Replication Data Codebook for Inequality and the Dynamics of Public Opinion: The Self-Reinforcing Link Between Economic Inequality and Mass Preferences

Nathan J. Kelly
Department of Political Science
University of Tennessee
email: Nathan.J.Kelly@gmail.com

and

Peter K. Enns
Department of Government
Cornell University

October 1, 2011

Introduction

This document contains details of the collection and coding of variables included in the replication dataset for our article: "Inequality and the Dynamics of Public Opinion: The Self-Reinforcing Link Between Economic Inequality and Mass Preferences," published in the *American Journal of Political Science*. The dataset is available in Stata format at `http://dvn.iq.harvard.edu/dvn/dv/nkellydata`. The data are time series data, with year being the time operator.

Variables and Descriptions

mood

The mood series was taken from Stimson (1999), with updated data taken from his website (`http://www.unc.edu/~jstimson/resource.html`). We utilize the 2006 annual version of his mood series. Older versions of the measure are not archived on his website, but a detailed description of the creation of the series is there, along with the most recent update to the series.

unemployment

This is the percent of the civilian labor force unemployed using the official Bureau of Labor Statistics annual estimate. Unemployment data can be found at `http://www.bls.gov`.

inflation

This is the official rate of inflation calculated based on the CPI-U series. We calculate the inflation rate as the percent change in the year over year average CPI-U index with the index equal to 100 in 1982-1984. Inflation data can be found at `http://www.bls.gov`.

policy

The policy series is taken from Erikson, MacKuen & Stimson (2002), updated through 2006 by the authors. This measure examines important policy change by focusing on the crucial public laws identified by David Mayhew (2005). From this list, laws related to domestic policy with national impact are coded as to whether they were viewed as expanding (liberal) or contracting (conservative) government at the time they were passed. Laws that were ambiguous in their expansion versus contraction of government were coded as neutral and do not contribute to the policy change captured in this measure. Liberal legislation is counted +1, conservative legislation -1, and exceptionally important laws (as defined by Mayhew) are counted +2 or -2. Each year since 1947, a score is produced by summing liberal minus conservative legislation—this is annual policy change. The

current level of policy is produced by accumulating annual policy change over time. A net liberal shift in policy produces a positive change in this policy measure. Since the late 1940s, the most important policy changes have usually led to government expansion. In essence, then, the debate in the United States has not been literally about the contraction versus expansion of government, but about whether government should expand in response to the problems that develop in an increasingly complex society. Given this, it is better to examine the accumulation of policy relative to the long-term trend of government expansion. Our variable policy is a detrended version of cumulative policy change.

gini

This is the gini coefficient of family income inequality. Values range from 0 to 1, with higher scores indicating more inequality. The data are from the U.S. Census Current Income Reports Historical Tables available at `http://www.census.gov`.

welfare

This series is taken from the General Social Survey from 1973-2006 (years in which the question was not asked are interpolated). The question asked: We are faced with many problems in this country, none of which can be solved easily or inexpensively. I'm going to name some of these problems, and for each one I'd like you to tell me whether you think we're spending too much money on it, too little money, or about the right amount. Welfare, are we spending too much, too little, or about the right amount on welfare? We calculate the percentage replying "too little" of those responding either "too little" or "too much."

dempres

This variable is coded 1 if the president is a Democrat and 0 if a Republican.

mood_lowinc and mood_highinc

In order to generate a measure of the public's policy mood by income group, we used all questions used in Stimson's (1999) measure of policy mood for which individual data are available. Relying on data from the Roper Center, the General Social Survey, and the American National Election Study we were able to obtain 76 distinct questions, which were each asked multiple times between 1956 and 2006. In total, our measure of mood combines information from 1,032 survey questions.

Prior to generating mood by income level, we first generated an aggregate mood series using these questions. Following Stimson, we calculated the percent offering a liberal response divided by the percent offering a liberal response plus the percent offering a conservative response and combined the series into a single overtime measure with

Stimson's dyad ratios Wcalc algorithm.[1] Our aggregate series correlates with Stimson's mood series at an impressive r=0.85.[2] Thus, we are confident that the questions we have used produce a valid proxy of Stimson's policy mood.

In order to generate a separate mood series for each income group, for each survey question we calculated the percent of liberal respondents divided by the percent of liberal respondents plus the percent of conservative respondents for the lowest, middle, and highest income groups. These income groups correspond with the lowest 20 percent, middle 60 percent, and highest 20 percent income levels reported in the survey. Of course, income group categories offered in surveys do not always correspond exactly with income quintiles. Thus, in most cases we actually selected income categories that corresponded with lower incomes than the lowest 20 percent and higher incomes than the highest 80 percent of respondents. This strategy was used to ensure that whenever possible, our income groups are based on more extreme income percentiles than we suggest. In the few surveys where the lowest and highest income groups needed to extend beyond the lowest or upper quintile, coding was done to ensure that the number of respondents in the high or low income category never exceeded 30 percent.

The variables in our dataset refer to the lowest 20 percent (mood_lowinc) and highest 20 percent (mood_highinc). As with general mood, higher values indicate liberal mood, and scores can theoretically range from 0 to 100.

The Policy Mood by income group series were originally developed for two projects. We suggest citing both works when using these data.

Kelly, Nathan J. and Peter K. Enns. 2010. "Inequality and the Dynamics of Public Opinion: The Self-Reinforcing Link Between Economic Inequality and Mass Preferences." *American Journal of Political Science.* 54(5): 855-870.

Enns, Peter K. and Christopher Wlezien. 2011. "Group Opinion and the Study of Representation." in *Who Gets Represented?* ed. Peter K. Enns and Christopher Wlezien. New York: Russell Sage Foundation.

[1]The algorithm scales each question series to a common metric and then uses a factor analytic approach to produce a measure based on the common overtime variation of the component series. Full details on Wcalc can be found in Stimson (1999) and `http://www.unc.edu/~jstimson/resource.html`.

[2]The correlation is based on the years 1956 to 2006 using Stimson's Release 2 (5-29-2009) mood5208.prn, available at `http://www.unc.edu/~jstimson/time.html`.

References

Erikson, Robert S., Michael B. MacKuen & James A. Stimson. 2002. *The Macro Polity.* New York: Cambridge University Press.

Mayhew, David R. 2005. *Divided We Govern: Party Control, Lawmaking and Investigations, 1946-2002.* 2nd ed. New Haven: Yale University Press.

Stimson, James A. 1999. *Public Opinion in America: Moods, Cycles, and Swings.* 2nd ed. Boulder, CO: Westview Press.