The Spatial and Temporal Dynamics of Voter Preference Determinants in Four U.S. Presidential Elections (2008 - 2020)

Ziqi Li[1*], A. Stewart Fotheringham[2]

1. School of Geographical and Earth Sciences, University of Glasgow, United Kingdom
2. Spatial Analysis Research Center, School of Geographical Sciences and Urban Planning, Arizona State University, United States

***Corresponding Author: Ziqi Li** (Ziqi.Li@glasgow.ac.uk)
Address:
East Quad, University of Glasgow
Glasgow, United Kingdom G12 8QQ
Email: Ziqi.Li@glasgow.ac.uk

Funding
This work is supported by the US National Science Foundation under grant award no. 2117455.

Conflict of Interest

The authors declare that there is no conflict of interest.

The Spatial and Temporal Dynamics of Voter Preference Determinants in Four U.S. Presidential Elections (2008 - 2020)

Abstract

Political and social processes that shape people's voting preferences might be linked with geographical location, varying from place to place, and operating at local, regional and national scales. Here, we use a local modeling technique, multiscale geographically weighted regression (MGWR), to examine spatial and temporal variations in the influences of county-level socioeconomic factors on voter preference during the 2008 – 2020 US Presidential elections. We argue that the local intercept in the MGWR model is an indicator of the effect of spatial context on voter preference and not only can this be separated from the effect of other socio-economic factors, it needs to be in order to prevent misspecification bias in the indicators of these other factors. We identify strong and consistent divisions across the country in how context shapes election results.

Keywords: *Election dynamics, voter preference; multiscale geographically weighted regression; mgwr; scale*

1 Introduction

Examination of the social and political processes that shape voters' political preferences has been carried out systematically since the 1960s (*inter alia,* Cox 1969), and there is a substantial body of literature focused on the spatial nature of elections from both compositional (characteristics of individuals) and contextual (place) effects (Taylor 1973; Taylor and Johnson 1979; Johnson et al., 1990; Warf and Leib; Agnew, 1994; Agnew, 2014). Despite King (1996)'s view that political behavior can be solely explained by individual factors, political geographers generally believe that location and place play a vital role in shaping voters' preferences, and that political and social relationships may not bet uniform over space. The latter view has been reinforced by analyses using survey and polling data which suggest that individual characteristics are not sufficient to fully explain voting preferences, and there exists a certain amount of dependency between voting behavior and location and place. (O'Loughlin et al., 1994; Flint, 1996; Agnew, 1996; Sui and Hugil, 2002; Lappie and Marschall, 2018). However, the importance of context has not been well quantified in political geography (O'Loughlin, 2018) and the examination of a local contextual effect on voting behavior has been largely qualitative through the investigation of culture, history, personalities, media and other influences (*inter alia,* Morrill et al. (2011) on the 2004 and 2008 US Presidential elections). Here, we argue that contextual effects can be identified robustly through local spatial models and that this supports a 'place-based' political geography which emphasizes the importance of location in affecting voting behavior.

Recognizing that spatial variations in the determinants of voting behavior might exist, the use of classic global regression models, such as Ordinary Least Squares (OLS) and various types of spatial regression models, seems overly simplistic because such models provide no or limited insights into possible spatially varying behavior. To account for possible spatial heterogeneity in processes, various forms of local models have been developed such as the spatial expansion model (Casetti, 1972), spatial regimes models (Thioulouse et al., 1995), Geographically Weighted Regression (GWR) (Fotheringham et al., 2002), Bayesian spatially varying coefficients models (SVC) (Gelfand, 2003) and spatial filtering methods (Griffith, 2008). Compared to the spatial expansion and regimes models, GWR-type models are able to capture continuous spatial heterogeneity in a flexible way without needing any *a priori* knowledge of the geographic pattern of such heterogeneity (Fotheringham et al., 2002). The GWR framework produces local parameter estimates that are either comparable or superior to those produced by Bayesian SVC and spatial filtering methods and is arguably more intuitive, flexible and extendable (Wolf et al., 2018; Oshan et al., 2018). Additionally, GWR has been shown to be a useful technique in electoral geography to explore localized relationships and geographic variations in the determinants of voting behavior. For example, Shin and Agnew (2008) apply GWR to electoral and census data in a case study of the Italian Lega Nord Party and argue that it is crucial to take spatial issues into consideration when conducting statistical analyses of voting patterns. Darmofal (2008) employs GWR to account for the spatially varying effects of voting behavior and demographics during the 1928-1936 Democratic realignment period. Cho and Gimpel (2009) investigate the local impact of economic hardship in the 2008 presidential election using a similar model. Warf (2011) reports the non-stationary influence of class, ethnicity, and region on Obama's support in the 2008 Presidential election. Using GWR, Manley and Demšar (2015) analyze the spatial relationship between voter turnout and socio-economic and political variables in the 2012 London mayoral election, and Miller and Grubesic (2021) discover local halo effects and spatially heterogeneous relationships between socioeconomics and Republican support in the 2016 US Presidential election.

Despite the fact that GWR is able to capture localized relationships, the model is restricted by assuming that spatial variation in relationships operates at the same spatial scale for all processes by employing only one spatial weight matrix to account for the spatial structure. This is probably unrealistic in many real-world scenarios and almost certainly in voting behavior. What motivates and influences an individual to cast a particular vote is the consequence of a wide range of processes that may vary over space and operate at different geographic scales, from the very local to the regional or the global (Miller 1994). More specifically, Johnston and Pattie (2006) point out eight spatial scales, from the household to the neighborhood to the country that may have a compositional effect on people's voting behavior (Forest, 2017). The possible existence of multiple processes operating at different spatial scales has recently been recognized by Fotheringham et al. (2017) in a model formulation termed multiscale geographically weighted regression (MGWR) which allows the estimation of multiscale processes in a single local model. This is achieved by using a vector of covariate-specific bandwidths in contrast to the single bandwidth for all covariates that is employed in GWR. In this way, MGWR is able to account for potentially different degrees of spatial heterogeneity exhibited by the different processes represented in a model. The covariate-specific bandwidths control the amount of bias and variance in the local parameter estimates associated with each covariate and provide a comparable indicator of the spatial scale over which different processes are relatively stable (Fotheringham et al., 2017; Wolf et al. 2018; Li et al. 2020). A set of local parameter estimates obtained from a small optimized bandwidth indicates that the associated spatial process represented by the parameter estimates varies over relatively short distances, while a set of local parameter estimates obtained from a large bandwidth is indicative of a relationship that exhibits variation only over large distances or is constant over space. Recent research in different fields ranging from public health to environmental analysis supports the view that different processes may exhibit different degrees of spatial heterogeneity which can be modeled with MGWR (Fotheringham et al., 2019; Cupido et al., 2019; Yang et al., 2019; Oshan et al., 2020).

A recent study by Fotheringham et al. (2021) applies MGWR to examine spatial variations in the determinants of voter preferences in the 2016 US Presidential election. The study also indicates that the influence of spatial context in shaping voter preferences can be quantified and separated from the influences of other socio-economic factors. The results suggest that spatial context is a significant factor in determining voter preference and exhibits interesting regional variations which match our intuition about ‘red’ versus ‘blue’ parts of the country. However, the study examines voter preferences in a single election, and it remains uncertain whether the results are an artefact of the data set used or whether they represent spatial patterns that are relatively stable over time. Assessing the temporal stability of the determinants of voter preferences will help clarify this and will also help in understanding the historical geography of presidential elections and may offer insights into future elections. If the spatial variations in the determinants of voter preferences can be shown to be relatively stable over time, this would allow greater confidence in the conclusion that the spatial trends in the determinants of voter preferences reflect real processes and are unlikely to be artefacts of a single time period.

Building on the discussion and modeling of voter preferences in Fotheringham et al. (2021) which focused solely on the 2016 election, we apply the same model calibrated with MGWR to the four most recent US Presidential elections in 2008, 2012, 2016, and 2020 which took place at times of varying political sentiment and hence which provide a good test of the robustness of the results obtained for the 2016 election.

The 2008 election saw a Democratic landslide victory with Obama beating McCain. The 2012 election resulted in a weakened Obama victory over Romney. The 2016 election resulted in a Republican upset victory with Trump beating Hilary Clinton. and then Trump lost his re-election bid to Biden in the 2020 election. Given the different contexts of each of these four elections, it is of interest to examine to what extent the determinants of voter preferences varied over both space and time. Accordingly, the paper proceeds as follows. In Sections 2 and 3, the data and methods are introduced. Section 4 describes the results of calibrating both a global OLS model and a local MGWR model of the determinants of voter preferences in each of the four elections. The paper concludes in Section 5 with a discussion of the results and implications for future analysis of voter preferences.

