Deciphering the Relationship Betweeen Race, Income, and Political Support in the 2020 United States Election*

Liban Timir

March 12, 2024

In this paper, we delve into the intricate landscape of American politics using the comprehensive Cooperative Election Study (CES) of 2020, focusing on how race and income shape political support and preferences across the nation. Through a detailed analysis, I uncovered significant correlations between these socioeconomic factors and political allegiances, revealing the multifaceted nature of political engagement. Specifically, they both play crucial roles in influencing political behaviors and preferences, highlighting concerns that drive American political engagement. This research not only provides a deeper understanding of the dynamics at play in the current political climate but also emphasizes the importance of considering the varied perspectives and experiences within the American populace to grasp the full spectrum of political support in the United States.

1 Introduction

In the contemporary political landscape of the United States, marked by increasing polarization, understanding the foundations of political support has never been more crucial. The Cooperative Election Study (CES) of 2020 (Schaffner, Ansolabehere, and Luks 2021) presents a dataset to explore this phenomenon. This dataset not only captures a wide array of sociopolitical attitudes and behaviors but also provides insights into the demographic and socioeconomic factors that influence political support. This offers a valuable lens through which to examine the complex dynamics of political engagement.

This study aims to fill a critical gap in the existing literature on political support in the United States by leveraging the CES 2020 data to examine the role of sociodemographic

*Code and data are available at: https://github.com/libant/US-Political-Support-Analysis

factors in shaping political preferences and behaviors. Previous research, while extensive, has often focused on more generalized measures of political engagement or on specific demographic groups, without fully addressing the multifaceted nature of political support across the nation's diverse populace. This paper seeks to dissect the nuanced ways in which factors such as race and income predict political support, thereby providing a more detailed understanding of American political behavior. The estimand is the correlations between socioeconomic factors such as income and political support.

The findings reveal significant correlations between these characteristics (income and race) and political support, highlighting the complexity of political engagement in the United States. These results not only highlight the importance of demographic factors in predicting political preferences but also illustrate the varied nature of political support across different segments of the American population, offering new insights into the priorities and concerns that drive political behavior in the United States.

The paper begins with an overview of the CES 2020 dataset and the methods employed in the data analysis, followed by a presentation of results and key findings. I conclude the paper with a discussion that contextualizes the results within the broader literature on political support in the United States. The analysis was conducted using the statistical programming language R (R Core Team 2023), utilizing packages such as tidyverse (Wickham et al. 2019), dplyr (Wickham et al. 2023), ggplot2 (Wickham 2016), dataverse (Kuriwaki, Beasley, and Leeper 2023), janitor (Firke 2023), rstanarm (Goodrich et al. 2022), modelsummary (Arel-Bundock 2022), broom.mixed (Bolker 2022), knitr (Xie 2023) and additional papers such as Arceneaux's "Do Campaigns Help Voters Learn? A Cross-National Analysis" (Arceneaux 2005), Laurison's "Voting Intersections: Race, Class, and Participation in Presidential Elections in the United States 2008–2016" (Laurison, Brown, and Rastogi 2021), and Rhodes's "Is America More Divided by Race or Class? Race, Income, and Attitudes among Whites, African Americans, and Latinos"(Rhodes, Schaffner, and McElwee 2017).

2 Data

2.1 Measurement {sec-measurement}

2.2 Race

Schaffner, Ansolabehere, and Luks (2021) asked validated American voters "what racial or ethnic group best describes them" based on an 8-item demographic distribution. The survey respondents indicated their responses based on the following options:

1. White
2. Black
3. Hispanic

4. Asian
5. Native American
6. Middle Eastern
7. Two or more races
8. Other

The distribution of survey respondents by race offers a generally diverse portrait of the study's participants, which is crucial for ensuring that the findings are reflective of a diverse population. The largest group of respondents identified as White (N = 44,128), followed by Black (N = 6,952), Hispanic (N = 5,180), and Asian (N = 1,831), with smaller representations from Native American (N = 471), Middle Eastern (N = 84), those identifying with two or more races (N = 1,349), and other racial categories (N = 1,005), totaling a respondent pool of roughly 61,000.

