# Replication Instructions

There are two different sets of replication materials, one to replicate our re-analysis of the LAPOP 2007 survey originally analyzed by Bohn, and one to replicate our analysis of other surveys taken in 2002 and 2006.

1) Re-analysis of the 2006 Lapop survey:
    a. We have included one R syntax file called "_replic.R" that reads in the Lapop survey data (provided by lapop in .dta format), recodes some variables, performs several different analyses, and produces tables reported in the paper and in the webappendix.
    b. Note that we have **not** provided the actual data file because this is owned by LAPOP, and distributed under restricted use policies.

2) Analysis of other 2002 and 2006 surveys.
    a. We have included the data file for this analysis in two formats (.RData and .R). This is a data file with **selected** variables from six surveys taken in 2002 and 2006. These include data on the municipality of respondents, which was not included in the original surveys, obtained from CESOP/UNICAMP.
    b. We have also included another .RData file with predicted vote for Lula in 2002 and 2006, which is included in the graphs explained in item "c", below. These were computed from the LAPOP 2007 survey. Code to produce this estimates is included in "_replic.R", but we supply this for convenience.
    c. We have included one R syntax file called "_replic.20022006.R" that reads in the data file mentioned in "a." It uses these surveys to predict vote for lula in each election, loads similar results from LAPOP2006 (mentioned in "b"), and plots them into figure 1. The file also analyzes the relationship between the human development index of municipalities and vote for Lula in each year, based on four of these six surveys and produces Figure 2 and the tomography plots in the webappendix.