Econometric Analysis of Voting Behaviour: Socioeconomic Drivers of the 2016 U.S. Presidential Election

Author: Rishit Pothu | University College London

MSc Data Science and Public Policy (Economics)

Date: 17 January 2025

Executive Summary

This report explores the county-level factors influencing the Republican shift in the 2016 U.S. Presidential election, comparing them to the 2012 election by analysing demographic, economic, and turnout-related variables through regression modelling and simulation techniques.

- **Methodology**:

  - Employed Ordinary Least Squares (OLS) regression to secure accurate estimates, backed by thorough diagnostics and bootstrapping.
  - Carried out Monte Carlo simulations to verify the stability of coefficients and examine sampling processes variability.

- **Key Findings**:

  - **Demographics:** An increase in the population proportions of individuals aged 20-29, males, and whites was strongly linked to swings toward the Republican vote.
  - **Economic Factors:** Higher unemployment rates were positively correlated with Republican support, whereas increasing per-capita income decreased Republican vote share.
  - **Voter Turnout:** Higher turnout significantly impacted Republican successes, especially in rural and economically disadvantaged regions.
  - **Geographic Trends:** Rural counties tended to favour Republican candidates, while urban areas exhibited a slight resistance to this trend.
  - **Education:** Unlike usual trends, an increase in the number of individuals earning bachelor's degrees showed a minor yet positive correlation with Republican voting trend swings.
  - The model accounted for 19% of the variation in voting behaviour, with diagnostics affirming the robustness of its results. Monte Carlo simulations and bootstrapping reinforced the stability of coefficient estimates.
  - Some limitations exist, such as slight deviations from OLS assumptions and possible effects of unmeasured factors. These findings provide important insights into electoral dynamics, emphasising the complex interactions among demographic, economic, and geographic influences on voter behaviour.

Contents Page

Introduction

The 2016 U.S. Presidential election marked a notable change in voting behaviour from 2012. Donald Trump (Republican) defeated Hillary Clinton (Democrat), overturning Barack Obama's Democratic victory over Republican Mitt Romney in 2012. This shift prompts intriguing questions about the factors contributing to the Republican Party's resurgence. Understanding these dynamics is essential for scholarly evaluation and guiding future political strategies and policymaking.

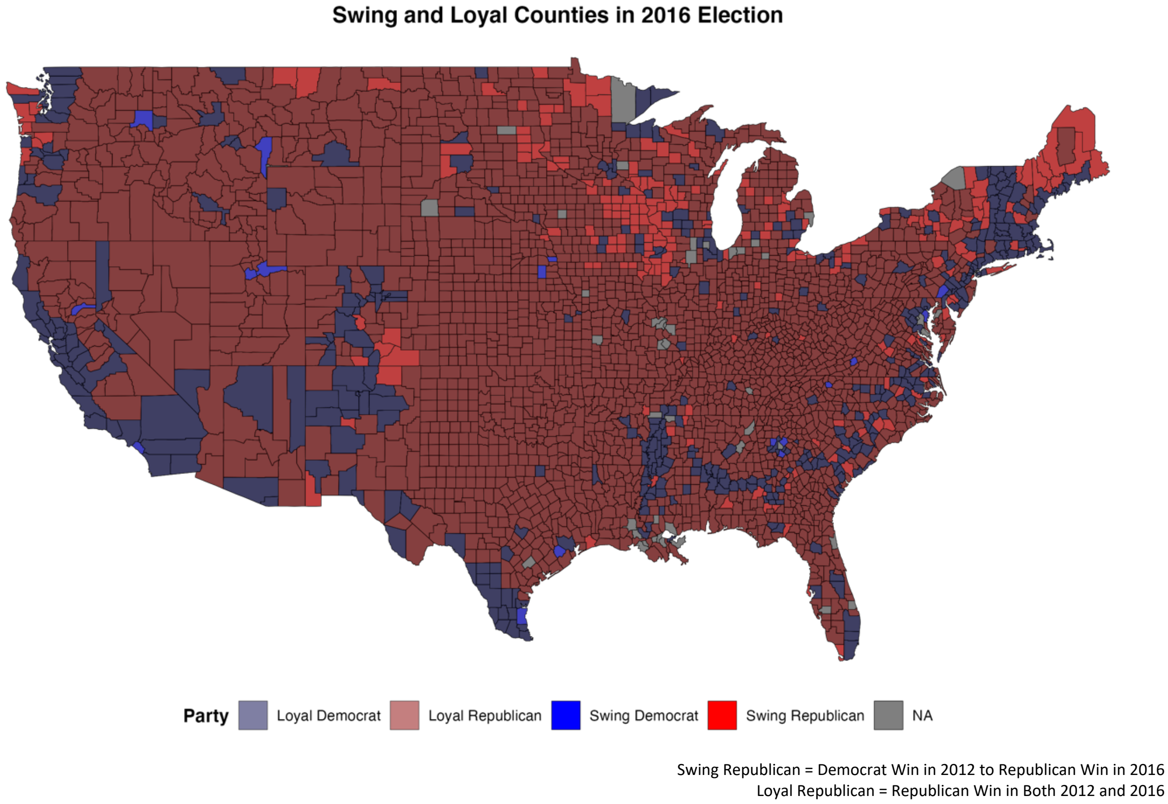

Figure 1: County-Level Election Results Map showing Swing Counties (which changed party from 2012 to 2016) and Loyal Counties (which remained with the same party in both elections)

Bankert (2020) emphasises that the 2016 election was characterised by significant negative partisanship, where voters were driven more by their disapproval of the opposing party than by their endorsement of their own. This trend fuelled polarisation and led to greater acceptance of unconventional candidates such as Trump, as illustrated in Figure 1, which shows that several states that voted Democrat in 2012 shifted to Republican in 2016.

This report delves further into the factors driving the swing toward the Republican Party in the 2016 U.S. Presidential election at the county level. It commences with constructing a comprehensive dataset that integrates voting data from 2012 and 2016. A regression model is subsequently proposed to analyse the relationship between these variables and changes in voting behaviour, identifying key drivers such as demographics, income, and turnout. Finally, a simulation study investigates the sampling distributions of the model’s coefficients, offering insights into the reliability and variability of the estimates.

1 Data

1.1 Dependent and Control Variables

All data is available at the county level from 6 November 2012 to 8 November 2016, which are the election dates. To keep this model simple but interpretable, these key variables are chosen:

Dependent Variable:
- Republican Swing Rate: Change in the Republican vote share.

