Codebook - Distancing from Corruption-Tainted Parties: Swing-Voter Dependence and Party-Organization Availability

Thiago do Nascimento Fonseca

Getúlio Vargas Foundation (Brazil)

Luís de Sousa

University of Lisbon (Portugal)

Table 1 - Description

Variable	Description (short label)	Type / Scale	Coding / Construction	Notes / Source
ANO_ELEICAO	Municipal election year	Integer (year)	Values observed in panel: 2004, 2008, 2012, 2016, 2020	Brazil holds municipal elections every 4 years; 2016 is the first post-scandal municipal election considered.
SIGLA_PARTIDO	Party abbreviation	Categorical (string)	Standard TSE party acronyms (e.g., PT, MDB, PSDB, PDT, PL, etc.)	Party metadata come from TSE/FGV CEPESP electoral records.
NUMERO_PARTIDO	Party numeric code	Integer	TSE ballot code for each party	From TSE/FGV CEPESP data.
COD_MUN_IBGE	Municipality code (IBGE)	Integer (ID)	7-digit IBGE identifier for municipality *m*	Used to identify municipalities in the panel and for clustering.
UF	State abbreviation	Categorical (string)	Two-letter state code (e.g., SP, RJ, BA)	From TSE/FGV CEPESP records.
vote.counc	Council (city council) vote share of party *p* in municipality *m*, election *t*	Proportion in [0,1]	Party's share of valid votes in the councilor election	Vote share is one of the electoral outcomes analysed.
cand.counc	Share of councilor candidates affiliated with party *p* in *m*, *t*	Proportion in [0,1]	Party's fraction of councilor candidacies among all council candidates in the municipality	Used to test whether parties less reliant on swing voters show weaker distancing.

Variable	Description (short label)	Type / Scale	Coding / Construction	Notes / Source
cand.mayor	Party *p* nominates a mayoral candidate in *m*, *t*	Binary {0,1} or proportion	In most implementations: 1 if the party fielded ≥1 mayoral candidate in the municipality; 0 otherwise (some tables may report proportion)	The **primary** outcome is the (share/probability of) mayoral candidacies by party.
vote.mayor	Mayoral vote share of party *p* in *m*, *t*	Proportion in [0,1]	Party's share of valid votes for mayor in the municipality	Used to support anticipation/punishment logic.
nep	Effective Number of Parties (municipal)	Continuous (≥1)	Laakso–Taagepera ENP computed at the municipality–election level (from vote or seat distributions; used as time-varying covariate)	Included as a covariate in doubly-robust DiD models.
nep.is	Effective party indicator (binary)	Binary {"0", "2016"}	For each municipality–election, order parties by seat share from highest (rank = 1) to lowest. Let nep be the Effective Number of Parties (Laakso–Taagepera) for that municipality–election. Set nep.is = 1 if the party's rank ≤ nep; otherwise nep.is = 0.	Computed within each municipality–election using council (legislative) seat shares (or the series you use to compute nep) to define the ranking.
pib.def	Deflated (real) municipal GDP	Continuous	Municipality-level GDP deflated to constant prices	Used as a time-varying covariate in the models.
pib.def.log	Log of deflated municipal GDP	Continuous (log)	Natural logarithm of **pib.def**	Included as covariate to capture economic scale.
id	Panel unit identifier	Integer/ string	Identifier for the panel unit used in the DiD routines (in your scripts, used for clustering at the municipality level)	Each observation corresponds to a party–municipality–year; errors clustered by municipality.
time	Panel time index (municipal election year)	Integer (year)	Typically takes values 2004, 2008, 2012, 2016, 2020	Matches municipal election calendar; 2016+ is the post-treatment period.
treatment	Scandal exposure (party-level)	Binary {"0", "2016"}	=1 for PT (Workers' Party); =0 for all other parties, in all municipalities	PT is the treated group; others are controls. First affected municipal election is 2016.