The Power of Partisanship in Brazil:

Evidence from Survey Experiments

(Replication Documentation)

David Samuels
University of Minnesota
dsamuels@umn.edu

Cesar Zucco
Fundação Getúlio Vargas
cesar.zucco@gmail.com

The following replication material is available at http://hdl.handle.net/1902.1/21485.

- `_replication_documentation.pdf` This file
- `.pdf` Supplemental information with extended results and additional analysis
- `_experimentvalid.RData` File containing the internet survey-experiment data
- `_experimentvalid.dta` Same file as above, provided in state format for convenience
- `_experimentBEPSdata.RData` File containing the subset of data from BEPS study
- `_replic.R` Main replication file
- `_replicfunctions.R` Main replication file

1 The Data Files

The main data file (`_experimentvalid.RData`) was produced by eliminating invalid answers, recoding, and renaming variables present in a raw data file that was produced by directly by the online survey. This raw data file (`77517_data_20111223mac.txt`) is available from the authors upon request.

The data file `_experimentvalid.RData` contains a R data.frame named `d`, with 3212 observations of 20 variables, as follows:

- T : Factor w/ 4 levels, indicating treatment condition
- TT : Factor w/ 2 levels, indicating pooled treatment condition
- QPresal : Factor w/ 3 levels: Answer to Pre-sal question, coded as:
  77 Don't know
  45 Agreement with the PSDB
  13 Agreement with the PT

- QEmpresas : Factor w/ 3 levels: Answer to Gvt. Financing of Private Companies question, coded as QPresal
- QInvestimento: Factor w/ 3 levels: Answer to Foreign Investment question, coded as QPresal
- QVenezuela : Factor w/ 3 levels: Answer to Venzuela question, coded as QPresal.
- QSalario : Factor w/ 3 levels: Answer to Minimun Wage question, coded as QPresal
- party : Character, party name
- P : Factor w/ 6 levels: recoding of party to express preferences as follows:
  0 Non-partisan
  13 PT
  15 PMDB
  45 PSDB
  45 PV
  66 Other
- PP : Factor w/ 4 levels: recoding of party including only 0, 13, 45, and 66
- Idade : numeric: Age
- IdadeBracket : Factor w/ 4 levels: Age brackets
- Sexo : Factor w/ 2 levels: Gender
- R : Ord.factor w/ 5 levels: Income brackets
- iA : num: Political Participation index
- CEPRegion : Factor w/ 5 levels: Macro-region
- President : Factor w/ 6 levels: Vote for president in last election
- Tptypre : logical: Whether party ID question was shown before or after Q questions
- Diff : Factor w/ 3 levels: Does respondent see any difference between parties?

The data file `_experimentBEPSdata.RData` contains a R data.frame named `dbeps`, with 1221 observations of 11 variables, as follows:

- TT : Factor w/ 2 levels, indicating treatment condition or control
- T : Factor w/ 2 levels Alternative labeling of conditions, for compatibility with FB survey
- P : Factor w/ 5 levels: recoding of party to express preferences as follows:
  0 Non-partisan
  13 PT
  45 PSDB
  45 PV
  66 Other
- PP : Factor w/ 4 levels: recoding of party including only 0, 13, 45, and 66
- Idade : numeric: Age
- Sexo : Factor w/ 2 levels: Gender
- R : Ord.factor w/ 5 levels: Income brackets
- iK : int: Political knowledge index..
- region : Factor w/ 5 levels: Macro-region
- vote20101: Factor w/ 4 levels Vote for president in last election
- Qempresas: Factor w/ 3 levels Answer to Gvt. Financing of Private Companies question, coded as:
  77 Don't know
  45 Agreement with the PSDB
  13 Agreement with the PT

2 Replication Instructions

The R replication file `replic.R` reads in the data source(s), performs the analysis and creates graphs and tables used in the paper by employing the different functions that are declared in `replicfunctions.R`

For basic replication of the results and figures in the paper, simply save the `.RData` and `.R` files in your computer. Open the `replic.R` file and specify the path to the folder in which the data, and the other R file with the functions are. If desired, declare the name of the folder where graphs should be saved to (this is clearly marked at the start of the file).

2.1 Overview

We first transformed the five outcome variables with answers to the five experimental items into two different dummy variables: agreement with the PT or with the PSDB. Because the survey allowed for a "Don't know" option, analysis of agreement with either party is not exactly symmetrical (but is close). On each of the two variations of the dependent variable we fit two different regression models: a linear probability model where the only regressors are party dummies and treatment condition dummies and a logistic regression including control variables. Finally, we estimated each set of models using the four treatment conditions separately (control + PT cue + PSDBcue + DBLcue) and by pooling the treatment conditions together (control + some cue).

This means that for each of the five questions in the online survey there are eight regression models being estimated (i.e. 2 for agreement with the PT with 4 conditions, 2 for agreement with the PSDB with 4 conditions, 2 for agreement with the PT with 2 conditions, 2 for agreement with the PSDB with 2 conditions). And, some of these models are also fit to the the (simpler) single item that was fielded in BEPS.

In order to deal with the repetitive nature of the analysis, we created functions that essentially repeat the same steps in the analysis for each of the five questions: The estimating functions perform the basic estimation (pooled or multiple cues); the effects functions use the output of the estimating functions to computed predicted values and treatment effects; and the plotting functions display the data graphically. Beyond invoking these functions, the code simply conducts some balance tests, and manually merges BEPS results with FB survey results.