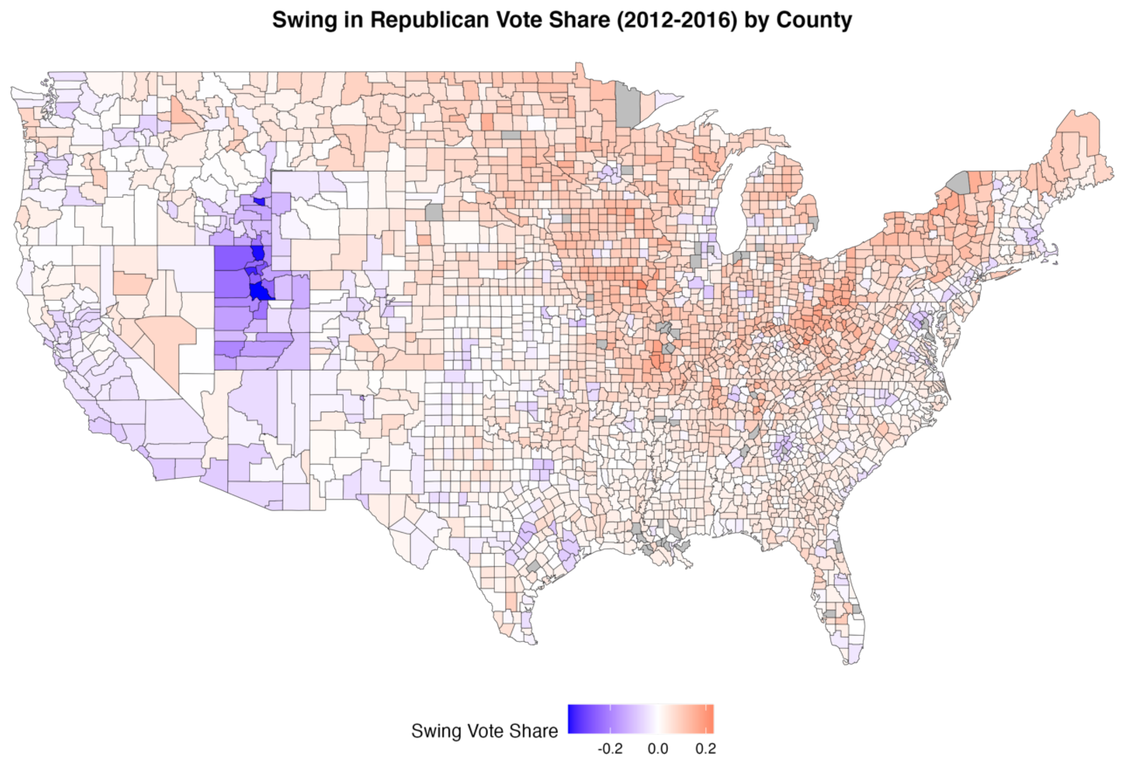

Figure 2: Republican County Vote Share Change (Swing) from 2012 to 2016

Control Variables:
1. Demographic Variables:
    - Age: Change in the proportion of the population aged 20-29.
    - Gender: Change in the proportion of the male population.
    - Race: Change in the proportion of the white population.

2. Voter Turnout: Change in proportion of the population who voted.

3. Geographic Variable: Urban/Rural classification of counties.

4. Economic Variables:
    - Unemployment: Average across election dates.
    - Income: Change in per-capita personal income

5. Education: Change in the proportion of the population with a bachelor's degree or higher

The data is expressed as decimals representing percentage changes, except for the geographic variable, which is binary. I opted for the unemployment metric to be a four-year average because it offers a more stable and representative view of economic conditions throughout the period. This way, voting behaviour may mirror perceptions of long-term economic trends instead of solely reflecting the immediate situation 2016. However, this four-year averaging could obscure recent improvements or declines in unemployment, potentially diluting the link to voting behaviour.

1.2 Rationale for Variable Selection

Demographic Variables

Age greatly influences voting behaviour, revealing unique trends among various age groups. According to Waiphot Kulachai et al. (2023), younger voters tend to favour more progressive ideologies, backing liberal parties or candidates, while older voters often prefer conservative choices. This contrast stems from generational experiences, life-cycle effects, and differing priorities at various stages of life. As people age, their political socialisation and life experiences may alter their voting preferences.

Deckman and Cassese (2019) examine the influence of party affiliation, social class, and perceptions of masculinity on voting behaviour, presenting the idea of "gendered nationalism". Silva, N.M. (2020) highlights that voters' perspectives on race and gender were crucial in determining the 2016 election results. The study reveals that individuals with elevated levels of racial resentment and sexism were more inclined to support Trump, and these attitudes proved to be stronger predictors of voting behaviour than economic discontent.

Voter Turnout

Morgan S and Lee J (2017) highlight the significance of voter turnout, noting that increased mobilisation in rural and deindustrialised areas was crucial in changing the outcome of key swing states. In these rural and economically challenged regions, voter turnout and Republican support surged compared to 2012. In contrast, urban areas saw a slight drop in Democratic turnout, which intensified the impact of rural mobilisation.

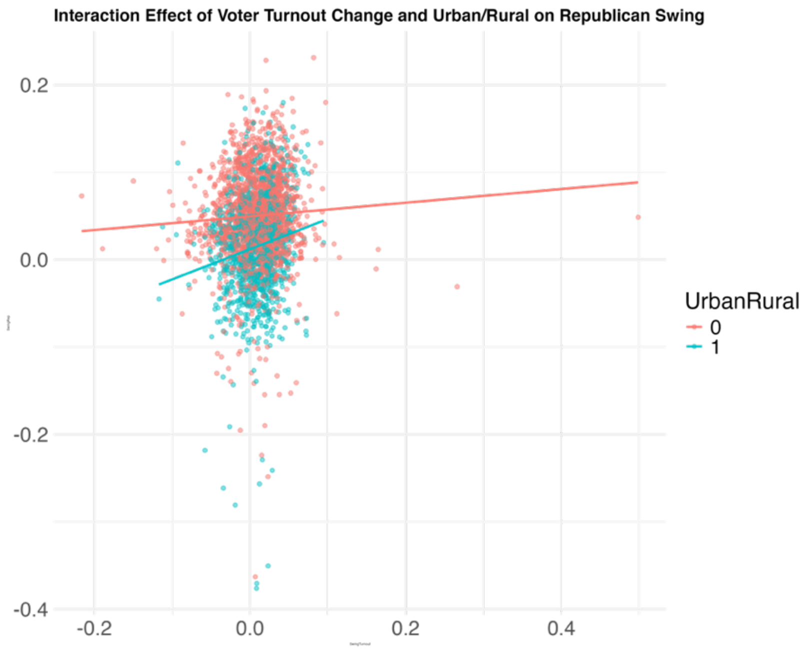

Figure 3: Voter Turnout and Urban/Rural Change Interaction Effect on Republican Swing

Geographic Variable

According to Brown T.E. (2023), voting patterns can differ significantly between urban and rural areas due to variations in socioeconomic conditions. Since the 1970s, there has been a gradual shift in rural areas towards the Republican Party, contrasting with urban areas that have increasingly supported the Democrats. This trend is clearly demonstrated by Figures 4 and 5 from the 2016 election, which show that rural counties significantly backed Donald Trump, while urban counties favoured Hillary Clinton.

