Annals of the American Association of Geographers

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/raag21

Scale, Context, and Heterogeneity: A Spatial Analytical Perspective on the 2016 U.S. Presidential Election

A. Stewart Fotheringham , Ziqi Li & Levi John Wolf

Scale, Context, and Heterogeneity: A Spatial Analytical Perspective on the 2016 U.S. Presidential Election

A. Stewart Fotheringham,[*] Ziqi Li,[†] and Levi John Wolf[‡]

[*]*Spatial Analysis Research Center, School of Geographical Sciences and Urban Planning, Arizona State University*
[†]*Spatial Analysis Research Center, School of Geographical Sciences and Urban Planning, Arizona State University, and Department of Geography and Geographic Information Science, University of Illinois*
[‡]*School of Geographical Sciences, University of Bristol*

This article attempts to identify and separate the role of spatial "context" in shaping voter preferences from the role of other socioeconomic determinants. It does this by calibrating a multiscale geographically weighted regression (MGWR) model of county-level data on percentages voting for the Democratic Party in the 2016 U.S. presidential election. This model yields information on both the spatially heterogeneous nature of the determinants of voter preferences and the geographical scale over which the effects of these determinants are relatively stable. The article, perhaps for the first time, is able to quantify the relative effects of context versus other effects on voter preferences and is able to demonstrate what would have happened in the 2016 election in two scenarios: (1) if context were irrelevant and (2) if every county had exactly the same population composition. In addition, the article sheds light on the nature of the determinants of voter choice in the 2016 U.S. presidential election and presents strong evidence that these determinants have spatially varying impacts on voter preferences. *Key Words: context, local modeling, multiscale geographically weighted regression, voter preferences scale.*

Identifying the drivers of voter preferences for one party over another has clear and important implications not just for understanding the outcomes of electoral processes but also for being able to influence the outcomes of future elections by targeting key groups of voters. Consequently, a great amount of research has been undertaken to try to identify the factors that play a role in influencing how voters vote (Sigelman and Sigelman 1982; Powell 1986; Hillygus and Jackman 2003; Leigh 2005; Mutz 2018; Schaffner, MacWilliams, and Nteta 2018). Two contrasting methods aimed at identifying the determinants of voter preferences dominate the literature. Both have advantages and disadvantages.

The first uses detailed surveys on a limited number of individuals and elicits information on how people voted or intend to vote along with a set of individual-level characteristics (Branton 2003; Hillygus and Shields 2005; Guth et al. 2006; Payne et al. 2010; Mutz 2018). Relationships are then established between voter preferences and the individual-level characteristics. This is a powerful methodology in that the relationships identified are at the individual level. It is limited, however, in terms of the number of people that can be surveyed, so sample sizes are small. This limitation is a particular problem if the determinants of voter preference vary geographically—the limited sample size typically does not allow any spatially disaggregated analysis to be undertaken.

The second uses large-scale voter preferences recorded typically as percentage share of the vote for one party within some aggregated spatial unit such as electoral district or county (Kim, Elliott, and Wang 2003; Levernier and Barilla 2006; Scala, Johnson, and Rogers 2015; Miller and Grubesic 2020). Such data have the disadvantage of being prone to the ecological fallacy whereby the results might not reflect the behavior of the individuals within the aggregated spatial unit. Because the data tend to be more comprehensive, however, often representing all voters, they allow for more detailed spatially disaggregated analysis. Given that the focus of this study is on identifying the relative role of spatial context in determining voter preference for one party over another using spatially disaggregated models of voter choice, we employ aggregated data on voter preferences.

Annals of the American Association of Geographers, 0(0) 2020, pp. 1–20 © 2020 by American Association of Geographers
Initial submission, January 2020; revised submissions, May and July 2020; final acceptance, July 2020
This article has been corrected with minor changes. These changes do not impact the academic content of the article.
Published by Taylor & Francis, LLC.

The article proceeds as follows. A traditional global model of voter preference is constructed based on county-level data on votes for the Democratic Party in the 2016 U.S. presidential election. This model is then calibrated by multiscale geographically weighted regression (MGWR; Fotheringham, Yang, and Kang 2017), which allows potential spatial nonstationarity in the determinants of voting for the Democratic Party to be identified. This model also generates information on the spatial scales over which the relationships between voter preference and various socioeconomic attributes vary. In addition, the local intercept estimated in the calibration of an MGWR model allows the identification and quantification of contextual effects in the determination of voter preference. Finally, once contextual effects have been identified and measured, various "what if?" scenarios can be examined. Here, we are able to answer two intriguing questions related to the 2016 U.S. presidential election:

1. What would have happened if every county in the United States had exactly the same population composition?
2. What would have happened if spatial context had played no role in influencing voter preference?

We now consider the role of spatial context in human behavior in general and then follow that by considering its specific role in influencing voting behavior. We then describe a model of voter preference for the Democratic Party in the 2016 U.S. presidential election and report the local varying results of calibrating this model by MGWR. In this we focus on the role of the local intercept as a measure of the influence of context and answer the two questions just raised. To our knowledge, this represents the first time that the role of context has been measured, and we also report on the spatial scales at which relationships affecting voter preference vary across the United States.

Spatial Context and Human Behavior

We use the phrase *spatial context* (henceforth simply *context*) to describe the impact of a person's location in space on how they behave in space. In essence it is a shorthand term for the impact largely unmeasurable effects of one's location might have on one's preferences and actions, as shown in Figure 1. In some cases, there might be identifiable and

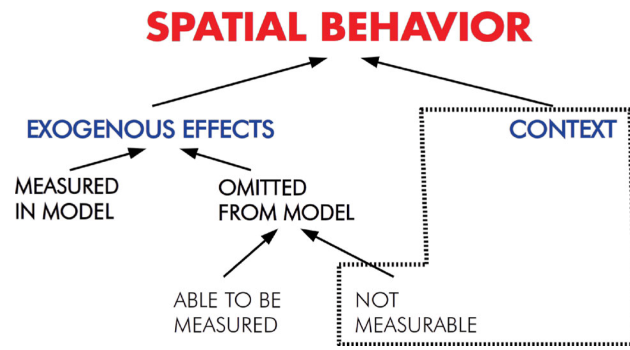

Figure 1. The role of context in determining spatial behavior.

measurable aspects of context we could include in a model as explicit covariates, but often this is not the case and then context is a catch-all term for multiple effects on behavior caused by locale or place.

Examples of what might cause contextual variations in spatial behavior include the following:

1. Traditions, customs, lifestyles, and daily practices common to an area that affect social norms, which in turn affect individual behavior (Golledge 1997).
2. Ecological influences—features of the physical environment that might affect people's thoughts, feelings, and behavior (Gomez, Hansford, and Krause 2007; van de Vlivert 2008; Altman and Wohlwill 2012).
3. Economic conditions such as persistently high levels of unemployment, which can affect thoughts and feelings (Darity and Goldsmith 1996; Clark 2003; Ayllón 2013).
4. Selective news representation by local media (McCombs et al. 1997; Garz 2018).
5. The influence of friends, family, and local "opinion leaders," often referred to as "social imitation" (Braha and de Aguiar 2017).
6. Differences in personality types across space (Krug and Kulhavy 1973; Rentfrew et al. 2013; Rentfrow, Jokela, and Lamb 2015).

Whatever its cause, there are many studies that recognize the impacts of location on various types of human behavior (inter alia Escobar 2001; Plaut, Markus, and Lachman 2002; Enos 2017; Coffee et al. 2020). As Enos (2017) stated, "Context—or, more precisely, social geography—can directly affect our behavior and is therefore tremendously important" (78).

As convinced as some researchers are that context plays an important role in human behavior, to date there has been no research that has managed to separate contextual effects from other determinants of behavior and to quantify the relative contributions of each. As Enos (2017) stated, "Nobody doubts that context can affect behavior and careful studies of 'neighborhood effects' have strongly suggested it can. However, the exact nature of contextual effects—how much they really matter—is elusive to researchers" (120). We now turn to consider the role of context in voting behavior.

Spatial Context and Voting Behavior

The role of spatial context in political analysis has long been recognized as a significant factor in the politics of the United States (Agnew 1996; Escobar 2001), originating perhaps with the seminal study of Key and Heard ([1949] 1984), although see King (1996) for a dissenting view. Traditionally, political science literature assumes that spatial context operates within some predefined social–spatial frame of analysis. For instance, one long-standing literature on contextual effects in political science arises in "American political development" through the analysis of sectionalism and sectional politics (Bensel 1987; Archer 1988). Recent studies of U.S. voting behavior have now recognized that context can be very influential, resulting in effects (in aggregate or at the individual level) that are not the same at all locations (Gelman et al. 2007). Complicating matters, some drivers of voting behavior are recognized to be driven by a spatial–social context where both demographic and geographic factors are in play at the same time (Pantoja and Segura 2003; Rocha and Espino 2009; Hersh and Nall 2016).

Although this recognition of the role of context in influencing political attitudes represents progress, many of these studies rely on the assumption that the geography of this context is known and exogenous. Such an assumption is both highly questionable and very restrictive. Recent work in political science has demonstrated that not only are spatial dimensions important in electoral analysis (Cho and Gimpel 2012) but contextual effects must be treated as unknown and their spatial extent and distribution be estimated for data rather than assumed a priori (Calvo and Escolar 2003; Darmofal 2008; Cho and Gimpel 2010; Clemens, Crespin, and Finocchiaro 2015; Manley and Demsar 2015). This exploratory approach to understanding the spatial structure of the electorate suggests that context can emerge from analyses of political processes even when not explicitly modeled.

Even given the increased recognition of the role of context in affecting voting behavior, however, major problems remain in terms of both separating the effect of context from other effects and quantifying the relative contribution of context. As O'Loughlin (2018) stated in his summary of thirty-five years as an editor of *Political Geography*, "If context has remained a mantra in political geography, how do we measure its importance?" (148).

In this article, we argue that MGWR is capable of measuring the importance of context across spatial processes that affect election outcomes. We now describe a set of results from calibrating a traditional global model of voter preference followed by the results of calibrating the same model by MGWR, which is used to both separate the contextual effect from other socioeconomic determinants of voter preference and quantify its relative contribution to voting behavior.

