Are Demographic Trends Favoring Republicans or Democrats in Future Elections? Republicans and Democrats will need to adjust their political strategies as their coalitions shift.
By Zachary Spahr

Introduction:

The 2020 US Presidential Election, won narrowly by Joe Biden, was much closer than many political observers were expecting. As the results were counted, it became obvious that the bias of the electoral college had only grown since 2016, meaning that Republicans could have won the overall election while losing the popular vote by about four percentage points. Given that there has been a lot of discussion about how this point should impact the two parties' electoral strategies, I wanted to look at how well Democrats and Republicans performed in different congressional districts around the country. Inspired by a similar analysis done by Lakshaya Jain at the website Split Ticket, I decided to create a multivariate regression that evaluated the relationship between the demographic characteristics of a congressional district (race, education, and level of urbanization) and Biden's vote share in that district. This allowed me to examine where Biden and Trump performed better or worse than expected, providing some insight into how the two political parties should evaluate the strength of their coalitions going forward.

Data Collection and Model Methodology:

I collected the data needed to make the regression model from three main sources. First, I downloaded the presidential results by congressional district from Daily Kos Elections, which publishes this data every election cycle. I then used Python's requests library to query district-level race and education data from the US Census API. Lastly, I used City Lab's congressional district urbanization index as a measure of population density. I cleaned this data with Python's Pandas library to build the regression model. The values of the coefficients in the regression are in the table below. The R Squared value for this regression model is 0.754.

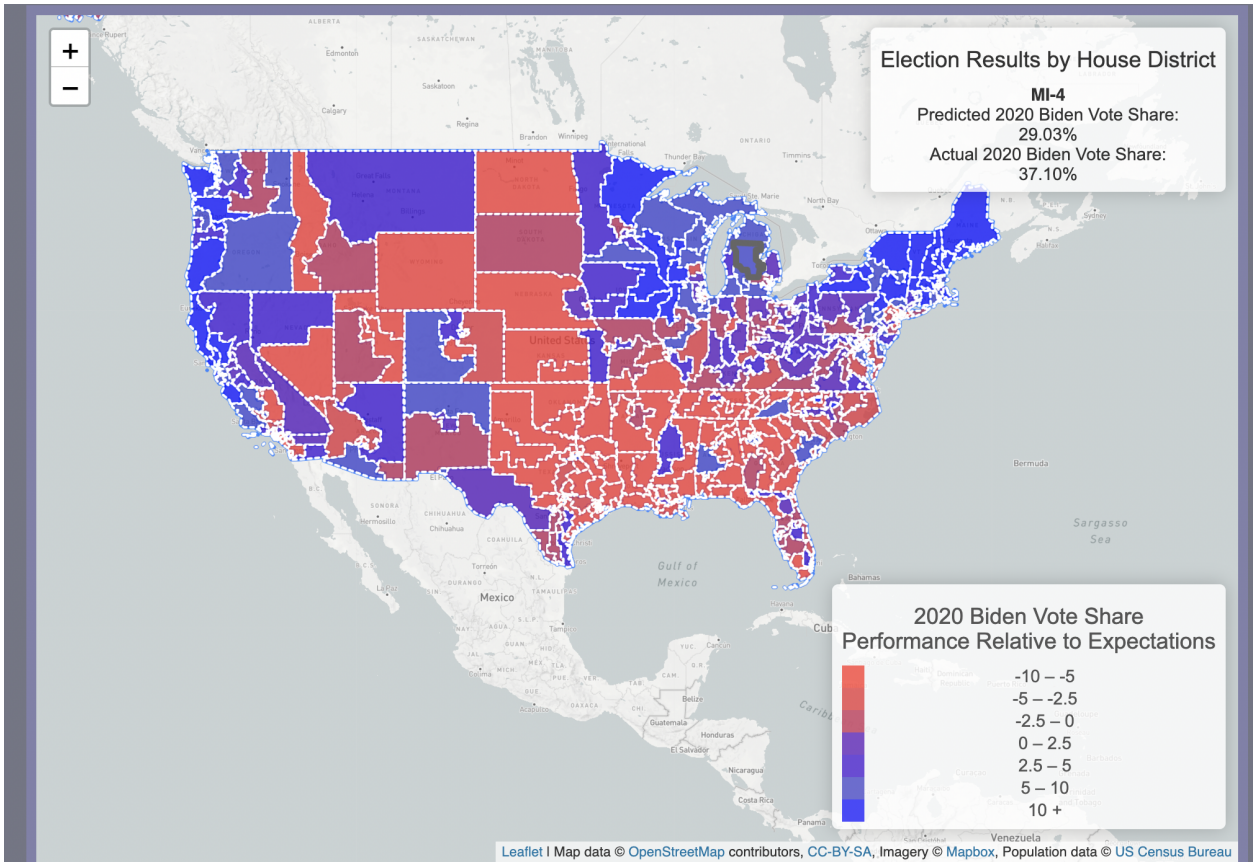
Variables	Coefficients	Standard Error	T Stat	P-Value	Lower 95%	Upper 95%	Significant?
constant	67.7089	14.806	4.573	0.000	38.607	96.811	Yes
white_non_college_pct_pop	-0.5778	0.155	-3.731	0.000	-0.882	-0.273	Yes
white_ba_pct_pop	0.1439	0.160	0.898	0.370	-0.171	0.459	No
black_pct_pop	0.0927	0.153	0.606	0.545	-0.208	0.393	No