This racial distribution is significant in ensuring that analyses of societal issues, such as political preferences or policy impacts, are examined through the lens of racial diversity. By acknowledging the racial composition of the respondents, the study strengthens the validity of its conclusions across the different racial and ethnic experiences present within the sample population.

2.3 Income

Schaffner, Ansolabehere, and Luks (2021) asked validated American voters "what was their family's annual income" based on a 17-item demographic distribution. The survey respondents indicated their responses based on the following options:

1. Less than $10,000
2. Between $10,000 and $20,000
3. Between $20,000 and $29,999
4. Between $30,000 and $39,999
5. Between $40,000 and $49,999
6. Between $50,000 and $59,999
7. Between $60,000 and $69,999
8. Between $70,000 and $79,999
9. Between $80,000 and $99,999
10. Between $100,000 and $119,999
11. Between $120,000 and $149,999
12. Between $150,000 and $199,999
13. Between $200,000 and $249,999
14. Between $250,000 and $349,999
15. Between $350,000 and $499,999
16. $500,000 or more
17. Prefer not to say

The distribution of survey respondents by income offers a representation across various income categories, capturing the economic diversity within the participant pool. The largest group of respondents identified their income as between $50,000 and $99,999 (N = 18,014), followed by less than $30,000 (N = 14,163), between $30,000 and $49,999 (N = 10,789), between $100,000 and $199,999 (N = 9,780), and $200,000 and above (N = 2160), along with those that preferred not to say (N = 6,075), totaling a respondent pool of roughly 61,000.

This income distribution is significant in contextualizing the survey results, as economic status often correlates with perspectives on policy, access to resources, and overall life satisfaction. It serves as a critical variable for examination with other demographic factors such as race, enhancing the depth and applicability of the study's findings across different income groups. By encompassing a wide range of income levels, the survey ensures a comprehensive analysis of the correlation between income and key factors such as education and health, thus supporting nuanced interpretations of the data collected.

2.4 Voter Registration, Turnout, and Political Preferences

Schaffner, Ansolabehere, and Luks (2021) asked validadted American voters "are they registered to vote" based on a 3-item demographic distribution. The survey respondents indicated their responses based on the following options:

1. Yes
2. No
3. Don't know

Schaffner, Ansolabehere, and Luks (2021) also asked "which candidate did they prefer for President of the United States" based on a 4-item demographic distribution. The survey respondents indicated their responses based on the following options:

1. Joe Biden (Democratic)
2. Donald Trump (Republican)
3. Other
4. Not sure

In regards to voting turnout, Schaffner, Ansolabehere, and Luks (2021) asked "which of the following statements best describes them" based on a 5-item demographic distribution. The survey respondents indicated their responses based on the following options:

1. I did not vote in the election this November
2. I thought about voting this time – but didn't
3. I usually vote, but didn't this time
4. I attempted to vote but did not or could not
5. I definitely voted in the November 2020 General Election

The distribution of survey respondents captures the voting distribution within the participant pool. Voter turnout was high (N = 45,660), along with a high voter registration (N = 48,225). In regards to the vote itself, out of 5,861 respondents, the majority voted for Biden (N = 2,324).

The significance of this is that there is a considerable level of civic engagement among the survey participants, which is crucial for drawing conclusions about electoral behaviors and political preferences. Additionally, the data shows a significant majority of the survey population are not just eligible voters but are also taking the necessary steps to exercise their voting rights. When it comes to the political preferences themselves, the data is critical as it can be used to inform the public's response to political platforms, policy preferences, and social issues that resonate with the candidate. High turnout and engagement in this context could also point to a heightened political awareness and mobilization around the time of the survey, presumably influenced by the events at the time.

