**School of Politics, Security, and International Relations**

**MA Comprehensive Examination, Spring 2024**

**Please read the following instructions before beginning the exam**

* This exam is in two parts. Part I will involve a critique of a political science article. Part II consists of three questions on data analysis and interpretation, involving your completing several aspects of data analysis using STATA.
* You should answer **all** the questions
* All of your answers should be completed in this Word document.
* This is an **open-book, open-notes/course materials, etc. exam**. However, all of the work must be your own. There should be no talking or consulting between students taking the exam.
* When you have completed the exam, please email this document to Kyrie Ottaviani at: [GraduatePoliSci@ucf.edu](mailto:GraduatePoliSci@ucf.edu). You should also email a copy to your own Knights email account. Ensure that you save the Exam file often so that you do not lose your work!
* You have until **5:00 pm** to complete and submit the exam.

**Part I**

Provide a critique of the political science journal article that you have been assigned.

The article critique should not simply summarize the article. Please use the following subheadings to help you organize your critique:

* Research problem/question/hypotheses
* State of prior research
* Data and Methods
* Findings and implications
* Future research agenda

A well-constructed and reasoned critique will consist of approximately 4-5 pages (regular font, double spaced).

Start your critique on the next page

**Part II**

**Important: Provide *relevant* STATA output (copy and paste) for each of the questions below.**

**Use the dataset “GSS” for questions 1 and 2. Use the dataset “NES” for question 3. These are the 2020 iterations of those datasets.**

**QUESTION 1 (10 POINTS): GSS DATA**

One of the perennial hypotheses in voting behavior is the idea that there is a gender gap, such that women as a whole tend to express more liberal policy views than do men as a whole. This is a complex hypothesis involving several different concepts. We will test this relationship here. Using views about whether the government should or should not mandate permits to purchase guns (**gunlaw**), determine if gender (**gender**) impacts respondents’ expressed views on that question. Compute and explain the difference in the proportion of respondents adopting either view. Does it appear that females express more liberal views than males on this question?

Then, control for the respondents’ ideology (**polviews**). Once one controls for the respondent’s ideology, do there appear to be differences among men and women such that women express more liberal views on gun control than men as a whole? Explain your answer.

**QUESTION 2 (10 POINTS): GSS DATA**

One of the key defining factors among the U.S. electorate is race. Is an individual’s race (**raceacs1**) related to their confidence level expressed towards the Congress (**conlegis**)? Use the variables ***conlegis*** and ***raceacs1*** to respond to this question to determine if there is a significant difference in the mean levels of confidence expressed towards Congress between those respondents who are white as opposed to those who are non-white. Explain your answer in full with reference to the statistical output generated.

**QUESTION 3 (80 POINTS): NES DATA**

A staple of political research is the concept of feeling thermometer in which respondents are asked how warmly or cooly they feel with respect to political candidates and/or issues. Use the feeling thermometer values (0-100) about Democratic presidential candidate Joe Biden post-election (**ft\_biden\_post**) as the dependent variable and complete the below analyses. Do those who are more liberal systematically (**libcon7**) give Biden warmer ratings on the feeling thermometer once several control variables are included in the specification, *ceteris parabus*?

For **libcon7**, ensure that you exclude any values from your analysis that do not include substantive responses to the question about the respondent’s ideology.

1. See the list of variables in the table below.

* 1. Locate the variable descriptions in the data file for the independent, dependent, and control variables.
  2. Write the variable descriptions into the empty cells in the table.