hispanic_pct_pop	-0.2046	0.155	-1.323	0.187	-0.508	0.099	No
asian_pct_pop	-0.1681	0.186	-0.906	0.366	-0.533	0.197	No
City_Lab_Congressional_Index_Dense suburban	10.9477	1.719	6.368	0.000	7.569	14.327	Yes
City_Lab_Congressional_Index_Pure urban	21.4604	2.502	8.577	0.000	16.542	26.378	Yes
City_Lab_Congressional_Index_Rural-suburban mix	2.2625	1.282	1.765	0.078	-0.256	4.781	No
City_Lab_Congressional_Index_Sparse suburban	6.1349	1.599	3.837	0.000	2.992	9.277	Yes
City_Lab_Congressional_Index_Urban-suburban mix	14.9066	2.118	7.039	0.000	10.744	19.069	Yes

Note: The values of the city lab urbanization coefficients are based on using pure rural districts as a reference group.

Interpretation of this regression equation:

The main point of making this model was to look at the difference between the expected and actual vote shares of the two candidates in each district. To do this, I looked more at the residuals of the model than the values of the coefficients. The actual coefficients and p values are unlikely to be reliable because multiple variables in the model are highly correlated with one another, leading to a collinearity problem that biases these values. However, this does not affect the overall predictions of the model, which are still useful to look at. These predictions displayed the difference between Biden's expected and actual vote share based on the district's demographic makeup. The map below displays these residuals, with redder districts indicating a strong Republican overperformance and bluer districts indicating a strong Democratic overperformance. This allowed me to look at the strength of the parties in the different regions of the country.



The difference between expected and actual vote shares in the 2020 Presidential election.

Democrats are still overperforming in the Midwest relative to the rest of the country

After the 2016 election, there was a lot of discussion about what went wrong for Hilary Clinton in the Midwest. Many of the states in this region voted for Obama both in 2008 and 2012 by relatively large margins before voting for Trump in 2016. However, Biden was able to counteract this trend quite significantly in the states of Michigan and Wisconsin. Most of the districts in these two states were a darker shade of blue on the map above, indicating Democratic overperformance by at least 5 points. Looking at the table below, Biden averaged a 6.75 percent overperformance in each congressional district in Wisconsin and Michigan relative to what this regression model would predict.

District	Biden Actual Vote Share	Biden Expected Vote Share	Biden Performance Residual (Over/Under Performance)
WI-5	41.7	46.061428834832675	-4.361428834832672
MI-11	51.6	52.529025790689694	-0.9290257906896926

MI-8	48.8	49.22154881349383	-0.42154881349383544
MI-10	34.4	33.20487151808472	1.195128481915276
WI-1	44.7	43.47139720266157	1.2286027973384321
MI-14	79.5	75.23779797132447	4.2622020286755316
MI-9	55.9	51.54808559629428	4.35191440370572
MI-2	43.2	38.41457541982726	4.785424580172744
MI-3	47.4	42.10587399559714	5.294126004402855
WI-8	41.3	35.14193293000909	6.158067069990906
WI-6	41.6	35.39121663662299	6.208783363377009
WI-4	76.2	69.35524739663441	6.844752603365592
MI-4	37.1	29.02773878598709	8.072261214012912
MI-6	46.8	38.55549680252343	8.244503197476568
MI-7	41.4	32.90684049369766	8.49315950630234
MI-1	40.6	31.902766594975994	8.697233405024008
WI-7	39.3	29.915332230126385	9.384667769873612
MI-5	51.4	41.66556089023431	9.734439109765688
MI-13	78.8	66.79675147624732	12.003248523752674
MI-12	64.2	52.07001205794348	12.129987942056523
WI-3	46.8	31.44030171397755	15.359698286022446
WI-2	69.4	47.44819299747684	21.951807002523168
Averages	51.00454545	44.24599982	6.75854563

