



Data Article

Ideological positions in the chilean chamber of deputies (2002–2026): A legislative roll-call dataset

Jorge Fábrega

Center for the study of social complexity. Universidad del Desarrollo, Chile

ARTICLE INFO

Article history:

Received 14 July 2025

Accepted 6 October 2025

Available online 14 October 2025

Dataset link: [Ideological Estimates of the Chilean Chamber of Deputies, 2002–2026 \(Reference data\)](#)

Keywords:

Legislative behavior

Roll-call votes

Ideal point estimation

Political ideology

Chilean congress

Dynamic scaling

ABSTRACT

This dataset provides ideological position estimates for members of the Chilean Chamber of Deputies across six consecutive legislative periods, spanning from 2002 to 2026 in four-year intervals. The estimates are based exclusively on roll-call votes related to bill proposals debated and voted on in the plenary sessions. It excludes other types of votes, such as those on non-binding agreements or investigative commissions, which do not directly contribute to lawmaking.

To estimate ideological positions, the dataset includes results from three distinct methods. The first uses the classic DW-NOMINATE model, primarily for comparative purposes. The second applies a Bayesian model to estimate latent ideological positions within each legislative period. The third introduces a dynamic model suitable for large scale data that links estimates across time, allowing direct comparisons of ideological positions even when legislators served in different periods.

In addition to ideological estimates, the dataset includes information on each deputy's party affiliations throughout their time in office, as well as their age and date of birth. A major electoral reform enacted during the middle of this time frame increased the number of deputies from 120 to 155. As a result, earlier periods contain fewer legislators than later ones—a structural change that enhances the dataset's value for studying institutional evolution and its effects on legislative dynamics.

E-mail address: jfabrega@udd.cl

Social media: [@jorgefabrega](#)

<https://doi.org/10.1016/j.dib.2025.112163>

2352-3409/© 2025 The Author. Published by Elsevier Inc. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>)

Each methodological approach offers a different lens to examine ideological variation and change within the Chilean legislature. The dynamic model, in particular, supports longitudinal research by enabling consistent ideological comparisons across more than two decades. Researchers can use this dataset to explore topics such as political polarization, coalition dynamics, party-system development, and the relationship between institutional reforms and representation. The structure and transparency of the dataset make it suitable for integration with other legislative or electoral databases, offering additional opportunities for analysis in political science, public policy, and institutional studies.

© 2025 The Author. Published by Elsevier Inc.
This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>)

Specifications Table

Subject	Social Sciences
Specific subject area	The evolution of legislative ideology in the Chilean Chamber of Deputies based on roll-call voting data (2002–2026).
Type of data	Raw, Tables, Processed data and graphs
Data collection	Roll-call votes were obtained from the official website of the Chilean Chamber of Deputies. The data were processed and cleaned using R (v4.3.0). Ideal points were estimated using three methods to generate a one-dimensional ideological scale, implemented through the wnominate, pscl, and emIRT packages, respectively. The first two methods produce estimates independently within each legislative period, while the third method enables comparisons across periods by incorporating a dynamic structure. For the Bayesian and dynamic models, confidence intervals were also computed.
Data source location	Institution: Cámara de Diputadas y Diputados de Chile (Chilean Chamber of Deputies) City: Valparaíso Country: Chile Primary data source: Official website of the Chilean Chamber of Deputies – https://www.camara.cl Type of data: Secondary data (roll-call votes) Latitude and longitude: Approx. –33.0458° S, –71.6197° W The dataset was compiled from publicly available roll-call votes published by the Chilean Chamber of Deputies on its official website.
Data accessibility	Repository name: Ideological Estimates of the Chilean Chamber of Deputies, 2002–2026 Data identification number: https://doi.org/10.7910/DVN/FOXOIT Direct URL to data: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/FOXOIT
Related research article	None

1. Value of the Data

- This dataset provides systematically estimated ideological positions for all members of the Chilean Chamber of Deputies across six legislative periods (2002–2026), enabling the study of ideological change and legislative polarization over time.
- It offers high-resolution information on intra-party and inter-party dispersion, which is useful for assessing discipline, fragmentation, and the effects of institutional reforms.
- The dataset supports comparative research across countries using similar roll-call scaling methods, especially for those studying Latin American legislatures.

- Policymakers and analysts can use this resource to evaluate the cohesion and ideological clarity of coalitions and blocs in specific legislative contexts.
- It also serves as a foundation for future studies linking electoral dynamics, political representation, and legislative behavior in Chile.

