God’s Country:
The Proliferation of Evangelical Christendom and its Impact on Voting Behavior in America

Claire Williams
Honors Politics Undergraduate Thesis
New York University
Class of 2019

Abstract
Since Ronald Regan’s presidential campaign of 1980, the “evangelical vote” has been considered essential to a successful Republican presidential bid. Observing that evangelical Christians demonstrate a tendency to vote in higher proportions for Republican presidential candidates than mainline-Protestants and other Christians, this study aims to answer the question: What is the impact of the proliferation of evangelical Christendom on voting behavior in America? The study examines the effects of the spread of evangelical Christian churches on a county’s tendency to vote for Republican presidential candidates. To analyze the physical impact of a religious institution’s opening or closing on voting behavior at the county level, the study utilizes a difference in differences design including time and geographically fixed effects for 2,710 counties in the United States over seventeen presidential elections. This study finds a statistically significant, positive relationship between greater shares of evangelical churches and higher vote shares for Republican presidential candidates within counties between the 1980 and 2012 Presidential elections. The study also finds that this relationship is emphasized further in southern counties than in non-southern counties.
I. Introduction

Each election cycle, the “evangelical vote” becomes a subject of great interest to pollsters interested in predicting outcomes for Republican presidential candidates. Who exactly are the evangelicals and why are they so important to the Republican Party? Evangelicals are a subset of Protestant Christians who believe in the personal conversion experience of being “saved” or “born again,” they believe in textualism of the Bible, that Jesus, through his death and resurrection, is the only source of salvation and forgiveness of sin, and they are strongly motivated to share the gospel through individual or organized missions (FitzGerald 2018). Most also believe in the rapture, and believe that only those who follow God will be saved. One key aspect, which differentiates Evangelicals from other mainline Protestant groups, is that they evangelize. That is, their mission to spread the gospel, to “Christianize” the nation (Burge, Lewis 2012).

In Ronald Reagan’s 1980 campaign, the Republican Party recognized the importance of the evangelical voters as a reliable base and actively sought their backing. The Republican platform began to include planks supporting organized prayer in public schools and defining human life as beginning at conception, and the party started to embrace the term “pro-family” to describe its agenda. Evangelical voters responded, providing strong support for Reagan in 1980 and 1984. By the 1988 presidential election, abortion was a centerpiece of Christian Right politics (Wormald 2014). Since Reagan’s election, the Evangelical vote has remained strongly Republican, with Donald Trump most recently receiving 82% of their vote in the 2016 election (Wormald 2014).

Nearly half (48%) of white evangelical Protestants and 60% of highly committed evangelicals say their religious beliefs frequently affect their electoral choices, compared with
15% of white mainline-Protestants (Wormald 2014). As of 2012, America is no longer majority mainline-Protestant, and evangelical-Protestantism, alongside non-Christian religions, continues to rise in popularity, with nearly 25% of the populous self-identifying as evangelical-Protestant Christian (Liu 2013). This phenomenon of shifting religious influence is critical to the study of politics in the United States and understanding why the evangelical vote is considered such an important aspect of a successful Republican campaign. Knowing that evangelicals vote disproportionately for Republicans, questions arise about how big, exactly, their influence is on the American electoral landscape. This background provided, my research question becomes: What is the impact of the proliferation of evangelical Christendom on voting behavior in America?

II. Literature Review

Bartels’ 2005 study, *What’s the Matter with Kansas?*, uses a “trend” variable to observe voting in presidential elections since 1984 (Bartels 2005). He separates the data by class and uses indices of economic and social policy preferences constructed with the NES data. For frequent churchgoers, the strength of these relationships roughly doubled between 1984 and 2004. For people whose religious beliefs provide “a great deal” of guidance in their day-to-day lives the increases were smaller but still substantial. Bartels notes that social voting does not discount economic voting, even amongst the religious — though he never talks about evangelicals, specifically (Bartels 2005). While Bartels’ study offers some interesting insights into the strength of religion on individual voting behavior, one notable issue in this study is his use of social policy preferences and indices to evaluate the relationship between religion and voting, leading to endogeneity problems and failing to utilize a causal mechanism in his design.
Hirsch, Booth, Glenna, and Green’s 2012 study, *Politics, Religion, and Society: Is The United States Experiencing a Period of Religious-Political Polarization?*, demonstrates findings that religious identity influenced voter choice and that this influence increased significantly and substantially across the study period (Hirsch, Booth, Glenna, Green 2012). Over the study period, the effect of these two constructs on voter choice increased, and the implication is that, over time, religiously influenced voters became more likely to vote Republican, whereas more secular individuals became more likely to vote Democratic (Hirsch, Booth, Glenna, Green 2012). What this survey does not account for is the division between evangelical and mainline-Protestants, though it offers a sort of rough scale which places “belief in biblical authority” as a kind of indicator of the strength of ideology amongst Protestants (Hirsch, Booth, Glenna, Green 2012). Using multiple cross sections to empirically model presidential voting over the period of 1980 to 2008, the study reports an analysis of two surveys (General Social Survey and Cornell National Social Survey). The GSS/CNSS independent variable, "religious identity," is operationalized by cross-classifying two measures: biblical authority and religious tradition. The dependent variable for the GSS/CNSS analyses is the binary voting preference for the Republican versus Democratic Presidential candidate. The trend line observed in Protestant partisanship between 1980 and 2008 suggests growing religious-political polarization, where liberal interpreters of biblical authority became stronger Democratic partisans, and literal interpreters of biblical authority became stronger Republican voters. The authors attempted to address endogeneity concerns by including gender, income quartiles, racial, and geographic covariates. However, despite the efforts to separate these variables to evaluate their impact on confounding religion and voting behavior, this study, like the previous study by Bartels (2005),
analyzes trend lines, not discontinuous observable events, leaving a lack of causality as a central concern to the study’s attempts to link religion and voting behavior.