2 Data

The county-level election data from 2008-2016 were obtained from the MIT Election lab website[1]. The 2020 election data were accessed from a GitHub public repository[2] where the source data were retrieved from New York Times. The dependent variable in the regression models is defined as the percentage share of the vote that went to the Democratic Party out of the votes that went to either the Democratic Party or the Republican Party in each county. That is, any votes for the third-party candidates are ignored in the calculation of the percentage of Democratic vote. Maps of this two-way percentage of Democratic vote for the contiguous United States for all four elections are shown in Figure 1 with and counties in blue (red) being those where the majority of voters in the county voted for the Democratic (Republican) Party. Over the four time periods, the geographic patterns are very similar. Overall, the national two-way percentage vote for the Democratic Party decreased from 53.7% in 2008, 52.0% in 2012 to 51.1% in 2016, then bounced back to 52.3% in 2020.

[1] The data were retrieved from: https://electionlab.mit.edu
[2] https://github.com/tonmcg/US_County_Level_Election_Results_08-20

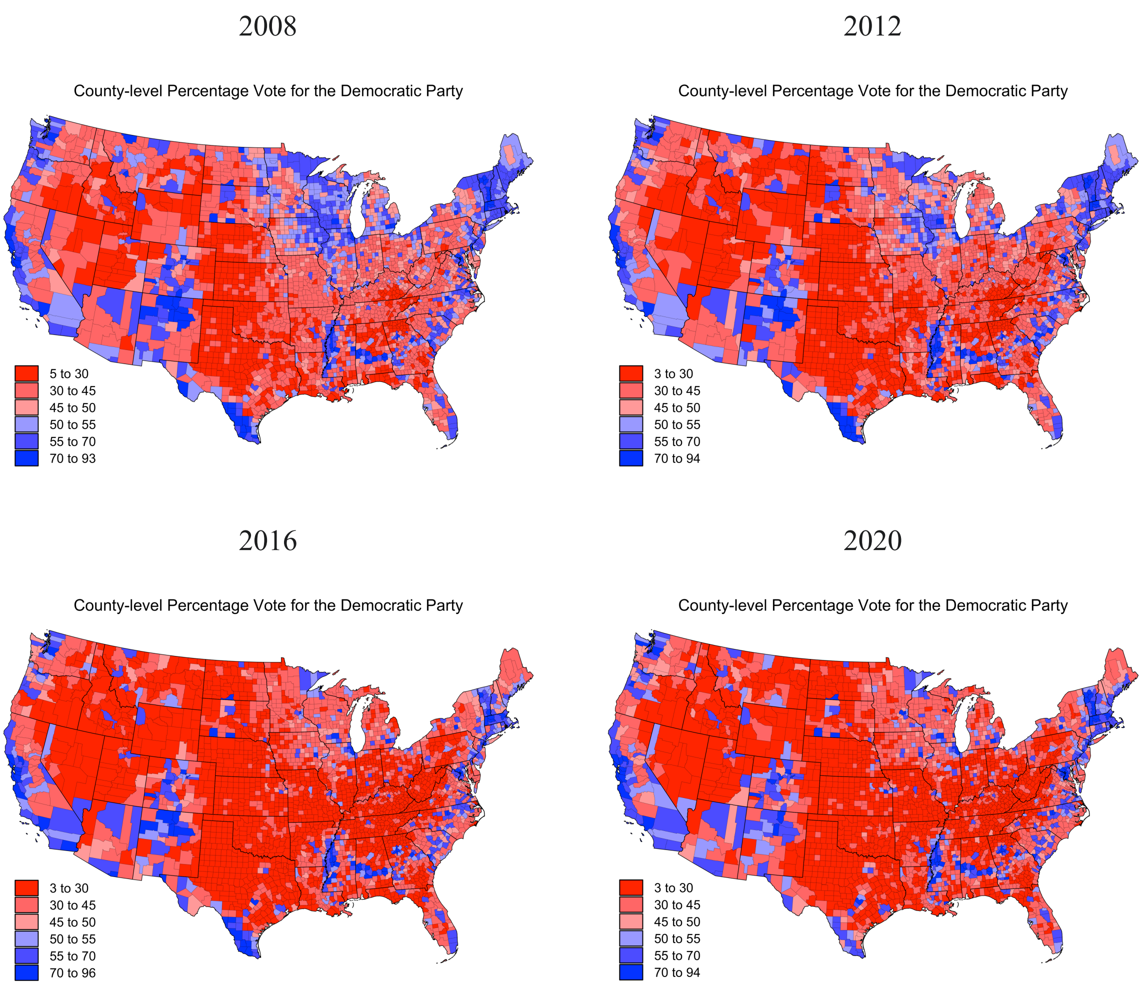

Figure 1. Observed county-level percentage of vote for the Democratic Party in the straight fight against the Republican Party for 2008, 2012, 2016 and 2020.

In order to understand the determinants of the above four distributions of voter preferences, we employ the same model used in Fotheringham et al (2021) who provide a justification for the model form and the set of explanatory variables included in the model. These are based on arguments made in the literature and mainstream media about the role of various socio-demographic and economic variables such as gender, race/ethnicity, age, income and education on voting preferences (Degan and Merlo, 2011; Leighley and Nagler, 2013; Scala et al., 2015; Tyson and Maniam, 2016). Further details can be found in Fotheringham et al. (2021). The independent variables for the model of voter preference are shown in Table 1 and their values for each county were obtained from the United States Census Bureau's American Community Survey (ACS) 5-year estimate datasets (2005-2009[3], 2008-2012, 2012-2016, and 2015-2019[4]) for the corresponding four presidential election years. Descriptive statistics of all the independent variables can be found in Supplementary Material (Table S1 – S4). The expectation was that the county-level percentage vote for the Democratic Party would increase with increasing concentrations of Black, Hispanic, and Foreign-born populations, with increasing concentrations of both inhabitants with a university education and population living in areas of high population density. Conversely, the expectation was that the Democratic vote would decrease as the proportion of elderly inhabitants within a county increased and as median income increased. There were no *a priori* expectations for the direction of the relationships between the Democratic vote and the remaining variables although all of these have been discussed either in the academic literature or the media as possible influences on voter preferences. For instance, an increase in voter turnout could have either a negative or positive impact on the Democratic vote depending on the socio-demographic composition of marginal voters in a county. Further, there was no expectation as to which, if any, of these factors might have a significant spatially varying influence on voter preferences. These are all issues for which there is little or no theoretical guidance, and which have largely been ignored.

To mitigate the sampling bias of modeling percentages based on small populations, counties with fewer than 5,000 population were not included in the analysis. Also, because a local model is calibrated

[3] The earliest available ACS 5-year estimate dataset (2005-2009) is used for the 2008 election.
[4] The latest available ACS 5-year estimate dataset (2015-2019) is used for the 2020 election.

which is based on the concept of 'borrowing' data from nearby locations, the counties in the non-contiguous states of Hawaii and Alaska were also removed from the analysis. Both the dependent and independent variables in the two models were standardized to have mean of zero and variance of one for each election. This allows the regression coefficients to be comparable across variables and over time, and the standardization is recommended in MGWR analysis to produce comparable estimated covariate-specific bandwidths (Fotheringham et al., 2017).

Table 1. The list of independent variables used in this study.

Variables	**Descriptions**
Sex_ratio	Sex ratio (the ratio of males to females)
Pct_age_18_29	Percentage of population aged 18 to 29
Pct_age_65	Percentage of population aged 65 and over
Pct_Black	Percentage of Black or African American alone
Pct_Hispanic	Percentage of Hispanic
Median_income	Median household income
Pct_Bachelor	Percentage of population with bachelor's degree or higher
Gini	Gini Index
Pct_Manuf	Percentage of population employed in manufacturing industry
Ln(pop_den)	Ln of population density (persons per square mile)
Pct_3rd_Party	Percentage of third-party vote
Turnout	Voter turnout
Pct_FB	Percentage of foreign-born population
Pct_Insured	Percentage of population with health insurance coverage

3 Methods

For each of the four election years, two models are calibrated using the data described above: (i) a global OLS model to act as a benchmark against which the local results could be compared; and (ii) a local MGWR model to examine the county-level associations between voting preferences and socio-demographic and socio-economic factors.[5] The OLS model is specified as:

$$y_{it} = \beta_{0t} + \sum_j \beta_{jt} x_{ijt} + \varepsilon_{it}$$

where y_{it} is the percentage of the Democratic vote for county *i* at year $t \in \{2008, 2012, 2016, 2020\}$, β_{jt} is the parameter for the j^{th} covariate x_{ijt} for county *i* at year *t*, and ε_{it} is the random error for county *i* at year *t*. Because the variables are all standardized, the intercept β_{0t} is zero. The MGWR model by comparison has location-specific parameters and is formulated as

$$y_{it} = \beta_{bw0t}(u_i, v_i) + \sum_j \beta_{bwjt}(u_i, v_i) x_{ijt} + \varepsilon_{it}$$

where y_{it} is the dependent variable for county *i* at year *t*, (u_i, v_i) are the centroid coordinates of the county *i*, $\beta_{bw0t}(u_i, v_i)$ is the local intercept for county *i* at year *t*, $\beta_{bwjt}(u_i, v_i)$ is the parameter for the j^{th} independent variable x_{ijt} for county *i* at year *t*, and ε_{it} is the random error for county *i* at year *t*. The covariate-specific bandwidths are noted as $bw_0, bw_1, \dots, bw_j$. Each of these controls the amount of data 'borrowed' from around location *i* in order to compute a set of local parameter estimates associated with a specific covariate (more details of this are given in Fotheringham et al 2017; Li et al. 2020). Larger optimized bandwidths denote relationships which are relatively homogeneous; smaller values denote relationships which are relatively heterogeneous.

