most conservative (most likely to give rejection among all standard error schemes that we consider in this paper). Other SE schemes give similar results. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Heterogeneity by population density

	Dep. var: registered (by population density)			Dep. var: voted (by population density)		
	High Density	Low Density	All	High Density	Low Density	All
	(1)	(2)	(3)	(4)	(5)	(6)
Border Zone Treatment	.023***	.0054	.023***	.022*	-.012	.022*
	(.0057)	(.010)	(.0057)	(.011)	(.015)	(.011)
Treatment × Low Pop. Density			-.018			-.034**
			(.013)			(.015)
County FE	Y	Y	Y	Y	Y	Y
Observations	206280	188898	395178	206280	188898	395178
Sample mean	.8	.8	.8	.42	.46	.44
Sample s.d.	.4	.4	.4	.49	.5	.5

Notes: Each column in this table originates from a separate RDD regression with 3-mile bandwidth and rectangular kernel. To divide the sample, we first counted number of people within each 0.1×0.1 lat/long grid. Grids with fewer than the median population are classified as low density, and Hispanics in those grids are used for Columns 2 and 5. Other Hispanics are included in Columns 1 and 4. The running variable (directed distance from 100-mile boundary) is included as a linear control, separately below and above the cutoff. We also control for latitude, longitude, age, and gender. For Columns 3 and 6, all controls are interacted with the indicator variable for low-population density grids. In parentheses below coefficients, spatial standard errors are shown with 100km cutoff. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Residual RD plots (blue = main results; red = demographic balance)



Notes: We first residualize each dependent variable with all controls discussed in the notes of Table 2 and county fixed effects. Circles indicate the averages of the residualized dependent variables for 1-mile bins. Dotted lines are fitted lines from local linear regressions, separately from below and above the cutoff. Panels (a)-(g) include only Hispanic voting-age population, whereas Panel (h) has all races. Democrat means registered democrat, and the partisan score is calculated by TargetSmart based on its proprietary prediction method—the score ranges from 0 to 100, where 100 means a devoted Democrat.

For online publication: Online Appendix

Table A1: No voting-age Hispanic population discontinuity at the boundary (McCrary test)

	Bin size: 0.05 miles			0.1 miles			0.2 miles		
	BW:1.5	BW:3	BW:4.5	BW:1.5	BW:3	BW:4.5	BW:1.5	BW:3	BW:4.5
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Border Zone	.0095	.016	.011	.011	.023	.016	.015	.028	.02
	(.013)	(.017)	(.019)	(.017)	(.02)	(.022)	(.023)	(.023)	(.024)
McCrary p-val	.46	.34	.56	.52	.24	.47	.52	.22	.41
County FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	28740	57480	86220	14370	28740	43110	6706	14370	21076
Sample mean	.2	.2	.2	.27	.27	.27	.36	.36	.36
Sample s.d.	.89	.88	.88	1	1	1	1.2	1.2	1.2

Notes: Each column in this table originates from a separate RDD regression with the specified bandwidth and rectangular kernel. We count the size of Hispanic voting-age population by distance bins (bin size specified above) within each county and run bin $\times$ county level regressions. The dependent variable is the inverse hyperbolic sine transformation of the number of voting-age Hispanics in each bin. The results are similar if we use number of voting-age Hispanics instead of asinh. Standard errors (in parentheses below coefficients) are clustered by county. We report the results from county clustering to be most conservative (most likely to give rejection among all standard error schemes that we consider in this paper). Other SE schemes give similar results. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: No treatment effects for whites

	Main outcomes: political participation		
	Registered (1)	Voted (2)	Vote Swing to Democrat (3)
Border Zone Treatment	-.0019	-.0026	.0075
	(.0025)	(.0047)	(.0055)
County FE	Y	Y	Y
Observations	2463248	2463248	2463248
Sample mean	.85	.62	-.02
Sample s.d.	.36	.49	.53

Notes: Each column in this table originates from a separate RDD regression with 3-mile bandwidth and rectangular kernel. The running variable (directed distance from 100-mile boundary) is included as a linear control, separately below and above the cutoff. We also control for latitude, longitude, age, and gender. Standard errors (in parentheses below coefficients) are clustered by county. *** p<0.01, ** p<0.05, * p<0.1. We report the results from county clustering to be most conservative (most likely to give rejection among all standard error schemes that we consider in this paper). Other SE schemes give similar results.