Given that Biden won Michigan and Wisconsin by less than three percentage points, this overperformance was the difference between Biden winning and losing these states. This data

reaffirms what other firms analyzing the election results have said about voters in Wisconsin and Michigan. For instance, Catalist, a Democratic polling firm, estimated that Biden won about 61% of white college voters and 40% of non-college white voters in Wisconsin. This was 7 points and 3 points higher respectively than the national average for these demographic groups, which tracks with Biden's overperformance in this state. Overall, Biden's strength in the Midwest was enough to help him carry these largely white states

Democrats are still underperforming in Texas and the Southeast

During most of my lifetime, it has been assumed that white voters in the south were reliably Republican. This was still mostly true in 2020; however, there have been signs that this is changing a bit in urban and suburban areas in this region. Democrats gained seats in suburban Dallas, Houston, Oklahoma City, Charleston, and Atlanta during the 2018 midterm election cycle and were able to hold a majority of these gained seats in 2020. A big reason for this was that white voters with degrees in these areas started voting more like their counterparts in northern states. With that said, Biden still performed much worse than the model would expect in this region. The table below contains the 20 congressional districts where Biden underperformed the most relative to the demographics of the district. 17 of these 20 congressional districts are in the Southeast or Texas.

District	Biden Actual Vote Share	Biden Expected Vote Share	Biden Residual (Underperformance)
NY-11	44.3	67.70146665053447	-23.401466650534474
FL-25	38.2	60.62085288273636	-22.420852882736355
TX-11	19.7	41.55510701771189	-21.85510701771189
TX-13	19.4	38.83611764434036	-19.436117644340364
TX-7	53.6	71.10205397247212	-17.50205397247212
TX-19	26.3	43.54183968933002	-17.241839689330018
FL-26	46.9	63.705854967217036	-16.805854967217037
GA-6	54.8	71.57470491829012	-16.774704918290126
AL-6	31.8	48.071222699477715	-16.271222699477715
TX-8	28.1	44.07001256262985	-15.970012562629847

NC-13	31.8	47.52024373158859	-15.720243731588592
UT-3	35.2	50.289910198532375	-15.089910198532372
TX-2	48.6	63.12655786473004	-14.526557864730037
TX-3	48.7	63.21484930314567	-14.514849303145667
GA-11	41.5	55.875088650583	-14.375088650583002
CA-48	49.7	63.970164143154264	-14.270164143154261
LA-1	30.1	44.12248306478711	-14.02248306478711
OK-2	22.1	36.103660980336734	-14.003660980336733
FL-27	51.3	65.25650020927826	-13.95650020927826
TX-1	27.2	40.70560141635394	-13.505601416353944
Averages	37.465	54.04821463	-16.58321463

This underperformance in many of these districts in the table above was the difference between Biden winning and losing in Texas and Florida. I thought it was particularly interesting to note that Biden still underperformed in GA-6 and TX-7 by approximately 17% even though both of these districts had flipped to Democrats in 2018, indicating there may be opportunities for Democrats to increase margins there in the future. To further this point, CNN exit polls had Biden winning 42% of white voters with a degree and 26% of white voters without a degree in Texas. These exit poll numbers were similar in Florida and Georgia among these groups as well, meaning that the underperformance shown in the model tracks well with the other data collected.

How should the two political parties think about this information?

As it stands currently, white voters in the North tend to vote more democratic than the national average and white voters in the South tend to vote more Republican than the national average. This poses an opportunity for both parties. If white voters without college degrees continue to swing towards the Republican party, it will be especially difficult for Democrats to win in states like Michigan and Wisconsin in the future. It was indeed the case that Biden overperformed by several percentage points to win these states, meaning that some mean reversion would give Republicans a good chance at winning them back. These states also have not been experiencing large amounts of population growth, meaning that it is less likely that an influx of different voters will massively change the potential coalition needed to win these states.

On the other hand, it is not out of the question that white voters with college degrees will continue to move towards Democrats as the parties continue to sort more along demographic lines, leaving Democrats with more opportunities in the South. If Republicans are unable to get over 55% of white college voters in states like Georgia and Texas, it may become harder for them to hold together the coalition they currently have that gave them a four percentage point electoral college advantage in 2020. Given that these states are growing and experiencing high levels of population growth, it is not out of the question that an influx of college graduates in Texas and the Southeast will vote more similarly to college graduates in other parts of the country, making these states more electorally competitive going forward. Overall, this data would suggest that the movement of white voters without degrees will continue to give the Republicans a growing edge in the electoral college; however, that advantage would swing back if Democrats limit their losses among white voters in the South and Texas.