In the MGWR calibration local parameters are estimated along with the covariate-specific bandwidths, $bw_0, bw_1, \dots, bw_j$, and this involves a back-fitting procedure following the Generalized

[5] A spatial error model was also calibrated as a comparator for the MGWR results, but the results were highly dependent on the arbitrary definition of the spatial weights matrix used to calculate residual covariance.

Additive Model (GAM) framework (Hastie and Tibshirani, 2002; Fotheringham, et al., 2017). Initial estimates of the local parameters were obtained from calibrating the model by GWR and then the back-fitting algorithm proceeds from the first to the last covariate by calibrating a univariate GWR against the partial additive components plus the current residuals as the dependent variable. In each univariate GWR, a covariate-specific optimal bandwidth is estimated. The back-fitting stops when the relative change in the parameter estimates is smaller than 1×10^{-5}, at which point a vector of covariate-specific bandwidths and a matrix of local parameter estimates are obtained from the last iteration of the back-fitting. The calculation of the covariate-specific bandwidths is based on optimizing a statistical criterion that includes a trade-off between parameter bias and parameter uncertainty, such as a corrected Akaike Information Criteria (AICc) (Yu et al. 2020). Beyond the bandwidth, adding further data to the local regression adds more bias to the local parameter estimates than it reduces the uncertainty of those estimates. Bias in local parameter estimates occurs because data not from location *i* are being used to estimate relationships at *i* and data from locations further away will produce more bias than data from locations in closer proximity to *i* if the relationships are spatially nonstationary.

Inference in MGWR is computed following the details in Yu et al. (2019) to obtain the standard errors of the parameter estimates and other model diagnostics such as model's corrected AICc. Confidence intervals (CI) around each optimized bandwidth are also calculated using the Akaike weights-based procedure introduced in Li et al. (2020). The CIs for the bandwidths account for uncertainties in the model and make it possible to make statements about the relative magnitudes of different bandwidths, and hence different levels of spatial heterogeneity, that might exist across processes. In the MGWR calibration, an adaptive bi-square kernel is employed which uses the number of nearest neighbors as the unit of the bandwidth, which means that the spatial weight equals one at each regression location and decays to zero at the bandwidth (Fotheringham et al., 2002). Data from counties lying beyond the bandwidth have zero weight. The OLS and MGWR models are calibrated by the MGWR 2.2 software[6] which was developed by the Spatial Research Analysis Center at Arizona State University (Oshan et al., 2019; Li and Fotheringham 2020). Additionally, multicollinearity among independent variables was examined using Variance Inflation Factor (VIF), and all VIFs are below 6 for each election model (see Table S5 in Supplementary Material for details).

4 Temporal and spatial stability of the determinants of voter preference

4.1 Global model results

The results of calibrating the OLS model are shown in Table 2 with the global parameter estimates and their associated standard errors reported for the relationships between Democratic vote percentage and various socio-economic variables for 2008, 2012, 2016 and 2020, respectively. The model performs relatively better for 2016 and 2020 with an R^2 values of 0.66 and 0.68 respectively and performs only modestly for 2008 and 2012 with R^2 values of 0.40 and 0.48 respectively. The residuals from these four calibrations are shown in the left column of Figure 2 and they display significant ($p < 0.01$) degrees of positive spatial autocorrelation based on the global Moran's I test[7] as shown in Table 3. This raises alarms because the spatially correlated residuals violate the assumption in OLS that they should be independent and identically distributed (i.i.d). As a result, OLS regression coefficients may be biased and the type I error rates inflated (Anselin and Bera, 1998, Dormann, 2007; Kuhn, 2007). To partially remedy this, we set a high significance threshold of 0.001 in Table 2. Since the variables are standardized, the estimates of the intercept are zero in each year and the parameter estimates (standardized coefficients) shown in the table can be interpreted as the relative importance of the associated variable in influencing preference for the Democratic Party. Race/ethnicity (percentage of Black and Hispanic) and education (percentage of people with a Bachelor's degree) have the strongest positive associations with voting Democratic in all four

[6] This software is freely available at https://sgsup.asu.edu/sparc/mgwr.
[7] Moran's I statistics are computed using *spdep* package in R.

elections and median household income has the strongest negative association. Other consistently positive associations with voting Democrat are population density, percentage of third-party voting, and turnout. The remaining variables are either insignificant (e.g. percentage of people working in the manufacturing sector and percentage of young population) in all four years or only marginally significant (Gini, gender, and percentage of elderly) in some years and insignificant in other years.

By examining the magnitudes of the standardized coefficients across four years, we can see the shifting relative importance of each factor on voting preferences from 2008 to 2020. For example, the impact of the percentage of black population within each county on voting for the Democratic Party increased over time with coefficients of 0.36 in 2008, 0.41 in 2012, and 0.53 in both 2016 and 2020. The impact of education in separating Democrats from Republicans rose largely in the recent two elections from the previous two elections (0.29 in 2008 and 0.25 in 2012 vs. 0.43 in 2016 and 0.46 in 2020), with more highly educated voters tending to side with the Democratic Party. Trump's deliberate appeal to less well-educated voters clearly worked but turned off large number of better educated voters. The impact of income on voter preference declined from 2008 to 2020 (-0.48 in 2008, -0.46 in 2012, -0.31 in 2016 and -0.22 in 2020) indicating that the traditional appeal of the Republican Party to higher-income voters and the Democratic Party to lower-income voters appears to be narrowing, again due to Trump's populist appeal in 2016 and 2020. Similarly, the differential between urban and rural counties in voter preference appears to be declining over time and although a higher turnout favored the Democratic Party in all four elections. The impacts of other factors remained relatively stable over the four elections.

Table 2. OLS regression coefficients for the 2008, 2012, 2016 and 2020 elections.

Variables	2008	2012	2016	2020
Intercept	0.00 (0.02)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)
Sex_Ratio	0.03 (0.02)	0.04 (0.02)	0.01 (0.01)	0.01 (0.01)

Pct_Black	0.36* (0.02)	0.41* (0.02)	0.53* (0.02)	0.53* (0.01)
Pct_Hispanic	0.27* (0.03)	0.28* (0.02)	0.28* (0.02)	0.26* (0.02)
Pct_Bachelor	0.29* (0.03)	0.25* (0.03)	0.43* (0.03)	0.46* (0.03)
Median_Income	-0.48* (0.04)	-0.46* (0.03)	-0.31* (0.03)	-0.22* (0.03)
Pct_Age_65	-0.06 (0.03)	-0.06 (0.02)	-0.01 (0.02)	0.06* (0.02)
Pct_Age_18_29	-0.04 (0.03)	-0.02 (0.02)	-0.02 (0.02)	0.03 (0.02)
Gini	-0.06 (0.02)	-0.04 (0.02)	0.03 (0.02)	0.02 (0.02)
Pct_Manuf	-0.01 (0.02)	-0.01 (0.02)	0.02 (0.01)	0.00 (0.01)
Ln(pop_den)	0.28* (0.02)	0.22* (0.02)	0.17* (0.02)	0.17* (0.02)
Pct_3rd_Party	0.19* (0.02)	0.23* (0.02)	0.16* (0.01)	0.18* (0.01)
Turnout	0.29* (0.02)	0.34* (0.02)	0.17* (0.02)	0.14* (0.02)
Pct_FB	0.18* (0.03)	0.20* (0.03)	0.19* (0.02)	0.11* (0.02)
Pct_Insured	0.27* (0.02)	0.23* (0.02)	0.17* (0.02)	0.16* (0.01)
N	2810	2814	2812	2807
R^2	0.40	0.48	0.66	0.68

Significance code * $p < 0.001$

However, these results assume that the relationships being modeled are constant over space – for example, a 10% increase in Hispanic population would lead to approximately a 2.8% increase in the vote for the Democratic Party in each county. Given the diversity of the Hispanic community, this is likely to be grossly misleading – many Hispanics from Cuba and Venezuela, for example, favor the Republican Party. Where the relationships being modeled exhibit spatial heterogeneity, the results from the global OLS model (or any other global model formulation) will simply be averages of potentially interesting spatial diversity. We now explore potential spatial heterogeneity in the determinants of voter preferences.

