

# Estimating Neighborhood Effects on Turnout from Geocoded Voter Registration Records\*

Michael Barber<sup>†</sup>

Kosuke Imai<sup>‡</sup>

First Draft: April 6, 2013  
This Draft: January 28, 2014

## Abstract

Do voters turn out more or less frequently when surrounded by those like them? While decades of research examined the determinants of turnout, little is known about how the turnout of one voter is influenced by the characteristics of other voters around them. We geocode over 50 million voter registration records in California, Florida, and North Carolina and estimate the effects of racial and partisan composition of small residential neighborhoods at the census block level. Through cross-section and panel difference-in-differences estimation, we address the general identification problem of neighborhood research: voters in different neighborhoods cannot be directly compared because both voters' individual characteristics and those of their neighborhoods differ. We find that a 10 percentage point increase in the out-group neighborhood proportion yields an approximately 0.5 to 2.5 percentage point decrease in the turnout probability. These neighborhood effects persist in non-competitive districts, suggesting that mobilization alone cannot explain their existence.

---

\*We thank Bruce Willsie, the president of Labels & Lists, Inc., for generously providing unlimited access to their database and answering numerous questions. We also thank Bill Guthe and Jonathan Olmsted for geocoding and computational assistance. Eitan Hersh, Marc Meredith, Ali Valenzuela, and seminar participants at Princeton University (Center for the Study of Democratic Politics and Center for the Study of Social Organization), Seoul National University, Vanderbilt University (Center for the Study of Democratic Institutions), and Waseda University provided useful suggestions.

<sup>†</sup>Ph.D. candidate, Department of Politics, Princeton University, Princeton NJ 08544.

<sup>‡</sup>Professor, Department of Politics, Princeton University, Princeton NJ 08544. Phone: 609-258-6601, Email: [kimai@princeton.edu](mailto:kimai@princeton.edu), URL: <http://imai.princeton.edu>

# 1 Introduction

Why do some citizens vote more often than others? This has been a central question in the study of democratic politics. Over the last several decades, much progress has been made in understanding the relationship between voters' turnout and their own characteristics (e.g., Campbell *et al.*, 1960; Wolfinger and Rosenstone, 1980; Aldrich, 1993; Verba *et al.*, 1995). We know, for example, that older, educated, and wealthy people vote often and White voters turn out more frequently than minorities. Nevertheless, relatively little is known about how one's turnout is influenced by the characteristics of other voters around them (but see Gimpel *et al.* (2004); Cho *et al.* (2006); Karpowitz *et al.* (2013) for notable exceptions). Do voters turn out more or less frequently when surrounded by those like them? It is this interaction between voters and their neighbors that we explore in this paper.

To estimate the neighborhood effects on turnout, we analyze unique data sets of over 50 million individual-level voter registration records from Florida, California, and North Carolina. The data sets are provided by Labels & Lists, Inc., a leading national non-partisan firm and the oldest organization in the United States that supplies voter data and related technology to candidates, political parties, pollsters and consultants for use in campaigns. We geocode the addresses of all registered voters in the data and identify their neighborhoods and other voters who live near them. In addition to its sheer size, another advantage of our data is that they consist of voter files collected at two different points in time. This enables us to conduct a panel data analysis for estimating the effects of changes in neighborhood characteristics on turnout over time.

We define each voter's neighbors as those who live nearby. We focus on *small residential neighborhoods* and as our measure use census blocks, which are designed to incorporate geographical, administrative, and other features of typical residential neighborhoods. This allows us to systematically measure neighborhoods for over 50 million voters and estimate neighborhood effects while holding electoral and other confounding factors constant. To be sure, such an objective measure

is likely to differ from its subjective alternatives. For example, Wong *et al.* (2012) asks each voter to identify their neighborhood on a map. While such a self-reported measure is certainly useful in other contexts, its application to our study is likely to induce endogeneity bias because the actual changes in residential neighborhoods can alter voters' subjective definition of their neighborhoods. In addition, collecting self-reported measures for each of over 50 million voters is infeasible.

Across our cross-section and panel analyses, consistent findings emerge. On average, a 10 percentage point increase (decrease) in the out-group proportion of one's neighborhood leads to an approximately 0.5 to 2.5 percentage point decrease (increase) in the probability of voting. This result appears to hold for both partisan and racial neighborhood effects. For example, all else equal, a Democrat who lives in a Republican neighborhood is less likely to vote than if he/she were to live in a Democratic neighborhood. Similarly, a Black voter turns out less frequently when surrounded by White neighbors. We show that these results are found for a variety of electoral environments, geographies, and time periods. The neighborhood effects are larger among those who did not vote in a previous election.

Our findings are consistent with the psychological theory of voter empowerment, which contends that voters turn out more (less) frequently when their neighbors are similar (dissimilar). This theory posits that turnout may increase because of feelings of increased efficacy or trust when voters are surrounded by those who share their views (e.g., Bobo and Gilliam Jr, 1990; Pantoja and Segura, 2003; Voss and Lublin, 2001). Additionally, turnout may decrease when voters have dissimilar neighbors because of their natural desire to avoid conflict and the moderating effect of exposure to opposing view (e.g., Campbell, 2006; Mutz, 2006). On the other hand, our results contradict with the threat theory, which predicts that neighborhood majority voters turn out more when the size of a neighborhood minority group increases because they compete for limited resources and representation (e.g., Key, 1949; Enos, 2010). Indeed, we find little systematic difference in estimated neighborhood effects between neighborhood majority and minority groups. For example, White voters appear to turn out less frequently when the number of non-White

neighbors increases regardless of their neighborhood majority status.

We also consider the mobilization theory of neighborhood effects. Campaigns may target certain neighborhoods where a large number of their potential supporters reside and contact these voters in order to maximize the efficacy of their mobilization strategies (Malchow, 2008; Fraga, 2012). Under this scenario, for example, Democrats in a Democratic neighborhood are more heavily mobilized than their Republican neighbors. Thus, the prediction of the mobilization theory is consistent with our finding that being surrounded by those who are not like you reduces your propensity to vote. However, we find that the neighborhood effects persist even in an uncompetitive electoral environment, suggesting that mobilization alone cannot explain the existence of neighborhood effects.

The fundamental methodological problem in identifying the neighborhood effects on political participation is that different voters live in different neighborhoods. This means that a simple comparison of two voters who reside in different neighborhoods confounds the effects of neighborhood characteristics with those of voter characteristics. We address this identification problem by employing difference-in-differences identification strategies. In our cross-section analysis, we first perform the comparison of voters within each neighborhood and then investigate how these within-neighborhood differences vary as a function of neighborhood characteristics.

While this removes much of the confounding factors specific to each neighborhood, (unobserved) systematic differences across individual voters within the neighborhood may still exist and bias our inference. In our panel analysis, therefore, we analyze how the turnout rate of a voter changes as his/her neighborhood characteristics change over time. This approach allows us to adjust for all time-invariant individual characteristics of voters. We conduct this panel difference-in-differences analysis for those who stayed at the same address in Florida and California.<sup>1</sup> The unique data we analyze make it possible for us to examine how robust our findings are under both

---

<sup>1</sup>Unfortunately, we do not have a panel dataset in North Carolina and hence our panel analysis is performed only for Florida and California.

cross-section and panel identification assumptions.

The rest of the paper proceeds as follows. In Section 2, we discuss both psychological and mobilization theory that may potentially explain how neighborhood characteristics influence individual's political participation. These theories imply different predictions, which we empirically test in the remainder of the paper. In Section 3, we use a cross-section difference-in-difference identification strategy to estimate the neighborhood effects. In Section 4, we conduct panel analysis by linking voters registration records across two different time periods. In Section 5, we conduct further empirical tests to adjudicate between the two competing theories that may explain our main findings. Finally, in Section 6, we given concluding remarks.

## 2 Theories of Neighborhood Effects

Scholars have theorized several ways in which voters' behavior may be influenced by those who live near them. In this section, we consider two main mechanisms, psychological and mobilization-based explanations, through which the demographic characteristics of one's neighbors may affect the probability of his/her voting.

### 2.1 Psychological Theories

We consider two psychological theories that predict how the demographic composition of one's neighborhood affects voters' political behavior. These two theories yield opposite predictions. First, the *empowerment theory* states that voters turn out more frequently when surrounded by people like them (e.g., Campbell *et al.*, 1960; Mutz, 2006; Campbell, 2006; Gimpel *et al.*, 2004; Cho *et al.*, 2006). The theory also implies that when individuals are surrounded by those who posses dissimilar demographic characteristics or differing political views, they are predicted to turn out less often. This depressing effect on the turnout of neighborhood minorities may arise due to the natural desire to avoid conflict, such as encountering neighbors with differing views at a polling place (e.g., Rosenberg, 1954; Verba and Nie, 1972; Huckfeldt, 1979; Mansbridge, 1983;

Eliasoph, 1998; Huckfeldt *et al.*, 2004). An alternative mechanism is that frequent exposure to opposing views makes voters ambivalent and uncertain of their own positions, and as a result they become less expressive (e.g., Feldman and Zaller, 1992; Zaller, 1992; Hochschild, 1993; Green *et al.*, 2000). Finally, several studies have found that voters are more likely to participate when feelings of trust, empowerment, or political efficacy are increased (Pantoja and Segura, 2003; Bobo and Gilliam Jr, 1990; Barreto *et al.*, 2004; Voss and Lublin, 2001). The majority of these results come from measures of coethnic *representation*, however, it could also be the case that empowerment increases when a voter's *neighbors* share her partisanship or ethnicity.

While these different mechanisms are difficult to disentangle empirically, recent studies provide evidence that is consistent with the empowerment theory. For example, Karpowitz *et al.* (2013) analyze the 2010 Cooperative Congressional Election study and find that voters who identify themselves as “neighborhood outliers” are less likely to vote. Similarly, through an analysis of the 2000 election in a sample of battleground counties of Florida, Gimpel *et al.* (2004) find that Republicans living in predominantly democratic neighborhoods were less likely to vote. These results extend beyond partisanship and apply to racial neighborhood effects as well. In a study of Asian American voters, Cho *et al.* (2006) present evidence that living among a large number of co-ethnics leads to high turnout.

Next, we consider the *threat theory* pioneered by V.O. Key (1949) who, more than a half century ago, found that white voters in predominantly black counties turned out at significantly higher rates than whites in predominantly white areas. According to this theory, neighborhood minorities see majority groups as a threat and are compelled to participate in politics in order to compete for limited resources and representation (e.g., Matthews and Prothro, 1963; Bobo, 1983; Carsey, 1995; Spence and McClerking, 2010). The prediction of the threat theory, therefore, is opposite of that of the empowerment theory: voters are more likely to turn out when surrounded by those who are not like them.

A large body of literature has studied the effects of white and black voter interactions in

terms of perceived threats across groups (e.g., Giles and Evans, 1985; Carmines and Stimson, 1989; Giles and Buckner, 1993; Giles and Hertz, 1994). Giles and Buckner (1993), for example, reinvestigate Key's (1949) results in the 1990 Louisiana Senate race and also find that white turnout is significantly higher in counties with larger Black populations. However, these results are based on pooled cross-section data, and Voss (1996) questions their validity by pointing out possible confounders such as income and education levels, which may be correlated with both the ethnic composition of county and its turnout level. Using a natural experiment, Enos (2010) addresses this methodological problem and finds that removal of Chicago public housing, which forced many black voters to move out of these neighborhoods, led to a decrease in white turnout. The author argues that the reduced racial threat contributed to this decreasing turnout rate. The threat theory may also extend to other racial (and non-racial) groups. For example, Enos (2011) presents experimental evidence that in Los Angeles, informing black voters of the turnout rate of nearby predominantly hispanic neighborhoods has a positive effect on turnout. This finding may be explained as the result of increased awareness of a racial threat.

## 2.2 Mobilization Theory

While the psychological theories discussed above focus on the social interactions between voters and their neighbors, we also consider the question of whether or not the strategies of elites, namely campaigns and candidates, lead to neighborhood effects. A large and well-established literature has shown that campaigns have a variety of methods for persuading and mobilizing voters (e.g., Gosnell, 1977; Gerber and Green, 2000; Hillygus, 2005; Arceneaux and Nickerson, 2009; Alvarez *et al.*, 2010). In recent years, campaigns employ increasingly sophisticated data-driven strategies to target only their supporters through mailings, phone calls, or door-to-door canvassing (e.g., Alvarez *et al.*, 2010; Hersh, 2011; Franz, 2013).

We consider the *mobilization theory*, which states that given the budget constraint faced by campaigns, they identify and choose neighborhoods with a large number of core supporters as

	Psychological Theories	Mobilization Theory
Empowerment	Threat	
—	+	—

Table 1: Hypothesized Neighborhood Effects for Psychological and Mobilization Theories. The empowerment theory hypothesizes that voters are less likely to turn out when surrounded by those who are not like them. The hypothesized direction of neighborhood effect is the same under the mobilization theory. In contrast, the threat theory gives the opposite prediction that voters are more likely to turn out when surrounded by those who are not like them.

their primary grounds for mobilization rather than blanketing the entire district with mobilization treatments (Cho *et al.*, 2006; Fraga, 2012). This may provide campaigns with a more cost-effective mobilization strategy than targeting co-partisans in every part of the district (Malchow, 2008; McNamara, 2008). Under this scenario, Democrats in a Democratic neighborhood, for example, are expected to be mobilized heavily when compared to their Republican neighbors. Similarly, Democratic campaigns may target African American (Latino) voters in predominantly African American (Latino) neighborhoods and so these voters turn out at a higher rate than their White neighbors. Thus, the theory predicts that voters are more likely to turn out when surrounded by other voters who are like them.

### 2.3 Summary of Hypotheses and Effect Heterogeneity

Table 1 summarizes the predictions of the three theories discussed above. Among the two psychological theories, the empowerment theory hypothesizes that voters are less likely to turn out when surrounded by those who are not like them. The hypothesized direction of neighborhood effect is the same under the mobilization theory. In contrast, the threat theory gives the opposite prediction that voters are more likely to turn out when surrounded by those who are not like them.

These neighborhood effects are “pure” in a sense that in our analysis we do not consider the possible interaction effects between voters and their electoral environment. For example, scholars have debated whether majority-minority districts, which were created under the Voting Rights Act, influence turnout among racial minority voters (e.g., Gay, 2001; Barreto *et al.*, 2004; Henderson

*et al.*, 2012). More generally, there exist decades of research on how coethnic candidates mobilize minority voters (e.g., Bobo and Gilliam Jr, 1990; Gilliam, 1996; Washington, 2006; Barreto, 2007; McConaughy *et al.*, 2010). While these scholars also rely on psychological and mobilization theories similar to those discussed above, we focus on the estimation of pure neighborhood effects by adjusting for these candidate-voter interaction effects. For example, we hold constant relevant factors such as candidate characteristics and items on ballots.

Finally, we hypothesize that the magnitude of neighborhood effects vary across voters in a systematic manner. Previous work has found that voters with low propensity to turn out are more susceptible to voter mobilization (Gerber and Green, 2000; Nickerson, 2008; Arceneaux and Nickerson, 2009). We argue that under all of the three theories neighborhood effects are expected to be greater among these voters.

## 2.4 Methodological Challenges of Neighborhood Research

Empirical investigation of neighborhood effects must overcome a fundamental methodological challenge. The fact that those who live in different neighborhoods tend to have dissimilar characteristics means that neighborhood effects are typically confounded by the effects of individual characteristics. As a result, studies that rely on small cross-section data make it difficult for researchers to adjust for these confounding factors (e.g., Bobo, 1983; Giles and Hertz, 1994). In addition, because of insufficient data, previous researchers often use counties, city limits, or census tracts, which are quite large geographically as proxies for residential neighborhoods and contain a large number of voters (e.g., Campbell, 2006; Mutz, 2006; Gimpel *et al.*, 2004; Cho *et al.*, 2006).

