

CODEBOOK

Chilean Political Landscape Dataset

CPLD • Version 3.0 • Variable Definitions

FILE	cpld_v3_0.csv
VERSION	3.0
OBSERVATIONS	6,119
VARIABLES	42
COVERAGE	1989–2025
AUTHOR	Kenneth Bunker
CONTACT	kenneth.bunker@uss.cl
DATAVERSE	doi.org/10.7910/DVN/NGRY3R
WEBSITE	labdemgob.github.io/cpld
LICENSE	CC BY 4.0

HOW TO READ THIS CODEBOOK

Each entry names a variable, gives its definition, and where applicable the formula used to compute it, its type and unit, and its full source. Variables are numbered sequentially from the identification block onward. Entries sourced from the Servicio Electoral de Chile (SERVEL) are the author’s elaboration based on certified electoral results.

Unit of observation. Each row represents one electoral unit in one election. All variables are unit-level aggregates; the number of lists competing in the unit is itself recorded in the variable **list**. The electoral unit varies by election type: **district** for Diputados and Constituyentes, **region** for Senadores and Gobernadores, **province** for Cores, and **commune** for Alcaldes and Concejales. Row counts therefore equal unit counts: Diputados 2021 has 28 rows (28 districts); Alcaldes 2024 has 345 rows (345 communes).

ANATOMY OF AN ENTRY

Number · name · type	Running variable number, the column name, and its data type or unit.
Definition	What the variable measures and how it is constructed.
Formula	The computation, shown in a tinted box where one applies.
Source	Full APA reference, repeated in each entry for self-contained citation.

NOTATION USED IN FORMULAS

v_i – vote share of list i	s_i – seat share of list i
p_i – district vote proportion	P_i – national vote proportion
w_i – weight (vote share)	x_i – ideological score of list i
dm – district magnitude	n – number of lists in the unit

01 IDENTIFICATION VARIABLES

01 **type** STRING

Type of election. Values: Diputados, Senadores, Constituyentes, Alcaldes, Concejales, Cores, Gobernadores, Municipales (elections prior to the split of Alcaldes/Concejales).

Source – Author’s elaboration based on Servicio Electoral de Chile (SERVEL). (n.d.). *Certified electoral results*.

02 **type2** STRING

Electoral system group. Identifies sets of elections sharing the same rules (district boundaries, magnitudes, seat totals) – e.g. Diputados 1989–2013 (60 districts, binominal); Diputados 2017–2021 (28 districts, proportional); Senadores Binominal (Pares/Impares); Cores 2021.

Source – Lijphart, A. (1994). *Electoral systems and party systems: A study of twenty-seven democracies, 1945–1990*. Oxford University Press.

03 **type3** STRING

Election label. Concatenation of type and election year, identifying a specific election within a type. Example: “Diputados - 2021”.

Source – Author’s elaboration.

04 **region_2024** NUMERIC

Regional identifier using 2024 regional boundaries, enabling longitudinal comparison across elections despite historical boundary changes. Units from elections under older boundaries are assigned to the region that contains them under the 2024 definition. Notable assignments: Diputados district 42 → region 16; Senadores circ. 1 (Arica y Tarapacá) → region 15; Senadores circ. 13 (Bio-Bio Cordillera) → region 8.

Source – Author’s elaboration.

05 **unit** INTEGER / STRING

Identifier of the electoral unit. Integer for district- and region-based elections; string for commune-based elections.

Source – Author’s elaboration based on Servicio Electoral de Chile (SERVEL). (n.d.). *Certified electoral results*.

06 **election** INTEGER · 1989–2025

Year of the election.

Source – Author’s elaboration based on Servicio Electoral de Chile (SERVEL). (n.d.). *Certified electoral results*.

02 LIST STRUCTURE VARIABLES

07 **list** INTEGER

Number of electoral lists competing in the unit. Counts all distinct lists including IFP candidates; each IFP candidate constitutes one list.

Source — Author's elaboration based on Servicio Electoral de Chile (SERVEL). (n.d.). *Certified electoral results*.

