Release Notes for the Manifesto Project Dataset: South America

Updates (2015a – 2017b), changes, corrections, and known errors

Known Errors:

- The pervote variable is missing for some Argentine elections as it is not possible to calculate an accurate vote share based on the election reports that we have.

Update 2017b: from December 2017

Elections added:

- Argentina 1995, 1999

New variable:

- We added the variable "id_perm" as a last column to the dataset. This variables contains six-digit character vectors that permanently identify observations across different versions of the dataset. In most cases this can also be achieved by using party-date as identifier. However, we sometimes correct or change party codes between different dataset versions.

Update 2017a: from July 2017

Elections added:

- Brazil 1994, 1998, 2002 (presidential elections)

Smaller corrections:

- Corrected the variables testresult and testeditsim for the 2003/2007 Argentine elections

Update 2016b: from December 2016

Elections added:

- Argentina 2003, 2007, 2009,
- Bolivia 2014
- Chile 2013
- Brazil 2014

Further changes:

- A new variable *datasetversion,* constant for all observations, indicates the version of the dataset. It helps keep track of the dataset version even if, for example, the filename is changed.

- We completed the standardization of the calculation of the coding reliability by slightly changing the coding of the entry test master copy of manual 4 and 5, and recalculated the reliability scores (Krippendorff's Alpha, variable *testresult*) of the coders that were trained with manual 4 or 5. This affects 456 cases. The change in Krippendorfs alpha for these cases ranges from -0.13 to +0.23 with a median of 0.003. We also added a new variable to measure the coding reliability called *testeditsim*. This measure compares the sequence of codes from a coder's entry test to the sequence of codes from the master copy of the entry test. The new variable *testeditsim* indicates the relative Levensthein distance between the two sequences of codes, and subtracts it from 1 (a value of 0 indicates no matching codes, 1 indicates all codes match).

- We added a new variable called *corpusversion* that indicates the version of the Manifesto Corpus that contains the documents from which the code frequencies. For more information on the Manifesto Corpus, see https://manifesto-project.wzb.eu/information/documents/corpus

Smaller corrections:

- We corrected the coding of pervote, absseat and totseats in Argentina as far as we could find information on electoral results. In the first version of the dataset absseat and totseats sometimes indicated the share of seats won by a party or the seats a party has in parliament. As in Argentina at regular parliamentary elections only half the seats in parliament are re-elected this produced an inconsistency. We harmonized this. Now, absseat describes a party's absolute number of seats in parliament after the election and totseats indicates the total number of seats in parliament. Pervote still indicates the share of votes von at the election.
- We corrected the edate variable for the Argentinan election in 2013 which was wrong for two observations. The date variable was correct.

Remark:

- This version is called 2016b (despite the fact that there is no 2016a version of the dataset) as it was published at the same time as the 2016b version of the main dataset. We attempt to publish future versions of both datasets in the same update rhythm.

Version 2015a: from 1 June 2015

First release.

Elections covered:

- Argentina 2011, 2013
- Brazil 2006, 2010
- Chile 1989, 1993, 1999, 1995, 2009