The Question of the "Trump Effect": Basic Voting Characteristics, Economic Indicators, and Migration Flows in the 2016 Presidential Election

Pablo A. Ordóñez Bravo
Advisor: W. Bowman Cutter

Pomona College

Abstract— **This paper examines the correlations between basic voter characteristics, economic indicators, migration flows and the change in Republican presidential vote share in the 2016 election. Using county-level data from 2,784 counties and county-equivalents in the United States, the paper develops a regression analysis that finds these correlations do in fact exist. The correlations reveal that 1) basic voter characteristics have strong explanatory power, 2) poverty and unemployment have a statistically significant near-zero effect on voting behavior, and that 3) different immigrant characteristics have conflicting impacts on the Republican presidential vote. I conclude that an inexplicable "Trump effect" does not appear to exist, and that the 2016 election appears to be significantly understandable through regression analysis.**

I. Introduction

The 2016 United States presidential election, perhaps more so than others in recent memory, was subject to unexpected turns. Donald Trump's upset victory on election night was anticipated by few—not even Trump himself.[1] Many explanations for Trump's win have been advanced *ex post facto* by political analysts and pundits, including the notion of a "Trump effect": some broad characteristic of Donald Trump's victory that breaks all of the rules and is impervious to conventional explanation. As the deciding components behind the 2016 presidential election result remain largely unexamined by existing economic literature, my paper attempts to test these broader claims.

Econometric analysis through an economic voting framework appeared to be the most appropriate way to approach this question. Using county-level data, this study uses three broad categories to test for correlations: basic voting characteristics, economic indicators, and migration flows. In contrast to solely using the results of the 2016 presidential election, I chose to correlate to the change in voting shares from the previous election in 2012. Under this methodology, the coefficients of the corresponding independent variables work as a referendum on both the outgoing Democratic administration (based on the 2012 results) and the incoming Republican administration (based on the 2016 results). I also test the econometric model through two different dimensions—counties throughout the United States and counties in the top ten swing states. Doing so allowed me to better gauge the influence of these variables in different contexts.

Basic voting characteristics serve as a foundation for the model, and include the general voting descriptors most commonly associated with voting behavior. County-level economic indicators are also incorporated. In the face of the election's heightened level of discussion around immigration, I additionally test if changes in county migration flows can be tied to changes in the Republican presidential vote share from 2012 to 2016. This regression helps examine whether changes in immigration levels resulted in higher or lower levels of support for Donald Trump.

The results for these regressions were statistically significant across several dimensions. The basic voting controls proved to be in line with results across other literature, and explained voter behavior to a large degree. These results demonstrate that race and education played a significant role in the 2016 election, revealing a tilt against the Republican presidential vote share by a large and small margin, respectively. The economic indicators, when statistically significant, revealed a zero or near-zero effect on the share of the Republican presidential vote. This was a surprising result and has certain implications about the effects of poverty and unemployment on voting behavior. Lastly, migration flows, when statistically significant, showed three results, one of which was unexpected. First, the regressions revealed that an increase in a county's domestic inflow was tied with a lower share of Republican presidential votes. Second, an increase in international migration was also tied to a lower share of Republican votes. Third, and most surprising, an increase in a county's population of foreign-born people—immigrants—had a slightly positive correlation to the share of Republican votes. These results provide some preliminary insights into how different indicators reveal the potential for different political attitudes among immigrants.

The paper's most meaningful result is the clear existence of a statistical correlation between voter characteristics and the change in the 2016 Republican presidential vote share from 2012. This analysis strongly indicates that the 2016 presidential election is not inexplicable and can in fact be understood fairly well through standard econometric analy-

[1] Jennifer Jacobs and Billy House, "Trump Says He Expected to Lose Election Because of Poll Results," Bloomberg.com, December 13, 2016, accessed July 08, 2017, https://www.bloomberg.com/news/articles/2016-12-14/trump-says-he-expected-to-lose-election-because-of-poll-results.

sis. Contrary to notions advanced by pundits and analysts, a "Trump effect" does not necessarily appear to exist. Through the econometric analysis of basic voting characteristics, economic indicators, and migration flows, it becomes evident that the election, like any other, can be explained and understood through patterns in numbers.

II. Literature Review & Justification

Much literature exists on the effect of economic conditions on American voting behavior during presidential elections. This field is broadly known as "economic voting" and analyzes voting as "straightforward political demand for the amelioration of economic grievances" (Weatherford, 1983). A common starting point in economic voting literature is Ray Fair's "The Effect of Economic Events on Votes for President" (1978). Fair's model attempts to connect different economic factors to how votes might be distributed in a presidential election. His paper uses various economic specifications and concludes that "change in real economic activity [does] appear to have an important effect on votes for president" (Fair, 1978). Bartels (1997) claims that this conclusion has a consensus, such that "the clearest and most significant implication of aggregate election analyses is that objective economic conditions...are the single most important influence upon [voting shares]." Despite general agreement on the significance of economic conditions in elections, economic voting models are imperfect—Fair's model, in fact, when tested with various election results, often comes up short (Bartels, 1997). Some authors attribute this deficiency to a degree of "specification uncertainty" that they claim is inherent to the study of presidential elections (Leamer, 1978). In the face of this probable uncertainty when studying economic voting, models must be thus guided by both goodness-of-fit concerns as well as *a priori* political and economic considerations[2] (Bartels, 1997). Scholars can thus take a variety of considerations into account when building economic voting models. This latitude is helpful for exploring new possibilities, as once having accounted for a given set of consensus economic considerations, researchers can explore all justifiable *a priori* considerations through econometric examination in relation to presidential voting outcomes. My economic voting model, like others, must too define those parameters by which it treats "political demand" and "economic grievances."

I define political demand by directly measuring county-level changes in Republican presidential voting shares from 2012 to 2016. This definition is conventional and appears in different variations through the literature. As an example, DeSimone and LaFountain (2007) define political demand as the share of votes for President George W. Bush in the 2004 presidential election. While my definition for political demand is standard, my definition for economic grievances requires greater justification. Unlike political demand, no clear consensus definition exists for economic grievances.

Before selecting variables to define economic grievances, however, the class of economic data used in this paper is important to justify. In economic voting, two broad data methodologies are widely used by scholars. One body of scholarship advocates for the use of "egocentric voting,"[3] which uses survey data to capture voters' personal opinions about their economic conditions. Using this type of data, researchers can measure economic activity by placing polls and asking voters questions about their financial realities. DeSimone and LaFountain (2007) used egocentric voting to ask voters whether they felt that they were better off, worse off, or equally well-off as they had been four years ago. Egocentric voting presumes that how voters feel about their own personal economic conditions is the greatest determinant of how they vote. Another body of scholarship adheres to "sociotropic voting," which posits that concrete economic conditions are better correlated to actual voting behaviors. Sociotropic voting advocates for the use of tangible economic factors because voters are more correctly informed by macroeconomic fluctuations (Kinder & Kiewiet, 1979). Other scholars also concur that voters are sophisticated enough to utilize "objective economic information and forecasts" to some capacity (Holbrook & Garand, 1996).

