

ORIGINAL ARTICLE

# Does residential sorting explain geographic polarization?

Gregory J. Martin<sup>1\*</sup> and Steven W. Webster<sup>2</sup>

<sup>1</sup>Stanford Graduate School of Business, Stanford, CA, USA and <sup>2</sup>Weidenbaum Center on the Economy, Government, and Public Policy, Washington University, St. Louis, MO, USA

\*Corresponding author. Email: gjmartin@stanford.edu

## Abstract

Political preferences in the United States are highly correlated with population density, at national, state, and metropolitan-area scales. Using new data from voter registration records, we assess the extent to which this pattern can be explained by geographic mobility. We find that the revealed preferences of voters who move from one residence to another correlate with partisan affiliation, though voters appear to be sorting on non-political neighborhood attributes that covary with partisan preferences rather than explicitly seeking politically congruent neighbors. But, critically, we demonstrate through a simulation study that the estimated partisan bias in moving choices is on the order of five times too small to sustain the current geographic polarization of preferences. We conclude that location must have some influence on political preference, rather than the other way around, and provide evidence in support of this theory.

Speaking at the 2004 Democratic National Convention, Barack Obama, then a candidate for the US Senate, famously declared that “there’s not a liberal America and a conservative America; there’s the United States of America.” Obama then went on to decry political pundits who “like to slice and dice our country into red states and blue states” (Obama 2004). The implication of Obama’s speech was that the perception of geographic polarization of the country into reliably Democratic and Republican areas was not based in fact, but was instead a false narrative imposed by the media.

Obama’s rhetoric of unity and homogeneity across party lines notwithstanding, there is substantial evidence that cultural and lifestyle preferences correlate strongly with political tastes. If such correlations between political and lifestyle preferences are strong enough, the implication is clear: we should expect to find a geographically polarized electorate of exactly the kind that Obama’s speech sought to deny. This pattern is indeed strikingly evident in American voting behavior: locations with higher population density—and associated lifestyle characteristics such as smaller housing units, proximity to cultural amenities, and access to mass transit—have higher proportions of Democratic voters than do locations with lower density—and associated larger homes, more greenery and greater privacy from neighbors (Rodden 2010).

What is the causal relationship underlying this robust empirical pattern? Does it emerge, as some scholars have suggested, as a result of “residential sorting:” partisans’ choices to preferentially move to neighborhoods with like-minded neighbors (Bishop 2008)? Or might the causal arrow be reversed, with the physical and social environment of one’s residential location influencing one’s political beliefs? In this paper, we contribute empirical and simulation-based evidence measuring *how much* of the observed geographic polarization in political preferences can be explained by the former mechanism of selection into politically congruent neighborhoods. We show that residential sorting, on its own, cannot explain a level of geographic polarization anywhere close to that observed in the US today, implying that the latter mechanism must have some empirical power.

Our analysis begins with a direct measurement of the influence of political preferences on voters' moving decisions, relative to nonpolitical factors. We precisely measure partisan biases in location choice on both direct political attributes and partisan-correlated nonpolitical attributes, using two large-*N* datasets of registered voters drawn from public voter files. We then test, through a simulation analysis, whether the partisan bias in moving decisions that we measure is sufficient to sustain the observed level of geographic polarization. Our simulation also allows us to assess the effect of residential mobility on two important normative standards of democratic performance: electoral competitiveness and legislative malapportionment. We conclude by examining whether individuals, conditional upon moving, tend to adopt the partisan makeup of their new locale.

Our results suggest that political tastes are, in fact, predictive of where individuals choose to move. However, compared to nonpolitical factors, political preferences have relatively small influence on residential location choices. Registered Democrats moving within state, for instance, on average move to locations about five log points more dense than observably comparable registered Republicans, a differential which, though precisely measured and statistically different from zero at the 99 percent level, is less than 4 percent of 1 SD of the change in log density of within-state moves in our sample.

The small partisan bias in moving decisions that we measure is not sufficient even to sustain, much less to increase, the level of geographic polarization that we observe in the American electorate today. Americans move frequently enough, and the partisan biases in location choice we measure are small enough, that repeated rounds of sorting quickly homogenize the geographic distribution of partisanship. In a hypothetical world where individuals' partisan identities were fixed and residential sorting the only causal force operating on the geographic distribution of preferences, the equilibrium level of geographic polarization would be much lower. In fact, a calibration exercise shows that maintenance of the current level of geographic polarization by sorting alone would require partisan biases in location choice to be roughly five times larger than what we actually observe in the data.

Relatedly, our simulation results also show that the effects of residential mobility on the normative performance of elections are benign. Residential mobility, on its own, increases the fraction of competitive districts, reduces the likelihood of legislatures whose composition differs substantially from that of the voting population and increases local political heterogeneity. To the extent that these criteria have worsened in the United States in recent decades, the culprit lies elsewhere.

Finally, we conclude by presenting data consistent with the idea that one's choice of geographic residence influences her political tastes. Our analyses indicate that voters who move to less politically congruent locations are more likely to subsequently change their party affiliation to match that of the new location. Thus, our results suggest that social influences exert influence on individuals' political preferences.

### **Partisanship, lifestyle preferences, and geographic polarization**

Political scientists have demonstrated that political ideology and party ID are predictive of a variety of lifestyle and consumption preferences. These political factors have explanatory power in areas as varied as mate selection (Hitsch, Hortacsu and Ariely 2010; Huber and Malhotra 2017); media consumption (Prior 2007; Levendusky 2009); and housing decisions (Tam Cho, Gimpel, and Hui 2013; Gimpel and Hui 2017; Mummolo and Nall 2017). Rodden's (2010) survey shows that this predictive power extends to geography: both within and across metropolitan areas, and at surprisingly small geographic scales, locations with higher population density have higher proportions of Democratic voters than do locations with lower density.

Geographic polarization—the phenomenon wherein partisans cluster together in politically homogeneous neighborhoods—is not merely a political curiosity. In fact, it has serious

consequences for at least three normative standards of democratic performance. First, increasing the local homogeneity of citizens' political preferences will tend to produce more homogeneous legislative districts, leading to more districts where election outcomes are in little doubt. As electoral competitiveness is an essential ingredient in voters' ability to exert control over their representatives (Ferejohn 1986; Gordon, Huber, and Landa 2007), geographic polarization poses a potentially serious threat to elections' accountability function. Second, the specific pattern of one party's voters concentrating themselves in densely populated areas generates a bias in representation that reduces that party's seat share in a single-member-district legislature, relative to its prevalence in the overall electorate (Chen and Rodden 2013). This bias makes the composition of Congress and state legislatures unrepresentative of voters' aggregate partisan preferences. Third, geographic polarization makes it less likely that citizens encounter others whose political views differ from their own in their daily lives. Hence, it is a potential cause of increasing affective polarization in the United States since the 1980s.

Given the substantive importance of geographic polarization, political scientists have posited several theoretical explanations for the phenomenon. In one prevalent theory, lifestyle tastes are the exogenous factor driving the emergence of the observed geographic pattern of partisanship. Political tastes are simply brought along for the ride as conservatives pining for a three-car garage move to suburbs and liberals wanting to walk to work move to central cities. An alternative theory, originating in the sociological concept of homophily, posits that desire to live among politically like-minded neighbors is the generative force driving geographic sorting (Bishop 2008; Lang and Pearson-Merkowitz 2014). In this conception, partisans are actively "seeking politically compatible neighbors" when making choices about where to live (Gimpel and Hui 2015), rather than simply happening to find themselves living among co-partisans who share their tastes for non-political housing characteristics.

