Homophily Turnout

Network effects and voter turnout

Taylor J Weidman

December 7, 2023

Abstract. *At the heart of public choice lies the question of why voters turnout. This question has produced empirical and theoretical results suggesting a role for social norms and neighborhood spillovers. This paper uses a panel of voters and their nearest neighbors, finding that a voter treated with a new in-party neighbor increases their turnout rate by roughly 1/3% in Municipal elections, 1/4% in Primary elections, and 1/20% in Presidential elections when compared to a similar but untreated nearby voter. This "Homophily Turnout" occurs most strongly among a voter's closest neighbors, in low-turnout elections, is negative for out-party treatments, and is intensified with treatments of the same race. All but the Statewide panel survive a balance test, consistent with the parallel trends assumption, suggesting the estimates are not due to the to residential sorting or a registration effect. Together this suggests that homophily networks shape the flow of information and norms in the voter's turnout decision.*

I. Introduction

A striking feature of democratic political systems is the voluntary participation of voters who have little individual impact and significant participation costs. The US political landscape has residentially clustered to a level not seen in recent history raising questions about residential spillovers for the decision to turnout *[Kaplan et al. (2022), Weidman (2023)]*. If social norms indeed transmit most densely over homophily networks (*networks of similar people*), turnout norms may depend not just on the residential closeness of two voters, as in *Gerber et al. (2008)*, but also on their party membership. A line of turnout models support this idea, showing how peer networks could influence turnout through a strategically designed party turnout norm that is enforced through peer monitoring *[Downs (1957); Harsanyi (1980); Coate and Conlin (2004); Feddersen and Sandroni (2006); Favarelli and Walsh (2015); Levine and Mattozzi (2020)]*.

This paper develops a simple framework relating the shape of residential networks to turnout behavior, building on the observations that those who are geographically nearer and politically and demographically similar will be more likely to interact, and that one conforms to the beliefs, attitudes, and behaviors of those one is closest to. This *Homophily Turnout* framework (*developed next*) suggests that the shape of spillovers across political networks within the voter's neighborhood should increase turnout among voters treated with a new in-party neighbor, that the response should be greater for nearer treatments, in lower turnout elections, among racially similar voters, with a null or negative impact of an out-party treatment. This also suggests that the recent residential clustering of the electorate has the potential to impact aggregate turnout. While more work is required to estimate the counterfactual aggregate turnout, since one's neighbors are unlikely to be one's closest ties, the aggregate impact of homophily networks is likely much larger over the electorate's full network structure than just one's residential network.

These types of highly local residential networks are observable in a new panel of voter's and their point-neighborhoods developed in *Weidman (2023)*. However, measuring spillovers at the residential level has presented an empirical difficulty that neighborhoods are themselves selected, both on political preferences and voter participation. This 'reflection problem' has been a longstanding issue in the study of observational relationships between neighborhoods and behavior, like turnout. To avoid these issues, this paper makes use of the observation that a new neighbor may choose their neighborhood based on political characteristics at any point in time, but they will have difficulty anticipating how their potential neighbors' political engagement *will change* between elections relative to nearby observationally similar voters. This allows a test of whether homophily networks in residential neighborhoods transmit information, norms, and beliefs. This approach finds that a voter treated with a new in-party neighbor increases their turnout rate by roughly 1/3% in Municipal elections, 1/4% in Primary elections, and 1/20% in Presidential elections. This *Homophily Turnout* occurs most strongly among a voter's closest neighbors, in low-turnout elections, is negative for out-party treatments, and is intensified with treatments of the same race. The timing of treatment makes it possible to show that in all panels except Statewide elections the direction of causality goes from treatment to turnout, satisfying the parallel trends assumption, and is not due to residential sorting on turnout trends or a registration effect. This paper contributes to our understanding of information spillovers, voter behavior, homophily networks, and deploys a flexible set of tools to a new dataset.

II. Conceptual Framework

Three empirical regularities motivate the paper's empirical model: 1) turnout *rates* are positively related to in-group neighborhood composition, a relationship that is 2) hyperlocal, attenuating with geographic distance, and 3) appears in nearly every available election *[Weidman (2023)]*. These empirical patterns result from many intersecting forces. One type of mechanism responsible for these relationships is the residential sorting of voters into neighborhoods with like-minded neighbors alongside the switching of parties toward the majority affiliation of their neighbors, both for reasons directly and indirectly related to politics.

But as *Levine and Mattozzi (2020)* suggest, another type of mechanism may be at play: network spillovers. The types of residential spillovers shown by *Gerber et al. (2008)* have long been studied as important for the emergence of equilibrium behavior, with much direct evidence pointing to a complicated behavioral role for norms in one's social context. For example, *Carrell et al. (2011)* used a clever natural experiment to show how information and beliefs are contagious over networks. *Galesic et al. (2018)* show that the intentions to vote among one's social network is a better predictor of turnout one's own intention, suggesting a potential role for social influence. Together, beliefs, attitudes, and behaviors are broadly influenced by those in one's immediate social environment. In the residential context, spillovers in turnout are likely due both directly to neighborhood social networks and indirectly to information channels like yard signs.

Further, the role of residential neighborhood spillovers interact with the tendency for individuals to form networks based on similarity and identities *[Currarini et al. (2009)]*. Members of homophily networks, networks of similar people, tend to spend more time together and exchange more information. Notably, *Bertrand et al. (2000)* use this type of interaction between homophily networks and residential proximity to demonstrate that geographic ethnic networks transmit information about public services. Homophily networks even form on political views within one's family network: *Chen and Rohla (2018)* evocatively show that voters avoid psychologically costly political disagreements by spending less time on Thanksgiving Day with family members who belong to an opposing political party. This type of cost avoidance leads to denser network connections among like-minded voters, a pattern that may also appear among one's neighbors. Homophily networks may also lead to the reinforcement of one's beliefs. *Sunstein (1999)* shows that members of homogeneous groups tend to move toward more extreme beliefs. Just like one's choice of information sources, homophily networks may shape political attitudes in ways that create a landscape of polarization *[DellaVigna and Kaplan (2007)]*.

The composite result from information spillovers across political homophily networks concentrated within residential networks describes *Homophily Turnout* (*HT*), a positive relationship between the density of a voter's residential homophily network and their turnout decision. Similar to the framework developed in *Bertrand et al. (2000)*, *HT* suggests that both residential proximity and homophily shape the flow of information and norm setting.

Networks are more dense both among those who live nearby and among those with similar political affiliations, suggesting a particular shape for the flow of information, beliefs, and norms. For the reasons mentioned above, one becomes similar to, behaves similarly to, and can enforce norms more effectively among those one interacts with most. A nearby neighbor likely has a larger impact on one's turnout decision than a voter in a neighboring county. A nearby neighbor one knows will have a larger impact than a neighbor one does not. Peer monitoring of social norms as suggested by *Levine and Mattozzi (2020)* make it easier to enforce the party-designed turnout norm, leading to higher turnout when in-party networks are most dense. It may also be more difficult to vote against the interest of those who are closer on the network, suggesting a lower cost to voting when the neighborhood contains more likeminded voters.