The model discussed in this paper uses a sociotropic voting framework. Centered on voter opinion as opposed to tangible economic data, egocentric voting may not measure economic changes that voters face to a sufficiently reliable degree. While economic data may not be fine-grained enough to discern the economic fortunes of individual voters, hard economic indicators are likely still more accurate than voter opinions that have been subjected to a variety of biases.[4] As survey data may be too reliant on context and the underlying presumptions of voters, and can suffer the risks of skewed conclusions or the inability to replicate results, this paper instead examines the existence of strictly numerical relationships between economic conditions and voting behaviors. Hard economic indicators, while imperfect in accuracy, are numerical certainties that to a great degree reach beyond opinion bias. This model seeks not to "rely on survey data, but rather examine the impact of economic conditions on the true outcome" (Brunner et al., 2008).

My chosen variables for "economic grievances" are greatly informed by Bloom and Price (1975), whose focus on sociotropic economic conditions attempts to correlate voting shares as a variable dependent on party identification and the percent change of real per capita income. While Bloom and Price's definition is not a wholly apt fit for this paper, the emphasis on income informs my model and provides a solid foundation expanded upon by the inclusion of other *a priori* conditions that I test on voter behavior. This paper is novel as it places variables examined individually across other papers in the literature together in a unified model.

Among the variables examined in this paper are education and unemployment. While these variables have previously

[2]*A priori* here refers to considerations that are derived from theoretical deduction as opposed to statistical convention or practice.

[3]In the literature, egocentric voting is also referred to as "pocketbook voting."

[4]Bowman Cutter commented on an earlier draft of this claim.

been explored, evidence of their effects on voting in the literature has been mixed. Education is often linked to other economic markers (Bailey & Dynarski, 2011). Gelman & Su (2010), however, suggest that "despite patterns we see for education levels, voting by income remains strongly patterned along traditional lines." Higher levels of unemployment has been linked to higher levels of voter turnout and consequent voter share (Burden & Wichowsky, 2012). Other works, by contrast, suggest that unemployment might have potentially negative or neutral effects (Rosenstone 1982). As little on the effects of these variables appears to have been settled, education and unemployment should be examined carefully.

In addition to customary variables, my paper attempts to further contribute to the literature by considering 2016 presidential election voting in relation to migration flows. The inclusion of migration flows presents a departure from mainstream inquiry. Though scholars have recently begun to explore the potential effects of global movement on voting, their current explorations have primarily concerned the movement of goods rather than people. Jensen et al.'s "Winners and Losers in International Trade: The Effects on U.S. Presidential Voting" (2016) examines the "employment effects of expanding ports" on changing vote shares with a strictly focus on trade flows. Autor et al. (2016) similarly considers labor markets with varying levels of trade exposure in relation to voting behaviors. While models are not directly applicable to this paper, they inform its development with regard to voting behavior, economic conditions, and global movements.

It is worth noting that at the time of this writing, few papers have been published on the 2016 presidential election. Though political scientists and economists alike often wait until data is more complete, part of the pause may be able to be attributed to Donald Trump's astounding upset. Trump's victory confounded many existing economic and political conventions, and scholars are currently determining the degree to which Trump's win invalidates some of those considerations. This paper is in part motivated by this problem—I believe it is worth exploring an economic voting model with migration flows that may help explain the increase in vote share for Trump that won him the election.

III. Methodology

The economic voting model used in this paper considers multivariable linear regressions that examine cross-sectional demographic data. This approach involves study of county-level economic variables from 2010 to 2015 that together build an approximate image of the correlations between economic considerations and votes during the 2012 and 2016 presidential elections. The basic model, estimated through robust regression, is the following:

$$\begin{aligned}\Delta\% \textit{ Republican Vote Share} &= \beta_0 + \beta_1(\%\textit{Race}) \\ &+ \beta_2(\%\textit{Equation}) + \beta_3(\%\textit{Female}) \\ &+ \beta_4(\%\textit{Household Income}) + \beta_5(\%\textit{Age}) \\ &+ \beta_6(\%\textit{Population Density}) + \mu\end{aligned}$$

This simple form of the model best captures the broader theoretical framework this study applies. It delineates the basic voting controls the literature has collectively agreed on having the strongest correlation on voting in presidential elections. These "consensus economic considerations" allow for the accurate examination of a priori factors like migration flows. If the model did not include these consensus factors, the coefficients for the a priori factors may erroneously draw explanatory power from the exclusion of these well-documented, existing correlation. This study thus attempts to include "controls that are particularly linked to proxy for voter preferences and therefore absorb their spurious link with reported changes in economic well-being." (DeSimone & LaFountain 2007).

Many economic voting models that examine changes between presidential elections often include the share of votes from the previous election as an independent variable in their models. I have chosen not to do this. I argue that its inclusion does not necessarily help expand the models true correlations but would instead skew the model towards a disingenuously higher R-squared value. As this skewed R-squared value may convey more statistical significance than might plausibly exist, in the interest of accuracy I have avoided this practice.

Though Equation (1) is helpful for understanding this papers framework, it is insufficient in expressing the statistical nuance regression analysis requires. Equation (2) is thus expanded to include the models specifications as written in the data model:

$$\begin{aligned}\textit{Republican\_Change\_Perc} &= \beta_0 + \beta_1\textit{(Black\_Perc)} \\ &+ \beta_2\textit{(AmIndian\_Perc)} + \beta_3\textit{(Asian\_Perc)} \\ &+ \beta_4\textit{(PacificIslander\_Perc)} + \beta_5\textit{(Latino\_Perc)} \\ &+ \beta_6\textit{(Bachelors\_Perc)} + \beta_7\textit{(Fem\_Perc)} \\ &+ \beta_8\textit{(Bachelors\_Perc)} + \beta_9\textit{(perc017)} + \beta_{10}\textit{(perc1824)} \\ &+ \beta_{11}\textit{(perc2539)} + \beta_{12}\textit{(perc4054)} + \beta_{13}\textit{(perc65plus)} \\ &+ \beta_{14}\textit{(Pop\_Density)} + \mu\end{aligned}$$

$Republican\_Change\_Perc$ represents the change in the percent of Republican presidential vote from 2012 to 2016. $Black\_Perc$, Am_Indian_Perc, $Asian\_Perc$, $PacificIslander\_Perc$, and $Latino\_Perc$ represent the yearly reported percentages of each racial group by county. I have excluded Non-Hispanic white individuals as they represent the largest racial group and are part of the models constants. $Bachelors\_Perc$ is the average yearly percent of people that hold Bachelors degrees in each county. Fem_Perc is the yearly percent of women in each county. $Med\_Household\_Income$ is the yearly median household income in each county. The variables $perc017$, $perc1824$, $perc2539$, $perc4054$, and $perc65plus$ indicate the percentages of people in each county aged 0 to 17, 18 to 24, 25 to 39, 40 to 54, and 65+. I have excluded people ages 55 to 64 as they represent one of the largest population sizes and are also part of the models constants. $Pop\_Density$ is each countys yearly population density, calculated by dividing a countys population by its land area.

