

# Critical Analysis Project

*Paper & Analysis Due: MAY 8<sup>TH</sup>, 9:00 A.M.*

## 1 THE ASSIGNMENT

---

In 2022, millions of Americans lined up outside of school gymnasiums, community buildings, and churches to vote in federal, state, and local elections. In the United States, voting is a time-honored and sacred form of **political participation**—a mechanism by which citizens can express their political opinions and affect political outcomes. The right to vote is not the only way Americans can work to influence politics: volunteering in campaigns, putting up lawn signs, and calling elected officials are all forms of political participation. **Donating to political campaigns** is another critical, although hotly contested, form of participation in politics. Many people think that money equals power, and that those individuals who make more money will donate more and, in turn, have a greater say in the political process. According to the [Pew Research Center](#) (2018), two-thirds of Americans think that those who donate more money have better representation in government. This tension begs the question: **is the commonly held conception that wealthy citizens give more in political donations reflected in real-world data?**

*H1<sub>0</sub>: There is no difference in the amount of political donations made by individuals with a higher household income and individuals with a lower household income*

*H1<sub>1</sub>: Individuals with a higher household income give more in political donations than individuals with a lower household income*

For your final project, you will be using real data from the [2022 Congressional Cooperative Elections Study](#) to investigate disparities in the political giving behavior of Americans. This assignment will be composed of three parts: (1) a quantitative analysis, and (2) a written memo, and (3) a supplementary appendix.

## 2 QUANTITATIVE ANALYSIS

---

Using your knowledge of linear regression and coding skills in R, your task is to conduct a robust quantitative analysis to determine whether certain kinds of Americans donate greater amounts to political campaigns. In this assignment, the main dependent variable is `ces$donation`, which takes on a range of values from \$0 to \$1,000,000. The main independent variable is `ces$faminc_new` which takes on a range of values from \$9,000 to \$600,000. Data and a codebook are provided, please use only these materials to complete the quantitative portion of your assignment:

### 2.1 HOW MUCH DID AMERICANS DONATE BY EDUCATION LEVEL?

Individuals who possess higher educational attainment tend to generate higher income ([see here](#)). Additionally, individuals who possess higher education attainment tend to be more involved in politics ([see here](#)). Why might this pose a problem for our analysis? Explain in your memo. To provide support, empirically and graphically demonstrate the relationship between education and income, as well as the relationship between education and political giving. Describe the trends you uncover. Include any plots produced in your memo. What should we do to rectify this issue? Explain.

### 2.2 PREPARING YOUR DATA FOR ANALYSIS

In the attached codebook, a number of variables have been included for this analysis. You must decide which of these variables warrant inclusion in your regression. **In addition to `ces$faminc_new` and**

`ces$educ`, include at least two other control variables from the codebook below in your analysis. Be sure to provide clear reasoning for your modeling decisions in the memo you produce.

It is also your job to consider how your chosen control variables should be cleaned or simplified prior to running your regression. Consider the following questions when deciding on your model specification:

- Which of these variables may be confounders with income and political donations?
- Which of these variables may induce multicollinearity?
- Can these variables be simplified (e.g., can the number of levels in a factor variable be reduced)? If so, what data-driven methods can I use to support my recoding of this variable?

You should discuss any and all data cleaning choices in the supplementary appendix of your paper. Plots, summary statistics, or frequency tables may be helpful here as support for your decision-making.

### 2.3 RUNNING THE REGRESSION & REPORTING THE RESULTS

Run two separate models. In the first, investigate a bivariate relationship between political donations (`ces$donation`) and household income (`ces$faminc_new`). In the second, include your primary IV (`faminc_new`) as well as any control variables you have identified. Use the results to produce a nicely formatted regression table, where there is one column for model 1 (bivariate) and a second column for model 2 (multivariate). Your table must include the model coefficients, the number of observations, and the r-squared statistic. You can include the standard errors if you wish, but this is not required. Please ensure you bold or place a star next to those coefficients that are statistically significant at an alpha level of 0.05. This table should be included in the supplementary appendix of your memo.

In the body of your memo, provide a substantive interpretation for the (1) household income variable (`faminc_new`). Compare this model coefficient to that from your bivariate analyses. What does this tell us? Based on our findings, do we reject or fail to reject our null hypotheses?