2. Background

Between the legislative periods beginning in 2002 and 2022, the Chilean Chamber of Deputies experienced substantial changes in its composition and institutional structure. This period witnessed a growing diversification of political parties represented in Congress, shifts in the relative strength of party coalitions, and key electoral reforms—most notably, the 2015 reform that expanded the number of deputies from 120 to 155. These developments occurred alongside broader transformations in the national political landscape, including increased fragmentation, changes in electoral behavior, and the emergence of new political actors.

The dataset was compiled to provide systematic estimates of the ideological positions of legislators over this two-decade span, using roll-call votes on bills as the empirical foundation. The theoretical motivation lies in contributing to the study of legislative behavior and representation under conditions of political and institutional change. The methodological contribution of this dataset is the combination of static and dynamic estimation techniques that allow for both within-period and cross-period comparisons of ideological positions.

By focusing on a period marked by both institutional reforms and broader political shifts, the dataset offers a foundation for future research on party systems, representation, and polarization in comparative contexts.

3. Data Description

The repository contains two types of files focused on roll call data, and ideological estimates, respectively. In the first kind of files, there are six datasets with roll-call votes from plenary sessions on bills debated in the Chilean Chamber of Deputies. Each file is named following the format `matriz__periodo_20XX_YY.csv`, where XX and YY correspond to the start and end years of the legislative period. The periods covered are: 2002–2006, 2006–2010, 2010–2014, 2014–2018, 2018–2022, and 2022–2026. For the most recent period (2022–2026), data collection includes votes up to June 2025. See [Table 1](#).

Each file in [Table 1](#) has the same structure as described in [Table 2](#): rows represent individual legislators, while columns contain metadata and roll-call vote records. The first five columns provide identifiers and names for each legislator. Starting from column 6, each subsequent column corresponds to a unique roll-call vote associated with a legislative bill. Cell values indicate the voting behavior of each legislator for that bill (e.g., "Afirmativo", "En Contra", NA, etc.).

The raw roll-call datasets provided in the first folder are sufficiently structured and detailed to allow researchers to implement their own ideological estimation methods, depending on

Table 1
Structure of the roll-call vote dataset for each legislative period.

Period	File name	Deputies (rows)	Roll-call votes (columns – 5)
2002–2006	<code>matriz__periodo_2002_06.csv</code>	121	1950
2006–2010	<code>matriz__periodo_2006_10.csv</code>	123	1990
2010–2014	<code>matriz__periodo_2010_14.csv</code>	123	3082
2014–2018	<code>matriz__periodo_2014_18.csv</code>	121	3267
2018–2022	<code>matriz__periodo_2018_22.csv</code>	163	3520
2022–2026*	<code>matriz__periodo_2022_26.csv</code>	157	3393

* Note: Data for the 2022–2026 period includes roll-call votes up to June 2025.

Table 2
Structure of the roll-call vote dataset for each legislative period.

Column Range	Description
Column 1	Row index (sequential number for reference only)
Column 2	Unique ID assigned to each legislator
Column 3	First name of the legislator
Column 4	Paternal surname of the legislator
Column 5	Maternal surname of the legislator
Columns 6 onward	Individual roll-call votes, one column per vote ID (e.g., V12345)
Cell values	Vote choice: "Afirmativo" (yes), "En Contra" (no), NA (not applicable or absent), among other possible labels such as "Abstención" or "Pareado" (if applicable)

their specific research goals or methodological preferences. These data include complete voting records, and legislator identifiers, enabling replication or adaptation using alternative modeling strategies.

For convenience and comparative purposes, a second set of files includes pre-computed ideological estimates derived from the roll-call data using three commonly used methods, as described below:

- DW-NOMINATE, a frequentist approach that generates within-period static estimates implemented in R using the *wnominate* package [1].
- Bayesian ideal point estimation, implemented using the *pscl* package's *ideal()* function, also producing within-period estimates [2].
- Dynamic ideal point estimation, implemented with the *dynIRT()* function from the *emIRT* package, which allows for cross-period comparisons by linking legislators' positions over time [3].

Each method produces one-dimensional ideological scores for all legislators present in a given period, along with additional metadata such as standard errors or confidence intervals (where applicable). The outputs are available in both wide and long formats, facilitating different types of analysis—whether focused on individual periods or on ideological trajectories across legislatures.

In addition to ideological scores, the datasets include information on each legislator's gender, date of birth, and the political parties to which they were affiliated during each legislative period. This metadata allows for more detailed analyses of representation, party dynamics, and demographic patterns within the legislature.

