Codebooks and Description of Replication Data of the Paper "Media Coverage of Campaign Promises Throughout the Electoral Cycle"

This document describes the R scripts and all variables that are included in the replication datasets. Note that the scripts for retrieving the promise-related news articles and the full texts cannot be shared for copyright reasons. The dataset, however, contains the text of all sentences that have been classified as relating to political promises and non-political promises (the latter set of sentences is not relevant and omitted from the analysis).

R Scripts

- `01_replication_media_coverage_promises.R` contains the code to reproduce all plots and tables reported in the main manuscript.
- `02_replication_media_coverage_promises_supporting_information.R` contains the code to reproduce all plots and tables reported in the Supporting Information.

Datasets

`data_timeseries.rds` contains the variables required to reproduce the analyses relating to the classification and sentiment of newspaper articles in the main manuscript and the Supporting Information.

Variable name in data_timeseries.rds	Short description
`text_id`	Unique ID for each newspaper article
`date`	Date when newspaper article was published
`year`	Year of publication of newspaper article
`countryname`	Name of country in which the newspaper was published
`newspaper`	Name of newspaper
`tabloid_broadsheet`	Indicator whether newspaper is classified as a broadsheet or a tabloid
`cycle`	ID for each electoral cycle

country_cycle	Combination of country name and the electoral cycle in which a newspaper article was published
document	Identifier for each document
class_inductive	Class of sentence (using the inductive classification)
class_deductive	Class of sentence (using the deductive classification)
pledge_ongoing	Count variable that indicates the number of matches from the inductive/deductive dictionary for ongoing promises
pledge_broken_inductive	Count variable that indicates the number of matches from the inductive dictionary for broken promises
pledge_fulfilled_inductive	Count variable that indicates the number of matches from the inductive dictionary for fulfilled promises
pledge_broken_deductive	Count variable that indicates the number of matches from the deductive dictionary for broken promises
pledge_fulfilled_deductive	Count variable that indicates the number of matches from the deductive dictionary for fulfilled promises
ntoken	Number of tokens of the sentence
ndoc_cycle	Number of newspaper articles (separately for each electoral cycle)
nsentences_cycle	Number of sentences from newspaper articles (separately for each electoral cycle)
sentence	Text (sentence) that was extracted using the initial dictionary classification
mention_party_mp_context	Variable indicating whether a relevant politician or political party is mentioned in the sentence
elecdate	Date of next election
priorelecdate	Date of previous election
gdp_change	Change in GDP growth, compared to previous quarter
gdp_change_lag	Change in GDP growth, compared to previous quarter (lagged by one quarter)
log_gdp_change_lag	Logged variable of gdp_change_lag to account for skewness
gov_type	Type of government (single- or multiparty government). Source: ParlGov
gov_type_aus_single	Type of government (single- or multiparty government), but treating governments between the Liberal Party of Australia and the National Party of Australia as a single-party cabinet
gov_type_maj_min	Indicator whether government is a majority or minority government, based on proportion of seats by parties coded as government parties. Source: ParlGov
poll_change_to_election	Change in public support in a given quarter compared to the previous election (averaged for coalition governments across all government parties). Source: Müller and Louwerse (2018)
positive	Number of 'positive' words from Lexicoder Sentiment Dictionary in the sentence
negative	Number of 'negative' words from Lexicoder Sentiment Dictionary in the sentence
neg_positive	Number of 'negated positive' words from Lexicoder Sentiment Dictionary in the sentence
neg_negative	Number of 'negated negative words from Lexicoder Sentiment Dictionary in the sentence
sentiment	Sentiment score using the four sentiment categories and aggregated using the formula described in Proksch et al. (2019). Note that terms which were used for the classification of promise-related sentences are excluded from the Lexicoder Sentiment Dictionary.

data_evaluations.csv is the dataset used for the meta-analysis of existing studies on citizens' perceptions of election promises. The data were collected using the descriptions in the respective papers or the replication materials.

Variable name in data_evaluations.csv	Short description
study	Name of study (authors and year of publication)
pledge	Summary of the pledge to be evaluated by respondents
fulfilment	Indicator whether pledge has not been fulfilled, partially fulfilled, or fully fulfilled
incorrect	Proportion of respondents who gave incorrect answer
dont_know	Proportion of respondents who answered 'don't know' or did not reply
respondents	Number of valid respondents in survey
country	Country in which a study was conducted

data_sample_ongoing.xlsx contains the manual hand-coding of 400 sentences classified as 'ongoing' using the inductive classification. Sentences have been coded by the author of this paper.

Variable name in data_sample_ongoing.xlsx	Short description
sentence	Randomly sampled sentence classified as 'ongoing' by the dictionary (100 per country)
info_about_pledge	Manual coding whether sentence contains information on a political promise
info_broken_fulfilled	If a sentence was classified as relating to a political promise, this variable indicates whether the sentence mentions the breaking or fulfillment
info_broken_fulfilled_vague_concrete	If the sentence mentions the breaking or fulfilment of a promise, this variable indicates the vagueness/concreteness of the statement regarding the fulfilment or breaking of a promise
countryname	Name of country in which the newspaper was published
text_id	Unique ID for each newspaper article
newspaper	Name of newspaper
sentence_pre	Sentence that occurred before the statement classified by the dictionary as 'ongoing'
sentence_post	Sentence that occurred after the statement classified by the dictionary as 'ongoing'

`data_crowdcoded.csv` contains the raw coding retrieved from the crowdsourcing platform *Figure8*. The variables used for the assessment of the dictionary classification are created in the replication script. The following variables from the raw dataset are relevant for the reproduction of the results.

Variable name in data_crowdcoded.csv	Short description
`_unit_id`	Identifier for each sentence to be coded by crowd workers
`fulfilment`	A crowd worker's assessment regarding information on the fulfilment or breaking of a promise
`pledge`	Binary variable that indicates whether a crowd worker regarded a statement as a political promise
`sentence`	Sentence to be coded
`_golden`	Indicator whether an observation is the coding of a 'gold/test questions' which is used for the continuous quality control. These sentences and codings are omitted from the subsequent analysis.
`trust`	'Trust score' of worker which equals the proportion of correctly answered 'gold/test questions' throughout the coding job