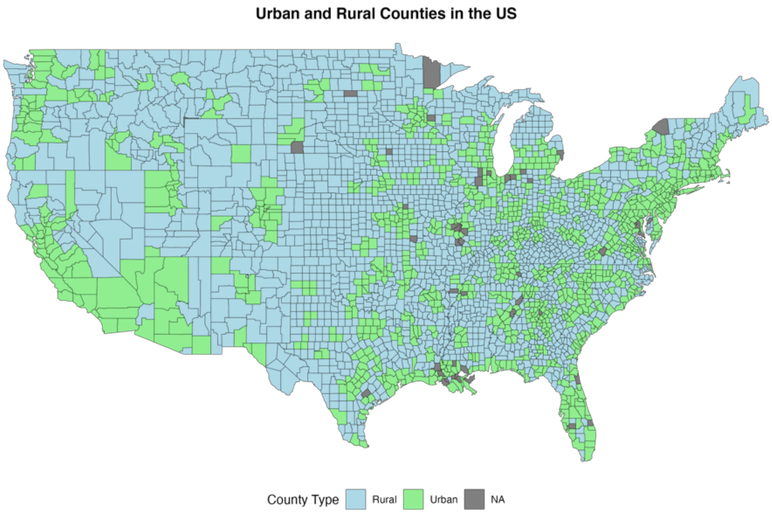

Figure 4: Urban/Rural Counties

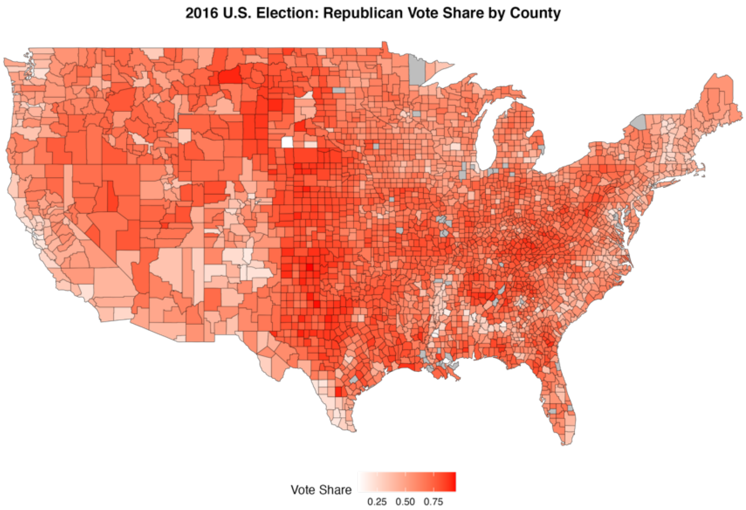

Figure 5: Republican County Vote Share

Economic Variables

Anson, I.G. (2015) showcases that voters often respond to economic trends, rewarding or punishing incumbents based on economic performance. Economic distress and changes in income have been linked to increased support for populist or opposition candidates. Incorporating variables like income and unemployment captures these dynamics.

Education

Educational achievement correlates with socioeconomic status, influencing cultural values and political behaviour. Individuals with higher education adopt progressive views due to exposure to diverse perspectives, supporting liberal attitudes and Democratic candidates. Conversely, those without a degree focus more on security concerns, aligning with Republican ideologies.

1.3 Merging Data

County-level demographic information was obtained from the National Cancer Institute, which provides population figures for each county. I integrated this with election data from Harvard Dataverse, which includes the total votes cast in each county, and then divided that by the population to calculate voter turnout. Furthermore, the election data contains Republican voting figures, from which I determined the respective shares. Urban and rural classifications, along with education statistics, were sourced from the USDA, while personal income per capita data came from the BEA. These datasets were combined using state and county FIPS codes to accurately identify unique counties. Finally, unemployment data was retrieved via the FRED API and merged using state FIPS and county names.

1.4 Check for Multicollinearity

Table 1: Variance Inflation Factor for Voting Model

	VIF
Proportion Change of Ages 20-29	1.139983
Proportion Change of Males	1.153483
Proportion Change of White Population	1.102268
Voter Turnout Change	1.061921
Unemployment	1.094170
Urban/Rural	1.055613
Proportion Change of Per-Capita Personal Income	1.135122
Proportion Change of Population with Bachelor's Degree or Higher	1.016595

The low VIF values indicate that multicollinearity does not pose a problem for this model. Consequently, all predictors present can accurately measure their effects on the dependent variable without significant issues related to redundancy or inflated standard errors. This enhances the credibility of the regression coefficients' interpretability and assures that the model's findings are statistically sound.

2 Voting Behaviour Model

2.1 Rationale for Model Selection

All the relevant variables are continuous—except for one, which is binary. Therefore, I employed an Ordinary Least Squares (OLS) regression model tailored for continuous outcomes. OLS minimises squared error loss, aligning with maximum likelihood estimation (MLE) under the assumption of normally distributed residuals, which provides interpretable coefficient estimates that measure the effect of each predictor on changes in vote share, facilitating the explanation of how demographics, economics, and turnout-related factors influence voting behaviour. This avoids the need for explicitly implementing MLE or optimisation techniques like gradient descent, as they are already embedded in the computational algorithms used in R. Numerous studies examining voting behaviour and electoral dynamics utilise OLS to assess the influence of critical predictors on changes in vote share (e.g., Tuttnauer, O. and Wegmann, S. (2022)).

Initial checks revealed perfect multicollinearity among age-related predictors (swingages_20_29, swingages_30_44, swingages_45_64, and swingages_65_plus). To address this, only swingages_20_29 were retained based on their statistical significance and theoretical relevance. This is also the reason for choosing only male % for gender and white % for race. The VIF analysis confirmed that the final set of predictors had low multicollinearity, ensuring stable and reliable coefficient estimates, making alternatives such as the Lasso model redundant to use as there is no need to impose any additional regularisations.

The following model follows these key assumptions:

1. **Linearity**: The relationship between Republican Swing and the predictors is linear.
2. **Homoscedasticity**: The variance of residuals (ϵ) is constant across all values of the predictors.
3. **Independence**: Observations (counties) are independent of one another.
4. **Multicollinearity**: Predictors are not highly correlated.