A Global Model of Voter Preference

Data Used

We use the 2016 presidential election data from the MIT Election Lab for our analysis of electoral context.[1] Any third-party votes have been ignored so that our dependent variable is defined as (the number of votes cast for the Democratic Party/Number of votes cast for either the Democratic or Republican Party) $\times$ 100. Further, we use the 2012–2016 five-year American Community Survey estimates at the county level for the exogenous factors relevant to the election in 2016. The independent variables used in this study are described in Table 1.[2] Data from the noncontiguous states of Hawaii and Alaska have been removed, as have data from counties with fewer than 5,000 population to reduce small-number bias in percentages. This left a total of 2,813 counties.

Figure 2 describes the distribution of the dependent variable. Two things are immediately evident. One is that there is a tremendous spatial variation in support for the Democratic Party with the percentages varying from 3 percent to 96 percent. The second is that the spatial clustering of similar values strongly suggests that this pattern of voting has not resulted from a purely random process—there must be other factors influencing people's voting behavior. The question is what these factors are and whether their effect on voting behavior is constant over space. A further question is this: Is context or

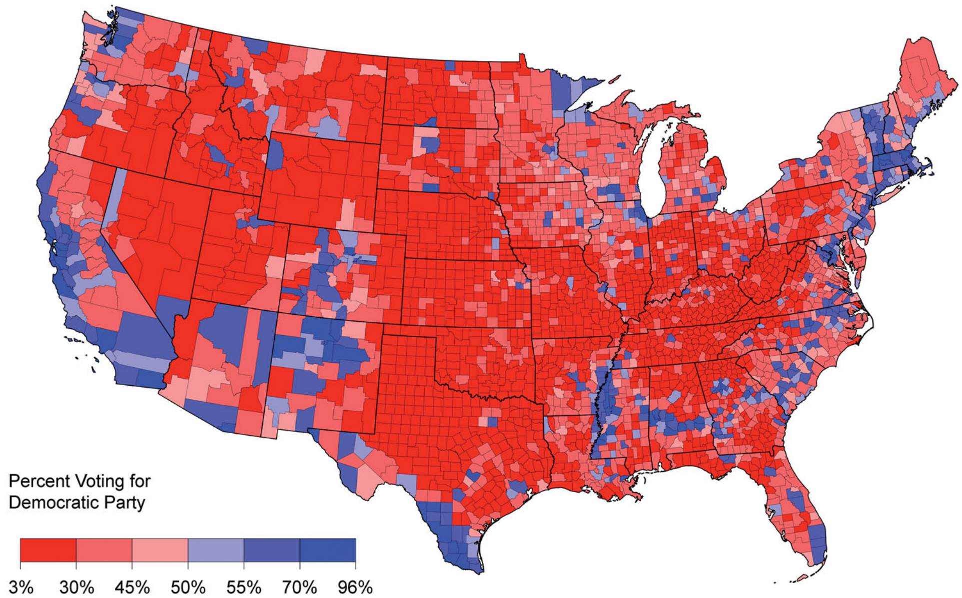

Figure 2. County-level percentage vote for the Democratic Party.

"place" one of these factors and, if so, how important is it?

To examine which attributes of each county most strongly associated with the vote for the Democratic Party, we first calibrated a traditional linear model of the following form using data on fourteen covariates as shown in Table 1:

$$y = \alpha + \sum_j \beta_j x_j + \varepsilon. \quad (1)$$

In this model, epsilon (ε) is the classic independent and identically distributed error term. The selection of the covariates was strongly influenced by discussion of the 2016 election in both the academic literature and the media, but the choice was also guided by statistical testing for multicollinearity using variance inflation factors and for heteroscedasticity.[3] All fourteen covariates and the dependent variable were standardized to have a mean of 0 and variance of 1. The results of calibrating this global model are shown in Table 1.

The model performs reasonably well in replicating 66 percent of the variance in percentage Democratic vote and many variables are significant at the 1 percent significance level. Because the variables are unit normalized, the estimate of the intercept is zero and the absolute magnitude of the remaining parameter estimates can be used as an indicator of effect strength. The results suggest that five of the covariates have no significant influence on voting behavior. These are gender ratio, percentage of young voters, percentage of older voters, income disparity, and percentage employed in manufacturing. Only one covariate associates significantly negatively with the size of the Democratic vote, and that is income: As anticipated, places with higher incomes have a greater tendency to vote for the Republican Party. The remaining eight covariates all associate highly significantly positively with voting for the Democratic Party. In order of effect, the vote for the Democratic Party within a county increases with increasing proportions of black population, people with at least a bachelor's degree, Hispanic

Table 1. Results of global ordinary least squares model

Variables	Estimate	*SE*	*t* Value	Pr(>|*t*|)
(Intercept)	0.000	0.000	0.000	1.000
Sex_Ratio	0.010	0.010	0.808	0.419
Pct_Age_18_29	–0.021	–0.021	–1.094	0.274
Pct_Age_65	–0.014	–0.014	–0.737	0.461
Pct_Black	0.528*	0.528	35.974	0.000
Pct_Hispanic	0.283*	0.283	15.005	0.000
Median_Income	–0.310*	–0.310	–12.333	0.000
Pct_Bachelor	0.425*	0.425	16.530	0.000
Gini	0.027	0.027	1.669	0.095
Pct_Manuf	0.019	0.019	1.426	0.154
Ln(Pop_Den)	0.166*	0.166	9.811	0.000
Pct_3rd_party	0.164*	0.164	11.376	0.000
Turnout	0.168*	0.168	8.853	0.000
Pct_FB	0.192*	0.192	9.322	0.000
Pct_Insured	0.174*	0.174	10.692	0.000
Model R^2	0.66			

*Significant at 95 percent level.

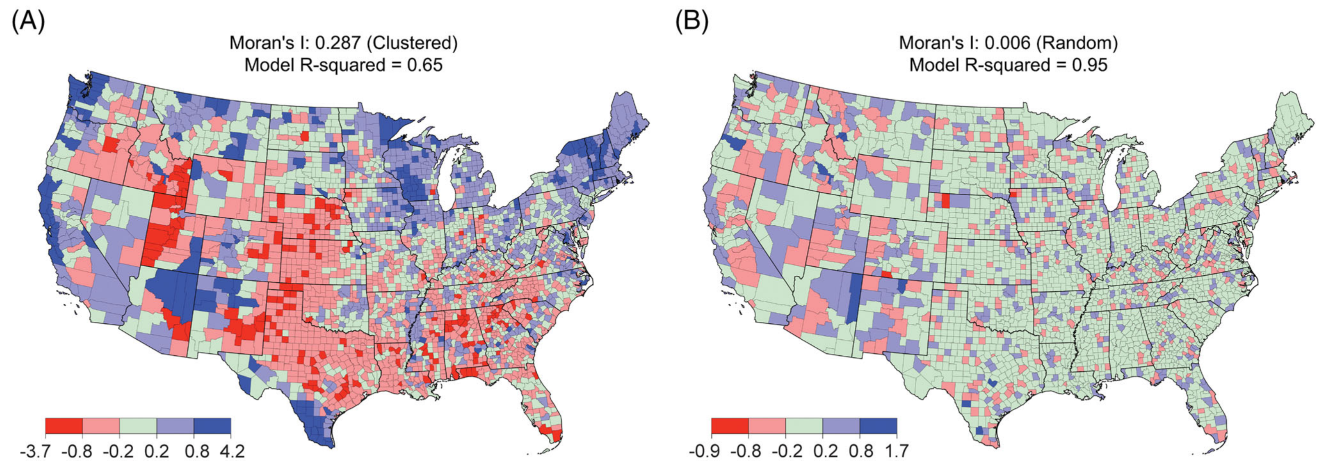

Figure 3. (A) The spatial distribution of the residuals for the global ordinary least squares model. The expected value of Moran's *I* if the distribution were random is 0.0003 and the observed value is 0.287 with a *p* value of 0.000. (B) The spatial distribution of the residuals for the multiscale geographically weighted regression model. The expected value of Moran's *I* if the distribution were random is 0.0003 and the observed value is 0.006 with a *p* value of 0.109.

population, foreign-born population, people with health insurance, voter turnout, the log of population density (a measure of "urbanness"), and votes for a third-party candidate.

The results just described represent averages across the country, however. Just as representing rainfall across the United States by a single, average value would be highly misleading, the parameter estimates in Table 1 would be equally misleading if the strength of the effect of each determinant of voting behavior were not constant across the country. We must allow the possibility that the strength or (indeed) even direction of each of the determinants of voting behavior included in Table 1 might vary spatially—just as the effect of income could vary between rich states and poor states (Gelman et al. 2007), other effects driving the percentage of Democratic votes in counties in Alabama might not be exactly the same as in counties in Oregon. Direct evidence for questioning the results presented in Table 1 is provided in Figure 3A, which shows the spatial distribution of the residuals for the global ordinary least squares (OLS) model. It is clear that the residuals exhibit strong spatial dependency, which invalidates the inferences made from the results. This is borne out statistically, because the Moran's *I* value computed from the residuals is extremely unlikely had their distribution been random.

A Local (MGWR) Model of Voter Preference

To examine potential spatial variations in the effect strength of various determinants of voting behavior, we calibrate the model in Equation 2 by MGWR (Fotheringham, Yang, and Kang 2017) using the MGWR 2.0 software freely available at https://sgsup.asu.edu/sparc/mgwr (Oshan et al. 2019). For the calibration, a spatially adaptive kernel function was selected using a bi-square weighting function with the number of nearest neighbors indicating the magnitude of the optimized bandwidth. Not only does the MGWR calibration technique generate far superior replications of the actual voting percentages in each county (the R^2 value increases from 0.66 to 0.95 and the corrected Akaike information criterion value decreases from 5,008 to 1,045), it typically yields residuals with little or no spatial dependence, as illustrated in Figure 3B. The Moran's *I* value for these residuals is 0.006, compared to 0.287 in the global model, and is not significantly different from zero. It is clear from Figure 4 that most of the values of the residuals from the MGWR calibration are close to zero.