2.5 Data Source, Measurement and Characteristics

This paper utilizes data published by the Cooperative Election Study (CES) based on the U.S 2020 election (Schaffner, Ansolabehere, and Luks 2021). The study aimed to collect representative data from a diverse cross-section of the American population, adhering to the highest ethical standards to ensure participant confidentiality and data integrity. The dataset was last updated on February 14th 2022. The interest in understanding the dynamics of American political behavior during the pivotal 2020 presidential election led to the creation of the CES 2020 dataset. However, a notable limitation of this dataset comes from its method of data collection, which was done by 60 research teams composed of a variety of academic institutions and organizations. Consequently, the dataset primarily captures a broad perspective on the nation's politics, potentially overlooking more granular, local-level political dynamics. This limitation highlights the inherent challenges in capturing comprehensive data on political behavior that span across the diverse and increasingly divided United States. Additionally, the dataset is designed to ensure the anonymity of its respondents, as it collects information on political preferences, voting behavior, and demographic characteristics without including any personally identifiable information, thereby maintaining the confidentiality of individuals' political beliefs.

The CES 2020 dataset contained a sample size of approximately 61,000 cases. This extensive collection reflects a broad spectrum of variables, covering political, demographic, and socioeconomic data points. Nevertheless, it was noted that certain variables had missing values or were not directly related to my examination. In the initial stages of data cleaning, attention was concentrated on crucial variables such as voter identification, voter registration status, presidential preferences, voter turnout, race, and income. Variables that closely corresponded to this paper, including those related to demographics and electoral participation, were prioritized. Whereas, variables with less direct relevance to the paper were omitted to efficiently analyze the dataset.

2.6 Voter Registration

As outlined in Table 1 (**?@tbl-registrationsummary**), there is a notable level of engagement among the electorate with a substantial proportion completing the registration process prior to the election. The calculated statistics, including the total number of participants, the count of registered voters, and the percentage of those registered, elucidate the readiness and eagerness of the population to participate in the electoral process. The median registration rate, not directly calculated in the code but inferable from the distribution of the registration percentages across different groups, likely serves as a robust indicator of this preparedness. The variations in registration rates, potentially inferred from the analysis by race, income, and other demographics (as structured in the code for race and income, but could be extended to other variables), may reveal discrepancies in registration practices and levels of electoral engagement among diverse segments of the population. This detailed breakdown by key demographics highlights the disparities and patterns of voter registration, offering insights into the dynamics of electoral participation across different states and demographic cohorts.

2.7 Presidential Preferences

The analysis of presidential preferences by race and income, as outlined in Table 2 (**?@tbl-presidentialdistribution**), sheds light on the relationship between socio-economic factors and electoral choices in the 2020 election. Within this visualization, presidential candidates' support is categorized by family income levels and further dissected across different racial groups, employing a color-coded scheme to differentiate the candidates. Observing the distribution, it's apparent that the majority of the electorate's preferences are polarized along both racial and income lines, with distinct patterns emerging within each subgroup. It shows a significant tilt towards one candidate (Biden), over others. However, there is a varying scales of support for presidential candidates overall across income and racial lines.

2.8 Voter Turnout

The analysis of the 2020 U.S Election voter turnout, as outlined in Table 3 (**?@tbl-voterturnout**), suggests intriguing patterns of voter turnout across various races and income levels. The bar charts, delineated by race, reveal the proportions of individuals within each income bracket who participated in the voting process. Notably, higher income levels appear to correlate with increased voter turnout across most racial groups, indicating a potential socioeconomic impact on electoral engagement. The racial facet of the data exhibits distinct turnout trends, potentially reflecting differing barriers to entry or motivational factors influencing political participation. For instance, some minority groups show a relatively consistent turnout across income levels, suggesting the possibility of a strong community-driven voter mobilization that transcends economic status. On the contrary, in other racial categories, there is a more pronounced gradient of turnout correlating with

income, highlighting socioeconomic status as a potentially more significant factor in these communities' political participation. These insights can drive future studies to explore the underlying causes of such patterns and develop targeted strategies to increase voter turnout across all segments of society.

3 Results

3.1 Summary Statistics of Demographic Factors

As outlined in Table 4 (**?@tbl-demographicsummary**), the largest race represented in the voters was white and the largest group of income was between $50,000 and $99,999. The variations in registration rates in regards to race and income may reveal discrepancies in registration practices and levels of electoral engagement among diverse segments of the population. This detailed breakdown by key demographics highlights the disparities and patterns of voter registration, offering insights into the dynamics of electoral participation across different states and demographic cohorts.