This network shape suggests turnout can vary with a spatial reorganization of the electorate, offering five specific predictions orthogonal to the core voter turnout models. First, *HT* suggests that a voter's turnout should increase with the number of politically aligned neighbors. A voter should increase their turnout after a new in-group neighbor moves in next door. Second, *HT* suggests that the relationship between turnout and in-party membership should weaken as the residential distance to a new in-party neighbor grows. The relationship between in-party neighborhood membership and treatment should appear most strongly between voters who live next door, since voters are much more exposed to those in their neighborhood than to those in a neighboring county. As the distance between two voters increases, their influence on each other should weaken.

Third, the opposite side of this framework is that a voter's turnout should not increase with the number of politically misaligned neighbors. While having more misaligned neighbors may motivate a voter to turnout more, after controlling for instrumental incentives it should be harder to vote against the interest of an opposing voter that is nearer than one who is further away. If peer monitoring is the mechanism for *HT*, the voter's turnout should be independent of the number of one's out-party neighbors. If this were the mechanism, increasing the number of out-party neighbors should do nothing to the cost of monitoring. If the costliness of voting against or benefit of voting for the interests of one's network are responsible for *HT*, the voter should turnout at a lower rate after an out-party neighbor moves into the neighborhood. Assuming one cares about the interests of one's neighbors, it should be more difficult to vote for a ballot when the new neighbor wishes it would not pass.

Fourth, the *HT* relationship should be strongest in low-turnout elections. Large elections tend to saturate the information landscape with advertising, media, and direct targeting in very different ways than smaller elections. In low-information, low-turnout elections, there should be a relatively larger role for information spillovers that are less obscured by channels outside the neighborhood environment.

Finally, if this pattern is due to neighborhood homophily networks, *HT* should appear on dimensions of homophily other than political affiliation. It has long been known that homophily networks form around racial identities. If networks are truly the mechanism responsible for the positive relationship between turnout and in-party neighborhood share, the *HT* relationship should be strongest when one's in-party neighbors are of one's own race. While this heterogeneity of *HT* on race cannot rule out all potential explanations, it would strengthen the idea that the effect is indeed about networks.

These patterns play out on a changing political geography. The American electorate has undergone a substantial residential clustering of voters into neighborhoods with a disproportionate share of copartisans over the course of the last 50 years, reaching a level not seen in recent history *[Sussel (2013); Kaplan et al. (2022)]*. Clustering not only has consequences for drawing district lines, but may also have spillover effects on political behavior itself *[Wilkinson (2019); French (2020)]*. We are only beginning to understand how these recent geographic changes affect voter behavior.

III. Methods

The *Homophily Turnout* framework is taken to a linked geocoded administrative panel constructed from a sequence of snapshots of North Carolina voters between 2010 and 2020 *[Weidman (2023)]*. Voter files are maintained by each states' secretary of state in response to the US Congress adopting the *Help America Vote Act* of 2002 over concerns about consistency and accuracy of voter records raised by the 2000 Gore-Bush election *[Igielnik et al. (2018)]*. North Carolina maintains administrative snapshots of all registered voters since 2005 with individual turnout histories for voters in more recent years. The panel contains a sequence of each voter's individual turnout, party affiliation, demographics, and a vector of their geographically closest neighbors. Each voter's sequence is split when they move or reaffiliate. At the expense of efficiency, this offers many statistical benefits, such as avoiding the need to model voters' residential and party choices.

Figure M1 shows the demographic composition of the panel in 2010 and 2020 along with the number of voters by party at each snapshot date. The Republican party is composed of nearly all White voters, an equal share of male and female voters, and a nearly constant total membership over the decade. The Democratic party is composed of nearly equal shares of White and Black voters, a slight lean toward female, and a decline in total membership over the decade. This decline is not due to a decline in statewide registration, but comes at the same time as a dramatic rise in voters affiliating with non-major parties [Weidman (2023)].

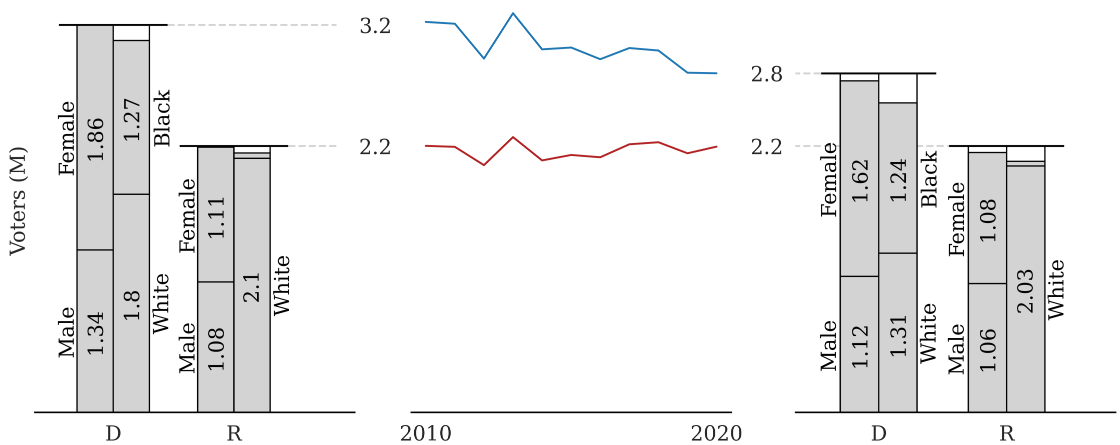

Figure M1. *Demographics of data snapshots*

Each voter's geocoded residence locates them at a point on the map. Their point-neighborhood is defined using a radius around their residential address such that the resulting circle contains the voter's k nearest neighbors in the first year the voter appears in the data. This time-consistent *k-nearest-neighbors* approach provides better statistical resolution than fixed geographic approaches that tessellate the plane. This makes it possible to observe the changes in the voter's nearest neighbors through time [Weidman (2023)]. A simulated point-neighborhood at a point in time is shown on the left side of Figure M2, with the innermost ring containing the voter's 5 nearest neighbors and the next largest ring containing the voter's 10 nearest neighbors. The dataset contains these point-neighborhoods of the voter's k nearest neighbors for each $k \in \{5, 10, 25, 50, 100, 200, 400, 800\}$.

The radial structure of point-neighborhoods make it possible to construct an annulus around the point-neighborhood as a statistical control. To avoid potential interactions between those in the annulus and the point-neighborhood, the annulus begins at an inner radius containing

800 voters, chosen to be both small and unrelated to the turnout decision. To visualize this setup, Figure M3 shows an exaggeratedly small buffer of voters in light gray. An inner ring at 800 and an outer ring at $800+k$ for each k defines the corresponding annulus. For example, the $k = 5$ point-neighborhood has a corresponding annulus of roughly equal population defined by the 800 and 805 rings. Holding equal the population of the annulus and the point-neighborhood in this way avoids the statistical issues related to population variability *[Cui et al. (2015)]*. The empirical distribution of point-neighborhood radii is shown on the right side of Figure M2.