Though they differ on the ultimate cause, both theories rely on a common element: the willingness and ability of partisans to differentially move away from areas that are a worse match to their tastes—political or lifestyle—and into areas that are a better match. If voters' residential choices are highly constrained, then no matter how strong is the correlation between lifestyle and political tastes, or how strong is the desire to live near others with similar political preferences, residential sorting cannot produce a geographically polarized pattern. The degree to which voters' housing choices are constrained by external factors, then, is a crucial determinant of the ability of residential sorting to generate geographic polarization.

A few closely related works directly examine constraints on partisans' residential sorting. Nall (2015) argues that the construction of the Interstate Highway System, by opening up new tracts of land for development that were previously too far away to allow commuting to central-city jobs, generated an abrupt expansion in the feasible set of housing options available to white-collar workers. This dramatic relaxation of constraints on sorting generated a measurable increase in partisan segregation along "an urban-to-suburban continuum." Mummolo and Nall (2017) show that while Democrats and Republicans have different preferences over housing and neighborhood characteristics, their flexibility in neighborhood choice is limited by universal desires like affordability, low crime rates, and quality of education.

Unlike Mummolo and Nall (2017) and other studies that use survey data to measure the degree to which partisans would *like to* sort in a hypothetical situation, our analysis measures how much they *do* sort in reality. For instance, we show that the partisan differential in stated preference for the ability to walk to work measured by Mummolo and Nall's (2017) survey corresponds to a partisan differential in behavioral outcomes: preference for moving to areas with high Walk Scores.<sup>1</sup> Relative to the partisan difference in unconstrained stated preferences, however, the partisan difference in actual observed moving choices is very small.

<sup>1</sup>Walk Scores are a proprietary measure—produced by the Redfin real estate listing service—that aims to measure the "walkability" of a geographic location on a 0–100 scale, by computing the number of restaurants, retail locations, offices,

We build on the observation that partisans' housing choices are constrained to show, via a series of simulations, that the constraints on sorting we measure are tight enough to make residential mobility effectively a de-polarizing force in American politics. Our simulation analysis translates the measured partisan bias in moving decisions into quantifiable substantive effects on the performance of elections. We show that in spite of measurable partisan differences in location preference, residential mobility, on net, has positive implications for the accountability and representativeness of electoral institutions and increases the preference diversity of local political units.

### Measuring partisan sorting and its determinants

To begin, we briefly describe the data employed in our analyses. The primary datasets are drawn from public voter registration files in states which have partisan registration. There are two main advantages of this data source. First, we observe very large sample sizes, in the millions of individual voters, which allows precise estimates of the quantities of interest: partisan influence on moving patterns. As will become clear in the simulation section, estimating the precise magnitude of influence of partisanship on moving choices is critical for understanding the aggregate electoral consequences of geographic sorting: small effects and large effects, even if they point in the same direction, can produce very different implications for substantive outcomes. Second, our data allow us to identify the same individual voter's residential location before and after a move, as well as his or her party of registration before and after. This individual-level data avoid problems of ecological inference that would plague an analogous design relying on local aggregates.<sup>2</sup>

Our data comprise four primary sources: (1) a random sample of individual registered voters who moved from one state to another prior to the 2016 election cycle; (2) the universe of registered voters in the state of Florida between 2006 and 2012, among which our primary analyses focus on the subset who moved from one residence to another between 2008 and 2010; (3) precinct-level vote totals for the 2008 elections; and (4) population characteristics from the US Census bureau, generally at the level of the census tract. The first dataset contains 50,000 individual voters who moved across state lines, sampled from the database of the political consulting firm and data vendor Catalist, which is privately maintained but relies on public voter files. These voters moved between states with partisan registration, such that we observe party of registration on both ends of the move; moves in this data occurred between 2005 and 2016.<sup>3</sup> The second dataset is drawn directly from the publicly available voter files maintained by the Florida secretary of state, and comprises a total of 1,435,698 voters<sup>4</sup> who appeared in both 2008 and 2010 voter files and whose residence address changed between 2008 and 2010. The third dataset is taken from Ansolabehere, Palmer, and Lee (2014); the fourth were downloaded from the database maintained by the Minnesota Population Center, and appended with additional information we collected. Detailed descriptions of the datasets, as well as summary statistics and details of the data cleaning and joining process, are presented in the Online Appendix.

### Estimating the magnitude of partisan residential sorting

We first measure the strength of partisan influences on moving decisions, conditional on other observable characteristics of the voter and of their initial place of residence. We estimate

---

parks, etc. within walking distance of a location. Walk Scores are commonly used by real estate agents, buyers, and renters to evaluate homes and apartments.

<sup>2</sup>With aggregate-level data, we would see only an average rate of in-migration and out-migration and an average rate of change in voting preference, without observing the actual composition of the moving population.

<sup>3</sup>The Online Appendix contains information on the distribution of origin and destination states in the Catalist data.

<sup>4</sup>Because our dependent variables are measured at either the census tract or precinct level, in most of our analyses we include only movers who moved across precinct or tract boundaries: 1.12 million individuals moved across precincts and 1.06 million moved across tracts.

regressions of this form for both within-state (Florida data) and cross-state (Catalist data) moves. We focus attention on three dependent variables: the logged population density, the Walk Score, and the 2008 Republican presidential vote share of the voter's *destination* precinct.

We focus on these dependent variables because each sheds light on a different aspect of the motivating question. Measuring the degree to which Democrats are attracted to (and Republicans repelled by) high-density precincts when they move allows us to estimate the contribution of residential mobility to the empirical correlation between density and partisanship, and the associated malapportionment of seats in Congress or state legislatures. Walk Scores are a highly salient housing characteristic on which liberal and conservative voters' preferences appear to diverge sharply (Mummolo and Nall 2017); hence, measuring the differential impact of partisanship on choice of Walk Score allows us to measure the degree to which divergence in stated preference in surveys translates into divergence in behavioral outcomes. Finally, the difference in Republican presidential vote share of the destination precinct between Democratic and Republican voters is informative about the increase in district homogeneity, and associated decline in competitiveness, that can be attributed to residential sorting. Comparison of the estimated effects in this regression with those in the Walk Score regression additionally allows us to compare the performance of the housing-characteristic-driven versus partisanship-driven theories of partisan sorting.

In all regressions, we include on the right-hand side the value of the dependent variable for the same individual's *origin* precinct. The logic is that the origin precinct value captures unobserved preferences or constraints of the individual that are likely to be preserved both before and after a move. What we are attempting to measure are differential effects of partisanship on moving decisions that predict variation in the outcome variables above and beyond what would be expected if voters' post-move neighborhoods simply preserved the characteristics of their pre-move neighborhoods.

Each analysis is conducted with no fixed effects, with fixed effects for the voter's destination county of residence, and with fixed effects for the voter's destination zip code of residence. The logic for including fixed effects for destination geographies is to understand whether even among voters who moved to the same county (zip code), registered partisans chose measurably different neighborhoods within that county (zip code). Thus, we estimate the following empirical models:

$$y_{i,\text{post}} = \alpha y_{i,\text{pre}} + \beta_D D_{i,\text{pre}} + \beta_R R_{i,\text{pre}} + \gamma_1 X_{i,\text{pre}}^1 + \gamma_2 X_{i,\text{pre}}^2 + \delta' Z_{i,\text{post}} + \varepsilon_i \quad (1)$$

where  $y_{i,\text{post}}$  and  $y_{i,\text{pre}}$  are one of the three dependent variables described above before or after the move, respectively;  $D_{i,\text{pre}}$  and  $R_{i,\text{pre}}$  are dummies for Democratic and Republican partisan registration prior to the move (at most one of which can equal one);  $X_{i,\text{pre}}^1$  are individual-level and  $X_{i,\text{pre}}^2$  are tract-level covariates for voter  $i$  or her origin tract of residence; and  $Z_{i,\text{post}}$  is a set (possibly singleton, in the pooled regression) of dummies for destination geography of residence.