In one of the most compelling pieces for this study’s theory on church attendance and voting behavior, a study by Brad Lockerbie (2013) establishes that the more frequently people show up to their church, the more they might be pushed to vote certain ways (Lockerbie 2013). Lockerbie’s study uses regression and logit analysis of the American National Election Studies of 1992 through 2008. He finds that the relationship between religion and voting behavior in presidential elections is more complex than previously thought, and that church attendance might be an indicator of how seriously one takes one's particular religious faith. He notes that “if this is true, and the major means by which religion influences politics is through voting, we should see that this item is statistically significant and powerful, and the other religious variables are either insignificant or significant but comparatively weak (Lockerbie 2013).” He also asserts that the Republican leadership's frequent invocation of religious themes might pull those for whom religion is important, as demonstrated by attendance, toward the conservative side (Lockerbie 2013). The difficulty of analyzing a causal relationship between evangelism and voting behavior from Lockerbie’s study is that his conclusions center on political attitude analysis. From this analysis, he makes predictions and speculations about voting behavior based on background studies, but there can be no sound causal conclusions from his analysis. However, this study offers groundwork for my theories on the activities of individuals in religious institutions, and the subsequent impact of that institutional activity in the public sphere.

Bradberry’s 2016 study, *The Effect of Religion on Candidate Preference in the 2008 and 2012 Republican Presidential Primaries*, shows that Republican candidates who most explicitly appealed to religious voters (Mike Huckabee in 2008 and Rick Santorum in 2012) were the
preferred candidate of Republican respondents who attended religious services at the highest levels, and that as attendance increased, so did the likelihood of preferring that candidate (Bradberry 2016). He also found that identification as a born-again Christian mattered to candidate preference. Specifically, born-again Christians were more likely than non-born again Christians to prefer Huckabee to Mitt Romney, John McCain and Ron Paul in 2008, and Santorum to Romney in 2012. The overall findings of the paper provide evidence that religion variables can add to understanding of why voters prefer one candidate to another in presidential primaries (Bradberry 2016). In determining which Republican candidate in 2008 most explicitly (or most frequently) appealed to religious voters, Bradberry analyzes two data sources: nationally-televised presidential debates and the candidates’ television ads. Bradberry looked for and counted any overt appeals by the candidates to religious voters, whether by explicit references to religion, faith, God, or by any emphasis on issues such as abortion and gay marriage using a list of phrases as criteria. He limited the final analysis to the final four candidates standing—Huckabee, McCain, Romney, and Paul. Among Republican respondents who seldom or never attend religious services, Huckabee registers a low 12%. However, Huckabee’s support increases among moderate and high attenders, and ultimately reaches an impressive 47% among those who attend religious services more than once a week. Notably, Huckabee even surpasses McCain among the highest attending respondents (Bradberry 2016). Despite the study’s insights into church attendance rates and voting, this study leaves the door open to issues of endogeneity. Confirming that Bradberry was successfully able to analyze the target audience (Protestant Christians) that these political adds aimed to capture is not something that can be confirmed, and though it seems that these trends align in a significant way, it could be
the case that other factors influenced the votes— not the television ads for presidential candidates.

There are numerous studies, papers, and surveys that help to illuminate the impact of evangelical Christians’ role in voting patterns in America, but these studies lack causal mechanisms and rely instead on trends to establish their findings. As discussed above, each of these studies, despite their robust research designs, utilize individual attitude survey data, leading to a generic issue of endogeneity. This is the primary issue I pose to address in my research design. By examining the number of physical religious institutions in a geographical unit over time rather than attempting to track and measure trends of the individual attitudes of religious voters over time, I aim to more clearly examine the impact of *institutionalized* religion and the impact of such a physical entity like a church on the causal relationship between active congregational worship activities and voting behavior. While surveys can effectively measure individual attitudes and make assumptions about what these attitudes have on voting, they consistently run the risk of getting tangled in issues of endogeneity, relying on human units, rather than institutional units of measure to draw conclusions which may not be causally sound.

