From Winning an Election to Political Participation*

Sahar Parsa (New York University)

October 2020

*New York University, sahar.parsa@nyu.edu. I would like to thank Debraj Ray, Karna Basu, Sam Boysel, Thomas Fujiwara, Siddharth Hari, Kelsey Jack, Camilo Garcia-Jimeno, Marco Migueis, Arturas Rozenas, James Snyder, Julia Schwenkenberg, Lore Vandewalle, and Jeffrey Zabel.

Abstract

How do people react when the candidate they supported wins or loses an election? I explore the consequences of having supported a winning (or losing) candidate on future political participation. I begin by presenting two empirical patterns: 1) there is a positive relationship between county vote share to the winning party and future voter turnout in that county, 2) there is no significant change in the party *composition* of votes, suggesting that the increase in political participation percolates to voters of all persuasions. I provide evidence for a combination of an "individual" and a "community" effect to explain this joint pattern. While winning voters are "encouraged" by the win of their preferred candidate compared to losing voters, losing voters appear to be galvanized by being surrounded by winners. In the data, the two effects happen to balance, highlighting the significance of peer effects on voter participation.

JEL classification: D72, D83, P16

Keywords: Voting, Voter turnout, Voting behavior, Learning

1 Introduction

Ever since the "paradox of voting" gave formal voice to a well-known puzzle, scholars from a wide range of fields have tried to understand the driving forces behind voter participation, often turning to behavioral and cultural explanations.[1] This paper investigates the effect of past electoral experience; specifically, that of being on the winning or losing side of a previous election, on subsequent voter participation. Being on "one side or the other" can take many shapes. I focus on a specific definition: an individual experiences an electoral win if the candidate *supported by* that individual wins the election. This definition includes individuals who supported a winning candidate independently of the act of voting, and so it encompasses both voters and non-voters. The hypothesis I test formally in this paper is that winning an election will be accompanied by an increase in voter participation in subsequent elections.

Conceptually, being on the winning or losing side of an election has both individual and community-level (or peer) effects. A supporter of the winning party might feel energized by their success, which is an *individual-level* effect. It could come from disparate factors. First, an individual supporting an elected official could become more engaged politically, as her candidate will represent her voice or preference in office. Such heightened engagement could translate into an increase in her political participation; see Zipp (1985). Second, and on a related note, the party in office could skew the distribution of funds towards counties where it received greater electoral support, see Ansolabehere and Snyder (2006). Third, voters could feel more engaged politically from the intrinsic effect of being on the winning side, as emphasized in the theoretical approach of Bendor, Diermeier and Ting (2003). Borrowing from the reinforcement learning literature, an electoral win would be seen as a successful outcome, worthy of repetition, while an electoral loss could be seen as a failure, meant not to be repeated. Finally, we should note there are also opposing arguments. For instance, the

[1] The "paradox of voting" — or the Downs paradox — states that a rational voter should not vote, because her costs of voting will exceed the expected benefits.

winning side could feel complacent about their win, while losers are spurred on to participate more actively.

Moreover, winning or losing an election is also a *collective* experience. The positive experience of winning an election is often shared with neighbors, colleagues, friends and acquaintances, leading to more political discussion and engagement, which then translates into a change in turnout.[2] Again, there are arguments for a positive or negative change. Socially generated complacency might be even more virulent than its individual counterpart. Or a loss could galvanize a losing community into trying harder the next time around. In their review essay on aspirations, Ray and Genicot (2019) provide a number of examples of this sort of behavior.

To identify these effects, I turn to close elections, where the identity of the winner behaves as if it were the outcome of a coin toss experiment. Tight races are a relevant setting to investigate the above effects, for two reasons. First, I emphasize the competitive nature of an election, not unlike other competitive activities such as team sports. And close elections are likely to heighten this notion of competition from the voter standpoint, where one party *is pitted against* the other, increasing the feeling of winning or losing. Using a sports analogy, soccer matches that pit a weak team such as the United States against a strong team like Germany differ fundamentally from matches that pit, say, Spain against Germany. In this paper, I focus on the set of electoral races that resembles the latter, and not the former. My choice of close elections puts both groups of voters on the same level playing field. In elections with larger margins between the winner and the loser, the effect of being on the winning or losing side has *other* informational, and psychological content that could be worth studying, but we abstract from these here.[3]

[2]The importance of social interactions in voting behavior have been suggested in a number of recent studies in a variety of contexts, see Madestam et al. (2013), Fujiwara, Meng and Vogl (2016) and Madestam et al. (2011). Most notable, Madestam et al. (2013) shows that socializing during the 4th of July increases the voter turnout while leaning towards the Republican party.

[3]For instance, it is possible that in the presence of higher victory margins (less competitive races), inspiration translates into frustration, which is beyond the scope of this study (see Ray and Genicot (2019) and the references cited there for a discussion)

Second, close elections assign the treatment status of "winner" or "loser" to individual voters essentially at random. Given the unpredictable and competitive nature of close elections, they are likely to elicit sincere voting from individuals, and therefore the random assignment of winner or loser status is likely to be psychologically salient, in the sense of having the result truly in (or out of) line with individual sentiments.

I focus on the effect of close gubernatorial elections between 1990 and 2015 on future voter turnout. Gubernatorial elections are particularly useful, as they are often highly competitive, and therefore not infrequently tight.[4] I then look at voter turnout in subsequent *presidential* and *senatorial* elections. For more detail on this empirical strategy and the choice of elections, see Section 3.1.

I begin by presenting two empirical patterns, using county-level electoral data. I divide counties in two groups: those counties with majority support for the gubernatorial winner — or *winning counties* — and counties with majority support for the gubernatorial loser, or *losing counties*. This is only an expository step: there is nothing special about the specific division of counties based on majority. We could have used supermajority thresholds, and indeed my later analysis will use smoothly varying county-level vote shares rather than any cutoffs at all.[5]

I then show that winning counties exhibit greater voter participation in subsequent elections, compared to losing counties. This effect on voter turnout is economically sizable. To give a sense of magnitude, the estimates imply a differential participation between winning and losing counties of 4 percentage points on an average turnout base rate of around 52%. Alternatively, the estimate represents a third of the standard deviation of the change in voter turnout.

Observe that this first outcome is compatible with any of the individual-level stimuli listed in the second paragraph of this introduction. However — and this is my second result

[4]Moreover, gubernatorial elections do not suffer from territorial changes and redistricting.

[5]The share of voters in a county that supported the winning candidate is plausibly positively correlated with the actual county support for a candidate, whether expressed in the ballot box or not. As such it can be seen as a proxy for the experience of an electoral win or loss, as defined in the first paragraph.

— I show that the increase in voter participation is not accompanied by a change in the party composition of votes cast. What might account for this symmetry in participation across winners and losers? One possibility, as suggested by Ansolabehere and Snyder (2006) in a related but distinct context, is that governing parties skew the distribution of funds towards counties that provided them with electoral support, but in so doing, they increase the participation of *both* supporters and opponents.[6] But this argument begs the question of why political leaders would pursue such a neutral but costly policy in the first place. More generally, the mechanisms emphasized above, such as having one's preferences represented in office or the presence of positive reinforcement effects from winning or losing, are not enough to rationalize this dual pattern of higher participation with no compositional change. The only exception is the possible conjunction of positive and negative forces referred to earlier in the introduction, to which I return. Specifically, I explore the presence of community and individual winning effects, captured in the estimated coefficients from the county-level regressions, which combine to yield this dual pattern.[7]

Parsing individual and neighborhood effects empirically is challenging for two reasons. First, I do not observe — for long time durations and across multiple elections — the voting behavior of the same individuals. Instead, I use more aggregate units; specifically, county-level data.[8] The difficulty — and the second empirical challenge — in trying to tease these

[6]Ansolabehere and Snyder (2006) examine the distribution of intergovernmental transfers of funds across counties from 1957 to 1997. In line with our results, the authors argue that the targeting of funds is correlated with an increased voter mobilization of all political stripes, rather than a shift of vote shares towards the elected party.

[7]That said, I do not take a firm stand on the precise channels that drive these effects. These include psychological winning effects, or the consequence of better funding, or policies that represent past winner preferences, not to mention a variety of peer effects that range from individual aspirations to collective action.

[8]To be more precise, for the purpose of this study, we need information about future and past electoral behavior for the same unit. Given the exogenous source of variation is at the state level — close gubernatorial elections — we are going back to the early 1990s in order to gain power and expand the scope of the study. There is no individual level (or even precinct level) longitudinal data with electoral information, going as far back as 1990. The use of counties is common, see Kanazawa (1998). Counties are not politically relevant subdivisions in this context, either to candidates or to voters. States don't have county-based electoral colleges for gubernatorial elections and, for gubernatorial candidates, winning by running up vote totals in safe counties is a viable strategy provided that safe counties have large enough population. This helps with the interpretation of the results. In other words, the county is mainly representing a specific geographic partition of a state.

twin effects apart is that while I am using *county-level* data, it goes without saying that I do not know what the exact community is. There is no reason to believe that community and county coincide. How then can I reconcile and isolate the presence of a community winning effect using county-level data? To achieve this, I introduce a framework to understand the underpinnings for the use of county-level data, in a setting where we — as analysts — are unsure of the true community structure. This framework will further help to interpret and extend the two county-level motivating results already discussed.

