

# Presidential Voting and County Composition; A Research Note on the 2020 Election

**Gregg Smith**

*University of Tasmania, Data Solutions\**

## Abstract

*We investigate the 2020 Presidential Election by using the county as the unit of analysis ( $n = 3,112$ ) and examine the variance in the county percentage of votes cast for Biden and Trump as dependent variables. Our seven independent variables are conceptually related to racial diversity, educational attainment, wellbeing (includes economic, health, and quality-of-life indicators), and rural/urban classification. Using OLS regression, the model explains about 73% ( $R^2$ ) of the variance in county vote outcome for both the Biden and Trump models. The regression results show that the counties that Trump won tended to be populated with people who are more white, less educated, score lower on a composite wellbeing measure and are located in a more rural setting compared to the counties that Biden won. However, the urban/rural effect is so small that we question whether there is a real effect after controlling for the effects of other variables.*

Keywords: 2020 Election, County analysis, Presidential voting

\* This project was funded by *Data Solutions*. Questions and comments should be directed to the author at [gs253253@bigpond.net.au](mailto:gs253253@bigpond.net.au). The original dataset is freely available, contact the author. Gregg Smith received his PhD in political science from the University of Wisconsin – Milwaukee. He has taught at universities in the United States and Australia, and has done research for the Australian Government at the executive-level. Sussan Best, Southern Cross University, provided helpful comments. The author has no competing interests to declare.

## Introduction

This research examines the 2020 county-level presidential vote.<sup>1</sup> There are 3,141 counties or county equivalents (Boroughs, Parishes, Independent Cities, and 1 District) in the United States. We had to exclude Alaska because of incompatibility between vote and county data, and one county in Hawaii, Kalawao, is excluded due to lack of data, which leaves us with 3,112 counties or county equivalents. We believe that the observed variance in votes cast for president at the county-level can be largely explained by the county's compositional demographics. The four conceptual factors we are interested in are racial diversity, educational attainment, overall *wellbeing*, which is based on economic and health indicators, and where a county is located along a rural/urban continuum.

Voting is fundamentally individual behaviour and the bulk of empirical research tries to untangle voting patterns for an election based on a psychological model in which vote choice is a function of attitudes about party, issues and candidates (see Abramowitz 2018 for a comprehensive description of changes in voting patterns, partisanship and coalitions from 1952 to 2016; see Finkle et al. 2020 for a current description of the partisan landscape). In this research we are looking at patterns of vote outcome when all votes casted are aggregated to the county-level.

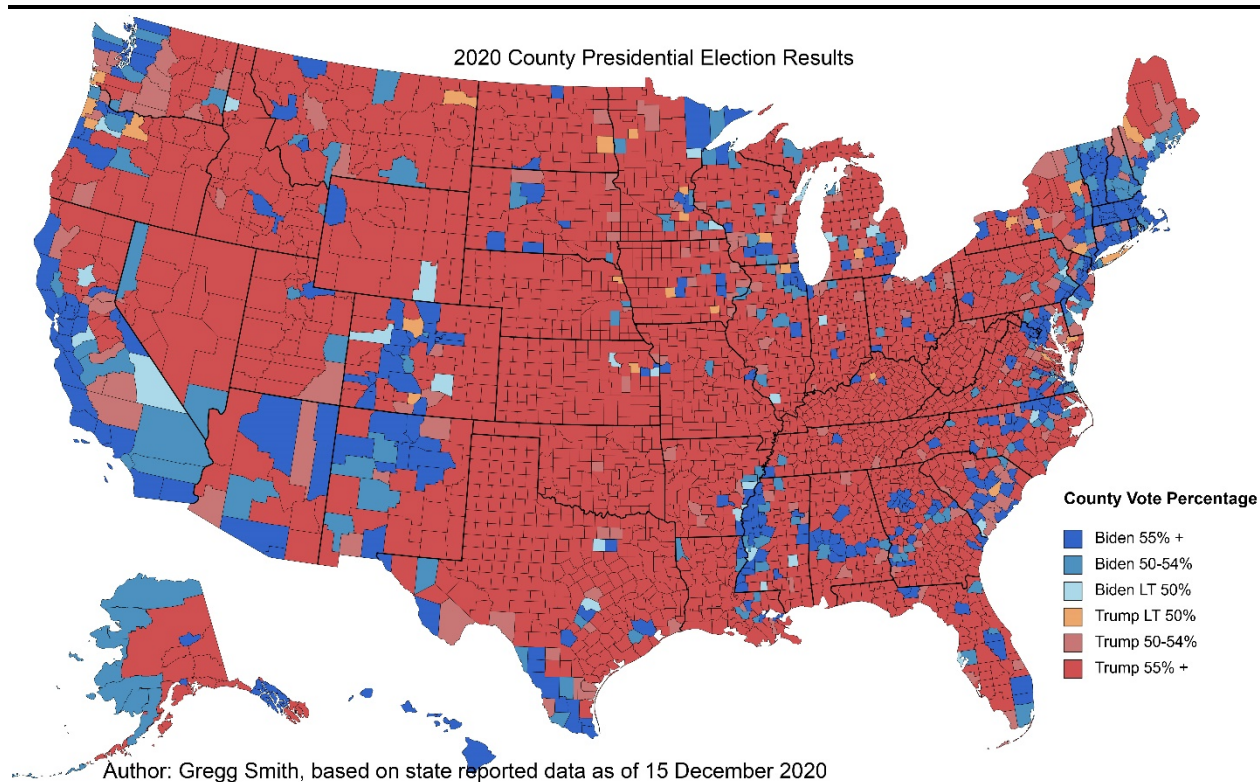
Using the county as the unit of analysis provides a bridge between the characteristics of individuals, who do the voting, and emergent county-level phenomena. For example, consider the differences between asking an individual about employment as an issue and calculating an unemployment rate. We think that having a theoretical-level above the individual will prove useful in understanding some aspects of the 2020 presidential election. Studying counties also provides a means for assessing the effect of rural/urban location as it divides the country into over 3,000 geopolitical units.

