Assignment II: Datasets in Scholarly Writing

Assignment Purpose:

One of the challenges facing authors who use large datasets to study complex topics is ensuring that there is some congruence between the theoretical concept of interest to them and the variables that are available. For example, "democracy" is hard enough to describe purely qualitatively. Does democracy consist of free and fair elections? Does it require some measure of civil and political liberties?

Even if we can decide on a definition, translating that definition into an objective measure can be tricky. How do we capture civil liberties, for example? The number of newspapers in circulation (e.g., a free press)? Whether a particular state has a bill of rights? How about political freedom—the number of political parties? The number of political dissidents in prison? This issue of making sure a variable measures what the author intends it to measure is one of *validity*. If a measure is valid, it captures the essential facets of a concept.

In this assignment, you will use your knowledge of two datasets to explore the challenges around validity when we work with ambiguous or complex concepts in large-N analyses.

Instructions:

1. Go to the course website and open the "Data Articles" folder. There, you will find a list of articles that have been written using the datasets from the last assignment. Pick one article that corresponds to one of the datasets you worked with last time.

2. Find the article (this gives you good practice in looking up scholarly articles using various article databases) and read it, noting carefully the way in which the author defines and uses the theoretical concept in his or her argument.

3. Drawing on your knowledge of this variable from your last paper, think about how well the variable as defined and measured in the dataset corresponds to the theoretical discussion in the article. Is the variable a valid measure given the author's argument? Why or why not?

4. Finally, assess whether the same variable in the other dataset you examined might have any advantages over the chosen variable in terms of validity. You might also consider whether the alternative variable has any other advantages (e.g., more appropriate geographic/chronological coverage, etc.)

5. Write up the results of your investigation in a 3-4 (double-spaced) paper. The beginning of the paper should include a very brief summary of the author's argument and how s/he defines the concept before moving to an assessment of validity and a comparison of the two datasets.