We calculated two metrics to further assess the goodness of fit of the MGWR model: (1) a binary prediction accuracy measure that is the percentage of counties that are correctly predicted in terms of the majority of voters favoring the Republican or Democratic Party and (2) the mean absolute error between observed and predicted Democratic vote share. Both metrics are weighted by the total votes received in each county in the 2016 presidential election to account for the varying sizes of the counties. The results are shown in Table 2. The weighted prediction accuracy is 93.9 percent for all counties and 94.5 percent and 93.3 percent for Democratic and Republican counties, respectively. In terms of

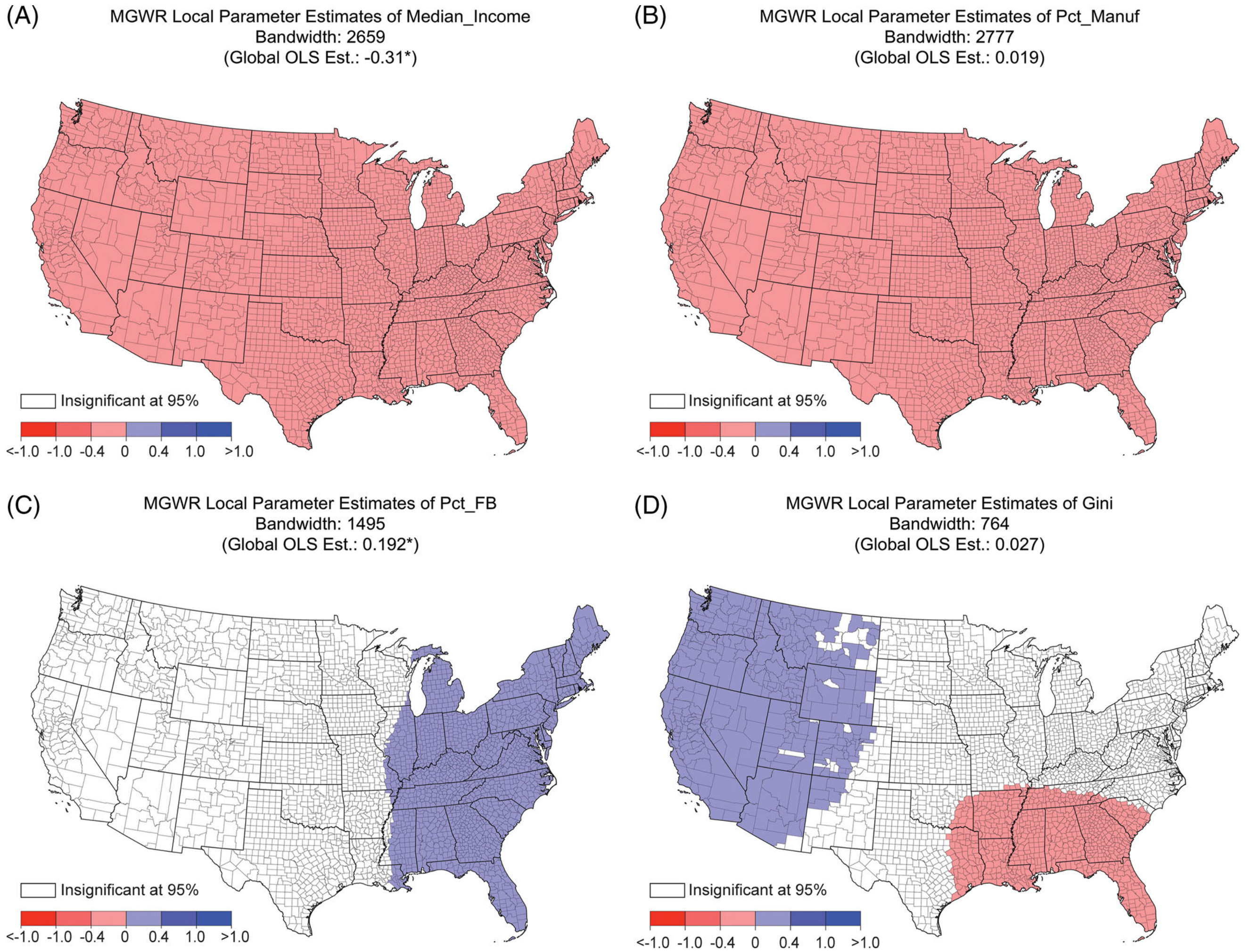

Figure 4. County-specific parameter estimates for (A) income, (B) manufacturing employment, (C) foreign-born population, and (D) income diversity within a county. *Note:* MGWR = multiscale geographically weighted regression; OLS = ordinary least squares.

the prediction error, the MGWR- predicted Democratic percentage vote only deviates by an average of 2.8 percent from the observed vote share regardless of whether a county voted Republican or Democrat. The performance of the model by state can be seen in Appendix A. Although the statistics suggest that the model performs very well in the vast majority of states for both Republican and Democratic counties, there are a couple of exceptions, such as Arizona and New Hampshire, where the binary statistics are low. In both cases this is a quirk of the statistics themselves. In Arizona, for example, a state with only fifteen counties, the model predicts Maricopa County to be won by the Democrats but it went marginally Republican. Given that Maricopa County contains approximately two-thirds of the state's population, the population-weighted binary measures become highly skewed.

In addition to producing superior estimates of the known voting percentages, the MGWR calibration also produces county-specific estimates of all of the parameters in the model, as opposed to a single, average value for the whole country, and it provides a covariate-specific estimate of the bandwidth, which is an indicator of scale over which different processes exhibit spatial heterogeneity. These bandwidths are robustly estimated and independent of the order in which the covariates are entered into the calibration when the variables are standardized, as they are here. Evidence of the stability of the bandwidths can be found in Appendix B, in which the order of covariates is shuffled in three calibrations yet the optimized bandwidths and their confidence intervals (CIs) remain essentially unchanged. Further discussion on the optimized bandwidths and their CIs can be found in Li et al. (2020). The

Table 2. Weighted binary prediction accuracy and mean absolute prediction error of multiscale geographically weighted regression

	Within all counties	Within Democratic counties	Within Republican counties
Weighted binary prediction accuracy	93.9%	94.5%	93.3%
Weighted mean absolute prediction error	2.8%	2.8%	2.8%

covariate-specific bandwidth estimates and CIs are shown in Table 3 and the county-specific parameter estimates are shown in Figures 4 through 7.

The covariates in Table 3 are listed in order of magnitude of their estimated bandwidths. Here, each bandwidth represents the number of nearest neighbors to each county from which data have been borrowed and down-weighted according to the distance from the regression focus. Larger bandwidths indicate more globally stationary relationships; smaller bandwidths indicate more locally varying relationships. It is clear from Table 3, for example, that the influence of the percentage of people employed in manufacturing within a county on the percentage vote for the Democratic Party is virtually the same in every county across the country, whereas the influence of young voters, African Americans, and the percentage who are insured varies considerably across counties. To see this more clearly, we now describe each of the fifteen sets of local parameter estimates because these demonstrate the relative stability or instability of the effects of each determinant on voting tendency across the country and also describe what is being missed by a traditional global modeling approach. In each map, only those county-specific estimates that are statistically significant are displayed—in each case, statistical significance has been adjusted to account for multiple hypothesis testing (Da Silva and Fotheringham 2016).

Figures 4A and 4B show the distribution of the county-specific parameter estimates for manufacturing employment and income, which, as their large optimal bandwidths suggest, are essentially constant across space. Both are significantly negatively related to voting for the Democratic Party in every county. The local parameter estimates for income align with their global counterpart, which is also significantly negative. The global parameter estimate for manufacturing, however, is not significantly different from zero and yet every local estimate is significantly negative. This is a spatial example of Simpson's paradox. Contradictory results are found in the same data analyzed at different geographical scales;

Table 3. Estimated covariate-specific optimal bandwidth from MGWR

Variables	MGWR bandwidths	95% bandwidth CI
Pct_Manuf	2,777	[2,158, 2,810]
Median_income	2,659	[2,158, 2,717]
Pct_FB	1,495	[1,100, 1,754]
Gini	764	[696, 1,100]
Pct_Age_65	656	[600, 850]
Sex_Ratio	603	[446, 850]
Pct_Hispanic	543	[446, 601]
Ln(Pop_Den)	387	[292, 446]
Pct_Bachelor	208	[174, 233]
Pct_3rd_party	160	[137, 173]
Turnout	117	[110, 137]
Pct_Age_18_29	58	[51, 65]
Pct_Black	43	[43, 45]
Pct_Insured	43	[43, 45]
Intercept	43	[43, 45]

Note: MGWR = multiscale geographically weighted regression; CI = confidence interval.

although manufacturing employment appears to have no significant impact on voting preference at the national level, local analysis shows that counties with higher proportions of people employed in manufacturing tend to record higher percentages of votes for the Republican party, *ceteris paribus*.

Figure 4C shows the county-specific parameter estimates for the variable percentage of foreign-born residents. Interestingly, the global estimate is significantly positive at $\alpha = 0.001$, but the county-level estimates are only significant for counties roughly east of the Mississippi. Either variations of foreign-born residents have no effect on voter preference west of the Mississippi or the variation in the data itself is too small to register any significant effect. What is interesting is that the global results give a very misleading picture of what is happening at the local level.

The local estimates for the Gini coefficient indicating income diversity within each county highlight the problem of relying on the global parameter estimate to provide accurate information that relates to all parts of the region of analysis. In this case the global estimate is insignificant, leading one to the conclusion that income diversity has no influence on voter preference anywhere in the county. The results of the MGWR calibration in Figure 4D belie that conclusion, though. In the far western states, the greater the income diversity, the greater the vote for the Democratic Party. Across the Southeast, though, greater income diversity is associated with voting for the Republican Party. Arguably, in the far West, individuals living in counties with greater income diversity are more sensitized to the plight of the less fortunate and tend to vote Democrat, whereas in the Southeast the opposite is the case.

In the global model, the effect of older voters and gender ratio on voter preference by county is insignificant, a result that runs counter to general intuition, but here the results are conditioned on other variables. As can be seen in Figures 5A and 5B, the local results largely support the nonsignificance of these two variables on voter preference but do highlight two interesting regional diversions from the national trend. In the western states there is a significant trend associating Republican voting and older population, whereas in the Northeast and the eastern parts of the Midwest there is a trend associating Republican voting with an increase in male population share. These relationships are not seen in any other parts of the country.

As suggested by the optimized bandwidths of 543 and 387 nearest neighbors, respectively, the spatial distributions of the local parameter estimates for percentage of Hispanic population within a county and the natural logarithm of population density both exhibit broad regional trends (see Figures 5C and 5D). Increasing proportions of Hispanics within a county has a strong and significant positive impact on the Democratic vote across the country, but the effect is strongest in the Southwest. The impact of population density, which is one measure of the urbanness of a county, is significantly positive on the Democratic vote in most of the middle and western parts of the country but insignificant along all of the East Coast and throughout much of the South. This result indicates that a division exists between rural and urban voters in the middle and western parts of the United States but along the East Coast and in the South, no such divergence exists.