Having controlled for basic voting variables, the model then expands to include *a priori* economic conditions examined under an sociotropic voting framework. The model is augmented as thus:

$$\begin{aligned}\Delta\% \textit{ Republican Vote Share} &= \beta_0 + \beta_1(X)\\ &+ \beta_2(\% \textit{ Poverty}) + \beta_3(\% \textit{ Civilian Labor Force})\\ &+ \beta_4(\% \textit{ Unemployment}) + \mu\end{aligned}$$

The X variable represents the variables previously presented in Equation (1). The other elements are chosen variables to represent the potential economic conditions that may be correlated to voting. A poverty variable is included to get a sense of the countys wealth distribution in relation to voter behavior. Civilian labor force and the unemployment rate are included to serve as parameters to measure the change in each countys labor force. An expanded form of Equation (3) that includes the models statistical specifications is written as follows:

$$\begin{aligned}\textit{Republican\_Change\_Perc} &= \beta_0 + \beta_1(X)\\ &+ \beta_2(\textit{Poverty\_Perc}) + \beta_3(\textit{LaborForce\_Perc})\\ &+ \beta_4(\textit{Unemployment\_Rate}) + \mu\end{aligned}$$

Here, $Republican\_Change\_Perc$ and X have the same definitions as listed previously in Equation (2). $Poverty\_Perc$ is the yearly percent of people in poverty per county, $LaborForce\_Perc$ represents the yearly percent of people in the civilian labor force per county, and $Unemployment\_Rate$ is the yearly unemployment rate by county as calculated by BLS.

Apart from the consensus specifications explored in Equation (1) and the economic indicators expressed in Equation (3), this paper also explores the correlation of migration flows to changes in presidential voting. Consequently, another equation explores these migration flows. This equation is expressed as follows:

$$\begin{aligned}\Delta\%\textit{Republican Vote Share} &= \beta_0 + \beta_1(X)\\ &+ \beta_2(\Delta\% \textit{ Domestic Migration})\\ &+ \beta_3(\Delta\% \textit{ International Migration})\\ &+ \beta_4(\% \textit{ Foreign Born Population}) + \mu\end{aligned}$$

As with Equation (3), I have kept the consensus indicators from Equation (1) and consider them in concert with newly introduced migration variables. My model considers domestic and international migration separately, as adding both flows into a single yearly migration variable could potentially occlude separate, individual correlations that domestic and international migration might have. To examine the degree to which non-domestic immigrants are a part of the communities in these counties, the model also includes a foreign-born population variable to measure how many people per county are foreign-born. Though there is some risk of collinearity with the foreign migration variable, I argue it is worth including to potentially catch separate correlations.

Statistically specified, this equation is written as such:

$$\begin{aligned}\textit{Republican\_Change\_Perc} &= \beta_0 + \beta_1(X)\\ &+ \beta_2(\textit{Change\_Dom\_Perc}) + \beta_3(\textit{Change\_Int\_Perc})\\ &+ \beta_4(\textit{ForeignBorn\_Perc}) + \mu\end{aligned}$$

where $Republican\_Change\_Perc$ and X have the same definitions as listed previously, $ForeignBorn\_Perc$ is the average yearly percent of foreign-born people living in each county, and $Change\_Dom\_Perc$ and $Change\_Int\_Perc$ represent the 2010 to 2015 respective change in county domestic and international migration.

My final equation examines whether the correlations found in my previous equations hold when considered all together. Succinctly, this is expressed as:

$$\begin{aligned}\Delta\%\textit{Republican Vote Share} &= \beta_0\\ &+ \beta_1(X) + \beta_2(Y) + \beta_3(Z) + \mu\end{aligned}$$

where X represents the basic voting controls from Equation (2), Y represents the economic conditions listed in Equation (4), and Z represents the migration flow variables from Equation (5).

Altogether the complete statistically specified equation in the model is:

$$\begin{aligned}\textit{Republican\_Change\_Perc} &= \beta_0 + \beta_1\textit{(Black\_Perc)}\\ &+ \beta_2\textit{(AmIndian\_Perc)} + \beta_3\textit{(Asian\_Perc)}\\ &+ \beta_4\textit{(PacificIslander\_Perc)} + \beta_5\textit{(Latino\_Perc)}\\ &+ \beta_6\textit{(Bachelors\_Perc)} + \beta_7\textit{(Fem\_Perc)}\\ &+ \beta_8\textit{(Bachelors\_Perc)} + \beta_9\textit{(perc017)} + \beta_{10}\textit{(perc1824)}\\ &+ \beta_{11}\textit{(perc2539)} + \beta_{12}\textit{(perc4054)} + \beta_{13}\textit{(perc65plus)}\\ &+ \beta_{14}\textit{(Pop\_Density)} + \beta_{15}(\textit{Poverty\_Perc})\\ &+ \beta_{16}(\textit{LaborForce\_Perc}) + \beta_{17}(\textit{Unemployment\_Rate})\\ &+ \beta_{18}(\textit{Change\_Dom\_Perc}) + \beta_{19}(\textit{Change\_Int\_Perc})\\ &+ \beta_{20}(\textit{ForeignBorn\_Perc}) + \mu\end{aligned}$$

in which all variables in this equation retain their previous definitions.

IV. Analysis

The 2016 presidential election results used in this model are drawn from the work of entrepreneur Gary Hoover, who compiled data sets from Michael Kearney's open-source election night results. The data presents the presidential voting results of 3,112 counties and county-equivalents in the United States. Kearney compiled the election night voting tallies from official sources for the two major-party nominees, Sec. Hillary R. Clinton (D) and Mr. Donald J. Trump (R). As these results are from the election night, they do not capture the additional Democratic votes that led to Clinton's 2.86 million vote surplus in the popular vote.[5] For the purposes of this paper, however, these missed votes are not consequential. Clinton's additional votes were overwhelmingly from large counties that had already voted in

[5] https://transition.fec.gov/pubrec/fe2016/2016presgeresults.pdf

her favor (e.g. Los Angeles County). As my paper weighs the data by county population size, my paper is not significantly affected. The dataset remains a reliable representation of the 2016 presidential election results. Also used in this model are the official results from the 2012 presidential election between Pres. Barack H. Obama (D) and Gov. W. Mitt Romney (R). The availability of both 2012 and 2016 presidential election results allow me to calculate the change in vote percentage.

The demographic datasets are compiled from the U.S. Department of Agriculture's Economic Research Service and the U.S. Census Bureau's 2010-2014 American Community Surveys (ACS). These data broadly cover population statistics such as population size, density, race, gender, education, and age. Economic statistics such as median household income, poverty level, unemployment rate, civilian labor force, and migration flows are also drawn from this dataset. Finally, the number of average yearly number of foreign-born persons per county was provided by the Migration Policy Institute after a brief e-mail exchange.

Of the original 3,112 counties in the unified dataset, 328 were removed to facilitate analysis. A small number of counties were removed due to missing or inaccurate data across different statistics. Two states, Alaska and Delaware, were removed because the county data available for them were not reliable. The remaining removals were performed to balance the dataset. The United States has a number of counties at the extremes of the population distribution. Their inclusion would have resulted in inaccurate coefficients. Accordingly, counties in the highest and lowest five percent of the population distribution were removed from the dataset. After these removals, the number of counties observed was reduced to a final number of 2,784. An additional step was taken to limit the influence of extreme values across the dataset through weighting all summary statistics and regression analysis by population size. All statistics are rounded to the nearest tenth decimal.