The three sets of results for within-state moves, with various combinations of fixed effects and control variables, are presented together in visual form in Figure 1.<sup>5</sup> Analogous results for cross-state moves are in Figure A3 and Tables A8–A10 in the Online Appendix. The vertical labels on the right-hand side of the coefficient plots denote the dependent variable in each model.

As the pattern of directionality and statistical significance for the Catalist sample is extremely similar to that reported for the Florida movers data, the discussion that follows focuses on the Florida estimates. The only significant difference between the two sets of results is that the magnitudes in specifications without fixed effects for destination geographies are much larger in the Catalist sample. This is a consequence of the fact that there is much more possible variation across all of the 29 partisan registration states than within the single state of Florida. Once we add fixed effects so that we compare only among voters moving to the same county or zip code, the magnitudes become very similar to those reported for the Florida movers dataset. This replication of results in two independent large samples demonstrates that the effects we measure are not specific to the state of Florida or to the 2008–10 time period.

<sup>5</sup>Results are also presented separately in tabular form in the Online Appendix Tables A5–A7.

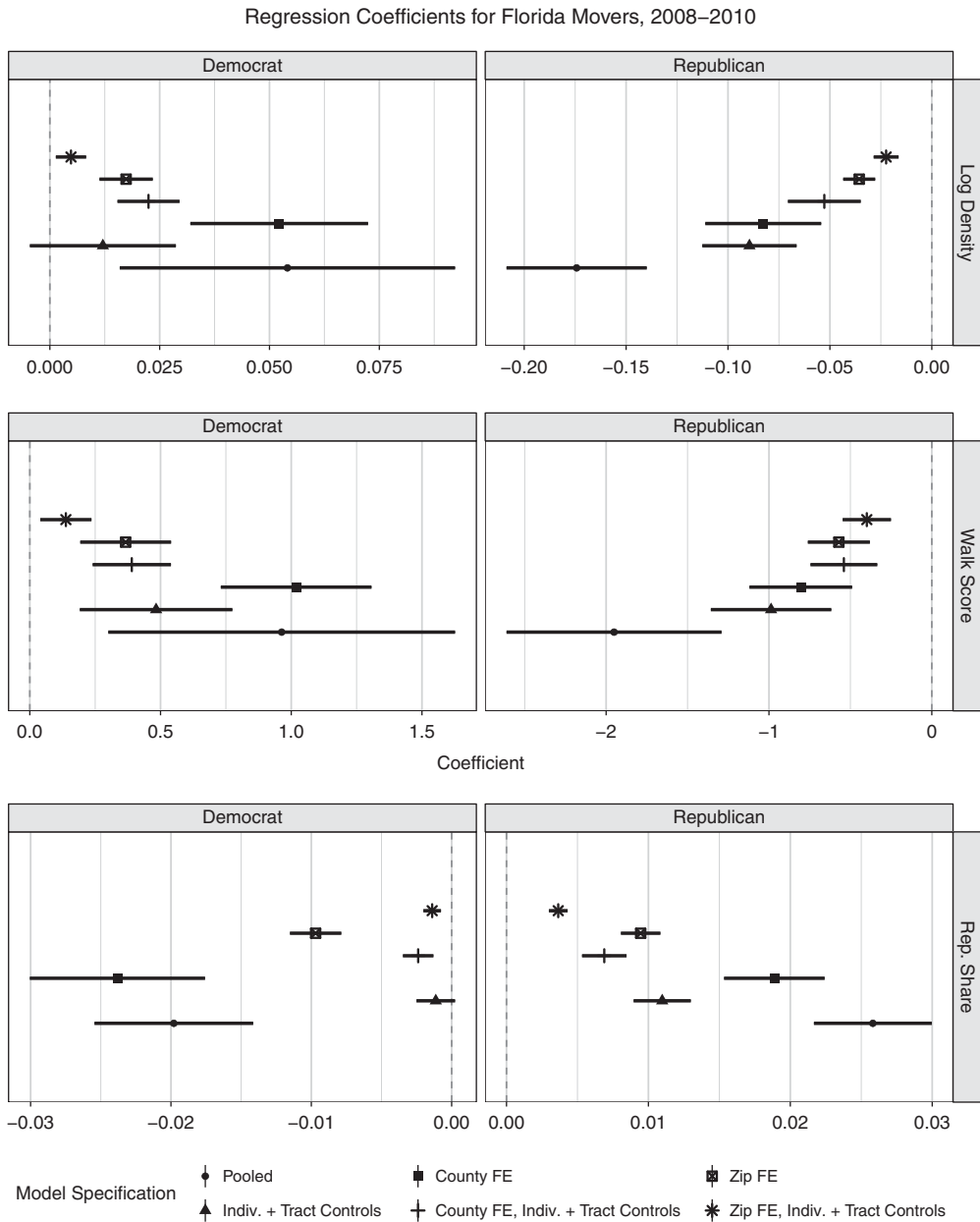


Fig. 1. Coefficient estimates for the Florida movers sample.

**Density**

In the first model shown in Figure 1, the dependent variable is the log density of the Census tract to which an individual moved in 2010. The figure displays coefficient estimates and 95 percent confidence intervals<sup>6</sup> of dummy variables for Democratic and Republican partisan registration in 2008.<sup>7</sup>

<sup>6</sup>Because the “treatment” variables are assigned not at the individual level but at the geographic-area level, in all analyses we compute cluster-robust standard errors, using 2010 county of residence as the clustering variable.

<sup>7</sup>Registered voters without a stated partisan affiliation in 2008 are the excluded category.

The results show that there is a measurable partisan correlation in moving decisions. According to the base model estimates, a registered Republican voter is expected to move to a new location that is 17.5 log points (i.e., approximately 17.5 percent) *less* dense than an independent voter moving from a comparably dense origin location. Registered Democrats move to locations 5.3 log points *more* dense than independents from comparably dense origins. The magnitudes decline somewhat when individual- and origin-tract-level covariates are included, to  $-8.9$  and  $+1.2$  log points, respectively; with errors clustered at the county level only the Republican dummy remains significantly different from zero. For comparison, the 25th percentile census tract in Florida has log population density of 6.04; the corresponding 75th percentile is 7.51. The observed partisan differences are thus fairly small compared to the overall variation present in the data.

When we include fixed effects for destination geographies, the magnitudes decline—mechanically, because there is by definition less possible variation in density within a particular county, and even less within a particular zip code, than there is within the entire state<sup>8</sup>—but precision increases substantially. We can reject the hypothesis that  $\beta_R$  and  $\beta_D$  are zero for all specifications at the 99 percent level. The magnitudes indicate that if we observed two voters, one a registered Republican and the other a registered Democrat, moving to the same zip code from origin locations with similar observable characteristics, we would expect the Republican voter to choose a tract approximately three log points less dense than the Democrat.

The magnitude of these differences is small relative to the possible variation in density available, which suggests that the ability of partisans to sort on population density is highly constrained by other factors. Nonetheless, the partisan differences in location choice are statistically significant and all in the direction that would tend to amplify the existing correlation between partisanship and density. Repeated over several election cycles, this observed rate of sorting has the potential to meaningfully increase the concentration of Democrats in dense areas and attendant malapportionment of legislative seats, a hypothesis which we evaluate through a simulation exercise presented in the next section.