Opinion surveys are useful in many respects, but they often fail to capture a representative proportion of the population, leading to selection bias issues, and also pose problems of response bias. Individuals cannot always be counted on to reply to surveys with complete accuracy because of the way that questions are posed and external pressures to respond in particular ways to the sometimes contentious questions. A study by McAndrew, Siobhan, and Voas, *Measuring Religiosity Using Surveys* (2011), illustrates the difficulties of using survey responses to measure religious affiliation. They find that responses are heavily influenced by the wording and context of the question and that asking individuals whether or not they attend
religious services does not always yield results that should be considered reliable (McAndrew, Siobhan, Voas 2011).

Additionally, there is not a significant amount of literature that attempts to distinguish evangelicals from their mainline counterparts in regards to voting mobilization. I believe there should be further emphasis on the differences between the two groups, considering the gap in partisanship demonstrated in previous Presidential elections. Social scientists do not focus enough on the physical impact of congregational proliferation on voting behavior. They recognize civic engagement and its ties to congregation attendance, they measure the ideology of evangelicals and their voting trends, but they are not looking at real outcomes of congregational activity on voting behavior.

III. Causal Model

One phenomenon which can help to explain the hypothesize that evangelicals will be more likely to vote for Republicans is commonly referred to as “the God Gap.” The God Gap illustrates the tendencies of voters who identify as religious to vote for Republicans, and voters who identify as secular or non-religious to vote for Democrats (Sheets 2011). Amongst these religious voters, the Protestant base is the largest in America. This base, which prior to 2012 was composed of a majority of mainline-Protestants, is now composed mainly of evangelical-Protestants, at 55% (Wormald 2014).

The party platform of the Republicans is preferable to evangelicals because they view the party as more friendly towards their values (Liu 2013). Some of these values include traditionally conservative ideas of marriage, a strong affinity to be “pro-life,” a belief that the Republican party is more friendly towards religious freedoms, and a distrust of Democratic party, which Evangelicals view as less religious and more socially liberal (Liu 2013).
Bartels’ 2005 study provides substantial evidence of an increasingly close alignment between the social issues positions of frequent church-goers—highly religious people—and their partisanship and voting behavior (Bartels 2005). Researchers seem to agree that it is possible to establish causality between religious affiliation and voting behavior, and stress that identity-voting is strong amongst evangelicals (Bartels 2005). This study relies on the validity of identity-voting, which shows that voters, who are otherwise centrist, move toward the parties that align with their identities — in this case, being an evangelical Christian (Stephen & Ansolabehere 2016).

Additionally, Tea Party supporters are much more likely than registered voters as a whole to say that their religion is the most important factor in determining their opinions on these social issues— and they draw disproportionate support from the ranks of white evangelical Protestants (Stephen & Ansolabehere 2016).

IV. Problems With Causal Inference

When observing voting behavior, there are many confounders that need to be addressed. Social identities such as economic class, ethnicity, gender, race, and religious preference can all affect voting behavior. Additionally, class, ethnicity, gender, and race can affect religious preference. Hirsch, Booth, Glenna, and Greens’ 2012 study attempts to reconcile some of these issues, finding that Protestant partisanship is relatively less divided by social class, and more divided by biblical belief in comparison to Catholic partisanship that is more divided by social class, and less divided by biblical belief (Hirsch, Booth, Glenna, Green 2012). They speculate that this difference is related to how political parties appeal to these two religious traditions. The study findings suggest that the influence of religious identity on Presidential voter choice strengthens during periods of rising economic inequality. They also find that the effect of gender
on partisanship is less pronounced and overshadowed by social class and religious identity. Although women in the aggregate are net Democratic partisans, upper class, white, biblical literalist, Protestant women are strong Republican partisans (Hirsch, Booth, Glenna, Green 2012).

Another issue of causal inference in many studies observing religion and voting behaviors is the attempt to measure continuous, rather than discontinuous events. Though individual voting behavior and individual religious affiliation may trend together over time, there is no way to look at both of these phenomena and soundly determine that there is a causal relationship between them. Even if there were, there is also no way of identifying which is causing the other, providing no causal mechanism. As mentioned above in the literature review, while speculations about voting behavior using individual religious attitude surveys or trend lines can offer some potential insights, they can never untangle the issue of endogeneity that accompany a lack of sufficient causal mechanisms in an experiment. Individual religious behaviors are difficult to quantify and group, and making inferences from religious attitudes to voting behavior is problematic. Therefore, measuring ideology changes over time based on trends is not an effective causal method to examine voting behavior.