Making the (plausible) assumption that the relevant geographical community for an individual is not split across counties, I show that a community effect (no matter what the community is, as long as the restriction above is met) is explicitly captured by the coefficient on the squared vote share *at the county level* for the winning candidate. The influence of a squared term is not foreign to settings emphasizing similar peer effects (see, e.g., Esteban and Ray (2011)). In any community-level effect, a typical reduced form model will have the vote share as an influence on individual behavior, which is then squared via aggregation across all individuals supporting one party or another. The additional mileage afforded by my model is that data from a larger aggregate created by the analyst — a county in this case — is shown to be admissible, and captures both the individual and community effects described earlier.

The framework gives us formal permission to run a regression of subsequent voter turnout on the strength of voter support for the winning candidate as well as the square of that strength at the county level, and permits a deeper insight into individual and community effects. To be precise, it emphasizes three effects behind the jump in voter turnout observed in the data: (i) an individual winning effect, (ii) a community effect on the losers and (iii) a community effect on the winners. The aforementioned square term specifically identifies the differential community effect on individuals supporting the winning candidate and once estimated, reveals that effect to be negative. The estimated negative coefficient on the squared term suggests that all else equal, living in a winning community spurs on the participation

of losing voters relative the winning voters in that community.

To sum up, I explore the consequences of winning (or losing) an election on future political participation. I show that winning counties experience (i) an increase in voter participation compared to losing counties, but (ii) no significant change in the composition of votes, suggesting that this increase in voter participation extends to voters of all persuasions, both ideologically winning and losing voters.[9] However, I emphasize a strong asymmetry between the two. Specifically, I show that winning voters directly benefit from the fact that they were on the right side of the election, and participate more in subsequent elections. But the losers participate more too, for different reasons: they appear to be galvanized by being around winners. Overall, the results suggest a combination of an "individual" and a "community" effect to explain this dual pattern. That brings me to the result (iii): while winning voters are "encouraged" by the win of their preferred candidate compared to losing voters, losing voters are galvanized by the *community* level win. In the data, the two effects happen to balance, highlighting the significance of peer effects on voter participation.

This somewhat serendipitous cancellation is further explored in a numerical simulation that rests on the estimated coefficients of our regression. I use the no-composition effect to pin down that individual and peer effect, and show that it is in line with a strong positive individual win effect, an aggregate negative peer effect for the winners, and a positive peer effect for the losers, given our setting.

These patterns are interesting per se, but as I discuss at the end of the paper, one can draw two potential lessons from them, worth exploring further. First, they suggest an important role for political diversity as a driver of political engagement and participation. Just as a sunny or rainy day could affect voter turnout (Madestam et al. (2013), Madestam et al. (2011), and Fujiwara, Meng and Vogl (2016)), the experience of an electoral win at

[9]This is interesting to the extent that the result follows from close elections. Notice that close elections could be consistent with a highly contested environment where the differences between parties are very salient and/or minimal. Fraga and Hersh (2010) show that competitive elections differ substantially from less competitive elections. Their study suggests that voters are less likely to be on the fence about their decision to vote or abstain as elections become more competitive.

a community level could generate an environment favorable to higher political engagement and participation, if the community is politically diverse to start with. In these diverse environments, where losers and winners live together, interact, discuss and argue, the winners don't suffer as much from "complacency" out of living next to a lot of winners, *and* the losers benefit from living next to many enthused winners, sharing their political engagement while being aware of their differences. Second, from the perspective of a campaign manager who is trying to tilt voter turnout in their desired direction, an important question is whether to signal strength or shortcomings, through polling data, in order to deter the entry of rivals. In an analogy with Fudenberg and Tirole (1984), campaign managers might want to display a lean-and-hungry look to avoid the fact-cat effect, .

2 Related Literature

This paper fits into the broader literature on the determinants of voter participation. A wide range of studies have highlighted behavioral and cultural factors to explain the "Downs paradox",[10] from the role of "civic duty" to ethical considerations, pro-sociality, or social pressure, see Riker and Ordeshook (1968); Harsanyi (1977); Feddersen and Sandroni (2006); Benabou and Tirole (2006) and Ali and Lin (2013), with empirical and experimental studies coming from Fowler (2006); Fowler and Kam (2007); Dawes, Loewen and Fowler (2011); Gerber, Green and Larimer (2008) and DellaVigna et al. (2016).

This paper is tightly connected to the aforementioned literature, as also the studies in which voters are presumed to invoke adaptive rules of thumb. Building on Bush and Mosteller (1955) and Kanazawa (1998), Bendor, Diermeier and Ting (2003) introduce a model where voters are adaptively rational and generate turnout decisions using rules of thumb over their past turnout decisions and election outcomes. With reinforcement learning, actions that are successful today are more likely to be reproduced tomorrow, while unsuccessful ones are less likely. Using related arguments to Bendor, Diermeier and Ting (2003), recent studies

[10]The Downs paradox, also known as the "paradox of voting", states that for a rational, self-interested voter, the costs of voting will normally exceed the expected benefits, so she should not vote.

examine the impact of transitory shocks and their persistent effects on voter participation; see Kaplan and Mukand (2011), Madestam et al. (2013, 2011) and Fujiwara, Meng and Vogl (2016). For instance, Fujiwara, Meng and Vogl (2016) show how low turnout driven by rainy days tend to lead to low turnout in the future. Relatedly, Kaplan and Mukand (2011) find persistence from other shocks, showing that citizens registered to vote right after September 11, 2001, are more likely to be registered as Republicans even half a decade after the terrorist attacks. Madestam et al. (2013, 2011) use rainfall on Tax Day and Independence Day to estimate the effect of participating in Tea Party protests and independence day celebrations on political preferences and behavior. An electoral win (loss), like a sunny (rainy) day, seems to generate an environment favorable to higher political engagement that transcends support for the winning or the losing party.

There is a nacent literature exploring psychological mechanisms behind voting behavior; see Mullainathan and Washington (2009) or Healy, Malhotra and Mo (2010).[11] We emphasize the psychological effect of winning or losing a competition. Psychologists have long studied the effect of success and failure in other settings. As far back as Thorndike (1927)'s law of effect, it has been recognized that any behavior that is followed by pleasant consequences is likely to be repeated, and any behavior followed by unpleasant consequences is likely to be stopped. The law of effect is related to the "winner effect", a well-researched phenomenon in different competitive fields, such as chess, sports, and business, see Robertson (2012) and Coates (2012). The winner effect focuses on the biological effect of winning, making one more likely to overcome increasingly harder challenges in the future. In Bernhardt et al. (1998) testosterone samples are taken from fans before and after the world final between Brazil and Italy, showing how the levels of testosterone of the Brazilian supporters remained high, while the Italian ones went down after the match. Coates (2012) suggests how these hormonal changes could be communicated to other people, and lead large groups of people

[11]Mullainathan and Washington (2009) provide evidence consistent with cognitive dissonance theory among voters, and Healy, Malhotra and Mo (2010) provide evidence behind local college football games outcomes in electoral behavior. Healy, Malhotra and Mo (2010) argue that voters' well-being may influence voting decisions on a subconscious level.

to experience an "upward spiral of confidence". Electoral competition, by virtue of their competitive nature, could have similar effects.

Finally, this paper connects to the wide literature on close elections in the United States.[12] Although I am using close elections, the empirical strategy slightly differs from previous work on narrow electoral margins. Most studies in the existing literature on close elections compare close winners to close losers across state-, district-, or municipal-level elections. These comparisons are typically made between different regions and predicate on differences between candidates such as party affiliation or gender. In this paper, I am interested in the set of close elections per se, and I will invoke the empirical specification inherent to the RDD literature, contributing to the large body of knowledge on close elections and electoral behavior.

3 From Winning an Election to Voter Participation

I start with two motivating empirical patterns. Using aggregate county-level electoral data, I show that 1) counties leaning towards the gubernatorial winning party exhibit a higher voter turnout in subsequent (presidential and senatorial) elections, compared to counties leaning towards the gubernatorial election loser, but 2) these counties do not exhibit any compositional change, on average, in close elections. Specifically, counties leaning towards the winning party are defined as counties with more than 50% of their vote share to the winning candidate, and counties leaning away from the winning party are defined as counties with less than 50% of their vote share to the winning candidate.[13] There is nothing particularly special about this choice, which is only intended as an expository example. The formal

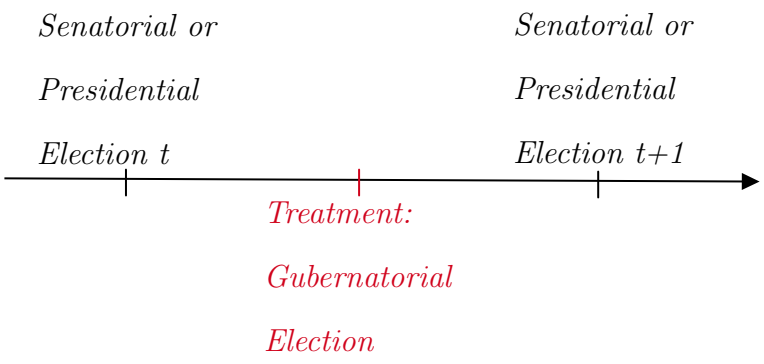
[12] Within this literature, close gubernatorial elections are still under-exploited compared to other offices; see Erikson, Folke and Snyder Jr (2015).