08 **lists_national** INTEGER

Number of distinct electoral lists competing nationally in the election, excluding IFP candidates. Measures the national menu of partisan options.

Source — Author's elaboration based on Servicio Electoral de Chile (SERVEL). (n.d.). *Certified electoral results*.

09 **listmin** NUMERIC · %

Minimum vote share (%) obtained by any list in the unit.

Source — Author's elaboration based on Servicio Electoral de Chile (SERVEL). (n.d.). *Certified electoral results*.

10 **listmax** NUMERIC · %

Maximum vote share (%) obtained by any list in the unit.

Source — Author's elaboration based on Servicio Electoral de Chile (SERVEL). (n.d.). *Certified electoral results*.

11 **candidates** INTEGER

Total number of individual candidates competing in the unit, excluding null and blank votes.

Source — Author's elaboration based on Servicio Electoral de Chile (SERVEL). (n.d.). *Certified electoral results*.

12 **ifp_candidates** INTEGER

Number of Independientes Fuera de Pacto (IFP) candidates competing in the unit – independents running outside any coalition or party list.

Source – Author's elaboration based on Servicio Electoral de Chile (SERVEL). (n.d.). *Certified electoral results*.

13 **ifp_seats** INTEGER

Number of seats won by IFP candidates in the unit.

Source – Author's elaboration based on Servicio Electoral de Chile (SERVEL). (n.d.). *Certified electoral results*.

14 **percent_ind** NUMERIC · %

Percentage of total candidates who are IFP candidates.

$$\text{percent_ind} = (\text{ifp_candidates} / \text{candidates}) \times 100$$

Source – Author's elaboration.

03 ELECTORAL SYSTEM VARIABLES

15 **dm** INTEGER

District magnitude – number of seats allocated in the unit. For majoritarian elections (Alcaldes, Gobernadores), dm = 1; for proportional elections, dm varies by unit.

Source – Author's elaboration based on Servicio Electoral de Chile (SERVEL). (n.d.). *Certified electoral results*.

16 **as** INTEGER

Assembly size (local). Equals dm for most proportional election types. For Cores, equals the total seats allocated within the region to which the unit belongs.

Source – Author's elaboration.

17 **as2** INTEGER

Assembly size (national). Total seats allocated nationally in the election – the sum of dm across all units.

Source – Author's elaboration.

18 **sp** NUMERIC

Seat product – key predictor of fragmentation in the Seat Product Model.

$$sp = dm \times as$$

Source – Shugart, M. S., & Taagepera, R. (2017). *Votes from seats: Logical models of electoral systems*. Cambridge University Press.

19 **spm** NUMERIC

Seat Product Model predicted ENP – theoretical prediction of the effective number of parties from the seat product.

$$spm = (dm \times as)^{(1/6)} = sp^{(1/6)}$$

Source – Shugart, M. S., & Taagepera, R. (2017). *Votes from seats: Logical models of electoral systems*. Cambridge University Press.

20 **logmag** NUMERIC

Natural logarithm of district magnitude.

$$\logmag = \ln(dm)$$

Source – Author's elaboration.

04 PARTY SYSTEM FRAGMENTATION

21 **enpv** NUMERIC

Effective number of parties by votes at the district level – parties weighted by vote share. IFP candidates included.

$$ENPV = 1 / \sum p_i^2$$

Source – Laakso, M., & Taagepera, R. (1979). "Effective" number of parties: A measure with application to West Europe. *Comparative Political Studies*, 12(1), 3-27. <https://doi.org/10.1177/001041407901200101>

22 enps NUMERIC

Effective number of parties by seats at the district level. IFP candidates included.

$$\text{ENPs} = 1 / \sum s_i^2$$

Source — Laakso, M., & Taagepera, R. (1979). "Effective" number of parties: A measure with application to West Europe. *Comparative Political Studies*, 12(1), 3-27. <https://doi.org/10.1177/001041407901200101>

