Final Project Checkpoint 2: Codebook and Data Description
DUE October 12, 11:59 PM

Your second checkpoint will explain how you will collect and compile your data and detail your variables with a codebook. "Collecting" your data may involve finding a source of data (a survey, replication data from a previous study, existing datasets on the phenomena you're interested in, etc.) that contains the variables you need to test your theory. Although you can also collect some data on your own, I recommend against this, simply because it takes so much time. At this point, don't worry about doing any data analysis, but you should 1) know that the data are available, 2) know that they can be merged together, and 3) if possible, you should be in the process of actually compiling the data into a dataset that you can analyze. Although we will talk more about controls later, your dataset should include data on your dependent variable, your primary independent variable, and other important variables that you think also affect your dependent variable (and especially if you think they may be correlated with both your independent and dependent variable).

All you need to submit for this checkpoint is your <u>codebook</u>. This codebook should describe *each* of your variables. More specifically, for each variable, you should identify its name, its level of measurement (interval, nominal, ordinal), the source for the variable, how the variable is coded, and any other notes you believe might be important. Here is an example of a codebook entry for an interval-level variable and a categorical variable (note, you don't have to format your entries exactly this way, but codebooks often do take a tabular form).

Judicial Ideology (variable name) My measure of judicial ideology is categorical and based on the party of the appointing president. This requires the assumption that presidents appoint justices who share their ideology. Categorical Value = 1 if appointed by a Republican president; 0 if appointed by a Dem. president; . = missing data Source: Epstein, Lee, Jeffrey A. Segal, Harold J. Spaeth, and Thomas G. Walker. 2012. *The Supreme Court Compendium: Data Decisions, and Developments.* Washington, DC: Congressional Quarterly.
Judicial Ideology (variable name) My measure of judicial ideology is Martin and Quinn's Bayesian measure of judicial ideology based upon the votes of U.S. Supreme Court justices. Lower values indicate more liberal justices. Interval Values = range from -8 to 4.5, with higher values indicating more conservative justices, and lower values indicating more liberal justices; . = missing data. Source: Andrew D. Martin and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10: 134-153.

Please submit your codebook via CANVAS by 11:59PM EST on Thursday, October 12.
Please print your hypothesis, in bold, at the top of your submission!