4.2 Local Model Results

A local model of voter preferences employing the same set of covariates as described above was calibrated by MGWR for the four Presidential elections. The diagnostics for each calibration are shown in Table 3 and include a comparison with the OLS results. As expected, the MGWR models outperform the OLS models for all four years with R^2 values exceeding 0.90. The superiority of the MGWR model results is better identified by their much smaller AICc values as this goodness-of-fit measure takes into account the increased complexity of the MGWR formulation. Maps of residuals from the MGWR calibrations are compared with these from the OLS models in Figure 2. By using the same color scheme, it can be seen that the magnitude of the residuals is generally much smaller for the four MGWR calibrations and the MGWR residuals have random patterns for all four elections, as shown by the values of Moran's I in Table 3. These summary results strongly support the idea that there is some spatial non-stationarity in the processes affecting voter preference and these are captured by the MGWR model, but not by the OLS model. We explore this non-stationarity in further detail below. The comparison of the residuals also suggests that the non-stationarity in relationships which is not captured in the OLS model exhibits itself through strongly dependent residuals.

Table 3. Comparison of model performance between OLS and MGWR.

	2008		**2012**		**2016**		**2020**	
	OLS	MGWR	OLS	MGWR	OLS	MGWR	OLS	MGWR
R2	0.40	0.91	0.48	0.92	0.66	0.95	0.68	0.95
AICc	6578	2448	6188	2139	5008	1045	4762	924
Moran's I	0.19*	0.05	0.26*	0.11	0.23*	0.04	0.27*	0.002
Spatial Pattern	Clustered	Random	Clustered	Random	Clustered	Random	Clustered	Random

Significance code * $p < 0.01$

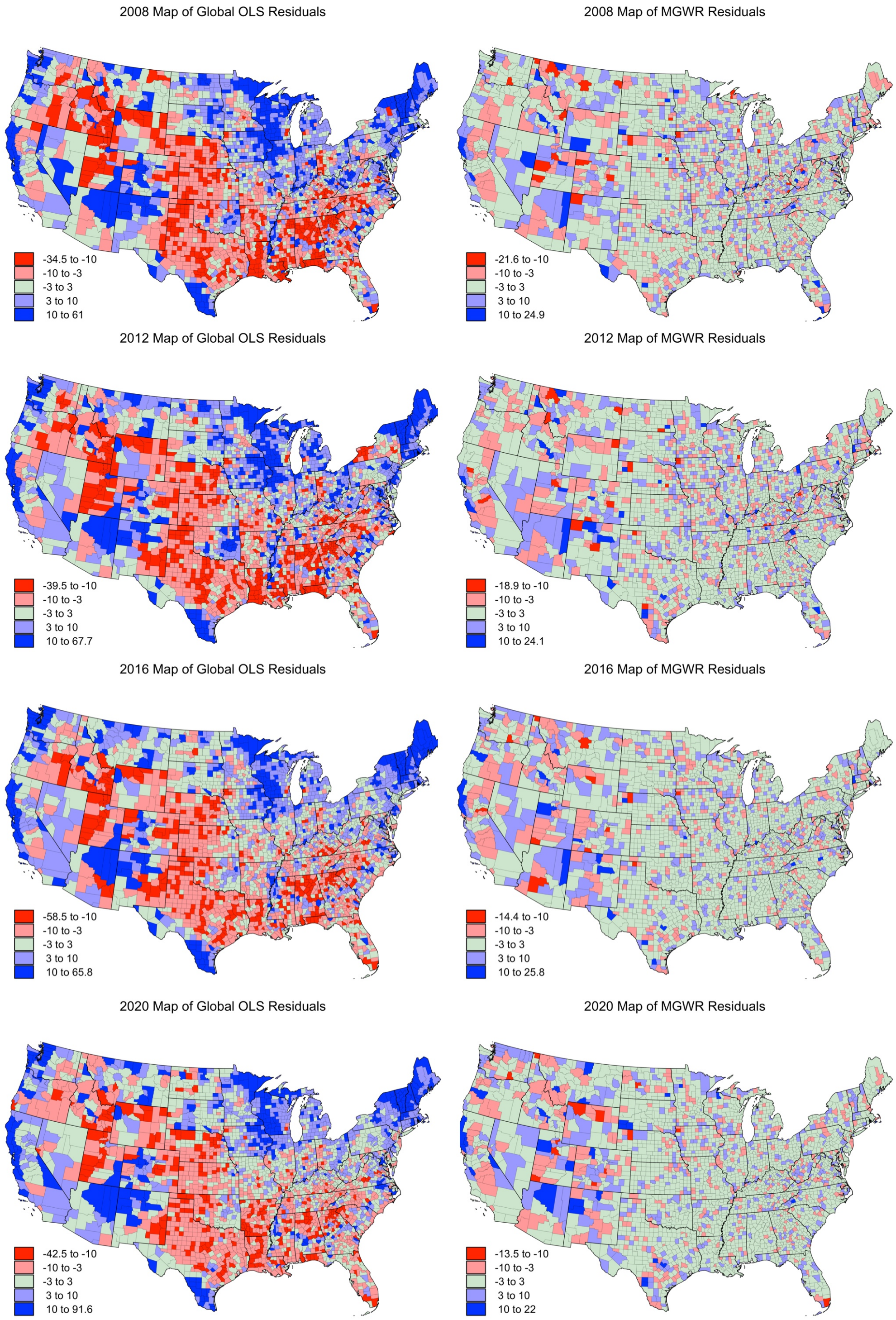

Figure 2. Maps of OLS residuals (left) and MGWR residuals (right) for 2008, 2012, 2016 and 2020.

As well as being a more accurate model of the determinants of voter preferences, MGWR also provides optimized covariate-specific bandwidths which provide comparable measures of the spatial scales over which the relationships being modeled vary. The optimized bandwidths in MGWR represent the number of weighted data points used in each local regression. Adding data from locations beyond the bandwidth increases the bias in the parameter estimates more than it reduces the uncertainty about their values. If a process is relatively stable over space, the bandwidth will contain a large proportion of the data points in the study area (in the extreme case of a global model where the processes are all constant, the bandwidth will tend to infinity) because adding data points to the local regressions does not induce bias but will increase uncertainty in the parameter estimates. Conversely, if a process varies relatively rapidly over space, the optimized bandwidth will be a small proportion of the data points in the study area because adding data from locations that are not in close proximity will increase the bias more than it reduces uncertainty. Consequently, the optimized bandwidth for each covariate indicates the spatial scale over which the conditioned process linking change in independent variable to change in the outcome is relatively stable. More details on the computation of bandwidths can be found in Li et al. (2020).

Table 4 summarizes the optimal bandwidths and their 95% confidence intervals (CI) in terms of the number of nearest counties from which data are borrowed in the estimation of each set of local parameters. The CIs are computed based on the methodology developed by Li et al. (2020), which uses Akaike weights to quantify the amount of bandwidth selection uncertainty taking account of the sampling variation of the data. The covariate-specific bandwidths and CIs are useful to determine if there are any significant differences in the spatial scales over which different processes operate and if there is any significant change in the spatial heterogeneity exhibited by the parameter estimates over time. If two bandwidth CIs do not overlap, this indicates that the difference in the covariate-specific bandwidths is statistically significant. For example, in 2016, the bandwidth for the local parameter estimates associated with median income is significantly larger than the bandwidth for the percentage of black population (the CIs are 2158, 2717 and 43, 45, respectively. This suggests that the influence of median income on voter preferences is much more spatially homogeneous than is the influence of the percentage of black population. A similar conclusion is drawn from comparing the CIs of the bandwidths associated with these two covariates in the 2012 and 2008 elections. Other bandwidths which are very stable over the four elections are those associated with the intercept, education level, income disparity, and population density; the remaining bandwidths exhibit significant variation across the elections.

Table 4. Optimized covariate-specific bandwidth and 95% confidence interval in MGWR.