Discuss any other noteworthy findings based on the control variables you decided to include. Provide the substantive interpretation for these control variables, as well. In the supplementary appendix, provide the mathematical equation from your model using all available data from your regression output (alpha value, coefficients, etc.)

## 3 Memo

---

The memo you produce should include a short introduction. It should explicitly state the question of interest and hypotheses being investigated. In the body of the essay, discuss the data descriptively, including all plots and statistics produced. Report your findings as described above.

In the final section, please provide a thoughtful and **thorough** discussion of the policy implications of these results. What are these results' implications on the democratic process? What recommendations might you make to experts? Finally, in your conclusion, think back to your modeling approach and provide some thoughts. What improvements could have been made? What could you have done differently?

### 3.1 FINDING (ACADEMIC) SOURCES

If you would like to include help academic sources for your paper (this is not a requirement, but could be useful in your policy discussion) there are a number of good search engines to help. The three that I suggest that you use are:

1. JSTOR (<http://www.jstor.org.libproxy.lib.unc.edu/action/showAdvancedSearch>)
2. Notre Dame Library (<https://www.library.nd.edu/>)
3. Google Scholar (<https://scholar.google.com/>)

As with all essays, any and all outside information used should be included in a bibliography/references section. It does not matter which citation style you use, so long as you are consistent. Plagiarism will be penalized to the fullest extent possible through the UNC Honor System.

## 4 CHECKLIST --- REVIEW BEFORE SUBMISSION

---

### **R script**

- Includes all code used to generate plots, produce summary statistics
- Includes all code used to re-code variables of interest
- Includes all code used to produce linear regressions

### **Memo**

- Introduction:
  - Describes political problem
  - Explicitly states question of interest
  - States hypotheses
- Model Specification:
  - Describe rationale for educ variable inclusion (be sure to include any plots or empirical statistics produced in your quantitative analysis here)
  - Describe rationale for other variable inclusion (including plots may be appropriate here)
- Results:
  - Substantively interpret bivariate and multivariate regression output for cesfaminc\_new
  - Discuss differences between bivariate and multivariate analysis
  - Substantively interpret multivariate regression output for all other control variables
  - State whether we reject or fail to reject null hypothesis
  - Discuss policy implications for the results presented
- Conclusion:
  - Reflect on model specification and research design; provide thoughts on how both could be improved. Using the r2 statistic here could be helpful.

### **Supplementary Appendix**

- Provide rationale for data cleaning choices
- Provide regression output table
- Provide mathematical equation for specified regression model

## 5 CODEBOOK

<b>Variable</b>	<b>Details</b>
caseid	A unique number for each respondent
pid3	Respondent's answer to the question "Generally speaking, do you think of yourself as...?"  1 Democrat 2 Republican 3 Independent 4 Other 5 Not Sure
educ	Respondent's answer to the question, "What is the highest level of education you have completed?"  1 Did not graduate from high school 2 High school graduate 3 Some college, but no degree (yet) 4 2-year college degree 5 4-year college degree 6 Postgraduate degree (MA, MBA, MD, JD, PhD, etc.)
faminc_new	Respondent's self-reported income, ranges from 9000-600000  NAs indicate respondent preferred not to say
gender4	Respondent's self-reported gender, coded as follows:  1 Man 2 Woman 3 Non-binary 4 Prefer not to say

birthyr	Respondent's self-reported birth year
race	Respondent's answer to the question, "What racial or ethnic group best describes you?"  1 White 2 Black or African-American 3 Hispanic or Latino 4 Asian or Asian-American 5 Native American

	<p>6 Two or More Races  7 Other  8 Middle Eastern</p>
votereg	<p>Respondent's answer to the question, "Are you registered to vote?"</p> <p>1 Yes  2 No  3 Don't Know</p>
contact	<p>Respondent's answer to the question, "Did a candidate or political campaign organization contact you during the 2022 election?"</p> <p>1 Yes  2 No</p>
urbancity	<p>Respondent's answer to the question, "How would you describe the place where you live?"</p> <p>1 City  2 Suburb  3 Town  4 Rural area  5 Other</p>
hadjob	<p>Respondent's answer to the question, "At any time over the past five years, have you had a job?"</p> <p>1 Yes  2 No</p>
child18	<p>Respondent's answer to the question, "Are you the parent or guardian of any children under the age of 18?"</p> <p>1 Yes  2 No</p>