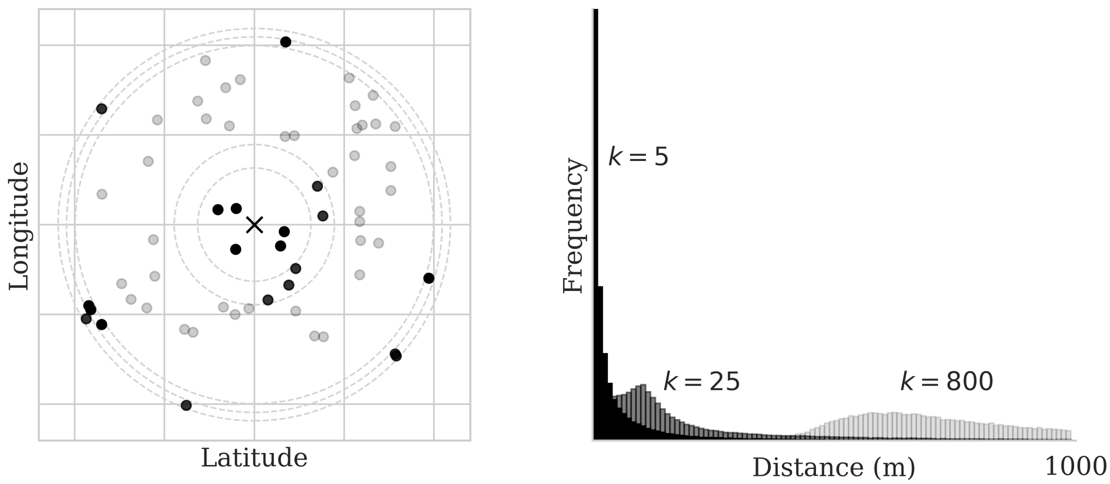

Figure M2. *Neighborhoods and their sizes*

Elections are grouped into three disjoint panels of similar types of elections. Figure M3 shows the distribution of turnout rates for each panel of elections. The Presidential panel is run only with presidential elections. The Primary panel includes all primary elections. The Municipal panel includes all municipal elections. The number of individual observations is the sum of the individual voters included in each election, with most voters appearing in multiple elections. For example, an electorate with 100 voters who appear in all three of the same presidential elections would produce 300 observations.

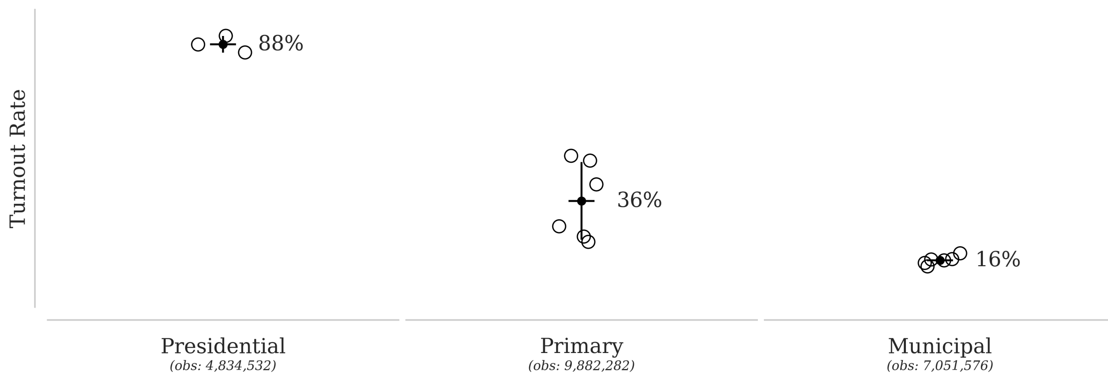

Figure M3. *Turnout rates of disjoint panels*

The 'reflection problem' makes it impossible to interpret the observational relationships between turnout and neighborhood introduced in *Weidman (2023)*. Does a voter's behavior depend on the behavior and characteristics of their neighbors, or does a voter behave similarly to their neighbors because they are similar and subject to the same shocks? As Bertrand et al. (2000) point out, this problem can be viewed as the result of three related omitted variable biases. First, unobserved voter characteristics may be correlated with turnout shocks. For example, active voters appear to choose residential neighborhoods with disproportionately more members of their own party, leading to higher treatment rates in neighborhoods with many likeminded neighbors. Second, unobserved neighborhood characteristics may be correlated with turnout. For example, if political campaigns target precincts with more receptive voters, neighborhoods with many likeminded voters may see an increase in turnout rates. Third, treatment rates may be correlated with regional realignments. For example, voters may move into neighborhoods in which the neighbors have become more politically aligned, neighborhoods with turnout rates that are increasing for reasons unrelated to the neighbors. These biases are all likely positive, meaning a positive relationship between turnout and in-party neighbors cannot be interpreted as evidence for *Homophily Turnout*.

III.a Empirical Approach

To avoid these omitted variable biases, the empirical approach uses *changes* in the voter's turnout decision and their point-neighborhood to compare a treated voter to a nearby voter who is observably similar but untreated. Then this estimate further compares a treatment in their point-neighborhood to a treatment in the annulus. Equation (1) formalizes this approach. Voter i is a member of cohort c and makes a turnout choice in election t, y_{itc} (*observed*), the cumulative treatments in their k point-neighborhood, N^k_{it} (*observed*)[1], and

[1] To further minimize the possibility that treatment is related to *changes* in the voter's political engagement, N_{it} isolates the changes in the neighborhood due only to new in-party neighbors moving into the

the more distant neighbors in the residential annulus Z_{it}^k (*observed*). Time-invariant characteristics, like race, gender, and family history impact the voter's turnout decision and neighborhood composition through α_i (*unobserved*). Time-varying characteristics like age, income, and unobserved election-specific factors like party targeting and media environment impact the voter's turnout decision and neighborhood composition through α_{tc} (*unobserved*). Finally, to compare elections with different turnout rates, the score adopted here divides the voter's turnout decision, y_{itc}, by the baseline turnout rate of all other voters $j \neq i$, R_{it}: $w_{itc} = y_{itc}/R_{it}$. This captures the idea that the role of treatment should be proportional to baseline turnout, scoring turnout more highly in lower turnout elections.[2] Therefore, Equation (1) estimates β as an elasticity, which should be interpreted as the change in the election's turnout *rate* if every voter were to receive one in-party neighbor.

$$w_{itc} = \alpha_i + \alpha_{tc} + \beta^k N_{it}^k + \gamma Z_{it}^k + \epsilon_{itc} \tag{1}$$

This approach controls for many common omitted variable biases. First, the voter fixed effects absorb each voter i's unchanging characteristics like gender, race, and family background, strong predictors of turnout and residential choice. This leaves only the variation *around* the voter's mean observables. Second, cohort-election fixed effects absorb election-specific shocks common among nearby and observationally identical voters who have participated in all the same elections. A cohort c is defined as the set of voters in the same census block, affiliated to the same party, with a flexible set of demographics, who appear in all the same elections. This makes cohort-election fixed effects election-specific, absorbing heterogeneous shocks common among observationally idential voters within the block in any election t. For example, cohort-election fixed effects compare treated and untreated Black male Democrats who have lived in the same Census block between 2016 and 2020. This absorbs the possibility that a block becomes less residentially attractive to a demographic group within a party while also becoming more politically engaged through campaign targeting. Cohort-election fixed effects leave only the *unshared* changes. Third, the model uses the idea that comparing a treatment in the point-neighborhood to a treatment in the annulus will further minimize the possibility of unobserved heterogeneity within the cohort.

neighborhood, ΔN_{it}: $N_{it} = \sum_{T=0}^{t} \Delta N_{it}$. The estimates produced using all neighborhood changes are similar but are roughly 33% larger than those produced using only new neighbors. There is still the possibility that a voter is shocked and leads a new neighbor to update their registration from a past address to their current address. This may complicate the interpretation, but it does not undermine the idea that this relationship is about residential spillovers. Future work could use the intensity of the new voters' political activity to separate these channels.