### Walk Scores

The second set of results depicted in Figure 1 focuses on the Walk Score of the voter's destination precinct. The specification is, again, as shown in Equation 1, with post-move (2010) Walk Score appearing on the left-hand side and the pre-move (2008) value of Walk Score appearing on the right-hand side of the equation. Mummolo and Nall (2017) find that there are significant partisan differences in preferences over the walkability of a neighborhood, with Democrats valuing this neighborhood attribute relatively heavily and Republicans relatively much less. Conditional on a move, then, we expect Democrats to be more likely to select into more-walkable neighborhoods and Republicans to be more willing to sacrifice this characteristic for other desirable features such as home size or affordability.

Much as in our models of log density, the results show a distinct partisan correlation with preference for high Walk Scores. The models in the second row of Figure 1 show that Democratic partisans tend to sort into precincts with higher Walk Scores and Republican partisans tend to move to precincts with lower Walk Scores. Even among the set of voters who moved to the same county (zip code), the Republicans sought out less walkable precincts of the county (zip code) and Democrats more walkable precincts of the county (zip code).

The magnitude of the partisan effects is such that a registered Republican voter is expected to move to a new location with Walk Score about two points less than an independent from a comparably walkable origin location; a registered Democratic voter is expected to move to a location with 1 point higher Walk Score. Including individual- and tract-level covariates reduces

<sup>8</sup>Standard deviations of log density for Florida as a whole, within county, and within zip code are 1.56, 1.14, and 0.85, respectively.

the magnitude of these effects by about half, although all remain significant at the 99 percent level. Including fixed effects for destination counties or zip codes reduces the magnitude of the effects (again, mechanically) but leaves the direction and significance level unchanged.

Again, the partisan difference is consistent and robust to the inclusion of numerous control variables, but its magnitude is small relative to the overall variation present in the data: about one-eighth of a standard deviation. The Republican-to-Democrat difference is on the order of typical differences between houses in the same neighborhood, not the difference in average score between central-city and suburban locations.

### **Republican Presidential Vote Share**

The models we have shown thus far indicate that there are consistent, though fairly small, partisan differences in revealed preferences over certain housing characteristics. There are at least two distinct sorting mechanisms by which this partisan difference could emerge. In one, political tastes and housing tastes are correlated but have no causal relationship, perhaps because both are driven by the same unobserved personality factor. In the other, voters' desire to live near co-partisans generates the observed correlation: for historical reasons, large numbers of Democrats tend to live in denser parts of cities, and hence when Democrats try to live near each other they invariably choose to live in denser locations, preserving the historical pattern.

To distinguish these two mechanisms, we also present in Figure 1 models of a third dependent variable in the Florida mover sample: the 2008 Republican presidential vote share in the destination precinct. That is, we regress the 2008 Republican presidential vote share of the precinct in which the voter resided *in* 2010 on the 2008 Republican presidential vote share of the same voter's 2008 precinct, plus dummies for partisanship and additional conditioning variables and fixed effects. The model estimated here is:

$$s_{i,2010} = \alpha s_{i,2008} + \beta_D D_{i,2008} + \beta_R R_{i,2008} + \gamma_1 X_{i,2008}^1 + \gamma_2 X_{i,2010}^2 + \delta' Z_{i,2010} + \varepsilon_i \quad (2)$$

Here,  $s_{i,2008}$  is the Republican 2008 presidential vote share of the voter's precinct of residence in 2008, and  $s_{i,2010}$  is the Republican 2008 presidential vote share of the voter's precinct of residence in 2010. The only difference between this model and the specification outlined in Equation 1 is that the tract-level attributes are measured for the destination (2010) tract rather than the origin (2008) tract. We include these to make possible an evaluation of whether or not the partisan effect persists once we account for the physical and population characteristics of a location. In other words, we would like to know whether an unexpectedly Democratic precinct could still act as an attractor of Democrats even if it had none of the housing characteristics we typically expect Democrats to prefer.

A comparison of the regression coefficients in specifications with and without controls for non-political housing and neighborhood characteristics is informative in distinguishing the driving factor behind the observed partisan sorting patterns.

For example, imagine a precinct whose housing stock is comprised mostly of small apartments and which is densely populated with recent college graduates, but unlike most precincts fitting that description has a large population of registered Republicans. We are interested in whether or not such a precinct would attract a disproportionate share of Republican in-migrants. A comparison of the regression coefficients  $\beta_D$  and  $\beta_R$  using the raw versus residualized Republican presidential vote share—which is exactly what is achieved by including housing and population characteristics of the destination tract on the right-hand side of the equation—can distinguish between these two possibilities.

The results in the third row of Figure 1 show that there is some degree of homophily in moving decisions: registered Republicans tend to move to more Republican precincts, and Democrats to less, compared to independents moving from precincts with similar voting patterns. Comparisons of the coefficients on the Democratic dummy between the models with and



without tract-level control variables show that the apparent Democratic homophily in moving decisions is essentially eliminated by inclusion of controls. The coefficients shrink by an order of magnitude, and in the pooled regression lose significance, when tract demographic and housing controls are added, suggesting that Democrats preferentially move to less Republican precincts only to the extent that Republican vote share correlates with other, nonpolitical neighborhood attributes like walkability, housing stock, and so on.

The Republican dummy also shrinks toward zero when controls are added, but the difference is much less dramatic: the magnitudes decline by about half. Though the Democratic and Republican effect sizes are very similar in magnitude in the unconditional regression, when destination neighborhood controls are added the Republican dummy ranges from four to ten times as large, depending on the set of fixed effects included. For partisans of both types, nonpolitical neighborhood characteristics appear to be the primary factor driving location decisions. For registered Republicans, however, a high percentage of co-partisans remains an independently attractive feature of a neighborhood, regardless of that neighborhood's nonpolitical physical and demographic attributes.

As before, the effect sizes are fairly small but non-negligible. A registered Democrat is expected to move to a precinct that is about 2 percent less Republican than an independent voter moving out of a precinct with similar voting patterns. The corresponding effect for Republicans is slightly larger, at 2.6 percent, and in the opposite direction. As Republicans are moving to more Republican precincts and Democrats to more Democratic precincts, presumably making the Democratic precincts more Democratic and the Republican precincts more Republican in the next election cycle, the results shown in the last row of Figure 1 are consistent with increasing geographic polarization, and a concomitant decline in the competitiveness of sub-state-level elections, over time. In the next section, we introduce a simulation exercise to address the question of the rate at which precinct-level polarization is increasing and electoral competitiveness declining due to the observed sorting patterns.

### Simulating substantive effects

The results presented thus far are consistent with small but consistent partisan homophily in relocation decisions. In this section, we conduct a simulation analysis that allows us to translate the effects we measure on residential location choices into their consequences for electoral competitiveness, bias in legislative representation, and the geographic polarization of political preferences. We simulate a scenario where the observed rates of partisan sorting over a single election cycle (measured from our dataset of Florida voters in 2008–10) are repeated over several election cycles, and measure the consequences for these three substantive outcomes. We treat individuals voters' preferences as fixed and simulate only changes in their location, in order to isolate the effect of residential sorting alone on electoral outcomes.

The simulation consists of three steps, each based on regression estimates from the Florida dataset. In the first step, we estimate the probability that every voter registered in Florida in both 2008 and 2010 moves between 2008 and 2010, conditional on observed individual- and tract-level demographic information. We fit this step using a linear probability model of the decision to move or not, presented in Online Appendix Table A11, column (2). This model contains both the party registration dummies as well as the same set of individual- and tract-level covariates used in the regression models presented previously.<sup>9</sup> Using each individual registered voter's estimated probability of moving, we draw a simulated moving decision from a Bernoulli distribution. In each round of our simulation, approximately 12 percent of individuals decide to

<sup>9</sup>The party dummies have null effect in this model, unlike the other two models employed in the simulation. Other covariates such as individual age and urbanity of the individual's census tract are much more highly predictive of decisions to move.

move, consistent with the fraction of 2008 registered voters in Florida who moved from 2008 to 2010.