To address these limitations of previous studies, we geocode over 50 million individual voter registration records in Florida, California, and North Carolina. The sheer size of data allows us to use census blocks as small residential neighborhoods while enabling the precise estimation of neighborhood effects and adjusting for confounding factors. We also analyze unique panel data sets and examine how the changes of voters' neighborhood characteristics over time affect their

turnout. This approach adjusts for unobserved and time-invariant characteristics of individual voters. In what follows, we first describe our cross-section analysis and then present the results from our panel analysis.

### 3 Cross-Section Analysis

In this section, we describe our cross-section analysis and present its results. We begin by introducing our data based on voter registration records and explaining how we validate our measures of partisanship and race. We then explain our identification and modeling strategy before presenting our empirical findings.

#### 3.1 Data

We analyze voter registration data from Florida, California, and North Carolina, which are summarized in Table 2. The data come from Labels & Lists, Inc., a leading non-partisan firm and the oldest organization in the United States that supplies voter data and related technology to candidates, political parties, pollsters and consultants for use in campaigns. The data are based on statewide voter files, which, for every active registered voter in the state, contain the registration information including the voter's birthdate, original registration date, address, district, precinct, and vote history. For each state, the voter files also contain a record of each voter's party registration. Furthermore, Labels & Lists includes the gender, and ethnicity of every voter in each state. In Section 3.2, we discuss the details of these partisanship and race measures and validate their accuracy.

For Florida, Labels & Lists provided us with statewide voter files from May 2004 and July 2012, and we analyze turnout in the 2002 and 2010 elections separately using each file. There are approximately 10 million voters in Florida divided into 25 congressional districts. In California, Labels & Lists has provided us with a statewide file from August 2006 and July 2012 and we analyze turnout in the 2004 and 2010 elections separately from each data set. Slightly more than

	Florida		California		North Carolina
Elections	2002	2010	2004	2010	2010
Acquisition date of voter file	May 2004	July 2012	Aug 2006	July 2012	July 2013
Registered voters, in millions	9.4 (9.3)	10.1 (11.2)	15.9 (16.6)	15.4 (17.3)	5.6 (6.2)
Turnout (%)	52.1 (55.3)	49.6 (48.7)	75.9 (76.0)	60.5 (59.6)	44.8 (43.7)
Congressional districts	25	25	53	53	13
Census blocks (2010 Census)	293,056		383,892		179,879
Partisanship (%)	Republican	38.4 (38.7)	36.1 (36.0)	34.5 (34.7)	30.8 (31.2)
	Democratic	41.7 (42.4)	40.9 (41.1)	43.1 (43.0)	43.6 (44.1)
	Independent	17.4 (18.8)	23.0 (22.8)	17.8 (17.7)	25.6 (20.3)
Race (%)	Black	11.7 (14.6)	13.4 (16.5)	1.7 (7.4)	1.8 (7.2)
	Latino	11.4 (16.8)	16.7 (22.9)	19.7 (32.4)	22.6 (37.6)
	White	60.0 (65.4)	53.6 (57.5)	55.1 (63.4)	51.2 (61.6)

Table 2: Summary of the Voter Registration Records Data Used in Our Analysis. For Florida and California, we analyze two voter files, one new and the other old. In North Carolina, only a recent voter file is available. For each of the two elections we analyze, we compare the number of registered voters and partisanship compositions with the corresponding official figures from the Secretary of State office (presented in parentheses). In addition, we compare the racial compositions of registered voters with those of all state residents taken from the 2000 and 2010 U.S. Census (presented in parentheses).

15 million voters live in California, and these voters are divided into 53 congressional districts. In North Carolina, Labels & Lists provided us with a statewide file from July 2013 and we analyze turnout in the 2010 election. North Carolina has approximately 6 million registered voters divided into 13 congressional districts.

As shown in Table 2, the state-wide summary statistics of our data largely match with the corresponding official statistics released by the Secretary of State office (shown in parentheses). What explain some discrepancies that exist in the total number of registered voters and turnout among registered voters? First, the voter files we use are not obtained immediately after each election when the official statistics are released. In addition, Labels & Lists removes duplicate records, which often occur in the state maintained database, as well as voters who have been classified as inactive by the Secretary of State's office.<sup>2</sup> We also compare the overall partisanship

<sup>2</sup>Inactive voters are voters who have not voted in the past several general elections. The state will send a postcard to the address listed on the voter file requesting the voter confirm her address and registration. If there is

compositions with the corresponding official figures from the Secretary of State office (presented in parentheses) and find that they are reasonably close to each other. For racial compositions, our data exhibit a pattern similar to the one found in the 2000 and 2010 census (shown in parentheses) though these two measures have important differences: (1) our data include only registered voters while the census counts all residents, and (2) they are also collected at different points in time.

### 3.2 Validating Partisanship and Race Measures

We validate our partisanship and race measures based on the Labels & Lists data. First, to validate our partisanship measure, we compare it against the official vote returns at the precinct level. While partisanship and vote choice are not always perfectly correlated, research and exit poll data suggest that it is a powerful predictor of vote choice (e.g., Campbell *et al.*, 1960; New York Times, 2012). In Figure 1, for each precinct in Florida, California, and North Carolina, we plot the proportion of registered Republican voters (our partisanship measure) against the percent of voters casting a ballot for the Republican presidential candidate in the 2008 election (Anscombe and Rodden, 2013).<sup>3</sup> In all three states, the correlation between party registration and election returns is high, considering that our data are obtained four years after the election.<sup>4</sup>

---

no response, the voter is classified as inactive, but not removed from the file (California Secretary of State, 2013; Florida Secretary of State, 2013; Voter Registration Statistics, 2013). In the Florida 2002 voter file, there are four congressional districts for which the data provided by Labels & Lists has much lower turnout than the official results. The remaining 21 districts are similar to the official results (as shown in Table 2). We omit these districts from our analysis because we suspect the data in these four districts are incorrectly coded. However, even if we include them, the results do not change in any substantial way.

<sup>3</sup>The 2012 election data at the precinct level are not yet available.

<sup>4</sup>In fact, as seen in Table 2, Florida and California have roughly the same overall distribution of party registration. However, Florida is consistently a battleground state while California has been a Democratic stronghold. This may be the result of Independents in Florida voting Republican more than Independents in California. If this is the case, we would expect to see a greater deviation from the 45 degree line in Figure 1. This is exactly what we observe. Similarly, in North Carolina, we see most precincts reporting larger Republican vote returns than proportions of Republicans in the precinct, suggesting that Independents in North Carolina are voting mainly for the Republican

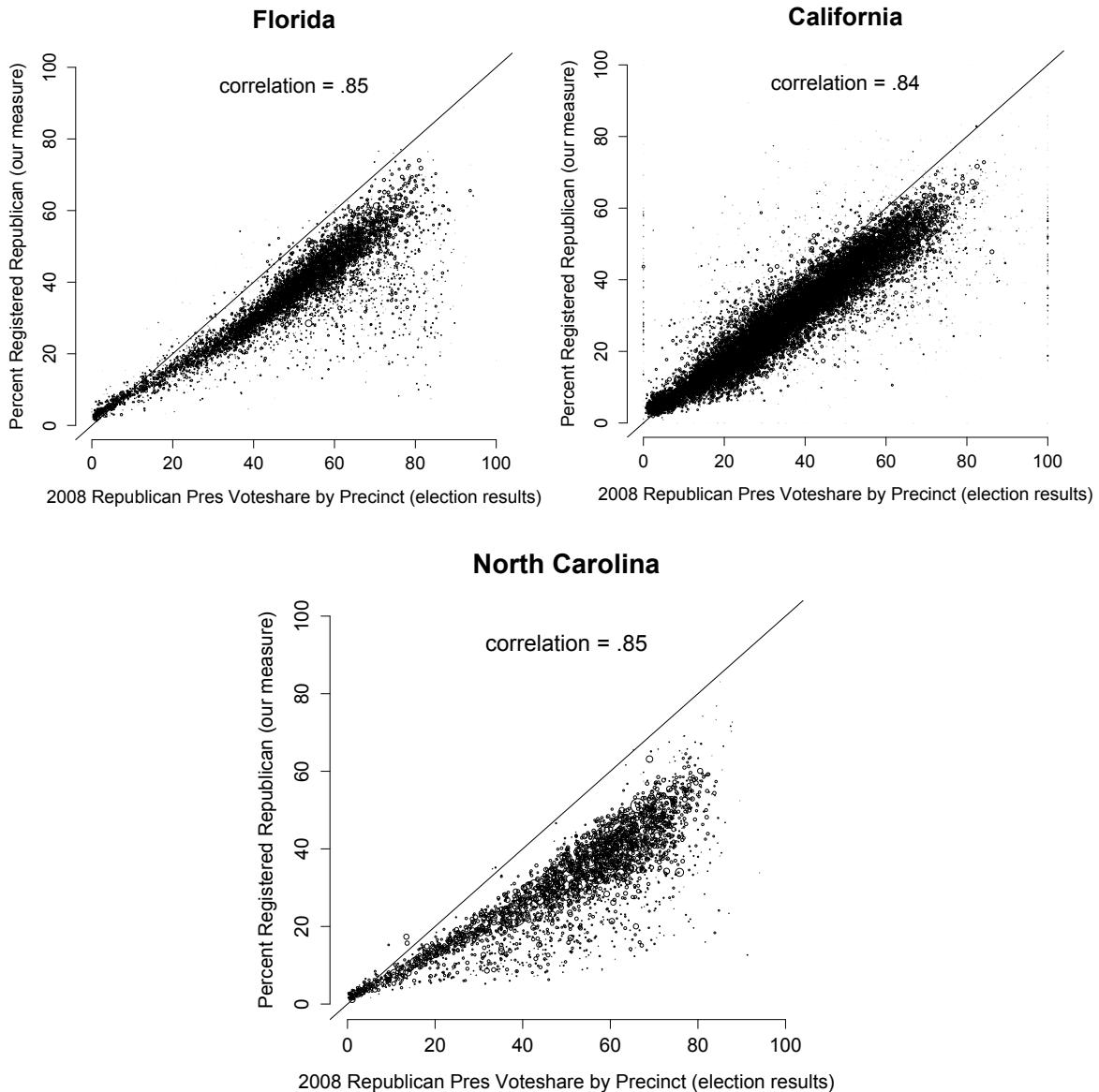


Figure 1: Validating Our Partisanship Measure against the Official Election Results at the Precinct Level. For each precinct in Florida, California, and North Carolina, we plot the proportion of registered Republican voters (our partisanship measure) against the percent of voters casting a ballot for the Republican presidential candidate in the 2008 election. In all three states, the correlation between party registration and election returns is high, suggesting that registration records are a reasonable measure of partisanship.

The result suggests that our measure of partisanship is closely aligned with candidate choices.

Second, we validate our race variable by comparing it against the official statistics at the census block level. Inferring the race of a voter is easier in some states than in others. In Florida and North Carolina, voters are allowed to report ethnicity on their registration form, and Labels & candidate and that possibly some Democrats in North Carolina are voting for the Republican candidate.

Lists primarily relies upon this information. In California, however, this option is not available, and thus determining the ethnicity of voter is much more difficult. In these difficult cases, Labels & Lists makes a prediction based on a variety of information including voters' first and last names and the racial composition of the census blocks they live in. This suggests that our racial composition measure is likely to be more reliable in Florida and North Carolina than in California. In particular, in California, we expect our Black voter variable to be a conservative measure, failing to identify many Black voters while minimizing falsely classifying non-Black voters as Black.

In Figure 2, for each census block in Florida, California and North Carolina, we plot the proportions of registered Black (left column), Latino (middle column), and White (right column) voters against the 2010 official census statistics that measure the proportions of corresponding racial groups. We do not expect perfect correlation because the census include all residents, not just registered voters. Since our analysis concerns neighborhood effects on turnout among registered voters, we focus on the racial composition of registered voters in the neighborhood rather than all residents. In addition, the timing of data collection is not identical. Nevertheless, if our racial composition measures are accurate, we should see reasonably high correlations.

The figure shows that with the exception of Blacks in California, the correlations between two measures are high. In Florida where ethnicity is recorded with voter registration, the two sources of data are highly correlated for each ethnic group (Black: .87, Latino: .79, White: .75). The same is true in North Carolina (Black: .83, Latino: .33, White: .81), with the exception of Latinos. In this case, the correlation between the percent Latino in each block and census percentages for those same blocks is quite low. We suspect that this is not due to poor coding in the voter file since North Carolina voters are asked to report their ethnicity when registering to vote. Rather, we conjecture that this low correlation is due to low registration among Latinos in North Carolina. While 2010 census data show that 7% of the North Carolina population is Latino, only 1.2% of registered voters in North Carolina are Latino according to the North Carolina Secretary of State. Thus, our subsequent analyses include Latinos in North Carolina.

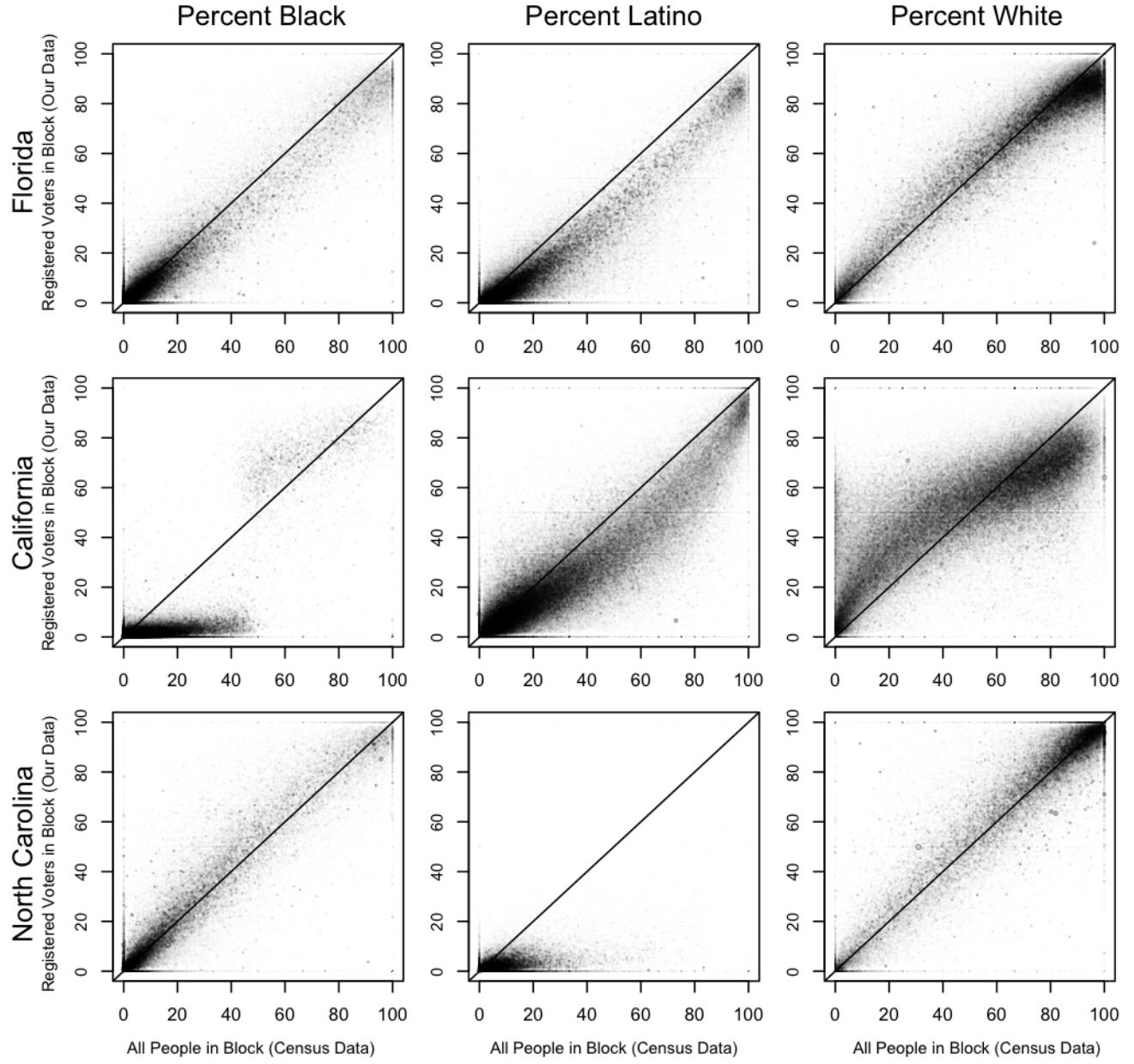


Figure 2: Validating Our Racial Composition Measures against the Official Census Measures at the Census Block Level. For each census block in Florida, California, and North Carolina, we plot the proportions of registered Black (left column), Latino (middle column), and White (right column) voters against the 2010 official census statistics that measure the proportions of corresponding racial groups. With the exception of Blacks in California and Latinos in North Carolina (see text for discussion of these anomalies), the correlations between two measures are high, suggesting that our racial composition measures are reasonably accurate.

