## 1. OVERVIEW

This project includes two datasets. The National Level Party System Dataset (NLPS) includes all the free elections between 1946 and 2016 that occur in countries that have power alternations. For consistency purposes, we exclude countries that are democratic but have not yet had power alternations, such as South Africa and Namibia. For the long-term democracies with simple systems, which are used for data analyses in the Shugart and Taagepera (2017) book, some early 2017 elections are also included. Scholars who want to use all country elections that occurred in a given year may want to drop the 2017 elections before running cross-national analysis. In NLPS, each observation describes one assembly election and, in case of presidential and semi-presidential system, the concurrent or immediately prior presidential election.

The District Level Party System Dataset (DLPS) has similar case selection criteria as NLPS, but it does not include multi-tiered systems. In DLPS, each observation describes one district in one legislative election.

The two datasets can be merged using *country*, *year*, and *second_election* variables.

## 2. VARIABLES IN NLPS

**Case identification variables**

*country*: The name of the country.

*year*: The year the legislative election is held.

*date*: The day the legislative election is held.

*second_election*: A dummy variable that takes the value of 1 if the legislative election is the second one in the same year and 0 if not.

**Institutional variables**

*presregime*: 1the executive format is (pure) presidential; 0 if other format

*SP*: 1 if the executive format is semi-presidential; 0 if other format

*non_SP_hybrid*: 1 the executive format is a hybrid but not semi-presidential; 0 if other format

*assemb*: The total number of seats in the lower house of the legislature.

*uppershare*: The share of total assembly seats in the upper tier(s)

*simple*: 1 if the electoral system is simple, as defined in Shugart and Taagepera (2017); 0 if not.

*compens*: 1 if the electoral system is compensatory (via the upper tier); 0 if not compensatory.

*parallel*: 1 if the basic tier seats and the upper tier seats are separately calculated, in other words, the electoral system is parallel; 0 if not parallel.

*fptp*: 1 if the assembly election uses a first-past-the-post electoral system; 0 if other system.

*mmp*: 1 if the assembly election uses a mixed-member proportional system; 0 if other system.

*sntv*: 1 if the assembly election uses a single non-transferable vote system; 0 if other system.

*two_rd*: 1 if the assembly election uses a two-round system; 0 if other system.

*legal_thresh*: If the electoral system has a legal threshold, the share of national votes a party has to earn to be entitled to any seat.

*listtype*: If there is a list tier, the type of the list.

0: non-list system
1: open list
2: flexible (semi-open) list
3: closed list
4: other list type, specifically
        Colombia 1991–2002, personal list (quasi-SNTV)
        Colombia 2006-2014, party option of closed or open list
        Denmark, party option of open or various types of flexible list
        El Salvador 2012-15, free list (*panachage*)
        Honduras, 2005-2013, free list (*panachage*)
        Luxembourg, free list (*panachage*)
        Switzerland, free list (*panachage*)
        Uruguay, factional list (sub-party pooling)

**Legislative election results**

*Ns0*: The total number of parties in the legislature, regardless of party size. Each independent is considered one party.

*Ns*: Effective number of seat-winning parties.

*Nv*: Effective number of vote-earning parties.

*s1*: The largest assembly party's seat share.

*v1*: The largest assembly party's vote share.

*v1_e*: The largest electoral party's vote share.

*plu_reversal*: A dummy variable that takes the value of 1 if the party that wins the most votes fails to win the most seats (v1 ≠ v1_e) and 0 if not.

*pty_n_s1*: The name of the party with the largest seat share.

*pty_n_s1_dist_match*: The name of the party with the largest seat share, matched to the district party name if in the DLPS.

*disprop*: Gallagher Index of disproportionality.

**Previous presidential election results**

The following variables describe the presidential election that is either concurrent with the legislative election in the same row or, in case there is no concurrent election, the latest before that assembly election.

*date_pres_prior*: Date of the presidential election.

*Np*: The effective number of presidential candidates.

*presc11, presc12, presc13, presc14, presc15, presc16*: The vote share of the six largest vote winners in the first round of the presidential election.

*presc21, presc22*: The vote share of both presidential candidates in the second round.

*pty_n_pres*: The party name of the winning presidential candidate.

*pty_n_pres_dist_match*: The party name of the winning presidential candidate, matched to the district party name if in the DLPS.

*pres_largest*: A dummy variable that takes the value of 1 if the presidential party is also the largest party in the assembly and 0 if not.

*prespty_s_nat*: Presidential party's seat share in the assembly.

*prespty_v_nat*: Presidential party's vote share in the assembly election.

*runoff_comeback*: A dummy variable that takes the value 1 if the presidential candidate with a plurality of the votes in the first round loses the second round and 0 if not.

*pty_n_runner_up_pres*: The party name of the runner-up presidential candidate.

*pty_n_runner_up_pres_dist_match*: The party name of the runner-up presidential candidate, matched to the district party name if in the DLPS.

### 3. VARIABLES IN DLPS[1]

*country:* The name of the country.

*ctr*: The numerical code of the country consistent with Kollman et al.'s (2007) Constituency-Level Elections Archive.

*year:* The year the legislative election was held.

*cst_n*: The name of the district.

*cst*: The numerical code of the district.

*mag*: The district magnitude.

*Ns_dist*: Effective number of seat-winning parties of the district.

*Nv_dist*: Effective number of vote-winning parties of the district.

*Ns0_dist*: The number of parties that win at least one seat in the district.

*Nv0_dist*: The number of parties that win at least one vote in the district.

*v1_dist, v2_dist, v3_dist, v4_dist, v5_dist*: The vote shares of the five largest parties in the district.

*s1_dist, s2_dist, s3_dist, s4_dist, s5_dist*: The seat shares of the five largest parties in the district.

*pty_n1_dist, pty_n2_dist, pty_n3_dist, pty_n4_dist, pty_n5_dist*: The names of the five largest parties in the district.

*mn_mag*: The mean district magnitude of the election.

*min_mag*: The smallest district magnitude of the election.

*max_mag*: The largest district magnitude of the election.

*sd_mag*: The standard deviation of the district magnitude of the election.

*second_election:* A dummy variable that takes the value of 1 if the legislative election is the second one in the same year and 0 if not.

---

[1] The variables that overlap with the NLPS are not listed in this section.