Data Exploration: The Political Psychology of Trump Indictment

Your name here

October 15, 2025

In this Data Exploration assignment, we will be working with the replication data from Ballard-Rosa et al (2021). The actionable part of each question is **bolded**.

Ballard-Rosa et al (2021) asks whether *authoritarianism* is a stable characteristic of a person, developed during early socialization and fixed throughout adulthood, or an attitude that can arise contemporaneously in reaction to economic threats, with an idea that economic threats can activate a frustration-aggression mechanism. The replication data combines their original survey of 1913 British people in 2017 with the economic characteristics of respondents' local communities. Note that the community characteristics are at the "Travel to Work Area (TTWA)" level, similar to commuting zones in the U.S.

Data Details:

- File Name: `ballard_rosa.dta`

- Source: Ballard-Rosa et al (2021); This file is a rearranged version of their replication data at Harvard Dataverse. Below are select variables in the data:

Variable Name	Variable Description
`caseid`	Unique respondent ID
`...`	See below for variables measuring authoritarianism.
`age`	Age
`female`	Binary variable for gender
`universityEduc`	Binary variable for university education
`higherCert`	Binary variable for higher certification
`isMarried`	Binary variable for whether married
`pcnuk2015`	% non-UK born in R's Travel to Work Area
`chg_immig`	Percent point increase in % non-UK born in R's TTWA
`ipw_1991_2007_weighted`	A measure of local labor market exposure to Chinese import competition, exposure in R's TTWA
`...`	Other variables in the survey. See survey_codebook.pdf

```
#Load the data
df <- read_csv("ballard_rosa.csv")
```

```
## Rows: 1913 Columns: 464
## -- Column specification -------------------------------------------------
## Delimiter: ","
## chr  (23): q14, q10_a_t, q41_other, q43_other, comment_stan, postcode, reven...
## dbl (435): caseid, forceNecessaryGroupsThreat, forceNecessaryIndivThreat, po...
## lgl   (6): temp_emp2002, temp_emp2003, emprate2002, emprate2003, unemp_rate2...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Authoritarianism is measured in the 18 columns to the right of the `caseid` column. The authors define *authoritarian values* as "individual preference for order and conformity and belief that these value outcomes should be achieved by force if necessary". The authors view authoritarianism as a concept comprised of three separate sub-dimensions: authoritarian aggression, submission, and conventionalism. Each dimension is measured with 6 questions below. Respondents answered how much they agree with each statement at a five-point scale from 'strongly agree' to 'strongly disagree'.

Authoritarian Aggression:

- Strong force is necessary against threatening groups.
- It is necessary to use force against people who are a threat to authority.
- Police should avoid using violence against suspects.*
- People should avoid using violence against others even when ordered to do so by the proper authorities.*
- Using force against people is wrong even if done so by those in authority.*
- Strong punishments are necessary in order to send a message.

Authoritarian Submission:

- We should believe what our leaders tell us.
- Our leaders know what is best for us.
- People should be critical of statements made by those in positions of authority.*
- People in positions of authority generally tell the truth.
- People should be skeptical of all statements made by those in positions of authority.*
- Questioning the motives of those in power is healthy for society.*

Authoritarian Conventionalism:

- People emphasize tradition too much.*
- Traditions are the foundation of a healthy society and should be respected.
- It would be better for society if more people followed social norms.
- Traditions interfere with progress.*
- People should challenge social traditions in order to advance society.*
- People should respect social norms.

Note that * indicates that it is reverse coded in the data.

Then the authors use the average of the 6 questions to construct each `avgAggression`, `avgSubmission`, and `avgConventionalism`. Then, the average of these three are used to construct `asc` (ASC, Aggression-Submission-Conventionalism), which they use as the main measure of authoritarianism.

Question 1

What is the distribution of the authoritarian values among respondents? Look at the mean, median, and standard deviation of ASC. Plot the full distribution using geom_density or geom_histogram.

Question 2

How about the distribution of each of three dimensions? On average, are people likely to agree with the statements above? Are people more likely to agree on some of these questions than others? How do these dimension correlate with each other? Feel free to explore further within each dimension.

Question 3

Do you think these 18 questions measure well the proposed concept of authoritarianism? Does each dimension capture something distinct about authoritarian values? What other questions or experiments could we use to better measure authoritarian attitudes?

Question 4 (Data Science Question)

The authors argue that some localities are more negatively impacted by Chinese import and people in those localities would develop more authoritarian values due to economic frustration. `ipw\_1991\_2007\_weighted` measures how much the local employment of a place depended on industries that were negatively affected by Chinese import. For anyone curious, a detailed explanation on how this variable is constructed can be found in Ballard-Rosa (2021) on pages 2333 and 2334.

PART A

Make a scatterplot to look at the relationship between `ipw_1991_2007_weighted` and the ASC measure of authoritarian attitudes. You can use `geom_point()` for this purpose. Do they seem to correlate with each other? Draw a line that summarizes these points. You can use `geom_smooth()` for this purpose.

PART B

Use single variate linear regression to test the relationship between `ipw_1991_2007_weighted` and the ASC measure of authoritarian attitudes. Interpret the results. Does the local exposure to Chinese trade shock increase authoritarianism?

Hint: The `lm()` function is used to calculate linear regressions in R. Here is a guide to linear regression in R that may be helpful.

Question 5 (Data Science Question)

Is the effect of the trade shock on authoritarianism robust to the inclusion of control variables? **Fit a multivariate linear regression controlling for individual-level demographic variables: age, gender, education level, and marriage status. What if you further include controls for other local characteristics such as the percent non-UK born population in respondents' TTWA or percent point increase in the non-UK born population in respondents' TTWA? Present the results in a regression table and comment on your findings. Make sure to comment on the main explanatory variable but also include the substantive interpretations of coefficients for the control variables.** The Ballard-Rosa data have a lot of interesting variables. If you'd like, feel free to explore further with more control variables.

Hint: The `stargazer` package makes good regression tables; set `type = 'text'`.