23 enpv_national NUMERIC

Effective number of parties by votes at the national level, computed on aggregated national vote totals, excluding IFP candidates.

$$\text{ENPv_national} = 1 / \sum P_i^2$$

Source — Laakso, M., & Taagepera, R. (1979). "Effective" number of parties: A measure with application to West Europe. *Comparative Political Studies*, 12(1), 3-27. <https://doi.org/10.1177/001041407901200101>

24 enps_national NUMERIC

Effective number of parties by seats at the national level, excluding IFP candidates.

$$\text{ENPs_national} = 1 / \sum S_i^2$$

Source — Laakso, M., & Taagepera, R. (1979). "Effective" number of parties: A measure with application to West Europe. *Comparative Political Studies*, 12(1), 3-27. <https://doi.org/10.1177/001041407901200101>

25 dif NUMERIC

Difference between the raw number of lists and ENPv at the district level — how many "effective" lists are lost to fragmentation of small parties.

$$\text{dif} = \text{list} - \text{enpv}$$

Source — Author's elaboration.

26 dif2 NUMERIC

Reduction in effective parties from translating votes into seats. Values below zero are set to zero to correct for rounding error.

$$\text{dif2} = \max(0, \text{enpv} - \text{enps})$$

Source — Author's elaboration.

05 DISPROPORTIONALITY INDICES

27 **Rae**

Rae disproportionality index over vote shares (v_i) and seat shares (s_i) across the n lists; higher values indicate greater disproportionality.

$$\text{Rae} = (1/n) \cdot \sum |v_i - s_i|$$

Source – Rae, D. W. (1967). *The political consequences of electoral laws*. Yale University Press.

28 **LH**

Loosemore-Hanby disproportionality index.

$$\text{LH} = 0.5 \cdot \sum |v_i - s_i|$$

Source – Loosemore, J., & Hanby, V. J. (1971). The theoretical limits of maximum distortion: Some analytic expressions for electoral systems. *British Journal of Political Science*, 1(4), 467–477.
<https://doi.org/10.1017/S000712340000925X>

29 **Lijphart_1**

Largest deviation from proportionality among lists holding at least one seat.

$$\text{Lijphart}_1 = \max |v_i - s_i| \text{ for } s_i > 0$$

Source – Lijphart, A. (1990). The political consequences of electoral laws, 1945–1985. *American Political Science Review*, 84(2), 481–496. <https://doi.org/10.2307/1963530>

30 **Lijphart_2**

Largest deviation from proportionality among all lists.

$$\text{Lijphart}_2 = \max |v_i - s_i| \text{ for all } i$$

Source – Lijphart, A. (1994). *Electoral systems and party systems: A study of twenty-seven democracies, 1945–1990*. Oxford University Press.

31 **Gallagher**

Gallagher least-squares disproportionality index.

$$G = \sqrt{0.5 \cdot \sum (v_i - s_i)^2}$$

Source – Gallagher, M. (1991). Proportionality, disproportionality and electoral systems. *Electoral Studies*, 10(1), 33–51. [https://doi.org/10.1016/0261-3794\(91\)90004-C](https://doi.org/10.1016/0261-3794(91)90004-C)

32 Cox_Shugart

Regression-based measure of the relationship between vote and seat shares. A value of 1.0 indicates perfect proportionality; values above 1.0 indicate a bonus to larger parties.

$$\text{Cox_Shugart} = b, \text{ where } s_i = a + b \cdot v_i \text{ (OLS)}$$

Source — Cox, G. W., & Shugart, M. S. (1991). Comment on Gallagher's "Proportionality, disproportionality and electoral systems." *Electoral Studies*, 10(4), 348-352. [https://doi.org/10.1016/0261-3794\(91\)90025-N](https://doi.org/10.1016/0261-3794(91)90025-N)

06 PARTY SYSTEM NATIONALIZATION**33 Mainwaring_Jones** NUMERIC · [0,1]

Nationalization score — homogeneity of a party's vote distribution across units. 1 = perfectly uniform; lower = greater regional concentration. Aggregated across parties weighted by national vote share.