Although the analysis is undertaken with county-level data (the model recognizes no state boundaries), the demarcation line between the significance and nonsignificance of the impact of urban-ness on voting preference is striking. The line almost precisely separates counties in Ohio (significant impact) from counties in the neighboring states of Pennsylvania and West Virginia (no significant impact), counties in Indiana (significant impact) from counties in Kentucky (no significant impact), counties in Oklahoma (significant impact) from counties in Arkansas (no significant impact), and counties in Texas (significant impact) from counties in Louisiana (no significant impact).

As the proportion of residents within a county with a bachelor's degree increases, the proportion of people voting Democrat increases significantly, *ceteris paribus*, in every county of the United States, although the impact of education on favoring the Democrat party is greater in the Northeast and in the West. Again, although the analysis is undertaken at the county level with no state variables, the boundaries where the strength of the relationship changes are sometimes very closely aligned with state boundaries—consider, for example, the Texas–New Mexico boundary in Figure 5E.

The two covariates relating to voting behavior, the percentage of votes cast within a county for third-party candidates and the percentage of eligible voters who actually voted in each county (turnout), reveal interesting spatial variations in their impacts on voting share for the Democratic Party across the country. Although the global parameter estimates for both variables are significantly positive, the local estimates indicate how misleading the global results can be. Third-party candidates are often varied and have a regional, rather than national, attraction to voters. This is clearly seen in Figure 6A, where across much of the eastern part of the country, and particularly in the Southeast, third-party candidates drew more heavily from Republican voters, increasing the Democratic vote share in a straight fight between Democrats and Republicans. The opposite is true in western states centered on Utah, where the third-party candidate drew more heavily from traditional Democratic voters and therefore increased the share of the Republican votes at the expense of Democratic votes. This seems entirely reasonable: Third-party candidates typically have regional appeal and arise from different ends of the political spectrum, so it would be very surprising if their appeal were

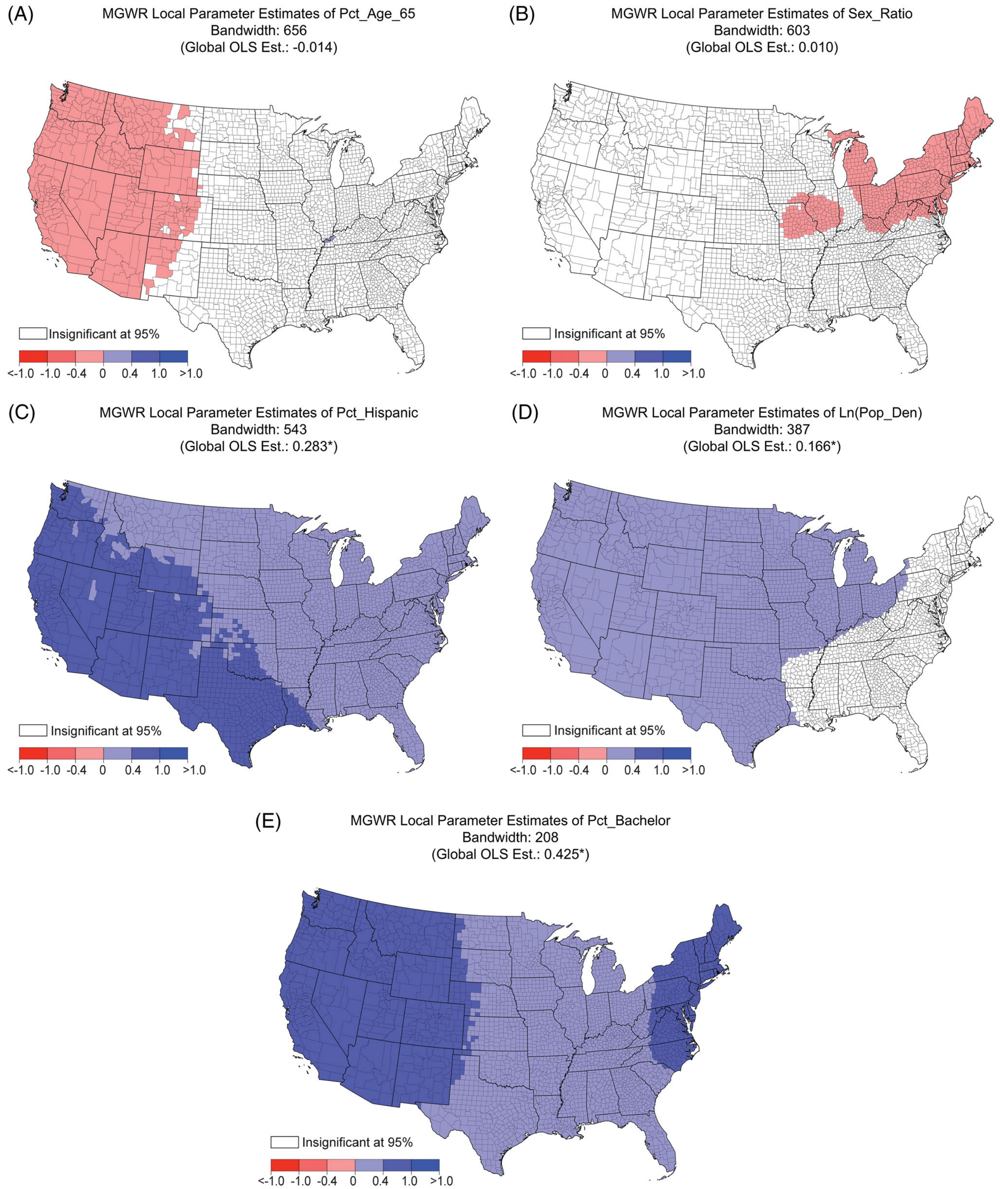

Figure 5. County-specific parameter estimates for (A) older voters, (B) gender balance, (C) percentage Hispanic, (D) ln(population density), and (E) education. *Note:* MGWR = multiscale geographically weighted regression; OLS = ordinary least squares.

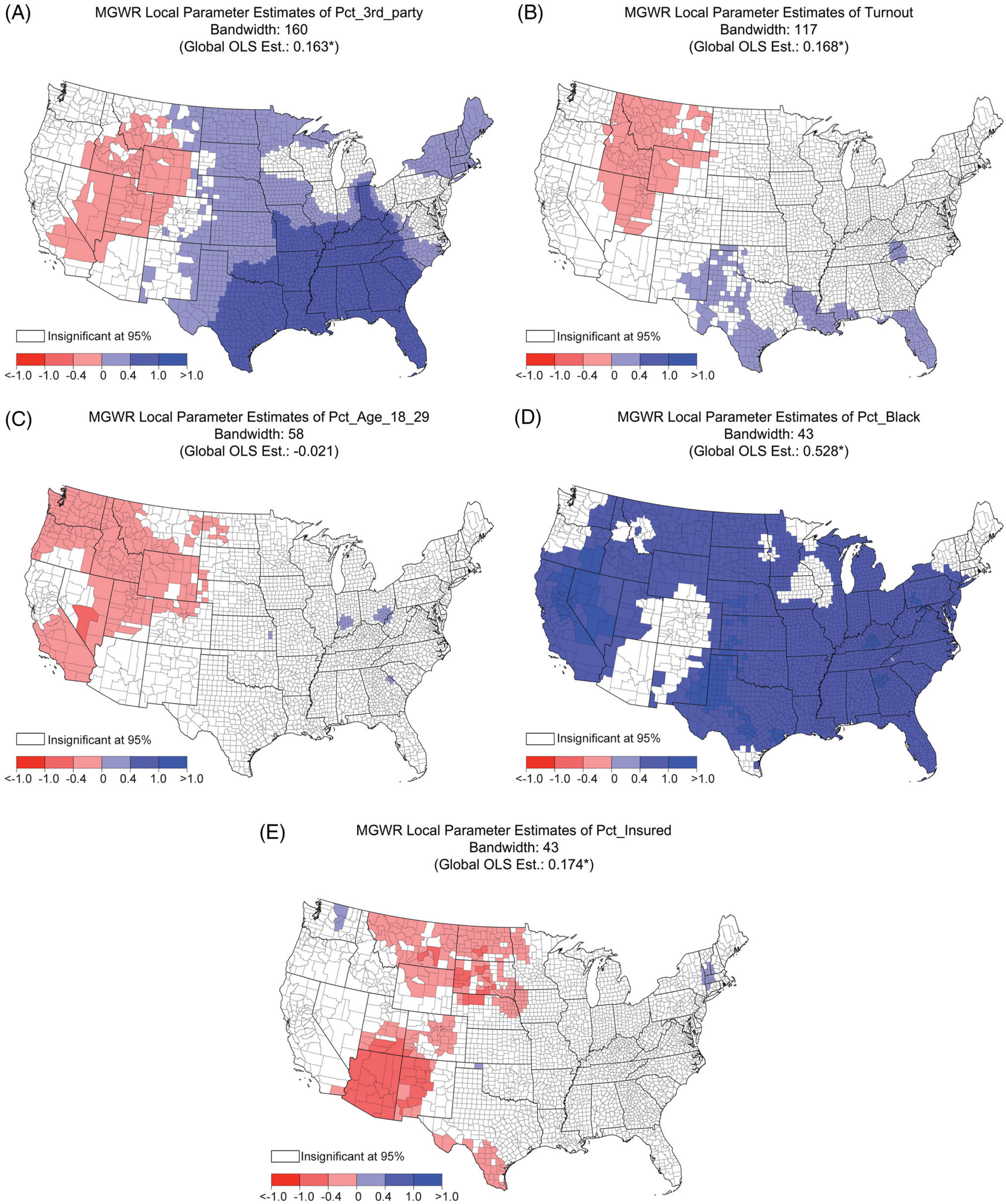

Figure 6. County-specific parameter estimates for (A) percentage third-party votes, (B) turnout percentage, (C) young voters, (D) percentage black, and (E) percentage insured. *Note:* MGWR = multiscale geographically weighted regression; OLS = ordinary least squares.

