Pedro Henrique Borges da Silva

Uma Análise de Redes do Senado Brasileiro Usando Dados de Votações e de Discursos

Brasília, Brasil

2024

Pedro Henrique Borges da Silva

A Network Analysis of Brazilian Senate Using Roll Call Votes and Speech Data

Brasília, Brasil

2024

Pedro Henrique Borges da Silva

A Network Analysis of Brazilian Senate Using Roll Call Votes and Speech Data

Dissertação apresentada ao Curso de Mestrado Acadêmico em Economia, Universidade de Brasília, como requisito parcial para a obtenção do título de Mestre em Economia

Universidade de Brasília - UnB

Faculdade de Administração Contabilidade e Economia - FACE

Departamento de Economia - ECO

Programa de Pós-Graduação

Orientador: Daniel Oliveira Cajueiro, Dr.

Brasília, Brasil

2024

Pedro Henrique Borges da Silva
A Network Analysis of Brazilian Senate Using Roll Call Votes and Speech Data/ Pedro Henrique Borges da Silva. – Brasília, Brasil, 2024-
58p. : il. ; 30 cm.

Orientador: Daniel Oliveira Cajueiro, Dr.

Dissertação de Mestrado – Universidade de Brasília - UnB
Faculdade de Administração Contabilidade e Economia - FACE
Departamento de Economia - ECO
Programa de Pós-Graduação, 2024.

1. Redes Complexas. 2. Discurso político 3. Senado brasileiro. II. Universidade de Brasília. III. Faculdade de Administração, Contabilidade e Economia - FACE. IV. Departamento de Economia IV. A Network Analysis of Brazilian Senate Using Roll Call Votes and Speech Data

Pedro Henrique Borges da Silva

A Network Analysis of Brazilian Senate Using Roll Call Votes and Speech Data

Dissertação apresentada ao Curso de Mestrado Acadêmico em Economia, Universidade de Brasília, como requisito parcial para a obtenção do título de Mestre em Economia

Trabalho aprovado. Brasília, Brasil, 2 de abril de 2024:

Daniel Oliveira Cajueiro, Dr.
Orientador

Bernardo Pinheiro Machado Mueller, Dr.
Convidado 1

Matheus Schmeling Costa, Dr.
Convidado 2

Brasília, Brasil
2024

Acknowledgements

I would like to express my heartfelt gratitude to my family and friends whose unwavering support, encouragement, and love sustained me throughout this journey. I am profoundly grateful for your understanding and encouragement, which have enabled me to pursue my academic aspirations. Thank you for being my pillars of strength and for standing by me every step of the way.

Resumo

Esta dissertação realiza uma análise de rede do Senado Brasileiro, analisando discursos e votações nominais da 51ª à 56ª legislaturas para explorar a conectividade intra e interpartidária. Os resultados indicam que, embora as redes de votação sejam altamente conectadas, as conexões são em sua maioria fracas, com uma força de ligação mediana em torno de 11% das votações contenciosos. A análise da rede de discursos revela um aumento na força de conexão entre os senadores até a 53ª legislatura, seguido por um declínio para os níveis iniciais da 56ª legislatura. O Partido dos Trabalhadores (PT) demonstra a maior coesão interna, tanto em critérios de discurso quanto de votação, com o PSB e o PSDB também mostrando conectividade interna significativa. O PT também emerge como o partido mais interconectado com outros partidos em ambas as análises de votação e discurso, sugerindo seu papel influente nas atividades parlamentares. O PSB e o PSDB também mostraram níveis significativos de conexão interpartidária. No entanto, comparar as redes de discurso e votação provou-se um desafio, revelando a necessidade de um exame mais profundo da sensibilidade das métricas de rede aos filtros de peso de conexão. Este estudo contribui para a compreensão da dinâmica parlamentar através de redes complexas, oferecendo um novo *framework* para analisar o comportamento legislativo e destacando a importância da análise de rede no estudo da atividade parlamentar.