Counties are rich in data as a result of their legal status as one of the primary administrative units of the United States, and consequently the national government collects and reports a vast amount of county data. The number of counties per state ranges from three in Delaware to 254 in Texas. The specific governmental powers of counties vary widely between the states. Counties have significant functions in all but two states, Rhode Island and Connecticut. Political parties have county-level organizations, which become a workhorse of the party system during elections. County population and composition vary widely. More than half

of the population live in just five percent of the counties. Less than 200 people live in Loving County (Texas) and 66 of them voted in 2020, in contrast to the 4,263,443 people who voted in Los Angeles County. About 14.2% of the national population is African American, but this conceals the variation at the state-level. Idaho has a black population of just 1.2% in contrast to Mississippi's 38.9%, but the state data also conceals that Elmore County (ID) has 3.1% and Tishomingo (MS) has just 2.5%.

### Overview of the 2020 Outcome

One of the most iconic and widely published images of presidential elections is the blue and red map showing the county vote result, which is usually seen on the front page of newspapers the day after the election, but was mostly absent or incomplete in 2020. Our version of this map is shown in Figure 1 and based on official results as of 15 December 2020. By convention, the redder the color, the greater the Trump victory and the shades of blue show the margin of victory for Biden.



**Figure 1: 2020 County Vote Result, Blue (Biden), Red (Trump)**

Biden won the national election by getting 306 of the 538 state elector votes (57%), which came from 25 states and the District of Columbia, and included five states that Trump

won in 2016; Arizona, Georgia, Michigan, Pennsylvania and Wisconsin. Biden won one of Nebraska's five electors and three of Maine's four electors, which are the only two states that don't use a winner-take-all system. Biden won the popular vote: 81,283,098 (51.3%) to Trump's 74,222,958 (46.8%). Other candidates received 2,926,539 votes, or 1.8%, which was considerably down from the 5.2% they got in 2016.

At the county-level, Trump received more than 50% of the vote in 2,545 counties, got more votes than Biden in a further 28 counties with less than 50% of the vote and won all the counties in West Virginia and Oklahoma.<sup>2</sup> Biden won all the counties in Hawaii, Massachusetts and Rhode Island. Thus, Trump beat Biden in 83% of the 3,112 counties in this analysis which is slightly down from the 84% he won in 2016. Most of the counties that Trump won in 2016, he also won in 2020. Eighty counties flipped in 2020 with 64 of those going to Biden and 16 to Trump. County turnout rates averaged 63.3% in 2020, which was an increase of 6.4% over the average 2016 turnout, which was 56.9% (Smith and Young 2020). Turnout rates varied from 19% in Chattahoochee, Georgia, to 93% in Leelanau, Michigan. Table 1 shows the counties that Trump won in 2020 compared to 2016 in five key states. Note that four of the five states had an average increase in turnout from 2016 that exceeded the national county average, which was 6.44%.

TABLE 1: Trump County Vote Result in Key States 2016, 2020

<i>State (number of counties)</i>	<i>Average Increase In Turnout from 2016 (%)</i>	<i>Counties Trump Won in 2016</i>	<i>Counties Trump Won in 2020</i>	<i>Difference between 2020 and 2016 (county)</i>
Arizona (15)	4.23	11	10	-1 (Maricopa)
Georgia (159)	8.22	128	129	+1 (Burke)
Michigan (83)	9.87	75	72	-3 (Saginaw, Leelanau, Kent)
Penn (67)	8.81	56	54	-2 (Northampton, Erie)
Wisconsin (72)	7.52	60	58	-2 (Door, Sauk)
	<i>Average All, 6.44</i>			

### Variable Description

We are looking to explain the variance in two dependent variables: the percent of the county vote received by Biden and the percent by Trump. The operational independent variables are conceptually related to racial diversity, educational attainment, *wellbeing* and *rural/urban location*.