Table 1 above shows the basic voting controls from Equation (1). I have included most demographic indicators but have withheld age and race to detail those in Tables 2 and 3. The variables presented have been weighed by county population and thus represent the values found after adjusting for differences in population concentration. Similar to mechanisms such as the Electoral College, adjusting for differences in population concentration decreases the weight of urban populations and increases the weight of rural populations. This adjustment is performed to better capture the effect that adding a given number of people with a certain characteristic may have on the voting share in a given county. As Republican candidates have generally outperformed Democratic candidates in rural areas during recent presidential elections, both Mr. Trump in 2016 and Gov. Romney in 2012 perform better than their Democratic counterparts weighted by county, although both candidates lost the popular vote. Republicans, however, increased their presidential voting share from 2012 to 2016 by about 1.6%, which is consistent with the weighted results seen.

Table 1. General characteristics by county.

VARIABLES	(1) N	(2) Mean	(3) Standard Deviation	(4) Min	(5) Max
1. Approx. Voting Results					
Δ% Republican Vote Share***	2,784	1.64	0.56	-3.6	2.3
2016 Presidential Election					
Sec. Hillary R. Clinton (D)	2,784	28,977	29,087	96	133,833
Mr. Donald J. Trump (R)	2,784	34,871	27,664	256	145,519
2012 Presidential Election					
Pres. Barack H. Obama (D)	2,784	29,918	28,829	119	129,229
Gov. W. Mitt Romney (R)	2,784	33,183	26,918	179	161,567
2. General Demographics					
Population Size	2,784	155,630	122,482	2,954	446,753
Population Density	2,784	336.7	663.9	0.474	10,214
Median Household Income	2,784	51,587	13,728	22,640	125,635
% Female	2,784	50.0	1.77	27.4	60.4
% with Bachelor's or higher	2,784	24.68	10.02	5.847	75.09

*Observations have been weighted by county population size in 2015, with a weighting value of 153,030,410.

**Unless otherwise noted, values represent a yearly average or an approximate estimate.

***Change here is measured from the 2012 and 2016 presidential elections.

The average weighted county in this dataset has a population of 155,630. Mean median household income per county is close in line with the national average of $51,587. The number of female persons per county is also near the national average at 50%. Although the mean proportion of individuals with bachelor's degrees is approximately 25%, there is great variance across counties—some counties have as few as 5.9% with a four-year degree whereas others up to 75.2% of the applicable population holds a four-year degree.

Table 2. Percentages of race by county.

VARIABLES	(1) N	(2) Mean	(3) Standard Deviation	(4) Min	(5) Max
% Non-Hispanic White	2,784	71.6	19.8	1.06	100.0
% Black or African-American	2,784	10.4	13.4	0	86.9
% American Indian or Alaska Native	2,784	1.04	4.19	0	91.3
% Asian	2,784	1.98	2.54	0	34.6
% Native Hawaiian or other Pacific Islander	2,784	0.01	0.06	0	12.5
% Hispanic or Latino	2,784	9.22	11.7	0	93.2

*Observations have been weighted by county population size in 2015, with a weighting value of 153,030,410.

**Names of categories taken from U.S. Census Bureau classifications.

Table 2 above displays the racial distribution that was omitted from Table 1. Non-Hispanic white individuals are the largest racial category, comprising an average of 71.60% of a county's population when weighed by population size.

On average, Black or African-Americans comprise 10.40%, Latinos 9.22%, Asian people 1.98%, Native Americans or Alaska Natives 1.04%, and Native Hawaiians or Pacific Islanders 0.01% of a county's population. In several counties, there is very high racial homogeneity. Many counties have an almost exclusively Non-Hispanic white population, whereas in others Black Americans, Native American, and Latinos represent as much as 90% of a county on average.

Table 3. Age percentages by county.

VARIABLES	(1) N	(2) Mean	(3) Standard Deviation	(4) Min	(5) Max
% 0-17	2,784	23.0	3.09	7.75	40.3
% 18-24	2,784	10.1	3.90	1.41	58.3
% 25-39	2,784	18.3	2.56	6.5	36.0
% 40-54	2,784	20.2	2.08	5.95	28.3
% 55-64	2,784	13.1	1.81	3.79	26.1
% 65+	2,784	15.4	4.11	3.30	50.9

*Observations have been weighted by county population size in 2015, with a weighting value of 153,030,410.

Table 3 above presents a similar breakdown like Table 2 but with population by age. People aged 0 to 17 are most represented, comprising on average 23% of a county's population when weighted by population size. The next largest group is people aged 40 to 54, at 20.2% of a typical county's population. Some areas have very high proportions of senior citizens, with individuals over age 65 comprising up to 50.9% of the population in the oldest counties.

Beyond the basic characteristics from Equation (1) represented in Tables 1, 2, and 3, I also examine the economic sociotropic metrics from Equation (2) in Table 4.

Table 4. Economic sociotropic metrics by county.

VARIABLES	(1) N	(2) Mean	(3) Standard Deviation	(4) Min	(5) Max
% in Poverty	2,784	15.7	5.75	3.2	47.4
% in Labor Force	2,784	47.7	5.46	19.3	87.1
% in Unemployment	2,784	5.40	1.63	1.8	21.8

*Observations have been weighted by county population size in 2015, with a weighting value of 153,030,410.

Controlling by population size, an average of 15.7% of people are in poverty and 47.7% are in the civilian labor force by county. Civilian labor force is presented alongside the unemployment rate to provide a better measure of the status of a county's workforce. The United States Census Bureau's Supplemental Poverty Measure (SPM) is widely considered a more accurate indicator of poverty than the official poverty rate as it calculates cost-of-living using a more comprehensive set of factors, including the cost of housing.[6] Unfortunately, this study did not have access to county-level SPM data.

Table 5. Migration flows by county.

VARIABLES	(1) N	(2) Mean	(3) Standard Deviation	(4) Min	(5) Max
% Foreign-born	2,784	5.78	4.99	0	40.8
Δ% Domestic Migration	2,784	0.07	0.97	-6.7	21.4
Δ% International Migration	2,784	0.14	0.15	-0.12	1.67

*Observations have been weighted by county population size in 2015, with a weighting value of 153,030,410.

Finally, the model examines the migration flows from Equation (3), represented above in Table 5. An average of 5.78% of a county's population was foreign-born in this dataset. This statistic, however, has great variance—some counties have almost no foreign-born residents, while in others nearly 40.8% of the population was born outside the country. This variance extends to change in domestic migration. Though on average a county experienced a 0.07% change in domestic migration, some counties saw changes as varied as -6.7% and 21.4%. In contrast, much lower variance existed for international migration, with counties only experiencing variance between -0.12% and 1.67%.