Palavras-chave: Processamento de linguagem natural, discurso político, votações parlamentares, redes complexas

Abstract

This dissertation conducts a network analysis of the Brazilian Senate, analyzing speech and roll call votes from the 51st to the 56th legislatures to explore intra- and inter-party connectivity. Results indicate that while voting networks are highly connected, connections are mostly weak, with median strength around 11% for contentious votes. Speech network analysis reveals increasing connection strength among senators up to the 53rd legislature, then a decline to initial levels by the 56th. The Partido dos Trabalhadores (PT) demonstrates the highest internal cohesion, both at speech and voting criteria, with PSB and PSDB also showing significant connectivity. PT also emerges as the most interconnected party across both voting and speech analyses, suggesting its influential role in parliamentary activities. PSB and PSDB also showed significant levels of inter-party connectivity. However, comparing speech and voting networks proved challenging, revealing the need for deeper examination of network metric sensitivity to weight filters. This study contributes to understanding parliamentary dynamics through complex networks, offering a novel framework for analyzing legislative behavior and highlighting the importance of network analysis in parliamentary activity studies.

Keywords: Natural Language Processing, political speech, roll call votes, complex networks

List of Figures

List of Tables

Contents

1 Introduction

Representative or indirect democracies rely on the population's selection of representatives. Some reasons for voting for a candidate include their policy platforms, their past accomplishments, their leadership skills, and their party affiliation and also their positions on relevant issues revealed by speeches. The electorate expects their chosen candidates to follow through on their campaign promises and to be properly represented by their parliamentary activity. Although some studies have shown that politicians tend to maintain their political views over time (Poole, 2007), it remains unclear whether there is coherence between their discourse and action, represented by their votes in parliament. This issue is particularly relevant to the public, as it is of interest to determine whether legislators' speeches are consistent and coherent with their voting patterns.

There is extensive research on the estimation of the ideology of parties or legislators. The majority of studies in this field focus on estimating the ideal points of individual legislators on a ideology scale, relying on data obtained from either roll call votes or speech records. However, the differences between these estimates have not been thoroughly explored. There are several reasons to suspect that the estimates derived from these different sources of data are not equivalent, for example, there may exist external pressures that influence votes, beyond politicians' true beliefs (Power; Zucco, 2009). In this sense, measures based on roll call votes may reveal additional dimensions that can contaminate parliamentary voting, such as a strong executive-legislative dynamic. This phenomenon is especially pertinent in Brazilian politics, with its presidential system (Leoni, 2002) (Morgenstern, 2004).

This study aims to analyze parliamentary activity in the Brazilian Senate employing a methodology that can address some of the limitations inherent in ideological measures, thereby enriching the analysis of parliamentary dynamics. We introduce an approach grounded in network theory, enabling a new understanding of political positions of Brazilian legislators through the lens of complex networks. Our preference for the upper house stems from the considerable political power and influence of senators, who represent states with a smaller membership compared to the lower house [1]. This strong representation may influence their speech and voting patterns, thereby providing deeper insights into wide-ranging political issues. Additionally, the relatively fewer members in the Senate likely afford senators greater opportunities to elaborate on their ideas within their parliamentary speeches, potentially leading to a richer and more extensive discourse data.

To explore this, we introduce a network based on speech patterns, linking politicians

[1] Each state has only three senators whereas the number of deputies can reach up to 70 in the case of São Paulo, for instance

whose speeches exhibit similarity. This analysis is facilitated by employing Natural Language Processing (NLP) techniques to process the speeches delivered by Brazilian senators over the years. Concurrently, a network reflecting roll-call voting behavior in the Senate is constructed using roll call votings data, with connections formed between senators who demonstrate alignment in their votes on contentious bills—those prompting significant senatorial division. The intensity of these links is quantified by the frequency of their agreement on such divisive issues. Subsequently, we compare these networks using a distance metric within a network space, aiming to ascertain the congruence between the senators' voting behavior and their rhetorical alignments.