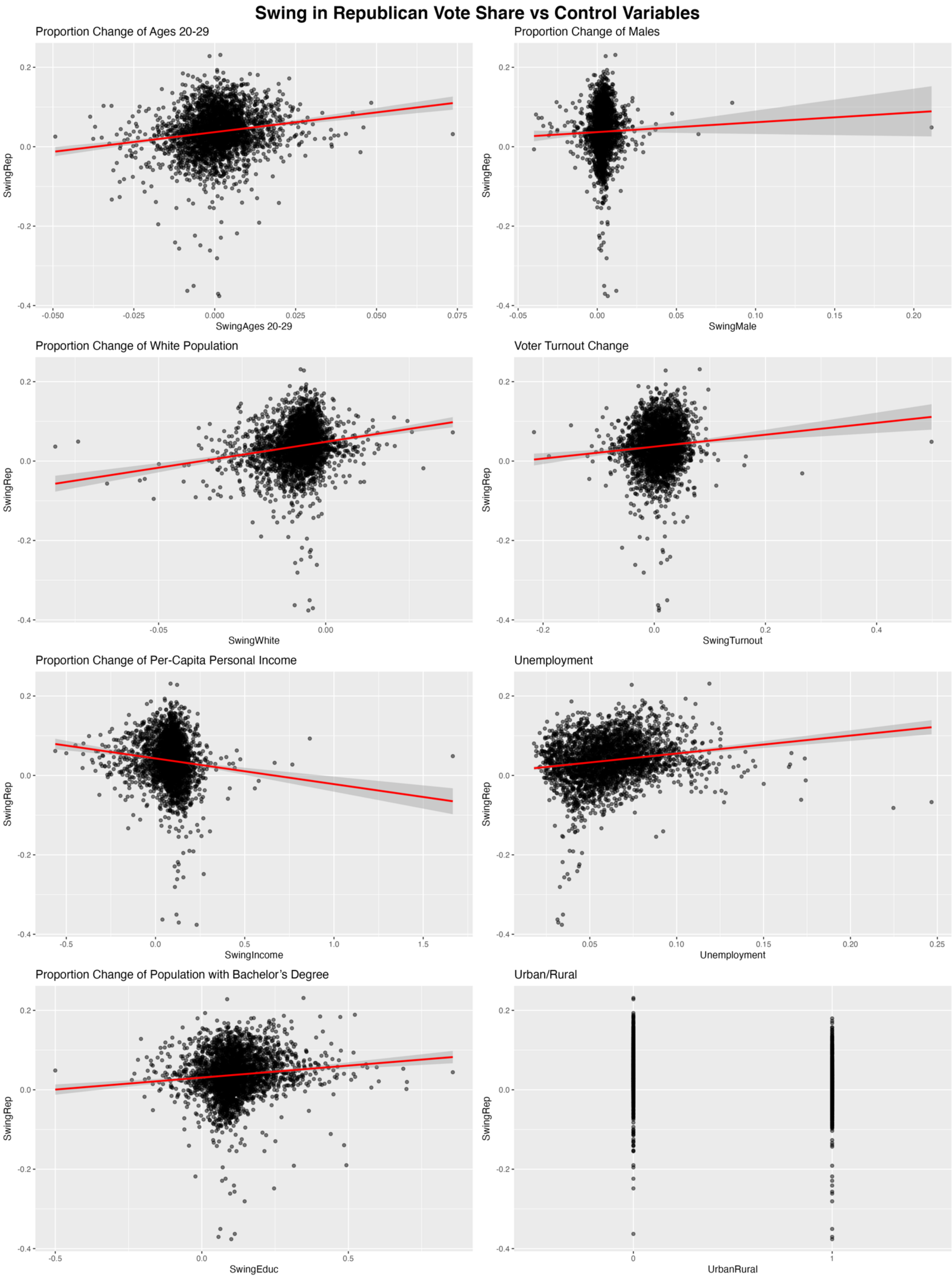

Figure 6: Republican Vote Share Swing and Control Variables Scatter Plot

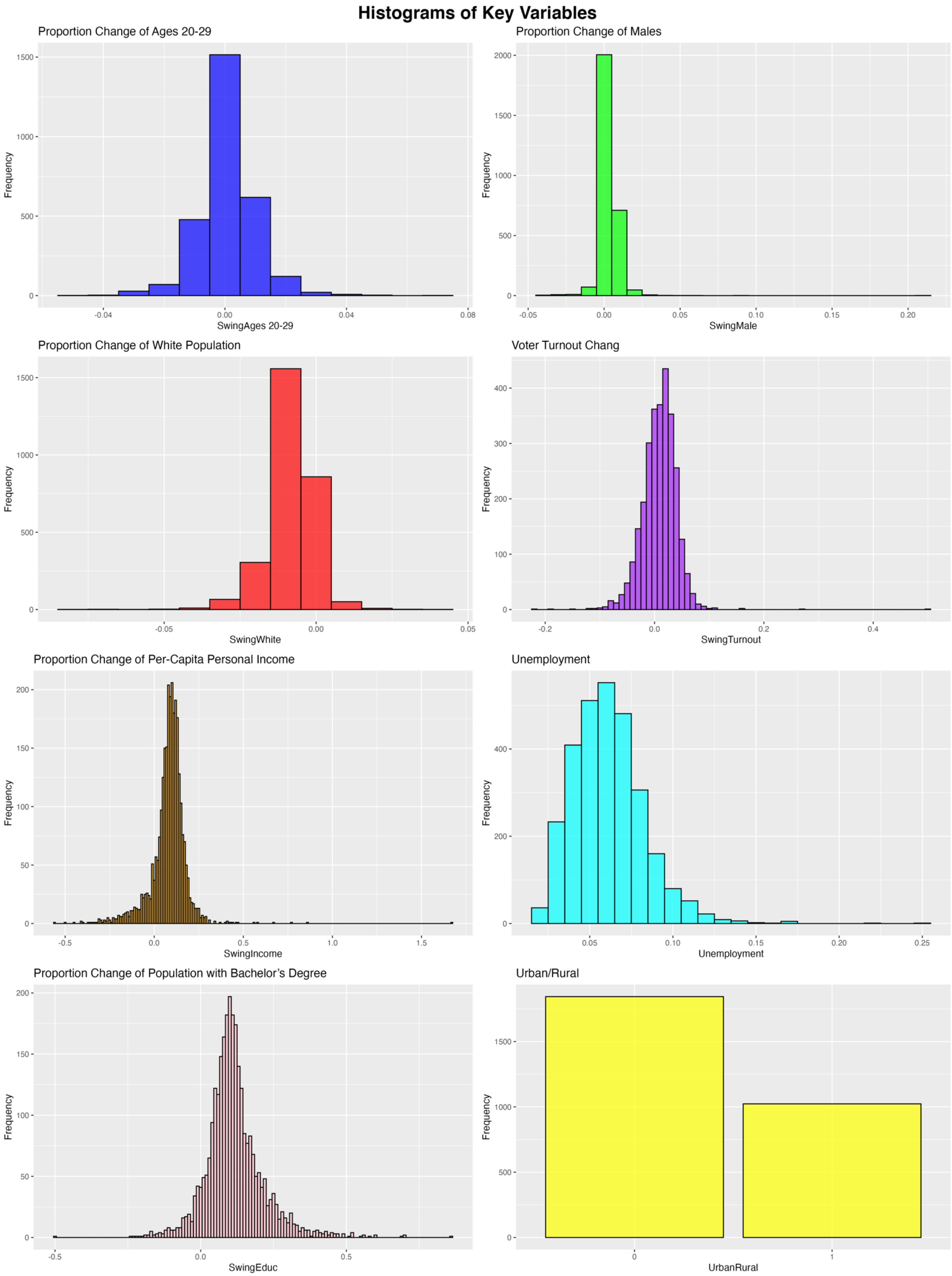

Figure 7: Histograms of Control Variables

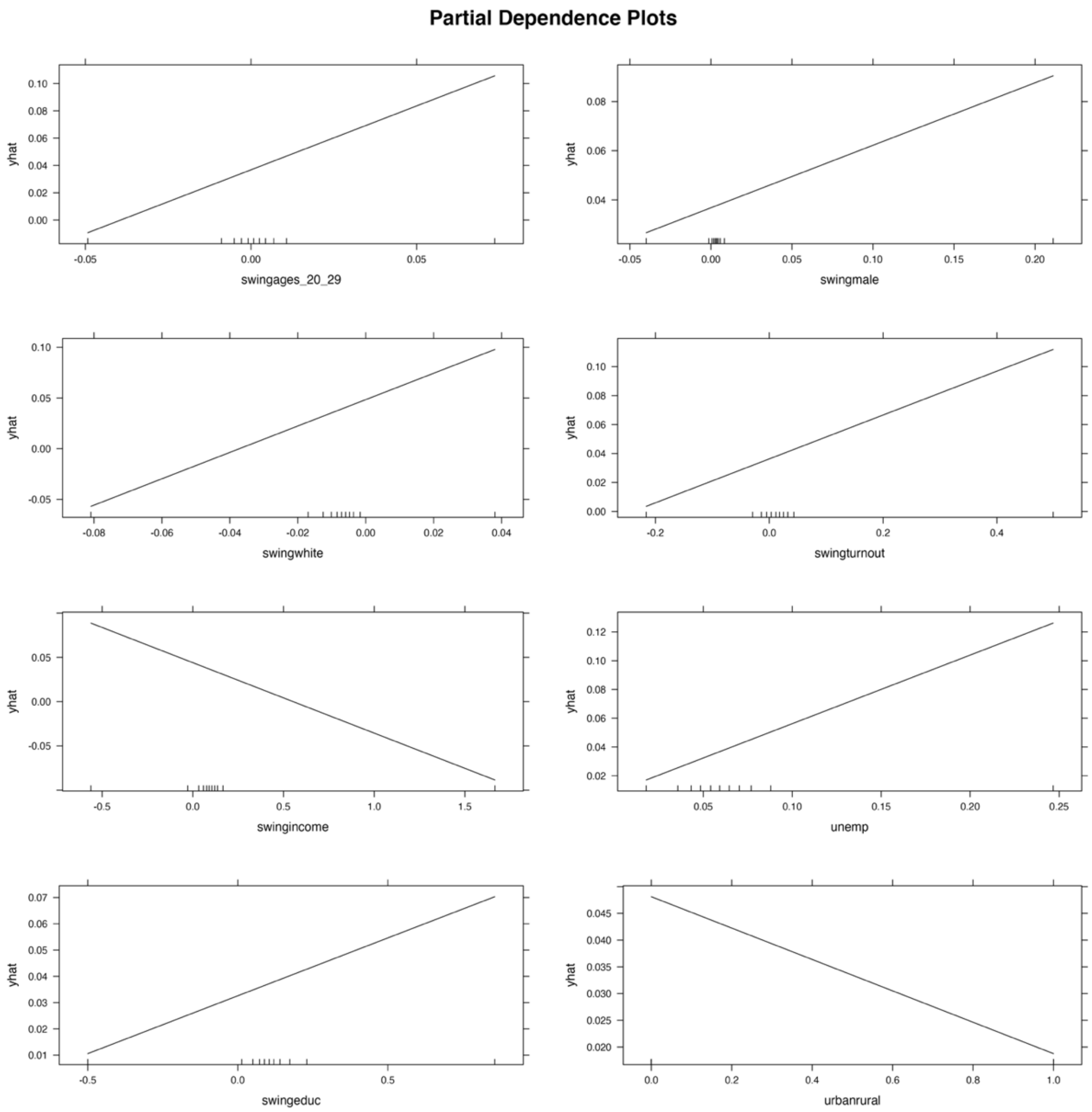

Figure 8: Partial Dependence Plots of Control Variables on Republican Vote Share Swing

The scatter plot in Figure 6 reveals positive trends among the 20-29 age group, the white population, and the male population, indicating that an increase in these demographics corresponds to a higher swing in Republican votes. Additionally, a positive trend suggests increased voter turnout and educational attainment are associated with higher Republican vote swings. Furthermore, rural counties display a higher swing towards Republicans, while income displays a negative slope, suggesting that wealthier areas swung less Republican. Lastly, a positive correlation highlights that higher unemployment correlates with a Republican swing. The partial dependence plots in Figure 8 further reinforce Figure 6's results.