V. Results

In analyzing the relationship between the variables described and changes in presidential voting share, I ran four corresponding regressions on the complete dataset of county data available. For the data analysis, I used analytic weights by population size to neutralize the influence of disproportionately populated counties. As this paper concerns county-level effects, regressions were run with fixed effects to absorb the influence of broader state correlations. Further, robust regressions were used instead of Ordinary Least Squares (OLS) regressions. The OLS method is traditionally very sensitive to data outliers, and in the context of the often-volatile indicators in my county-level data robust regressions was the better technical choice. Finally, as the regressions evaluate twenty independent variables, I have additionally included adjusted R-Squared values. Adjusted R-Squared values decrease if independent variables do not possess explanatory power, so they prove useful to determining whether the variables examined are statistically significant.

My model's analysis can be seen on the following page, in Table 6.

[6]US Census Bureau, Data Integration Division, "Poverty - Experimental Measures," Supplemental Poverty Measure Latest Research - U.S Census Bureau, accessed June 27, 2017, https://www.census.gov/hhes/povmeas/methodology/supplemental/research.html.

Table 6. Regression results for every county in the dataset.

VARIABLES	(1) Δ% Republican Vote Share	(2) Δ% Republican Vote Share	(3) Δ% Republican Vote Share	(4) Δ% Republican Vote Share
% Black	-3.45***	-5.15***	-3.58***	-5.20***
	(-4.996)	(-6.457)	(-5.142)	(-6.478)
% American Indian or Alaska Native	2.29	-0.44	2.11	-0.53
	(0.908)	(-0.174)	(0.846)	(-0.209)
% Asian	-0.32	-4.07	2.53	-1.64
	(-0.047)	(-0.703)	(0.335)	(-0.249)
% Native Hawaiian or Pacific Islander	25.22	28.89	22.31	27.67
	(0.739)	(0.856)	(0.674)	(0.830)
% Hispanic or Latino	-4.15***	-5.09***	-3.01**	-3.93***
	(-4.015)	(-5.469)	(-2.428)	(-3.366)
% with Bachelor's degree or higher	-0.35***	-0.34***	-0.34***	-0.34***
	(-20.572)	(-20.886)	(-19.604)	(-19.962)
% Female	7.40	12.28***	3.59	9.10*
	(1.600)	(2.657)	(0.754)	(1.818)
Median Household Income	-0.00	0.00	-0.00	0.00
	(-1.514)	(1.303)	(-1.525)	(1.426)
% Aged 0-17	-36.75***	-41.36***	-34.03***	-37.67***
	(-4.314)	(-4.752)	(-3.966)	(-4.263)
% Aged 18-24	-10.35	-15.15*	-9.42	-13.64
	(-1.285)	(-1.817)	(-1.168)	(-1.631)
% Aged 25-39	-8.33	-7.96	-8.83	-7.28
	(-0.913)	(-0.857)	(-0.956)	(-0.780)
% Aged 40-54	6.98	1.75	9.74	5.56
	(0.576)	(0.145)	(0.808)	(0.465)
% Aged 65+	-5.64	-10.66	-2.45	-5.97
	(-0.510)	(-0.925)	(-0.223)	(-0.520)
Population Density	-0.00	-0.00	-0.00	-0.00
	(-0.891)	(-1.344)	(-0.627)	(-1.173)
% in Poverty		0.00***		0.00***
		(2.901)		(3.150)
% in Civilian Labor Force		-3.97		-2.48
		(-1.558)		(-0.957)
% Unemployed		0.00		0.00
		(0.807)		(0.919)
Δ% Domestic Migration			-0.19**	-0.11
			(-2.013)	(-1.195)
Δ% International Migration			0.36	0.81
			(0.377)	(0.850)
Δ% Foreign-born			-0.05	-0.06*
			(-1.505)	(-1.782)
Constant	0.19**	0.17*	0.19**	0.15*
	(2.246)	(1.937)	(2.256)	(1.720)
Observations	2,784	2,784	2,784	2,784
R-squared [adjusted]	0.838 [0.834]	0.842 [0.838]	0.839 [0.835]	0.842 [0.838]

†Robust t-statistics in parentheses. 48 categories of fixed effects absorbed. *** p<0.01, ** p<0.05, * p<0.1

The first regression corresponds to Equation (1) and concerns only the basic voting controls. The variables in Equation (1), per the literature, have often demonstrated a strong correlation to voting behavior. Regression (1) thus serves as a base that exhibits the explanatory power of the "consensus economic considerations" that I mentioned earlier in the paper. This base regression enables a clear distinction between those consensus variables and the *a priori* considerations I have additionally proposed. This clear distinction further allows measurement of the degree to which the *a priori* variables, presented in Equations (2) and (3), possess explanatory power for voting behavior.

Regression (1) strongly supports the existence of a statistical relationship between the characteristics of voters and presidential voting behavior. Though the literature has consistently supported this conclusion, it is important that this model arrives at this conclusion individually. Regression (1)'s value of 0.838 demonstrates very strong support for the significance of the relationship. The regression's adjusted R-Squared value, at 0.834, reaches a similar conclusion. These results imply that the basic voter controls included hold substantial explanatory power on voting behavior.

Many of Regression (1)'s independent variables have high t-stat values that suggest statistical significance. Increases in the percentage of African Americans, Latinos, and individuals with a bachelor's degree or higher result in a decrease to the share of Republican presidential vote. These results are consistent with conventional wisdom and other results in the literature.

Regression (2) maintains the basic voting controls of Equation (1) but adds economic indicators from Equation (2). The addition of these economic indicators appears to strengthen the relationship with the change in Republican presidential vote share, with both coefficients and t-stats increasing significantly. The coefficient for percentage of African Americans, for instance, rises from -3.45 to -5.15, and its t-stat from -4.996 to -6.457. The addition of these economic indicators also appears to be an improvement to the strength of the model's statistical correlation. Regression (2)'s R-Squared value increases to 0.842 from the 0.838 of Regression (1). Its adjusted R-Squared value increases in parallel—from 0.834 to 0.838. Three additional variables also gain statistical significance. The percentage of female persons and the share of people in poverty are statistically significant at the 99% confidence level, and the percentage of persons aged 18-24 is now significant at the 90% confidence level.

Regression (3) keeps the basic voting controls from Equation (1) and adds the migration flows from Equation (3). The addition of migration flows appeared to strengthen the relationship with voting behavior, but to a smaller extent than the addition of economic indicators in Regression (2). Regression (3)'s R-Squared and adjusted R-Squared values increased to 0.839 and 0.835 respectively. The majority of coefficients and t-stats in Regression (3) increase slightly from Regression (1)'s values, but none by more than the results of Regression (2). The exceptions are percentage of Latinos, which remains statistically significant but experiences a decreased coefficient, and the percent of female persons and persons aged 18-24, which become statistically significant and experience decreased coefficients compared to the results of Regression (1). Regression (3) suggests that while international migration is not statistically significant, domestic migration is at the 95% confidence level with a coefficient of -0.19.