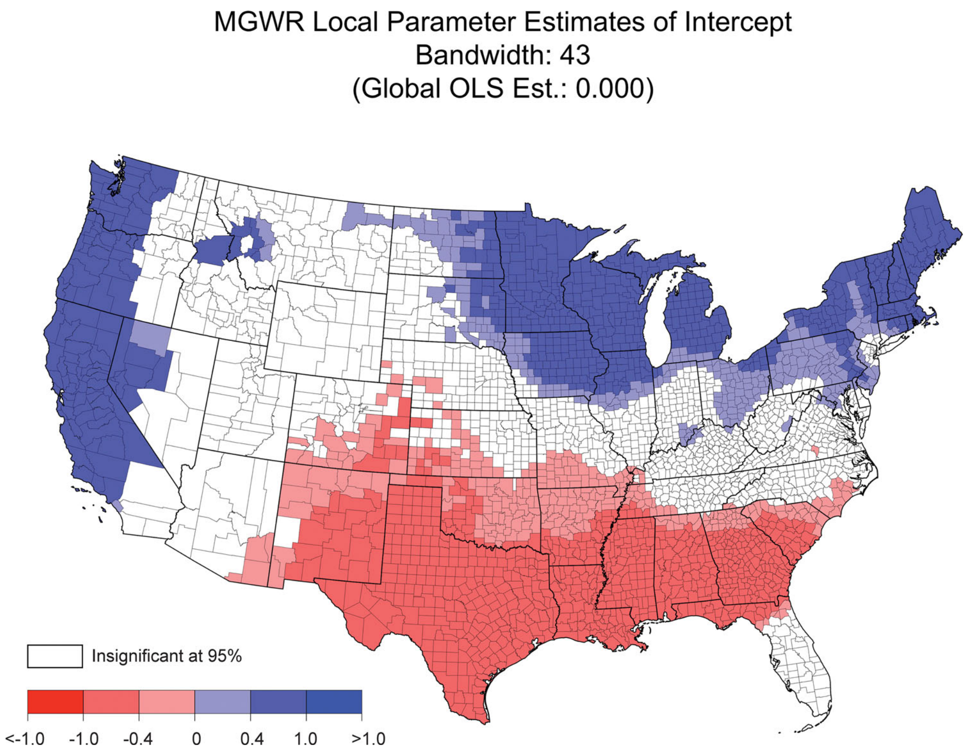

Figure 7. County-level parameter estimates of the intercept. *Note:* MGWR = multiscale geographically weighted regression; OLS = ordinary least squares.

uniform across the country as implied by the global parameter estimate.

Voter turnout is included in the model so that its effect can be accounted for and so the remaining parameters in the model are not obviously biased by its omission—there is no implication that turnout affects individual voter preference. The results, however, are quite striking. Contrary to the global model finding that high turnout rates favor the Democratic Party uniformly across the country, the local results suggest that it has an impact in only very limited parts of the country (see Figure 6B). The Democratic Party should be encouraging voter turnout in Florida, Louisiana, and the counties in southern Texas, and the Republican Party should be encouraging voter turnout in Idaho, Montana, and Utah; elsewhere turnout has little impact on the share of the vote.

Contrary to media reporting, the global results suggest that counties with larger shares of younger voters (aged eighteen to twenty-nine) do not exhibit any significant tendency to vote either Democrat or Republican. The local results in Figure 6C bear this out, with the vast majority of county-specific parameter estimates being insignificant. Exceptions to this are counties in Southern California and throughout much of the Northwest, where younger voters, *ceteris paribus*, show a significant inclination to vote Republican. Recall that these results are conditioned on many other factors, including education and race, whereas the polls about voting intention that are reported in the media typically are not. There are two very small clusters of counties in Indiana and Ohio centered on Indiana University and Ohio State University where there is a significant positive relationship between younger voters and the Democratic share of the vote. Such local variations are captured by the very small bandwidth reported for this relationship (fifty-eight nearest neighbors).

The tendency for African American voters to favor the Democratic Party is well known and is borne out by both the global result and the local parameter estimates displayed in Figure 6D. Most counties across the United States have extremely strong relationships between the percentage of African Americans in each county and the Democratic vote percentage. The only exceptions are swathes of counties in the Northwest, in the Southwest, the upper Midwest, New England, and the southern tip of Texas. What unites all of these counties is a very low proportion of African American residents, so there is very little variation locally in the percentage of African Americans.

Although the distribution of parameter estimates in Figure 6D appears fairly uniform, this is an artefact of the shading scheme employed and there are very strong local variations in the strength of the relationship between the percentage vote for the Democratic Party and the proportion of African Americans in each county. The optimized bandwidth for this relationship is forty-three nearest neighbors, the minimum allowed by the software for statistical purposes.

The relationship between the percentage of residents with health insurance in a county and percentage vote for the Democratic Party is significantly positive in the global model, yet the local parameter estimates shown in Figure 6E are largely insignificant. The one state showing a huge exception to this is Arizona, where the relationship is significantly negative for every county. The impact of the state borders with Nevada and California is again very noticeable. A significant negative relationship is also found in the northern Great Plains states, in the counties in southern Texas, and in the counties of New Mexico and Utah that border Arizona. In these counties, as the proportion of people with health insurance increases, the share of the vote for the Democratic Party decreases. Despite the global relationship being significantly positive, such a relationship is found for only a handful of counties, again suggesting that a variant of Simpson's paradox is present in the results. How one interprets the relationship between the proportion of people with health insurance and support for the Democratic Party very much depends on the scale of the analysis.

Along with the covariate-specific optimized bandwidths, the local estimates of the intercept are perhaps the most interesting output from MGWR. These indicate the intrinsic levels of the dependent variable holding everything else in the model constant. In this case, the local intercept estimates indicate the intrinsic support for the Democratic Party (positive) or the Republican Party (negative). In essence this is a measure of context.[4] Another way of interpreting the local intercept estimates is that if all counties had exactly the same mix of population, the local intercept estimates indicate how each county would then vote due to the unmeasurable effects of place. In this case, the results suggest that intrinsically leaning Democratic counties are found throughout New England, the upper Midwest, and down the Pacific Coast. Intrinsically Republican-leaning counties are found throughout the South, except for Florida. These findings conform very well to the general preconceptions of voter preferences in the United States. Here, for the first time, not only have we been able to identify these counties but we can also measure their degree of "intrinsic Republicanism" or "intrinsic Democraticism" and translate this into actual percentage votes for one party or the other (see later). Note the very detailed nature of the results (reflected in the optimized bandwidth being only forty-three). The technique is able, for instance, to separate the intrinsically Democrat-leaning counties in western Oregon and Washington from the intrinsically neutral counties in the eastern half of both states. Similarly, it separates northern (intrinsically Democrat-leaning) and southern (intrinsically neutral) counties in Illinois; intrinsically neutral West Virginia from intrinsically Democrat-leaning Ohio; intrinsically neutral North Carolina from intrinsically Republican-leaning South Carolina; the intrinsically Republican-leaning Deep South of South Carolina, Georgia, Alabama, and Mississippi from the intrinsically neutral counties to the north; Michigan from Indiana; and the panhandle and northern counties of Florida from the rest of the state. Figure 7 represent a new political geography—that of context in influencing voter preferences.

Separating the Role of Context from Socioeconomic Effects on Voter Preference

The ability to estimate a local intercept in MGWR allows the separation of contextual effects on voter preferences from socioeconomic effects. It further allows the influence of both of these contributions to the way people vote to be quantified. To see this, consider the model being calibrated by MGWR:

$$y_i^* = \alpha_i + \sum_j \beta_{ij} x_{ij}^* + \varepsilon_i, \qquad (2)$$

where y_i^* is the standardized value of y_i for county i, x_{ij}^* represents the standardized value of the jth covariate x for county i, α_i is the intercept for county i, and β_{ij} represents the slope coefficient for the jth covariate for county i. This equation can be rewritten as

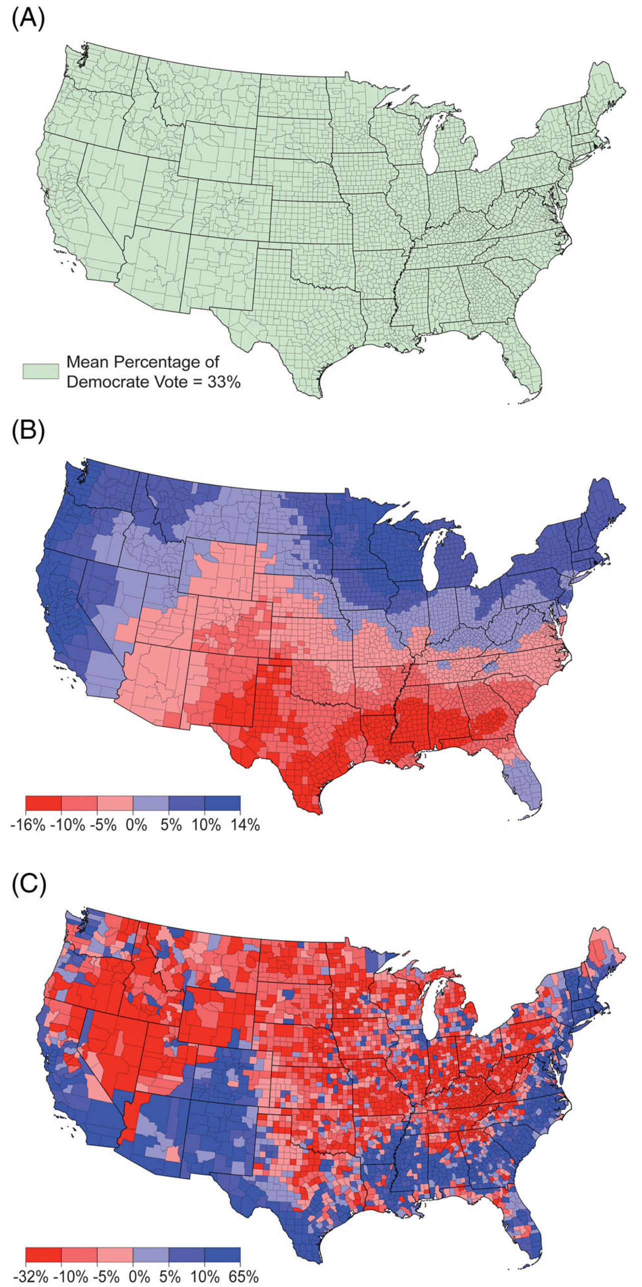

Figure 8. Contributions to the overall vote for the Democratic Party in each county: (A) mean county vote, (B) contribution due to context, and (C) contribution due to population mix.

$$(y_i - \bar{y})/\sigma_y = \alpha_i + \sum\nolimits_j \beta_{ij}\ (x_{ij} - \bar{x}_j)\ /\sigma_{x_j}, \qquad (3)$$

where $\bar{y}$ denotes the mean county value of y_i and $\bar{x}_j$ denotes the mean of the jth covariate over all counties. The symbol σ represents the standard deviation of y or the jth covariate over all counties. This equation can be expanded and rearranged to produce an expression for the percentage vote for the Democratic Party in each county, y_i, as

$$y_i = \bar{y} + \alpha_i\sigma_y + \sigma_y \sum\nolimits_j \beta_{ij}\ (x_{ij} - \bar{x}_j)\ /\sigma_{x_j}, \qquad (4)$$

which has three components. $\bar{y}$ is the mean vote for the Democratic Party across all counties, $\alpha_i\sigma_y$ is the proportion of the vote for the Democratic Party due to location or context, and $\sigma_y \sum_j \beta_{ij}\ (x_{ij} - \bar{x}_j)\ /\sigma_{x_j}$ is the proportion of the vote for the Democratic Party due to the mix of population within each county. We can map these three components as shown in Figure 8, where the predictions are compared to actual county-level votes.