[2] The resulting model is not sensitive to how turnout rate R_{it} is defined, and has the natural property that scores *nearly* sum to one no matter the turnout rate.

Treatment in the annulus is somewhat correlated with treatment in the point-neighborhood, but appears unrelated to the voter's turnout decision, introducing minimal concern of bias to zero *[Weidman (2023)]*.

The model leaves only the variation in turnout shock and treatment shock *around* the cohort mean, testing how a treated voter's turnout changes relative to those who are untreated in their cohort. The core identifying assumption is that treatment must be independent of turnout *trends*. If true, within-cohort heterogeneity would be orthogonal to the error term. Under this parallel trends assumption, the true voter-election shock, α_{itc}, will not systematically differ from the cohort-election shock, α_{tc}, used in Equation (1): $(\alpha_{itc} - \alpha_{tc}) \perp \epsilon_{itc}$. In all three panels there is strong evidence against two remaining mechanisms for bias.

Restating the core identifying assumption in the context of residential sorting, a voter's turnout decision must not change their residential response to in-party treatment. The estimate would be biased upward if a change in the voter's actively level changes their probability of moving in response to treatment. This would not only positively bias estimates in the main period, but also positively bias estimates using treatment *after* the election, and negatively bias estimates soon after the *previous* election. Further, restating the core identifying assumption in the context of treatment, a voter's probability of treatment must not change after becoming more politically active. The estimate would be biased upward if recently engaged voters persuade their new neighbors to update their registration to reflect their new address or if they were to attract more politically aligned neighbors at a higher rate with their new yard signs. This bias would be weaker in the pre-period, when new neighbors will be less influenced to update their address in advance of an election in which the ballot has not even been established. They will also place fewer yard signs a year and a half in advance than half a year in advance. This bias would also be present in the post period. As discussed later, neither bias appears to be present.

IV. Results

This empirical approach allows testing the model's main predictions. First, *HT* suggests that β^k, the post-treatment turnout shock, should be positive for treatments within the voter's smallest point-neighborhoods. Figure R1 plots β^5 for each panel across the full set of model specifications, with standard errors clustered on the radius of the point-neighborhood. Figures for the other β^k will be included in the appendix. First, the estimate of voter fixed effects, $\hat{\alpha}_i$, captures the role of unchanging characteristics in both the voter's turnout and neighborhood, absorbing most of the observational relationship (*not shown*) in each panel. The voter's baseline turnout is highly related to their neighborhood composition, and estimating the annulus control, $\hat{\gamma}$, with $\hat{\alpha}_i$ absorbs very little to none of the estimate.

Second, the estimate of cohort-election fixed effects, $\hat{\alpha}_{tc}$, absorbs the component of the post-treatment turnout shocks common among cohort members. Except in the Presidential panel, the variation between *voters* explains more than the variation between *cohorts*. In Primary and Municipal panels, the post-treatment turnout response is largely explained by the unchanging characteristics of the voter, showing a larger estimate with $\hat{\alpha}_{tc}$ than with $\hat{\alpha}_i$. Since the cohort captures fewer voter-level characteristics than $\hat{\alpha}_i$, this suggests either that the strongest regional shifts occur in Presidential elections or that the variability of $\hat{\alpha}_i$ within the cohort is very small.

Third, since the cohort is defined using unchanging characteristics, adding $\hat{\alpha}_{tc}$ to $\hat{\alpha}_i$ helps to disentangle regional shifts from within-cohort variation. The marginal effect of adding $\hat{\alpha}_{tc}$ to $\hat{\alpha}_i$ substantially reduces the estimate in Presidential elections, pointing to strong regional shifts within the cohort in these elections. Regional shifts appear much weaker in Municipal elections, as most of the variation within the cohort is accounted for by the voter's unchanging characteristics.

Lastly, participation in the Primary panel appears to be most impacted by regional shifts. Voter fixed effects on their own produce a *negative* estimate, meaning that treated voters are *less* likely to participate in Primaries. Neighborhoods that receive the most in-party treatments experience a decline in turnout rates in Primary elections. However, the estimate is positive when comparing the turnout rates of treated and untreated voters in the cohort: treated voters turnout at higher rates than their untreated cohort members. Further, defining the cohort without demographics returns a much smaller estimate than when defining the cohort using demographics. This suggests that not only are trends heterogeneous between demographic groups, this heterogeneity on demographics is obscured when looking only at Census blocks. This demographically heterogeneous effect also appears to a smaller degree in Municipal elections. Not only does heterogeneity not bias up the estimate, but it hides a more positive underlying relationship. Adding $\hat{\gamma}$ to the $\hat{\alpha}_i + \hat{\alpha}_{tc}$ model absorbs no meaningful variation, further suggesting minimal concern that unobserved heterogeneity biases the estimate.

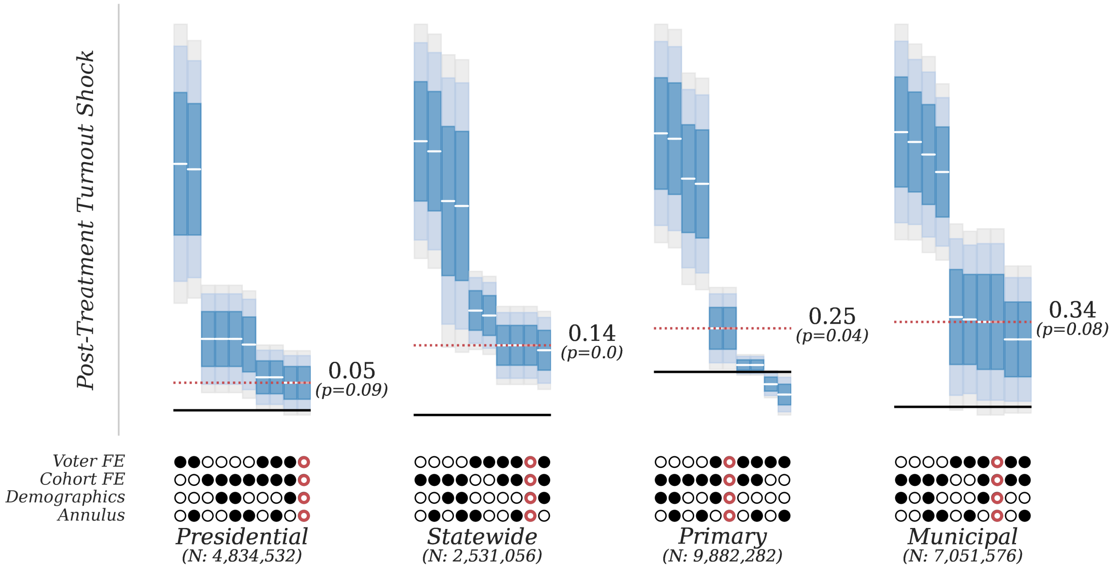

Figure R1. *Each panel's estimates of β^5 are shown across model specifications. The preferred specification is shown in the specification table. The estimate of the preferred specification is indicated by the red dashed line and printed with it's p-value beside the specification chart. Error bars represent standard errors, and significance levels of $p < 0.5$ and $p < 0.10$.*

The preferred specification estimates that after being treated with a new in-party neighbor in their point-neighborhood ($k = 5$), a voter will increase their turnout rate by roughly 1/3% in Municipal elections, 1/4% in Primary elections, and by a smaller rate in the Presidential panel. Estimates have greater significance using the higher treatment rates in $k = 10$.