In an effort to address the difficulties of establishing causality, my research design uses a discontinuous treatment. I will observe all Christian congregations in the United States over a sixty year period, using all mainline-Protestant, Catholic, Orthodox, and other non-denominational groups in a placebo group compared to evangelical Congregations in the treatment group. This method aims to compare growth between the placebo and treatment groups and those groups’ respective voting behaviors. Looking at institutions instead of attitudes allows
for a more robust causal design and, in addition, will help to avoid making assumptions about individual voting behavior, by instead focusing on religious institutional activity.

Because the potential for endogeneity issues is always at risk in natural experiments, I aim to address the issue by including time-varying covariates, controlling for confounding factors like disparities in income, class, racial composition, and population differences. I will use dummy variables indicating each year in the sample, with one omitted as a reference group for the election year. My coefficients on these dummies (time-effects) will capture temporal changes in religious institutional activities common to all counties over time in a multi-county analysis or “county-year” panel (Angrist & Pischke 2015). The same method will be applied to capture county-effects. This model presumes that, in the absence of a treatment effect, evangelical congregational growth rates would not deviate from common year effects following the linear trend captured by the coefficient. Evidence for trends should come as sharper deviations from otherwise smooth trends. My difference in differences research design (DID) captures treatment effects in the face of uncommon trends. The sharper the deviation, the more likely this study will be to uncover any sort of causal effect between the religious institutional activity of evangelical churches on voting behavior in American counties over time (Angrist & Pischke 2015).

V. Testable Hypotheses

I hypothesize that the shift away from mainline-Protestantism and the growing influence of evangelical-Protestantism will lead to an increased probability of voting for Republicans in presidential elections. I define my independent and dependent variables and the expected relationship between them as follows: The independent variable will be the number of
evangelical congregations within a given county.\textsuperscript{1} The dependent variable will be the vote share for Republicans in presidential races within a given county. The expected relationship between the variables is that as the proportion of evangelical congregations to other Christian congregations in a county increases, the county will be more likely to vote for a Republican presidential candidate. I also hypothesize that mainline-Protestant shares within counties will not have significant impact on a county’s tendency to vote in higher proportions for Republican presidential candidates.

In order to better analyze the effects of changing religiosity on voting behavior and ensure the impact is properly measured, fixed effects will be employed in the model. A fixed effects regression model helps to control time invariant factors. It is important to account for these confounders to avoid omitted variable bias. My models will also include time fixed effects for year-to-year differences in voting tendencies and geographic fixed effects to account for particular voting behaviors, such as political party strength in each county. Time varying covariates will also be included. These time varying covariates will be including using US Census data, outlined in my empirical methods section, applied to the election years my study examines. The covariates account for confounding issues such as differences in education levels, demographic makeup of counties, household income disparities, and population growth and decline.

VI. Description of Data

The Presidential election data originates from Data Planet by Sage Publishers, a social science organization that complies political data. From this, the study specifically uses the Dave Congregations defined in the Independent Variables Section

\textsuperscript{1} Congregations defined in the Independent Variables Section
Leip’s Political Atlas Presidential Election Returns. Party vote share in Presidential elections is used as the key dependent variable with the unit of analysis defined at the county-level. At the county level, the vote share for the top-two candidates are recorded. I purposefully exclude other candidates from my analysis, observing only the Republican and Democratic candidate vote shares. The collection years for the presidential elections are 1952-2016, reporting each presidential election every four years.

To measure the effects of congregation shares of evangelical churches on voting behavior, I use data from the Association of Religion Data Archives (ARDA). The ARDA hosts collections of surveys, polls, and other data submitted by researchers and made available online by the ARDA. Specifically, this study utilizes county-level congregational census surveys to measure evangelical congregation shares. These surveys are conducted by The Association of Statisticians of American Religious Bodies (ASARB), who invite all Judeo-Christian religious bodies listed in the *Yearbook of American and Canadian Churches* to participate (*Churches and Church Membership in the United States*, ARDA). The data collected rely on studies of self-reporting churches in the United States. While this collection process does allow for the possibility of potential selection issues, the distribution of churches listed in the dataset resembles previously established percentages of religious institutions collected by Pew Research and the U.S. Census. It may be the case that some types of churches do not wish to answer the surveys, and while this is a limitation that remains insurmountable, it should also be noted that, overall, the datasets represent roughly ninety percent of all congregations in the United States, and 80% in the 1971 collection year (*Churches and Church Membership in the United States*, ARDA). These sets are estimated to contain about 55% of the American population in each response cycle. When churches are sent invitations to participate in the surveys, denominational
offices were tasked with collecting information from all churches within their county. The reporting was then sent to the ASARB (the principle investigators), double checked for any irregularities or significant changes in the numbers from the previous reporting cycle, and returned to the denominational offices to be fixed if necessary. The congregations are coded numerically, with the number of specific congregations in each county recorded as numbers.