- Racial diversity is operationalized with an estimate for the percentage of the county that is that is non-white.<sup>3</sup>
- Educational attainment is operationalized with an estimate for the percentage of the county population that has at least a bachelor's degree.
- *Wellbeing* is more complicated because it is a multidimensional concept. In brief, it assigns a value to how well or poorly a county population is doing. We use four surrogate variables to capture this. One economic item (percent of the county in housing distress), two health indicators (changes in life expectancy and physical fitness)<sup>4</sup> and an overall *quality of life* measure developed by the Robert Wood Johnson Foundation and the University of Wisconsin Population Health Institute.<sup>5</sup>
- *Rural/urban location* refers to how the county population is distributed within the county's geographic boundaries (see Ghelfi and Parker 1997, 2004). Where the county is located on a rural/urban continuum determines the availability of goods and services, and opportunities for employment. It is operationalized with a nine-point ordinal scale developed by the Department of Agriculture.<sup>6</sup>

Table 2 is a statistical description of the variables in this analysis. It includes the variables in the model and variables that are discussed but excluded from the model. Table 2 also shows the bivariate Pearson correlations between the vote outcome dependent variable and the set of county composition variables. We prefer to do a separate multivariate analysis for the Biden and Trump outcome, but for purposes of showing the zero-order correlations, we create a vote outcome variable by subtracting the percentage of the county vote received by Biden from the percentage received by Trump. Thus, positive values are associated with a Trump victory and negative values with a Biden win.

The bivariate correlations are consistent with previous research that found a strong relationship between race and vote (e.g., Carmines and Stimson 1989; Tate 1991; Leighley 2001; Frey 2014; Cook et al. 2017). Counties with more racial diversity (non-white population) voted

more for Biden than Trump ( $r = -.55$ ). Galston and Hendrickson (2016) reported that the impact of educational level on presidential vote choice was greater in the 2016 election than in previous elections and our findings for 2020 indicate that this continues to be the case. The correlation for the percent of the county that has at least a bachelor's degree and vote outcome is a moderately strong ( $r = -.54$ ).

TABLE 2: Descriptive Statistics of County Variables

<i>County-Level Variable (2020)</i>	<i>Zero-Order Correlation (r) With Vote Outcome</i>	<i>Mean</i>	<i>Median</i>	<i>Std. Deviation</i>	<i>N</i>
Biden Votes (%)		33.27	29.98	15.98	3,112
Trump Votes (%)		64.97	68.29	16.13	3,112
Vote Outcome (Rep. % - Dem. %)		31.70	38.44	32.10	3,112
<b>Independent Variables</b>					
Non-white Population (%)	-.55	15.31	8.60	15.89	3,112
Ed, bachelor's degree or higher (%)	-.54	21.56	19.20	9.43	3,112
Quality of Life measure (z-score)	-.14	-.002	-.033	.395	3,059
Physical Inactivity (% of population)	.36	51.41	51.90	6.44	3,112
Changes in Life Expectancy (years)	-.46	2.72	2.70	1.39	3,112
Severe Housing Problems (%)	-.57	13.74	13.30	4.15	3,112
Rural / Urban Continuum (ordinal value)	.41	4.99	6.00	2.70	3,112
<b>Non-model Variables</b>					
Population Density (pop per sq mile)	-.27	275	45	1,794	3,112
Median Age (years)	.24	41.8	41.6	5.37	3,112
Median Household Income (\$)	-.27	52,686	50,530	13,838	3,112
Poverty (%)	-.04	15.17	14.10	6.11	3,112

*Wellbeing* is conceptually defined as a county characteristic that is analogous to an overall attitude about how well or how poorly the county is doing. Individual attitudes on economic issues (e.g., Fiorina 1978; Evans and Anderson 2006) such as jobs, wages, taxes; social issues such as crime (e.g., Arnold and Carnes 2012), welfare (e.g., Gilens 2005), environment, immigration (e.g., Citrin et al. 2007), drugs and war (e.g., Grose and Oppenheimer

2007); and service availability and delivery issues such as healthcare, education (e.g., Holbein and Ladd 2015) and transportation will coalesce into a generalised county attitude about how well things are going. Different combinations of economic, social and service delivery attitudes, which collapse into a sense of *wellbeing* that affects vote outcome are nuanced, not necessarily linear and contextualised by affections for parties and candidates (e.g., the rise in *negative partisanship*, Abramowitz and Webster 2016).

We operationalize *wellbeing* with four variables. A summary composite item, *quality of life*, which according to its authors, “Examining quality of life can tell us a lot about how people perceive their health – whether they feel healthy and satisfied. When communities have higher rates of those who do not feel healthy, it can influence other factors of health including mortality rates, unemployment, poverty, and the percentage of adults who did not complete high school.”<sup>7</sup> A surrogate economic indicator (housing distress); two health-related surrogates, the county’s change in life expectancy from 1985 to 2010 (see Wang et al. 2013) and the percentage of the county that is physically inactive (see Dwyer-Lindgren et al. 2013).<sup>8</sup> Note that the percent of the county with severe housing problems has the highest correlation with vote outcome ( $r = -.57$ ).