Result 1. *Turnout rates increase after being treated with a new in-party neighbor.*

These positive estimates show a clear relationship between the new in-party treatments and the voter's turnout rate. However, to be interpreted as evidence for residential network spillovers, a residentially closer treatment must have a larger impact than one further away. Figure R2 plots the estimates using the preferred specification for increasingly large neighborhood sizes (k). The estimates for the post-treatment turnout response go to zero as k grows. The increase in turnout after an in-party treatment is largest within the smallest the point-neighborhoods and attenuates to zero as the size of the point-neighborhood increases.

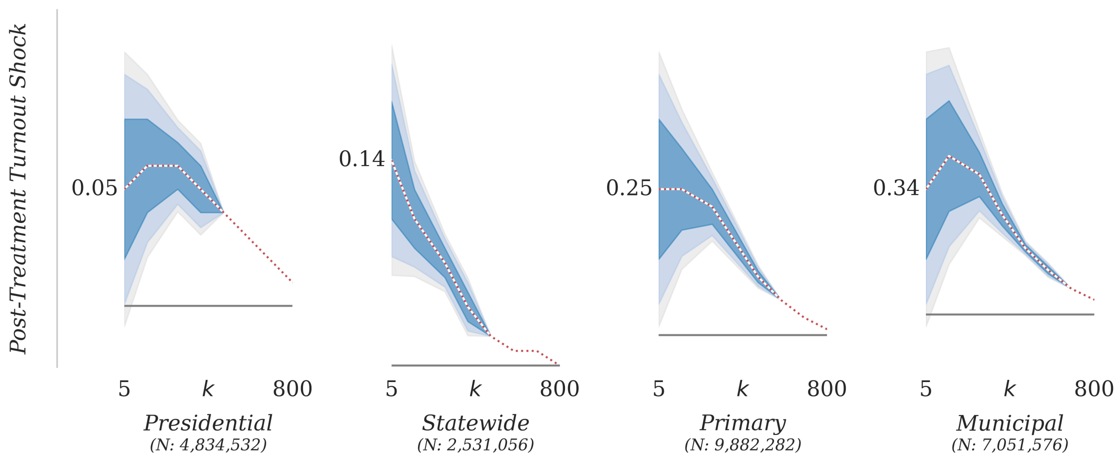

Figure R2. *Estimates of each panel's preferred model specification are shown across increasing k. The horizontal axis is scaled according to $ln(k)$. The kinks occur at each k estimates are produced:* $[5, 10, 25, 50, 100, 200, 400, 800]$. *Error bars represent standard errors, and significance levels of $p < 0.5$ and $p < 0.10$.*

This relationship between the role of the in-party treatment and residential proximity strongly suggests that the relationship in Result 1 is directly about neighbors and not about the broader environmental factors in the residential neighborhood. This attenuation limits the scope of plausible explanations to factors related only to the closest neighbors.

Result 2. *The impact of a new neighbor on turnout attenuates to zero in residential distance.*

This finding is consistent with the spillovers in *Gerber et al. (2008)*, who show that turnout norms have spillovers within the residential environment. But for these estimates to be truly about networks, a new neighbor of a different party cannot have the same positive effect, since they would not be as closely connected on the homophily network or transmit the same kinds of information. Figure R3 plots the main specification across election panels for both in-party and out-party treatments. In the panels with more political inflection (Presidential and Primary), the role of the in-party and out-party treatment are statistically different with differing signs. In Municipal elections, when the political inflection is very different, the estimate of an out-party treatment is indistinguishable from zero but is separated from the in-party estimate. The more significant estimates for the closest 10 reflect the larger sample of treatment in larger point-neighborhoods, with some smaller estimates for the closest 5 being smaller and higher variance.

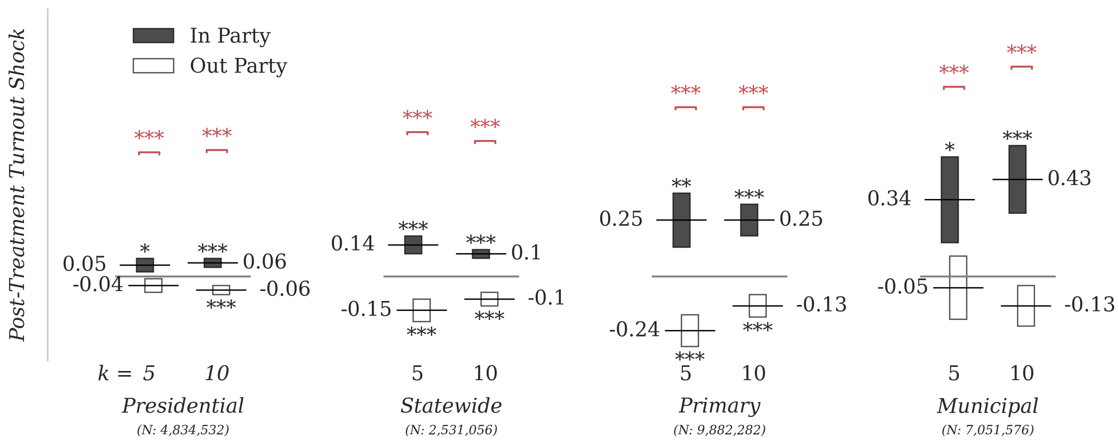

Figure R3. *Estimates across election types and party membership (standard errors, * p<0.1, ** p<0.05, *** p<0.01)*

Together, this strongly confirms that the party affiliation of the new neighbor, not just that a voter moved in next door, is an important part of the relationship to turnout.

Result 3. *The post-treatment turnout shock is smaller in out-party treatment, and negative in more party-salient panels.*

These point estimates of the post-treatment turnout shock, β, are largest in lower-information elections, like Municipal elections and Primaries, and weaker in higher-information elections, like Presidential elections. The left panel of Figure R4 systematizes the relationship between turnout and the estimate elasticity, β, by plotting the coefficients and standard errors for all three disjoint panels against their average in-sample turnout rates. The shaded region fits a second degree polynomial over the standard errors and point estimates. The negative relationship between the change in baseline turnout and the response to treatment strengthens the idea that the estimate of the post-treatment turnout response, β, is indeed related to the neighbors and not the broader political environment.

Estimating the post-treatment turnout response as an elasticity captures the idea that the impact of a new neighbor should be proportional to one's existing probability of turnout. But baseline turnout is very different across these elections. A post-treatment turnout response in absolute turnout would need to be larger in Presidential elections to have the same elasticity as in Municipal elections. The right panel of Figure R4 multiplies the estimate of the post-treatment turnout shock, β, by the panel's average in-sample turnout rate, R. This total response is no larger in Municipal than Presidential elections, with a somewhat higher total response in the Primary panel where turnout is nearer to 50%. However, while this is suggestive of a consistency in the total turnout response, it obscures the impact of treatment on the margin of the election. The most informative view interprets the post-treatment turnout response as being proportional to the election turnout rate, estimating the increase in the baseline turnout rate.