Figure 8A simply denotes the average county-level vote for the Democratic Party, which was approximately 35 percent. The share of the popular vote for the Democratic Party was actually just over 50 percent—the discrepancy is due to the uneven population sizes of counties across the United States and the fact that counties with larger populations (urban areas) tend to strongly favor the Democratic Party.

Figure 8B transforms the parameter estimates shown in Figure 7 into votes per county gained or lost by the Democratic Party due to context. This figure paints a vivid picture of deep-seated political alignments that exist across the country and are independent of the mix of population within each county. This is not a map of how counties voted but of their underlying favoritism of one party over another. In some counties the latent preference for the Democratic party is worth 14 percentage points; in others the latent preference for the Republican Party loses the Democratic Party up to 21 percentage points. What is illuminating in Figure 8B is the distribution of the near politically neutral counties that essentially fill in the central part of the country, folding down to Arizona but also including Florida. Counties in this belt essentially decide the election.

Figure 8C shows the impact of the population composition in each county on the share of the vote gained by the Democratic Party. It shows a voting pattern perhaps unfamiliar to most people—that of how each county would have voted if context were not important and voting behavior was determined solely and uniformly by the type of voters in each county. It shows, for instance, that most counties in the South should be carried by the Democratic Party because of their population composition but are generally Republican strongholds. It shows that many counties in states such as Minnesota and Wisconsin should vote Republican but typically vote Democrat. It highlights very starkly that New Mexico and

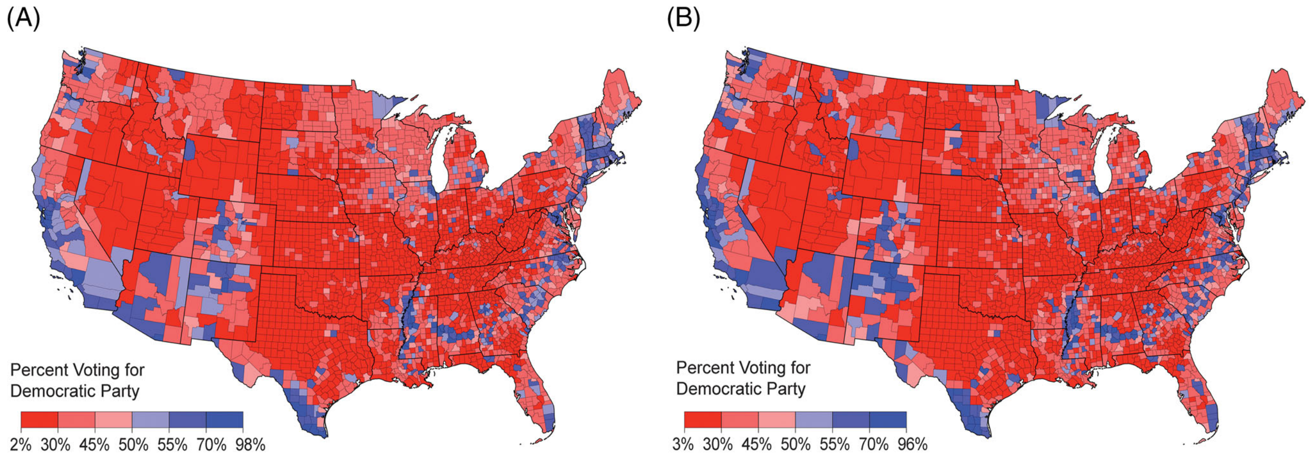

Figure 9. (A) Predicted and (B) actual percentage votes for the Democratic Party by county.

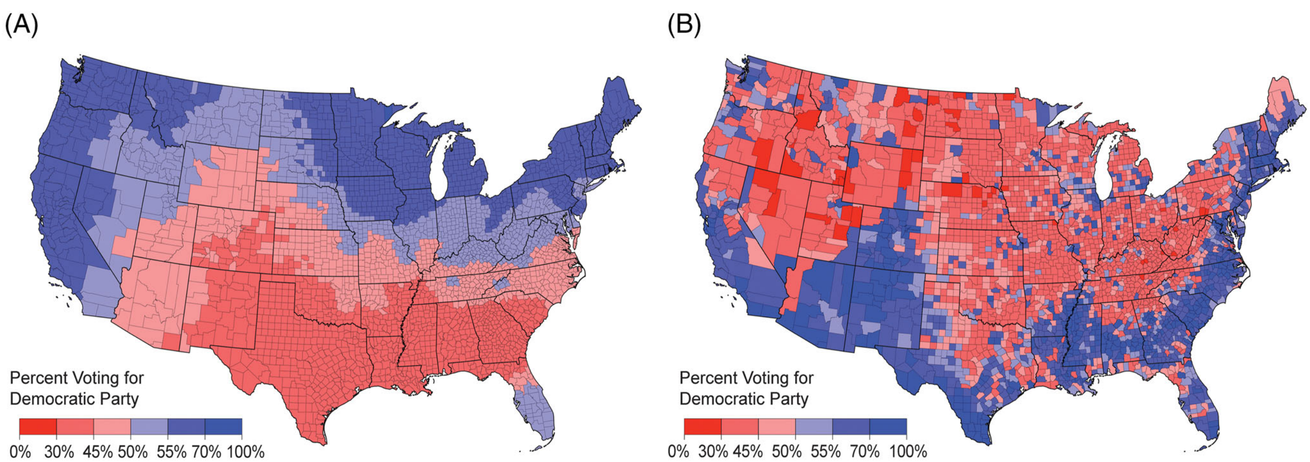

Figure 10. How each county would have voted (A) if population composition were constant across the country and (B) if geographical context did not matter.

Arizona should be Democratic strongholds but are not. This is a political geography that has not been seen before and its viewing is only possible through the calibration of local models such as MGWR.

If the predicted votes given in Figures 8A through 8C are summed, the predicted vote for the Democratic Party in each county is obtained. These values are shown in Figure 9A alongside the actual votes in Figure 9B. The similarity of these two distributions indicates the accuracy of the local MGWR model in replicating the voting patterns in the 2016 U.S. presidential election ($R^2 = 0.95$).

Finally, the results of the local modeling calibration allow two interesting, hypothetical, questions to be answered:

1. What would have been the result of the 2016 U.S. presidential election if every county had the same population composition?
2. What would have been the result of the 2016 U.S. presidential election if geographical context did not matter?

Both questions can be answered from Equation 4. If population composition were uniform across the country, the third term on the right side of Equation 4 disappears because all of the coefficients would be zero, so that percentage vote for the Democratic Party is calculated as

$$y_i = \bar{y} + \alpha_i \sigma_y. \quad (5)$$

This distribution is shown in Figure 10A with the overall Democratic share of the vote being 51 percent. The election would have been a close call, essentially dividing the country into two roughly equally geographical regions, but probably shaded by the Democrats winning states such as New York, California, Illinois, Ohio, and Pennsylvania with large numbers of electoral college votes.

Conversely, if geographical context had not been a factor in the 2016 election, how would counties have voted? This can be answered by deleting the second term on the right side of Equation 4. The resulting distribution of votes for the Democratic Party is shown in Figure 10B. In this case, the result would be a much clearer win for the Democratic Party, carrying most of the states with large populations. The pattern of votes for the Democratic Party would have been a fascinating near-reversal of what actually happened, with the South going largely for the Democrats and the North going largely for the Republicans.

Conclusions

This article addresses two major issues in understanding voting behavior. First, it questions the assumption that the determinants of voting behavior are constant across space. The use of the MGWR framework not only allows local parameter estimates to be estimated but also generates covariate-specific bandwidths that describe the spatial scale over which the processes affecting voting behavior operate. Some effects on voting behavior, for example, are shown to be global, whereas others operate over very local scales. The joint impact of modeling this more complex behavior is a far superior model both in terms of replicating voting patterns and in terms of generating information on the processes affecting how people vote. As such, this article presents a major new avenue of research into electoral geography. Second, the use of the MGWR local modeling framework allows the identification and measurement of the effect of geographical context on voting behavior. Both the presence and extent of this effect have been a source of debate for decades in the political geography arena, and this article again presents a major impetus to research on this topic. Arguably for the first time it is now possible to measure the strength of the influence of place on behavior, providing an important bridge between two previously dichotomous approaches to the study of human geography: the place versus space or the idiographic versus nomothetic debate. This article represents the use of a statistical model to capture the local effects of place on behavior and suggests that place is an important driver of voter preference.

Altogether, the ability to separate geographical context from socioeconomic drivers of voter behavior helps us understand the implicit geography of U.S. politics. Although the red state–blue state understanding of U.S. electoral geography is both pervasive and parsimonious, it must be acknowledged that U.S. politics is more than fifty state-based elections in parallel: There are substate and cross-state patterns in socioeconomic and contextual effects on how people vote. One could, for example, claim that the effects described here might be captured by the addition of various state-level variables to the model, but there are three reasons why this would not be useful. The first is exemplified by a binary variable such as 1 if the state has a Republican governor and 0 otherwise. For many local regressions this variable would have no variation, because the surrounding counties would all have a value of 1 or 0 and so no local parameter estimates would be forthcoming in such instances. States that have Republican governors tend to have a high degree of spatial clustering (e.g., the South), so this would be a major problem. Secondly, suppose that a state-level continuous variable were added that superficially would not appear to cause such a problem as identified previously (the values for all counties in the same state would still be the same but counties in neighboring states would have different values). A problem would still exist because the variable would no longer be continuous—it would have several levels in each local model and possibly just two levels in many cases where counties in the local regressions are drawn from two contiguous states. A third reason is that even if these statistical issues did not arise, there would still be a problem of assuming that the processes being investigated ended at state boundaries.