The *Rural/urban Continuum* variable subdivides the Office of Management and Budget (OMB) metro and nonmetro categories into three metro and six nonmetro categories. Each county is assigned a value from one to nine (see note 6). Previous research has established the increasing importance of population density in explaining voting behaviour (see, e.g., Frey 2014; Enos 2017; Abramowitz 2018). According to Wilkinson (2019), “urban-rural economic divergence has put many lower-density areas in dire straits, activating a zero-sum, ethnocentric mindset receptive to scapegoating populist rhetoric” (1). The bivariate correlation between rural/urban location and 2020 vote outcome is about .4, which indicates that Trump did better in rural counties.

If you extrapolate the bivariate correlations to describe a typical Trump county, one could say that they are populated with older white people from non-metro areas who are neither rich nor poor, are less educated, are more unhealthy as they are less physically fit and have shorter expected lifespans compared to counties that supported Biden.<sup>9</sup>

### Multivariate Analysis

The correlations in Table 2 indicate that all seven predictor variables are significantly related to the percent of the county vote for each candidate. What we want to know is what happens when the variance of the county vote is regressed on the variance of the input variables. To answer this, we use ordinary least squares (OLS) regression. OLS creates estimates that minimise the squared distances between the county vote predicted by the model and the observed county vote. This method will estimate the effect of each independent variable on the dependent variable while controlling for the effects of the other independent variables. OLS also assesses the overall fit of the model; that is, how much of county vote is explained by the regression. More specifically, the SPSS regression procedure was used with stepwise entry in which the probability ( $p$ ) of the  $F$ -ratio to enter the model was specified at  $p \leq .001$  and the probability to be removed must be  $p \geq .005$ . The variable with the smallest probability will be entered at each step provided it meets the ( $p$ ) selection criteria.<sup>10</sup>

Tables 3 and 4 describe the results of the regression analysis for the Biden and Trump vote. The results, of course, are similar with opposite signs but are not identical. The independent variables are listed in order of their influence based on the absolute strength of the standardised beta coefficients ( $\beta$ ). The betas can be compared to each other; the larger the absolute value of the beta, the greater the influence of the variable. The step the variables were entered is listed in the column on the left. The adjusted coefficient of determination ( $R^2$ ) is cumulative and can be interpreted as the percent of the variance in the dependent variable that is explained by the independent variables. For both models,  $R^2$  is about .73, which means that 73% of the variance in the percentage of votes cast for Biden and Trump is explained with seven independent variables. The tolerance statistic is the percent of the variance in each predictor that cannot be explained by the other variables in the model. It can be used to assess the extent to which multicollinearity is a problem. A value close to zero would indicate that a predictor's influence can be explained by the other variables. The tolerance value for all seven variables is sufficiently above zero for us to conclude that each of them is making a unique contribution. The  $t$ -ratio is the unstandardized coefficient divided by its standard error, and is used to assess the statistical significance of each variable. The significance of the  $t$  ratios for all regression coefficients in both models is at least  $p < .0001$ .



Table 3 presents the results of the regression analysis when the percent of the vote for Biden is the dependent variable. Non-white population has the greatest effect and is positive ( $\beta = .41$ ), which means that as racial diversity increases, the percentage of the vote going to Biden also increases. Educational attainment is the next most important ( $\beta = .35$ ) and is also positive; as the percent of the county population that has at least a bachelor's degree increases, so does the percentage of the vote cast for Biden. The other variables are significant and in the expected direction. As quality of life, housing problems and life expectancy increase, so does the Biden vote; as the physical fitness of the population decreases, so does the Biden vote. The negative beta for rural/urban location ( $\beta = -.12$ ) means that the more rural the county, the less likely it is to vote for Biden.