*Descriptions*

|  |  |  |
| --- | --- | --- |
|  | Variable Name | Description |
| Dependent | **ft\_biden\_post** | [Given] |
| Independent | **libcon7** | [Given] Scale ranging from 1 (extremely liberal) to 8 (extremely conservative) |
| Control | **hh\_income\_pre** |  |
| Control | **educ\_5cat** |  |
| Control | **race\_ethnicity** |  |
| Control | **age** |  |
| Control | **bible\_god\_man** |  |
| Control | **partyid7** |  |
| Control | **gender** |  |

1. Suppose you suspect that ***ft\_biden\_post***would differ by race. **Conduct a difference of means test**. Use ***race\_ethnicity***. *To do so, one must first recode the original values of* ***race\_ethnicity*** *into two values (white = 0; non-white = 1). Name the new variable that you have created as a result of the variable transformation* ***race\_2cat****. Once you completed the recode, compute the difference of means analysis for this question for the two groups*. Report here all the relevant information you use to conduct the means test (including number of observations, the sample means, difference between the means, *t*-score, and *p*-value). Interpret your results.
2. **Descriptive Statistics**
3. What are the mean, median, mode, standard deviation, and skewedness of the variables below? Fill in the table provided below.
4. Are the independent and dependent variables normally distributed? Explain.
5. For each of the variables, which central tendency measure is best to use when describing the distribution, the mean, the median, or the mode? Explain.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Mean | Median | Mode | Standard Deviation | Skewedness |
| **ft\_biden\_post** |  |  |  |  |  |
| **libcon7** |  |  |  |  |  |
| **educ\_5cat** |  |  |  |  |  |
| **hh\_income\_pre** |  |  |  |  |  |
| **age** |  |  |  |  |  |
| **bible\_god\_man** |  |  |  |  |  |

1. **Statistical Significance and Bivariate testing**
   1. Estimate Pearson Correlations of the dependent variable with the independent variable and each of the control variables ***below***. Fill in the table below.

*Correlation Coefficients*

|  |  |
| --- | --- |
|  | **ft\_biden\_post** |
| **libcon7** |  |
| **educ\_5cat** |  |
| **hh\_income\_pre** |  |
| **age** |  |

* 1. Based on these estimated correlations, which variable seems to have the **strongest one-to-one relationship** with the Biden feeling thermometer? Describe.
  2. What is the correlation between the Biden feeling thermometer and ideology? Describe.
  3. Based on the reported correlation between the Biden feeling thermometer and ideology, do we reject or fail to reject the null hypothesis describing the relationship between the dependent variable and the key independent variable? Explain your reasoning fully.

1. **Bivariate Regression**
   1. Specify and estimate a Linear Regression for the effect of respondent ideology on the Biden feeling thermometer.

* 1. List the coefficient estimates (and constant), corresponding standard errors, *p*-values, and *R2* estimate for the model.
  2. In this context, is respondent ideology a statistically significant predictor of the Biden thermometer rating? Explain.
  3. Interpret the effect of a one-unit increase in ***libcon7*** on the dependent variable*.*
  4. What is the percentage of variation in ***ft\_biden\_post*** that can be explained by respondent ideology?
  5. Based on this regression model, do we conclude that respondent ideology is systematically associated with the Biden thermometer levels? Explain your answer.

1. **Multiple regression**
   1. Now, specify and estimate a multivariate Ordinary Least Squares (OLS) regression model to test the effect of ***libcon7*** on ***ft\_biden\_post***, but in this model include **all** ofthe control variables listed in 3.A above.
   2. Report the regression equation with all of the estimated slope coefficients. Interpret and report here **each** of the estimated coefficients.
   3. Which coefficients, if any, are statistically significant? Which coefficients, if any, are not statistically significant? Explain.

* 1. Interpret the Adjusted R-squared (*Adj. R2*).
  2. In this model, what are the *key* predictors of respondents’ feelings toward Joe Biden as a presidential candidate? Explain.

* 1. Based on this regression model, can we confirm the hypothesis formulated above? Explain. Are there any confounding effects there that should be addressed?
  2. Does the multivariate specification here more systematically explain Biden’s evaluations than does the bivariate regression model reported in question E above? Explain your reasoning.
  3. Now, repeat the multivariate specification listed above in part 1), but this time use the **robust option** to correct for several specification and/or measurement errors. What does this revised analysis tell us about the relative impact of each of the regressor variables?