Equally, one could try to explain away some of the context by adding a Republican or Democratic indicator to each state, but that would not explain why such indicators exist in the first place; that is, why do some states perennially have Republican governors and others have Democratic governors despite the socioeconomic compositions of their populations? Our argument is that this can be attributed to context—an intrinsic leaning toward one political party—and our model provides a data-driven measurement of this otherwise nebulous, yet important, concept. In addition, the research interrogates both substate and cross-state patterns in the partisan alignment of places and finds some significant and important results.

First, the use of MGWR pivots the analysis away from the state-focused lens used elsewhere in U.S. electoral studies. This allows us to show that the contextual factors affecting voting outcomes in counties do not always neatly follow state lines. Instead, they have a distinctive geographical pattern. As can be seen in the estimate maps for the effect of income inequality, turnout, third parties, or health insurance coverage, socioeconomic factors might not have the same effect over the entire nation, or even within a single state. Past studies examining how these effects change from state to state typically assume that effects are constant within each state. This assumption might not hold in general.

Second, the separation of socioeconomic factors from contextual and geographic ones allows us to isolate the expected pattern of votes in a hypothetical election where context does not count. This hypothetical geography is wholly unlike the one observed in modern electoral cycles, suggesting that contextual effects are integral to the outcome of elections. An inversion of the usual north–south voting alignments would be extremely unusual. The urban–rural geography expressed in these factors is recognizable, though, so the urban–rural divide likely will remain important in understanding the interplay between socioeconomic determinants and contextual effects on voting.

Third, many traditional "battleground states," such as Ohio, Florida, or Virginia, are neatly split by contextual factors but have consistent socioeconomic voting influences at the state level. This suggests that these battlegrounds are drawn primarily by conflicting contextual effects acting within a state, rather than by divisions in the socioeconomic factors.

Altogether, this suggests that although state geography is useful for understanding U.S. electoral outcomes (i.e., who wins at the end of the day), state geography obscures the actual variation in socioeconomic and contextual effects that drive the outcome (why they win). The geographical scale and structure of variation can be different depending on the socioeconomic factor under analysis. Further, contextual effects vary significantly from state to state. They lie at the foundations of many battleground states, often having a positive impact in one area of the state but a negative impact elsewhere. In an election where context does not count, the resulting geography of partisan alignment would look very different from what occurred in 2016.

Finally, our results for the first time quantify a map of the geographical impacts of context on voting preferences by county (Figure 8B). This is a map of geographic partisanship that shows the depth of Republicanism across the Southeast (except central and southern Florida) countered by a "blue wall" of Democratic support across the Northeast, upper Midwest, and down the West Coast. The counties forming the border between these two blocs represent those areas where voters are most willing to change their voting preferences due to changes in prevailing conditions. It would be illuminating to reproduce a similar map for previous presidential elections to assess the changes in both the geography and degree of electoral partisanship over time. It is hoped that this article will stimulate many such studies.

Notes

1. The data are available from the MIT Election Data and Science Lab (see https://electionlab.mit.edu/data).
2. The variables we selected were the product of a long period of experimentation, discussion with political scientists and political geographers, a thorough reading of the literature in this area, and common sense. We cannot, however, claim that the selection is theory based, because there is little acknowledged theory in the social sciences and particularly when it comes to examining the determinants of individuals' voting preferences. This should not denigrate the research, however; much can be learned from empirical experimentation. Confidence in the variable selection is gained from the very strong replication power of our model both at the global level and at the local level (see below).
3. The data were gathered from the American Community Survey 2012–2016 five-year estimates. No variance inflation factor was greater than 4.
4. There is little evidence to suggest that any major explanatory variable has been omitted in our model, nor what such a variable would be to "explain" the spatial pattern of what we call context here. As described earlier, what is meant by context here is an otherwise unmeasurable effect on people's preferences for one political party over another based solely on locality (the influence of family and friends, local media, etc.), which is independent of any other measurable effect on voting preference (e.g., income, ethnicity, age, etc.). The pattern that emerges is very similar to that of most people's mental images of political leanings across the United

States. Here, possibly for the first time, we have been able to quantify this effect.

ORCID

Ziqi Li http://orcid.org/0000-0002-6345-4347
Levi John Wolf http://orcid.org/0000-0003-0274-599X

References

Agnew, J. 1996. Mapping politics: How context counts in electoral geography. *Political Geography* 15 (2):129–46.

Altman, I., and J. F. Wohlwill. 2012. *Human behavior and environment: Advances in theory and research*. Vol. 2. Boston, MA: Springer Science & Business Media.

Archer, J. C. 1988. Macrogeographical versus microgeographical cleavages in American presidential elections 1940–1984. *Political Geography Quarterly* 7 (2):111–25. doi: 10.1016/0260-9827(88)90023-7.

Ayllón, S. 2013. Unemployment persistence: Not only stigma but discouragement too. *Applied Economics Letters* 20 (1):67–71. doi: 10.1080/13504851.2012.681020.

Bensel, R. F. 1987. *Sectionalism and American political development*, 1880–1980. Madison: University of Wisconsin Press.

Braha, D., and M. A. M. de Aguiar. 2017. Voting contagion: Modeling and analysis of a century of U.S. presidential elections. *PLoS ONE* 12 (5):e0177970. doi: 10.1371/journal.pone.0177970.

Branton, R. P. 2003. Examining individual-level voting behavior on state ballot propositions. *Political Research Quarterly* 56 (3):367–77. doi: 10.1177/106591290305600311.

Calvo, E., and M. Escolar. 2003. The local voter: A geographically weighted approach to ecological influence. *American Journal of Political Science* 47 (1):189–204. doi: 10.1111/1540-5907.00013.

Cho, W. K. T., and J. G. Gimpel. 2010. Rough terrain: Spatial variation in campaign contributing and volunteerism. *American Journal of Political Science* 54 (1):74–89. doi: 10.1111/j.1540-5907.2009.00419.x.

Cho, W. K. T., and J. G. Gimpel. 2012. Geographic information systems and the spatial dimensions of American politics. *Annual Review of Political Science* 15 (1):443–60. doi: 10.1146/annurev-polisci-031710-112215.

Clark, A. E. 2003. Unemployment as a social norm: Psychological evidence from panel data. *Journal of Labor Economics* 21 (2):323–51. doi: 10.1086/345560.

Clemens, A. C., M. H. Crespin, and C. J. Finocchiaro. 2015. The political geography of distributive politics. *Legislative Studies Quarterly* 40 (1):111–36. doi: 10.1111/lsq.12067.

Coffee, N. T., T. Lockwood, P. Rossini, T. Niyonsenga, and S. McGreal. 2020. Composition and context drivers of residential property location value as a socioeconomic status measure. *Environment and Planning B: Urban Analytics and City Science* 47 (5):790–807. doi: 10.1177/2399808318805489.

Darity, W. A., and A. H. Goldsmith. 1996. Social psychology, unemployment and macroeconomics. *Journal of Economic Perspectives* 10 (1):121–40. doi: 10.1257/jep.10.1.121.

Darmofal, D. 2008. The political geography of the New Deal realignment. *American Politics Research* 36 (6):934–61. doi: 10.1177/1532673X08316591.

Da Silva, A., and A. S. Fotheringham. 2016. The multiple testing issue in geographically weighted regression. *Geographical Analysis* 48 (3):233–47. doi: 10.1111/gean.12084.

Enos, R. 2017. *The space between us*. London: Cambridge University Press.

Escobar, A. 2001. Culture sits in places: Reflections on globalism and subaltern strategies of localization. *Political Geography* 20 (2):139–74.

Fotheringham, A. S., W. Yang, and W. Kang. 2017. Multiscale geographically weighted regression (MGWR). *Annals of the American Association of Geographers* 107 (6):1247–65. doi: 10.1080/24694452.2017.1352480.

Garz, M. 2018. Effects of unemployment news on economic perceptions—Evidence from German federal states. *Regional Science and Urban Economics* 68:172–90. doi: 10.1016/j.regsciurbeco.2017.11.006.

Gelman, A., B. Shor, J. Bafumi, and D. Park. 2007. Rich state, poor state, red state, blue state: What's the matter with Connecticut? *Quarterly Journal of Political Science* 2 (4):345–67.

Golledge, R. G. 1997. *Spatial behavior: A geographic perspective*. New York: Guilford.

Gomez, B. T., T. G. Hansford, and G. A. Krause. 2007. The Republicans should pray for rain: Weather, turnout, and voting in U.S. presidential elections. *The Journal of Politics* 69 (3):649–63. doi: 10.1111/j.1468-2508.2007.00565.x.

Guth, J. L., L. A. Kellstedt, C. E. Smidt, and J. C. Green. 2006. Religious influences in the 2004 presidential election. *Presidential Studies Quarterly* 36 (2):223–42. doi: 10.1111/j.1741-5705.2006.00300.x.

Hersh, E., and C. Nall. 2016. The primacy of race in the geography of income-based voting: New evidence from public voting records. *American Journal of Political Science* 60 (2):289–303.

Hillygus, D. S., and S. Jackman. 2003. Voter decision making in election 2000: Campaign effects, partisan activation, and the Clinton legacy. *American Journal of Political Science* 47 (4):583–96. doi: 10.1111/1540-5907.00041.

Hillygus, D. S., and T. G. Shields. 2005. Moral issues and voter decision making in the 2004 presidential election. *PS: Political Science & Politics* 38 (2):201–09.

Key, V., and A. Heard. [1949] 1984. *Southern politics in state and nation*. Knoxville: University of Tennessee Press.

Kim, J., E. Elliott, and D. M. Wang. 2003. A spatial analysis of county-level outcomes in U.S. presidential elections: 1988–2000. *Electoral Studies* 22 (4):741–61. doi: 10.1016/S0261-3794(02)00008-2.