On the other hand, in California where ethnicity is not recorded with registration, our measures, which are based on predicting ethnicity using surname and other demographics, correlate well for Latinos, but not for Blacks and Whites (Black: .68, Latino: .77, White: .65). This is due to the fact that traditional Latino surnames are more distinct than African American and White surnames (Barreto *et al.*, 2004), making predictions much easier for Latino voters. However, looking at Figure 2 for California (middle row), we see that despite similarly low correlations (Black: .68, White: .65), the plot for Whites follows the 45 degree line much more closely than the plot for Black voters, suggesting that the measure is noisy but is roughly unbiased. Given the poor classification in California for African Americans, therefore, we exclude them from our subsequent analyses.

### 3.3 Census Blocks as Small Residential Neighborhoods

We define each voter's neighbors as those who live nearby. We focus on *small residential neighborhoods* and use census blocks as our measure. We choose not to rely on subjective measures such as a self-reported neighborhood proposed by Wong *et al.* (2012). While such measures can be useful in other contexts, their application in our study is likely to induce a potential endogeneity problem because voters' perception of neighborhood is likely to be influenced by the actual changes in their residential neighborhood. The massive size of our data set also makes it impossible to collect such measures by survey. We emphasize that census blocks represent small residential neighborhoods. Indeed, in all cases shown in Wong *et al.* (2012), the census block of a voter is contained in his/her self-reported neighborhood boundaries.

There are additional reasons for our choice of census blocks as residential neighborhoods. First, unlike other definitions such as those solely based on distance, census blocks incorporate various geographical and administrative features that often characterize one's small residential neighborhoods. According to U.S. Census Bureau (2010), census blocks are defined as “areas bounded on all sides by visible features, such as streets, roads, streams, and railroad tracks, and

by invisible boundaries, such as city, town, township, and county limits” (p. A-10). This means, for example, all voters in a census block belong to the same electoral precinct. Therefore, they vote at the same polling location, cast ballots for the same offices, and choose among the same set of candidates (p. A-27). This allows us to hold the electoral environment constant. Second, voters in the same census block are assigned to the same public school districts (p. A-27), and tend to be homogeneous in terms of their demographics. This enables better adjustment for unobserved confounders. Finally, census blocks are the smallest administrative units where the demographic information is available from the Census Bureau. As shown in Section 3.2, we exploit this fact to validate the key individual-level measures used in our analysis.

We geocode every voter in our voter files by calculating the latitude and longitude from their addresses. We use Esri’s ArcGIS software to transform addresses into latitude and longitude coordinates.<sup>5</sup> We then assign each voter to the appropriate census block using the 2010 Census shape files. As shown in Table 2, in Florida, there are approximately 300 thousand census blocks. In California there are roughly 380 thousand blocks, and in North Carolina, there are nearly 180 thousand blocks. Census blocks represent relatively small neighborhoods, and on average one block contains fewer than 100 registered voters. Figure 12 in Appendix presents the distribution of the number of registered voters per census block by congressional district for Florida, California, and North Carolina. We see that the distribution of census block size is similar across congressional districts and states. Other geographical units such as census block groups and census tracts are much larger, exceeding typical residential neighborhoods in their size and containing a large number of heterogenous voters.

Once all registered voters are assigned to unique census blocks, we calculate the percent of registered voters in a block that belong to each party and racial group. These measures of parti-

---

<sup>5</sup>Specifically, we use Esri’s ArcMap 10.0 Service Pack 5. The addresses were geocoded against the ArcLogistics Route 2011 dataset TTNA2011\_R1. The Locator USA Geocoding Service with Alternate Address (included with ArcLogistics Route) was used to generate matches.

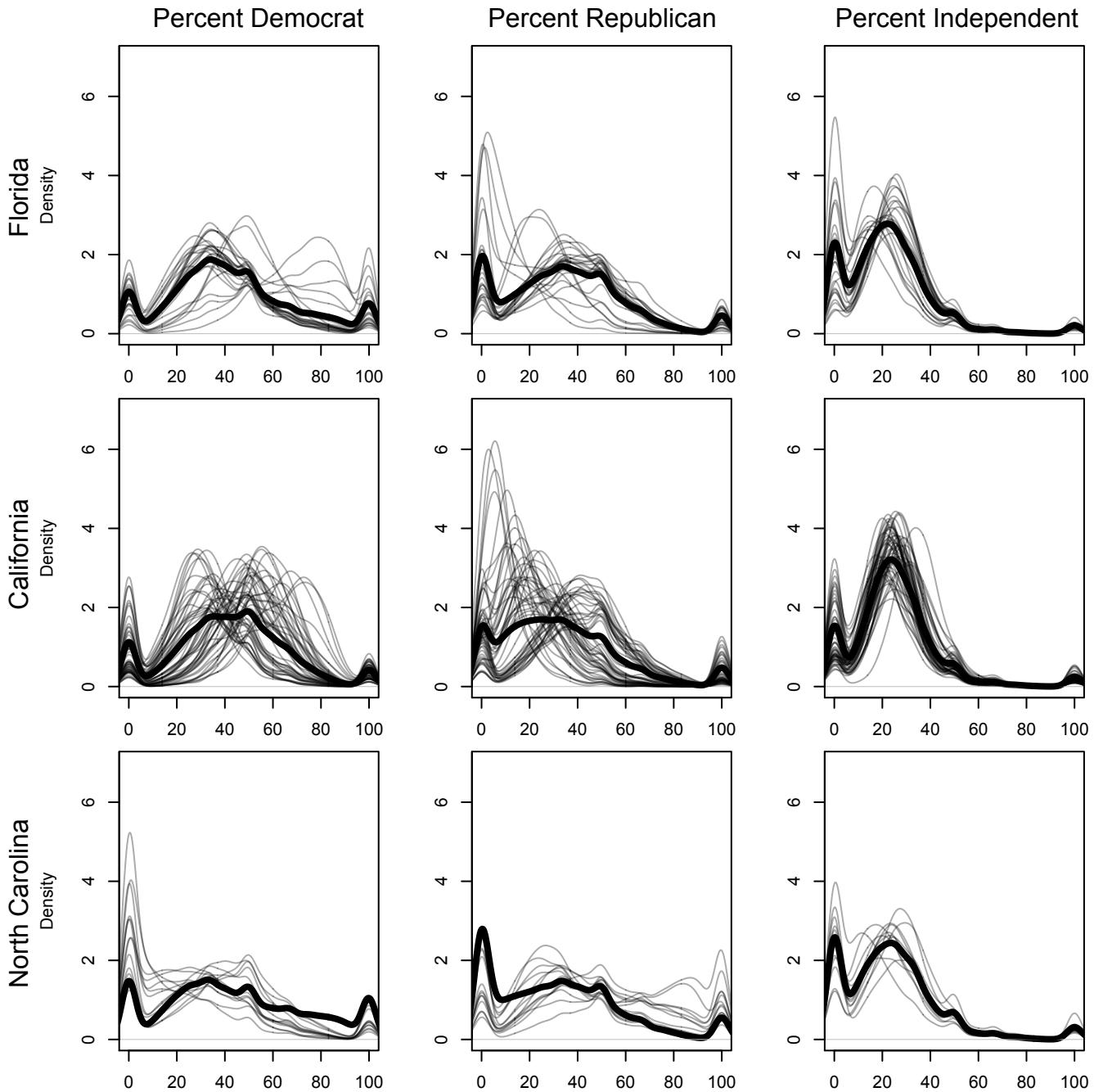


Figure 3: Distributions of Partisan Composition Among Registered Voters in Census Blocks by Congressional District. For each congressional district in Florida, California, and North Carolina, we plot the distribution of the percentages of Democratic (left column), Republican (middle column), and Independent (right column) voters contained in a census block (thin lines). The thick line represents the partisanship distribution for the entire state. The figure shows that there is large variation in the partisanship across census blocks and congressional districts.

sanship and racial composition form the basis for our analysis of neighborhood effects. Figure 3 presents the distribution of partisanship of census blocks by congressional district (thin lines) as well as the distribution for the entire state (thick lines). The figure shows that there is a large variation in the partisanship across census blocks as well as congressional districts. For example, some blocks contain nearly all Republicans, nearly all Democrats, whereas others contain roughly equal proportions of partisan and independent voters. There is also variation across congressional districts.

Figure 4 displays the distribution of racial composition among registered voters in census blocks by congressional district. The variation in racial composition is less obvious than that in partisanship both across census blocks and congressional districts. Nevertheless, a small number of congressional districts contain a significant number of predominantly Black or Latino neighborhoods. In addition, while many census blocks have a large proportion of White voters, some blocks contain majorities of Black or Latino voters. These variations in partisanship and racial composition enable us to identify the neighborhood effects on turnout.

### 3.4 Identification and Modeling Strategies

Using the data described above, we test three theories set forth in Section 2. As discussed in Section 2.4, the fundamental methodological problem of neighborhood research is that people who live in different neighborhoods are dissimilar in their individual characteristics. This means that if we simply compare voters across different neighborhoods our estimates of neighborhood effects are confounded by the effects of individual characteristics. To overcome this methodological challenge, we consider difference-in-differences strategies in both cross-section and panel settings.

Figure 5 illustrates the difference-in-differences identification strategy for our cross-section analysis. In panel (a), we have a Democratic neighborhood where the majority of registered voters are Democrats (blue balls). In panel (b), on the other hand, we have a Republican neighborhood where the majority of registered voters are Republicans (red balls). The simple comparison of turnout

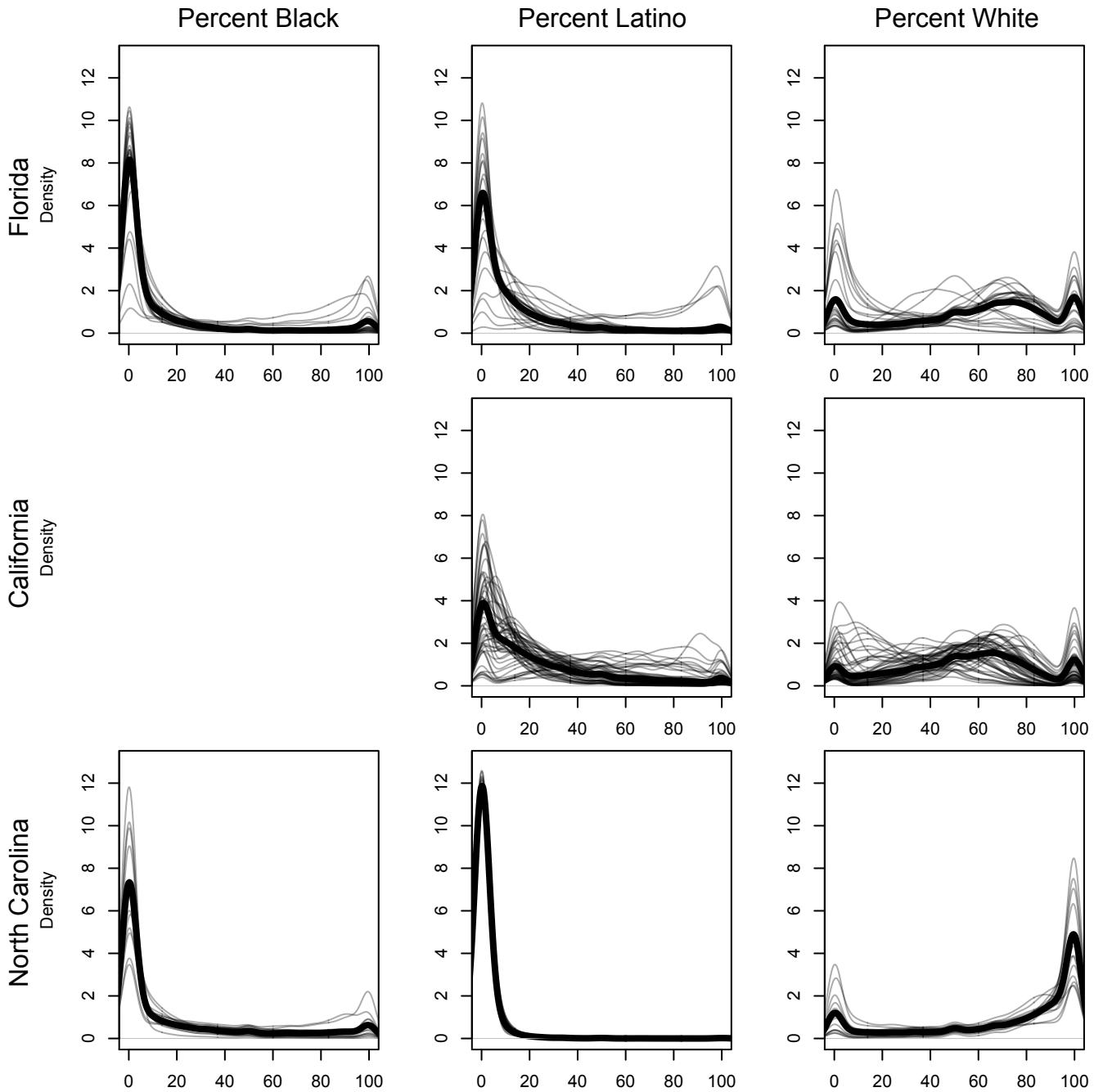


Figure 4: Distributions of Racial Composition Among Registered Voters in Census Blocks by Congressional District. For each congressional district in Florida, California, and North Carolina, we plot the distribution of the percentages of Black (left column), Latino (middle column), and White (right column) voters contained in a census block (thin lines). The thick line represents the partisanship distribution for the entire state. While there are many blocks with a small percentage of minority voters, other blocks contain majorities of Black and Latino voters. As discussed in Section 3.2, we do not analyze Black voters in California.

