Codebook for the datasets used in the article: Rethinking Public Funding of Parties and Corruption: Confronting Theoretical Complexity and Challenging Measurement

Description

In this research, we use three datasets with the same variable names when they are present in all datasets. "impact.dta" is the primary dataset used for the analysis in the article and online appendices. Explanatory and control variables are aggregated based on Business Environment and Enterprise Performance Survey (BEEPS) rounds; hence they are not uniformly aggregated since the BEEPS surveys were conducted at uneven time intervals.

Additionally, we employ two other datasets, "impact_lags_aggr.dta" and "impact_lags.dta". "impact_lags_aggr.dta" contains fewer variables, although with the same names for the identical ones. However, explanatory and control variables are aggregated using a uniform 3-year lag structure. Explanatory and control variables are aggregated by averaging their values for the previous three years. This dataset is used to estimate Models 3 and 4 presented in Tables 3 and 4 in the article and Tables B15 and B16 in the online appendix B.

"impact_lags.dta" is the most disaggregated dataset. It is used to estimate models with 1- to 5-year lag structure presented in Table 5 of the article. However, it does not aggregate the explanatory and control variables. See the replication files for the construction of lags.

Table 1 below provides variable names and data sources for variables used in the study.

Table 1: Codebook for variable names and sources in all datasets

Variable	Variable name	Source
country	Country code	Not applicable
country_name	Country name	Not applicable
year	Year when the BEEPS round was released	Not applicable
elections_dummy	Whether there was or not a parliamentary contest organized in the year when the respective BEEPS round was released	Authors' coding
elections_frequency	Number of parliamentary elections held between BEEPS rounds	Authors' coding
dpfrv_total	Nominal amount of direct public funding per registered voter	Lipcean (2021)
dpfrv_total_infl	Amount of direct public funding per registered voter, inflation-adjusted	Authors' calculations based on Lipcean (2021)
dpfv_dprc_infl	Amount of direct public funding per vote, inflation-adjusted and discounted	Ibidem.

dpfv_dprc_infl_lag	Amount of direct public funding per vote, inflation-adjusted, discounted and lagged	Ibidem.
dpfv_nomin_infl	Amount of direct public funding per vote, inflation-adjusted	Ibidem.
dpfv_nomin_infl_lag	Amount of direct public funding per vote, inflation-adjusted and lagged	Ibidem.
dpfv_total	Nominal amount of direct public funding per vote	Ibidem.
dpfv_total_lag	Nominal amount of direct public funding per vote, lagged	Ibidem.
capture_impact_1	Party and legislative corruption	Authors' estimations based on BEEPS data
capture_impact_1_lag	Party and legislative corruption lagged	Ibidem.
capture_impact_1_norm	Party and legislative corruption: normalized	Ibidem.
capture_impact_1_norm_lag	Party and legislative corruption: normalized and lagged	Ibidem.
capture_impact_2	Party, legislative and central government corruption	Ibidem.
capture_impact_2_lag	Party, legislative and central government corruption: lagged	Ibidem.
capture_impact_2_norm	Party, legislative and central government corruption: normalized	Ibidem.
capture_impact_2_norm_lag	Party, legislative and central government corruption: normalized and lagged	Ibidem.
capture_impact_3	Party, legislative, central and local government corruption	Ibidem.
capture_impact_3_lag	Party, legislative, central and local government corruption lagged	Ibidem.
capture_impact_3_norm	Party, legislative, central and local government corruption: normalized	Ibidem.
capture_impact_3_norm_lag	Party, legislative, central and local government corruption: normalized and lagged	Ibidem.
v2x_libdem	Democracy level-VDEM	Coppedge et al. (2021)
democracy_scaled	Democracy level-VDEM rescaled	Authors' rescaling based on Coppedge et al. (2021)
democracy_scaled_lag	Democracy level-VDEM rescaled lagged	Ibidem.
elect_syst	Electoral system: 1-Majoritarian, 2-Mixed, 3-Proportional	Bormann & Golder (2013);

		International IDEA (2021)
elect_syst_lag	Electoral system: 1-Majoritarian, 2-Mixed, 3-Proportional: lagged	Ibidem.
ethn_fr	Ethnic fractionalization	Drazanova (2020)
ethn_fr_lag	Ethnic fractionalization: lagged	Ibidem.
gdppc_ppp_2017	GDP per capita in purchasing power parity	Gapminder dataset
gdppc_ppp_2017_lag	GDP per capita in purchasing power parity: lagged	Ibidem.
gdppc_ppp_2017_log	Log of GDP per capita	Ibidem.
gini_disp	Inequality: Gini index (disposable income)	World indequality database
gini_disp_lag	Inequality: Gini index (disposable income): lagged	Ibidem.
gov_size_imf	Size of the governmental sector: percentage of GDP	International Monetary Fund
gov_size_imf_lag	Size of the governmental sector: percentage of GDP: lagged	Ibidem.
govern_type2	Regime type: 1-Presidential, 2-semi-presidential, 3-parliamentary	Cruz et al. (2021); Elgie (2018)
govern_type2_lag	Regime type: 1-Presidential, 2-semi-presidential, 3-parliamentary: lagged	Ibidem.
judicial_scaled	Judicial independence	Staton et al. (2019)
press_freedom_rev	Press Freedom reversed: higher scores, more freedom	Freedom House
private_agreg	Political donations restrictions	Authors' coding from national regulations
private_agreg_lag	Political donations restrictions: lagged	Ibidem.
unique_id	A unique identifier used for observation-based Jackknife variance estimations	Not applicable

References

Bormann, N.-C. & Golder, M. (2013). Democratic Electoral Systems around the world, 1946–2011. *Electoral Studies 32*(2): 360–369.

Coppedge, M., Gerring, J., Knutsen, C.H., Lindberg, S.I. & Teorell, J. (2021). *V-Dem Dataset*. Varieties of Democracy (V-Dem) Project. Retrieved from https://www.v-dem.net/en/data/data/v-dem-dataset-v111/

Cruz, C., Keefer, P. & Scartascini, C. (2021). *The Database of Political Institutions 2020*. Inter-American Development Bank. Retrieved from https://publications.iadb.org/en/node/29752

Drazanova, L. (2020, September 12). Historical Index of Ethnic Fractionalization Dataset (HIEF). Harvard Dataverse. Retrieved from https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/4JQRCL

Elgie, R. (2018). *Semi-presidentialism – A new country-years dataset*. Retrieved from https://presidential-power.net/?p=7869

International IDEA. (2021). *Electoral System Design Database*. Retrieved from https://www.idea.int/data-tools/data/electoral-system-design

Lipcean, S. (2021). Dataset on direct public funding of political parties across post-communist regimes 1990-2020. *Harvard Dataverse*. Retrieved from https://doi.org/10.7910/DVN/TNHZML

Staton, J., Linzer, D., Reenock, C. & Holsinger, J. (2019). Update, A Global Measure of Judicial Independence, 1900-2015. Harvard Dataverse.