King, G. 1996. Why context should not count. *Political Geography* 15 (2):159–64. doi: 10.1016/0962-6298(95)00079-8.
Krug, S. E., and R. W. Kulhavy. 1973. Personality differences across regions of the United States. *The Journal of Social Psychology* 91 (1):73–79.
Leigh, A. 2005. Economic voting and electoral behavior: How do individual, local, and national factors affect the partisan choice? *Economics & Politics* 17 (2):265–96.
Levernier, W., and A. G. Barilla. 2006. The effect of region, demographics, and economic characteristics on county-level voting patterns in the 2000 presidential election. *Review of Regional Studies* 36 (3):427–47.
Li, Z., A. S. Fotheringham, T. M. Oshan, and L. J. Wolf. 2020. Measuring bandwidth uncertainty in multiscale geographically weighted regression using Akaike weights. *Annals of the American Association of Geographers* 110 (5):1500–1520. doi: 10.1080/24694452.2019.1704680.
Manley, E., and U. Demsar. 2015. Space matters: Geographic variability of electoral turnout determinants in the 2012 London mayoral election. *Electoral Studies* 40:322–34.
McCombs, M., J. P. Llamas, E. Lopez-Escobar, and F. Rey. 1997. Candidate images in Spanish elections: Second-level agenda-setting effects. *Journalism & Mass Communication Quarterly* 74 (4):703–17. doi: 10.1177/107769909707400404.
Miller, J., and T. Grubesic. 2020. A spatial exploration of the halo effect in the 2016 U.S. presidential election. *Annals of the American Association of Geographers*. https://doi.org/10.1080/24694452.2020.1785271
Mutz, D. C. 2018. Status threat, not economic hardship, explains the 2016 presidential vote. *Proceedings of the National Academy of Sciences of the United States of America* 115 (19): E4330–E4339. www.pnas.org/cgi/doi/10.1073/pnas.1718155115.
O'Loughlin, J. 2018. Thirty-five years of political geography and *Political Geography:* The good, the bad and the ugly. *Political Geography* 65:143–51.
Oshan, T. M., Z. Li, W. Kang, L. J. Wolf, and A. S. Fotheringham. 2019. mgwr: A Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS International Journal of Geo-Information* 8 (6):269. doi: 10.3390/ijgi8060269.
Pantoja, A. D., and G. M. Segura. 2003. Does ethnicity matter? Descriptive representation in legislatures and political alienation among Latinos. *Social Science Quarterly* 84 (2):441–60. doi: 10.1111/1540-6237.8402014.
Payne, B. K., J. A. Krosnick, J. Pasek, Y. Lelkes, O. Akhtar, and T. Tompson. 2010. Implicit and explicit prejudice in the 2008 American presidential election. *Journal of Experimental Social Psychology* 46 (2):367–74. doi: 10.1016/j.jesp.2009.11.001.
Plaut, V. C., H. R. Markus, and M. E. Lachman. 2002. Place matters: Consensual features and regional variation in American well-being and self. *Journal of Personality and Social Psychology* 83 (1):160–84. doi: 10.1037/0022-3514.83.1.160.
Powell, G. B. 1986. American voter turnout in comparative perspective. *American Political Science Review* 80 (1):17–43. doi: 10.2307/1957082.
Rentfrow, P. J., S. D. Gosling, M. Jokela, D. J. Stillwell, M. Kosinski, and J. Potter. 2013. Divided we stand: Three psychological regions of the United States and their political, economic, social, and health correlates. *Journal of Personality and Social Psychology* 105 (6):996–1012. doi: 10.1037/a0034434.
Rentfrow, P. J., M. Jokela, and M. E. Lamb. 2015. Regional personality differences in Great Britain. *PLoS ONE* 10 (3). doi: 10.1371/e0122245.
Rocha, R. R., and R. Espino. 2009. Racial threat, residential segregation, and the policy attitudes of Anglos. *Political Research Quarterly* 62 (2):415–26. doi: 10.1177/1065912908320931.
Scala, D. J., K. M. Johnson, and L. T. Rogers. 2015. Red rural, blue rural? Presidential voting patterns in a changing rural America. *Political Geography* 48:108–18. doi: 10.1016/j.polgeo.2015.02.003.
Schaffner, B. F., M. MacWilliams, and T. Nteta. 2018. Understanding white polarization in the 2016 vote for president: The sobering role of racism and sexism. *Political Science Quarterly* 133 (1):9–34. doi: 10.1002/polq.12737.
Sigelman, L., and C. K. Sigelman. 1982. Sexism, racism, and ageism in voting behavior: An experimental analysis. *Social Psychology Quarterly* 45 (4):263–69. doi: 10.2307/3033922.
Van de Vliert, E. 2008. *Climate, affluence, and culture*. London: Cambridge University Press.

A. STEWART FOTHERINGHAM is a Regents' Professor of Computational Spatial Science in the School of Geographical Sciences and Urban Planning at Arizona State University, Tempe, AZ 85281. E-mail: stewart.fotheringham@asu.edu. His research interests include spatial data analytics, spatial processes, and spatial interaction modeling.

ZIQI LI is a Visiting Assistant Professor in the Department of Geography and Geographic Information Science, University of Illinois, Urbana–Champaign, IL 61820. E-mail: ziqi@illinois.edu. His research interests include GIScience, spatial data science, spatial statistical learning, and their applications in multidisciplinary fields.

LEVI JOHN WOLF is a Senior Lecturer (Assistant Professor) in the School of Geographical Sciences at the University of Bristol, UK. E-mail: levi.john.wolf@bristol.ac.uk. His research interests include probabilistic programming, machine learning for applications in urban geography, politics, sociology, and economics.

Appendix A. Votes weighted binary prediction accuracy and absolute prediction error for each state

	Votes weighted binary prediction accuracy			Votes weighted absolute prediction error		
State	Overall (%)	Democratic counties (%)	Republican counties (%)	Overall (%)	Democratic counties (%)	Republican counties (%)
AL	99.1	96.2	100.0	1.5	1.3	1.5
AR	100.0	100.0	100.0	2.4	3.4	2.2
AZ	38.3	100.0	22.9	7.2	2.9	8.3
CA	97.4	99.8	70.0	3.1	2.9	5.3
CO	77.3	74.6	81.7	4.2	2.8	6.5
CT	100.0	100.0	100.0	0.9	0.7	3.0
DC	100.0	100.0	No Republican counties	2.2	2.2	No Republican counties
DE	100.0	100.0	100.0	1.4	1.5	1.3
FL	100.0	100.0	100.0	2.8	3.0	2.7
GA	99.1	98.8	99.4	2.2	2.5	1.9
IA	96.8	100.0	94.7	2.9	2.5	3.2
ID	97.4	37.1	100.0	3.8	7.2	3.7
IL	99.6	99.4	100.0	3.0	3.2	2.5
IN	96.0	85.0	100.0	2.3	1.3	2.6
KS	75.5	100.0	73.3	3.2	8.9	2.7
KY	100.0	100.0	100.0	2.3	2.3	2.3
LA	99.1	96.5	100.0	1.8	2.8	1.5
MA	100.0	100.0	No Republican counties	2.6	2.6	No Republican counties
MD	90.3	86.6	100.0	1.8	1.6	2.3
ME	81.5	66.3	100.0	2.1	2.8	1.2
MI	89.3	94.9	84.5	2.2	1.7	2.6
MN	94.5	89.5	100.0	2.7	3.1	2.2
MO	96.9	88.5	100.0	2.3	1.5	2.7
MS	98.0	92.3	100.0	1.9	1.6	2.0
MT	84.2	45.0	100.0	5.4	6.9	4.8
NC	94.3	90.9	97.6	2.3	2.4	2.1
ND	99.6	75.3	100.0	4.0	12.1	3.9
NE	83.9	63.8	100.0	3.5	3.2	3.7
NH	46.9	100.0	20.7	3.3	1.7	4.1
NJ	98.6	97.9	100.0	2.2	2.0	2.6
NM	85.1	86.5	81.2	5.6	5.8	5.2
NV	100.0	100.0	100.0	2.5	2.3	4.1
NY	93.4	97.3	85.0	2.4	2.5	2.4
OH	97.4	94.0	100.0	2.1	1.9	2.3
OK	100.0	No Democratic counties	100.0	2.5	No Democratic counties	2.5
OR	76.4	79.9	71.4	4.2	5.5	2.4
PA	91.5	84.0	100.0	2.3	2.2	2.4
RI	82.3	100.0	0.0	2.8	3.2	1.0
SC	95.8	100.0	94.4	1.8	1.1	2.1
SD	98.5	40.8	100.0	3.7	7.3	3.6
TN	100.0	100.0	100.0	2.3	3.2	1.9
TX	96.9	93.4	99.8	3.6	5.1	2.3
UT	63.0	4.7	100.0	7.9	12.3	5.0
VA	98.7	99.6	97.9	1.8	1.7	2.0
VT	100.0	100.0	100.0	3.1	3.1	4.2
WA	98.3	97.7	100.0	3.6	3.4	4.2
WI	93.7	82.3	100.0	3.6	4.9	2.8
WV	100.0	No Democratic counties	100.0	2.7	No Democratic counties	2.7
WY	100.0	100.0	100.0	2.7	7.4	2.4

Appendix B. Orders of covariates in the calibration and optimized covariate-specific bandwidths and confidence intervals

Covariate	Order in the model	Bandwidth	95% CI	Order in the model	Bandwidth	95% CI	Order in the model	Bandwidth	95% CI
Intercept	1	43	[43, 45]	1	43	[43, 45]	1	43	[43, 45]
Sex_Ratio	2	603	[446, 850]	10	550	[446, 850]	3	603	[446, 850]
Pct_Black	3	43	[43, 45]	11	43	[43, 45]	4	43	[43, 45]
Pct_Hispanic	4	543	[446, 601]	12	543	[446, 696]	11	543	[446, 696]
Pct_Bachelor	5	208	[174, 233]	13	208	[174, 233]	12	208	[174, 233]
Median_Income	6	2,659	[2,158, 2,717]	14	2,658	[2,158, 2,717]	5	2,658	[2,158, 2,717]
Pct_Age_65	7	656	[600, 850]	15	656	[600, 850]	6	656	[600, 850]
Pct_Age_18_29	8	58	[51, 65]	2	58	[51, 65]	13	58	[51, 65]
Gini	9	764	[696, 1,100]	3	763	[696, 1,100]	9	764	[696, 1,100]
Pct_Manuf	10	2,777	[2,158, 2,810]	4	2,777	[2,158, 2,810]	7	2,777	[2,158, 2,789]
Ln(Pop_Den)	11	387	[292, 446]	5	387	[292, 446]	8	387	[292, 446]
Pct_3rd_party	12	160	[137, 173]	6	139	[124, 160]	10	160	[137, 173]
Turnout	13	117	[110, 137]	7	139	[124, 160]	14	117	[110, 137]
Pct_FB	14	1,495	[1,100, 1,754]	8	1,487	[1,100, 1,754]	15	1,495	[1,100, 1,754]
Pct_Insured	15	43	[43, 45]	9	43	[43, 45]	2	43	[43, 45]

Note: CI = confidence interval.