Variables	2008	2012	2016	2020
Intercept	43 (43, 44)	43 (43, 44)	43 (43, 45)	43 (43,44)
Sex_Ratio	2807 (2156, 2808)	2813 (2409, 2813)	603 (446, 850)	2740 (2154, 2748)
Pct_Black	43 (43, 45)	43 (43, 45)	43 (43, 45)	43 (43, 45)
Pct_Hispanic	43 (43, 47)	1314 (1101, 1505)	543 (446, 601)	870 (870, 1098)
Pct_Bachelor	195 (174, 233)	210 (151, 197)	208 (174, 210)	490 (445, 490)
Median_Income	2368 (1753, 2560)	2090 (1755, 2409)	2659 (2158, 2717)	2715 (2154, 2717)
Pct_Age_65	85 (78, 101)	101 (78, 106)	656 (600, 850)	43 (43, 45)
Pct_Age_18_29	460 (387, 542)	494 (387, 542)	58 (51, 65)	168 (160, 196)
Gini	1363 (1099, 1753)	1246 (946, 1350)	763 (696, 1100)	563 (540, 695)
Pct_Manuf	2809 (2156, 2809)	424 (387, 542)	2809 (2158, 2810)	295 (291, 295)
Ln(pop_den)	410 (351, 542)	383 (351, 447)	387 (292, 446)	298 (291, 300)
Pct_3rd_Party	312 (292, 446)	318 (292, 388)	160 (137, 173)	2746 (2154, 2748)
Turnout	2765 (2156, 2787)	415 (387, 542)	117 (110, 137)	182 (173, 196)
Pct_FB	2809 (2406, 2809)	1448 (1101, 1755)	1424 (1100, 1754)	2518 (2154, 2558)
Pct_Insured	87 (78,101)	43 (43, 47)	43 (43, 45)	101 (101, 101)

For each covariate, we can map the local parameter estimates obtained from MGWR for each of the four elections to examine the stability of the spatially varying effects of each covariate on voter preference. There are 15 sets of local parameter estimates maps as shown in Figure 3 and Figure 4, and we present them in terms of the magnitude of their bandwidth from processes that are global (with large bandwidths) to processes that are highly local scale (with small bandwidths). Counties with significant positive or negative local parameter estimates are colored in blue and red accordingly, indicating that an increase in a given variable would increase the share of the vote for the Democratic (blue) or Republican (red) Party in that county. The same color scale is employed for all the parameter estimates maps, and the parameter estimates with higher absolute magnitudes have a darker color, indicating they have relatively more influence over voter preference. Local parameter estimates for counties with fewer than 5,000 population, which are not included in the model, are interpolated using the Inverse Distance Weighting (IDW) technique with a power of 2. Counties with insignificant parameter estimates (at the 95% level) are shaded in grey, and the significance has been adjusted to account for multiple hypothesis testing (da Silva and Fotheringham, 2016; Yu et al., 2019).

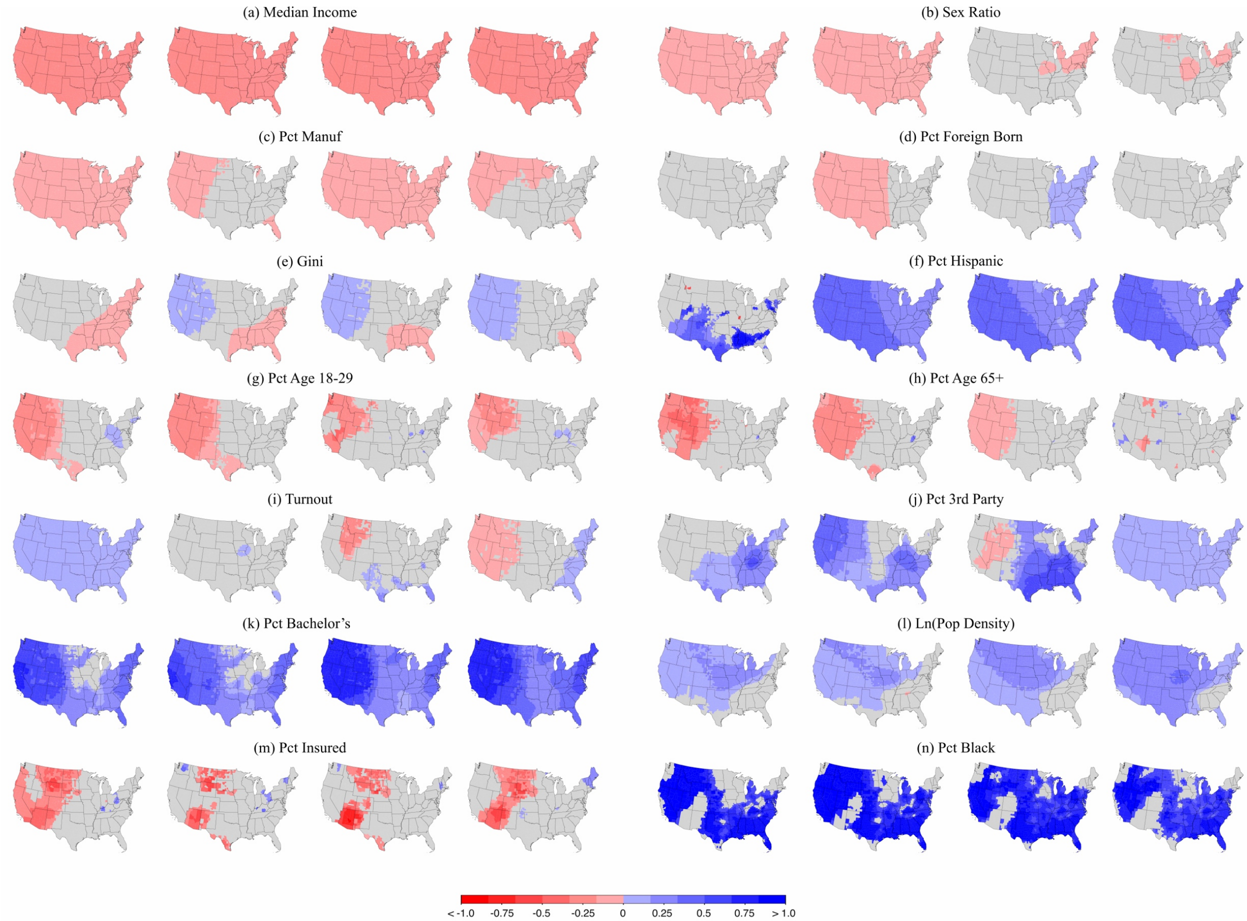

Figure 3. Standardized local parameter estimates maps obtained from MGWR. Counties with significant positive or negative local parameter estimates are colored in blue and red. Counties with insignificant parameter estimates are shaded in grey.

Figure 3 (a) depicts the local parameter estimates for household median income. The MGWR local parameter estimates are aligned with their global counterparts: in all four elections, median income is significantly negatively associated with voter preference for the Democratic Party for all counties with bandwidths in each case in excess of 2,000 nearest neighbors indicating global relationships in all four elections. That is, when median income within a county increases, the proportion of votes cast for the Republican Party increases, with virtually the same degree of sensitivity for all the counties in the United States. Across the four elections, the influence of income on voter preference remained relatively constant.

The local parameter estimates for sex ratio (males to females) are shown in Figure 3(b). In 2008 and 2012, sex ratio has a global bandwidth indicating a national-level significant negative association with voting for the Democratic Party. That is, when there were more males than females in a county, voter share for the Democratic Party decreased. However, in the 2016 and 2020 elections the influence of gender was only significantly negative in part of the New England and Midwest. In the remainder of the country, gender ratios appeared to have little or no influence on voting. The MGWR results in 2008 and 2012 contradict those of the global model where the global relationship is insignificant. In 2016 and 2020, there is alignment in general between the global and local results with both generally being insignificant, but the global result clearly hides the locally significantly negative association between males and the Democratic vote in the north-east and parts of the mid-west.

The local parameter estimates for the percentage of the population employed in manufacturing are depicted in Figure 3(c). Across the four elections, the local parameter estimates are consistently negative, indicating that as the proportion of manufacturing employment in a county increases, the share of the Republican vote increases. This relationship was virtually constant across the US in 2008 and 2016 but in 2012 it was only found in the western states and in Florida. One explanation for the exceptional spatial pattern in 2012 and 2020 is that the counties with insignificant parameter estimates are largely in areas where manufacturing experienced a great deal of distress due to both of the 2008 economic crisis and the 2020 COVID-19 pandemic. These findings question the widely held view in the media that manufacturing workers tend to vote Democrat: when accounting for the influence of other covariates and taking a local view, this does not appear to be the case. While states with relatively large manufacturing bases do tend to vote Democrat, the vote for the Democratic Party appears to be linked to factors other than employment in manufacturing.