The dynamic model includes scores even for legislators with non-consecutive terms; these cases were manually reviewed to remove interpolated estimates for periods when the legislator was not in office (see *Limitations* section for details).

The ideological estimates are provided in a long-format table, where each row corresponds to a legislator during a specific legislative period. The dataset consolidates the results from the three estimation methods—dynamic, Bayesian, and DW-NOMINATE—into a single structure. For each legislator-period combination, the table includes point estimates and confidence or credible intervals (when available), along with party affiliations and demographic metadata. [Table 3](#) describes the structure and content of this long-format file in detail.

The dataset is also provided in wide format, where each row corresponds to a unique legislator, and columns contain ideological estimates for each legislative period and estimation method. This format facilitates longitudinal analysis of individual trajectories, allowing users to compare how a legislator's ideological position changes over time. Each estimation method—DW-NOMINATE, Bayesian, and dynamic—has its own set of period-specific variables. [Table 4](#) summarizes the structure of the wide-format dataset.

To assess the relationship between the three estimation methods, [Fig. 1](#) displays pairwise scatterplots and correlation coefficients for the ideological scores produced by DW-NOMINATE, the Bayesian model, and the dynamic model. Each point represents a legislator-period com-

Table 3

Variable structure of the ideological estimates dataset (long format).

Column Name	Description
n	Row index (sequential number for internal reference only)
Id	Unique identifier assigned to each legislator
Name	Full name of the legislator
Period	Legislative period (e.g., "Periodo_1" corresponds to 2002–2006)
dim_IC_low	Lower bound of the 95 % confidence interval from the dynamic model estimate (dynIRT)
dim_ideology	Point estimate of the legislator's ideological position from the dynamic model (dynIRT)
dim_IC_high	Upper bound of the 95 % confidence interval from the dynamic model estimate
wnom_ideology	Static ideological estimate from the DW-NOMINATE model for that period
bay_ic_low	Lower bound of the 95 % credible interval from the Bayesian model (ideal function)
bay_ideology	Posterior mean of the legislator's ideological position from the Bayesian model
bay_ic_high	Upper bound of the 95 % credible interval from the Bayesian model
quant_parties	Number of different political parties to which the legislator was affiliated in that period
party_1	First political party affiliation in the given period
party_2	Second political party affiliation (if any) during the same period
party_3	Third political party affiliation (if any) during the same period
party_4	Fourth political party affiliation (if any) during the same period
Sex	Self-reported gender of the legislator ("Masculino" or "Femenino")
Birth_date	Legislator's date of birth in ISO format (YYYY-MM-DDTHH:MM:SS)

Table 4

Variable structure of the ideological estimates dataset (wide format).

Column Name (Prefix)	Description
Id	Unique identifier assigned to each legislator
Name	Full name of the legislator
Sex	Self-reported gender ("Masculino" or "Femenino")
Birth_date	Date of birth in ISO format (YYYY-MM-DDTHH:MM:SS)
party_Px	Party affiliation in period x ($x = 1$ to 6)
dim_Px	Dynamic estimate for period x (from dynIRT)
dim_low_Px / dim_high_Px	Lower and upper bounds of 95 % confidence interval for dynamic estimate in period x
wnom_Px	DW-NOMINATE estimate for period x
bay_Px	Bayesian posterior mean for period x
bay_low_Px / bay_high_Px	Lower and upper bounds of 95 % credible interval for Bayesian estimate in period x

ination. The figure shows the degree of alignment between the different methodological approaches.

Density plots of ideological estimates from the dynamic model (dynIRT) across six legislative periods. Each ridge represents the distribution of legislator scores in a given period. The horizontal axis corresponds to the estimated ideological position on a single dimension (Fig. 2).

For completeness, additional legislator-level metadata files are included as supplementary material in the Harvard Dataverse repository to allow for full replication. These consist of six CSV files—one for each legislative period—named using the format: `diputado_periodo_20XX_YY.csv`, where XX–YY corresponds to the years covered by each legislature. Each file contains demographic and party affiliation information for all legislators active in that period. Each file includes legislator ID and name, gender, date of birth, and party affiliations during each legislative period (as applicable). The same information in these files is already linked with ideological estimates in Tables 3 and 4.