(a) Democratic neighborhood      (b) Republican neighborhood

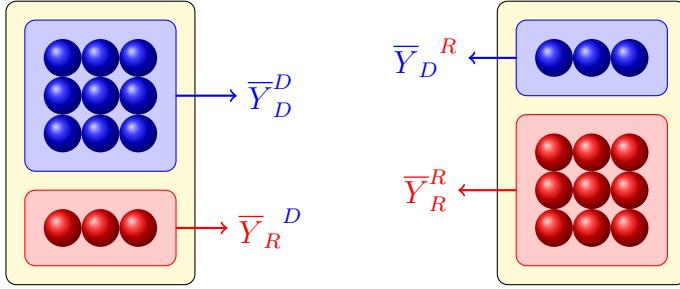


Figure 5: Difference-in-Differences Identification Strategy for Cross-Section Analysis. The figure depicts two neighborhoods, one Democratic (panel (a)) and the other Republican (panel (b)). For each neighborhood, we separately compute the difference in turnout rate between Republicans (red balls) and Democrats (blue balls), i.e.,  $\Delta^D = \bar{Y}_R^D - \bar{Y}_D^D$  for the Democratic neighborhood and  $\Delta^R = \bar{Y}_R^R - \bar{Y}_D^R$  for the Republican neighborhood. We then examine how the difference of these two quantities, i.e.,  $\Delta = \Delta^D - \Delta^R$ , varies as a function of neighborhood partisan compositions. The same identification strategy is also applied to the estimation of racial neighborhood effects.

rates among Democrats (Republicans) between the two neighborhoods, i.e.,  $\bar{Y}_D^D - \bar{Y}_D^R$  ( $\bar{Y}_R^D - \bar{Y}_R^R$ ), is unlikely to yield valid estimates of neighborhood effects because Democrats (Republicans) who live in a Democratic neighborhood are likely to differ in their individual characteristics from those Democrats (Republicans) who live in a Republican neighborhood. Thus, we cannot differentiate the effects of neighborhood characteristics from those of individual characteristics.

In the difference-in-differences framework, we first compare the turnout rate of Republicans and that of Democrats within each neighborhood, i.e.,  $\Delta^D = \bar{Y}_R^D - \bar{Y}_D^D$  and  $\Delta^R = \bar{Y}_R^R - \bar{Y}_D^R$  for Democratic and Republican neighborhoods, respectively. We then consider whether or not this within-neighborhood difference, i.e.,  $\Delta = \Delta^D - \Delta^R$ , varies as a function of neighborhood partisanship composition. The idea is that the within-neighborhood comparison will remove any observed or unobserved characteristics common to the voters who live in the same neighborhood and as a result it provides a better opportunity to isolate the effects of neighborhood characteristics.

This identification strategy enables a direct test of the theoretical predictions summarized in Table 1 while effectively adjusting for possible confounding factors. For example, we would expect this difference  $\Delta$  to be negative under the empowerment theory because, all else equal,

Republican voters in a Democratic neighborhood should feel disempowered. In contrast, under the threat theory, the prediction would be opposite: Republican voters in a Democratic neighborhood should feel compelled to express their opinion through the act of voting.

We base our estimation of the neighborhood effects on the linear probability model with fixed effects. We choose this model for two reasons. First, as shown by Imai and Kim (2012), under certain assumptions, the linear fixed effects model corresponds to the within-group comparison between the treated and control groups (i.e., Democrats and Republicans) where groups are defined by fixed effects (i.e., census blocks). Thus, the model directly operationalizes our identification strategy. Second, in our analysis, the number of fixed effects is quite large (because the number of census blocks is large), and thus the non-linear fixed effects models lead to the possible incidental parameter problem and computational difficulty. In contrast, the desirable statistical properties and an efficient computation algorithm are available for linear fixed effects models even when the number of fixed effects is large (we use an R package `wfe` (Kim and Imai, 2012)).

Specifically, for estimating the neighborhood effects for Democrats, we fit the following model to a subset of the data containing Democratic and Republican voters alone for each congressional district in each election,

$$Y_i = \alpha_{\text{group}[i]}^D + \beta^D \text{Dem}_i + \gamma^D \text{Dem}_i \times \overline{\text{Rep}}_{\text{block}[i]} + \delta_1^D \text{age}_i + \delta_2^D \text{age}_i^2 + \epsilon_i^D \quad (1)$$

where  $Y_i$  is the indicator variable for turnout,  $\text{Dem}_i$  is the indicator variable for being a Democrat,  $\overline{\text{Rep}}_{\text{block}[i]}$  is the proportion of Republican voters in voter  $i$ 's census block,  $\text{age}_i$  is the age of voter  $i$ , and finally  $\alpha_{\text{group}[i]}^D$  represents the fixed effects based on the full interaction of census block, gender (male or female), and race (Black, Latino, White, Asian, or others). This specification implies that we are comparing Democrats and Republicans who not only live in the same neighborhood but also have the same gender and race.

A similar modeling strategy is employed for estimating other neighborhood effects. For exam-

ple, for the neighborhood effects on Black voters, we fit the following model to the entire data,

$$Y_i = \alpha_{\text{group}[i]}^B + \beta^B \text{Black}_i + \gamma^B \text{Black}_i \times \overline{\text{Non-Black}}_{\text{block}[i]} + \delta_1^B \text{age}_i + \delta_2^B \text{age}_i^2 + \epsilon_i^B \quad (2)$$

where  $\text{Black}_i$  is the indicator variable for being a Black voter,  $\overline{\text{Non-Black}}_{\text{block}[i]}$  represents the proportion of non-Black voters in voter  $i$ 's census block, and finally  $\alpha_{\text{group}[i]}^B$  in this case represents the fixed effects based on the full interaction of census block, gender (male or female), and partisanship (Democrats, Republicans, and Independents). Thus, we compare Blacks and Non-Blacks who live in the same neighborhood and share the same gender and partisanship.

Under these model specifications, the coefficients  $\gamma$  for the interaction terms can be interpreted as the neighborhood effects of interest, representing how the proportion of out-group in your neighborhood influences your turnout. Specifically,  $\gamma$  represents percentage point increase in turnout probability when the proportion of out-group increases by one percentage point. For example,  $\gamma^D$  corresponds to the partisan neighborhood effect for Democrats while  $\gamma^B$  represents the racial neighborhood effect for Blacks. Thus, the sign of  $\gamma$  directly corresponds to those of the theoretical predictions presented in Table 1.

### 3.5 Empirical Results

**Average Effects.** We now present the results of our cross-section analysis. Figure 6 presents the estimated average neighborhood effects. As mentioned earlier, we fit a separate model for estimating each of partisan and racial neighborhood effects. Additionally, we analyze each election and congressional district separately in order to further account for electoral and district-level factors. We then present the average effect among all congressional districts for each state. We present the results separately by state to account for the possibility of difference in neighborhood effects by state. For example, as Southern states, Florida and North Carolina have unique racial and partisan histories that could influence neighborhood effects Key (1949). The first point (“FL”: solid triangles) in each column in Figure 6 shows the results for Florida. The 95% confidence intervals are also plotted though they are too short to be visible due to an extremely large sample

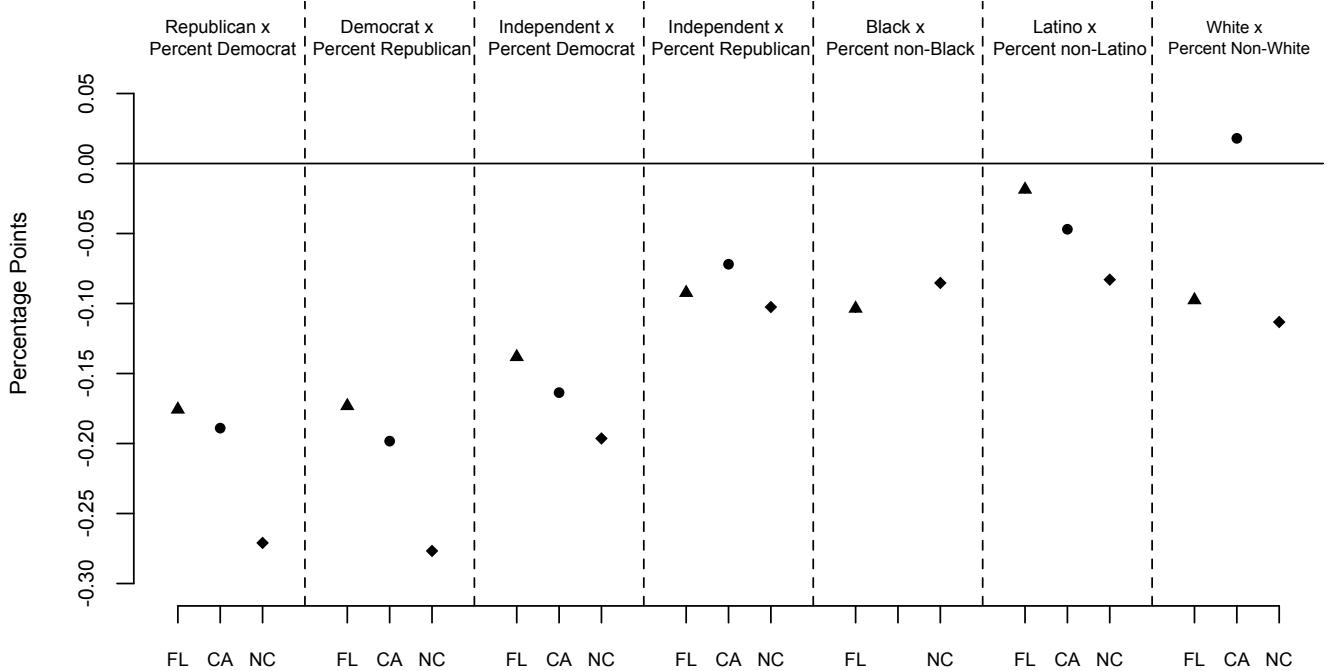


Figure 6: Estimated Average Neighborhood Effects from the Cross-Section Analysis. The results are presented for Florida (“FL”; solid triangles), California (“CA”; solid circles), and North Carolina (“NC”; solid diamonds). The 95% confidence intervals are plotted but are too short to be visible. On average, a 10 percentage point increase (decrease) in the out-group in a neighborhood leads to approximately 0.5 to 2.5 percentage point decrease (increase) in the probably of turning out. These effects exist for each partisan group and each racial group, except for Whites in California. We only show the results for Blacks in Florida and North Carolina due to the inaccuracy of the African American ethnicity measure for California.

size. The second (“CA”: solid circles) and third (“NC”: solid diamonds) points in each column correspond to the results for California and North Carolina averaged over congressional districts and elections.

The first four columns display the partisan neighborhood effects. In all three cases, we find that on average a 10 percentage point increase (decrease) in the partisan out-group in a voter’s neighborhood leads to an approximately 1 to 2.5 percentage point decrease (increase) in that voter’s probability of turning out. These findings hold for Republicans, Democrats, and independent voters. The remaining three columns show the racial neighborhood effects. The effect is strongest among Black and White voters in Florida and North Carolina, with a 10 percentage point increase in non-Blacks (or non-Whites in the case of White voters) in the neighborhood leading a Black (White) voter to be nearly 1 percentage point less likely to vote. This predicted decrease in turnout probability is smaller for Latinos in Florida and the effect is slightly positive

for Whites in California.

**Effect Heterogeneity.** We also separately examine those who voted in a previous election and those who did not. Previous get-out-the-vote studies have found greater treatment effects among those who are less likely to turnout. We hypothesize that neighborhood effects are greater among those who have not established a regular pattern of voting. Thus, we further restrict our comparison of voters to those who live in the same neighborhood, have the same turnout record in a previous election, and belong to the same gender and race (partisanship) category.<sup>6</sup>

We use the old voter files in California and Florida and match voters to their turnout in the 2004 and 2002 elections (2002 and 2000) in California (Florida). In North Carolina we do not have an old voter file, but the current voter file contains previous vote histories, which we use to code if the voter cast a ballot.<sup>7</sup> See Section 4 for a detailed discussion of our matching procedure. Figure 15 in the Appendix displays the results. Consistent with the previous findings, the neighborhood effect is greater, i.e., more negative, among those who did not vote in the previous election (open symbols). Among Latinos, racial neighborhood effects may even be positive for those who voted in a previous election. Thus, the neighborhood effects appear to be most noticeable among those who have not already established a regular habit of voting.

**Summary.** The consistently negative neighborhood effects support some of the theories outlined in Section 2 while they contradict others. Among the psychological theories we consider, the *threat* theory hypothesized that increase in the size of the out-group in a voter's neighborhood would lead to a boost in his/her turnout. Our results are inconsistent with this prediction. Rather, they align with the prediction of the *empowerment* theory. Voters are more (less) likely to vote

---

<sup>6</sup>This contrasts with our difference-in-differences analysis, which does not adjust for the lagged dependent variable. Note that neither approach is more general than the other. They rely on different identification assumptions.

<sup>7</sup>The previous vote histories are not as accurate as matching across voter files since a voter who was previously ineligible to vote (for example due to either not living in the state or being too young) will be coded as having not voted, while it was the case that the voter was in fact not capable of voting in the election in question.

	Florida	California
Voter files	2004, 2012	2006, 2012
Elections compared	2000 and 2008 2002 and 2012	2002 and 2010 2004 and 2008
Percent records matched	40	44
Percent voters moved (among matched)	33	30
Number of voters analyzed, in millions	2.4	4.4

Table 3: Summary of Panel Data for Florida and California. For the panel analysis, we link registration records between the two voter files and then compare turnout between two elections. We use exact matching based on the voter’s last name, gender, and birthdate for identifying the same voters in different voter files. Among those voters who are matched, we focus on those who stayed at the same address, but whose neighborhood characteristics changed.

when surrounded by people who are politically or racially similar (dissimilar) to them. Our results support the mobilization theory that campaigns strategically target mobilization efforts towards their supporters in areas where those supporters are more concentrated.

## 4 Panel Analysis

Our cross-section analysis described above adjust for the confounders that are common among voters who live in the same neighborhood. Given that census blocks define a small neighborhood, we expect a high degree of homogeneity among voters in the same block. Nevertheless, it is possible that remaining heterogeneity among voters in the same neighborhood biases our results. For example, White voters may be systematically different from their Black neighbors in aspects other than their ethnicity. To address this concern, we conduct a panel analysis by examining how the changes in the characteristics of a voter’s neighborhood over time affect his/her voting probability. Since these are relatively short-term neighborhood changes, their effects substantively differ from the neighborhood effects estimated in our cross-section analysis.

### 4.1 Data

Separately for Florida and California, we create a panel dataset by linking the registration records from the voter files obtained in the early 2000’s to those from the 2012 voter files. For North

Carolina, we do not have multiple voter files and hence cannot conduct panel analysis. Among the voters who exist in both voter files and did not move their residence between two time periods, we calculate the changes in their neighborhood characteristics between the two time periods. Table 3 summarizes our panel data set. In Florida, we link records between the 2004 and 2012 registration files and compare changes in turnout between the 2000 and 2008 presidential elections as well as between the 2002 and 2010 midterm elections. In California, we link records between the 2006 and 2012 registration files and compare changes in turnout between the 2004 and 2008 presidential elections as well as between the 2002 and 2010 midterm elections.

For linking records across two time periods in each state, we only consider exact matches based on a voter's last name, gender, and birthdate. In the case that there were duplicate matches, we selected randomly among the multiple matches. However, fewer than 0.1% of all records had duplicate matches. Using this criteria, in the 2012 files we are able to match 40% of all registration records in Florida and 44% in California. It is difficult to judge whether these matched rates are reasonable. The unmatched voters include those who have recently turned 18, those who have moved out of the state, those who moved into the state, and those who have deceased or become inactive. In addition, our procedure fails to match voters who changed their last name, due to marriage for example, between two time periods. It is also possible that clerical errors in transcribing a voter's name from the registration card to the electronic file sometimes leads to some voters' names appearing differently in the two different files. Despite these shortcomings, we use our conservative matching criteria in order to avoid as many false record linkages as possible.

Using the matched records, we determine whether or not each voter has moved to a new address or remained at the same address between the two time periods. Those voters who stayed at the same address are the ones included in our panel analysis while the voters who moved to another address are excluded. Among the matched records, we find that approximately 30% of registered voters in each state change their address more than 100 meters. We observe that the median move distance is approximately 30 kilometers. However, there are a number of people who move

significant distances within either state. Our analysis focuses on the 70% of matched voters who do not change their addresses during this time.

As a result, our panel analysis focuses on a total of 2.4 and 4.4 million voters in Florida and California, respectively, who did not move to another address between the two time periods. Such voters consist of approximately 25% and 29% of all registered voters. Given that one's decision to move is not random, the results of our panel analysis are only applicable to this subset of the population. Thus, the results of the panel analysis need to be interpreted with caution when compared with the results of our cross-section analysis (in addition to the fact that the neighborhood effects in our panel analysis represent short-term effects).