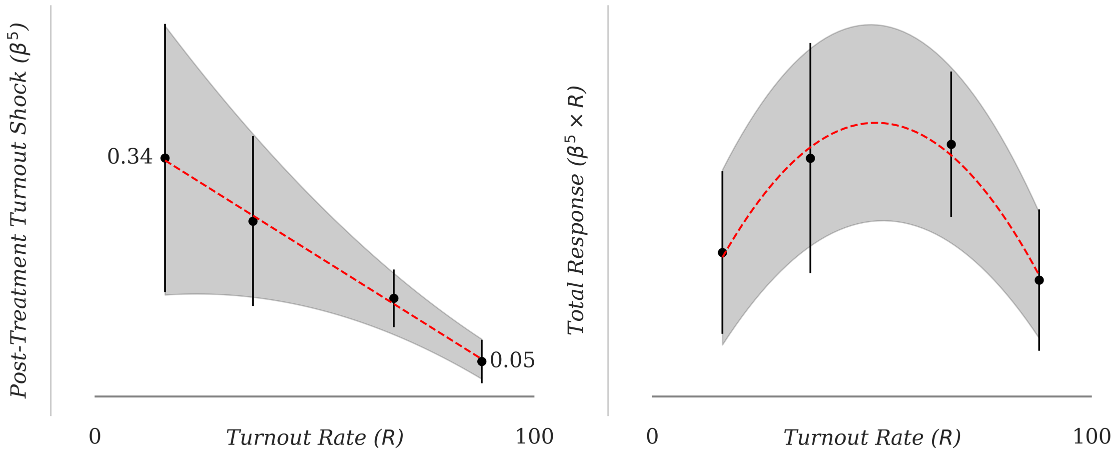

Figure R4. *Responsiveness and election size (standard errors)*

The negative relationship between the change in baseline turnout and the response to treatment strengthens the idea that the estimate of the post-treatment shock, β, is indeed related to the neighbors and not the broader political environment. If turnout were most responsive in high-profile elections like Presidential elections, the effect may be driven by factors like national news coverage or campaign advertising. This result points in exactly the opposite direction. Voters appear more responsive to their immediate neighbors in elections where they have *less* information about the candidates and issues at stake, relying more on the opinions and behaviors of those nearby. The fact that voters are most responsive in these elections suggests that it is truly related to the neighborhood and the social networks contained within.

Result 4. *Turnout rates are most responsive to treatment in lower turnout elections; total response is somewhat higher in elections closer to* 50% *turnout.*

These results may be driven by the flow of information and norms over two types of channels: 1) direct networks ranging from informal interactions to friendships, and 2) indirect networks like yard signs. The interaction between party affiliation and other dimensions of homophily, like race, can generate variation in the strength of the homophily networks to provide separation between indirect and direct networks. While varying the intensity of homophily networks within a party using the voter's race cannot fully rule out indirect mechanisms, it can strengthen the idea that HT is due to direct networks and not indirect mechanisms like yard signs. If HT is due to direct networks, the presence of a new in-party neighbor of the same race should have a stronger impact than a new in-party neighbor of a different race.

While the Republican Party is largely composed of White voters, the Democratic Party provides variation in the racial match between a voter and treatment, making it possible to test whether the strength of networks intensifies the estimate. This variation admits two related but distinct sets of tests. The first set tests whether the post-treatment turnout shock, β, is greater if the race of an in-party treatment is the same race as the voter, shown in Figure R5.B. If homophily networks shape the flow of information according to *HT*, then a same-race in-party treatment will have a greater impact than if the treatment were of a different race. The estimate of an in-party same-race treatment is larger than that of a in-party different-race treatment in many cases. A White voter will increase their turnout more after a White in-party treatment than a Black in-party treatment in all but the Presidential panel. A Black voter will increase their turnout more after a Black in-party treatment than a White in-party treatment in the Primary panel. This alone suggests the strength of direct homophily networks plays a role in *HT*.

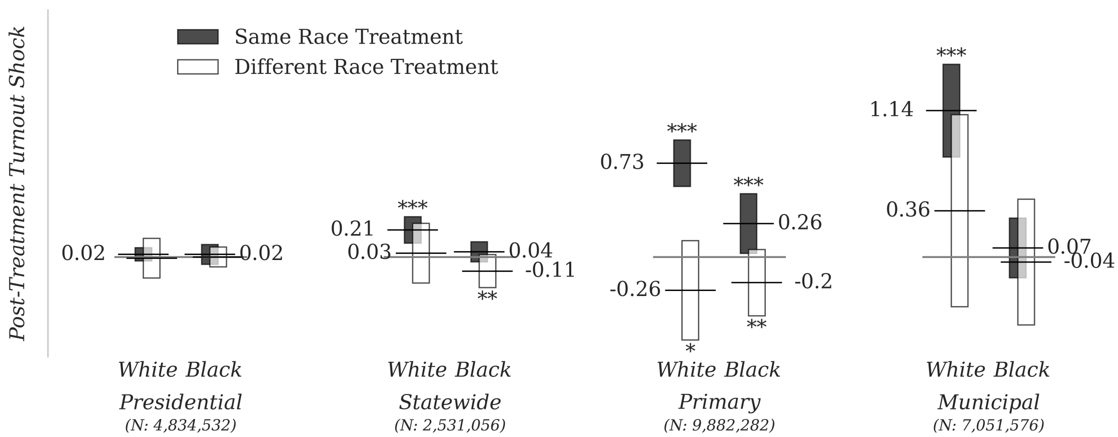

Figure R5.A *Racial variation within Democrats*
*(standard errors, * p<0.1, ** p<0.05, *** p<0.01)*

Despite these separations by race, there remains some possibility that differences in estimate could be due to the racial differences in political engagement of the new neighbors. For example, if White voters are more politically active than Black voters, it would not be surprising for a White voter to respond to a treatment by a White voter more than to a treatment by a Black voter. The second set tests whether the racial differences in treatment are the result of racial differences in the political engagement of new neighbors, shown in Figure R5.B. The voter's race indeed often is important in the post-treatment shock for the same new in-party neighbor. A White Democrat increases their turnout more than a Black

Democrat after the arrival of a new White Democrat in all but the Presidential panel. A Black Democrat increases their turnout more than a White Democrat after the arrival of a new Black Democrat only in Primary elections.