Capturing the number of evangelical congregations within each county is the primary independent variable of my study. I use seven separate data sets from 1952, 1961, 1971, 1980, 1990, 2000, and 2011 (Churches and Church Membership in the United States, ARDA). In order to group individual congregations into evangelical shares, I examined each individual congregation across every dataset and identified them by denomination. Each of the codebooks provided for the above-mentioned datasets listed individual congregations, not grouped by denomination, as variables. To identify the denomination to which each congregation belonged, I consulted a variety of sources including several individual church websites, the ARDA’s denominational categorization tools, and the Yearbook of American and Canadian Churches. These consultations acted in a two-fold process. The individual church websites and the Yearbook of American and Canadian Churches identify the congregations as members of a larger denomination, either nominally or through their “mission statements” which espouse the key tenets of the church. I would then re-confirm these identifications using the ARDA’s categorization tools. The ARDA categorizations also list key tenets of denominational groups, through which I was able to crosscheck the correct denomination to which the churches belonged. Using this method, I identified five groups of Christian churches within the data: evangelical-Protestant, mainline-Protestant, Amish and Mennonite churches, Catholic, and Orthodox churches. After the efforts to group the churches into denominational groups, each
group was then identifiable as a unique number out of the total number of congregations identified in the codebooks. For the purposes of this project, I will focus primarily on the Evangelical percentage. Evangelical congregations, as the independent variable, are the observations of primary concern in observing the proposed relationship between institutionalized religion and voting behaviors. All other congregations (including mainline-Protestant congregations) will act as a comparison group, or placebo, with which we will be able to observe potential differences between evangelical and mainline institutional effects on voting behavior within counties. The mainline-Protestant congregations, as noted in my literature review, tend to demonstrate voting patterns that do not resemble the higher likelihood of evangelicals to vote Republican. Whereas evangelical belief systems more vehemently support conservatism, emphasize biblical literalism, and stress the importance of the rapture, mainline Protestantism does not. Thus, as stated in my second hypothesis, I expect the mainline shares to function as a comparison point, or placebo, for the evangelical share coefficients. It should be noted that the comparison group does not include any non-Christian religious institutions.

I utilize numeric and percentage coding schemes for my variables. The numeric values are continuous, and the percentages are indicated as .X amount out of 1.00. All congregational data is numeric, apart from the evangelical share counts (coded as evangelical_share). Voting data is all coded as percentages. My covariates are also all coded as percentages, apart from median household income (avg_HH_income), which uses a numeric continuous value-coding scheme.

Because the Association of Religious Data Archives and the US Census conduct their surveys only once every ten years, I will use linear interpolation to align congregation numbers and census data with presidential election years when I begin the analysis process of my research.
to provide a more detailed view of congregational shares and census data than only once every ten years. There may be a few missing values for both the congregation and census data, because over the expansive amount of time that this study covers, counties have changed. In total, this study captures 2,710 counties out of 3,071 counties in the United States spanning this seventy year period, as redistricting since 1952 altered many county lines, and I was unable to capture all of these changes in the dataset.

As congregations do not always remain static entities, there are cases where certain congregations have switched denominational affiliation or changed names. This will not pose any major issues to my design because I have noted all congregations that change, their new respective titles, and the year in which the congregation switched denomination, if applicable. Additionally, as this study is largely uninterested in any one specific congregation, but rather groupings of mainline and evangelical denominations, this should not pose any major issues.

VII. Empirical Methods

This study’s empirical strategy incorporates a difference-in-differences design (DID). The DID design works by comparing the average change over time in the outcome variable for my treatment groups, compared to the average change over time for my control groups. The DID hypothesis suggests that the treatment variable impacts voting behavior. The dataset allows me to examine each individual county in the United States. The equation for this design is as follows:

\[ Y = a + B1 +B2 +B3 + covariates \]

In the equation, \( Y \) is the outcome variable (vote share for Republicans in presidential elections by county and time). \( a \) acts as the constant variable, \( B1 \) as the coefficient (percentage of evangelical congregations in county), \( B2 \) as the set of geographic fixed effects (counties over
time), B3 as the set of time fixed effects (election year over time), and the covariates will be
time-varying by geography and time.

My covariates are acquired from US Census data. These covariates include demographics
of each county, the median income, the population of each county, and education attainment
levels. Prior to 1980, the covariates collected are not consistently reported. Therefore, when
testing regressions to determine the model of best fit, I will run the entire equation without any
covariates for years 1952 to 2016, then again for years 1980 to 2016, then run the post-1976
elections with the covariates included and see what differences, if any, emerge on the outcomes.
If adding the covariates has a significant effect on the outcomes, the study will make note of this
and extrapolate upon the differences of adding covariates to the model. I expect the coefficients
in the model to show that an increase in the proportion of evangelical congregations will be
associated with an increase in the Republican vote share.

