Read Me: GoBifo Replication

Before running the replication:

- Download the "GBF_rep" file folder
- In "gbf_replication.do," change the directory to match the location where you have saved the GBF_rep folder
- Install user written command "estout"

What to expect:

- As the file runs, it generates both a log file and several outsheet and estout files in the GBF_rep folder. The latter two contain replications of estimates in the paper and appendix tables for ease of reference. To view the .out files, insheet them into Stata. To view the estout files, open them in Excel, choose option "fixed width" delimitation, adjust break lines to accommodate up to two stars after the coefficients, and then change all column data formats to "text."
- The "gbf_replication_log _master.pdf" contains the log of the entire replication.
- Variables in the datasets are labeled to correspond to the survey question (often truncated). The Excel file "Dataset key" contains the full labels without truncation, the variable name, hypothesis number, type of outcome (full sample or conditional), and row number to match up with Appendix J.
- The FWER adjustments to p-values are based on unseeded simulations, so vary by replication and will not generate exactly the same estimates as those in the tables.
- To replicate the FDR q-value estimates of Appendix J, Column 9, manually run Michael Anderson's "fdr_sharpened_qvalues.do" file. Check his website for updates (http://are.berkeley.edu/~mlanderson).
- The SUR estimations in Sections #6 and #10 take a long time to run. To expedite the processing, comment out these two sections.

Other things to note:

- All the de-identified "raw" data that corresponds exactly to the surveys is available for download on the Dataverse website (http://hdl.handle.net/1902.1/21708).

Discrepancies found between the replication file and main text tables:

- Tab 1, row 5, col 3: N = 235 in paper, is 236 in replication
- Tab 3, row 6, col 6: standard error in paper is 0.037, in replication is 0.038
- Tab 3, H10: coefficient 0.092** (s.e. 0.043) should only have only 1 * (significant at 97%)
- Tab 4, row 5, s.e. = 0.06 in paper, = 0.05 in replication; same in row 8, s.e. = 0.04 vs 0.03 in replication
- Tab 4, row 7, t.e. = 0.16** in paper, = 0.15** in replication
- Tab 4 row 10, col 4: N = 236 in paper, 233 in replication
- Tab 6, row 6, col 1: mean = 0.85 in paper, = 0.86 in replication
- Tab 6, row 9, col 1: mean =0.30 in paper, = 0.29 in replication
- Appendix Tab J: formula used to estimate p-values in the replication file does not exactly match the output in Column 7 (average difference ~ 0.002), however all TEs and SEs are identical

Questions, comments? Please email Kate Casey: kecasey@stanford.edu