In the second step, every voter who decided to move in step one moves to a new precinct according to a variant of the model of Equation 2 with fixed effects for *origin* rather than destination geographies, parameter estimates of which are presented in Online Appendix Table A11, column (3). This model generates a predicted Republican presidential vote share of the moved-to precinct, along with a prediction variance from the regression error. We draw a new Republican share from a beta distribution with mean equal to the model prediction and variance equal to the prediction variance, and assign the voter to the precinct whose Republican presidential vote share is closest to this draw. Finally, we recompute Republican presidential vote and registration shares in every precinct once moving is complete. We treat individuals' presidential votes and registration choices as fixed; the only thing that may vary over time is their residential location. The state-wide population of voters and hence state-wide Republican shares are fixed throughout the course of the simulation. All that may potentially change is the composition of lower-level geographic units such as precincts, state house, or congressional districts.

For partisan registration choices, we simply use the observed party of registration in 2008 from the voter file. For presidential votes, we estimate a model of 2008 Republican presidential vote probability given demographics and party of registration from the Florida voter file. This model is:

$$s_{i,2008} = \beta_D D_{i,2008} + \beta_R R_{i,2008} + \gamma_1 X_{i,2008}^1 + \gamma_2 X_{i,2008}^2 + \delta' Z_{i,2008} + \varepsilon_i \quad (3)$$

where  $s_{i,2008}$  is the Republican presidential vote share of voter  $i$ 's precinct of residence in 2008. Estimates of this model are presented in Online Appendix Table A.11, column (1). Using the predicted probabilities of Republican presidential vote from this model, we draw a simulated Republican presidential vote for each individual from the corresponding Bernoulli distribution. Again, these simulated votes are computed only once at the beginning (in 2008) and remain fixed throughout the course of the remaining cycles.

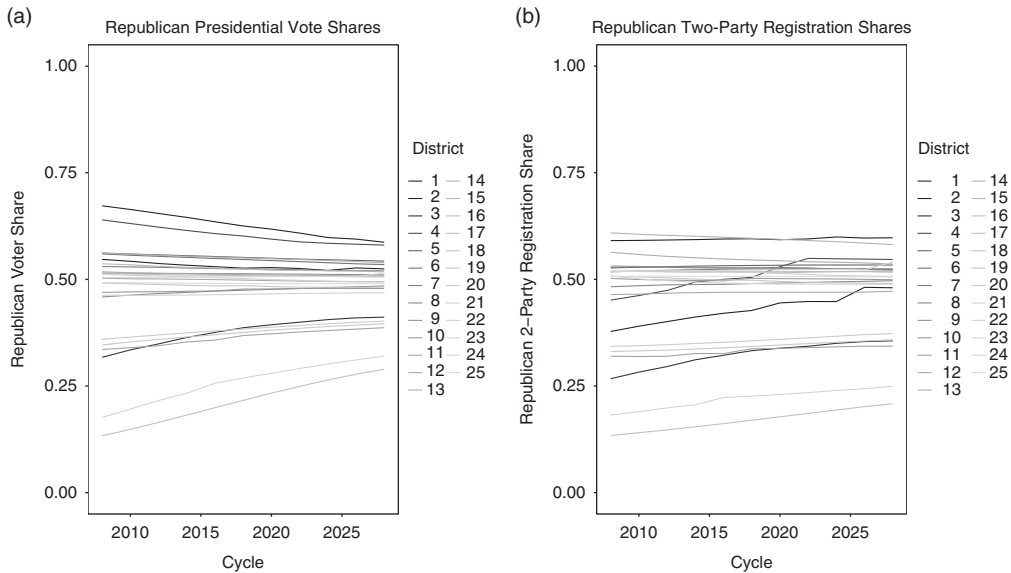
We repeat this process for ten iterations, at each step recording the new precincts of residence for every individual. Given the new precinct assignments, we can compute simulated quantities such as Republican vote or registration shares by congressional district and examine the trends resulting from repeated rounds of residential sorting.

### Competitiveness

We first examine the effects of sorting on the competitiveness of Congressional elections. In Figure 2, we plot two measures of the partisan composition of each of Florida's 25 Congressional districts over the course of the simulation: the fraction of voters in each district with simulated Republican presidential votes in 2008, and the fraction of Republican registrants in each district, among the district's registrants who registered as one of the two major parties.

Neither plot shows evidence of declining numbers of competitive districts. In fact, the pattern is just the opposite, displaying a clear trend of convergence towards partisan balance. The most heavily Democratic districts become less Democratic, and similarly the most heavily Republican districts become less Republican. Competitive districts remain competitive throughout.

This result indicates that, although Figure 1 demonstrated the existence of a pattern of selection into politically compatible neighborhoods, the magnitude of this politically motivated sorting is too small to sustain the 2008 level of (un-)competitiveness. Sorting on political attributes is swamped by sorting on other nonpolitical neighborhood characteristics as well as idiosyncratic individual reasons, leading to a pattern of partisan mixing and increasing homogeneity across districts. In the absence of other factors such as gerrymandering, incumbency advantages, changes in candidate recruitment, and so on, residential sorting alone cannot produce a decline in competitiveness, relative to the level prevailing in 2008.



**Fig. 2.** The result of ten cycles of simulated moving among registered voters in Florida on the partisan composition of Florida's 25 Congressional districts. The left panel (a) shows the fraction of voters in each district with simulated Republican presidential votes in 2008. The right panel (b) shows the fraction of Republican registrants in each district, among the district's registrants who registered as one of the two major parties.

### Malapportionment

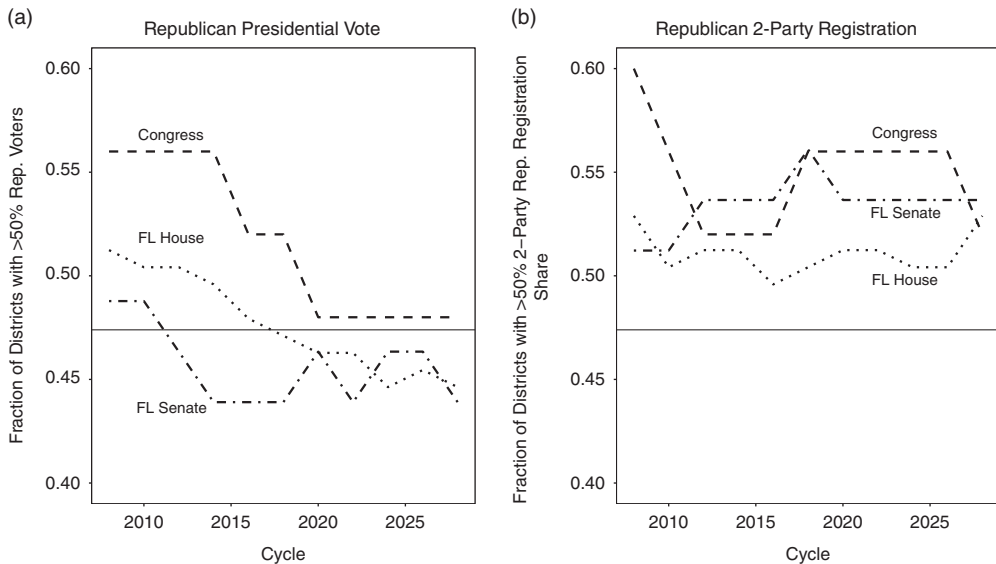
We next examine the effects of sorting on the correlation of partisanship with residential density, and the resulting malapportionment of Congressional and state legislative seats relative to the partisan shares in the state-wide population. Figure A4, in the Online Appendix, plots the correlation of tract-level log population density with Republican presidential votes, and with Republican and Democratic partisan registration. There is, as expected given the results of the sorting model of Table A5, Online Appendix, a weak trend of movement of registered Democrats (Republicans) into more (less) dense areas. However, partisan registration is not predictive enough of presidential vote choices<sup>10</sup> for this trend to be apparent in the correlation of Republican vote with population density. The latter quantity actually moves towards zero, though not dramatically, over the ten cycle course of the simulation. Again, while there are measurable partisan influences on choice of neighborhood density, these forces are smaller than the effects of non-partisan observables and individual unobservables. As a result, sorting on its own cannot sustain the observed 2008 correlation between vote choice and density.