VIII. Results

The following subsections examine the empirical analysis of two models. The first set of
models is a fixed effects model observing all recorded 2,710 American counties in the dataset.
The second set is also a fixed effects model, though it subsets the national model into southern
and non-southern counties to evaluate any potential regional differences. For both empirical
models, county and year fixed effects are incorporated, and standard errors are clustered at the
county level. The four national fixed effect equations are modeled with a base equation as
follows:
The four subsetted fixed effects equations are modeled with a base equation as follows:

\[ Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \beta_4 X_{4it} + \beta_5 X_{5it} + \ldots + u_{it} \]

Model one examines only the religious data in the regression equation, still taking into account both time and county fixed effects. All observations in the dataset are accounted for. The evangelical share shows a positive coefficient (0.0546) and a p-value significant at 0.001 level. The mainline share variable shows a small, negative coefficient at (0.0016) and no statistically significant relationship is demonstrated between mainline share and Republican vote share.

Without any control variables, it is difficult to assess the reliability of the model, but it is worth...
noting that at the simplest level, this regression holds in favor of the alternative hypothesis. It is also worth noting that the mainline share, showing no significant relationship in the model, supports this study’s secondary alternative hypothesis that mainline church shares do not significantly impact Republican vote shares. The model demonstrates an R-squared value of 0.24.

Model two includes all covariates in the dataset. The model also includes extrapolated data for the 2016 election. All of the variables show a statistically significant relationship with Republican vote share except for evangelical share. The religious variables, contrary to the previous results in model one, show an opposite correlation with Republican vote share outcomes. Evangelical share still indicates a positive, though smaller coefficient (0.014) though no statistically significant relationship. The mainline variable shows a negative coefficient (-0.042) and is statistically significant at the 0.001 level. This model reports an R-squared value of 0.30—the largest of the four models above. Despite there being no statistically significant relationship between the evangelical variable on Republican vote share, it still holds that there is no demonstrated negative relationship between evangelical share and the outcome in Republican vote share, and the mainline variable’s being negative is also interesting to note.

In an attempt to restrict model three, I refer to previous literature indicating the importance of the 1980 Presidential election and its impact on the potential relationship between evangelical share and Republican vote share. Model three shows a statistically significant association between evangelical share and republican vote share at 0.001 with a positive coefficient of 0.059, demonstrating a strong positive relationship. Mainline share also shows a positive coefficient of (0.012), but there is no statistically significant relationship demonstrated by this regression. In comparison to the two previous models, the third model was the closest to
matching the alternative hypothesis with robust results. These results, while promising for the alternative hypothesis, included data from the 2016 election. All data in the dataset had been anchored by the endpoint year of 2016, with the exception of the religious data, as the ARDA surveys only covered up to the 2012 election. The need to restrict the model further to exclude the 2016 election follows, testing the strength of the model to find the best fit while including in the regression only the “real” data, excluding extrapolated data. The extrapolated values assume the same growth rate between 2012 and 2016 as between 2008 and 2012. The possibility for overestimating the strength of the relationship due to inclusion of the 2016 data leaves too much room for error while attempting to draw accurate conclusions from the model. We also see in model three that there are fewer statistically significant coefficients than the previous models, suggesting that the model could be further refined to better evaluate any potential relationships between evangelical share and Republican vote share.

Model four restricts the data further to include presidential elections from 1980 to 2012. This model again shows a statistically significant relationship between evangelical share and republican vote share at the 5% level with a positive coefficient (0.033). Though the coefficient is slightly smaller than the previous model, the relationship remains strikingly robust. The mainline variable, again as in model three, remains statistically insignificant, though in this model it shows a negative coefficient of (-0.023). This model serves as the best fit for evaluating all counties in the United States over a thirty-two year time period. Looking at 1980-2012, this study is able to observe results that are more robust and of greater magnitude while maintaining that only real data is observed. The R-squared for this model is 0.23, losing only 0.01% of the observable data to the previous model. This model also maximized the statistical significance of the covariates.
These subsetted fixed effects models utilize a dummy variable, county_in_south, to evaluate possible regional differences between evangelical shares and Republican vote shares in southern and non-southern counties. For this study’s purposes, the United States was split into only “southern” states and “non-southern” states, grouped according to the US Census Bureau’s definition of American regions (U.S. Census Bureau). After determining the set of states to isolate, those southern state’s counties were identified for observation. The 2016 election was excluded for all four of these models after establishing in the previous regressions that including
the 2016 data could lead to overemphasizing the relationship between the dependent and independent variables. Models one and two examine the data with the dummy variable switched on, indicating a southern county, and models three and four examine non-southern counties, having the dummy variable switched off.

Model one’s examination of southern counties from 1952 to 2012 reveals negative coefficients on both religious indicator variables. Furthermore, neither of the variables shows a statistically significant relationship with the Republican vote share.

Model two examines southern counties between 1980 and 2012, reflecting the fourth model in the fixed effects regressions, though only isolating southern counties. In this model, both evangelical and mainline share variables show a strong statistically significant relationship with Republican vote share with coefficients of 0.128 and 0.089, respectively. Further analysis of these strikingly large coefficients is examined in the discussion section below.