Figure 3(d) shows the local parameter estimates for the percentage of the foreign-born population. The effect of this variable on voting preference is temporally unstable. In 2008 and 2020, the proportion of foreign-born population within a county was not a significant predictor of voting preference in any part of the country. In 2012 and 2016 the spatial patterns are actually quite similar with counties in the western half of the country exhibiting a negative association with voting for the Democratic Party and counties in the eastern half exhibiting a positive association. The distributions appear to be different only because the local parameter estimates in the west were significantly different from zero in 2012 while the estimates in the east were significant in 2016. It is possible that the countries of origin are somewhat different for foreign-born residents of counties in the east, a greater proportion of whom might be from Europe where many countries have political leanings that align more with the Democratic Party than with the Republican Party.

Across the four elections, income disparities have a positive association with Democratic vote support in the West (except for 2008) and a negative association in the Southeast (Figure 3e). One possible explanation for this spatially contradictory relationship lies in the issue of voter registration laws within various states: states in the south-east have typically made it more difficult for minorities to register to vote and a consequence of this is that in counties with large income disparities, the portion of the population with the lower incomes, who would tend to vote Democrat, would be under-represented in the list of eligible voters. Consequently, counties in such states with high income disparities would have a relatively higher vote for the Republican Party. Temporally, it is observable that the influence of income disparities has been expanding in the West but shrinking in the Southeast.

Figure 3(f) shows the local parameter estimates for the percentage of Hispanic population in each county. In the last three elections, the impact of the Hispanic population on voting preference has been strong and consistent across the country with significant positive association with votes for the Democratic Party being found in every county. The association is, however, stronger in the south and west where significant and long-term Hispanic populations can be found. Interestingly, in 2008, the relationship between the Hispanic population and Democratic vote was much more localized with a small bandwidth (43). Significant positive associations between Hispanic population percentage and share of votes for the Democratic Party were only found in the southwestern states, including Arizona, New Mexico, Texas, which had large proportions of Hispanics as well as in counties in New York and New Jersey. There was also a significant positive relationship in some counties in the Deep South from Louisiana to South Carolina.

Figure 3(g) and Figure 3(h) show the local parameter estimates for the percentage of young (age 18-29) and elderly (age 65+) population, respectively. Everything else being equal, higher percentages of younger-age voters in western counties are associated with lower vote percentages for the Democratic Party. There is no significant impact of younger-age voters in much of the rest of the country except for a couple of interesting outliers in 2016 and 2020 around Indiana University and Ohio State University where there is a significant positive relationship between younger voters and votes for the Democratic Party. This effect can also be seen in 2008 on a broader scale. It is interesting to note that the finding of a significant negative relationship between younger voters and votes for the Democratic Party in many western counties is at odds with the media reporting of younger votes favoring the Democratic Party but here we have conditioned for many other factors such as education, income and ethnicity whereas many polls used by the media report largely unconditioned figures. It is further interesting to note the decline in the counties where a significant negative relationship was reported in 2016 and 2020 compared to 2012 and 2008. The relationship between elderly voters and the Democratic share of votes is consistent across 2008-2016 elections and only exhibits a significant negative relationship for the western states. However, in 2020, there are local behaviors in some rural counties. There is no significant relationship in the rest of the country.

Figure 3(i) shows the association of the turnout rate (the proportion of those eligible to vote who actually voted) with people's voting preferences. In 2008, the Obama bandwagon effect meant that across the country, high turnout rates were associated with higher percentage votes for the Democratic Party. This effect was uniform as shown by the very large bandwidth. In 2012 enthusiasm for Obama had diminished and turnout had very little effect on voting preferences anywhere in the country, the only exceptions being in southern Florida and a few counties on the Illinois/Iowa/Missouri border where higher turnouts were beneficial to the Democratic Party. In 2016, the influence of voter turnout again changed with no significant impact being felt in the vast majority of counties but increasing voter turnout favored the Republican Party in counties in Montana, Idaho, and Utah while it favored the Democratic Party in Florida, Louisiana, and south Texas. In 2020, the east and west coasts showed an opposite relationship between voter turnout and the for the Democratic Party. That relationship is weak but significantly positive in western counties while negative in eastern counties.

Third-party voting is an interesting aspect of American elections. Essentially the vote is split between just two parties, the Democratic Party and the Republican Party. However, in some elections, a third-party candidate appears on the ballot who generates what can be a significant share of the votes in some counties. The spatial pattern of voting for third-party candidates is expected to be very election specific as third-party candidates are generally 'one-offs'. The local parameter estimates for the percentage of third-party votes cast in each county are shown in Figure 3(j). The covariate-specific bandwidth ranges approximately from 150 to 300, indicating a local-to-regional impact on voting preferences. In 2008, there is a significant positive association between the third-party vote and Democratic vote in the Northeast and the Southeast of the US. Most third-party votes were for either the independent candidate Ralph Nadar or the Libertarian candidate Bob Barr who received approximately 0.5% and 0.4% of the national vote, respectively. Interestingly, in 2008 nowhere did third-party voting favor the Republican Party. In 2012, positive associations between third-party voting and votes for the Democratic Party occurred in the West and the Southeast where the primary third-party candidate was Gary Johnson of the Libertarian Party who received 1% of the total national vote. In 2016, the country divided along third-party preferences with counties in New England, the Southeast, and much of the central US having a significant positive association between the third-party vote and the Democratic vote, while in the counties within and surrounding Utah, there was a significant negative association. Here, the third-party candidate was Evan McMullin, a conservative independent, and it would appear that he drew votes away from the Democratic Party due to strategic voting: McMullin had a good chance of defeating Donald Trump in Utah (McMullin received 21.5% of the votes) so that some democrats chose to vote for him rather than their own candidate who had little chance of winning the state. In 2020, the influence from the third party shows a positive but weak pattern across the nation. The best performing third party candidates Jo Jorgensen (Libertarian) and Howie Hawkins (Green) received much lower number of votes compared to their predecessors in 2016 and did not gain much of attention in the 2020 election national-wide.

Increasing levels of education have a strong significantly positive influence on the Democratic vote across the nation, as shown in Figure 3(k). This is a consistent result across all four elections but the ‘education divide’ in voter preferences appears to be getting stronger over time, a finding that is consistent with the political science literature (*inter alia,* Sides et al., 2017; Schaffner et al., 2017) and the mainstream media (*inter alia,* Silver, 2016; McGill, 2016; Tyson and Maniam, 2016). Although the education variable is always positively associated with votes for the Democratic Party, the relationship varies in intensity across the country being stronger in the Northeast, the West, and in the state of Florida but the increase in the educational divide in voting has taken place primarily in the West and Midwest.

Figure 3(l) shows the local parameter estimates of log of the population density, a variable acting as a measure of the rural/urban divide. The association between population density and voting Democratic exhibits a broad and consistent spatial variation with urban voters in the north and west exhibiting a significant preference for the Democratic Party compared to their rural counterparts, *ceteris paribus,* but with no significant difference in voter preference between urban and rural areas in the east and south-east of the country. Local parameter estimates related to the effect of having health insurance on voter preference are shown in Figure 3(m). In all four elections the relationship exhibits strong spatial heterogeneity and similar spatial patterns. Voters in Arizona and Montana who had health insurance appeared to have a strong preference for the Republican Party and those without insurance had a strong preference for the Democratic Party, *ceteris paribus.* The trend was more widespread in the 2008 election but was still confined to counties in the west. In all recent elections, ‘Obamacare’ was a significant issue separating the Democratic and Republican parties and it is surprising that the impact of having/not having health insurance on voter preference was so limited.

Figure 3(n) shows the impact of the percentage of African Americans within each county on voter preferences. Across the nation, there is a strong and positive relationship between the percentage of African Americans and the Democratic vote share. However, the intensity of this relationship varies spatially as indicated by the small bandwidths associated with this covariate. The only counties where the percentage of African Americans does not have a significant impact on voter preferences are those where the proportion of African Americans is very low such as counties in Washington, Arizona, New Mexico, Colorado, Wisconsin, and New England. Again, the results are consistent across all four elections.