Perhaps not surprisingly, then, the effects of residential mobility on the apportionment of legislative seats are benign. Figure 3 shows the fractions of legislative seats (in Congress, the Florida state senate, and the Florida house of representatives) with a majority of registered voters who either were simulated to vote Republican in the 2008 presidential election or were registered Republicans in 2008. The dashed lines in the figures show the corresponding state-wide population fractions, e.g., the fraction of 2008 Republican presidential voters in Florida as a whole. A consequence of Democrats' propensity to live in denser areas is that the initial apportionment has a majority of majority-Republican districts in Congress and the Florida state senate, despite the fact that Democrats make up a majority of the state-wide population.<sup>11</sup>

Over time, however, this divergence between legislative seats and population shares *declines* as a result of residential sorting. Both presidential vote and registration share measures show

<sup>10</sup>See Table A11 in the Online Appendix for estimates connecting the two quantities.

<sup>11</sup>This is the "unintentional gerrymandering" effect described by Chen and Rodden (2013).



**Fig. 3.** The result of ten cycles of simulated moving among registered voters in Florida on the apportionment of legislative seats in Florida. The left panel (a) shows the fraction of districts in each chamber with simulated Republican presidential vote shares greater than 50 percent. The right panel (b) shows the fraction of districts in each chamber with Republican two-party registration shares greater than 50 percent.

movement towards the state-wide benchmark, though the decline is larger in vote-share terms. Once again, sorting on its own cannot sustain the existing malapportionment of legislative seats, and in fact in isolation would reduce the level of such malapportionment due to correlation of partisanship with density.

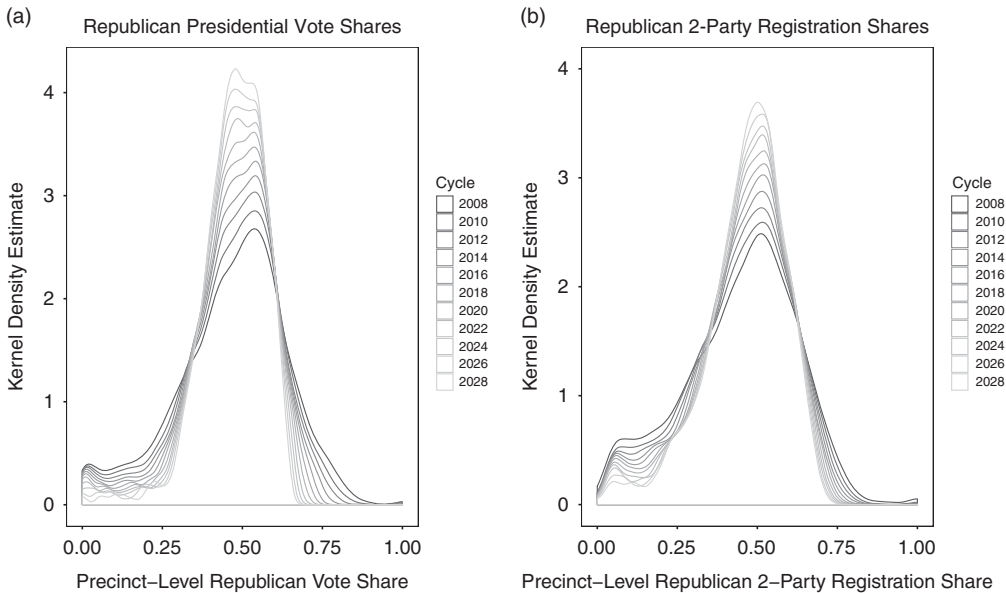
### Local Partisan Homogeneity

Finally, we examine the effects of residential sorting on local partisan homogeneity. We plot in Figure 4 the distribution of precinct-level Republican vote and registration shares. If residential sorting was a polarizing force, this exercise should produce an increasingly bimodal distribution over time, as somewhat Republican precincts become heavily Republican, and vice versa on the Democratic side. The actual pattern is the opposite: the mass of approximately evenly split precincts grows steadily, and the mass of extreme precincts falls.

A related measure, shown in Figure A5 in the Online Appendix, is the fraction of partisans living in precincts where the great majority of their neighbors vote or register for the same party. Sociologists and education scholars use an analog of this measure to study neighborhood and school segregation on racial or class lines. Unsurprisingly given the results in Figure 4, this measure of local partisan segregation steadily declines over repeated cycles of the simulation.

In sum, Floridians' residential mobility appears to be, on net, a force for greater electoral competitiveness, apportionment of legislative seats in line with population shares, and local integration across party lines. This overall pattern emerges *in spite of* the fact that partisans are more likely to select politically similar neighborhoods. The reason for this seeming disconnect is that the base rate of geographic polarization in 2008 is much higher than can be generated by the small partisan influence on moving patterns that we measure. The equilibrium level of geographic polarization produced by residential sorting alone is much lower than the level we actually observe.

In fact, a calibration exercise shows that the partisan bias in moving decisions would need to be on the order of five times larger than the value we estimate in order to be able to sustain the



**Fig. 4.** The result of ten cycles of simulated moving among registered voters in Florida on the precinct-level distributions of votes and registrations. The left panel (a) shows the distribution of simulated 2008 Republican presidential vote share at the precinct level in each of the ten simulation cycles. The right panel (b) shows the distribution of two-party Republican registration share at the precinct level in each of the ten simulation cycles.

2008 level as a steady state. Figure A6, in the Online Appendix, shows what happens to the distribution when we successively increase the partisan coefficients above their estimated magnitudes and repeat the same exercise as that in Figure 4(b). Even when we double or quadruple the magnitude of the partisan bias terms, we continue to see mass shifting towards mixed levels and away from the extremes. It is not until we reach six times the actual value that we begin to see appreciable increases in the polarization of the precinct-level distribution.

### Changes in partisanship

The results of the previous section show that partisan geographic sorting, in isolation, has a homogenizing—rather than polarizing—effect on the distribution of political preferences. What, then, explains the apparently stable degree of polarization in the current observed distribution of preferences? Why do we not see reversion to the mean among outlier precincts over time, as occurs in our simulation, as new residents move in?

One possibility is that individuals' political preferences are influenced by their social and physical environment. When a rural voter moves to an urban location or vice versa, his or her political preferences may change to fall in line with those of his or her new peers and neighbors, just as other cultural and lifestyle preferences are likely to change following such a move. Our simulation exercise explicitly rules this possibility out: individuals' political preferences are assumed to be completely fixed. Our simulation results show that, under this assumption of fixed preferences, residential mobility tends to produce politically homogeneous rather than segregated geographic distributions. The fact that we do not observe such a homogeneous pattern, then, suggests that the reverse causal channel (location changing preferences, rather than preferences changing location) has some bite.