The functions are not nearly as generic and portable as functions in R contributed packages. However, they allow for easier extended analysis of the data beyond what is shown in the paper. In

fact, `replic.R` includes many more graphs than presented in the paper. The figures in the paper are produced and saved by `replic.R`, with the following names:

- FIGURE 3A: `fig-barplot13-combined.pdf`
- FIGURE 3B: `fig-barplot45-combined.pdf`
- FIGURE 4A: `fig-barplot13inout.pdf`
- FIGURE 4B: `fig-barplot45inout.pdf`
- FIGURE 5A: `fig-barplot0-combined-with13.pdf`
- FIGURE 5B: `fig-barplot0-combined-with13.pdf`

2.2 Estimating Functions:

`exp.eachtreat()` and `exp.sometreat()`

Description Estimates the models with there either three treatment conditions (`exp.eachtreat`) or with a single pooled treatment condition (`exp.sometreat`)

Usage: `exp.eachtreat.fb(qq,partyvar="PP",bt=T,the.data=d)`

Arguments Function takes the following arguments

qq Name of the variable that contains answer to question
partyvar Name of the variable that defines partisanship
bt Whether to bootstrap the the linear probability model
the.data Data frame that contains the variables

Details Function estimates agreement with both PT and PSDB's position in the question qq as a function of partisanship (interacted with) and each of the four treatment conditions. It estimates a linear probability model without controls, and a logit model with controls. It also saves bootstrapped estimates of the parameters. This takes some time, but it makes it much easier to compute standard errors later (standard errors of effects are the product of interactions, and therefore not easily computed from the vcov matrix, especially in non linear models).

Value A list containing the following objects

reg13ctrl Logit model for agreement with PT, with controls
reg13 Linear probability model for agreement with PT
reg45ctrl Logit model for agreement with PSDB, with controls
reg45 Linear probability model for agreement with PSDB
reg13boot $\text{Matrix}_{[100]}$ of bootstrapped estimates for reg13
reg45boot $\text{Matrix}_{[100]}$ of bootstrapped estimates for reg45
reg13ctrlboot $\text{Matrix}_{[100]}$ of bootstrapped estimates for reg13ctrl
reg45ctrlboot $\text{Matrix}_{[100]}$ of bootstrapped estimates for reg45ctrl

2.3 Effects Functions:

`effects.cue()` and `effects.pooled()`

Description Computes baseline predicted values for control group and treatment effects and SE from lm, glm, and boostraped objects, which can then be passed on to the plotting functions.

Usage: `effects.cue(xx,boot.se=NULL)`

Arguments Takes the following arguments:

xx A list of regression results (lm or glm)

boot.se an matrix of bootstrapped estimates to be used for SE computation in the lm models

Details Function computes predicted values for each category of partisanship in the regression model (typically PSDBistas, PTistas, non-partians and others) in each treatment condition in the regression (either some treatment and control, or PTcue, PSDBcue, DBLcue and control). Standard errors of the linear probability model are either taken from the regressions, by combining the SE's of the interaction terms, or computed from the bootstrapped estimates matrix, if one is provided. Strandard errors for the logit regressions with control are bootrapped by function maBin (erer), and computed by averaging marginal effects over all observations (and not computing the marginal at the average value of the other variables). Original version `effects.cue.zelig` relied on Zelig, but Zelig 4.0 changed syntax considerably.

Value A list containing the following objects

pes "Effects" of treatment in each subpopulation of interest, relative to the baseline (below)

ses SE's of pes, either bootstrapped or "regular", depending on boot.se and on model choice

baseline Baseline predicted level of agreement (with selected party) for non-partisans in the control group. I.e. predicted values with all treatment and party ariables set at 0.

2.4 Plotting Functions

`plot.effects()` and `plot.effects.pooled()`

Description Produces line plots with point estimates and confidence intervals of the effects of different treatments, for all questions

Usage: `plot.effects(xx,leg=T,ommit.others=T,ommit.non=T,port=F)`

Arguments Takes the following arguments:

xx A list with predicted values, standard errors, and baseline estimates, produced as outcome from `effects.cue()`
leg If T, include legend in the plot
ommit.others If T, exclude "other partisans" from figure
ommit.non If T, exclude "non partisans" from figure
port If T, make labels and legend in Portuguese

Details Function takes effects.cue output and produces graphs such as the ones presented in the paper

Value Graphical output only

`plot.cue()` and `plot.cueH()`

Description Barplot comparing predicted levels of agreement in a baseline category against one or two different cues, either with vertical or horizontal bars (H).

Usage: `plot.cueH(x,pty,cue="dbl",x2=NULL)`

Arguments Takes the following arguments:

x A list of effects produced by `estimate.pooled()` or `estimate.cue`
pty Numeric, indicating the partisans of choice
cue Character, indicating which cue to plot ("dbl","pt", "psdb")
x2 Whether to plot a second effect, in which case the cue will be..

Details The function takes effects estimated for agreement with one of the two parties (this is defined by xx). It plots barplots for partisans defined by pty, and for the treatment indicated by cue. x2 adds a double cue to the same figure

Value Graphical output only