Figure 7 reveals that the 20-29 age group, white and male population, are centred around zero, indicating minimal change in these demographics across counties. Furthermore, the factors of Income and Education showcase a wider spread, suggesting greater variability in these characteristics among counties.

Table 2: Summary Statistics of all Variables (Between 2012 and 2016)

Observations	(N = 2867)
Swing in Republican Vote Share (%)	
Mean	3.76%
Median [Min, Max]	3.73% [-37.62%, 23.12%]
Proportion Change of Ages 20-29 (%)	
Mean	0.091%
Median [Min, Max]	0.081% [-4.93%, 7.36%]
Proportion Change of Males (%)	
Mean	0.33%
Median [Min, Max]	0.321% [-3.99%, 21.12%]
Proportion Change of White Population (%)	
Mean	-0.825%
Median [Min, Max]	-0.713% [-8.09%, 3.80%]
Voter Turnout Change (%)	
Mean	0.89%
Median [Min, Max]	1.16% [-21.63%, 49.88%]
Unemployment (Average Throughout 4 Years) (%)	
Mean	6.11%
Median [Min, Max]	5.92% [1.79%, 24.67%]
Urban/Rural (Binary, 1 = Urban, 0 = Rural)	
Mean (% of Counties classified as Urban)	35.68%
Proportion Change of Per-Capita Personal Income (%)	
Mean	7.96%
Median [Min, Max]	9.06% [-56.18%, 166.59%]
Proportion Change of Population with Bachelor's Degree or Higher (%)	
Mean	11.48%
Median [Min, Max]	10.48% [-50%, 85.63%]