[13] For more descriptive statistics about these counties, see Online Appendix Table 2 Panel A and B. It shows the mean and the standard deviation for the set of counties included in our sample, decomposed between "winner" and "loser" counties. The main difference between Panel A and Panel B is that Panel B restricts the sample further to the pool of counties where the governor won with less than a 5% margin. Overall, these tables demonstrate that the winner and loser counties do not differ along a set of socio-economic characteristics. We will come back to these socio-economic characteristics after the results, where we provide a full set of specification tests.

analysis in Section 4 will use the actual vote shares.

3.1 Data Description and Methodology

I investigate the change in voting behavior between t and $t+1$ in presidential and senatorial elections staggered in between a close gubernatorial election:



The main treatment uses *gubernatorial* elections. This has a number of practical advantages, as discussed in the introduction. Among others, the high number of close elections, essential to identifying the winning effect, increases the sample size and the generalizability of the results. The analysis does not rely on a small number of elections. In the time period covered in this study, 1990 to 2015, 42% of the gubernatorial races exhibited an electoral margin within 5% and all states have had at least one election with a margin of 5% or less with the exception of Delaware, North Dakota and South Dakota.[14]

However, while the winning treatment variable uses gubernatorial elections, I investigate the winning effect on the closest subsequent *presidential and senatorial* elections following the gubernatorial elections. Given the choice of gubernatorial elections as the "independent variable," the focus on subsequent presidential and senatorial elections has a number of advantages. First, presidential and senatorial elections are closer in time to the gubernatorial elections than, say, the *next* gubernatorial elections, and so help in capturing the power

[14]I define explicitly the margin of victory in equation (1). The Online Appendix Figure 1 illustrates the histogram of the gubernatorial election margin of victory variable. The distribution is left skewed with a mean of 0.049 and a standard deviation of 0.07. This is in line with Erikson, Folke and Snyder Jr (2015), although they use a longer sample.

of recent events. Furthermore, it minimizes the influence of a governor in office and the possible role played by traditional notions of incumbency advantage, decreasing the direct influence of the governor's behavior on her/his re-electability. Although the governor could still influence voting behavior on subsequent non-gubernatorial elections, the range of that influence is plausibly smaller than on gubernatorial elections. For instance, it is unlikely that the identity of the candidates running for other elections (presidential or senatorial) is a function of a specific county winning or losing the previous gubernatorial election. Each of these reasons helps in narrowing the possible interpretations of the results.

Dataset: I use electoral data from Dave Leip's Atlas,[15] from 1990 to 2015, for gubernatorial (state-level), senatorial and presidential elections, at the relatively disaggregated county level.[16] All U.S. governors serve four-year terms except those in New Hampshire and Vermont, who serve two-year terms. Senators serve terms of six years each, where approximately one-third of the seats are up for election every two years in a staggered fashion. Notice that both seats from a given state are only contested in the same general election when a mid-term vacancy is being filled.[17] Senators are elected by their state as a whole. Finally, presidential elections are held every four years, and the first election in the sample is 1992. Democrats control the governorship in approximately 44% of the 355 original number of elections, Republicans control approximately 53%, and the remaining 8 elections are won by a third party. There are 12 states, which had, at some point in time within the sample, a third party as either the governor or the runner-up. This accounts for 17 out of the 355 gubernatorial elections.[18] I exclude these elections from the final sample for ease of interpretation. Table 1 takes a first look at the observable characteristics for the remaining 338

[15]Leip, David. Dave Leip's Atlas of U.S. Presidential Elections. http://uselectionatlas.org (2019)

[16]I decided to exclude U.S. House of Representatives elections from this study, due to the complication added by the changes in congressional districts' boundaries. I recognize that counties also experience changes during the time period considered, but to a much lesser extent than congressional districts.

[17]For a number of years and states in the sample, I have more than one contested seat. In which case I proceeded in two ways: (i) I took the average vote shares to the party during the election, (ii) alternatively, I built a sample where I discard the years with a vacancy. The results are not sensitive to the method used. To increase the sample size, I kept the method with the average vote shares.

[18]The states are Alaska, Colorado, Connecticut, Hawaii, Louisiana, Maine, Minnesota, New York, Rhode Island, Texas, Utah, Vermont.

elections.[19] It reports the main statistics across various sub-samples defined by the gubernatorial margin of victory. The main takeaway is that states × elections do not differ in terms of the socio-economic characteristics in Table 1, as I restrict the gubernatorial margins of victory from all margins of victory to 5%, 2.5% and 1%.

Table 1: Mean and standard deviation of states socio-economic characteristics by gubernatorial margins of victory

	All states			5% margin			2.5% margin			1% margin		
	obs	mean	s.d.	obs	mean	s.d.	obs	mean	s.d.	obs	mean	s.d.
Margin of victory	338	0.09	0.07	140	0.02	0.01	77	0.01	0.01	28	0.00	0.00
Turnout	338	0.45	0.11	140	0.46	0.09	77	0.46	0.10	28	0.46	0.09
Total pop. (000)	338	5498	6254	140	5597	5247	77	5783	5186	28	5738	4890
% White	338	0.81	0.13	140	0.81	0.12	77	0.80	0.13	28	0.79	0.12
% African American	338	0.10	0.09	140	0.10	0.10	77	0.11	0.10	28	0.10	0.09
Median Income (000)	338	40	12	140	41	12	77	42	13	28	42	13
% Homeownership	338	0.67	0.05	140	0.67	0.04	77	0.67	0.05	28	0.67	0.05
% 4 year college	338	0.23	0.05	140	0.24	0.06	77	0.25	0.06	28	0.25	0.06
% Labor Force	338	0.64	0.04	140	0.64	0.04	77	0.64	0.04	28	0.65	0.04
% Unemployed	338	0.04	0.01	140	0.04	0.01	77	0.04	0.01	28	0.04	0.01

Notes: Table 1 reports the mean and the standard deviation of the states' socio-economic characteristics by the gubernatorial margins of victory (all and less than 5%, 2.5%, 1%). The socio-economic characteristics come from the decennial census, 1990 to 2010, and the electoral characteristics have been calculated using Dave Leip's Atlas.

There are a total of 21,379 county × year observations for the gubernatorial elections, with gubernatorial elections held every year in the sample. Along the same line, 13,610 counties x elections preferred a Republican candidate, 7538 counties x elections preferred a Democrat candidate, and 231 counties x elections supported a third party. Once merged with the subsequent presidential and senatorial election data and restricted to the 338 states, the final sample has 32193 observations, made of 15344 senatorial and 16849 presidential county × year election data.[20]

Finally, the county *voter turnout*, for the sample of presidential and senatorial elections pooled, has a symmetric distribution with a mean around 0.51 and a standard deviation of

[19] In the Online Appendix Table 1, you can find the main descriptive statistics including these 17 states. The exclusion of these 17 states doesn't alter these sets of socio-economic characteristics.

[20] The imbalance between the presidential and senatorial election data comes from the last gubernatorial election years, which have presidential election data for all counties in the sample, while many states didn't have the subsequent senatorial election yet.

0.12. The distribution of the county *vote shares to the gubernatorial winning party* is slightly skewed to the right with a mean of 0.56, while the standard deviation is around 0.13. See Online Appendix Figure 2 and 3 for the density function of these two variables.

Empirical Specification: Recall that we are interested in the differential change in voter turnout of *counties* leaning towards the winning party compared to *counties* leaning towards the gubernatorial election loser, for the set of close gubernatorial elections *at the state level*. To identify the set of close gubernatorial elections, I start by constructing the margin of victory of the gubernatorial elections at the state level, denoted $z_{s,t}$:

$$z_{s,t} \equiv \frac{vote_{s,t}^{\text{winning}}}{vote_{s,t}^{\text{winning}} + vote_{s,t}^{\text{runner-up}}} - 0.5, \tag{1}$$

where $vote_{s,t}^{\text{winning}}$ is the number of votes received by the elected party in state s in year t and $vote_{s,t}^{\text{runner-up}}$ is the number of votes received by the runner-up party during the same gubernatorial election.

Ideally, I want to estimate the relationship between the vote share of a county that accrues to the winning candidate, and the change in voter turnout in that county, for the set of gubernatorial elections where the gubernatorial winner's margin of victory is as small as possible. My main specification borrows from the regression discontinuity approach. I estimate the following equation:

$$\triangle T_{c,t} = \alpha_0 + \alpha_1 W_{c,t} + \alpha_2 W_{c,t} \times z_{s,t} + \alpha_3 z_{s,t} + \alpha_4 X_{c,t} + \eta_{c,t}, \tag{2}$$

where the sample is limited to the set of close gubernatorial elections, with a margin of victory in the neighborhood of 0.[21] The specification in equation 2 includes a local linear control on the *state-level margin of victory* (the equivalent of a running variable in the RDD literature). Including the local linear control accounts for the possibility of a relationship between the margin of victory and the outcome variable, even within the window of close

[21] see Hahn, Todd and Van der Klaauw (2001) and Lee and Lemieux (2010).

elections. I will present the results for four samples, namely elections with a margin of victory smaller than 5%, 2.5% and 1%, as well as for the optimal bandwidth.[22] In addition to the main specification, I will also present a robustness check which includes office, county, and year fixed effects, as well as the political party the county supported during the gubernatorial election. $X_{c,t}$ is the set of control variables. Finally, the outcome variable is standardized, and I use robust clustered standard errors at the state level.