## 4.2 Changes in Neighborhood Characteristics

We next examine the changes in neighborhood characteristics over time for Florida and California. Figure 7 shows distributions of neighborhood partisanship change for each voter by congressional district. While some neighborhoods underwent dramatic change in their partisan composition, others experienced little change during the period covered in our panel analysis. Interestingly, most distributions are symmetric around zero, suggesting that for some voters their neighborhoods have become more Republican while the opposite hold for other voters.

Figure 8 shows the distribution of changes in neighborhood racial composition for each voter by congressional district. When compared to the changes in neighborhood partisanship in Figure 7, neighborhood racial composition did not change as much. In addition, some neighborhoods in Florida had a relatively large increase of hispanic population during this time period. In Appendix, we plot the neighborhood racial and partisanship compositions at time  $t$  against those measured at time  $t + 1$  for each census block (see Figures 13 and 14). We observe a wide range of variation in how each neighborhood has changed over time. Some neighborhoods have undergone dramatic changes in partisanship and/or racial compositions whereas others have remained almost identical in those dimensions.

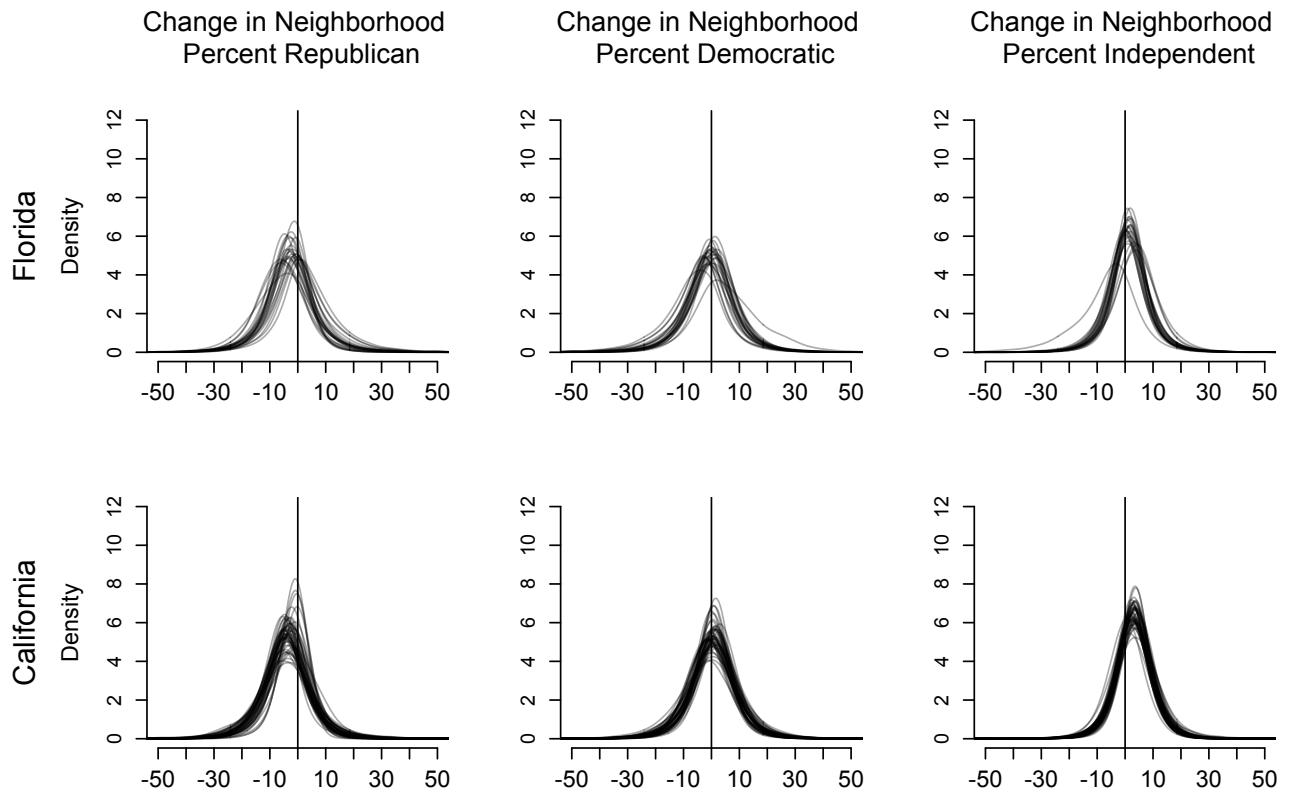


Figure 7: Changes in Neighborhood Partisanship. Each line shows the distribution of changes in neighborhood partisanship (percent Republican in the left column, percent Democratic in the middle column, and percent independent in the right column) by congressional district for Florida (upper panel) and California (lower panel). While some neighborhoods underwent dramatic change in their partisan composition, others experienced little change during the period covered in our panel analysis.

### 4.3 Identification and Modeling Strategies

We apply another difference-in-differences identification strategy in our panel analysis. The panel analysis allows us to examine within-voter changes over time, eliminating the confounding due to (time-invariant) observed and unobserved factors that are specific to each voter. Figure 9 illustrates our identification strategy, depicting the analysis of those voters who stay at the same address and therefore live in the same neighborhoods. Even if a voter does not move, however, his/her neighborhood characteristics may change. We focus on those who stayed at the same address because which neighborhood one moves to is likely to be endogenous and introduce selection bias.

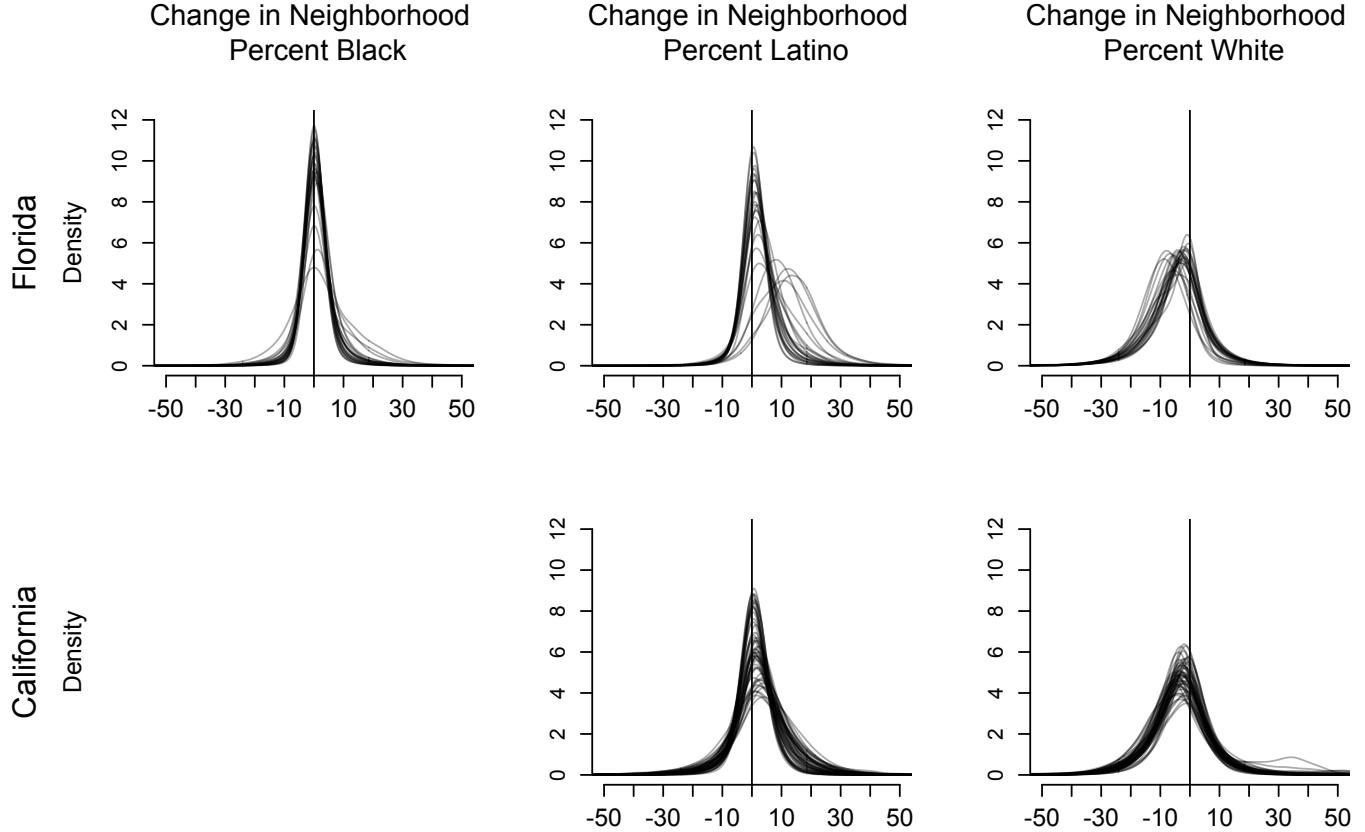


Figure 8: Changes in Neighborhood Racial Composition. Each line shows the distribution of changes in neighborhood racial composition (percent Black in the left column, percent Latino in the middle column, and percent White in the right column) by congressional district for Florida (upper panel) and California (lower panel). When compared to the changes in neighborhood partisanship (Figure 7), neighborhood racial composition did not change as dramatically.

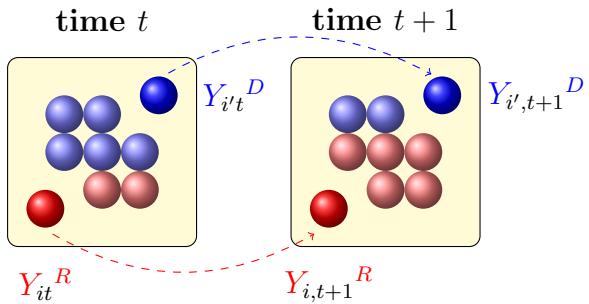


Figure 9: Difference-in-Differences Identification Strategy for Panel Data Analysis. The figure illustrates three identification strategies employed for our panel data analysis. We compare a Republican voter  $i$  and a Democrat  $i'$  who live in the same neighborhood. They do not move between two time periods but during this time their neighborhood becomes more Republican. The difference-in-differences estimator is  $(Y_{i,t+1}^R - Y_{i,t}^R) - (Y_{i',t+1}^D - Y_{i',t}^D)$ . This identification strategy is also applied to the estimation of racial neighborhood effects.

In this figure, a Republican voter  $i$  and a Democrat  $i'$  live in the same neighborhood and do not move between time  $t$  and time  $t + 1$ . During this time period, their neighborhood becomes more Republican. The difference-in-differences estimator is given by  $(Y_{i,t+1}^R - Y_{it}^R) - (Y_{i',t+1}^D - Y_{i't}^D)$ , which compares within-voter differences in turnout between the two voters. For testing the theories of neighborhood effects, we characterize this quantity as a function of the changes in neighborhood characteristics. For example, the empowerment theory would predict this quantity to be negative because, all else equal, a Republican voter whose neighborhood became more Republican would vote more frequently while his/her Democratic neighbor would become less likely to vote.

To estimate the neighborhood effects for those who did not move, we use the first differencing linear probability model with fixed effects. For example, for the partisan neighborhood effects among Democrats, we fit the following model to a subset of Democratic and Republican voters,

$$\begin{aligned} Y_{i,t+1} - Y_{it} = & \alpha_{\text{group}[i]}^D + \beta^D \text{Dem}_i + \gamma^D \text{Dem}_i \times (\overline{\text{Rep}}_{\text{block}[i,t+1]} - \overline{\text{Rep}}_{\text{block}[i,t]}) \\ & + \delta_1^D \text{age}_i + \delta_2^D \text{age}_i^2 + \eta_i^D \end{aligned} \quad (3)$$

where  $\overline{\text{Rep}}_{\text{block}[it]}$  is the proportion of Republican voters at time  $t$  in voter  $i$ 's neighborhood, and  $\alpha_{\text{group}[i]}^D$  represents the fixed effects based on the full interaction of census block, gender, and race. Similar to the cross-section analysis, this model specification implies that we compare Democratic and Republican voters who not only stayed in the same neighborhood but also belong to the same gender and race groups. The models used for racial neighborhood effects are similar. For example, we fit the following model for estimating the neighborhood effect among Black voters,

$$\begin{aligned} Y_{i,t+1} - Y_{it} = & \alpha_{\text{group}[i]}^B + \beta^B \text{Black}_i + \gamma^B \text{Black}_i \times (\overline{\text{Non-Black}}_{\text{block}[i,t+1]} - \overline{\text{Non-Black}}_{\text{block}[i,t]}) \\ & + \delta_1^B \text{age}_i + \delta_2^B \text{age}_i^2 + \eta_i^B \end{aligned} \quad (4)$$

where  $\alpha_{\text{group}[i]}^B$  represents the fixed effects based on the full interaction of census block, gender, and partisanship (measured at time  $t$ ). Thus, the comparison between Black and non-Black voters is made by focusing on the voters who stayed in the same neighborhood and share the same gender and partisanship at time  $t$ .

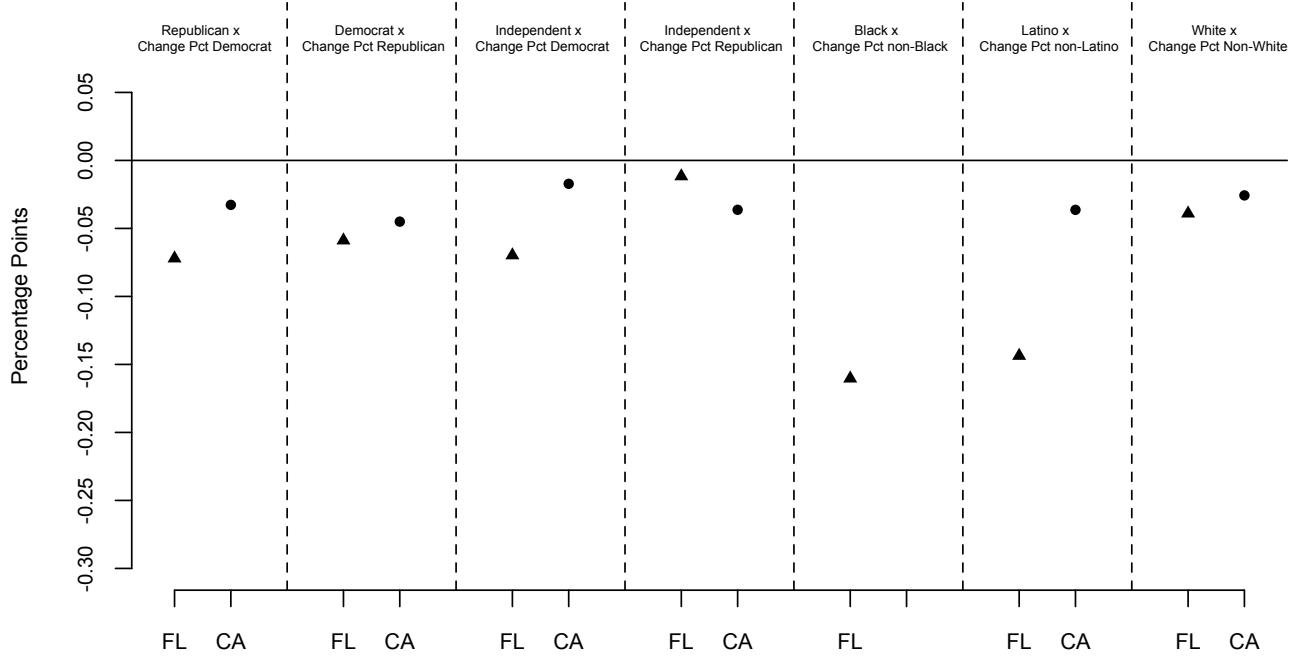


Figure 10: Estimated Average Neighborhood Effects under Panel Identification Strategy. We estimate each neighborhood effect by focusing on the comparison among those voters who stayed at the same address. The results are presented for Florida (“FL”; solid triangles) and California (“CA”; solid circles) separately. The 95% confidence intervals are also shown, but are too small to be seen. A negative estimate implies that an increase in the size of the out-group in a voter’s neighborhood leads to a decrease in the probability of voting. The magnitude and direction of these results is generally consistent with the cross section results shown in Figure 6. On average, a 10 percentage point increase in the size of the out group leads to an approximately 0.5 to 1.5 percentage point decrease in the probability of voting. We only show the results for Blacks in Florida due to the high level of noise in the California results.