3. Experimental Design, Materials and Methods

The dataset was constructed from publicly available roll-call vote records published on the official website of the Chilean Chamber of Deputies (www.camara.cl). These records were ac-

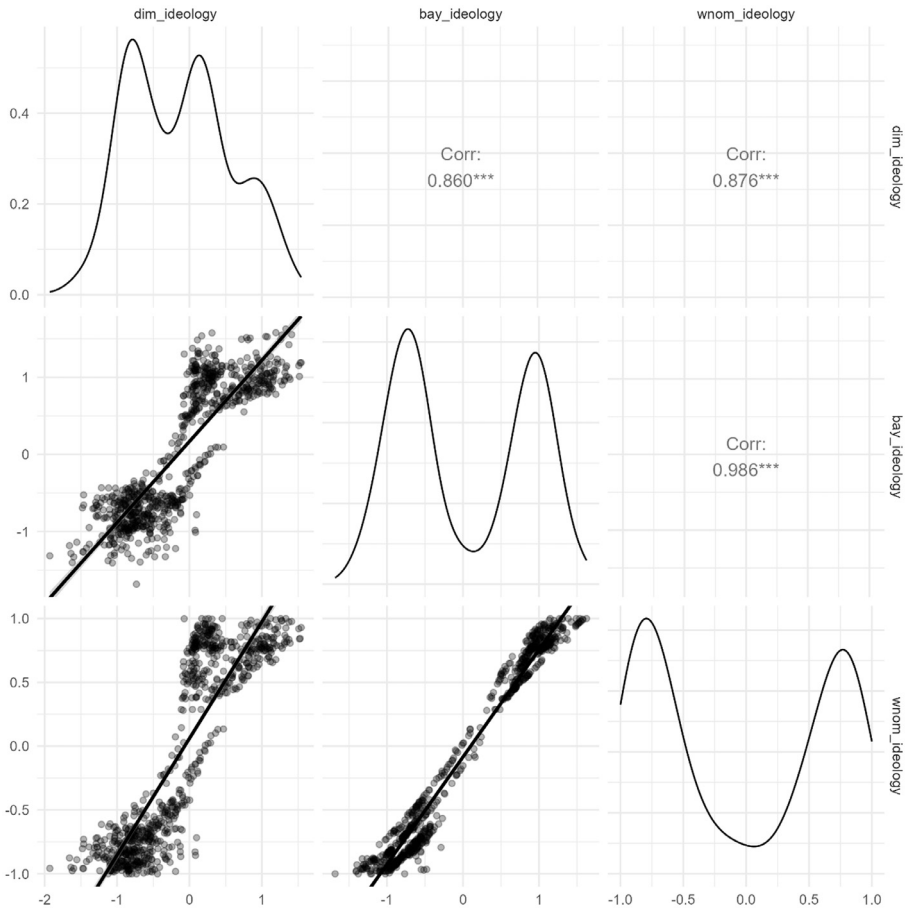


Fig. 1. Pairwise correlation of ideological estimates across methods.

cessed using a custom web scraping routine implemented in R (v4.3.0), relying on the rvest, httr, and xml2 packages to automate the download of vote results and associated metadata.

After collecting the raw data, we performed a filtering process to retain only votes corresponding to bills presented and voted on in plenary sessions. We excluded votes related to non-binding agreements, investigative commissions, and other types of parliamentary motions that do not directly affect lawmaking. The resulting dataset includes each legislator's vote (yes, no, abstain, or absent) for each qualifying bill, along with metadata such as bill ID, date, and title.

We then used three estimation procedures to derive one-dimensional ideological scores for each legislator:

- DW-NOMINATE, using the wnominate package, for static estimates within each legislative period [1].
- Bayesian ideal point estimation, using the pscl package and the ideal function, also computed separately for each period [2].
- Dynamic ideal point estimation, using the emIRT package and the dynIRT() function, to produce scores that are comparable across periods [3].

All processing, estimation, and data cleaning tasks were conducted in R, and outputs were exported in CSV format. Legislator-level metadata (such as full name, party affiliations throughout