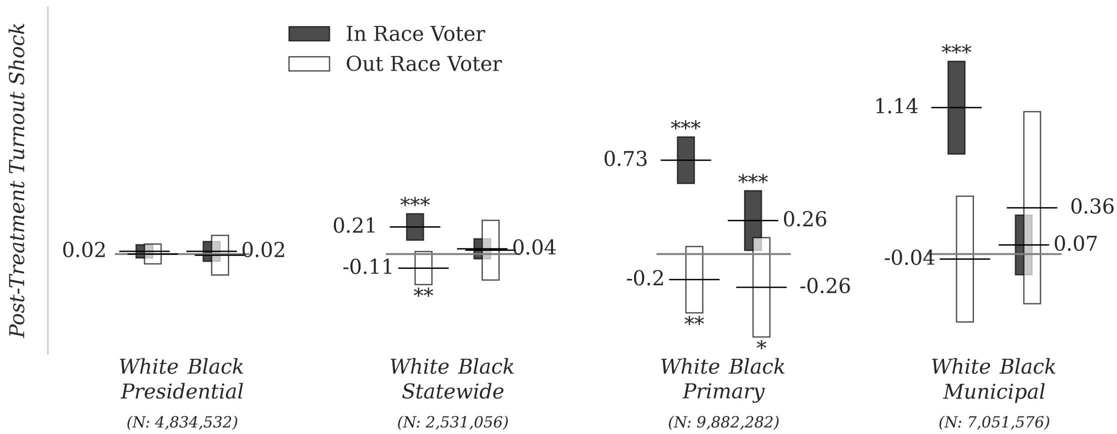

Figure R5.B *Racial variation within Democrats*
*(standard errors, * p<0.1, ** p<0.05, *** p<0.01)*

The separations on both margins suggest the intensity of the voter's homophily network on race matters for the impact of treatment on turnout. The effects are somewhat asymmetric, with White voters responding more than Black voters to the race of the in-party treatment. Taken together, these separations provide further support for *HT*: the denser the homophily network, the larger the estimate.

Result 5. *The impact of an in-party treatment is stronger within racial groups.*

These results are most consistent with direct networks but do not fully rule out the possibility that voters respond to a yard sign on the lawn of a same-race neighbor differently than the same yard sign on the lawn of a different-race neighbor. While more work is needed to fully separate the role of indirect and direct networks in *HT*, that intensity matters in this way strengthens the notion that it is the direct component of homophily networks that shape the flow of information, norms, and beliefs.

The cohort is defined by members of the same party. The model specifications with voter fixed effects fully absorb not just the effect shown in McCartney et al. (2021), that voters are more likely to make a residential move after an out-party voter arrives next door, but they absorb residential sorting on *any* unchanging characteristic. The model compares the turnout shocks of treated and untreated cohort members after absorbing their common shocks. A treated voter who decides to move out will not bias the estimate so long as their activity level is parallel to their cohort. There does not yet appear to be empirical evidence that even in-party treatment, like the out-party treatment in McCartney et al. (2021), is related to the voter's turnout decisions, which would not introduce bias, let alone to changes in their turnout decisions, which would.

Restating the core identifying assumption in the context of residential sorting, a voter's turnout decision must not change their residential response to in-party treatment. In other words, voters who move must turnout at rates parallel to their cohort. An upward bias would appear if a treated voter would have been more likely to move had they been *more* active. This type of bias could appear through attrition if a voter is more likely to move out of state (but not in state) after an out-party treatment, as these voters would appear to live at the address they left but would not vote. While voters' residential choices are certainly related to changes in the neighborhood, it is difficult to conceive of a mechanism in which the distance of the voter's move is related to differences in neighborhood changes between cohort members.

Residential sorting could also positively bias the estimate if a change in the voter's actively level changes their probability of moving in response to treatment. Voters may change their residential choices according to their activity level, with active voters being more likely to stay after an in-party treatment. This violation of parallel trends would bias up the estimate using treatment immediately preceding an election, but to the same degree would bias down the estimate using treatment right after the *previous* election. If the residential response to treatment depends on the voter's turnout, a voter having just voted will use that decision, not a shock in an unestablished future election. While this type of residential selection could be tested directly, the empirical evidence in this area does not resolve whether it should be a concern.

On the treatment side, an upward bias could appear if recently engaged voters persuade their new neighbors to update their registration to reflect their new address. This would lead to a higher rate of treatment than they would have experienced otherwise simply because they became more politically active.[3] A similar upward bias would appear if a voter who becomes politically engaged before an upcoming election attracts more politically aligned neighbors at a higher rate with their new yard signs. This issue weakens as the time between treatment and turnout grows, limiting the possibility that the next turnout shock could be reverse-causally related to treatments right after the previous election. This is for a number of reasons. Anticipating a future shock is more difficult when the future is further away. New neighbors will be less influenced to update their address in advance of an election in

[3]This is despite using only voters having moved from a previously registered address in the state.

which the ballot has not even been established. Fewer new yard signs are placed a year and a half in advance than half a year in advance. In this way, a violation of parallel trends would appear in the year after the election, with treatment more likely among voters who just increased their turnout.

Figure R6 reports estimates using the variation in the timing of treatment for $k = 5$, separating HT from residential selection and the reverse impact of a turnout shock on treatment. Treatment in the *pre* period occurs in the one year before the baseline estimates. Treatment in the *post* period occurs after the election. If the estimates are due to HT and not these two biases, estimates should vary with the timing of treatment in two ways. First, later treatments should not have a larger post-treatment turnout shock. Since the pre-period treatment is separated by one additional year, the new neighbor will experience more difficulty anticipating the voter's *future* election shock. A voter living nearby for half a year should have no more of an impact than a voter living nearby for a year and a half. The earlier the treatment, the weaker the relationship between the shock and other factors like yard signs and the updating of other neighbors' registrations. Further, residential selection would negatively bias estimates using treatment directly following the earlier of two elections. Figure R6 shows that the estimates are no greater when treatment happens nearer the second of any two elections. In all three panels, earlier treatments have an estimate that is not weaker. While the duration of treatment doesn't appear to be as important in Municipal elections as in Presidential and Primary elections, this offers strong support for the direction of causality going from the new neighbor to the voter's turnout decision.

Second, treatment *after* the election should be zero. A positive estimate using treatment after the election would violate the parallel trends assumption, showing that voter's who received a positive shock are more likely to be treated. Both residential selection and a registration effect would produce an upward bias in the post-period. As shown in Figure R6, turnout shocks are not positively related to future treatment, more evidence consistent with the parallel trends assumption.[4] Treatment *after* the election has no relationship to a turnout shock, minimizing the possibility that voters increase their residential response to treatment after a turnout shock or that recently engaged voters encourage their neighbors to register. Together, this strongly supports the interpretation that the arrow of causality goes from treatment to turnout shock.

Third, Statewide elections see a somewhat opposite relationship. Statewide elections observe a smaller estimate if treatment occurs *prior* to the election, and grows stronger *after* the

[4] Although Primaries have a somewhat negative estimate in the post-period.

election. This suggests Statewide elections might serve as a means of *registering* new in-party neighbors. This suggests reverse causality, where voters are more likely to be treated after having received an election shock. Since this analysis is restricted to voters who have moved, it suggests some of the baseline estimate is due to residential sorting or the re-registering of existing new neighbors. In Statewide elections new neighbors may either be inspired by their newly active neighboring voters to update their registration from their past address or choose residence based on things like yard signs. This doesn't rule out *HT* in Statewide elections, but suggests a role for this type of reverse causality.