As well as producing localized estimates of the parameters associated with each covariate in the model, calibration by MGWR also generates local estimates of the intercept which are shown in Figure 4. Given that all the variables in the model are standardized with mean zero and variance one, the local intercept measures the percentage of the vote that would be gained by the Democratic Party in each county *if that county had the average share of each covariate.* That is, the local intercept measures the intrinsic level of voter preference for the Democratic Party independent of all other contributing factors in the model; in essence it measures the degree to which a person's preference for one political party over another is due to where they live. Hence, the local intercept is a quantitative measure of context. From Figure 4, the spatial patterns of local intercepts show a strong degree of consistency across the four elections, suggesting that the role of spatial context in affecting voting preference has remained relatively stable over the four election periods. In all four elections, the West, Central North and the Northeast of the US have intrinsically blue counties (counties with an intrinsic leaning towards the Democratic Party) whereas the Southern US is largely comprised of red counties (counties with an intrinsic leaning towards the Republican Party). Note that the maps in Figure 4 do not depict *how counties voted* but show *how they would have voted* if they had populations with average demographic characteristics. Counties with insignificant local intercepts do not show any intrinsic leaning towards one Party although they are not necessarily neutral in that they may well have political leanings towards based on their population composition. That said, in this 'band of intrinsic neutrality', from Arizona in the West to Virginia in the East, there are many counties in which the majority vote in a presidential election is by no means certain. It is interesting to see, given that the state is often pivotal in elections, that southern and central Florida generally lies in this zone of intrinsic neutrality. Despite the overall consistent spatial pattern of geographical context in voter preference across the US, there are several local changes across the four elections. California and Oregon, for example, which are intrinsically blue states in 2008 and 2016 became intrinsically neutral in 2012. Utah had a surge in intrinsic Republicanism in 2012 with the presence of 'favorite son' Mitt Romney on the ballot, and there appears to be a contraction of the intrinsically Republican support in Arizona and Colorado balanced by a reduction in intrinsic Democrat support in parts of Illinois and Indiana. Southern Florida remains neutral from 2008 to 2016 but appears red in 2020 as the southern counties shifted further right, especially for Miami-Dade County where Trump narrowed the margin by more than 20 points.

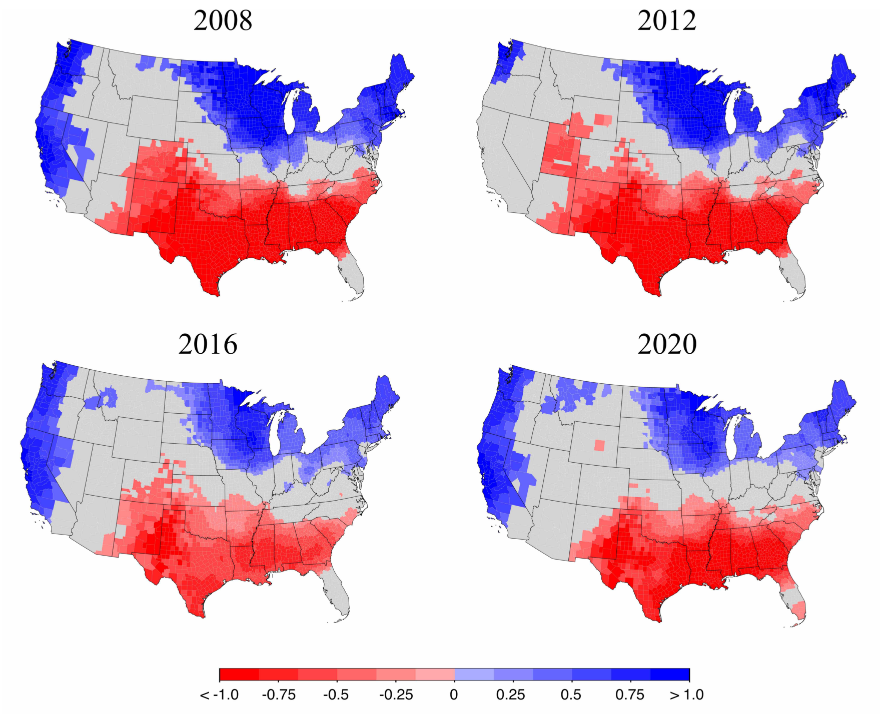

Figure 4. Local parameter estimates for the local intercept obtained in MGWR.

5 Summary

In this paper, we apply a recently developed local modeling technique, multiscale geographically weighted regression (MGWR), to examine the spatially and temporally varying nature of determinants of voter preferences in the 2008, 2012, 2016 and 2020 U.S. Presidential elections. The particular interest of this study is to investigate the spatiotemporal changes in the factors in determining voters' preferences. The results suggest that there are important spatial variations in how socioeconomic factors affect voter preferences and that ignoring such spatial variations will yield inadequate understanding of how people in different parts of the country make political decisions. Reassuringly, the spatial heterogeneity in how various factors affect voting behavior generally remains temporally stable although in some instances local and regional changes over time occur which yield interesting and novel insights into the nature of what determines voters' preferences. This can clearly be seen above in the results for the impact of third-party voting and turnout.

We are also able to report on the spatial scale over which variations in the determinants of voter preference occur. The results from the optimized bandwidths and their confidence intervals indicate that while the degree of spatial heterogeneity varies across processes, the operational scales of the socio-economic processes affecting voter preference are generally stable over the four elections. For example, while the influence of median household income on voter preference has been consistently global in the last four elections, there are large spatial variations in the strength of conditioned associations between the percentage of African-American population in each county and the share of the votes for the Democratic Party, which again though are reassuringly consistent over time. Perhaps most importantly, the spatial patterns of the local intercept, an indicator of the role of local context on voter preference, exhibit a consistent pattern of intrinsic preferences for both the Democratic and Republican parties across the country with the South being identified as intrinsically Republican and the west coast, North-east and parts of the upper Mid-west being intrinsically Democrat. Although it might be argued that these patterns are not unexpected, for the first time it has been possible to quantify the degree to which there are intrinsic preferences for either party and the geographical extent to which these preferences have a significant impact on the outcome of elections. It is important to recognize that the measures of local context are independent of the socioeconomic composition of counties and are normally hidden from sight. These measures indicate not how a county voted but how it would vote if its population had average demographic characteristics. The existence of significant contextual effects on voting suggests that even if all the counties in the country had the same socioeconomic composition, voting results would not be the same and nor would there be a random pattern of Blue vs Red counties. Rather, there would be a swathe of red counties across the southern USA and a wall of blue counties in the North-east, in parts of the Mid-west and along the west coast. Models of voter preference that do not explicitly measure this important contextual effect will be severely misspecified and will produce misleading estimates of the influence of various socioeconomic factors.

Although we are confident of the robustness of the major findings of this study, particularly in light of the stability of results over four elections, several criticisms can be leveled at it. Because of the limited temporal availability of the county-level American Community Survey datasets, we have only been able to apply the analysis to the four most recent Presidential elections. It would be interesting to look at the stability or variability of how socioeconomic processes and spatial context affected voting preferences over a much more extended period by utilizing the decennial census data. An even longer analysis of the dynamics of the determinants of voter preferences would identify which of the socioeconomic variables have had a relatively consistent influence on voter preference and which have varied significantly over time. However, to optimally match each election, which happen every four years, to the decennial census would need further examination.

It could be further argued that using county-level data might miss important spatial heterogeneity in the processes affecting voter preference within counties. This may be particularly relevant to the identification of contextual effects – is it possible with finer-scale data to detect even more local variations in the effect of context on voter preference? Unfortunately, there are several problems with undertaking a finer-scaled analysis of US voting behavior. Corresponding data on voting behavior and on socio-economic characteristics are not easily obtained by precincts and even if they were, there would be an issue with the increased volatility of percentage figures derived for small population areas. It has also been claimed that counties arguably provide the most appropriate geographic setting in which people are exposed to different social and physical influences that affect their voting preferences (Agnew and Shin, 2020). What is without much doubt is that using a local modeling approach with county level data is superior to using standard regions or states and is far superior to using a global modeling approach which ignores the possible presence of spatially nonstationary processes.

In summary, this study strongly suggests that the factors that affect a person's voting behavior may not be constant over space and therefore global models of voting behavior will be inappropriate in modelling such behavior. Local statistical models such as MGWR provide not just information on how the determinants of voter preferences vary over space but also on the spatial scale of this variation. Some determinants may have a global impact on voting behavior while others have very locally varying impacts. Without calibrating a local model of voting behavior such information is hidden behind the 'average' estimates of behavior that are produced by global models. In addition, MGWR, through its locally varying intercept, quantifies the degree to which voter preferences are determined by local context which is independent of the socio-economic determinants of such preferences. This is an extremely important result because not only does it indicate that where a person grows up is a major factor in how that person will vote, irrespective of their socioeconomic background, it validates and quantifies place-based geography. That a model calibrated in one location does not work well in another location should not be a surprise or a concern – the processes being modeled may well vary over space so that a single model will be unrepresentative of these dynamics. Models which do not explicitly account for local contextual effects and spatially varying behavior will producing misleading estimates of the parameters associated with other determinants of voting behavior.

References

Agnew, J. (1996). Mapping politics: how context counts in electoral geography. *Political geography*, *15*(2), 129-146.