Under these models, we interpret the coefficients  $\gamma$  for the changes in neighborhood characteristics as percentage point increase in turnout when the proportion of out-group increases by one percentage point. Thus, these coefficients map directly to the theoretical predictions summarized in Table 1. As in the cross-section analysis, we fit each model by congressional district in order to hold the electoral environment constant.

#### 4.4 Empirical Results

**Average Effects.** Figure 10 presents the results of our panel analysis. Specifically, we estimate each neighborhood effect by focusing on the comparison among those voters who stayed at the same address while their neighborhood may have changed. In terms of their direction of estimated

effects, the results are in agreement with those of the cross-section analysis shown in Figure 6. However, the effect sizes are generally smaller, perhaps due to the fact that these are short term effects.

On average, we find that an increase over time in the proportion of the out-group in a voter’s neighborhood leads to a decrease in the probability of turning out. The effect appears to be the greatest for Black voters in Florida, where a 10 percentage point increase in the size of the out group leads to an approximately 2 percentage point decrease in the probability of voting. The magnitude of these effects, however, is generally smaller than that of cross-section analysis. The difference in magnitude may reflect the fact that panel analysis focus on relatively short-term changes in neighborhood characteristics that occur within a decade. In contrast, the findings of cross-section analysis can result from a long-term exposure to certain neighborhood characteristics.

**Effect Heterogeneity.** We alter the dependent variable of the model to be a dichotomous variable indicating whether or not the voter participated in the 2010 or 2008 election, leaving everything else the same as in the previous difference-in-difference models described above. We then subset the data and separately estimate the neighborhood effect among those that voted in the previous election (2004 and 2000 in California, and 2002 and 2000 in Florida) and those that did not. Unlike the difference-in-differences analysis, this model further adjusts for the previous vote and estimates the heterogenous neighborhood effects with respect to this variable (see footnote 6). The results are shown in Figure 16 of Appendix. Similar to the results of the cross-sectional analysis (see Figure 15), the estimated neighborhood effects are greater among those that did not vote in the previous election (open symbols) than those who did (solid symbols). We also note that in California some of the estimated neighborhood partisanship effects are no longer negative even among those who did not vote. In contrast, the results in Florida are consistent with other findings presented in this paper. We leave the exploration of neighborhood effect heterogeneity across states to future work.

**Summary.** As with the cross-section analysis, the panel analysis in general provides evidence that is largely consistent with the empowerment and mobilization theories, while casting doubt on the threat theory. These results appear to be fairly robust as shown by the congruence of the cross section results with the panel results.

## 5 Further Empirical Testing of Theories

Sections 3 and 4 present the empirical evidence consistent with the predictions of both the empowerment and mobilization theories. In contrast, our findings contradict with the threat theory. In this section, we conduct additional empirical testing in order to determine whether the empowerment or mobilization theory is more consistent with our data.

Both the empowerment and mobilization theories predict that as voters are surrounded by those who are not like them, they should turn out less often. The empowerment theory attributes this to voters' internal, psychological factors such as feelings of discomfort or desire to avoid conflict with others who might hold dissimilar opinions. The mobilization theory contends, on the other hand, that the neighborhood effects are the results of mobilization strategies employed by campaigns. To maximize the efficiency of their mobilization, campaigns may target certain neighborhoods in which a large number of their potential supporters reside. Therefore, according to the mobilization theory, we would not expect neighborhood effects in districts with uncompetitive elections where candidates are unlikely to engage in mobilization campaigns (Schaffner, 2006; Jacobson and Kernell, 1983). In contrast, the empowerment theory predicts that the neighborhood effects persist even in uncompetitive elections because these effects are largely due to voters' psychological considerations.

**Average Effects.** Using both the cross-section and panel data, we examine whether the neighborhood effects persist even in uncompetitive districts. To do this, we estimate the neighborhood effects separately for competitive and uncompetitive districts. We define an uncompetitive district as one where the margin of victory is greater than or equal to 40 percentage points. That is, the

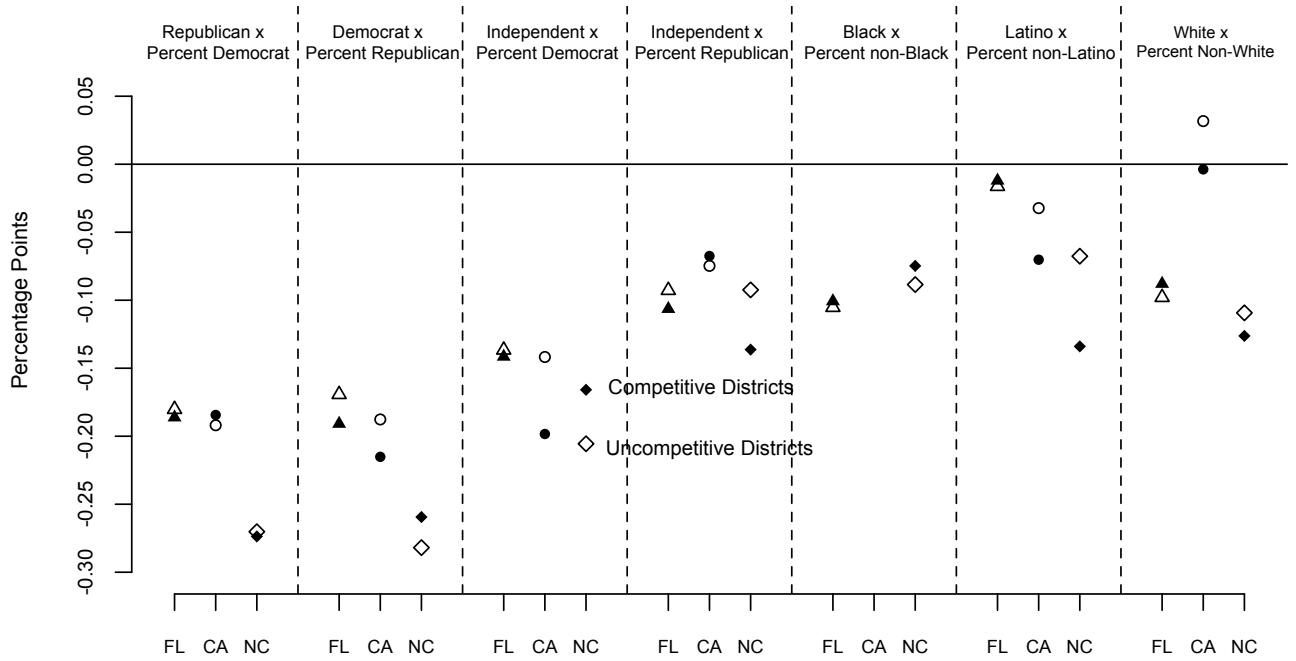


Figure 11: Estimated Average Neighborhood Effects in Uncompetitive Districts Using Cross Sectional Analysis. The figure plots the estimated average neighborhood effects (along with their 95% confidence intervals, which are too small to be seen) separately for uncompetitive districts (open symbols) and other districts (solid symbols) in Florida, California, and North Carolina. In many of the uncompetitive districts, a 10 percentage point increase in the size of the out group leads to an approximately 1 to 3 percentage point decrease in the probability of voting. Neighborhood effects persist even in uncompetitive districts.

winner obtained at least 70 percent of the two-party vote. Similarly, for the panel data, we define uncompetitive (competitive) districts as those districts whose margin of victory was greater (less) than 40 percentage points in both elections.

Figure 11 shows the results of the cross-sectional analysis and plot estimated neighborhood effects (along with their 95% confidence intervals, which are too small to be seen) separately for uncompetitive districts (open symbols) and other districts (solid symbols) in Florida, California, and North Carolina. The estimated effects are similar between uncompetitive and other districts, suggesting that neighborhood effects persist even in uncompetitive districts. Figure 17 of Appendix presents the results of our panel data analysis. Again, we find no systematic evidence that the neighborhood effects are smaller in uncompetitive districts. This result contradicts the prediction of the mobilization theory while remaining consistent with the empowerment theory.

**Effect Heterogeneity.** Finally, we conduct another analysis where we consider the possibility of nonlinearities in the effects of neighborhoods. Although our findings so far are inconsistent with the prediction of the threat theory, it is possible that the theory is supported in neighborhoods where we expect such threat mechanism to work. In particular, the logic of out-group threat may be more applicable to neighborhood majorities who feel threatened by the growing size of minority groups (e.g., Schelling, 1971; Card *et al.*, 2008). In contrast, the empowerment theory has been typically applied to neighborhood minorities (e.g., Mansbridge, 1983; Huckfeldt *et al.*, 2004)

To explore this link, we estimate the neighborhood effects separately based on baseline percentage of out-groups. We conduct the same analysis as presented earlier, but only for neighborhoods that contain between 0 to 34, 34 to 67, and 67 to 100 percent of the out-group. This allows us to identify if the effect is stronger or weaker in different types of neighborhoods. The results of the cross sectional and panel analysis are presented in Figures 18 and 19 of the Appendix. They show that the effects remain negative and do not vary in a systematic way. This again supports the empowerment theory and is inconsistent with the threat theory.

## 6 Concluding Remarks

In this paper, we have shown that voters' decision to turn out is conditional not only on their own demographics, but also the interaction between their demographics and those of their neighbors. Our findings suggest that these neighborhood effects persist in a variety of administrative, electoral, and geographical environments. We show that the robust neighborhood effects we identified are consistent with the predictions of the psychological theory of voter empowerment and cannot be explained by the mobilization theory alone. Of course, the analysis of voter registration data alone cannot directly test the causal mechanisms of these theories. We leave for future research this challenging task of investigating how exactly these neighborhood effects arise. Successfully answering this question is likely to require survey and experimental methodologies that directly measure and manipulate social and psychological factors related to neighborhood effects (Imai

*et al.*, 2011).

The existence of persistent neighborhood effects may mean that we need to call into question some of our most basic assumptions about which voters turn out more often. A vast majority of research on political participation is based on surveys of a relatively small number of respondents who live in a variety of neighborhoods and yet do not take into account neighborhood effects. Indeed, recent work using millions of registration records, Ansolabehere and Hersh (2011) have found, among other things, that many Black voters turn out at higher rates than Whites, and the varying correlation between income and vote choice may be better explained by race (Hersh and Nall, 2013) than aggregate income (Gelman, 2008). The presence of neighborhood effects may explain such discrepancies. As these examples illustrate, a better understanding of neighborhood effects may alter our basic understanding of political participation and other phenomena.

## References

- Aldrich, J. H. (1993). Rational choice and turnout. *American Journal of Political Science* 246–278.
- Alvarez, R. M., Hopkins, A., and Sinclair, B. (2010). Mobilizing pasadena democrats: measuring the effects of partisan campaign contacts. *The Journal of Politics* **72**, 01, 31–44.
- Anscombe, S. and Hersh, E. (2011). Gender, race, age, and voting: A research note. In *APSA 2011 Annual Meeting Paper*.
- Anscombe, S. and Rodden, J. (2013). 2008 Precinct-Level Election Returns. <http://projects.iq.harvard.edu/eda/data>.
- Arceneaux, K. and Nickerson, D. W. (2009). Who is mobilized to vote? a re-analysis of 11 field experiments. *American Journal of Political Science* **53**, 1, 1–16.
- Barreto, M. A. (2007). Si Se Puede! Latino candidates and the mobilization of Latino voters. *American Political Science Review* **101**, 3, 425–441.
- Barreto, M. A., Segura, G. M., and Woods, N. D. (2004). Mobilizing effect of majority-minority districts. *American Political Science Review* **98**, 1, 65–75.
- Bobo, L. (1983). Whites' opposition to busing: Symbolic racism or realistic group conflict. *Journal of Personality and Social Psychology* **45**, 6, 1196–1210.
- Bobo, L. and Gilliam Jr, F. D. (1990). Race, sociopolitical participation, and black empowerment. *The American Political Science Review* 377–393.
- California Secretary of State (2013). Inactive Voter Policy. <http://www.sos.ca.gov/elections/nvra/voter-reg-agencies/list-maintenance-requirements.htm>. [Online; accessed 06-April-2013].
- Campbell, A., Converse, P., Miller, W., and Stokes, D. (1960). *The American Voter*. University of Chicago Press.
- Campbell, D. E. (2006). *Why We Vote: How Schools & Communities Shape Our Civic Life*. Princeton University Press.

- Card, D., Mas, A., and Rothstein, J. (2008). Tipping and the dynamics of segregation. *Quarterly Journal of Economics* **123**, 1, 177–218.
- Carmines, E. G. and Stimson, J. A. (1989). *Issue evolution: Race and the transformation of American politics*. Princeton University Press.
- Carsey, T. M. (1995). The contextual effects of race on white voter behavior: The 1989 new york city mayoral election. *Journal of Politics* **57**, 1, 221–228.
- Cho, W. K. T., Gimpel, J. G., and Dyck, J. J. (2006). Residential concentration, political socialization, and voter turnout. *Journal of Politics* **68**, 1, 156–167.
- Eliasoph, N. (1998). *Avoiding politics: How Americans produce apathy in everyday life*. Cambridge University Press.
- Enos, R. D. (2010). What tearing down public housing projects teaches us about the effect of racial threat on political participation. *Unpublished manuscript* Retrieved from [http://ryandenos.com/papers/chicago\\_threat.pdf](http://ryandenos.com/papers/chicago_threat.pdf).
- Enos, R. D. (2011). Testing the elusive: a field experiment on racial threat. *Unpublished manuscript (Harvard University)* .
- Feldman, S. and Zaller, J. (1992). The political culture of ambivalence: Ideological responses to the welfare state. *American Journal of Political Science* 268–307.
- Florida Secretary of State (2013). Inactive Voter Policy. [http://election.dos.state.fl.us/publications/pdf/2011/2011-2012\\_VoterRegGuide.pdf](http://election.dos.state.fl.us/publications/pdf/2011/2011-2012_VoterRegGuide.pdf).
- Fraga, B. L. (2012). Candidates or influence? reevaluating the role of race in voter turnout. *Working Paper* .
- Franz, M. M. (2013). Targeting campaign messages. *New Directions in Media and Politics* 113.
- Gay, C. (2001). The effect of black congressional representation on political participation. *American Political Science Review* **95**, 3, 589–602.