Through the adoption of this methodology, the research aims to reveal the underlying connections among Senate members as influenced by their speech and voting patterns. Furthermore, it seeks to examine the variances within these networks and their evolution over time.

2 Literature Review

Political speech analysis has been an active research field with a focus on measuring the ideology of political parties or individual legislators. While earlier studies on this task used survey data or voting records to analyze the ideology within a congress (Power; Zucco, 2009) (Poole; Rosenthal, 1985), more recent studies have increasingly relied on language-based data for analysis (Diermeier et al., 2012), (Laver et al., 2003) (Lauderdale; Herzog, 2016). Unlike roll call voting, which can be influenced by external pressures such as party cohesion or logrolling with the executive, language-based expressions are often considered to be less susceptible to such constraints, potentially allowing representatives to express their true opinions more freely.

However, the differences between estimates based on roll call voting and speech records have not been fully explored, and there has not been many studies that have sought to address this gap. For instance, Schwarz compared political positions estimated through roll call voting to those estimated through legislative speeches, using an example of an energy policy debate from the Swiss legislature in 2002-2003 (Schwarz et al., 2017). The authors found that text scaling based on legislative speeches reveals a greater heterogeneity in intra-party preferences compared to roll-call scaling, indicating that legislative speeches may offer a more nuanced view of political ideology within parties. Moreover, the study concluded that the differences between voting and speech vary systematically according to constituency-level electoral preferences, with legislators voting with their parties but speaking to their constituents.

In recent years, the application of text analysis to political speech has gained popularity in Brazil as well. Several studies have employed natural language processing techniques to analyze Brazilian political speeches, with early papers including (Moreira, 2020), (Arnold et al., 2017), and (Izumi; Medeiros, 2021). While these studies have yielded valuable insights into the political landscape of Brazil, there is still much to be learned about the differences between voting and speech-based measures of political ideology.

One promising approach to investigating this difference is through the use of complex networks. Nery and Mueller implemented networks of legislators based on co-sponsorship relations to investigate the importance of caucuses (also known in portuguese as *bancadas temáticas*) in Brazilian politics (Nery; Mueller, 2022). Cajueiro et al. demonstrated the efficacy of this approach by utilizing published news articles on companies to model similarities between them and construct a network that reflects the strength of the association between these companies. By utilizing natural language processing techniques to analyze news texts and generate a mathematical representation of the texts, the authors were able to quantify the similarity between companies and gain insights into their relationships (Cajueiro et al., 2021). This methodology holds potential for extension to political speech data, presenting a promising avenue for exploring the dynamics of connections among senators. Furthermore, it facilitates the integration of these connections with measures derived from voting patterns, offering an approach to analyzing parliamentary behavior. A method for making a comparison between networks is evaluating a distance between them in a network space. (Andrade et al., 2008) proposes a measure that is specially appropriate for comparing networks with the same number of nodes, which is the case in our study.

3 Methodology

In this section, we describe the methodology employed to examine the speeches and voting patterns of senators using the framework of complex networks. Our approach is structured into the following steps, for both speech and voting network: (1) Initially, we start with data collection and preprocessing, utilizing computational tools to access public records of speeches and votes in the Brazilian Senate. (2) The subsequent phase involves the construction of networks, in which we establish the criteria for connections within the senators' network and set parameters to yield a network that is meaningful for our analysis. (3) This is followed by an adjustment phase, aimed at aligning the speech and voting networks to facilitate comparison. (4) Finally, we introduce a method for measuring speech and voting networks distance in a network space.

3.1 Data collection and preprocessing

The speech and voting data utilized in our study are publicly accessible and were retrieved from the website of the Federal Senate of Brazil. Our analysis aims to examine a substantial volume of data, covering the period from 1999 to 2023. To this end, we employed web scraping techniques to systematically download and organize this information into a dataset tailored for our research, incorporating variables of significance.