Agnew, J. (2014). *Place and politics: The geographical mediation of state and society*. Routledge.

Agnew, J., & Shin, M. (2020). The Counties that Counted: Could 2020 Repeat 2016 in the US Electoral College?. *The Forum* (Vol. 17, No. 4, pp. 675-692).

Anselin, L., & Bera, A. K. (1998). Introduction to spatial econometrics. *Handbook of applied economic statistics*, *237*.

Buja, A., Hastie, T., & Tibshirani, R. (1989). Linear smoothers and additive models. *The Annals of Statistics*, *17*(2), 453-510.

Cho, W. K. T., & Gimpel, J. G. (2009). Presidential voting and the local variability of economic hardship. In *The Forum*(Vol. 7, No. 1). De Gruyter.

Cox, K. R. (1969). The voting decision in a spatial context. *Progress in geography*, *1*, 81-117.

Cupido, K., Fotheringham, A. S., & Jevtic, P. (2019). Local Modeling of US Mortality Rates: A Multiscale Geographically Weighted Regression Approach. *Available at SSRN 3472830*.

Darmofal, D. (2008). The political geography of the new deal realignment. *American Politics Research*, *36*(6), 934-961.

Degan, A., & Merlo, A. (2011). A structural model of turnout and voting in multiple elections. *Journal of the European Economic Association*, *9*(2), 209-245.

Delaney, D., & Leitner, H. (1997). The political construction of scale. *Political geography*, *16*(2), 93-97.

Dormann, C., M. McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., ... & Kühn, I. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, *30*(5), 609-628.

Duncan, J. S., & Ley, D. (2013). *Place/culture/representation*. Routledge.

Flint, C. (1996). Whither the individual, whither the context?. *Political Geography*, *15*(2), 147-151.

Forest, B. (2018). Electoral geography: From mapping votes to representing power. *Geography Compass*, *12*(1), e12352.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons.

Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale geographically weighted regression (mgwr). *Annals of the American Association of Geographers*, *107*(6), 1247-1265.

Fotheringham, A. S., Yue, H., & Li, Z. (2019). Examining the influences of air quality in China's cities using multi-scale geographically weighted regression. *Transactions in GIS,* 23(6): 1444-1464.

Fotheringham, A. S., Li, Z. & Wolf, L. J. (2021). Scale, Context and Heterogeneity: A Spatial Analytical Perspective on the 2016 US Presidential Election. *Annals of the American Association of Geographers,* 111(6) 1602-1621.

Gelfand, A. E., Kim, H. J., Sirmans, C. F., & Banerjee, S. (2003). Spatial modeling with spatially varying coefficient processes. *Journal of the American Statistical Association*, *98*(462), 387-396.

Griffith, D. A. (2003). *Spatial autocorrelation and spatial filtering: gaining understanding through theory and scientific visualization*. Springer Science & Business Media.

Hastie, T. J., & Tibshirani, R. J. (1990). *Generalized Additive Models* (Vol. 43). CRC Press.

Johnston, R., Shelley, F. M., & Taylor, P. J. (1990). *Developments in electoral geography*. Routledge.

Johnston, R., & Pattie, C. (2006). *Putting voters in their place: Geography and elections in Great Britain*. Oxford University Press.

Johnston, R., Rohla, R., Manley, D., & Jones, K. (2019). Voting for Trump and the electoral mosaics of US metropolitan areas: Exploring changing patterns of party support by neighborhood. *Cities*, *86*, 94-101.

King, G. (1996). Why context should not count. *Political geography*, *15*, 159-164.

Kühn, I. (2007). Incorporating spatial autocorrelation may invert observed patterns. *Diversity and Distributions*, *13*(1), 66-69.

Lappie, J., & Marschall, M. (2018). Place and participation in local elections. *Political Geography*, *64*, 33-42.

Leighley, J. E., & Nagler, J. (2013). *Who votes now?: Demographics, issues, inequality, and turnout in the United States*. Princeton University Press.

Li, Z., & Fotheringham, A. S. (2020). Computational improvements to multi-scale geographically weighted regression. *International Journal of Geographical Information Science*, doi:10.1080/13658816.2020.1720692.

Li, Z., Fotheringham, A. S., Oshan, T. M., & Wolf, L. J. (2020). Measuring Bandwidth Uncertainty in Multiscale Geographically Weighted Regression Using Akaike Weights. *Annals of the American Association of Geographers*, doi:10.1080/24694452.2019.1704680.

Miller, B. (1994). Political empowerment, local—central state relations, and geographically shifting political opportunity structures: Strategies of the Cambridge, Massachusetts, Peace Movement. *Political Geography*, *13*(5), 393-406.

Miller, J. A., & Grubesic, T. H. (2021). A spatial exploration of the halo effect in the 2016 US presidential election. *Annals of the American Association of Geographers*, *111*(4), 1094-1109.

Morgan, S. L., & Lee, J. (2018). Trump voters and the white working class. *Sociological Science*, *5*, 234-245.

Morrill, R., Knopp, L., & Brown, M. (2011). Anomalies in red and blue II: Towards an understanding of the roles of setting, values, and demography in the 2004 and 2008 US presidential elections. *Political Geography*, *30*(3), 153-168.

O'Loughlin, J., Flint, C., & Anselin, L. (1994). The geography of the Nazi vote: Context, confession, and class in the Reichstag election of 1930. *Annals of the Association of American Geographers*, *84*(3), 351-380.

O'Loughlin, J. (2018). Thirty-five years of political geography and Political Geography: The good, the bad and the ugly. *Political Geography*, *65*, 143-151.

Oshan, T. M., & Fotheringham, A. S. (2018). A comparison of spatially varying regression coefficient estimates using geographically weighted and spatial-filter-based techniques. *Geographical Analysis*, *50*(1), 53-75.

Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., & Fotheringham, A. S. (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.

Oshan, T. M., Smith, J. P., & Fotheringham, A. S. (2020). Targeting the spatial context of obesity determinants via multiscale geographically weighted regression. *International Journal of Health Geographics*, *19*(1), 1-17.

Scala, D. J., Johnson, K. M., & Rogers, L. T. (2015). Red rural, blue rural? Presidential voting patterns in a changing rural America. *Political Geography*, *48*, 108-118.

Schaffner, B. F., MacWilliams, M., & Nteta, T. (2018). Understanding white polarization in the 2016 vote for president: The sobering role of racism and sexism. *Political Science Quarterly*, *133*(1), 9-34.

Shin, M. E., & Agnew, J. A. (2008). *Berlusconi's Italy: mapping contemporary Italian politics*. Temple University Press.

Sides, J., Tesler, M., & Vavreck, L. (2017). The 2016 US election: How Trump lost and won. *Journal of Democracy*, *28*(2), 34-44.

Smith, D. N., & Hanley, E. (2018). The anger games: Who voted for Donald Trump in the 2016 election, and why?. *Critical Sociology*, *44*(2), 195-212.

Sui, D. Z., & Hugill, P. J. (2002). A GIS-based spatial analysis on neighborhood effects and voter turn-out: a case study in College Station, Texas. *Political Geography*, *21*(2), 159-173.

Taylor, P. J. (1973). Some implications of the spatial organization of elections. *Transactions of the Institute of British Geographers*, 121-136.

Tesler, M. (2016). The education gap among whites this year wasn't about education. It was about race. *Washington Post*, *16*.

Thioulouse, J., Chessel, D., & Champely, S. (1995). Multivariate analysis of spatial patterns: a unified approach to local and global structures. *Environmental and Ecological Statistics*, *2*(1), 1-14.

Tyson, A. & Maniam, S (2016). Behind Trump's victory: Divisions by race, gender, education. Retrieved from https://www.pewresearch.org/fact-tank/2016/11/09/behind-trumps-victory-divisions-by-race-gender-education.

Warf, B., & Leib, J. (2011). Revitalizing electoral geography. Burlington, VT: Ashgate.

Yang, C., Zhan, Q., Lv, Y., & Liu, H. (2019). Downscaling Land Surface Temperature Using Multiscale Geographically Weighted Regression Over Heterogeneous Landscapes in Wuhan, China. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

Wolf, L. J., Oshan, T. M., & Fotheringham, A. S. (2018). Single and multiscale models of process spatial heterogeneity. *Geographical Analysis*, *50*(3), 223-246.

Yu, H., Fotheringham, A. S., Li, Z., Oshan, T., Kang, W., & Wolf, L. J. (2019). Inference in multiscale geographically weighted regression. *Geographical Analysis*.

Yu, H., Fotheringham, A. S., Li, Z., Oshan, T., & Wolf, L. J. (2020). On the measurement of bias in geographically weighted regression models. *Spatial Statistics*, *38*, 100453.