$$\text{PSN}_i = 1 - \text{Gini}(p_{i1}, \dots, p_{iK})$$

Source — Jones, M. P., & Mainwaring, S. (2003). The nationalization of parties and party systems: An empirical measure and an application to the Americas. *Party Politics*, 9(2), 139-166. <https://doi.org/10.1177/13540688030092002>

34 Chhibber_Kollman NUMERIC

Difference between national ENP and the average district-level ENP — how much more concentrated the national party system is than its local components.

$$\text{CK} = \text{ENPv_national} - \text{mean}(\text{ENPv_district})$$

Source — Chhibber, P., & Kollman, K. (2004). *The formation of national party systems: Federalism and party competition in Canada, Great Britain, India, and the United States*. Princeton University Press.

07 IDEOLOGICAL VARIABLES**35 av_id** NUMERIC · 1-10

Vote-weighted mean ideological score of the district, excluding IFP (score = 5). Scale: 1 (far left) to 10 (far right).

$$\text{av_id} = \sum(w_i \cdot x_i) / \sum w_i$$

Source — Bunker, K. (2025). Mapping party systems over time: A spatial distribution of ideological competition in Chile. *Journal of Political Ideologies*. <https://doi.org/10.1080/13569317.2025.2452179>

36 polarization_district NUMERIC

Ideological polarization at the district level – vote-weighted dispersion across the spectrum, excluding IFP. **NA** for two communes in Alcaldes 2021 with no matchable lists.

$$P = \sqrt{(\sum w_i \cdot (x_i - \bar{x})^2)}$$

Source – Dalton, R. J. (2008). The quantity and the quality of party systems: Party system polarization, its measurement, and its consequences. *Comparative Political Studies*, 41(7), 899-920.
<https://doi.org/10.1177/0010414008315860>

37 polarization_national NUMERIC

Ideological polarization at the national level – same formula applied to national vote shares, excluding IFP.

Source – Dalton, R. J. (2008). The quantity and the quality of party systems: Party system polarization, its measurement, and its consequences. *Comparative Political Studies*, 41(7), 899-920.
<https://doi.org/10.1177/0010414008315860>

08 CONTEXTUAL VARIABLES**38 ethnic_frac** NUMERIC · [0,1]

Ethnic fractionalization – probability that two randomly selected individuals from the unit belong to different ethnic groups (2017 census). 0 = homogeneous, 1 = maximally fractionalized.

$$F = 1 - \sum (n_g / N)^2$$

Source – Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2), 155-194. <https://doi.org/10.1023/A:1024471506938>

39 ethnic_presence NUMERIC · [0,1]

Share of the population identifying as belonging to an indigenous people (2017 census).

$$\text{ethnic_presence} = n_indigenas17 / \text{censo17}$$

Source – Instituto Nacional de Estadísticas. (2017). *Censo de Población y Vivienda*.

40 ethnic_number NUMERIC · %

Share of the population identifying as indigenous (2002 census). **NA** for some units due to gaps in the 2002 census.

$$\text{ethnic_number} = (n_indigenas02 / \text{censo02}) \times 100$$

Source – Instituto Nacional de Estadísticas. (2002). *Censo de Población y Vivienda*.

09 VOTE AND STRUCTURAL VARIABLES

ACCESS THE DATASET



SCAN FOR THE PROJECT PAGE

labdemgob.github.io/cpld

Downloads, codebook, version history, and replication materials for the CPLD.

MAINTAINER	Kenneth Bunker
LABORATORY	Laboratorio Democracia y Gobierno
FACULTY	Economía, Negocios y Gobierno
UNIVERSITY	Universidad San Sebastián
LOCATION	Santiago, Chile
CONTACT	kenneth.bunker@uss.cl
REPOSITORY	github.com/labdemgob
DATAVERSE	doi.org/10.7910/DVN/NGRY3R
FUNDING	USS-FIN-26-PNTE-07
LICENSE	CC BY 4.0