3.1.1 Voting data preprocessing

The preprocessing of voting data involved two indispensable filterings to refine the dataset for analysis. First, we selected only open (roll call) votings, leaving out the secret votings to ensure transparency about each senator's vote. Subsequently, we filtered for the voting sessions that demonstrated a significant level of division among senators. This distinction is important, as many subjects that undergo voting achieve unanimity or near-unanimous consensus among legislators.

Indeed, despite the ideological diversity among Senate members, numerous matters submitted for parliamentary voting are likely to achieve substantial consensus due to the general widespread agreement on their importance. Examples include proposals related to disaster relief funding, honorary resolutions, healthcare and education initiatives, key infrastructure projects, among others. Conversely, some issues carry less weight in political-ideological discussions, involving the ceremonial duties tied to senators' roles, like the confirmation of senior officials and directors within major government agencies. These, too, are subjects less likely to provoke dissent within the parliament.

Therefore, these votes offer limited insight into the senators' viewpoints or the structure of the underlying network they organize themselves into that our study aims to capture. So we start with a conservative filter to keep in our dataset only those votings that the winning result did not exceed 70%.

To enable party-wise analysis, we introduced a variable named "Senator-Party". It consists on a combination of the tuple [*senator*, *party*]. It accounts for individuals who changed their party affiliation during a legislature. Following a party switch, these individuals are treated as distinct entities, recognizing that such a change may reflect a shift in the senator's ideology or in the party's ideological stance.

3.1.2 Speech data preprocessing

Initially, we exclude speeches delivered by individuals who are not members of the Senate, including deputies from the lower house and other participants. Subsequently, speeches, for any reason, did not had its content available in the Brazilian Senate website are also omitted. Then, we employ Natural Language Processing (NLP) techniques to prepare the raw text data containing senators' speeches for further analysis. The preprocessing steps are delineated as follows:

- **Removal of of Introductory Text and Interjections**: each senator's speech begins with an introductory text, typically indicating the speaker's name, party, and state. This intro also marks interjections from other senators or attendees within the speech. The initial step involves isolating the main speech from these interjections and discarding the introductory segments.
- **Removal of Titles of Address**: parliamentary discourse frequently utilizes titles such as "ex^a", "ex^o" "*sr*", "sr^a", "*presidente*", "*presidenta*". Given these titles do not contribute meaningfully to our analysis, they are removed.;
- **Punctuation Removal**: All punctuation marks are removed to streamline text processing;
- **Stopwords Removal**: in natural language analysis, *stopwords* are words commonly used but considered to carry little to no semantic weight for the purpose of understanding or analyzing text;
- **Removal of Words With Fewer than 3 Characters**.
- **Lowercasing**: to ensure consistency in the data and to avoid treating words as distinct simply because of case differences, all the letters in all words are converted to lowercase.

Finally, the dataset is organized by legislature to facilitate analysis both within and across legislative periods. Subsequently, speeches attributed to each "senator-party" combination are consolidated into a unified document.

3.2 Network Model

A network is a structured representation of a set of objects (often referred to as *nodes* or *vertices*) that are interconnected by links (*edges* or *connections*). Nodes typically represent entities such as individuals or organizations while edges symbolize the relationships or interactions between these entities. In weighted networks edges are assigned a numerical value or weight that reflects the strength, capacity, or significance of the connection between nodes. In this study, aiming to understand the similarities among senators based on speech and voting patterns, we employ an analysis based on network theory where each node represents a unique "senator-party" combination within a given legislature. The connections between these nodes and their strength are determined by the similarity in their voting or speech patterns

We divide the data set by legislature, as our analysis will be within a legislature. So, in the method described below, we are taking into consideration only the speeches and votes within a given legislature.