The outcome variable in equation 2 is the difference in *voter turnout* between the closest election preceding, and the closest election following the gubernatorial election, denoted $\triangle T_{c,t}$. The voter turnout of county c in the presidential or senatorial election in year t is defined as the total number of votes divided by the voting age population of county c in year t.[23] In order to keep the notation simple, I use the index c and t without specifying the office (senatorial or presidential), even though $T_{c,t}$ is office specific and the formal analysis will explicitly control for it. The outcome variable is in first difference. Voter turnout is highly persistent and an important determinant is past voting behavior. I follow previous studies and use first differences with close elections. This helps to increase statistical efficiency and to deal with serial correlation; see Lee and Lemieux (2010) and de Benedictis-Kessner and Warshaw (2016).

Finally, I use a dummy variable, $W_{c,t}$, to identify the counties leaning towards the gubernatorial winning party, and the counties leaning away from the gubernatorial winning party. As mentioned at the beginning of Section 3, I use a majority cutoff to sort the counties in

[22] Our optimal bandwidth selection uses the bandwidth minimizing Mean Squared Errors (MSERD) for MSE-optimal point estimation using a common bandwidth on both sides of the cutoff, see Calonico et al. (2019).

[23] The voting age population is defined as the total population in county c above 18 years old measured using U.S. Census data. Given the decennial nature of the Census, I interpolate the voting age population of county c for intermediate electoral years using the available data. Consider two consequentive census years, Y and $Y+1$, I construct the annual growth rate of the population above 18 years old of a given county c as $g = (population18plus_{Y+1}/population18plus_{Y})^{1/10} - 1$, where $population18plus_{Y+1}$ is the population of county c in census year $Y+1$. Note that I annualize the growth rate by using $1/10$ as the census is every 10 years. For each year t in between two census years, Y and $Y+1$, I then apply the following formula to estimate the population above 18 years old of a given county c: $\widehat{population18plus_t} = population18plus_Y \times (1+g)^{t-Y}$. For instance, the interpolated population for county c in 1994 becomes $\widehat{population18plus_{1994}} = population18plus_{1990} \times (1+g)^4$

these two groups, where $W_{c,t}$ takes a value of 1 if the majority of its voters voted for the winning candidate and 0 otherwise. The coefficient of interest is α_1, i.e, the average differential change in turnout for the group of these winning counties compared to the group of losing counties, in close elections.

3.2 Two Motivating Patterns

1. Greater Voter Participation: The first result is summarized in Table 2. Column 1 to 4 cover the optimal bandwidth, 5%, 2.5% and 1% margins of victory, and column 5 shows the optimal bandwidth with the control variables. The estimates are all positive and statistically significant by conventional standards.

Table 2: First Difference in Voter Turnout on Winning Dummy, $W_{c,t}$

	1	2	3	4	5
	opt bw	5% bw	2.5% bw	1% bw	opt bw + control
$W_{c,t}$	0.436	0.479	0.626	0.426	0.412
s.e.	0.159	0.136	0.189	0.373	0.157
bandwidth	0.050	0.025	0.010	0.036	
obs	9244	12936	6611	2115	9244

Notes: Table 2 reports the main result for the first difference in turnout. Each column of the table represents a different specification or bandwidth. Column 1 to 5 cover the optimal bandwidth, 5%, 2.5% and 1% margins of victory, and column 5 shows the optimal bandwidth with controls. Standard errors are clustered at the state level. The variables are standardized.

The estimates oscillate around 0.5 of a standard deviation of the change in voter turnout, with a standard deviation of 0.09 and a mean close to 0. Using the coefficient on the optimal bandwidth (column 2) on the conservative side, the winning counties would enjoy a change in their turnout close to 0.436 of a standard deviation, relative to the losing counties. This corresponds to a change in turnout of 3.92%[$= 0.09 \times 0.436$], i.e., a change close to 7%[$= 3.92/51.5$] of the average turnout in our sample, 51.5%. Furthermore, the inclusion of the set of controls, county, office and year as well as the party the county supported fixed effects, does not seem to affect the prevailing result (see last column of Table 2). The point estimate with fixed effects and control is only marginally smaller from the baseline specification, reassuring us of the validity of the estimates.

The Online Appendix shows a number of alternative specifications. First, I could have used turnout in the subsequent election, while controlling for turnout in the previous election, instead of the first difference in turnout. In the Online Appendix Table 3 and Figure 6, I show the results remain unchanged for this alternative specification. This is expected given we are exploring close elections, and it is likely that, on average, winning and losing counties do not differ along their past turnout (and other socio-economic characteristics), which I explicitly investigate and show in the Online Appendix Section 4.[24] The consistency in the main patterns for this alternative specification affirms confidence in the estimates.[25] Second, the results hold for the sample of presidential elections or senatorial elections separately (without pooling them), see the Online Appendix Section 2, Table 4. Finally, I refer you to the Online Appendix Section 3 for a discussion and a vizualization of the data, and the source of the estimated coefficient.

2. The Constancy of Vote Composition: The just-described positive relationship between voter participation and the vote share to the winning party at the county level is consistent with the hypothesis that the supporters of the winning party become more politically engaged. This enhanced engagement could be either by virtue of winning an election or because the voter's party is in office and is therefore implementing her preferences in office. At the same time, losing voters could also be galvanized into participating more in a bid to counteract and reverse their past losses. It could also be that their social interactions with winning voters infuse them with the desire to take a more active role in the political sphere.

[24]This is an advantage of looking at close elections, which we specifically test in the Online Appendix Section 4. Although for a given election, winning counties will differ from losing counties, to the extent that they vote for another candidate, with all the differences in political preferences in might entail, aggregating accross all these close elections, the group of winning and losing counties are more likely to be balanced. For instance, one close election might confound being a democrat with being a loser, while for the set of close elections, democrats and republicans are as likely to be on both the winning and losing side. As such, the results are less likely to confound a specific alternative characteristic differentiating winning and losing counties.

[25]Finally, an alternative interpretation for the empirical strategy is that I am using a difference-in-difference strategy, where I compare the turnout of the winning and losing counties before and after the election, after restricting the sample to elections won with a close margin as the exogenous treatment mechanism of winning and losing.

I therefore turn to the effect of winning an election, not on overall participation, but on the *composition* of that participation — across vote shares for the winning (losing) party. It goes without saying that this is of intrinsic interest. But, in addition, an examination of the effect on vote composition will help us shed more light on the drivers of the increase in overall participation.

Specifically, let $v_{c,t}^{\text{winner}}$ be the senatorial or presidential vote share of county c assigned to the winning state governor party, i.e., the vote count to the party winning in the preceding gubernatorial election divided by the total number of votes in county c in the presidential or senatorial election t. I measure the first difference in the vote share to the winning party as $v_{c,t}^{\text{winner}} - v_{c,t-1}^{\text{winner}}$. Even though the first difference will give us information about the change in the composition of votes, it might suffer from the fact that winning and losing counties might start with different vote shares to the winning party by virtue of their winning and losing status. Specifically, winning counties might not have as much room for change compared to losing counties, as we expect them to start with a higher $v_{c,t-1}^{\text{winner}}$. To account for the aforementioned difference, instead of using a simple change in vote composition, I also provide the results for a measure of normalized change given by $\frac{v_{c,t}^{\text{winner}} - v_{c,t-1}^{\text{winner}}}{1 - v_{c,t-1}^{\text{winner}}}$. The denominator will correct for the fact that a winning county has a larger $v_{c,t}^{\text{winner}}$ in the baseline period. Note that the problem is symmetric for the losing county, they might not have as much room to move away from the winning party by virtue of having a smaller $v_{c,t}^{\text{winner}}$. So, I also report the percentage change in the vote shares to the winning party, $\frac{v_{c,t}^{\text{winner}} - v_{c,t-1}^{\text{winner}}}{v_{c,t-1}^{\text{winner}}}$. It should be noted that the measure of "normalized change" effectively scales up (down) the gain in vote share for winners (losers), because they have a relatively high (low) share to begin with, while the percentage change should scale the change down (up). Finally, I also show the result for a dummy variable equal to one if there is a change towards the winning party and zero otherwise. This measure will account for the direction of the change, but abstracting from the intensity of such change.

Table 3 summarizes the results for the optimal bandwidth and the optimal bandwidth

with control. I show the results for other bandwidths in the Online Appendix Table 7 — the results are unchanged. The first four columns show the results for the first difference of the vote shares to the winning party. These standardized coefficients are negative but small and mostly not statistically significant. As briefly discussed in the previous paragraph, this could be due to the fact that winning counties might not have as much room for change compared to losing counties. However, once we correct for this possibility by using the normalized change in column 7 and 8, the results are virtually zero. Overall, the table suggests a weak negative to no compositional change towards the winner, and mostly estimates that are all statistically insignificant.