Finally, Regression (4) keeps the basic voting controls from Equation (1), the economic indicators from Equation (2), and the migration flows from Equation (3). The relationship with presidential voting behavior is no stronger than the results found in Regression (2), with the same R-Squared and adjusted R-Squared values of 0.842 and 0.838. These results suggest that the addition of migration flows in Regression (4) does little to hinder or bolster the model with economic indicators from Regression (2). Regression (4), however, presents some deviations in independent variable statistical significance. In contrast to Regression (2), the percentage of female persons loses statistical significance and is now only significant to the 90% confidence level instead of the 99% level. In contrast to Regression (3), domestic migration is no longer statistically significant and the percent of foreign-born people is now significant in the 90% confidence level. With the complete model tested in Regression (4), it becomes clear that Equation (2)'s economic indicators add the greatest degree of explanatory power to the basic voting controls and Equation (3)'s migration flows do not significantly affect the strength of the correlation with changes in the Republican share of the presidential vote.

Table 6 considers 2,784 counties over 48 states.[7] This dataset is effective in conveying the broader implications of the model and the general effects of the independent variables. While this big picture is helpful, it is also worth exploring other more specific datasets that might reveal more acute insights.

In the context of the last presidential election, and given that the model's dependent variable is the change in the Republican presidential vote, it follows that a narrower dataset examining states most likely to be impacted by marginal changes in the vote should also be considered. In this study, these "swing states" are defined as those that have been pivotal to the results of the election such that 1) both candidates in recent elections have made consistent efforts to win them and 2) they are typically won by a small margin. For the purposes of this paper, ten states have been chosen that fit these qualifications: Colorado, Florida, Iowa, North Carolina, New Hampshire, Ohio, Pennsylvania, Virginia, Nevada, and Wisconsin. These ten states have been defined as swing states elsewhere in the literature (Jensen et al. 2012).

Narrowing the dataset to 665 counties within these ten states, I have again run my model's four regressions. As with the general dataset, I have used analytic weights by

[7] As I mentioned in the Data Analysis section, the two states that are not counted and have been removed for logistical purposes from the dataset are Delaware and Alaska.

population size and have run robust regressions with fixed effects. Both R-Squared and adjusted R-Squared values have again been included. The analysis of the ten swing states in the dataset can be seen on the following page, in Table 7:

The first regression corresponds to Equation (1)'s basic voting controls. The statistical relationship between these controls and presidential voting behavior still holds. Regression (1) exhibits consistently high R-Squared and adjusted R-Squared values, at 0.825 and 0.819, respectively. In this secondary analysis, basic voting controls again hold substantial explanatory power. As with the first analysis, this level of explanatory power supports the legitimacy of the relationships established by the following regressions which build on Equation (1).

Similar to Regression (1) in the first analysis, many independent variables hold high coefficients and are statistically significant to the 99% confidence level. Some differences are apparent, however. In this regression, the percent of Native Americans or Alaskan Natives gains significant statistical legitimacy. Additionally, the percent of persons aged 0-17, which was previously significant to the 99% confidence level in Table 6, is now only significant to the 90% confidence level. Most of these conclusions fall in line with other results in the literature, but the implications to swing states merit further discussion.

Regression (2) applied to swing states, as with the case of the general dataset, also demonstrates an increase in explanatory power compared to Regression (1). The R-Squared and adjusted R-Squared values rise from 0.825 and 0.819 in Regression (1) to 0.843 and 0.837 in Regression (2), respectively. While the statistical significance of existing basic voting controls remains consistent, the addition of economic indicators may provide new insights. Both the share of people in poverty and the proportion of individuals unemployed are statistically significant with near-zero coefficients.

Regression (3) shows similarities to the results of the primary analysis as well. Parallel to Table 6, Regression (3) has a stronger correlation than Regression (1) and a weaker correlation than Regression (2) in terms of adjusted R-Squared values. This regression does not include economic indicators, but intriguingly is the only regression in which median household income is statistically significant to the 95% confidence level. All other variables hold consistent to previous regressions. The newly included migration flow variables, however, appear to all be statistically significant to different degrees. Change in the percentage of domestic migration is significant to the 99% confidence interval, and both change in the percent of international migration and the percentage of foreign-born individuals are significant to the 95% confidence interval.

The final and fourth regression on swing states has very different results from the primary analysis. Although in the previous three regressions the percentage of Native Americans and Alaskan Natives was significant to the 99% confidence level, for this regression the variable decreases in statistical significance to the 95% confidence level. Although median household income was significant in Regression (3), it again loses significance in Regression (4). The economic indicators, though slightly less statistically significant, generally hold their values. While the share of residents in poverty decreases in significance to the 90% confidence level, the percentage of residents unemployed remains significant at the 99% confidence level. Migration variables experience a parallel decrease in statistical significance. The change in the percent of domestic migration reduces in significance to the 95% confidence level, and both the change in percent of international migration and the percent of foreign-born people are now significant to the 90% confidence level. Overall, the regression results indicate that all added *a priori* variables—the economic indicators from Equation (2) and migration flows from Equation (3)—are statistically significant to a substantial degree in the case of changes in swing state presidential voting shares. Regression (4)'s R-Squared and adjusted R-Squared values strengthen this conclusion. With values of 0.847 and 0.840, they represent the strongest correlation that the model achieves in any case examined in this study. These results suggest that in swing states the *a priori* variables included in this study demonstrate a statistically significant correlation to presidential voting behavior.

VI. Discussion

It is first important to note that in contrast to many studies in the existing literature, the model's dependent variable is not the 2016 Republican presidential vote share but rather the change in the Republican presidential vote share from the 2012 to the 2016 election. The coefficients of the independent variables are thus a measure of the degree to which the Republican presidential vote would have risen or fell in the context of how the party performed in the previous election. As an example, if a given independent variable had a statistically significant coefficient of -2.0, a one percent increase in that variable would suggest that Mr. Trump in the 2016 election would have performed two percent worse than Gov. Romney did in the 2012 election. The model presented in this paper does not claim to predict the 2016 presidential vote share but instead attempts to approximate the hypothetical shifts in the vote that could have occurred.

Also worthy of exploration is the underlying assumption that the regression results can be interpreted *ceteris paribus*[8]. The model presented in this paper divides voter characteristics into separate independent variables. In practice, however, this is not possible with individuals. A person, for example, is not just their race—the individual also possesses a gender and education level. The regressions in this model thus assume that a single independent variable can be changed without affecting the others, which is an impossibility. We assume when interpreting the regression results, however, that those influences are not disruptive enough to significantly skew the implications of the coefficients.

Additionally, there are limits to the results this paper presents due to the exclusion of the top and bottom five

[8]Latin. "All other things equal."

Table 7. Regression results for ten swing states in the dataset.