- Gelman, A. (2008). *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do*. Princeton University Press.
- Gerber, A. S. and Green, D. P. (2000). The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment. *American Political Science Review* 653–663.
- Giles, M. W. and Buckner, M. A. (1993). David duke and black threat: An old hypothesis revisited. *Journal of Politics* 55, 3, 702–13.
- Giles, M. W. and Evans, A. S. (1985). External threat, perceived threat, and group identity. *Social Science Quarterly* 66, 50–66.
- Giles, M. W. and Hertz, K. (1994). Racial threat and partisan identification. *American Political Science Review* 317–326.
- Gilliam, F. D. J. (1996). Exploring minority empowerment: Symbolic politics, governing coalitions, and traces of political style in Los Angeles. *American Journal of Political Science* 40, 1, 56–81.
- Gimpel, J. G., Dyck, J. J., and Shaw, D. R. (2004). Registrants, voters, and turnout variability across neighborhoods. *Political Behavior* 26, 4, 343–375.
- Gosnell, H. F. (1977). *Getting out the vote: An experiment in the stimulation of voting*. Greenwood Press.
- Green, M. C., Visser, P. S., and Tetlock, P. E. (2000). Coping with accountability cross-pressures: Low-effort evasive tactics and high-effort quests for complex compromises. *Personality and Social Psychology Bulletin* 26, 11, 1380–1391.
- Henderson, J. A., Sekhon, J. S., and Titiunik, R. (2012). Cause or effect? turnout in Hispanic majority-minority districts. Department of Political Science, University of California, Berkeley.
- Hersh, E. (2011). The public code of racialized electioneering. In *APSA 2011 Annual Meeting Paper*.
- Hersh, E. and Nall, C. (2013). The primacy of race in the geography of income-based voting: New evidence from public voting records. Working Paper.

- Hillygus, D. S. (2005). Campaign effects and the dynamics of turnout intention in election 2000. *Journal of Politics* **67**, 1, 50–68.
- Hochschild, J. L. (1993). Disjunction and ambivalence in citizens' political outlooks. *Reconsidering the democratic public* 187–210.
- Huckfeldt, R., Johnson, P. E., and Sprague, J. (2004). *Political disagreement: The survival of diverse opinions within communication networks*. Cambridge University Press.
- Huckfeldt, R. R. (1979). Political participation and the neighborhood social context. *American Journal of Political Science* 579–592.
- Imai, K., Keele, L., Tingley, D., and Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review* **105**, 4, 765–789.
- Imai, K. and Kim, I. S. (2012). On the use of linear fixed effects regression estimators for causal inference. Working paper available at <http://imai.princeton.edu/research/FEmatch.html>.
- Jacobson, G. C. and Kernell, S. (1983). *Strategy and choice in congressional elections*. Yale University Press New Haven.
- Karpowitz, C. F., Monson, J. Q., Nielson, L., Patterson, K. D., and Snell, S. A. (2013). The people in your neighborhood: How political minority status affects political participation. *Annual Meeting of the Southern Political Science Association* .
- Key, V. (1949). *Southern Politics in State and Nation*. A.A. Knopf.
- Kim, I. S. and Imai, K. (2012). wfe: Weighted linear fixed effects regression models for causal inference. available at the Comprehensive R Archive Network (CRAN). <http://CRAN.R-project.org/package=wfe>.
- Malchow, H. (2008). *Political targeting*. Predicted Lists, LLC.
- Mansbridge, J. J. (1983). *Beyond adversary democracy*. University of Chicago Press.

- Matthews, D. R. and Prothro, J. W. (1963). Social and economic factors and negro voter registration in the south. *The American Political Science Review* **57**, 1, 24–44.
- McConaughy, C. M., White, I. K., Leal, D. L., and Casellas, J. P. (2010). A Latino on the ballot: Explaining coethnic voting among Latinos and the response of White Americans. *Journal of Politics* **72**, 4, 1199–1211.
- McNamara, M. (2008). *The Political Campaign Desk Reference: A Guide for Campaign Managers and Candidates Running for Elected Office*. Outskirts Press.
- Mutz, D. C. (2006). *Hearing the other side: Deliberative versus participatory democracy*. Cambridge University Press.
- New York Times (2012). New York Times Exit Poll. <http://elections.nytimes.com/2012/results/president/exit-polls>.
- Nickerson, D. W. (2008). Is voting contagious? evidence from two field experiments. *American Political Science Review* **102**, 1, 49.
- Pantoja, A. D. and Segura, G. M. (2003). Does ethnicity matter? descriptive representation in legislatures and political alienation among latinos\*. *Social Science Quarterly* **84**, 2, 441–460.
- Rosenberg, M. (1954). Some determinants of political apathy. *Public Opinion Quarterly* **18**, 4, 349–366.
- Schaffner, B. F. (2006). The political geography of campaign advertising in us house elections. *Political Geography* **25**, 7, 775–788.
- Schelling, T. (1971). Dynamics models of segregation. *Journal of Mathematical Sociology* **1**, 143–186.
- Spence, L. K. and McClerking, H. (2010). Context, black empowerment, and african american political participation. *American Politics Research* **38**, 5, 909–930.
- U.S. Census Bureau (2010). Census Block Definition. <http://www.census.gov/prod/cen2010/doc/sf1.pdf#page=474>.

- Verba, S. and Nie, N. H. (1972). *Participation in america: Social equality and political democracy*. *New York: Harper& Row* .
- Verba, S., Schlozman, K. L., and Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Harvard University Press.
- Voss, D. S. (1996). Beyond racial threat: Failure of an old hypothesis in the new south. *Journal of Politics* **58**, 1156–1170.
- Voss, D. S. and Lublin, D. (2001). Black incumbents, white districts an appraisal of the 1996 congressional elections. *American Politics Research* **29**, 2, 141–182.
- Voter Registration Statistics (2013). North Carolina Secretary of State. [http://www.app.sboe.state.nc.us/webapps/voter\\_stats/results.aspx?date=11-02-2010](http://www.app.sboe.state.nc.us/webapps/voter_stats/results.aspx?date=11-02-2010).
- Washington, E. (2006). How Black candidates affect voter turnout. *Quarterly Journal of Economics* **121**, 3, 973–998.
- Wolfinger, R. E. and Rosenstone, S. J. (1980). *Who votes?*, vol. 22. Yale University Press.
- Wong, C., Bowers, J., Williams, T., and Simmons, K. D. (2012). Bringing the person back in: Boundaries, perceptions, and the measurement of racial context. *The Journal of Politics* Forthcoming.
- Zaller, J. (1992). *The nature and origins of mass opinion*. Cambridge university press.

## Appendix: Additional Empirical Results

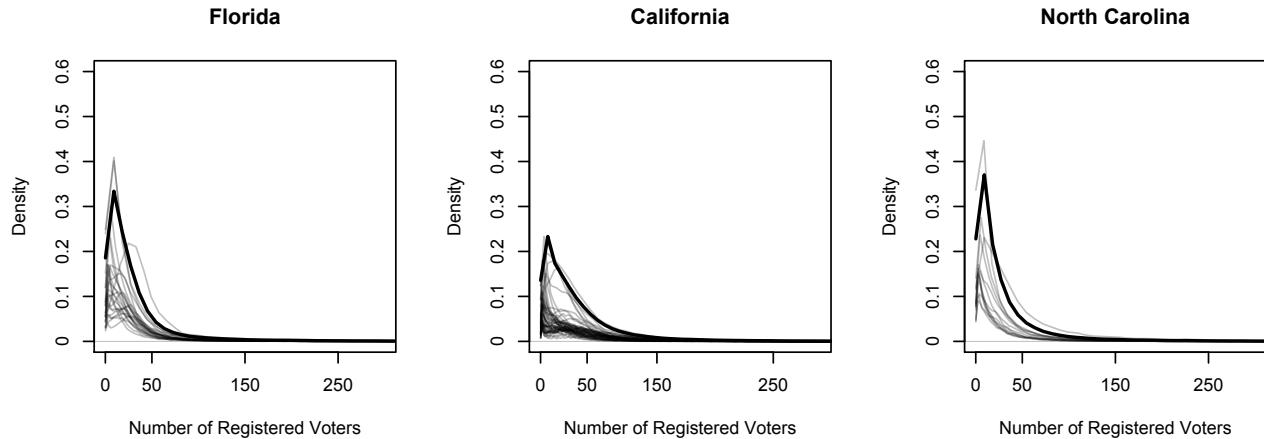


Figure 12: Number of Registered Voters in Census Blocks by Congressional Districts. For each congressional district in Florida, California, and North Carolina, we plot the distribution of the number of registered voters contained in a census block (thin lines). The plots show that census blocks are very small geographic units, usually containing fewer than 100 registered voters. The thick line shows the distribution for the entire state.

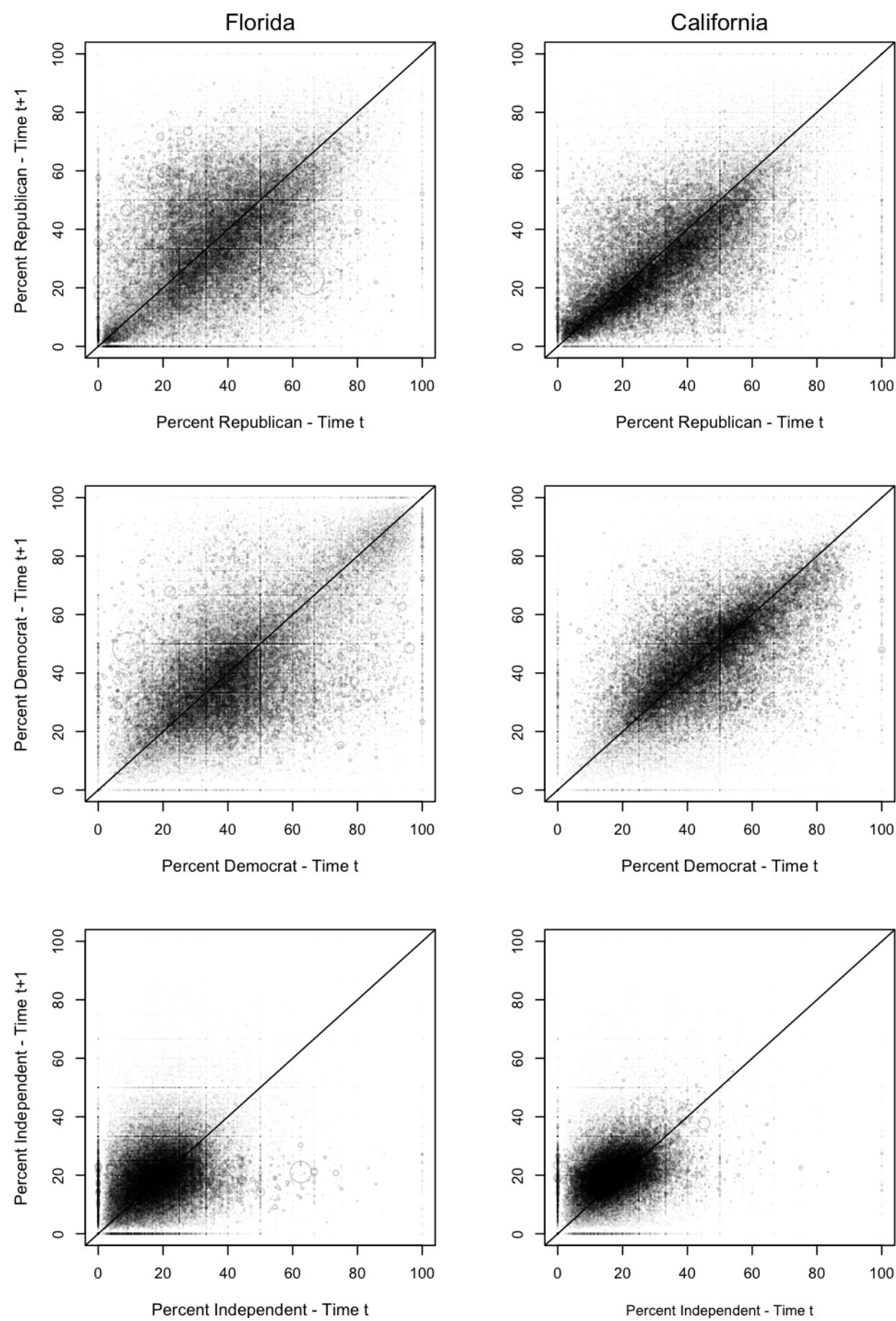


Figure 13: Change in Neighborhood Partisanship Composition over Time for Florida and California.

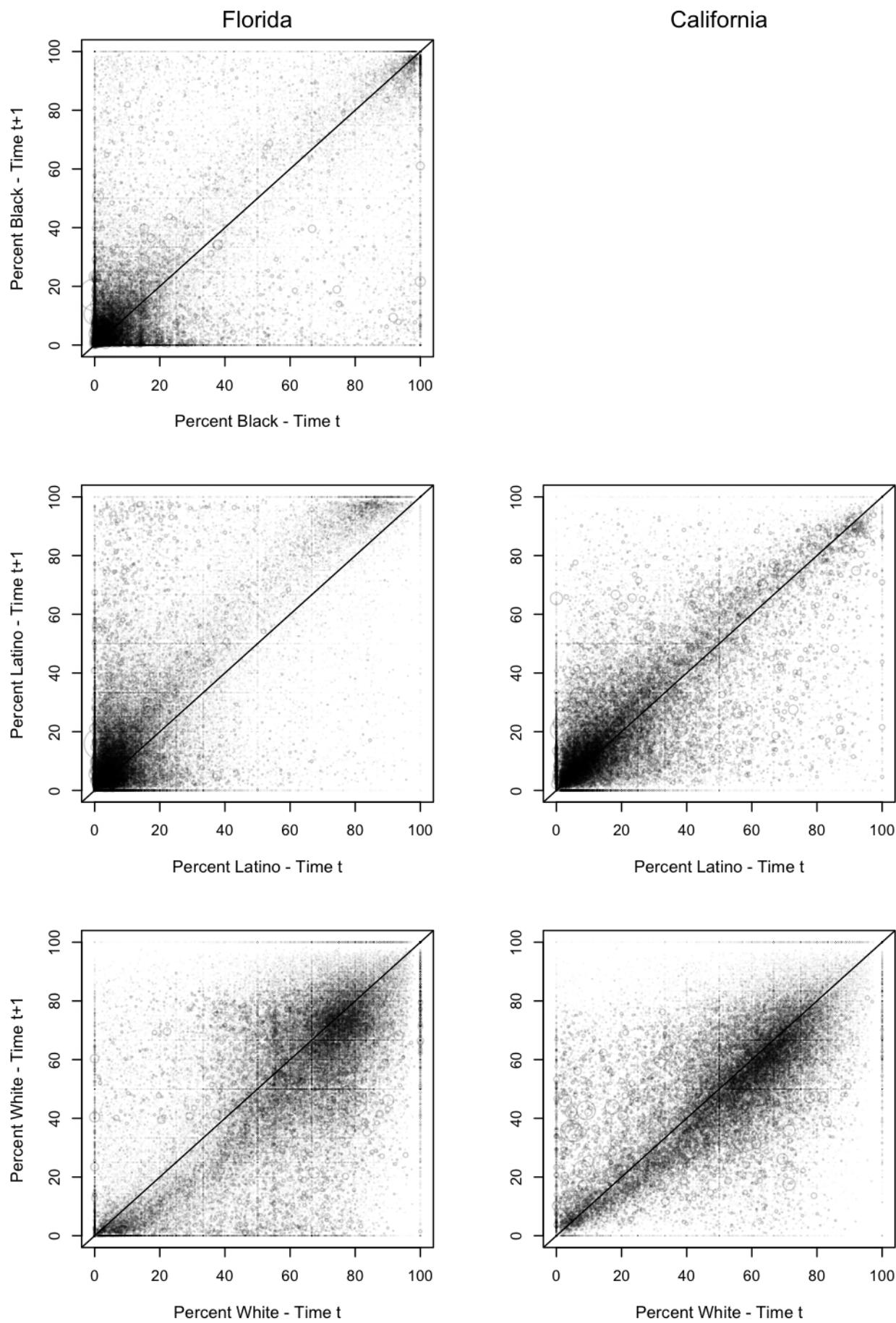


Figure 14: Change in Neighborhood Racial Composition over Time for Florida and California.