Table 3: County Composition and Biden County Vote Variance

Dependent Variable: Biden’s County Vote Percent 2020. Method: OLS		Unstandardized Coefficients		Standardized Coefficients	Tolerance	t ratio
		<i>B</i>	<i>Std.</i>	<i>Beta</i> ( <i>β</i> )		
			<i>Error</i>			
(Constant)		<i>17.414</i>	<i>1.582</i>			<i>11.00</i>
[Step Entered]						
[3]	Non-White Population (%)	<i>.410</i>	<i>.012</i>	<i>.408</i>	<i>.674</i>	<i>35.43</i>
[2]	Percent with a BA or higher	<i>.592</i>	<i>.024</i>	<i>.349</i>	<i>.446</i>	<i>24.66</i>
[5]	Quality of Life (Z-Score)	<i>8.979</i>	<i>.490</i>	<i>.222</i>	<i>.612</i>	<i>18.33</i>
[4]	Physically Inactive (%)	<i>-.495</i>	<i>.036</i>	<i>-.176</i>	<i>.551</i>	<i>-13.82</i>
[1]	Severe Housing Problems (%)	<i>.664</i>	<i>.046</i>	<i>.173</i>	<i>.641</i>	<i>14.59</i>
[6]	Changes in Life Expectancy (years)	<i>1.754</i>	<i>.133</i>	<i>.153</i>	<i>.664</i>	<i>13.17</i>
[7]	Rural / Urban Continuum Code	<i>-.697</i>	<i>.065</i>	<i>-.118</i>	<i>.744</i>	<i>-10.74</i>
Adjusted Coefficient of determination: $R^2 = .726$ Significance of all t ratios, $p < .0001$						

Table 4 presents the results when the Trump vote percentage is the dependent variable. Comparing Table 4 with Table 3 reveals (as expected) that the regression coefficients are very similar except, of course, the direction (+ or -) is reversed. The unstandardized regression coefficient for non-white population is  $B = -.40$  for Trump and  $B = .41$  for Biden; the coefficient for the percent with at least a college degree is  $B = -.62$  for Trump and  $B = .59$  for Biden. Collectively, the independent variables account for about 73% of the variance in the Trump vote.

For Trump, just like the Biden results, the variables that have the greatest impact are racial diversity and educational attainment.

Table 4: County Composition and Trump County Vote Variance

Dependent Variable: Trump’s County Vote Percent 2020. Method: OLS		Unstandardized Coefficients		Standardized Coefficients	Tolerance	t ratio
		<i>B</i>	<i>Std. Error</i>	<i>Beta (β)</i>		
(Constant)		80.309	1.600			50.20
[Step Entered]						
[3]	Non-White Population (%)	-.404	.012	-.398	.674	-34.44
[2]	Percent with a BA or higher	-.620	.024	-.363	.446	-25.54
[5]	Quality of Life (Z-Score)	-9.156	.495	-.224	.612	-18.48
[4]	Physically Inactive (%)	.522	.036	.184	.551	14.41
[1]	Severe Housing Problems (%)	-.664	.046	-.171	.641	-14.41
[6]	Changes in Life Expectancy (years)	-1.667	.135	-.144	.664	-12.38
[7]	Rural / Urban Continuum Code	.705	.066	.118	.744	10.75
Adjusted Coefficient of determination: $R^2 = .725$ Significance of all t ratios, $p < .0001$						

### Discussion of the Biden/Trump Regression Results

In both models, the bulk of explained variance is the result of the racial diversity and educational attainment within the county. These two variables are the primary predictors and by themselves explain about 51% ( $R^2$ ) of the Biden/Trump vote variance and account for about 72% of the model's predictive power. We operationalized racial diversity with just one race variable, percent non-white, because that variable is the best statistical predictor in the context of the other variables in the model. We considered but excluded other race variables such as percent white or Asian because race variables are highly correlated with each other so once one is entered, the remaining explanatory variance in the others is not significantly correlated with the dependent variable. Similarly, educational attainment is operationalized with just one variable, percent of population with at least a bachelor's degree. We excluded other education variables such as percent with some college or high school graduation rate because there is again too much multicollinearity among them. Thus, racial diversity is important in understanding vote outcome

and the percent of the county population that is not white is the best predictor. Level of education is important and the percent of the county that has at least a bachelor's degree is the best predictor.

Even though the remaining five variables collectively increase the  $R^2$  by a modest 20%, which is about 27% of the model's explanatory power, they are all theoretically important and statistically significant. The true influence of the five secondary predictors is likely depressed as an artifact of the equations used to create the model. As the independent variables are correlated with each other ( $r$  values range from 0 to .53), they are not truly independent and the problem of multicollinearity arises. For example, the correlation between non-white population and quality of life is .34. Any shared variance between those two variables and the dependent variable will be attributed to the not white population because it was entered first. However, multicollinearity does not affect the regression line or the model's goodness of fit. The five secondary predictors would explain 54% (adjusted  $R^2$ ) of the Biden/Trump vote variance if race and education were removed from the model.

Four of the secondary predictors measure the concept of wellbeing. When county vote is regressed on the wellbeing variables, about 50% ( $R^2$ ) of the vote variance is explained. We see this as evidence that a county composition model explaining vote outcome must include a *wellbeing* type variable, but not necessarily the surrogate indicators we selected.

The rural/urban classification variable is important for the same reasons, theoretical and statistical significance. The model shows that where the county is located affects vote outcome such that rural counties tended to support Trump and urban counties tended to Biden, and recall that the bivariate correlation with vote is  $\pm .41$ . However, most of the bivariate correlation between vote and location is spurious because it can be explained with other variables, but there is still something left in it to have a unique contribution to vote outcome. The rural/urban variance inflation factor statistic (VIF) is 1.34 and the related tolerance statistic is .744. The tolerance value means that 74% of the rural/urban effect cannot be explained by the other independent variables. This suggests that after controlling for the effects of the other variables, rural/urban location appears to exert a small effect.