Table 3: Change in vote shares to winning party on Winning and Losing Counties

	1	2	3	4	5	6	7	8	8	
	First Difference			Percentage		Normalized		Dummy		
	opt	5%	opt +	opt	opt +	opt	opt +	opt	opt +	
				control		control	opt	control	opt	control
$W_{c,t}$	-0.210	-0.204	-0.168	0.001	-0	-0.209	-0.199	-0.040	-0.021	
s.e.	0.185	0.163	0.160	0.003	0.003	0.163	0.128	0.087	0.087	
bandwidth	0.050	0.049	0.013	0.013	0.034	0.047	0.051	0.051		
obs	10352	12936	12572	3047	3687	8932	11847	13310	13310	

Notes: Table 3 reports the estimates for various measures of change in the vote share to the winning party running for presidential or senatorial elections. Each column of the table represents a different specification or bandwidth. Column 1 to 3 cover the first difference in the vote shares to the winning party for the OLS regression with all margins of victory across all states, column 1 to 3 cover the optimal bandwidth, 5%, and the optimal bandwidth with controls. Column 4 (optimal bandwidth) and 5 (optimal bandwidth + control) cover the percentage change in the vote shares to the winning party. Column 6 (optimal bandwidth) and 7 (optimal bandwidth + control) cover the normalized change in the vote shares to the winning party. Column 8 (optimal bandwidth) and 9 (optimal bandwidth + control) cover the dummy variable, equal to one if there is a change towards the winning party and zero otherwise. All the standard errors are clustered as the state level. See Online Appendix Table 7 for other bandwiths.

To sum up, I report (i) an increase in voter participation in winning counties compared to losing counties, but (ii) no change in the composition of votes, suggesting that (i) this increase in voter participation benefits all parties. How can I make sense of these combined patterns?

One possibility is that voters supporting the winning party in the gubernatorial elections support a different party in the subsequent senatorial or presidential elections. Switchers or split tickets do exist, with different parties supported in different offices. But the *relative*

magnitudes are small. Between 1990 to today, on average, 7.82% of individuals split tickets between the House and the Presidential Elections, while 5.1% split tickets between the Senate and the Presidential Elections.[26] These magnitudes are not insignificant, but to explain the absence of a compositional effect, they would have to be systematically biased in the direction "winners → losers," rather than the other way around. That could happen if voters were driven by considerations of political balance across parties, but that effect would have to be dominant for the switching argument to explain what we see, and even then the numbers are not big enough.

The explanation that I favor, and that I now proceed to develop in some detail, is that "winning an election" comes with both individual and peer effects. The positive experience of winning an election could be shared with neighbors, colleagues, friends and acquaintances, leading to more political discussion and engagement, translated in an increase in turnout. This increase in voter participation and engagement extend to and benefit all voters of all persuasions, both ideologically winning and losing voters, translating into a medley of reactions. The first that winning per se is good for participation: winning voters are reinforced in the direction of their win by wanting to vote in the future. But the second effect galvanizes losers, who realize that they must redouble their efforts the next time around. It is possible that in the presence of higher victory margins (less competitive races), this sort of inspiration translates into frustration, which is beyond the scope of this study (see Ray and Genicot (2019) and the references cited there for a discussion). But at any rate, it is an empirical question which demands a theoretical framework to discipline the data.

[26]See The American National Election Studies (ANES) at www.electionstudies.org. I can also gauge the changes between gubernatorial and other elections by looking at the party that respondents voted for in the last presidential, senatorial and house elections. This comparison is based on a smaller sample — less than 1000 across all years, due to the overlap of these questions in the ANES. The share of individuals voting for the same party across the gubernatorial and other elections is on average around 82% from 1990 to today, whereas 12% of respondents claim to vote for the other major party. I averaged the number across all elections. The remaining share $(1 - 0.82 - 0.12)$ includes both voting for the third parties and not voting.

4 Individual and Neighborhood Effects

The purpose of this section is two-fold. First, I seek to provide underpinnings for the use of county-level data, in a setting where I — the analyst — am unsure of the true community structure. Second, I seek to interpret and extend the two patterns discussed in the previous section, which have to do with positive turnout and unresponsive composition. I reconcile these two findings by positing and exploring two effects connected to winning an election: an individual and a community effect. I will explore these in some detail.

The reader might wonder why I am not using individual level data. First, information about the same individual's electoral behavior across multiple election years is rare, making it hard to study the impact of an electoral win on subsequent elections. Second, it is even harder to observe individual political support. The reason I would need that information is because an increase in turnout will largely come from the subset of non-voting individuals, and we would need to know their political preferences at that anterior stage. Finally, and most importantly, in the presence of peer effects, which we focus on, and in the absence of information about the relevant neighborhood for each individual, the use of individual data would generate an unavoidable bias. We could have used precinct level data, but their availability would have restricted us to more recent elections, creating power issues given that the treatment is at the state level. To circumvent these matters, I use aggregate data at the county level, where I can follow county-level electoral outcomes across many years and electoral offices. The framework I introduce will help us formally connect the individual and community winning effects to appropriate coefficients in a county level regression.

4.1 Conceptual Matters

Suppose that a continuum of individuals resides in the state, with each individual indexed by i. Let n_i be the relevant neighborhood of individual i, which is just a subset of agents — clustered by some attribute(s) — for which the individual experiences a sense of community. For the purpose of my analysis I will suppose that the n_i's *partition* the space of individuals;

that is, if $j \in n_i$, then indeed $i \in n_j$, and $n_i = n_j$. These could be tiny, so that each person is almost an island, or they could represent neighborhood blocks or perhaps a zip code, or still larger units, such as Greenwich Village or Lower Manhattan in New York City. What *is* important for our purpose is that these units are fully contained within counties. This is an assumption which may or may not be met, depending on the salient mode of communication and discussion. That is not to say that I disallow all other forms of interaction, such as geographically unrestricted communication over social media. But I *do* intend to imply that spatial clustering within the county is an important factor for social interaction. (Other non-spatial dimensions are not captured in the analysis.)

I track changes in turnout over two consecutive elections as a function of being in a "winning neighborhood," defined as a neighborhood which elected the winning candidate in a closely-contested state election. I assume that there are just two types of voters — those who support the winner and those who support the loser.[27] Specifically, if $\triangle T_i$ is the change in turnout of an eligible voter i in a neighborhood n for the next election of the same type in a neighborhood, then I posit that

$$\triangle T_i = \gamma_n + \alpha_d w_i + \alpha_v v_n + \alpha_I [w_i \times v_n] + \gamma_i + \epsilon_{n,i}, \tag{3}$$

where γ_n is a neighborhood-specific intercept term, w_i is an indicator function for whether the individual supported the winning candidate in the previous state-level election, *whether or not she voted*, α_d is the direct impact of supporting that winning candidate, v_n is the vote share obtained in i's neighborhood by the state-level winning candidate, α_v is the direct impact of the vote share in the neighborhood, α_I is an interaction term that captures the additional effect of the neighborhood vote share (for the gubernatorial winner) on a resident who was on the winning side, and γ_i is a person-specific effect. I take this last term to be iid across winners and losers in a neighborhood, with mean zero without loss of generality.[28]

[27] This structure can be extended to a set of undecided voters as well, but I do not explore the implications of that extension here.

[28] As already noted, this model includes past abstainers as well. The interpretation must be that abstainers

Aggregating (3) across individuals in a neighborhood n with m_n eligible voters, dividing by m_n, and now interpreting $\triangle T_n$ as the change in the neighborhood-level turnout rate, I see that

$$\begin{aligned}\triangle T_n &= \left[\gamma_n + (\int_{i\in n} \gamma_i)/m_n\right] + \alpha_d v_n + \alpha_v v_n + \alpha_I v_n^2 + \epsilon_n \\ &= \alpha_n + \alpha_w v_n + \alpha_I v_n^2 + \epsilon_n, \end{aligned} \tag{4}$$

where the old term γ_n as well as the person-specific effects $\{\gamma_i\}$ now all get subsumed into a neighborhood-level fixed effect α_n. The new coefficient α_w combines the previous α_d and α_v into a coefficient that captures the *sum* of the direct effect of winning (α_d) and any neighborhood effect (α_v) on the baseline eligible voter, here taken to be the losing eligible voter. The nonlinear squared term v_n^2 is the all-important signature for neighborhood effects on the winning voter.

This is the structure one would prefer to estimate, but as already mentioned, it can't be done, because (a) I do not know what the exact neighborhoods are, and (b) even if I did, I do not have data at the neighborhood level. I use counties instead, which — by my assumption — nest all neighborhoods. With that in mind, it is possible to connect (4) to a corresponding county-level specification.