2.2 Model Specification

The OLS regression model is as follows:

$$\text{Republican Swing}_i = \beta_0 + \beta_1 \text{Ages 20-29}_i + \beta_2 \text{Male}_i + \beta_3 \text{White}_i + \beta_4 \text{Voter Turnout}_i + \beta_5 \text{Unemployment}_i + \beta_6 \text{Urban/Rural}_i + \beta_7 \text{Income}_i + \beta_8 \text{Education}_i + \epsilon_i$$

With the results below:

2.3 Results of the Voting Behaviour Model

Table 3: OLS Regression Results

	OLS Voting Model
Intercept	0.027907 (0.003592) ***
Proportion Change of Ages 20-29	0.935444 (0.112624) ***
Proportion Change of Males	0.254150 (0.150042) •
Proportion Change of White Population	1.300654 (0.133563) ***
Voter Turnout Change	0.151597 (0.030916) ***
Unemployment	0.477059 (0.046328) ***
Urban/Rural	-0.029359 (0.002048) ***
Proportion Change of Per-Capita Personal Income	-0.079521 (0.010102) ***
Proportion Change of Population with Bachelor's	0.044049 (0.009681) ***
R^2	0.1907
Adj. R^2	0.1885
Observations	2858
Residual Std. Error	0.05113
F-Statistic	84.2 (8 and 2858 df) (p-value < 2.2e-16)

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; • $p < 0.1$

2.4 Takeaways of Model

All results are statistically significant except the male coefficient, which has a p-value less than 0.1, indicating that it is weaker than other factors. A one-unit rise in individuals aged 20-29 yields a 0.935 percentage-point swing in Republican votes, emphasising this age group's political impact. Conversely, a one-unit increase in the white population leads to a 1.3 percentage-point advantage for Republicans, highlighting racial composition's effect on voting behaviour. Voter turnout significantly influences results, with a one-unit increase prompting a 0.152 percentage-point swing toward Republicans, stressing the importance of participation. Economic factors also matter; a rise in unemployment corresponds to a 0.477 percentage-point shift in Republican support, showing the link between economic conditions and voter preferences. Urban counties, on the other hand, see a slight 0.029 percentage-point decrease in Republican vote swing compared to rural counties, reflecting established trends of rural areas leaning more Republican. Additionally, per-capita personal income increases correspond to a 0.08 percentage-point decrease in Republican support, suggesting that rising incomes, particularly in urban settings, may correlate with more liberal views. Furthermore, a one-unit increase in bachelor's degree attainment is associated with a 0.044 percentage-point swing towards Republican votes, albeit with a smaller effect size than other predictors.

The model explains approximately 19.07% of the variation in Republican vote swings, indicating that while notable, there are likely unobserved factors still influencing these outcomes.

2.5 Model Fit Evaluation

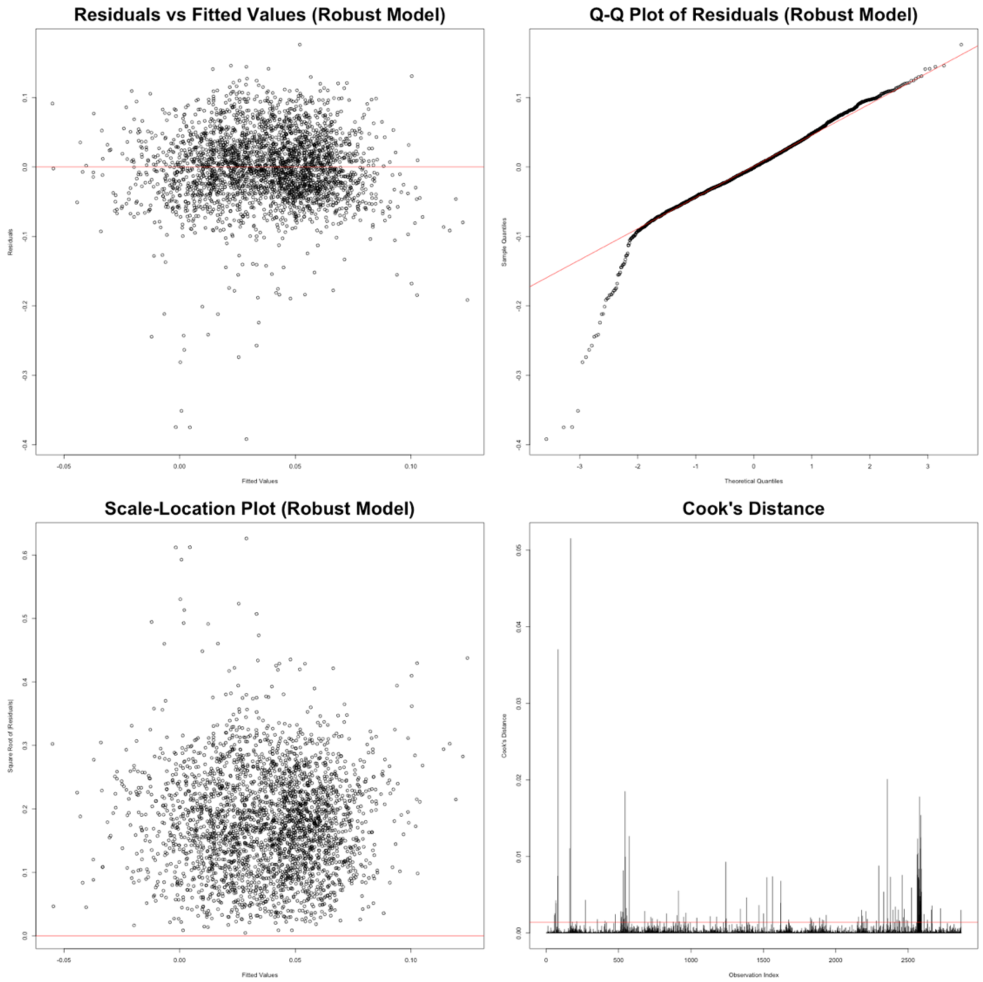

Figure 9: Residual Diagnostics of Robust OLS Model

After discovering mild heteroscedasticity and non-linearity, a robust model was incorporated by recomputing the standard errors using heteroscedasticity-consistent estimators on the original OLS model.

The Residuals vs. Fitted Values plot shows randomly distributed residuals around zero, confirming the validity of linearity and effective management of heteroscedasticity. The Scale-Location plot illustrates a consistent residual spread across fitted values, indicating constant variance. In the Q-Q plot, most residuals follow a normal distribution, closely aligning with the diagonal reference line, though slight deviations at extremes suggest some outliers. Lastly, Cook's Distance plot notes a few influential observations with higher leverage, but they remain within acceptable limits and do not distort the model significantly, indicating minimal influence from most observation values.