Alternatively, it may be that a rural/urban variable doesn't belong in a county composition model at all. It was entered on the last step of the regression and barely raises the coefficient of determination (adjusted  $R^2$  increase = .01). If population density had been used

to measure the rural/urban concept, it would not have been entered into the model as it would not have met the significance criterion to be included. It could be that location has no independent effect on vote outcome but is rather more like a canvass where other factors collect after being filtered and sorted. Extending the canvass analogy, reds tend to be painted with other reds and blues with blues. So, reds and blues are seen together in different locations but it is not the location that is important, it's the color that creates the effects. For now, we are content to keep our rural/urban variable in the model but stress that its true multivariate effect, if any, is much weaker than its bivariate correlation. Contrary to some published studies, we think it is an open question as to whether or not a population density type variable has any independent effect on vote choice and we don't think our concerns can be brushed away with an ecological fallacy argument.

### Conclusion

We think that studying what happens within the county contributes to our understanding of American presidential elections. Some variables of interest, such as health and education, are more useful when measured at the county-level rather than the state or individual level. It is advantageous to have a vote-choice dependent variable that takes on interval values, as opposed to one based on asking a voter who they voted for, which typically takes on two values.

We sought to create a parsimonious model in which the votes cast for Biden/Trump at the county-level could be explained with county composition variables that represent four theoretical concepts; racial diversity, educational attainment, wellbeing and rural/urban location. Seven surrogate independent variables were selected, one variable each for race, education and location; and four variables for wellbeing. About 73% of the variance in the county vote cast for Biden and Trump is explained with this model. The regression results show that the counties that Trump won tended to be populated with people who are more white, less educated, score lower on a *quality-of-life* measure, are less physically fit, have lower improvements in life expectancy, have marginally better economies and are located in a more rural setting compared to the counties that Biden won.

## References

- Abramowitz, Alan I. 2018. *The Great Alignment*. New Haven, CT: Yale University Press.
- Abramowitz, Alan I., and Steven Webster. 2016. "The Rise of Negative Partisanship and the Nationalization of US Elections in the 21st Century." *Electoral Studies* 41: 12–22.
- Arnold, R. Douglas, and Nicholas Carnes. 2012. "Holding Mayors Accountable: New York's Executives from Koch to Bloomberg." *American Journal of Political Science* 56 (4): 949–63. <https://doi.org/10.1111/j.1540-5907.2012.00603.x>.
- Blalock, Hubert M. Jr. 1963. "Correlated Independent Variables: The Problem of Multicollinearity". *Social Forces* 42 (2): 233–237. <https://doi.org/10.1093/sf/42.2.233>.
- Burden, Barry C., Jason M. Fletcher, Pamela Herd, Bradley M. Jones, and Donald P. Moynihan. 2017. "How Different Forms of Health Matter to Political Participation". *The Journal of Politics* 79 (1): 166–78. <https://doi.org/10.1086/687536>.
- Carmines, Edward G., and James A. Stimson. 1989. *Issue Evolution: Race and the Transformation of American Politics*. Princeton, NJ: Princeton University Press.
- Cook, Alexander C., Nathan J. Hill, Mary I. Trichka, Grace J. Hwang, and Paul M. Sommers. 2017. "Who Voted for Trump in 2016?" *Open Journal of Social Sciences* 5 (7): 199–210. <https://doi.org/10.4236/jss.2017.57013>.
- Citrin, Jack, Amy Lerman, Michael Murakami, and Kathryn Pearson. 2007. "Testing Huntington: Is Hispanic Immigration a Threat to American Identity?" *Perspectives on Politics* 5 (1): 31–48. <https://doi.org/10.1017/s1537592707070041>.
- Devos, Thierry, and Mahzarin R. Banaji. 2005. "America =White?" *Journal of Personality and Social Psychology* 88 (3):447–66. <https://doi.org/10.1037/0022-3514.88.3.447>.
- Dwyer-Lindgren, Laura, Greg Freedman, Rebecca E. Engell, Thomas D. Fleming, Stephen S. Lim, Christopher J. L. Murray, and Ali H. Mokdad. 2013. "Prevalence of physical activity and obesity in US counties, 2001–2011: a road map for action". *Population Health Metrics* 11: 7. <https://doi.org/10.1186/1478-7954-11-7>.
- Enos, Ryan D. 2017. *The Space between Us: Social Geography and Politics*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108354943>.
- Evans, Geoffrey, and Robert Anderson. 2006. "The Political Conditioning of Economic Perceptions." *Journal of Politics*, 68 (1): 194–207. <https://doi.org/10.1111/j.1468->

2508.2006.00380.x.