Table A3: Main results are robust to clustering in different levels

	Cluster: Congressional District			Cluster: County		
	Registered (1)	Voted (2)	Vote Swing to Democrat (3)	Registered (4)	Voted (5)	Vote Swing to Democrat (6)
Border Zone Treatment	.018***	.015*	.013*	.018***	.015*	.013**
	(.0058)	(.0087)	(.0068)	(.0049)	(.0079)	(.0052)
County FE	Y	Y	Y	Y	Y	Y
N(clusters)	103	103	103	191	191	191
Observations	395179	395179	395179	395179	395179	395179
Sample mean	.8	.44	.12	.8	.44	.12
Sample s.d.	.4	.5	.44	.4	.5	.44

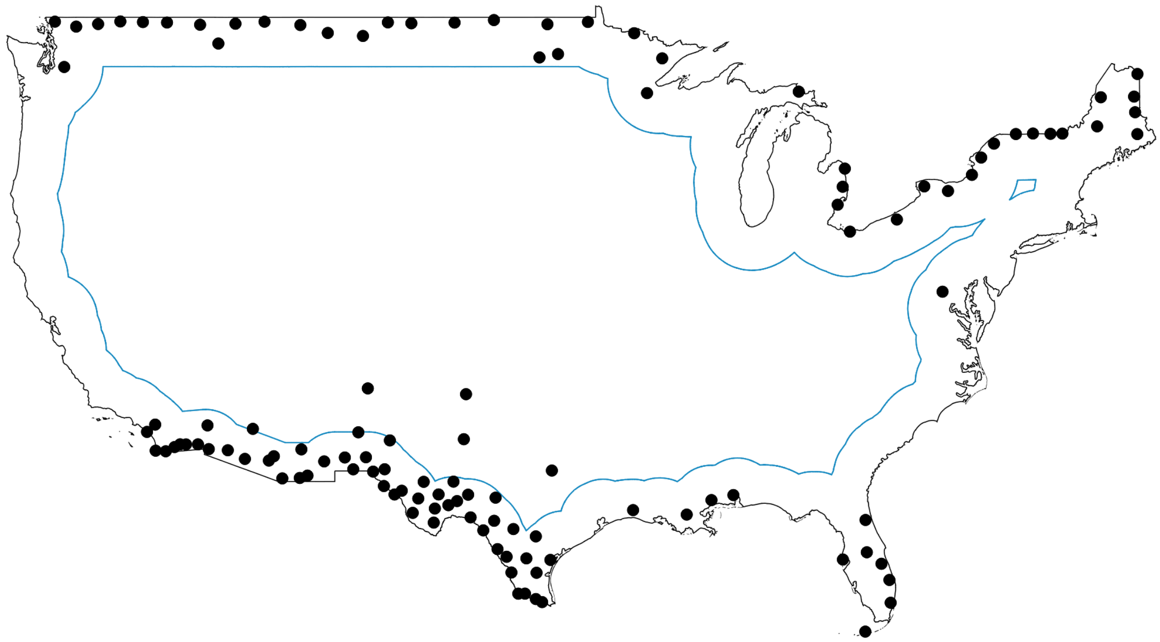
Notes: Each column in this table originates from a separate RDD regression with 3-mile bandwidth and rectangular kernel. The running variable (directed distance from 100-mile boundary) is included as a linear control, separately below and above the cutoff. We also control for latitude, longitude, age, and gender. Standard errors (in parentheses below coefficients) are clustered by group indicated above. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A4: Main results are robust to excluding top 10 most populated precincts

	Main outcomes: political participation		
	Registered (1)	Voted (2)	Vote Swing to Democrat (3)
Border Zone Treatment	.021***	.022**	.013**
	(.0065)	(.01)	(.0059)
County FE	Y	Y	Y
Observations	338996	338996	338996
Sample mean	.8	.45	.12
Sample s.d.	.4	.5	.44

Notes: Each column in this table originates from a separate RDD regression with 3-mile bandwidth and rectangular kernel. The running variable (directed distance from 100-mile boundary) is included as a linear control, separately below and above the cutoff. We also control for latitude, longitude, age, and gender. In parentheses below coefficients, spatial standard errors are shown with 100km cutoff. *** p<0.01, ** p<0.05, * p<0.1.

Figure A1: Duty locations of border patrol agents are strategically located



Notes: Duty locations (black dots in the figure) are geo-coded based on the figure in https://www.cbp.gov/careers/frontline-careers/bpa/duty-locations. Most duty locations inside the 100-mile boundary (blue line) correspond to major airports.