VARIABLES	(1) Δ% Republican Vote Share	(2) Δ% Republican Vote Share	(3) Δ% Republican Vote Share	(4) Δ% Republican Vote Share
% Black	-8.03***	-11.59***	-8.31***	-11.61***
	(-5.800)	(-7.475)	(-6.251)	(-7.680)
% American Indian or Alaska Native	15.26***	5.90**	14.19***	5.73**
	(7.769)	(2.240)	(7.693)	(2.240)
% Asian	19.83	9.01	13.78	5.08
	(1.202)	(0.762)	(0.805)	(0.369)
% Native Hawaiian or Pacific Islander	-81.21	-46.59	-62.97	-46.17
	(-0.643)	(-0.410)	(-0.499)	(-0.414)
% Hispanic or Latino	-5.56**	-8.51***	-8.62***	-10.76***
	(-2.303)	(-3.382)	(-2.783)	(-3.546)
% with Bachelor's degree or higher	-0.34***	-0.30***	-0.32***	-0.28***
	(-9.662)	(-9.051)	(-8.625)	(-7.911)
% Female	15.43*	3.65	7.08	-3.54
	(1.680)	(0.400)	(0.824)	(-0.395)
Median Household Income	-0.00*	-0.00	-0.00**	-0.00
	(-1.727)	(-0.414)	(-2.327)	(-1.071)
% aged 0-17	-24.10*	-10.00	-18.05	-5.43
	(-1.737)	(-0.689)	(-1.310)	(-0.377)
% aged 18-24	0.76	3.52	3.25	6.14
	(0.055)	(0.262)	(0.233)	(0.454)
% aged 25-39	14.46	19.56	14.31	19.08
	(0.874)	(1.180)	(0.868)	(1.151)
% aged 40-54	27.12	30.21	29.88	33.36
	(1.213)	(1.455)	(1.331)	(1.615)
% aged 65+	10.74	17.72	15.58	21.55
	(0.613)	(0.972)	(0.887)	(1.176)
Population Density	-0.00	-0.00	-0.00	-0.00
	(-0.889)	(-1.093)	(-1.006)	(-1.142)
% in Poverty		0.00**		0.00*
		(2.213)		(1.702)
% in Civilian Labor Force		7.91		7.52
		(1.455)		(1.365)
% Unemployed		0.01***		0.01***
		(4.954)		(4.879)
Δ% Domestic Migration			-0.52***	-0.46**
			(-2.967)	(-2.575)
Δ% International Migration			-2.37**	-1.79*
			(-2.050)	(-1.660)
% Foreign-born			0.15**	0.12*
			(2.114)	(1.654)
Constant	0.03	-0.11	0.04	-0.08
	(0.179)	(-0.702)	(0.281)	(-0.547)
Observations	665	665	665	665
R-squared [adjusted]	0.825 [0.819]	0.843 [0.837]	0.831 [0.824]	0.847 [0.840]

†Robust t-statistics in parentheses. 10 categories of fixed effects absorbed. *** p<0.01, ** p<0.05, * p<0.1

percent of counties by population in the analysis. Though at first glance the removal of the five percent largest counties by population may not appear significant, the nation's uneven population distribution means that this data adjustment removes nearly 50% of the United States population from analysis. I chose to remove these populous counties as I have found that the most and least populated counties in the United States are the most politically polarized, and thus their inclusion may skew the regression analysis. In Los Angeles County, California (the largest county in the country), for example, Sec. Clinton received 1,694,621 more votes than Mr. Trump.[9] Clinton's margin of victory in this single county was larger than the current populations of eleven U.S. states. Conversely, in Blaine County, Nebraska (one of the smallest counties in the United States), Mr. Trump's vote total (276) was over nine times greater than Sec. Clinton's (30).[10] As the model's regression analysis uses percentages, this large percentage difference would have also skewed the model's results. While I believe the exclusion of these very large and small counties on balance improved the explanatory power of my model, the omission of a large percentage of the United States population from analysis may limit the application of my model's results.

The results for the basic voting controls in the full dataset of 2,874 counties demonstrated statistical significance for most variables. This is an encouraging result, as it suggests that the model captures the broader implications of voter behavior correctly. The variables of *% Black* and *% Latino or Hispanic* consistently conveyed consistently large t-statistics of approximately -4.996 and -4.015 along with large coefficients of approximately around -3.45 and -4.15.[11] These values remain relatively consistent throughout the other three regressions. These values support familiar conclusions on the effect of minority populations on voting share, in which the increased presence of minority populations decreases the share of the Republican presidential vote. This result may also lend support to analyses of the 2016 presidential election that contend African American turnout could have potentially swung the election. The model's agreement with broader election analyses is also present in the *% with Bachelor's degree or higher* variable. With neutral coefficients consistently around -0.35, this result agrees that college educated individuals in general did not swing dramatically towards Sec. Clinton. The last statistically significant independent variable of the voting controls is *% aged 0-17*. These coefficients are very large (ranging from -34.03 to -41.36) and do not have a direct effect on voting behavior, as children and adolescents cannot vote. Although the independent variable may imply the effects of an increase in families, I have chosen not to pursue that line of thinking for fear of excessive extrapolation. Finally, the Population Density variable is consistently -0.00. This result is to be expected, as regressions have been weighted by population size.

[9]https://www.lavote.net/ElectionResults/Text/3496.

[10]https://www.nytimes.com/elections/results/nebraska.

[11]Respectively.

Beyond basic voter controls, economic indicators regressed in Regression (2) and Regression (4) also yield some statistically significant results for the full dataset. With t-statistics of 2.901 and 3.150, the regressions performed in this study consistently support that the share of residents in poverty has little to no effect on the change in Republican presidential voting share. This is a noteworthy conclusion that was unexpected. As economic voting literature links economic variables to voting outcomes, it would appear reasonable that poverty would have some correlation with voting outcomes. The regressions performed in this study, however, indicate this is not the case. This result may be an indirect implication of the various different population distributions across the United States: poverty is experienced differently in different parts of the country, and there may not be clear correlations to those experiences. Further investigation of this topic, however, is merited.

Finally, the regressions revealed that the majority of migration flow variables examined were not statistically significant for the primary analysis. In Regression (3), change in domestic migration is the only migration variable that is statistically significant, with a t-statistic of -2.013 and a coefficient of -0.19. This result suggests that counties to which people are moving to within the United States have a very slight negative correlation to the change in Republican presidential vote share. In Regression (4), *% Foreign-born* is statistically significant in the 90% confidence interval, with a coefficient of -0.06. This result, while perhaps somewhat unreliable due to the relatively small number of foreign-born people in the United States, also proves interesting. Among the objectives of this study was to examine if the increased presence of immigrants would be correlated with a higher share of Republican presidential votes. A statistically significant coefficient may have corroborated the idea of an immigration backlash that resulted in a tilt towards Mr. Trump. This hypothesis, however, is not supported by this model's results, which demonstrate a small coefficient with a very slight Democratic tilt. The longstanding popular view that counties with higher levels of immigrants correlate against the change in Republican presidential vote share may still hold.

VII. Table 7. Regression results for ten swing states in the dataset.

Narrowing the dataset from 2,784 counties around the United States to 665 counties in the ten U.S. swing states did not result in considerably different results. Most results for this secondary analysis hold consistent with the primary analysis, with a few notable exceptions that may provide some additional insights.

