The Global Populism Database

Codebook

Kirk A. Hawkins (Brigham Young University)

Rosario Aguilar (Centro de Investigación y Docencia Económicas)

Bruno Castanho Silva (University of Cologne)

Erin K. Jenne (Central European University)

Bojana Kocijan (Independent Researcher)

Cristóbal Rovira Kaltwasser (Universidad Diego Portales)

Please cite as:

Hawkins, K. A., Aguilar, R., Castanho Silva, B., Jenne, E. K., Kocijan, B., and Rovira Kaltwasser, C. 2019. “Measuring Populist Discourse: The Global Populism Database”. *Paper presented at the 2019 EPSA Annual Conference in Belfast, UK, June 20-22*.

Contact authors: Kirk Hawkins (kirk_hawkins@byu.edu) and Bruno Castanho Silva (bcsilva@wiso.uni-koeln.de).

Description

Starting in 2006, we began creating a dataset of populist discourse for political leaders using textual analysis of political speeches. The initial effort (Hawkins 2009) covered contemporary and historical Latin American presidents plus a few presidents and prime ministers from other regions. Since then, the database has been expanded several times to more countries and time periods. After recent efforts in 2018-2019 as part of The New Populism project at The Guardian, we rolled out a combined version of our dataset called the *Global Populism Database*.[1]

The Global Populism Database (hereafter, GPD) applies a technique known as holistic grading which was designed by educational psychologists to measure diffuse, latent aspects of texts such as tone, style, and quality of argument. The technique, originally used to grade essays in the College Board AP exams, has coders apply an integer grade scale and a rubric to identify rough attributes of texts at each grade. Coders are then trained by repeated exposure to anchor texts, or texts that benchmark scores in the rubric (White 1985; Sudweeks, Reeve, and Bradshaw 2004).

In our rubric, texts are initially assigned one of three scores, listed below with their descriptions. In more recent versions, coders have used a decimal scale (0.1, 0.2, etc.) in which 0.5 rounds to a 1 and 1.5 rounds to a 2.

> 2 A speech in this category is extremely populist and comes very close to the ideal populist discourse. Specifically, the speech expresses all or nearly all of the elements of ideal populist discourse, and has few elements that would be considered non-populist.
>
> 1 A speech in this category includes strong, clearly populist elements but either does not use them consistently or tempers them by including non-populist elements. Thus, the discourse may have a romanticized notion of the people and the idea of a unified popular will (indeed, it must in order to be considered populist), but it avoids bellicose language or references to cosmic proportions or any particular enemy.
>
> 0 A speech in this category uses few if any populist elements. Note that even if a speech expresses a Manichean worldview, it is not considered populist if it lacks some notion of a popular will.

The sample of texts is a quota sample consisting of four speeches for each term in office: a campaign speech (usually the closing or announcement speech), a ribbon-cutting speech (marking a commemorative event with a small, domestic audience), an international speech (given before an audience of foreign nationals outside the country), and a famous speech (one widely circulated that represents the leader at his or her best). Where many speeches

[1] The bulk of data collection projects over the years were funded by Brigham Young University, Universidad Diego Portales, The Guardian Foundation, and the Comparative Populism Project at Central European University. We are grateful for their assistance but emphasize that the resulting measures and any errors in them are our own

are available, we rely on the most recent speech with at least 1-2,000 words (extremely short speeches are difficult to code).[2]

In our technique, each text is read and coded in its original language; for the majority of cases, each text is coded by two individuals in order to ensure intercoder reliability.[3] Coders do not share their work with each other until it is complete. Discrepancies of .5 or greater are subjected to a reconciliation session in which coders can adjust their scores if they demonstrate an error in coding, but otherwise differing scores are retained and averaged for a final score on each document. The four scores are then averaged (unweighted) to provide a single score for each leader.

The GPD includes 1,113 speeches, with a total of 2,003 grades assigned. For the 886 speeches graded by at least two coders, Krippendorff's alpha (interval-level) is 0.824, indicating high reliability. This means that we have confidence in the scores assigned by a single coder when a second was not available.

The dataset is disaggregated at the coder level, containing every single score given to each speech by each coder who evaluated it. The unit of observation is the speech-coder. The unit of analysis is the leader-term, meaning that four speeches are found for each president or prime ministers' term (say, four speeches for Obama I and another four for Obama II).

[2] Consistent with previous findings (Hawkins 2009, Hawkins and Castanho Silva 2018), in the GPD campaign speeches seem to be the most populist, with an average score of 0.60, followed by famous (0.42), ribbon-cutting (0.21) and international (0.20).

[3] While we try to have two coders read and grade each speech, this is not always possible due to funding constraints or the lack of additional native speakers of the language; thus, in some cases there is only one coder. Individual coders still meet with the project coordinators to discuss the scores they assign.

List of variables

Country

Leader - name of president or prime minister;

Party - party or political group to which the leader is affiliated;

lr - A categorical indicator with -1 = left; 0 = center; 1 = right. Our default indicator is the left-right indicator (dw) from the Democratic Accountability and Linkages Project (DALP, Kitschelt 2013), with party averages split into three categories: left if the party is at least 0.5 standard deviation below the unweighted mean of the dataset, right if it is at least 0.5 standard deviation above the mean, and centre if the party is in-between 0.25 standard deviation above and 0.25 standard deviation below the mean. For borderline cases (parties coded between 0.25 standard deviation and 0.5 standard deviation either side of the mean), we adjudicate using either Chapel Hill Expert Survey for European Parties (Bakker et al. 2015) or the Political Representation, Parties, and Presidents Survey for Latin America (PREPPS) for Latin America. For a small number of observations not included in any of these datasets, our coding relies on online descriptions of parties and consultations with country experts.

president - binary. 1 if the leader is a president, 0 if a prime minister;

term - leader term in office, ordinal.

startofterm - date of start of respective term;

yearbegin - year of beginning of term;

endofterm - date of end of respective term;

yearend - year of ending of respective term;

speechtype - type of speech: international, campaign, ribboncutting, famous;

speechnum - number of that speech for that leader-term;

codernum - coder identifier for that speech-leader;

rubricgrade - grade of that speech by that coder on the 0-2 populism scale;

averagerubric - average populism grade for that speech. Arithmetic average across all coders for that speech;

totalaverage - average populism grade for that leader-term. Arithmetic mean for all speeches by all coders for that leader-term.

wb_region - World Bank global region classification.