Model three examines elections between 1952 and 2012 again, but for non-southern states. In this model, it appears that both religious indicator variables show positive coefficients that are statistically significant at 0.05 (evangelical) and 0.001 (mainline). The evangelical share shows a positive coefficient (0.018) and mainline shows a negative coefficient (-0.04).

Model four exhibits the closest reflection of model four in the national set of regressions. The relationship between evangelical share and republican vote share here is again statistically significant at 0.001 with a positive coefficient of 0.045 and mainline share showing a negative coefficient of -0.021 that is not statistically significant. This model demonstrates that the relationship between evangelical share and Republican vote share is both more statistically significant with a 1% larger coefficient when non-southern counties are isolated from southern counties.
IX. Discussion

The results of the regression models offer many points for discussion. As elaborated in this study’s literature review, previous research exhibits that attempts to determine a causal relationship between religion and voting behavior is not easily accomplished. Unlike these studies, my model makes methodological changes to evaluate the relationship between religion and voting by focusing on the institutional, physical power that a religious body possesses rather than attempting to base the study on individual attitudes. The results in six of the eight total models tested supported the alternative hypothesis that there is a statistically significant relationship between the proliferation of evangelical churches in America and voting behavior. Furthermore, the secondary hypothesis that mainline church proliferation is not causally related to voting behavior is also supported. However, the results of the regressions cannot guarantee to prove causality completely and leave room for a multitude of discussions.

In the national models, there emerged an interesting observation that until limiting the data to presidential elections post-1976, there was no statistically significant relationship between evangelical church shares and republican vote shares. As discussed in the literature review, the 1980 election was a critical indicator marking a shift in this relationship. The Reagan campaign made a targeted appeal to evangelical voters with a platform that focused on “traditional family values,” religious freedom as a priority for the nation, overt messages about God and country, and a pro-life stance at the center (Wormald 2014). As reflected in the results of the regressions, this message seems to have a substantial impact. There is the counter-argument, of course, that the causal relationship is reversed—that politics influence religion rather than religion influencing political choice. Again, previous literature has defended the causal relationship of religion impacting voting (Bartels 2005, Stephen & Ansolabehere 2016),
but it remains a challenge to confirm the direction of the phenomena. By restricting the model of best fit to the elections of 1980 to 2012, however, this study is able to establish a strong, positive correlation between the phenomena and pose the possibility of establishing a causal relationship.

The decision to eliminate data from the 2016 presidential election rested primarily on the fact that without an “anchor year” to interpolate between, the subsequent prediction for 2016 religious indicator variables (census data and voting data were real) was too big a prediction to make. However, despite the decision to omit this data, it remains likely that the upward trend in the proliferation of evangelical churches continued, based on trends established by polling data (Liu 2013). Perhaps when the ARDA releases the 2020 census on religious bodies in the United States, this study could more successfully evaluate the 2016 election and even the upcoming 2020 election to see if the results would differ.

The model of best fit, the regression that included elections from 1980 to 2012, supported the alternative hypothesis that evangelical share, in the county as a whole, impacts republican vote shares at the county level with a 3% impact on votes per 1% of evangelical churches added to a county since the previous election. On the other hand, as stated by the second hypothesis, the mainline churches did not have a statistically significant relationship with Republican vote share. The goal of including the mainline churches was to examine whether or not they could act as a “control” or placebo to compare with evangelical churches, and the results maintain that this placebo worked according to the predicted hypothesis. The census data also proved critical to include as covariates. A majority of these covariates returned statistically significant relationships on the outcome variable supporting the studies included in the literature review that indicated these variables as having an impact on Republican presidential vote shares. Including the covariates also lowered the coefficients on the religious indicator variables.
After analyzing the results of the national models, questions emerged regarding regional differences. Looking at the United States holistically, the results were indeed promising for the alternative hypothesis and accomplished the initial goals of the study. In order to test the strength of the national model and assess any regional differences that could emerge as exceptions to the national model, I employed a dummy variable that served to isolate the southern region of the United States. The results from the subsetted models offered two interesting results. Firstly, for what will be considered “historical data,” the data before the 1980 election we see in the southern states that there was no statistically significant relationship between religious indicator variables and republican vote share, contrary to the national model which established evangelical share as a significant indicator with a positive coefficient of 0.05.

Secondly, when the dummy variable was switched on for southern states in the 1980 to 2012 elections, the coefficients for both evangelical and mainline shares are both statistically significant at 0.01% and exhibit large coefficients of 0.128 and 0.089, respectively. This confirms the theory that southern states would be more apt to have evangelical share affect republican vote share, but this also demonstrates that, in southern counties, mainline shares affect Republican vote shares as well. When the dummy variable is switched off and only the non-southern states are examined in this time frame, we see more comparable results to the national analysis. The evangelical share exhibits a statistically significant relationship at 0.001 and possesses a positive coefficient of 0.045, slightly stronger than the national model, and mainline share demonstrating no significant relationship with Republican vote share and a negative coefficient of -0.021.