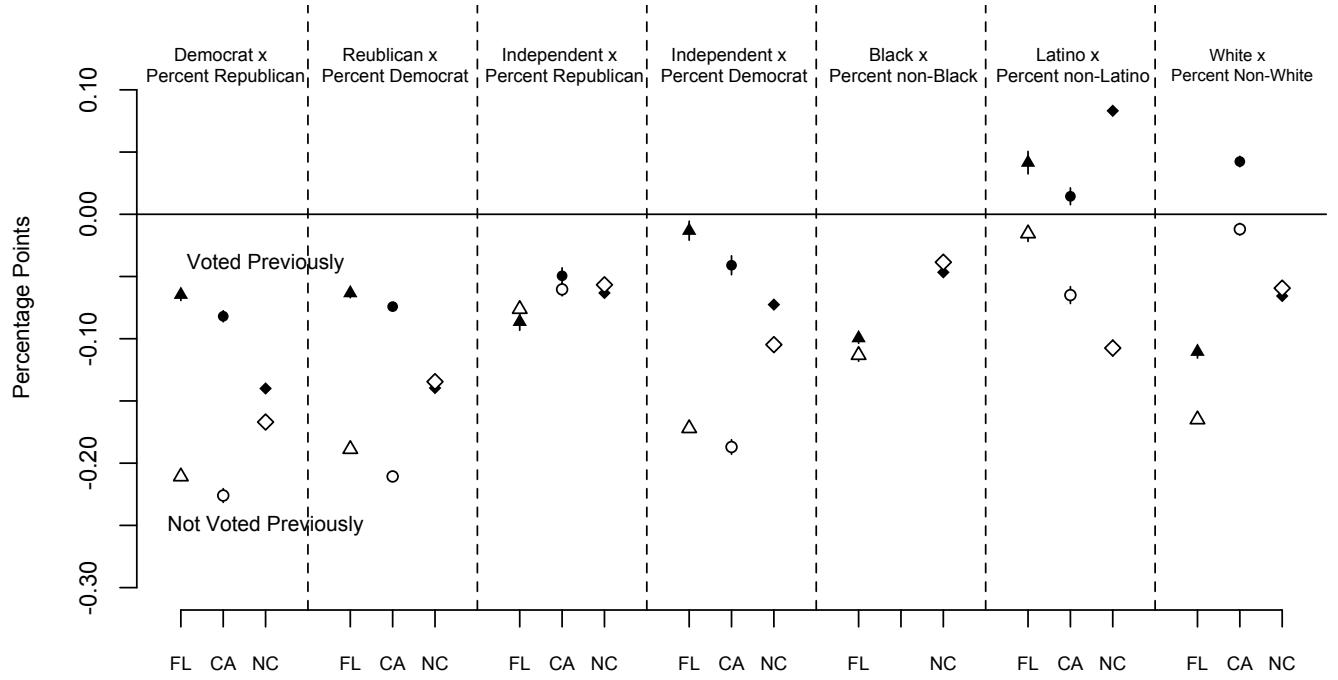


Figure 15: Cross Section Results when Adjusting for a Previous Vote. This figure shows the cross sectional model when considering only voters who voted in the previous election (solid symbols) and voters who did not vote (open symbols). Neighborhood effects are stronger among those that did not vote in the previous election.

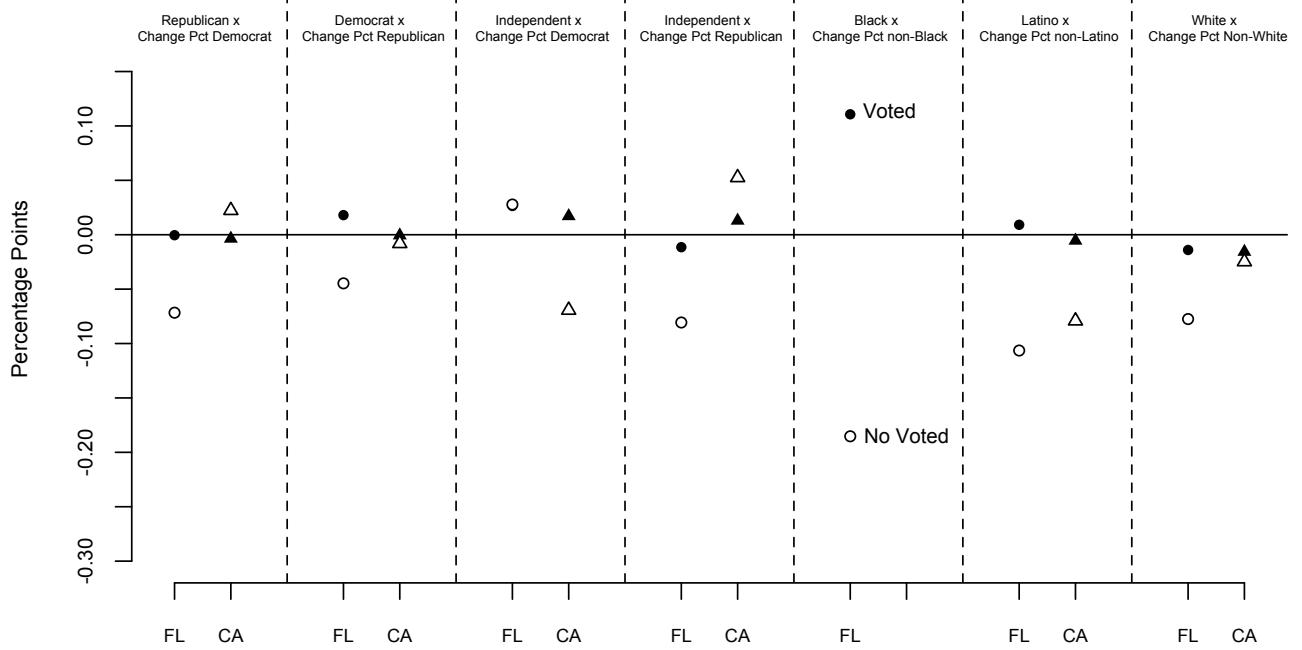


Figure 16: Panel Results when Adjusting for a Previous Vote. This figure shows the cross sectional model when considering only voters who voted in the previous election (solid symbols) and voters who did not vote (open symbols). In Florida, neighborhood effects are much stronger among those that did not vote in the previous election.

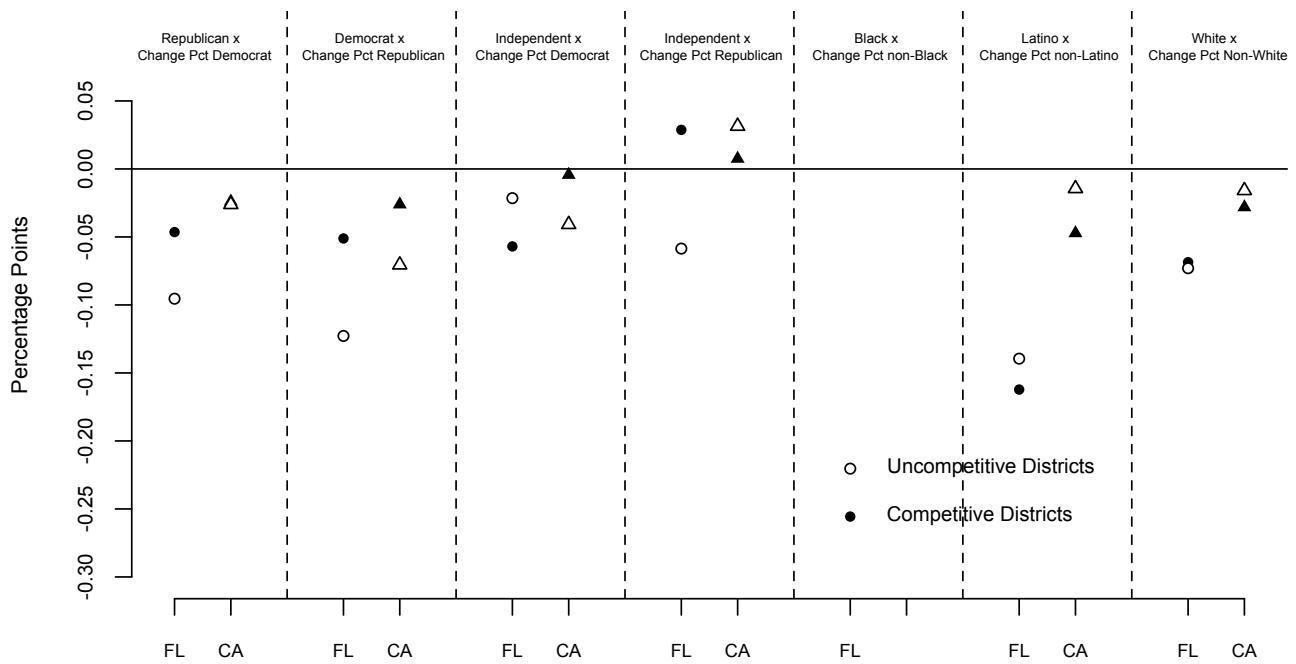


Figure 17: Estimated Average Neighborhood Effects in Uncompetitive Districts Using Panel Analysis. The figure plots the estimated average neighborhood effects (along with their 95% confidence intervals, which are too small to be seen) separately for uncompetitive districts (open symbols) and other districts (solid symbols) in Florida and California (see the caption of Figure 10 for details). In many of uncompetitive districts, a 10 percentage point increase in the size of the out group leads to an approximately 0.5 to 2.5 percentage point decrease in the probability of voting. Neighborhood effects persist even in uncompetitive districts.

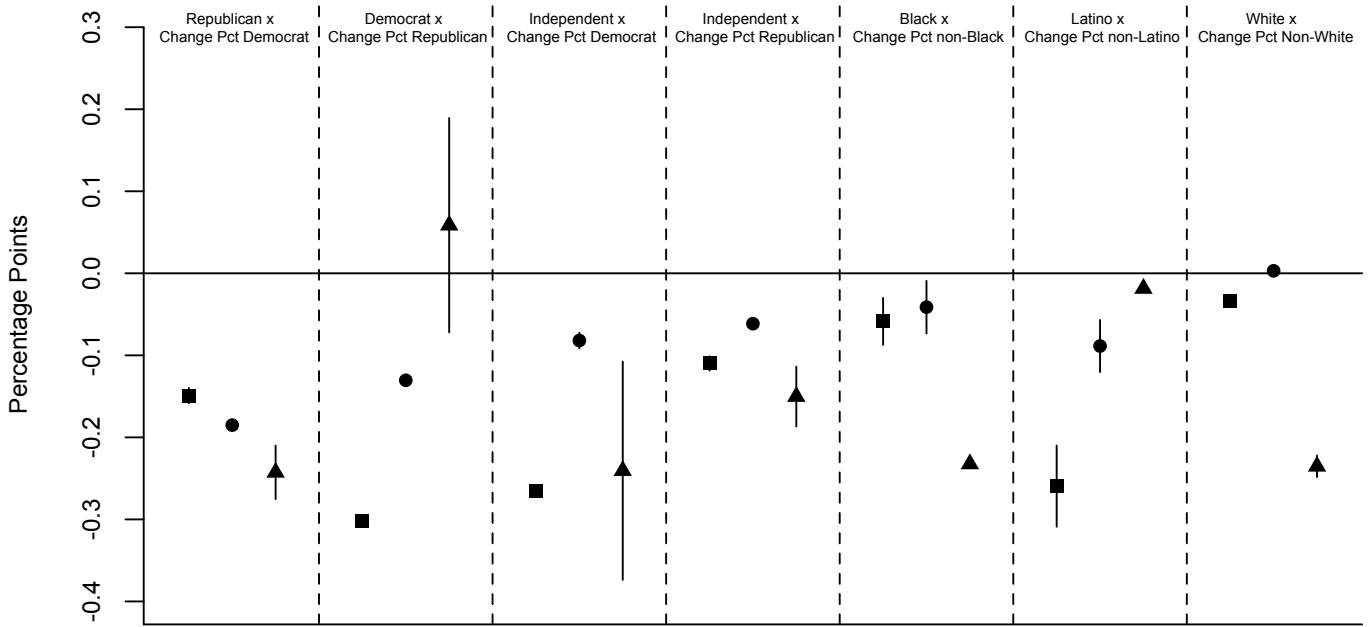


Figure 18: Cross Section Model Subset by Baseline Percentages of Outgroup. This figure shows the results of the cross sectional analysis divided into three separate categories based on the percentage of the out-group in the neighborhood. For example, the squares in each column refers to the estimates using only voters who live in neighborhoods with the out-group percentage ranging from 0 to 34 percent. Similarly, the circles in each column refers to the panel estimate based only on voters who live in neighborhoods with the out-group proportion ranging from 34 to 67 percent. Finally, the triangles refer to the estimates using only voters who live in neighborhoods with the out-group percentage ranging from 67 to 100 percent. The vertical lines represent 95% confidence intervals.

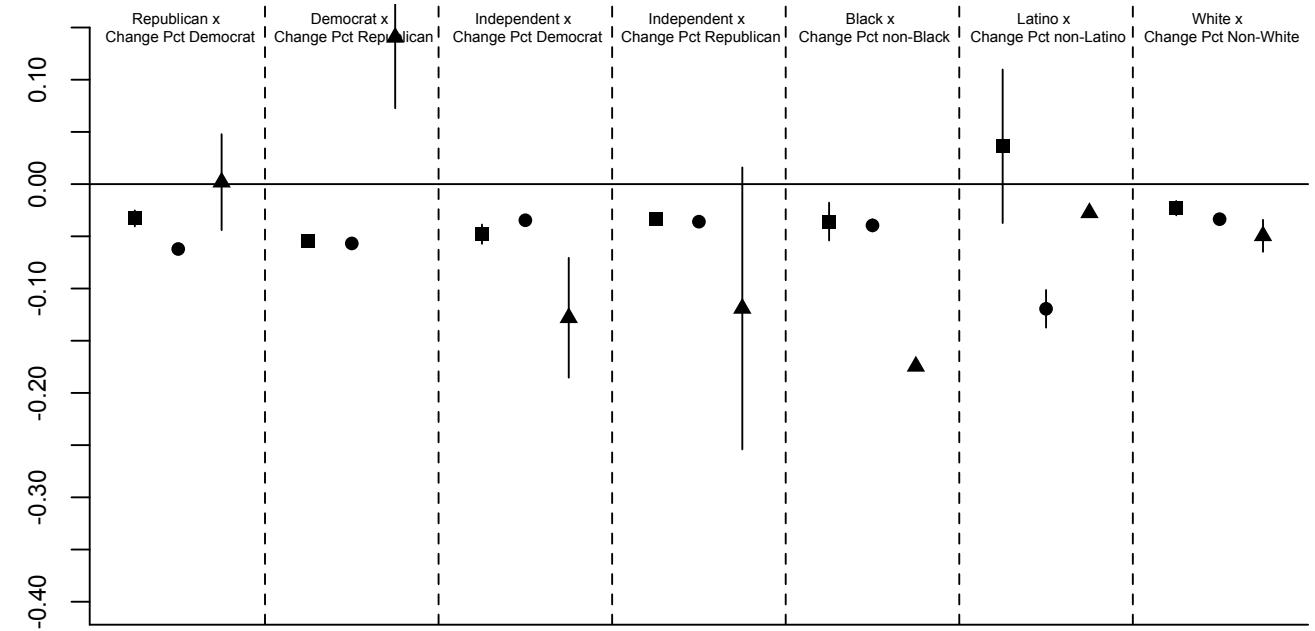


Figure 19: Panel Model Subset by Baseline Percentages of Outgroup. This figure shows the results of the panel analysis divided into three separate categories based on the percentage of the out-group in the neighborhood at the first period of the panel. For example, the squares in each column refers to the estimates using only voters who live in neighborhoods with the out-group percentage ranging from 0 to 34 percent. Similarly, the circles in each column refers to the panel estimate based only on voters who live in neighborhoods with the out-group proportion ranging from 34 to 67 percent. Finally, the triangles refer to the estimates using only voters who live in neighborhoods with the out-group percentage ranging from 67 to 100 percent. The vertical lines represent 95% confidence intervals.