- Finkel Eli J, Christopher A. Bail, Mina Cikara, Peter H. Ditto, Shanto Iyengar, Samara Klar, Lilliana Mason, Mary C. McGrath, Brendan Nyhan, David G. Rand, Linda J. Skitka, Joshua A. Tucker, Jay J. Van, Bavel, Cynthia S. Wang, James N. Druckman. 2020. "Political sectarianism in America." *Science* 370 (6516): 533-536. DOI: 10.1126/science.abe1715.
- Fiorina, Morris P. 1978. "Economic Retrospective Voting in American National Elections: A Micro-Analysis." *American Journal of Political Science* 22 (2): 426-43. <https://doi.org/10.2307/2110623>.
- Frey, William H. 2014. "Diversity Explosion: How New Racial Demographics Are Remaking America." *Brookings Institution Press* (10): Washington, D.C. <https://muse.jhu.edu/book/35812>.
- Galston, William A., and Clara Hendrickson. 2016. "The educational rift in the 2016 election." *Brookings Institution*. <https://www.brookings.edu/blog/fixgov/2016/11/18/educational-rift-in-2016-election/>.
- Ghelfi, Linda M., and Timonthy S. Parker. 1997. "A County-Level Measure of Urban Influence." *Rural America / Rural Development Perspective, United States Department of Agriculture, Economic Research Service* 12 (2). DOI: 10.22004/ag.econ.278818.
- Ghelfi, Linda M., and Timonthy S. Parker. 2004. "Using the 2003 Urban Influence Codes To Understand Rural America." *United States Department of Agriculture, Economic Research Service*. <https://www.ers.usda.gov/amber-waves/2004/april/using-the-2003-urban-influence-codes-to-understand-rural-america/>.
- Gilens, Martin. 2005. *Why Americans hate welfare: Race, media, and the politics of Antipoverty policy*. Chicago, IL: University of Chicago Press.
- Grose, Christian R., and Bruce I. Oppenheimer. 2007. "The Iraq War, Partisanship, and Candidate Attributes: Variation in Partisan Swing in the 2006 US House Elections." *Legislative Studies Quarterly* 32 (4): 531-57. <https://doi.org/10.3162/036298007782398495>.
- Holbein, John, and Helen F. Ladd. 2015. "Accountability Pressure and Non-Achievement Student Behaviors." *Calder Working Paper No. 122*. <https://caldercenter.org/publications/accountability-pressure-and-non-achievement->

student-behaviors.

- Leighley, Jan E. 2001. *Strength in Numbers? The Political Mobilization of Racial and Ethnic Minorities*. Princeton, NJ: Princeton University Press.
- Lewis-Beck, Michael S. 1978. "Stepwise Regression: A Caution." *Political Methodology* 5 (2): 213-40. <https://www.jstor.org/stable/25791533>.
- Ksiazkiewicz, Aleksander. 2020. "Conservative Larks, Liberal Owls: The Relationship Between Chronotype and Political Ideology." *The Journal of Politics* 82 (1): 367-71. <https://doi.org/10.1086/705927>.
- Pacheco, Julianna, and Jason Fletcher. 2015. "Incorporating Health into Studies of Political Behavior: Evidence for Turnout and Partisanship." *Political Research Quarterly* 68: 104-16. <https://doi.org/10.1177/1065912914563548>.
- Smith, Gregg, and Jazmin Young. (2020). "A Multivariate Analysis of the 2016 County-Level Presidential Vote and Turnout". *Advance* (preprint): <https://doi.org/10.31124/advance.12751913.v1>.
- Söderlund, Peter, and Lauri Rapeli. 2015. "Personal Health and Political Participation in the Nordic Countries." *Politics and the Life Sciences* 34: 28-43. <https://doi.org/10.1017/pls.2015.3>.
- SPSS Reports. 1985. "Regression." In *SPSS Statistical Algorithms*, 165-75. Chicago, IL: SPSS Inc.
- Tate, Katherine. 1991. "Black Political Participation in the 1984 and 1988 Presidential Elections." *American Political Science Review* 85 (4): 1159-76. <https://doi.org/10.2307/1963940>.
- Wang, H., A. E. Schumacher, C. E. Levitz, A. H. Mokdad, and C. J. Murray. 2013. "Left behind: widening disparities for males and females in US county life expectancy, 1985–2010." *Population Health Metrics* 11:8. <https://doi.org/10.1186/1478-7954-11-8>.
- Wilkinson, Will. 2019. "The Density Divide: Urbanization, Polarization, and Populist Backlash". *Niskanen Center*. Washington, DC: <https://www.niskanencenter.org/the-density-divide-urbanization-polarization-and-populist-backlash/>.

## Endnotes

<sup>1</sup> County 2020 vote for president source: Github,  
[https://github.com/tonmcg/US\\_County\\_Level\\_Election\\_Results\\_08-20/blob/master/2020\\_US\\_County\\_Level\\_Presidential\\_Results.csv](https://github.com/tonmcg/US_County_Level_Election_Results_08-20/blob/master/2020_US_County_Level_Presidential_Results.csv). Latest commit [f9b5f33](#) on Dec 15, 2020.

