Replication Readme For
"Are Samples Drawn from Mechanical Turk Valid for Research on Political Ideology?"

In order to replicate the results found in the paper, "Are Samples Drawn from Mechanical Turk Valid for Research on Political Ideology?" you need: the 2012 ANES Time Series Study data (anes_timeseries_2012_stata12.dta), the Qualtrics survey results (MTurk7.xlsx), the combined dataset (stackedmturkanes2.dta), the Cohen's *d* excel file (cohend.xlsx), the FDR excel file (qFDR_tables_q.05.xlsx), and the two R scripts (MturkRepublics.R and stacking_MTurk_and_ANES.R).

The Cohen's *d* and FDR excel files are used to calculate Cohen's *d* and FDR statistics found in our paper. These files use results calculated in our R script to produce these statistics that are later transferred into Latex.

The first R script (MturkRepublics.R) opens, recodes, and produces most of our results from both our initial datasets (anes_timeseries_2012_stata12.dta and MTurk7.xlsx). The second R script (stacking_MTurk_and_ANES.R) combines these two data sets into a new dataset (stackedmturkanes2.dta), recodes, and runs the remaining results.

To replicate the results open the first R script (MturkRepublics.R). After you open the script, clear the console, and load the packages load in the ANES data (anes_timeseries_2012_stata12.dta) on line 30. Lines 32 to 632 look at and recode the ANES data. Next, load in the Qualtrics survey that contains the results from the MTurk survey (MTurk7.xlsx) on line 638 (this dataset be slow to load into R). Then from lines 640 to 1175 look at and recode the MTurk dataset. Lines 1177 to 2137 produce our point estimate plots (Figures 1-3 and Figures A1-A3; you need to run these in the order they are presented in the script). The correlations found in these plots are calculated in the next R script (stacking_MTurk_and_ANES.R), but you will get to that later. These correlations can be calculated in the R script, but the code is run in the next R script. Lines 2141 to 232 run our models found in Tables A4 to A7, and they are used to plot Figures 4 and 4A. These models are rerun in the next R script (stacking_MTurk_and_ANES.R) in a format that is much easier to present, so for now just run these lines to reproduce Figures 4 and 4A. Next, run lines 2334 to 2646 to reproduce Figure 4 (you will need to run these plots in the order they are presented in). Then run lines 2653 to 2753 to create Figure 4A. The remaining code is this script looks at feeling thermometers in our dataset that we do not present in our paper, so you can now go on to the next R script (stacking_MTurk_and_ANES.R).

Now open the next R script (stacking_MTurk_and_ANES.R). Lines 7 to 72 can be run to subset, merge, and save the new stacked dataset. This will result in you producing the third dataset (stackedmturkanes2.dta), but you can skip to line 76 to load in this dataset if you already have it downloaded. Prior to running this data it is recommended that you clear the console, and reload the packages found in the previous R script (MturkRepublics.R). Now that you have just the stacked dataset (stackedmturkanes2.dta) loaded, run lines 78 to 94 for the remaining recoding. Run lines 96 to 217 to reproduce our models found in Tables A4 to A7. Next run lines 219 to 388 to reproduce Tables A8 to A11 and the p-values for the FDR excel file (qFDR_tables_q.05.xlsx). Lines 390 to 422 will give you the summary statistics included in our demographics table (Table A1). Lines 424 to 793 will give you correlations and t-tests for the Big Five, values, economic ideology, and social ideology. Run lines 813 to 851 to replicate Table A3. Finally, run lines 854 to 943 to produce statistics that were put into the Cohen's *d* excel file (cohend.xlsx) that was used to create Table A2.