Our results are summarized in **?@tbl-modelresults**.

Our model suggests white people were less likely to vote for Biden, and that there is a considerable effect of income on this.

The distributions for the model parameters provide insights into the effects of race and income on the likelihood of voting and the likelihood of a certain candidate over the other.

3.2 Effects of Demographic Factors on Voting Likelihood

3.2.1 Race

The effect of Race on voting likelihood was positive. Specifically, it was indicated that white individuals were more likely to vote than other individuals.

3.2.2 Income

Income showed a differential impact on voting likelihood, with median and low levels of income being associated with an increased probability of voting. The model estimates this for each income level, relative to the baseline category.

3.3 Examination of Results

These results highlight the crucial role demographic factors play in influencing electoral participation. White individuals and those with average levels of income are significantly more likely to vote, aligning with existing literature on voting behavior. The observed race difference in voting likelihood also contributes to ongoing discussions on race dynamics within political engagement.

4 Discussion

4.1 Effect of Race and Income on Voting

This model reveals the nuanced impact of race and family income on voting behavior. For instance, the analysis shows significant differences in voting patterns across racial groups, with non-white individuals being more likely to support Biden. Additionally, family income has a positive association with the likelihood of voting for Biden, potentially reflecting the economic priorities or perceptions of the candidate's policies among different income brackets.

The insights derived from this model are vital for several reasons. For one, they highlight how demographic factors like race and income influence political preferences, underscoring the importance of inclusive policy-making that addresses the needs of diverse population segments. They also can inform political campaigns about which demographics are more likely to support their candidate, allowing for more targeted outreach and engagement strategies. Moreover, they contribute to the broader understanding of social and economic determinants of political behavior, offering a lens through which to view the intersections of race, income, and political alignment.

A study that aligns with these findings is Laurison's (2021) Voting Intersections: Race, Class, and Participation in Presidential Elections in the United States 2008–2016". This research examines the role of socioeconomic factors in shaping electoral outcomes over several election cycles, providing evidence that both race and income significantly influence voter preferences and turnout. Specifically, they find that income strongly predicts White voting. However, the class gap in voting is not statistically significant among Black voters. In contrast to common characterizations of Black people as politically disengaged, lower income Black citizens are more likely to vote than their White counterparts. The findings corroborate the importance of considering these variables in political science research and suggest that understanding the complex interplay between socioeconomic status and political behavior is crucial for comprehending the dynamics of electoral politics.

4.2 Cultural Differences

The model highlights cultural differences in value systems across racial and income groups, influencing political alignment. For instance, certain races might place a higher emphasis on community support and social welfare programs, leading to a preference for candidates who prioritize these issues. Cultural differences in voting behavior might also reflect historical contexts and socioeconomic conditions unique to each racial or income group. These contexts can shape perceptions of political institutions and candidates, influencing electoral decisions.

The importance of these findings are that understanding cultural differences in voting behavior is crucial for political campaigns aiming to engage with a diverse electorate effectively. Tailoring messages that resonate with the unique values and concerns of different cultural groups can enhance voter outreach and participation. Additionally, insights into the cultural underpinnings of voting preferences can inform policy development, ensuring that it addresses the needs and expectations of culturally diverse constituencies.

An important study in this context is Rhodes (2017), "Is America More Divided by Race or Class? Race, Income, and Attitudes among Whites, African Americans, and Latinos". Their findings suggest that race and class intersect in different ways for different groups in society. Increasing income erodes differences in attitudes between Latinos and whites, but has no effect on the large gap in attitudes between African Americans and whites. Interpolating these insights with the context of the model's outcomes, we gain a nuanced understanding of the interplay between culture, race, income, and voting behavior, emphasizing the critical role of cultural diversity in shaping political processes.

4.3 Demographic Variable Effects

I found that potentially due to varying policy priorities, historical voting trends, or the influence of race-specific issues, voter demographics can reveal how different racial groups may have distinct voting patterns. Analysis shows that higher income levels correlate with a preference for Biden as president, while lower income brackets also supports Biden.