In order to examine this possibility, we examine changes in individual voters' partisan registration over time. Specifically, we ask if voters who moved to new locations that were very different than their old locations on politically salient dimensions subsequently changed their

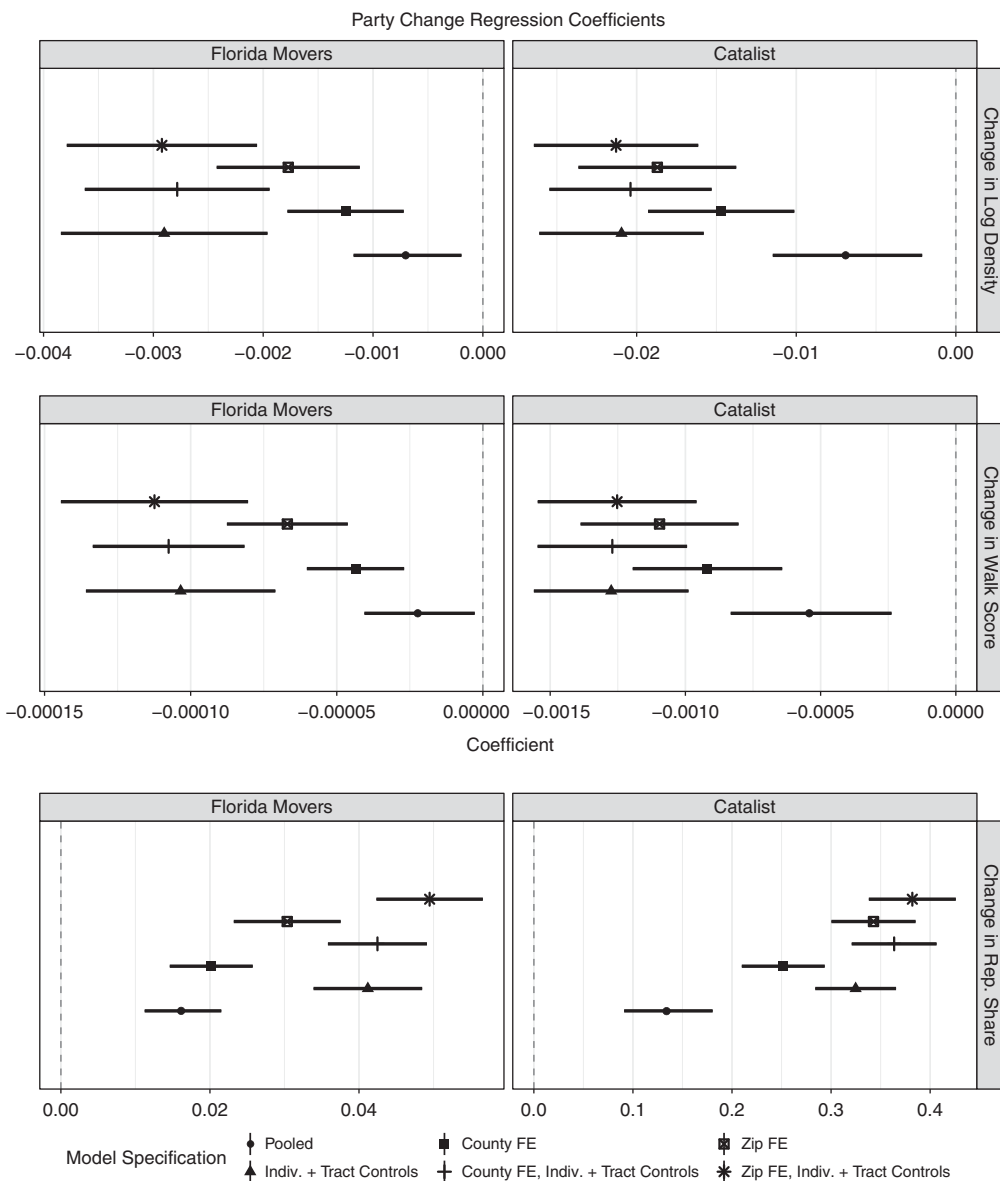


Fig. 5. Coefficient estimates of three independent variables for the change in the party of registration DV.

registration status in a direction matching the change in their environment. The model we estimate is:

$$\Delta p_i = \alpha \Delta c_i + \beta_D D_{i,pre} + \beta_R R_{i,pre} + \gamma_1 X_{i,pre}^1 + \gamma_2 X_{i,pre}^2 + \delta' Z_{i,pre} + \varepsilon_i \tag{4}$$

where the dependent variable  $\Delta p_i$  is a trichotomous measure of change in party registration ranging from  $-1$  to  $1$ . This variable is coded as a zero if there was no change in an individual's party registration after moving from one location to another; if an individual moves in a more liberal direction (i.e., she moves from Republican to Independent or if she moves from Republican *or* Independent to Democrat) the variable is coded as  $-1$ ; finally, if an individual moves in a more conservative direction (i.e., she moves from Democrat to Independent or if she

moves from Democrat *or* Independent to Republican) the variable is coded as 1. The main independent variable of interest  $\Delta c_i$  is the pre- to post-move change in the voter's precinct characteristics of interest:  $\Delta c_i$  may be the difference in either the log population density, walk score, or 2008 Republican presidential vote share between voter  $i$ 's pre- and post-move precincts of residence. We also include dummies for individuals' initial (pre-move) partisan affiliations to control for the fact that the value of  $\Delta p_i$  must be non-negative for initially Democratic registrants and non-positive for initially Republican registrants.

The results are presented in Figure 5 for both the Florida and Catalist datasets. Results are also in tabular form in the Online Appendix Tables A12–A14 for the Florida data and A.15–A.17 for the Catalist sample, respectively. Changes in all three independent variables are statistically significantly associated with changes in registration status in both datasets, in the direction that would be predicted if voters' preferences updated to match their new surroundings. For instance, moving to a census tract with 1 SD higher Walk Score is associated with an approximately 0.2 percentage point increase in the differential likelihood of a former independent registering as a Democrat versus registering as a Republican. A standard deviation increase in log density produces about a 0.4 percentage point increase in this same differential. Given that less than 8 percent of Florida movers changed their partisan registration status in any direction between 2008 and 2010, these are fairly substantial shifts.

The results for the Catalist sample show the identical pattern of directionality and statistical significance, but with substantially larger magnitudes. This is a result of the much higher baseline rate of party of registration change in the Catalist data: 31.4 percent of voters in the Catalist sample changed their partisan registration status after their move, compared to 7.6 percent of within-Florida movers.<sup>12</sup> Florida allows voters who move within the state to update only their address without re-filling the entire registration application, whereas cross-state movers must fill out a new registration form and actively choose a party affiliation. This lack of a default, no-change option for cross-state movers likely explains the much higher rate of party changes (and larger coefficients on the change in location attribute regressors) in the Catalist data.

Of course, the results in Figure 5 are not definitive evidence of social influence; it is possible that voters who change their party affiliation upon moving are already (unobservably, given that we have data only on party registration) less strongly attached to their pre-move party than the average voter. The necessity of re-registering after a move may provide these voters an opportunity to bring their formal party affiliation more closely in line with their voting preferences. However, note that magnitudes of the coefficients on pre- to post-move location change all *increase* substantially when additional individual- and tract-level controls are added. This pattern implies that such political-preference unobservables would have to be *negatively* correlated with the observable factors in our control set—such as the racial, income, and housing makeup of the individual's origin location—in order for these results to be purely driven by omitted variable bias.

Finally, we introduce the party change estimates into our simulation model to understand the effect of post-move party change on geographic polarization. Results, presented graphically in Figure A7 in the Online Appendix, show that allowing movers to change parties according to the pattern we observe in the data mitigates the trend towards de-polarization we observe in the baseline simulation. However, this reduction is not enough to keep the distribution stable at its 2008 level. We conclude that some changes in party ID in the direction of the local majority among *non*-movers are necessary to preserve the current level of geographic polarization.