Think of a county c containing one or more neighborhoods (one, if the county *is* the neighborhood, otherwise more). Let the county-level vote share for the winner (in the state election) be given by v_c. Think of each neighborhood n within a county c as a subunit, which is related to the county in some specific (deterministic) way; that is, its vote share is just a shift of the county vote share. Thus, if the neighborhood has a vote share v_n for the winning candidate of the state, then

$$v_n = v_c + \eta_n$$

where η_n is a neighborhood-specific deterministic shifter of the "voting climate" at the level

have pre-assigned ideologies; i.e., that they are "winners" or "losers" even if they did not vote.

of the county. Of course, by definition,

$$\sum_{n \in c} \pi_n \eta_n = 0, \tag{5}$$

where π_n is the population share of neighborhood n; after all, the neighborhood shares have to aggregate to the county vote shares ($\sum_n \pi_n v_n = v_c$). Therefore, if I aggregate (4) over all neighborhoods in the county, then I obtain:

$$\begin{aligned}
\triangle T_c = \sum_{n \in c} \pi_n (\Delta T_n) &= \sum_{n \in c} \pi_n \alpha_n + \alpha_w \sum_{n \in c} \pi_n v_n + \alpha_I \sum_{n \in c} \pi_n v_n^2 + \sum_{n \in c} \pi_n \epsilon_n \\
&= \sum_{n \in c} \pi_n \alpha_n + \alpha_w v_c + \alpha_I \sum_{n \in c} \pi_n (v_c + \eta_n)^2 + \epsilon_c, \\
&= \sum_{n \in c} \pi_n \alpha_n + \alpha_w v_c + \alpha_I v_c^2 + \alpha_I \sum_{n \in c} \pi_n \eta_n^2 + \epsilon_c \\
&= \alpha_c + \alpha_w v_c + \alpha_I v_c^2 + \epsilon_c,
\end{aligned} \tag{6}$$

where the penultimate equality uses equation (5), and where the last line defines

$$\alpha_c \equiv \sum_{n \in c} \pi_n [\alpha_n + \eta_n^2]. \tag{7}$$

Equation (7) shows that the squared shifters at the neighborhood level, as well as any county-level intercept term, all get absorbed into an overall county fixed effect α_c. Of course, the separate terms in (7) are not identified but I don't observe them, nor do I particularly care.[29]

Overall, equation (7) gives us a formal specification tying the vote share to the winning party and voter turnout in the presence of peer effects differentiated by the support of the individual (i.e., whether they supported the winner or not). In fact, in the presence of such peer effects, the use of a linear regression on the vote share to the winning party, omitting

[29]This procedure is not valid, if, instead of thinking of the η_n's as neighborhood-specific fixed effects, I think of them as stochastic perturbations of v_c. That would introduce a bias akin to measurement error, that I avoid. Even in such cases I would be all right if I had a large number of neighborhoods and so could take recourse to the law of large numbers applied to $\{\eta_n^2\}$. Alternatively, under the RDD strategy, one would expect (this would be an untestable assumption) balance along these stochastic perturbations at the cutoff.

the square term, is an unambiguous mis-specification.

4.2 Estimation of the Neighborhood Effect

I can directly estimate α_w and α_I using equation 6. Formally, I invoke the same RDD strategy to estimate the following regression under the optimal bandwidth, as well as 5%, 2.5% and 1% margins of victory:

$$\triangle T_{c,t} = \alpha_0 + \alpha_w v_{c,t}^{\text{winner}} + \alpha_I (v_{c,t}^{\text{winner}})^2 + [\alpha_w^z v_{c,t}^{\text{winner}} + \alpha_I^z (v_{c,t}^{\text{winner}})^2] \times z_{c,t} + \alpha_z z_{c,t} + \epsilon_{c,t}, \quad (8)$$

where α_w is the estimated coefficient on the strength of the support to the winning candidate (α_w), while α_I is the estimated neighborhood effect for the winning candidates (α_I).

Observe that this model differs from a standard RDD. In a traditional RDD setting, we are interested in estimating the "jump" in the outcome variable at the threshold. Instead, I am interested in identifying and estimating the relationship between turnout change and the vote share (and vote share squared) to the winning party, i.e., α_w and α_I, but around the "threshold". As in the RDD specification in the previous section, I estimate the coefficients on a subsample of close elections, where the identity of the winner is essentially the result of a coin toss. The reason behind the use of the RDD estimator is the same as for the earlier exercise: the level of support for the winning candidate could be correlated with the error term and the RDD buys us a source of randomness in the identity of the winning candidate. In addition, in the same vein as in the previous section, I allow an interacted term with the running variable (the margin of victory) on both sides of the threshold. All variables are standardized and robust clustered standard errors at the state level are used. Finally, I control for office, county, and year fixed effects, as well as the political party the county supported during the gubernatorial election, just as I did before. Note that the county fixed effect serves to capture α_c in our model.

The results are summarized in Table 4. They show that while α_w is positive, α_I is negative. The finding that $\alpha_I < 0$ suggests that the differential neighborhood effect *for the*

Table 4: First Difference in Voter Turnout on Vote Share and Vote Share Squared

	opt (3.75%)	5% bw	2.5% bw	1% bw
	(1)	(2)	(3)	(4)
Vote Share Winner	0.548	1.083	0.348	0.828
	(0.270)	(0.448)	(0.179)	(1.089)
Vote Share Winner sqd	-0.466	-0.982	-0.316	-0.675
	(0.282)	(0.430)	(0.180)	(1.142)
Observations	9,548	12,936	6,611	2,115

Notes: Table 4 shows the RDD result for different bandwidth of the first difference of voter turnout on the vote share to the winning party during the gubernatorial elections and the vote share squared. All the standard errors are clustered as the state level.

winning voters is negative. That is, holding everything else equal, a voter supporting the winning candidate is *less* likely to vote, the stronger the support for the winning candidate in his relevant neighborhood, compared to the voter living in the same neighborhood who voted for the losing candidate. One might interpret this as a "complacency effect": the knowledge that there were many supporters of the same political stripe might cause a winning voter to relax. I illustrate this relationship using a non-parametric fit between the vote share to the winning party and voter turnout, in Figure 1. I restrict the sample to the 5% margin of victory; see the Online Appendix, Figure 10 for other cutoffs.[30] Specifically, I use the locally weighted regression (loess) method. The method fits a smooth function representing the relation between the change in turnout and the vote share to the winning party, by leveraging for each vote share only the neighboring vote shares. This method is a non-parametric fit as it makes no assumptions about the form of the final relationship, and the final functional form is data-driven. It is interesting to notice that the final form of the non-parametric fit is single-peaked, and certainly compatible with a quadratic specification.

But of course, the negative coefficient on the squared term is not the only effect. There is additionally a positive estimated α_w that gives us the non-parametric fit in Figure 1 in

[30]The quadratic fit is robust accross all these cutoffs.

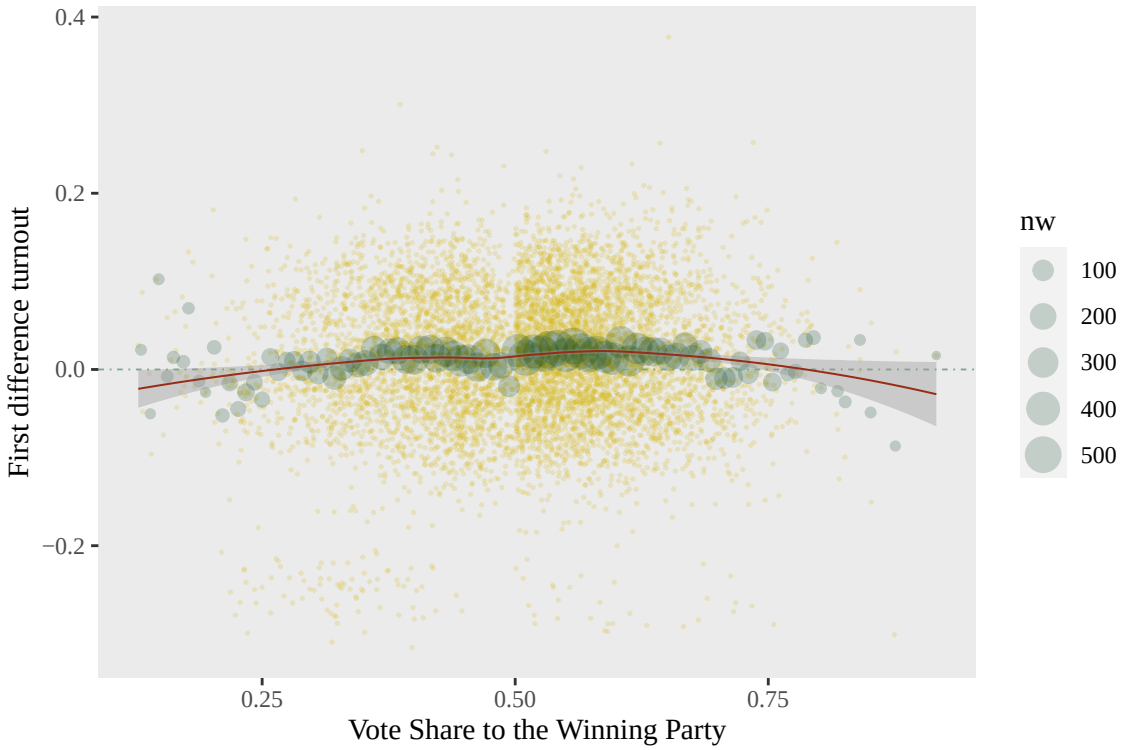


Figure 1: Non-parametric fit and linear fit of the change in voter turnout to the support to the winning party

Notes: Figure 1 is the first difference in voter turnout from election year t to election year $t+1$ for senatorial and presidential elections, i.e. $\triangle T_{c,t} = T_{c,t+1} - T_{c,t}$, where $T_{c,t}$ is the voter turnout for county c in election year t, on the vote share to the gubernatorial winning party for the 5% margin of victory. I created equal vote share bins. These bins have different number of counties treated within them, which is captured by the size of the circle. I overlaid a non-parametric fit using the locally weighted regression (loess) method. Standard errors are reported (shaded color around the line).

line with the coefficient estimated in Table 4. Recall that α_w is a composite variable, i.e., α_w can be expressed as the sum of the neighborhood effect for the baseline eligible voters, the loser, and the individual winning effect, from equation (4):

$$\alpha_w = \alpha_d + \alpha_v$$

Recall that α_d is the individual win effect, while α_v is the neighborhood effect on the baseline voter, taken to be the losing voter. Our positive value of α_w comes from some combination of the two. To put it more vividly, who turns out to a greater degree: the losing voters via the neighborhood effect, the winning voters via the individual win effect, or both? We

can shed some light on this by returning to the two main findings — a net positive shift in county voter participation, and no county compositional change, as we move from losing to winning counties. The latter finding implies that we need new voters — winners and losers — to vote in proportions similar to the earlier vote share of the county they live in. That suggests a *positive* neighborhood effect for the *losing* voters ($\alpha_v > 0$). But at the same time, because the neighborhood effect for winning voters is negative, this must also imply that $\alpha_d > 0$ as well; that is, the individual winning effect is positive.