2.6 Bootstrap

The model was further validated using bootstrapping to estimate the sampling distributions of the coefficients, providing robust insights into their variability and statistical significance.

Table 5: Bootstrap Results

	Original Coef of OLS Model	Mean	Bias	Std. Error
Intercept	0.0279	0.0278	-0.0001	0.0044
Proportion Change of Ages 20-29	0.9354	0.9398	0.0043	0.1183
Proportion Change of Males	0.2542	0.2587	0.0045	0.138
Proportion Change of White Population	1.3007	1.3082	0.0076	0.1288
Voter Turnout Change	0.1516	0.1511	-0.0005	0.0313
Unemployment	0.4771	0.4794	0.0023	0.0639
Urban/Rural	-0.0294	-0.0294	0	0.0021
Proportion Change of Per-Capita Personal Income	-0.0795	-0.0797	-0.0002	0.0107
Proportion Change of Population with Bachelor's Degree or Higher	0.044	0.0439	-0.0001	0.0104

The bootstrap findings validate the reliability of the OLS regression estimates. Most variables show minimal bias and acceptable standard errors, demonstrating the model's stability. Nevertheless, variables such as Proportion Change of Ages 20-29, Proportion Change of Males, and Proportion Change of White Population display somewhat increased variability, indicating their notable and context-dependent impact on the fluctuations in Republican vote share.

3 Simulation Study

3.1 Features of the Environment

1. **Sample Size (*n* = 100):**
   The number of observations in each simulation run is set to 100 to mimic realistic data scenarios where sample sizes are moderate. This choice ensures that the results reflect typical datasets used in voting behaviour studies.
2. **Number of Predictors (*p* = 8):**
   The predictors are based on the variables used in the original regression model.
3. **True Coefficients:**
   The true coefficients are derived from the original regression model estimates. This ensures that the simulation reflects the observed relationships in the actual dataset and provides insights into how these coefficients vary under repeated sampling.
4. **Residual Standard Error:**
   The residual standard error from the original model is used to generate random noise. This maintains consistency between the variability observed in the real data and the simulated data.

5. **Monte Carlo Iterations (*simulations* = 10,000):**
   Many simulation iterations provide precise estimates of the sampling distributions, ensuring the results are robust and statistically reliable.
6. **Data Generation Process:**
   Predictors are generated independently with standard normal distributions, except for urban-rural, which is a binary variable to reflect its real-world nature. The dependent variable is calculated as a linear combination of predictors, true coefficients, and random noise.

3.2 Rationale for Design Choices

This setup replicates real-world data features while enabling control over parameters. By aligning the true coefficients and error structure with the observed model, the simulation produces meaningful insights into the stability and reliability of the regression estimates. It also allows an examination of bias and variability across coefficients, which is crucial for understanding the robustness of the model.

3.3 Results and Comparison

Table 6: Simulation Results

	Original Coef of OLS Model	Mean	Bias	Std. Dev
Intercept	0.0279	0.0278	-0.0001	0.0076
Proportion Change of Ages 20-29	0.9354	0.9355	0	0.0053
Proportion Change of Males	0.2542	0.2542	0	0.0054
Proportion Change of White Population	1.3007	1.3997	−2.374e−6	0.0055
Voter Turnout Change	0.1516	0.1517	-0.0001	0.0055
Unemployment	0.4771	0.4771	0	0.0054
Urban/Rural	-0.0294	-0.0294	-2.6128e-8	0.0107
Proportion Change of Per-Capita Personal Income	-0.0795	-0.0794	0.0001	0.0055
Proportion Change of Population with Bachelor's Degree or Higher	0.044	0.0441	0	0.0054

The simulation results closely align with the original OLS estimates, with the mean values of the simulated coefficients matching the observed coefficients closely. Minor bias values indicate consistency with the original model, while the standard deviations showcase variability in the coefficients across the simulated datasets. The variable "Proportion Change of Ages 20-29" demonstrates the least variability, whereas the binary variable "Urban/Rural" shows the most.

In contrast, bootstrap standard errors are typically larger than those derived from simulations, especially for variables like "swingmale" and "urbanrural." This underscores Bootstrap's sensitivity to variability and outliers that are tied to specific samples, as it is directly dependent on the empirical dataset.

As shown in Figure 10, the sampling distributions from both approaches centre around the original OLS coefficients, confirming their consistency with the observed data and the presumed true model. Both methods produce unimodal and symmetric distributions; however, the bootstrap distributions occasionally exhibit heavier tails (for example, with swingeduc and urbanrural), indicating data-driven variability. The simulation's controlled environment results in narrower distributions for "Ages 20-29" and "White Population," while the binary nature of "Urban/Rural" leads to more concentrated bootstrap outcomes.

This dual-pronged methodology offers a thorough assessment of coefficient reliability: simulations evaluate variability under ideal conditions, whereas bootstrapping accounts for real-world dependencies and noise. Combined, they deliver strong insights into the stability of OLS estimates.

4 Limitations

The methods used in the analysis have notable limitations. OLS regression relies on strong assumptions. However, diagnostics indicated mild violations that could impact coefficient estimates and standard errors. Additionally, OLS is sensitive to outliers, which can unduly influence results, and it often oversimplifies complex relationships, especially in voting behaviour datasets.

Bootstrapping is vulnerable to the biases in the original sample, and its computational intensity can be significant with larger datasets. This method is strictly data-driven and lacks a theoretical framework.

Monte Carlo simulations simplify real-world complexities by making assumptions about true coefficients and their independence. These assumptions can accumulate inaccuracies from OLS models into simulated results. Monte Carlo simulations also assume fixed variance and normal errors, which might not reflect actual variability.

Lastly, while partially addressing assumptions like heteroscedasticity, robust model adjustments do not rectify non-normal residuals or model misspecification. These adjustments impact standard errors and confidence intervals but do not alter the coefficient estimates, potentially leaving the underlying issues unaddressed.