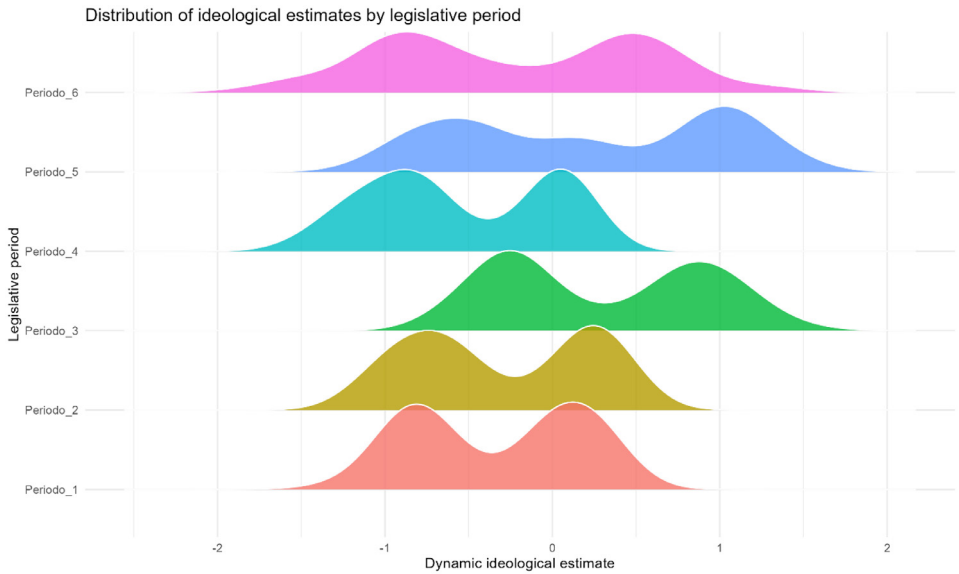


Fig. 2. Distribution of ideological estimates by legislative period (dynamic model).

the period, date of birth, and age at the start of each legislature) were integrated using supplementary public records from the Chamber's official site and cross-checked manually when inconsistencies appeared.

The raw vote data were obtained through a custom web scraping routine in R. Due to site-specific structure and maintenance concerns, the scraping code is not included, but all resulting raw data files are provided along with full code for data filtering, estimation, and final dataset construction. The vote records remain publicly accessible via www.camara.cl, and the full process from cleaned input to final outputs is fully reproducible using the files `A_Ideology_Estimation.R`, `B_Add_Variables.R`, and `C_Visualizations.R` located in the repository. Data and replication codes are available at: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/FOXOIT> [4]. You can also find the same information in my personal github account at https://github.com/jfabregalacoa/monitor_congreso/tree/master.

Limitations

During the data validation process, we identified inconsistencies in the assignment of legislator IDs across periods. Specifically, deputy Pablo Prieto Lorca appeared under two different IDs for separate legislative terms. After confirming via biographical data from the official congressional website, we unified both entries under a single ID.

In dynamic estimation using the `dynIRT()` function, a limitation arises when legislators serve non-consecutive terms. The algorithm assigns estimates for intermediate periods when a legislator served non-consecutive terms, interpreting absence as non-voting rather than non-membership. We manually removed such spurious estimates for multiple individuals to ensure consistency (e.g., Gabriel Ascencio in period 4, Carlos Kuschel in periods 2–4, among others).

Additionally, while the Chamber had 120 members from periods 1 to 4, the dataset includes >120 legislators per period due to mid-term replacements caused by deaths or appointments to executive positions. Similarly, for periods 5 and 6, the Chamber expanded to 155 members, yet the dataset includes 163 and 157 individuals, respectively, due to similar replacements.

These adjustments and manual corrections were necessary to preserve the accuracy and interpretability of the dataset.

Ethics Statement

The author confirms that he has read and comply with the ethical requirements for publication in *Data in Brief*. The current work does not involve human subjects, animal experiments, or any data collected from social media platforms.

CRediT Author Statement

Jorge Fábrega: Conceptualization, Data curation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing.

This is a sole author article.

Data Availability

[Ideological Estimates of the Chilean Chamber of Deputies, 2002–2026 \(Reference data\)](#) (Data-verse)

Acknowledgments

This research was partially funded by FONDECYT Regular Grant No 1231131, awarded by the National Agency for Research and Development of Chile (Agencia Nacional de Investigación y Desarrollo, ANID). No other contributors beyond the author meet the criteria for authorship.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] K.T. Poole, H. Rosenthal, *Congress: A Political-Economic History of Roll Call Voting*, Oxford University Press, New York, 1997.
- [2] J. Clinton, S. Jackman, D. Rivers, The statistical analysis of roll call data, *Am. Political Sci. Rev.* 98 (2) (2004) 355–370, doi:[10.1017/S0003055404001194](https://doi.org/10.1017/S0003055404001194).
- [3] K. Imai, J. Lo, J. Olmsted, Fast estimation of ideal points with massive data, *Am. Political Sci. Rev.* 110 (4) (2016) 631–656, doi:[10.1017/S0003055416000385](https://doi.org/10.1017/S0003055416000385).
- [4] Fábrega, Jorge, 2025, "Ideological Estimates of the Chilean Chamber of Deputies, 2002–2026", <https://doi.org/10.7910/DVN/FOXOIT>, Harvard Dataverse, V2, UNF:6:l8l6rXPOijdF5jQzFJ2a5w== [fileUNF]