There is literature to support why these differences in the southern, non-southern, and national models emerged. In a 2016 study by Hood analyzing U.S. Senate races in the American
south, Hood found that religious identification holds a more significant place in predicting voting outcomes in southern states than in “periphery southern states” or in non-southern states. Hood’s study looks again at individual voting behaviors, with evangelicals in the south identifying slightly more strongly as voting for Republicans candidates than mainline Protestants by a margin of 3%, which is reflected in this study’s subsetted southern model two’s results (Hood 2016).

The curiously large positive coefficients on the subsetted model two were initially difficult to explain. There were no errors in the dataset that could explain the large coefficients, and the regression commands remained the same across models with the exception of switching the dummy variable on and off. One possible reason for the high coefficients on both religious indicator variables is the heterogeneous effect on born-again Christians (evangelical Christians) in the south. A 2014 study by White found that “born-again Christian self-identification has a bigger impact on whites in states like Alabama, Mississippi, South Carolina, and Tennessee than it does in California, New York, Vermont, and Wisconsin (White 2014).” So, while the coefficients demonstrated in the subsetted model two are significantly higher than in the national models, there may be a heterogeneous effect affecting these coefficients, conflating multiple identities tied to being an evangelical Christian in the south, including whiteness and class (McKee & Springer 2015). While the statistical relationship is “real,” the model may exaggerate the coefficients. Another possible explanation for these large coefficients is that the density of churches in southern counties is higher than other counties.

The non-southern subsetted models more closely resemble the national models. For both the national model and the non-southern model, mainline share is not a significant indicator on Republican presidential vote share. Evangelical share is significant for both models, though at
different levels, with the national level having a slightly less significant relationship and a smaller coefficient.

These findings provide evidence for the hypothesis that as the proliferation of evangelical churches in America increases, the evangelical share of churches in a county will have a positive impact on Republican presidential vote shares. This finding is not statistically significant for elections preceding 1980, but the presidential election of 1980 and subsequent elections support the hypothesis. Furthermore, these effects are more pronounced in southern states than non-southern states. There remains the possibility that these effects are overestimated by the model and that other primary indicators of identity voting behavior like race—particularly a southern county’s tendency to indicate “whiteness” as a stronger self-identification method—to conflate the results of this particular model (McKee & Springer 2015). Finally, the use of time-varying covariates in both the national model proved to be essential to the regression, as demonstrated in the differences exhibited by choosing to include and exclude the covariates on the outcome variable of Republican presidential vote share.

X. Conclusion

This study examined how changes in evangelical-Christian church shares within counties in the United States have affected voting patterns in presidential elections since 1952. When the model extends to the country as a whole, a strong, positive relationship between the dependent and independent variables is demonstrated. These models also show that, as a whole, mainline-Christian church shares do generally not impact Republican vote shares within counties. In the case of the American south, the relationship between evangelical church proliferation and Republican vote shares is further emphasized, but there also arises a statistically significant
relationship between mainline church proliferation and Republican vote shares, leading to questions about the southern region’s voting behaviors and why they tend to differ from the rest of the country. The findings of this study in conjunction allow me to reject the null hypothesis that there is no relationship between evangelical church proliferation and republican voting share at the county level. Instead, these findings demonstrate a significant and robust relationship between the phenomena and support my alternative hypothesis that as the share of evangelical congregations in a county increases, so too will the vote share for Republican presidential candidates in that county.

In light of this paper’s empirical findings, I find further evidence to assert a relationship between changing religious institutions and subsequent changes in voting behavior. However, the discrepancy between national and regional voting behaviors indicates potential caveats in establishing this relationship. There remain many questions about causality regarding voting behavior, particularly when religion is used as an indicator. More data and empirical studies are needed to distinguish the causal pathways in the questions as mentioned above in the discussion section. Studies like this one may also benefit from the inclusion of online church activity or active “televangelism.” This aspect of religious activity went unaccounted for in my research due to a lack of available data and the study’s emphasis on physical religious institutions. That being said, future research aiming to examine religion and voting behavior in the United States may find that utilizing institutional data rather than individual survey data may lead to new and exciting results. For example, using this study’s methods, a researcher may also be able to analyze the effects of other religious institutions on vote shares for local elections or evaluate the influence of the “nons” (those who hold no religious affiliation and are the fastest growing religious graphic in the US) on voting. By introducing these theoretical expectations, this paper
hopes to provide an intriguing avenue for future research on the relationship between evangelical
church proliferation and its subsequent influence on voting behavior in America.
Bibliography


Churches and Church Membership in the United States. Collected by the Association of Statisticians of American Religious Bodies (ASARB) and distributed by the Association of Religion Data Archives (www.theARDA.com)