Let us examine this in more detail. Let T^+ and T^- be the two turnout rates among those who support the winners and losers respectively, at the county level. As a matter of simple accounting, it must be the case that

$$dv^{\text{winner}} = v^{\text{winner}}(1 - v^{\text{winner}})(dT^+/T^+ - dT^-/T^-), \tag{9}$$

where v^{winner} is the observed vote share to the winner, and $(1 - v^{\text{winner}})$ is the observed vote share to the loser. Equation 9 tells us that a zero compositional change in a county would require the percentage change in losing voters to equal the percentage change in winning voters.To match our finding, though, we ask for something weaker: we only need the *average* compositional change to be the same for the set of winning and losing counties. In turn, this average change depends on two factors: (i) the distributions of v^{winner} among winning and losing counties, and (ii) the division of α_w between the individual win effect α_d and the neighborhood effect on losers α_v. The former distributions are, of course, observed by the analyst, so if we take our estimates seriously, we can use the no-composition effect to impose discipline on the latter division in (ii). In particular, what is the decomposition of α_w in α_d and α_v such that the average compositional change is the same for the winner and the loser counties, given the distribution of vote shares to the winning candidate in our sample?

Figure 2 illustrates this issue for the optimal bandwidth sample in Table 4, where we show the average predicted compositional change as a function of the decomposition of α_w

between α_d and α_v, for the set of winning and losing counties separately.[31] We plot these averages against hypothetical values of α_d, where $\alpha_d \in [0, \alpha_w]$, and $\alpha_w = 0.548$. Notice that

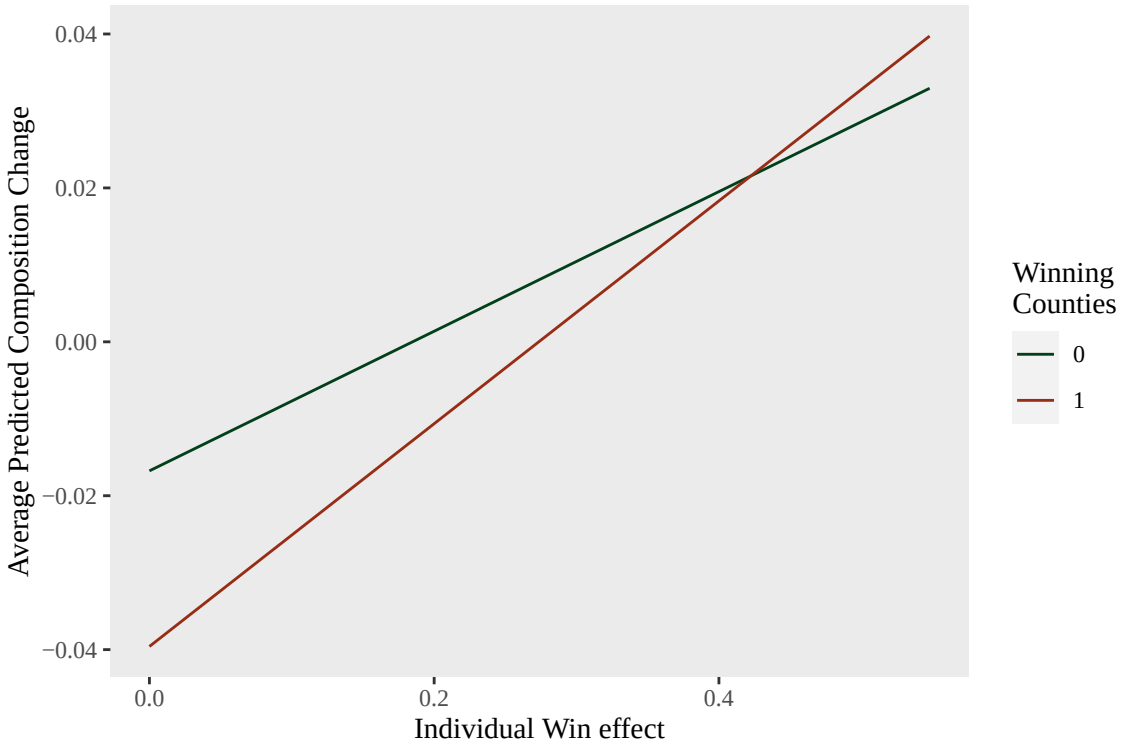


Figure 2: Predicted composition change as a function of the individual win effect by Winning and Losing Counties

Notes: Figure 2 is the average predicted composition change for winning and losing counties, from Table 4 — the optimal bandwidth, against the hypothetical individual win effect from $\alpha_d \in [0, 0.548]$, given the distribution of counties in the optimal bandwidth sample.

when $\alpha_d = 0$, the change in turnout mostly comes from the neighborhood effect. It follows that the average change in the vote share to the winner is greater in the losing counties, given the negative peer effect coefficient for winning voters. That is, the estimated compositional change is negative. However, as α_d increases, the differential compositional change decreases and eventually vanishes when $\alpha_d \simeq 0.4$.

Note that each $\alpha_d \in [0, \alpha_w]$ affects dT^+/T^+ and dT^-/T^- via equation 3, by changing the importance of the individual win effect α_d, or equivalently the baseline peer effect $\alpha_v = \alpha_w - \alpha_d$, and so predicts compositional change as a function of the vote share to the winning

[31] To be precise, the two lines represent $\frac{\sum_{i \in I_w} dv_i^{\text{winner}}}{N_l}$ and $\frac{\sum_{i \in I_l} dv_i^{\text{winner}}}{N_l}$, where I_w is the set of winning counties, and I_l is the set of losing counties in the optimal bandwidth sample.

candidate. Figure 3 illustrates this collection of functions, where each curve represents a different individual win effect from 0 to 0.548. Observe that the compositional change as a function of the vote share to the winner is nonlinear. That means that the *average* compositional change over all winning counties will depend on the *distribution* of counties across the vote share to the winning candidate. The final prediction for composition is ambiguous, while the results for turnout will unambiguously increase with the vote share to the winning party. In our setting, given the distribution of vote shares to the winning and losing counties, the dual pattern in the data (higher turnout and no compositional change), turns out to be in line with a strong individual win effect, a negative peer effect for the winning voters, and a positive peer effect for the losing voters.

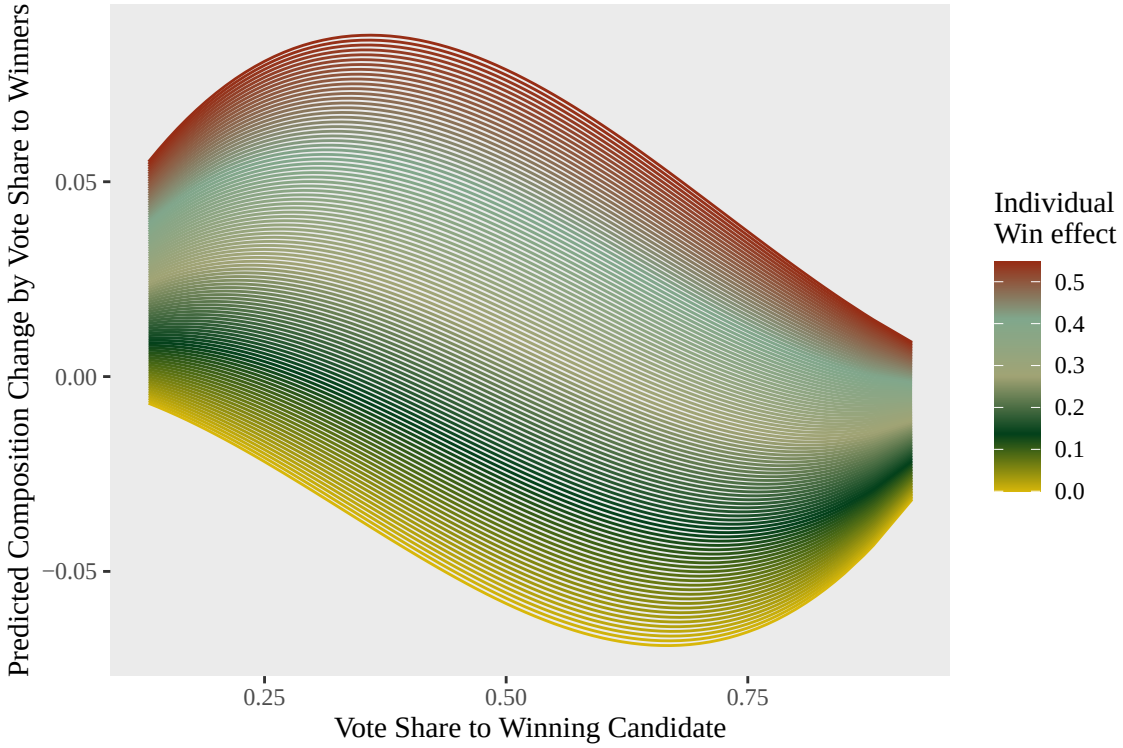


Figure 3: Predicted composition change as a function of the individual win effect by Winning and Losing Counties

Notes: Figure 3 is the predicted composition change as a function of the vote share to the winning candidate for the sample of optimal bandwidth, different functions represent the predicted value for $\alpha_d \in [0, 0.548]$.