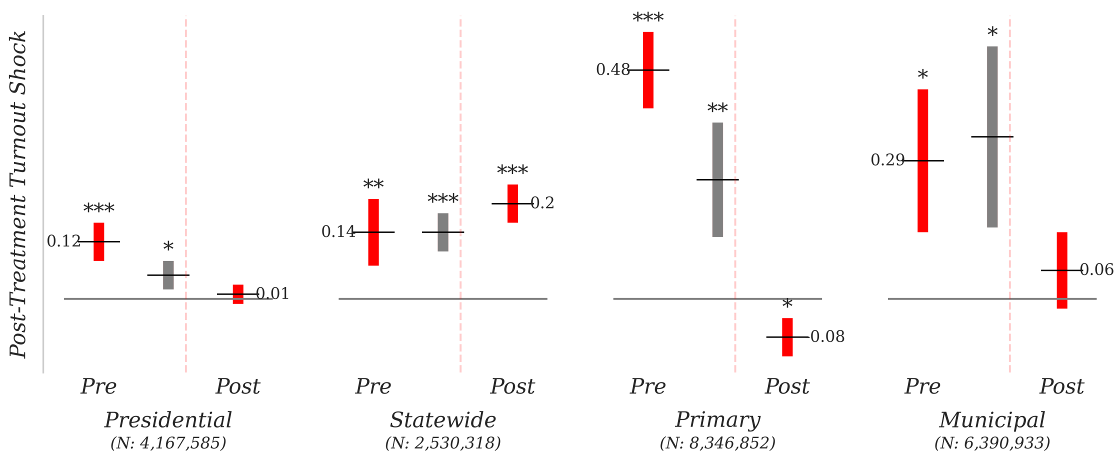

Figure R6. *Timing of treatment*
*(standard errors, * p<0.1, ** p<0.05, *** p<0.01)*

It does not appear that residential sorting or a registration effect drives the estimates in Presidential and Primary elections, when it would be most salient. All three panels experience a non-increasing relationship between the duration of treatment and post-treatment turnout response, and no positive response to treatment after the election, evidence consistent with parallel trends.

Result 6. *The post-treatment turnout response is non-increasing in treatment duration, with no reverse causality in Presidential, Primary, and Municipal panels.*

V. Conclusion

Network spillovers and social norms have long been studied in the emergence of equilibrium, with a growing body of theoretical and empirical connections to the turnout decision. The *Homophily Turnout* framework suggests a particular shape to network spillovers in voter turnout, which this paper develops and empirically tests. The spatial structure of the new administrative panel provides voter-level variation in time and space able to observe changes in the voter's party, address, point-neighborhood, and turnout decisions. The observation that new neighbors will have difficulty anticipating the difference in *future* turnout shock among two nearby observationally identical voters makes it possible to avoid the core empirical difficulties of selection and the reflection problem. An empirical approach using voter fixed effects and cohort difference-in-differences isolate the role of homophily networks in the turnout decision.

The empirical results confirm many of the framework's predictions about the shape of residential network spillovers in turnout behavior. First, a voter will increase their turnout rate by roughly 1/3% in Municipal elections after being treated with a new in-party neighbor in their point-neighborhood. Second, these coefficients attenuate to zero as the residential distance to treatment grows, limiting the scope of plausible explanations to those in the voter's local residential neighborhood. Third, the post-treatment turnout shock is negative or zero when treated with an out-party voter, meaning the response to a new neighbor depends on the political alignment of the voter and their treatment. Fourth, the voter's post-treatment turnout responsiveness is highest in lower turnout elections, like Municipal elections, and lowest in high turnout elections, like Presidential elections, while the total response is greatest in elections nearer to 50% in-sample turnout rates. This suggests voters are more responsive to their neighbors in elections with fewer competing information channels, strengthening the idea that HT is truly about the neighborhood and the social networks contained within. Fifth, party affiliation isn't the only dimension of homophily that matters. Voters often experience a greater post-treatment turnout shock after an in-party treatment that is also of the same race, further strengthening the idea that HT is due to direct network effects and not other information channels like yard signs. Lastly, treatment more than a year prior to the election produces estimates no weaker, and treatment after the election produces null or slightly negative estimates. This limits the scope of plausible explanations to a narrow window of time immediately following treatment and rules out residential sorting and a registration effect.

These findings go to the heart of why voters turnout in elections with no meaningful impact on the outcome. The *Homophily Turnout* framework suggests a particular shape to behavioral spillovers in residential networks that is consistent with the empirical and theoretical results for social norms in the turnout decision. These results emphasize that the shape of homophily networks around residential proximity, political alignment, and racial similarity matters in the flow of information, norms, and beliefs over homophily networks, particularly for voting behavior. As the electorate has clustered on characteristics like party and race,

HT also suggests the possibility for impacts on aggregate turnout with this type of spatial reorganization. These observations, framework, and approach provide a better understanding of voter behavior and also suggest the possibility of other spillovers within residential networks in our public decisionmaking process.

References

Bertrand, M., Luttmer, E. F. P., and Mullainathan, S. (2000). Network effects and welfare cultures. *The Quarterly Journal of Economics*, 115(3).

Carrell, S. E., Hoekstra, M., and West, J. E. (2011). Is poor fitness contagious?: Evidence from randomly assigned friends. *Journal of Public Economics*, 95(7-8).

Chen, M. K. and Rohla, R. (2018). The effect of partisanship and political advertising on close family ties. *Science*, 360(6392).

Coate, S. and Conlin, M. (2004). A group rule-utilitarian approach to voter turnout: Theory and evidence. *American Economic Review*, 94(5).

Cui, L., Favarelli, and Walsh, R. P. (2015). Foreclosure, vacancy and crime. *Journal of Urban Economics*, 87.

Currarini, S., Jackson, M. O., and Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4).

DellaVigna, S. and Kaplan, E. (2007). The fox news effect: Media bias and voting. *The Quarterly Journal of Economics*, 122(3).

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of political economy*, 65(2).

Favarelli, M. and Walsh, R. P. (2015). Mandate and paternalism: A theory of large elections. *Games and Economic Behavior*, 93.

Feddersen, T. and Sandroni, A. (2006). A theory of participation in elections. *American Economic Review*, 96(4).

French, D. (2020). *Divided We Fall: America's Secession Threat and How to Restore Our Nation*. St. Martin's Press.

Galesic, M., de Bruin, W. B., Dumas, M., Kapteyn, A., Darling, J., and Meijer, E. (2018). Asking about social circles improves election predictions. *Nature Human Behavior*, 2.

Gerber, A. S., Green, D. P., and Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review*, 102(1).

Harsanyi, J. C. (1980). Rule utilitarianism, rights, obligations and the theory of rational behavior. *Theory and Decision*, 2(1).

Igielnik, R., Keeter, S., Kennedy, C., and Spahn, B. (2018). Commercial voter files and the study of u.s. politics. *Pew Research Center*.

Kaplan, E., Spenkuch, J. L., and Sullivan, R. (2022). Partisan spatial sorting in the united states: A theoretical and empirical overview. *Journal of Public Economics*, 211(104668).

Levine, D. K. and Mattozzi, A. (2020). Voter turnout with peer punishment. *American Economic Review*, 110(10).

McCartney, W. B., Orellana, J., and Zhang, C. (2021). Sort selling: Political polarization and residential choice. *FRB of Philadelphia Working Paper*, 21(14).

Sunstein, C. R. (1999). The law of group polarization. *University of Chicago Law School, John M. Olin Law and Economics Working Paper*, 91.

Sussel, J. (2013). New support for the big sort hypothesis: An assessment of partisan geographic sorting in california, 1992–2010. *PS: Political Science and Politics*, 46(4).

Weidman, T. J. (2023). Panel nearest neighbors: A new algorithm applied to a realigning electorate. *Working Paper*.

Wilkinson, W. (2019). The density divide: Urbanization, polarization, and populist backlash. *Niskanen Center*.