Figure A2: Histogram of running variable

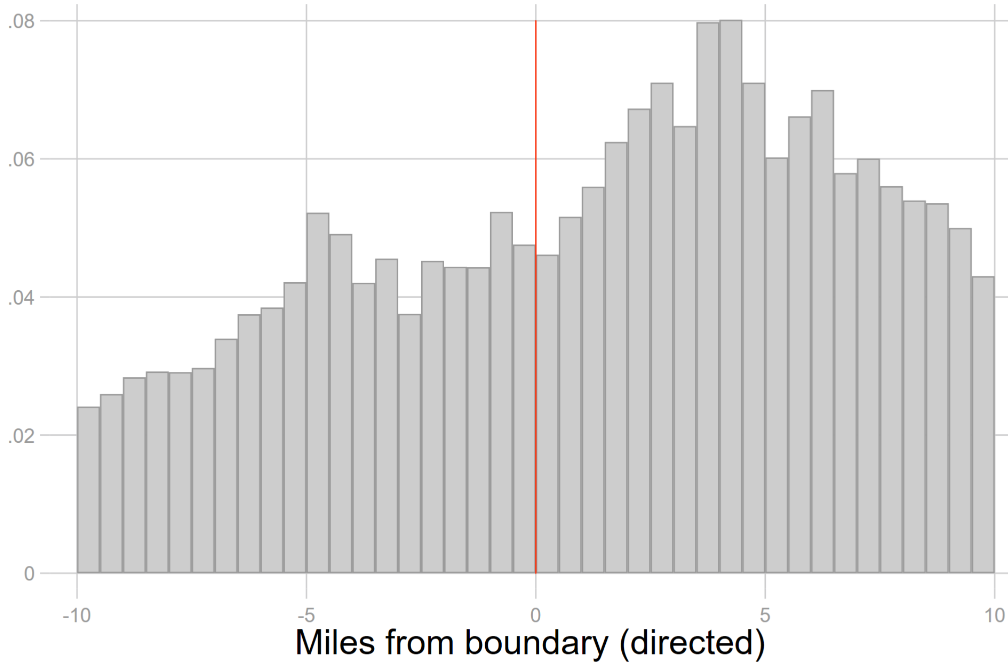
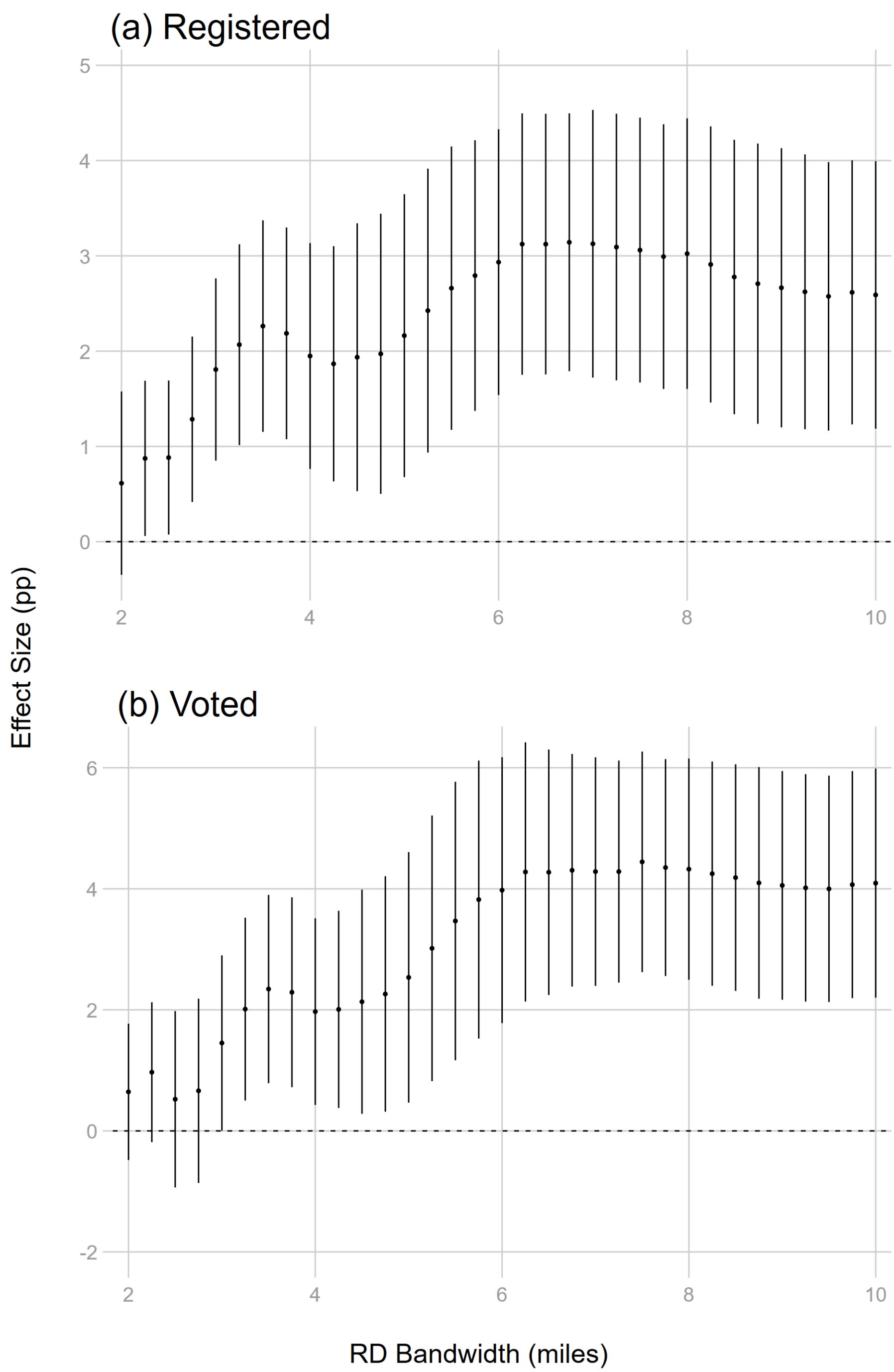


Figure A3: Political participation results using a range of bandwidths



Notes: Dots represent coefficient from running separate RDD regressions using different bandwidth (2-10 miles) with the same specifications as Columns 1 and 2 of Table 2. The lines show 90% confidence interval of each coefficient.