## Discussion

The American electorate is geographically divided. Partisan affiliations correlate with tastes for housing characteristics, with Democrats clustered in urban cores and Republicans spread out

<sup>12</sup>See Online Appendix Tables A1 and A3 for a breakdown of the patterns of party affiliation change in both datasets.

across suburban and rural areas. This pattern has some negative consequences for the functioning of elections as both preference aggregation and accountability mechanisms: urban Democrats' delegations in Congress and state legislatures are undersized relative to their share of the population, and they are represented by incumbents who face weak challengers, or no challengers at all, from the opposing party.

There are several plausible theories for the emergence and persistence of this pattern. In one, the correlation between urban-ness and preference for Democrats is a historical artifact<sup>13</sup> that has been sustained over time by individuals' desire to live near others who share their political views. In a second, preferences for housing attributes correlate with political preferences, and the observed political pattern is a side effect of selection on nonpolitical housing attributes. Finally, a third theory reverses the direction of causation, positing that political preferences are influenced by an individual's physical and social environment.

A critical distinction between the first two theories and the third is that the first two require the existence of a partisan bias in moving decisions: when partisans move from one location to another, they must preferentially select into areas containing more of their own co-partisans and/or more of the attributes associated with their party ID. Given that Americans as a whole are quite mobile—in our Florida sample, more than 10 percent of voters registered in both 2008 and 2010 moved to a different precinct in that two-year window—this partisan bias must be fairly strong in order to preserve the existing geographic pattern and prevent it from disappearing as the population shuffles around.

This paper directly measures the relative strength of this partisan bias in moving decisions. We find strong evidence for the existence of partisan sorting in two independent, large-scale datasets. Registered Democrats (Republicans) sort into areas that are more (less) dense, walkable, and Democratic-leaning in presidential elections. We find that the bias towards selecting politically congruent areas largely disappears once attributes of the destination's housing stock and population characteristics are accounted for. This distinction implies that partisans simply have correlated tastes for residence locations and are not selecting politically congruent neighborhoods per se. This finding is in line with previous work examining the link between political preferences and residential location (Brown 1988; Tam Cho, Gimpel, and Hui 2013; Gimpel and Hui 2017; Mummolo and Nall 2017).

However, the large sample sizes we have available allow us to precisely measure the magnitude and not just the direction of these effects. These magnitudes are, across the board, quite small. In other words, although partisans' tastes for politically salient attributes are correlated, they are not willing to act on those tastes if it means sacrificing proximity to jobs, school quality, housing affordability, or the myriad other idiosyncratic and non-partisan factors that influence voters' residential choices.

Our simulation analysis shows that this nonzero but small partisan bias is not sufficient to sustain a geographically polarized distribution of preferences. Because Americans move frequently, and the vast majority of the variation in moving decisions is non-partisan in nature, residential mobility on its own fairly quickly brings areas with concentrations of voters of one party or the other back towards the population mean. As a result, in spite of the existence of partisan bias in moving decisions, the net effect of residential sorting alone is to improve the competitiveness of typical Congressional and state legislative elections, to reduce the malapportionment of legislative seats due to the concentration of Democrats in high-density areas, and to reduce the fraction of voters living in very politically homogeneous neighborhoods.

If residential sorting alone would quickly generate a geographically homogeneous distribution of political preferences, then the fact that a very non-homogeneous pattern persists suggests that something else must be going on. Namely, the preferences of in-migrants must be adapting to match the modal preferences of their new neighborhoods. We show that this is in fact occurring,

<sup>13</sup>Due to, for example, the relative strength of the labor movement in urban areas during the industrial revolution (Rodden 2010) or the racial politics of post-war urban renewal and interstate highway construction (Nall 2015).



as voters who move to neighborhoods that are different from their previous residence on politically salient dimensions are much more likely to change their party affiliation to match that of their new neighbors.

Nonetheless, much work remains to be done to understand the mechanism by which this preference change occurs. Is there social pressure, or a desire to not stand out, that leads voters to adopt the affiliation of their new neighbors, whatever that may be? Or is there, as Lipset and Rokkan (1967) suggested, something inherent in urban life that encourages the adoption of a liberal ideology? We leave these questions for future work.

**Supplementary material.** To view supplementary material for this article, please visit <https://doi.org/10.1017/psrm.2018.44>

**Acknowledgments.** We thank Adam Glynn and Tom Clark for helpful comments and suggestions, Joo Eun Hwang for excellent research assistance, and Rob O'Reilly and the Emory Libraries for assistance in data collection.

## References

- Ansolabehere S, Palmer M and Lee A** (2014) *Precinct-Level Election Data*. Harvard Dataverse: <http://hdl.handle.net/1902.1/21919>
- Ansolabehere S and Rodden J** (2011) *Florida Data Files*. Harvard Dataverse: <http://hdl.handle.net/1902.1/16797>.
- Bishop B** (2008) *The Big Sort: Why the Clustering of Like-Minded America is Tearing Us Apart*. Boston: Houghton Mifflin Harcourt.
- Brown TA** (1988) *Migration and Politics: The Impact of Population Mobility on American Voting Behavior*. Chapel Hill: University of North Carolina Press.
- Chen J and Rodden J** (2013) Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures. *Quarterly Journal of Political Science* **8**(3), 239–69.
- Ferejohn J** (1986) Incumbent Performance and Electoral Control. *Public Choice* **50**, 5–25.
- Gimpel JG and Hui IS** (2015) Seeking Politically Compatible Neighbors? The Role of Neighborhood Partisan Composition in Residential Sorting. *Political Geography* **48**, 130–42.
- Gimpel JG and Hui I** (2017) Inadvertent and Intentional Partisan Residential Sorting. *The Annals of Regional Science* **58**(3), 441–68.
- Gordon SC, Huber GA and Landa D** (2007) Challenger Entry and Voter Learning. *American Political Science Review* **101** (2), 303–20.
- Hitsch GJ, Hortacsu A and Ariely D** (2010) Matching and Sorting in Online Dating. *American Economic Review* **34**, 130–63.
- Huber GA and Malhotra N** (2017) Political Homophily in Social Relationships: Evidence from Online Dating Behavior. *The Journal of Politics* **79**(1), 269–83.
- Lang C and Pearson-Merkowitz S** (2014) Partisan Sorting in the United States, 1972–2012: New Evidence from a Dynamic Analysis. *Political Geography* **48**, 119–29.
- Levendusky M** (2009) *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*. Chicago: University of Chicago Press.
- Lipset SM and Rokkan S** (1967) *Party Systems and Voter Alignments: Cross-National Perspectives*. New York: Free Press.
- Minnesota Population Center** (2011) *National Historical Geographic Information System. Technical Report Version 2.0*. Minneapolis, MN: University of Minnesota.
- Mummolo J and Nall C** (2017) Why Partisans Do Not Sort: The Constraints on Political Segregation. *The Journal of Politics* **79**(1), 45–59.
- Nall C** (2015) The Political Consequences of Spatial Policies: How Interstate Highways Facilitated Geographic Polarization. *Journal of Politics* **77**(2), 394–406.
- Obama B** (2004) *Transcript: Illinois Senate Candidate Barack Obama*. Available at <http://www.washingtonpost.com/wp-dyn/articles/A19751-2004jul27.html>, accessed 21 January 2016.
- Prior M** (2007) *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*. New York: Cambridge University Press.
- Rodden J** (2010) The Geographic Distribution of Political Preferences. *Annual Review of Political Science* **13**, 321–340.
- Tam Cho WK, Gimpel JG and Hui IS** (2013) Voter Migration and the Geographic Sorting of the American Electorate. *Annals of the Association of American Geographers* **103**(4), 856–70.