As a final remark, I must qualify this discussion to state that I have assumed little or

no ideological change, or a change from losing voters to winning parties. This is a plausible assumption in my opinion, and it is supported by the data in the previous section. It forces the compositional shift to come from voter participation as opposed to voters changing party affiliation.

To sum up, I show that winning counties experience (i) an increase in voter participation compared to losing counties, but (ii) no significant change in the composition of votes, suggesting that (iii) this increase in voter participation extends to voters of all persuasions, both ideologically winning and losing voters. I emphasize a strong asymmetry between *the two types of voters*. Specifically, I show that winning voters directly benefit by being on the right side of the election, and participate more in subsequent elections. But the losers participate more too, for different reasons: they appear to be galvanized by being around winners. Overall, the results suggest a combination of an "individual" and a "community" effect to explain this dual pattern. (iii) While winning voters are "encouraged" by the win of their preferred candidate compared to losing voters, losing voters are galvanized by the community level win.

5 Conclusion

In this paper, I investigate the effects of winning an election on political participation. By studying counties that showed support for a winner in close, state-level elections, I show that greater county support for a winning candidate results in more political participation by that county in subsequent elections. More intriguing is the fact that such participation is across the board — while participation rises, it does not appear to do so in a differential way for candidates that share the same political affiliation as the previous winner.

Much of the analysis concerns an understand of the sources of this participation effect, along with the zero-composition result. Might citizens who supported the winner be motivated in the same way as citizens who supported the loser? By definition, that cannot be — the former "won" and the latter "lost." I show that the sheer fact of winning (or the

"individual effect") is what appears to drive the later participation of winning voters. In contrast, the losing voters are galvanized by their community experience of the loss (a "peer effect"). The former shows up statistically as a linear effect, the latter as a quadratic effect.

Beyond the provision of evidence for both an individual and a peer effect, these results suggest an interesting role for political diversity in political participation. If my proposed framework is taken seriously, Table 4 highlights the implication that the winning effect is stronger precisely in counties with a mixture of voters from all persuasions. In fact, the change in turnout is weakest in a segregated world: one in which winning citizens live only among other winning citizens, and losers among losers. While winning citizens still experience the individual effect of victory, they are rendered relatively complacent by the large observed margin of their victory — the peer effect is negative. At the same time, losing citizens — those who might have been spurred on by the visible hype of winners among them — will also not be incentivized to participate in the future. As a consequence, in a world with spatial sorting or segregation along dimensions correlated with political preferences, the winning effect would be associated with a smaller increase in voter turnout.

Presumably, these effects can be explored even further, but statistical identification of these finer channels would require more data. One possibility is that if everyone thinks the same way and shares the same ideologies, there is less room for debate and discussion, resulting in anemic political engagement and a diminished incentive to vote. Alternatively, borrowing from the literature on reference-based behavior, highly competitive races could galvanize the losers precisely when they live close to the other group, see Ray and Genicot (2019). Overall, this could generate a relationship between residential segregation (or geographic sorting) and voter turnout. One could extend the argument to any form of distance that limits the interaction between individuals of different persuasions, be it polarization or geographical sorting. That said, a proper investigation of the effect of diversity on political participation is beyond the scope of this study.

References

Ali, S Nageeb and Charles Lin. 2013. "Why people vote: Ethical motives and social incentives." *American economic Journal: microeconomics* 5(2):73–98.

Ansolabehere, Stephen and James M Snyder. 2006. "Party control of state government and the distribution of public expenditures." *The Scandinavian Journal of Economics* 108(4):547–569.

Benabou, Roland and Jean Tirole. 2006. "Belief in a just world and redistributive politics." *The Quarterly journal of economics* 121(2):699–746.

Bendor, Jonathan, Daniel Diermeier and Michael Ting. 2003. "A behavioral model of turnout." *American Political Science Review* 97(2):261–280.

Bernhardt, Paul C, James M Dabbs Jr, Julie A Fielden and Candice D Lutter. 1998. "Testosterone changes during vicarious experiences of winning and losing among fans at sporting events." *Physiology & Behavior* 65(1):59–62.

Bush, Robert R and Frederick Mosteller. 1955. "Stochastic models for learning.".

Calonico, Sebastian, Matias D Cattaneo, Max H Farrell and Rocio Titiunik. 2019. "Regression discontinuity designs using covariates." *Review of Economics and Statistics* 101(3):442–451.

Coates, John. 2012. *The hour between dog and wolf: Risk-taking, gut feelings and the biology of boom and bust*. HarperCollins UK.

Dawes, Christopher T, Peter John Loewen and James H Fowler. 2011. "Social preferences and political participation." *The Journal of Politics* 73(3):845–856.

de Benedictis-Kessner, Justin and Christopher Warshaw. 2016. "Mayoral partisanship and municipal fiscal policy." *The Journal of Politics* 78(4):1124–1138.

DellaVigna, Stefano, John A List, Ulrike Malmendier and Gautam Rao. 2016. "Voting to tell others." *The Review of Economic Studies* 84(1):143–181.

Erikson, Robert S, Olle Folke and James M Snyder Jr. 2015. "A gubernatorial helping hand? How governors affect presidential elections." *The Journal of Politics* 77(2):491–504.

Esteban, Joan and Debraj Ray. 2011. "Linking conflict to inequality and polarization." *American Economic Review* 101(4):1345–74.

Feddersen, Timothy and Alvaro Sandroni. 2006. "Ethical voters and costly information acquisition." *Quarterly Journal of Political Science* 1(3):287–311.

Fowler, James H. 2006. "Altruism and turnout." *The Journal of Politics* 68(3):674–683.

Fowler, James H and Cindy D Kam. 2007. "Beyond the self: Social identity, altruism, and political participation." *Journal of Politics* 69(3):813–827.

Fraga, Bernard and Eitan Hersh. 2010. "Voting Costs and Voter Turnout in Competitive Elections." *Quarterly Journal of Political Science* 5:339–356.

Fudenberg, Drew and Jean Tirole. 1984. "The fat-cat effect, the puppy-dog ploy, and the lean and hungry look." *The American Economic Review* 74(2):361–366.

Fujiwara, Thomas, Kyle Meng and Tom Vogl. 2016. "Habit Formation in Voting: Evidence from Rainy Elections." *American Economic Journal: Applied Economics* 8(4):160–88.
URL: *http://www.aeaweb.org/articles?id=10.1257/app.20140533*

Gerber, Alan S, Donald P Green and Christopher W Larimer. 2008. "Social pressure and voter turnout: Evidence from a large-scale field experiment." *American Political Science Review* 102(1):33–48.

Hahn, Jinyong, Petra Todd and Wilbert Van der Klaauw. 2001. "Identification and estimation of treatment effects with a regression-discontinuity design." *Econometrica* 69(1):201–209.

Harsanyi, John C. 1977. "Morality and the theory of rational behavior." *Social research* pp. 623–656.

Healy, Andrew J, Neil Malhotra and Cecilia Hyunjung Mo. 2010. "Irrelevant events affect voters' evaluations of government performance." *Proceedings of the National Academy of Sciences* 107(29):12804–12809.

Kanazawa, Satoshi. 1998. "A possible solution to the paradox of voter turnout." *The Journal of Politics* 60(4):974–995.

Kaplan, Ethan and Sharun Mukand. 2011. The persistence of political partisanship: Evidence from 9/11. Technical report Mimeo, University of Maryland - http://econweb.umd.edu/~davis/eventpapers/KaplanPersistence.pdf.

Lee, David S and Thomas Lemieux. 2010. "Regression discontinuity designs in economics." *Journal of economic literature* 48(2):281–355.

Madestam, Andreas, Daniel Shoag, Stan Veuger and David Yanagizawa-Drott. 2011. "Shaping the nation: The effect of Fourth of July on political preferences and behavior in the United States." *Preprint - https://pdfs.semanticscholar.org/c747/9d24d4a2efd1606a6475cd5739a230cd497e.pdf* .

Madestam, Andreas, Daniel Shoag, Stan Veuger and David Yanagizawa-Drott. 2013. "Do political protests matter? Evidence from the tea party movement." *The Quarterly Journal of Economics* 128(4):1633–1685.

Mullainathan, Sendhil and Ebonya Washington. 2009. "Sticking with your vote: Cognitive dissonance and political attitudes." *American Economic Journal: Applied Economics* 1(1):86–111.

Ray, Debraj and Garance Genicot. 2019. "Aspirations and Economic Behavior." *Forthcoming, Annual Review of Economics* .

Riker, William H and Peter C Ordeshook. 1968. "A Theory of the Calculus of Voting." *American political science review* 62(1):25–42.

Robertson, Ian H. 2012. *The winner effect: The neuroscience of success and failure*. Macmillan.

Thorndike, Edward L. 1927. "The law of effect." *The American Journal of Psychology* 39(1/4):212–222.

Zipp, John F. 1985. "Perceived Representativess and Voting: An Assessment of the Impact of Choices vs.Echoes." *American Political Science Review* 79(1):50–61.