<sup>2</sup> Most partisan counties with at least 10,000 Population.

Most Democratic Counties with at least 10,000 Population		
<i>State</i>	<i>County</i>	<i>% Biden Vote 2020</i>
District of Columbia	District of Columbia	92.15
Maryland	Prince George's County	89.26
South Dakota	Oglala Lakota County	88.41
Virginia	Petersburg City	87.75
Maryland	Baltimore City	87.28
New York	New York County	86.78
Virginia	Charlottesville City	85.50
California	San Francisco County	85.27
Georgia	Clayton County	84.99
New York	Bronx County	83.48
Louisiana	Orleans Parish	83.15
Georgia	DeKalb County	83.12
Virginia	Richmond City	82.92
California	Marin County	82.34
Missouri	St. Louis City	82.24
Most Republican Counties with at least 10,000 Population		
<i>State</i>	<i>County</i>	<i>% Trump Vote 2020</i>
Alabama	Winston County	90.35
Georgia	Brantley County	90.24
Louisiana	LaSalle Parish	90.12
Alabama	Cleburne County	89.72
Alabama	Blount County	89.57
Texas	Gaines County	89.32
Kentucky	Jackson County	89.20
Florida	Holmes County	89.10
Kentucky	Martin County	88.82
Georgia	Banks County	88.57
Tennessee	Scott County	88.42
West Virginia	Grant County	88.42
Alabama	Marion County	88.40
Texas	Clay County	88.29
Utah	Duchesne County	88.14

<sup>3</sup> Unless otherwise noted, 2020 county population characteristics are sourced from the Census Bureau and based on estimates for July 2019 (last revised on June 22, 2020). Dataset:



Population, Population Change, and Estimated Components of Population Change: April 1, 2010 to July 1, 2019 (CO-EST2019-alldata). [https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-total.html#par\\_textimage](https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-total.html#par_textimage).

<sup>4</sup> Changes in life expectancy and are sourced from the *Center for Disease Control and Prevention*, <https://www.cdc.gov/nccdphp/dnpao/index.html>.

<sup>5</sup> See “2020 County Health Ranking” <https://www.countyhealthrankings.org/about-us> The County Health Rankings & Roadmaps program is a collaboration between the [Robert Wood Johnson Foundation](#) and the [University of Wisconsin Population Health Institute](#).

<sup>6</sup> The Rural/urban Continuum Codes form a classification scheme that distinguishes metropolitan counties by the population size of their metro area, and nonmetropolitan counties by degree of urbanization and adjacency to a metro area. Each county in the U.S. is assigned one of the 9 codes. <https://www.ers.usda.gov/data-products/rural/urban-continuum-codes.aspx>.

Code	Description
<i>Metro Counties</i>	
1	Counties in metro areas of 1 million population or more
2	Counties in metro areas of 250,000 to 1 million population
3	Counties in metro areas of fewer than 250,000 population
<i>Nonmetropolitan Counties</i>	
4	Urban population of 20,000 or more, adjacent to a metro area
5	Urban population of 20,000 or more, not adjacent to a metro area
6	Urban population of 2,500 to 19,999, adjacent to a metro area
7	Urban population of 2,500 to 19,999, not adjacent to a metro area
8	Completely rural or less than 2,500 urban population, adjacent to a metro area
9	Completely rural or less than 2,500 urban population, not adjacent to a metro area

We considered using population density as a rural/urban variable but as it is only weakly correlated with vote outcome ( $r = .27$ ) and decays to insignificance when the effects of the other variables in the model are considered.

<sup>7</sup> A full description of how the quality of life variable is created and the source of the quotation can be found at <https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources/county-health-rankings-model?componentType=health-outcome&componentId=6>.

<sup>8</sup> This variable is sourced from the 2020 County Health Rankings, see note 5. For a discussion on how health related variables have been used in political behaviour research, see Söderlund and Rapeli (2015); Pacheco and Fletcher (2015); Burden et al. (2017); and Ksiazkiewicz (2020).

---

<sup>9</sup> The attentive reader will note that we do not have an age or poverty variable. County median age and poverty were not included in the final model. The bivariate correlation for poverty rate is virtually zero. Median age is modestly correlated ( $r = +/- .24$ ) indicating that older voters trended to Trump. However, because most of age covariance with the dependent variables was shared with other independent variables, age became redundant and insignificant.

<sup>10</sup> We used SPSS, version 22. We have taken into consideration the criticism of this approach by Lewis-Beck (1978), problems with ordinal variables and multicollinearity (Blalock 1963). For a specification of the exact mathematical procedure, see SPSS Reports (1985), 165-75. All of the model variables are interval measurements except the rural/urban measurement, which is ordinal.