Of the basic voting controls, the *% American Indian or Alaska Native* variable becomes statistically significant in the 99% confidence interval. While it was my belief that the Native population was likely too small to be statistically significant, this was not the case. Surprisingly, this independent variable has a significant positive correlation with the Republican presidential vote share, running against the conventional wisdom that minority populations, including American Indians, typically vote against Republican candidates. While it may be possible that the population weighting the model's regressions use enable this correlation to be evident for the first time, this result merits further study. Of the economic indicators, the *% Unemployed* variable is statistically significant in the swing states analyzed, with a near-zero effect. This result is similar to the *% in Poverty* variable, which for both the full dataset and the swing state dataset is statistically significant with a coefficient of 0.00. It is again interesting that these regressions convey that these economic indicators have little to no effect. Economic voting may not be as straightforward as elements of the literature suggest.

Examining migration flows for the ten swing states, the *Δ% International Migration* variable becomes statistically significant with a coefficient of -2.37. This result is unexpected, indicating that an increase in international migration correlates against the change in the share of the Republican swing state presidential vote. This result sheds light on a central inquiry of this paper: the exploration of if immigrant flows contributed significantly towards the increase in Republican vote share in 2016 and Mr. Trump's victory. Many pundits and analysts suggested the idea of an immigrant backlash, which these results appear to argue against. Also surprising is the *% Foreign-born* variable's 0.15 coefficient, which suggests a Republican tilt. While these two characteristics may have been expected to be closely related, the relationship between them may be weaker than perhaps presupposed.

VIII. Conclusion

In concurrence with the literature before it, this paper supports the use of regression analysis to correlate between characteristics of voters and voting shares in presidential elections. Though some aspects of the particular model used in this study may invite debate, this broad conclusion appears to strongly hold. It is supported by the consistently large R-Squared and adjusted R-squared values that the regressions held across two data sets and four different regression conditions.

This conclusion does not necessarily assert that econometric models at the scale and methodology utilized in this study are effective or reliable election prediction tools. Even with adjusted R-Squared values consistently above 0.80, the coefficients still depict rough relationships that are imprecise. Presidential elections are typically decided at the margins, and predicting them correctly requires the precision that the regressions ran here cannot provide.

The statistical insignificance and near-zero neutrality of the economic variables included in the second and fourth regressions for both the general and swing state datasets is one of the most unexpected results from the regressions, and has significant implications to the efficacy of sociotropic economic voting. Had the economic indicator variables included not been statistically significant in any form of the model, that perhaps would have been a more inconclusive result. It is, however, the strong statistical significance of these near-zero coefficients that indicates a meaningful result. Though I had at first considered the sociotropic lens to be a more consistent, reliable approach to economic voting models, the benefits of egocentric voting may have become clearer. The egocentric approach offers the possibility of moving past statistical noise more effectively to better understand the correlations between economic characteristics and voter behavior. Ultimately, sociotropic economic indicators may still have some correlations to voting behavior, but that this model's approach was not effective in finding them.

The migration flow results from this study also present some interesting insights. The *Δ% Domestic Migration* variable suggests that areas where individuals are moving to within the nation are likely to have voted more Democratic in 2016 than in 2012. This metric may serve as a proxy economic indicator: counties with high domestic migration may experience such inflows due to robust economic performance, which could translate into voting for the incumbent party. On the other hand, with a single *Δ% International Migration* variable, it is not possible to discern which categories of immigrants are entering counties in which areas of the country. Having separate figures may reveal clearer correlations that this model misses.

The model does, however, produce a noteworthy insight—the coefficients of *Δ% Foreign-born* and *Δ%International Migration* consistently differ. Though at first glance both of these variables may be expected to have similar inclinations, this is not the case under this model. The *Δ% International Migration* variable has a coefficient that remains consistently around -2.37, and *Δ% Foreign-born* around 0.15. This difference between the two variables may be evidence of two separate immigration characteristics. The *Δ% Foreign-born* variable represents the degree to which a county has an immigrant population, whereas *Δ% International Migration* indicates the degree to which new immigrants are entering a county. It thus appears that counties that international immigrants are entering have a stronger Democratic lean, but counties that are developing centralized and established immigrant communities may have no strong inclination one way or another. This may be evidence of an immigrant generational shift, in which more established immigrant communities, with time and growth, become more politically balanced.

In aggregate, the model and its regressions reveal that the 2016 presidential election was not as arbitrary as some media narratives held it to be. Many statistically significant correlations between voter characteristics and changes in Republican presidential vote share, though subject to interpretation, exist and are robust. A "Trump effect"—some broad characteristic of Donald Trump's victory that breaks all the rules and resists all conventional explanations—appears to be more myth than reality. In any case, the results of this model strongly suggest that any such effect would be substantially overpowered by other voting behavior variables.

REFERENCES

[1] Weatherford, M. 1983. Economic Voting and the “Symbolic Politics” Argument: A Reinterpretation and Synthesis. The American Political Science Review, 77(1), 158-174. doi:10.2307/1956017

[2] Leamer, E. 1983. *Specification Searches: Ad Hoc Inference with Nonexperimental Data.* New York: John Wiley & Sons.

[3] Fair, R. 1978. The Effect of Economic Events on Votes for President. *The Review of Economics and Statistics*, 60(2), 159-173. doi:10.2307/1924969

[4] Autor, D., Dorn, D., Hanson, G., & Majlesi, K. 2016. Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure. NBER Working Paper Series. doi:10.3386/w22637

[5] Bailey, M., & Dynarski, S. 2011. Gains and Gaps: Changing Inequality in U.S. College Entry and Completion. doi:10.3386/w17633.

[6] Bloom, H., & Price, D. 1975. Voter Response to Short-Run Economic Conditions: The Asymmetric Effect of Prosperity and Recession. The American Political Science Review 69, no. 4: 1240-254. doi:10.2307/1955284.

[7] Burden, B., & Wichowsky, A. 2012. Unemployment and Voter Turnout. APSA 2012 Annual Meeting Paper.

[8] DeSimone, J., & LaFountain, C. 2007. Still the Economy, Stupid: Economic Voting in the 2004 Presidential Election. doi:10.3386/w13549.

[9] Kinder, D., & Kiewiet, D. 1979. Economic Discontent and Political Behavior: The Role of Personal Grievances and Collective Economic Judgments in Congressional Voting. American Journal of Political Science 23, no. 3: 495-527. doi:10.2307/2111027.

[10] Holbrook, Thomas, and James C. Garand. 1996. Homo Economus? Economic Information and Economic Voting. Political Research Quarterly 49, no. 2: 351. doi:10.2307/448878.

[11] Rosenstone, S. 1982. “Economic Adversity and Voter Turnout.” American Journal of Political Science 26, no. 1: 25-46. doi:10.2307/2110837.

[12] Jensen, Q., & Weymouth, S. 2016. Winners and Losers in International Trade: The Effects on U.S. Presidential Voting. doi:10.3386/w21899.

[13] Gelman, A., & Su, Y. 2010. Voting by education in 2008. Chance 23, no. 8: 8. doi:10.1007/s00144-010-0040-z.

[14] Brunner, E., Ross, S., & Washington, E. 2008. Economics and ideology: causal evidence of the impact of economic conditions on support for redistribution and other ballot proposals. Cambridge, MA: National Bureau of Economic Research. http://www.nber.org/papers/w14091.pdf