The importance of these insights are allowing political campaigns to tailor their messaging and policy proposals to better meet the needs and preferences of different demographic groups as you further understand voter demographics. Additionally, insights into the demographic breakdown of voter preferences can inform policymakers and political leaders, helping them to develop more inclusive and representative policies. By recognizing and addressing the diverse needs of the electorate, political actors can enhance engagement and participation across all segments of society, strengthening democratic processes.

A study that contextualizes the importance of demographic analysis in understanding electoral behavior is Arceneaux (2005), titled "Do Campaigns Help Voters Learn? A Cross-National Analysis". This paper further supports the hypothesis that campaigns 'enlighten' voters as the election draws near. Moreover, the article shows that some voters learn more from campaigns

than others. Campaign effects are more pronounced among individuals with low political sophistication and those living in party list systems. Implications for future research are explored, suggesting a ripe research agenda using under-tapped cross-national data. By leveraging such demographic insights, stakeholders in the political process can better understand and respond to the needs of an increasingly diverse electorate, fostering a more dynamic and responsive political landscape.

4.4 Limitations and Future Steps

One potential weakness of the CES2020 dataset is the representativeness of the dataset. While CES aims to capture a broad cross-section of the American electorate, challenges remain in fully representing underrepresented minorities, individuals without internet access, and those less inclined to participate in surveys, potentially skewing the dataset's reflection of the national population. Additionally, the accuracy of responses may be affected by measurement errors, including the framing of questions and social desirability bias, where respondents may provide answers they perceive as more socially acceptable rather than their true beliefs.

The cross-sectional design of CES 2020 limits the ability to track changes in political attitudes and behaviors over time, making it difficult to analyze trends or predict future shifts in the political landscape. Missing data, a common issue in large datasets, presents another challenge, potentially leading to biased outcomes if not adequately addressed.

Looking forward, enhancing the representativeness of the CES by exploring more diverse sampling strategies and addressing the challenges of capturing hard-to-reach populations could improve the dataset's comprehensiveness. Incorporating longitudinal elements or linking CES data with longitudinal studies could provide richer insights into how voter behaviors evolve, offering a dynamic view of political trends. Delving deeper into underexplored themes, such as the influence of local issues, misinformation, or detailed policy preferences, could unearth new understandings of voter motivations and behaviors.

Employing mixed methods, such as integrating qualitative data through interviews or focus groups, could complement the quantitative CES data, providing a more nuanced picture of the electorate. Advanced statistical techniques to better handle missing data can enhance the robustness of analyses, and exploring the reasons behind non-response may offer additional insights. Lastly, linking CES data with other datasets, like economic indicators or social media trends, could enrich analyses and offer a more holistic view of the factors shaping American politics. Addressing these potential weaknesses and pursuing these future steps could significantly augment the value of the CES dataset for political scientists and policymakers alike, contributing to a deeper and more nuanced understanding of the American electoral landscape.

References

Arceneaux, Kevin. 2005. "Do Campaigns Help Voters Learn? A Cross-National Analysis." *British Journal of Political Science* 36: 159–73. https://api.semanticscholar.org/CorpusID:153526075.

Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

Bolker, Ben. 2022. "Broom.mixed: Tidying Methods for Mixed Models." https://CRAN.R-project.org/package=broom.mixed.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Kuriwaki, Shiro, Will Beasley, and Thomas J. Leeper. 2023. *Dataverse: R Client for Dataverse 4+ Repositories*.

Laurison, Daniel, Hana E. Brown, and Ankita Rastogi. 2021. "Voting Intersections: Race, Class, and Participation in Presidential Elections in the United States 2008–2016." *Sociological Perspectives* 65: 768–89. https://api.semanticscholar.org/CorpusID:250056003.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rhodes, Jesse H., Brian F. Schaffner, and Sean McElwee. 2017. "Is America More Divided by Race or Class? Race, Income, and Attitudes Among Whites, African Americans, and Latinos." *The Forum* 15: 71–91. https://api.semanticscholar.org/CorpusID:148618045.

Schaffner, Brian, Stephen Ansolabehere, and Sam Luks. 2021. "Cooperative Election Study Common Content, 2020." Harvard Dataverse. https://doi.org/10.7910/DVN/E9N6PH.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.