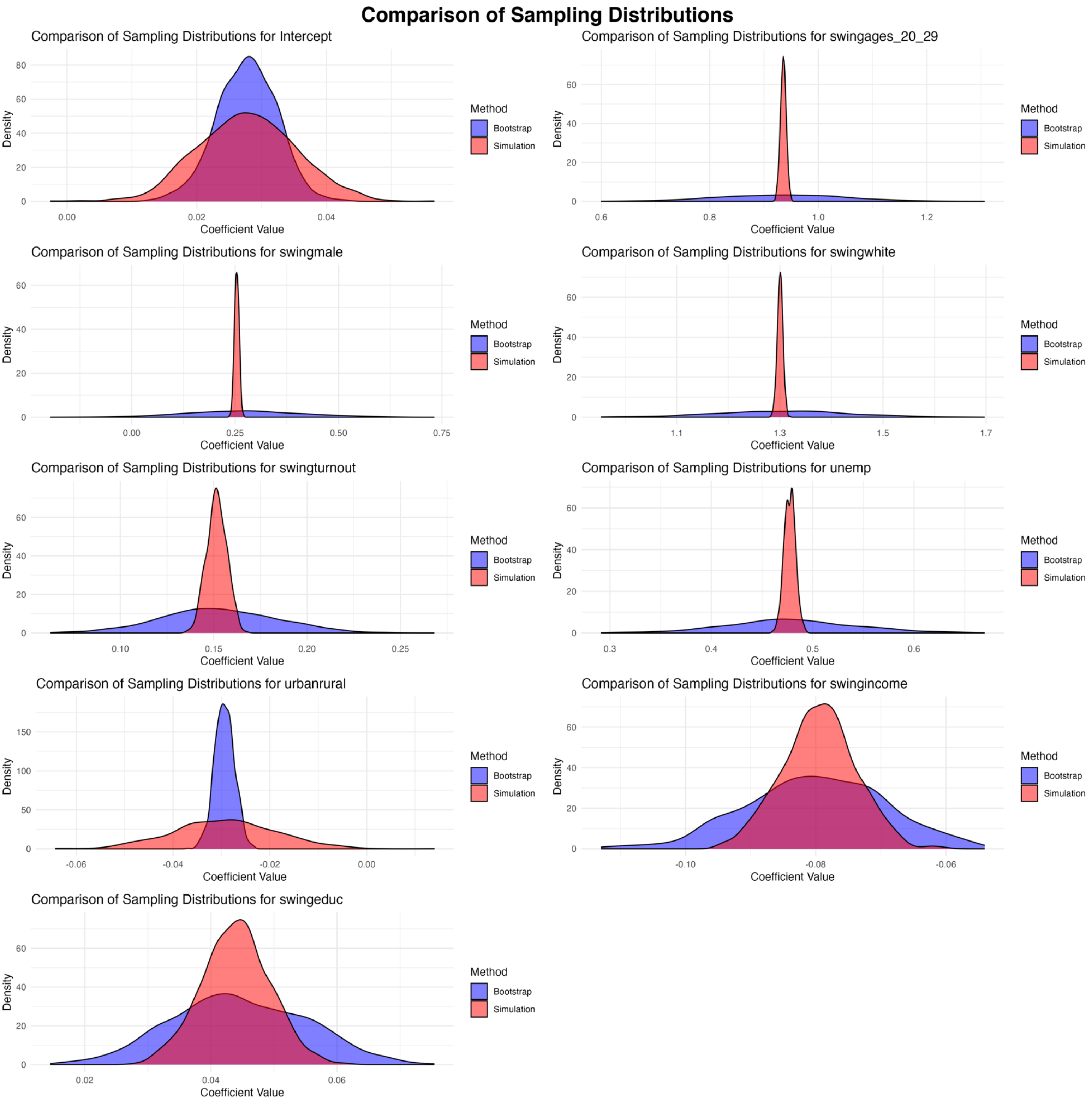

Figure 10: Sampling Distribution of Simulation and Bootstrap

References

Bankert, A. (2020). Negative and Positive Partisanship in the 2016 U.S. Presidential Elections. *Political Behavior*, 43. https://doi.org/10.1007/s11109-020-09599-1.

Webster, S.W. and Pierce, A.W. (2019). Older, Younger, or More Similar? The Use of Age as a Voting Heuristic*. *Social Science Quarterly*, 100(3), pp.635–652. https://doi.org/10.1111/ssqu.12604.

Waiphot Kulachai, Unisa Lerdtomornsakul and Patipol Homyamyen (2023). *Factors Influencing Voting Decision: A Comprehensive Literature Review*. [online] ResearchGate. Available at: https://www.researchgate.net/publication/373387538_Factors_Influencing_Voting_Decision_A_Comprehensive_Literature_Review.

Kim, S.S. and Zilinsky, J. (2022). Division Does Not Imply Predictability: Demographics Continue to Reveal Little About Voting and Partisanship. *Political Behavior*. https://doi.org/10.1007/s11109-022-09816-z.

Kulachai, W., Lerdtomornsakul, U. and Homyamyen, P. (2023). Factors Influencing Voting Decision: a Comprehensive Literature Review. *Social Sciences*, [online] 12(9), pp.1–16. https://doi.org/10.3390/socsci12090469.

Deckman, M. and Cassese, E. (2019). Gendered Nationalism and the 2016 US Presidential Election: How Party, Class, and Beliefs about Masculinity Shaped Voting Behavior. *Politics & Gender*, [online] 17(2), pp.1–24. https://doi.org/10.1017/s1743923x19000485.

Shook, N.J., Fitzgerald, H.N., Boggs, S.T., Ford, C.G., Hopkins, P.D. and Silva, N.M. (2020). Sexism, racism, and nationalism: Factors associated with the 2016 U.S. presidential election results? *PLOS ONE*, 15(3), https://doi.org/10.1371/journal.pone.0229432.

Morgan, S. and Lee, J. (2017). The White Working Class and Voter Turnout in U.S. Presidential Elections, 2004 to 2016. *Sociological Science*, 4, pp.656–685. https://doi.org/10.15195/v4.a27.

Brown, T.E. and Mettler, S. (2023). Sequential Polarization: The Development of the Rural-Urban Political Divide, 1976–2020. *Perspectives on Politics*, pp.1–29. https://doi.org/10.1017/s1537592723002918.

Anson, I.G. and Hellwig, T. (2015). Economic Models of Voting. *Emerging Trends in the Social and Behavioral Sciences*, [online] pp.1–14. https://doi.org/10.1002/9781118900772.etrds0090.

Hendrickson, C. and Galston, W.A. (2016). *The educational rift in the 2016 election*. [online] Brookings. Available at: https://www.brookings.edu/articles/educational-rift-in-2016-election/ [Accessed 17 Jan. 2025].

Tuttnauer, O. and Wegmann, S. (2022). Voting for Votes: Opposition Parties' Legislative Activity and Electoral Outcomes. *American Political Science Review*, [online] 116(4), pp.1357–1374. https://doi.org/10.1017/S0003055422000338.

For Variables (Republican Swing, Vote Turnout):
https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/VOQCHQ

For Variables (County Demographics):
https://seer.cancer.gov/popdata/download.html

For Variables (County Per-Capita Personal Income)
https://apps.bea.gov/regional/downloadzip.htm?_gl=1*1umjl9x*_ga*ODM2MzY4OTAwLjE3MzQ4ODIzMzc.*_ga_J4698JNNFT*MTczNjY1NTkwMi4xMC4xLjE3MzY2NTYyMzIuMTguMC4w

For Variables (Urban/Rural, Education):
https://ers.usda.gov/data-products/county-level-data-sets/county-level-data-sets-download-data