3.2.1 Speech network model

Borrowing some concepts of natural language processing, and to implement a vector space model for our analysis, let term w_i be a word or group of consecutive words identified by the unique index i. From now on, we are going to call it simply a *term*. The vocabulary $\mathcal{V} = \{w_1, \ldots, w_i, \ldots, w_{N_V}\}$ is the set of all distinct terms present in all speeches of the legislature and $I_V = \{1, \ldots, N_V\}$ is the set of all term indexes. The speeches $s_j = [w_{i_1}, \ldots, w_{i_k}, \ldots, w_{i_{L_j}}]$ consist of a list of L_j non-unique consecutive terms ($1 \leq k \leq L_j$ and $i_k \in I_V$), while $\mathcal{V}^{s_j}$ is the vocabulary that appears in the speech s_j. Finally, $\mathcal{S} = \{s_1, \ldots, s_j, \ldots, s_{N_S}\}$ is the set of all speeches.

Let $\mathcal{C}$ be the set of all senators and N_C the number of elements of this set. Since our objective is to identify the similarity between two senators' discourses, instead of measuring the properties of the individual speeches, we are interested in measuring the properties of s^k, the concatenation of all speeches given by a senator k for $k \in \mathcal{C}$. Based on this definition, let us also define $\mathcal{S}^C$ as the set of all s^k. Therefore, $\mathcal{S}^C$ is a set with N_C (concatenated) speeches, where each speech is associated with its senator k and each s^k is the concatenation of all speeches of $\mathcal{S}$ given by senator k.

The term-speech matrix $\mathbf{M}$ is a $N_V \times N_C$ matrix that presents the frequency of terms that occur in the collection of speeches of each senator. In a term-speech matrix, rows correspond to terms and columns correspond to the speeches s^k for $k \in \mathcal{C}$:

$$\begin{array}{c} \\ w_1 \\ w_2 \\ \vdots \\ w_{N_V} \end{array} \begin{array}{c} \begin{array}{cccc} s^1 & s^2 & & s^{N_C} \end{array} \\ \begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1N_C} \\ n_{21} & n_{22} & \cdots & n_{2N_C} \\ \vdots & \vdots & \cdots & \vdots \\ n_{N_V 1} & n_{N_V 2} & \cdots & n_{N_V N_C} \end{bmatrix} \end{array}, \tag{3.1}$$

where n_{ik} counts the number of times term i arises in the collection of speeches of senator k, for $k \in \mathcal{C}$.

We can represent the interactions between the senators by a network, where the nodes represent the senators. The connection between two senators k and l depends on the similarity of their discourses, which depends explicitly on the words that appear in the speech of both senators. In this network, let $\mathcal{N}_k$ be the set of neighbors of k, i.e., $l \in \mathcal{N}_k$ if there is a link between k and l.

One simple way to compress the information about the words that appear simultaneously in speeches of two different senators $k, l \in \mathcal{C}$ is given by the function

$$q_{k,l} = \cos(\theta_{k,l}) = \frac{\sum_{i=1}^{N_v} f^k(w_i) f^l(w_i)}{||f^k|| \; ||f^l||}, \tag{3.2}$$

where $\theta_{k,l}$ is the angle between the vectors f^k and f^l, for $k, l \in \mathcal{C}$, and $f^c(\cdot)$ represents a measure that accounts the importance of each word $w_i \in \mathcal{V}$ in s^c. There are several different forms for defining f^c, for c in $\mathcal{C}$, in the field of natural language processing. Here we are interested in building a network that matches the perception of people about the similarity between two senators' discourse . A recent work (Pincombe, 2004) has shown that the Entropy Model is the one that provides a similarity evaluation closer to the one provided by actual people in an experiment that compares the similarities generated by an algorithm to the ones evaluated by people. The Entropy model accounts for the importance of the word i in speech s^k, for $k \in \mathcal{C}$, by the normalized

$$f^k(w_i) = \omega_{\text{local}}(i, k) \times \omega_{\text{global}}(i) \tag{3.3}$$

where

$$\omega_{\text{local}}(i, k) = \log_2(n_{ik} + 1) \tag{3.4}$$

and

$$\omega_{\text{global}}(i) = 1 + \frac{\sum_{k=1}^{N_C} p_{ik} \log_2 p_{ik}}{1 + \log_2 N_C}, \tag{3.5}$$

with $p_{ik} = \frac{n_{ik}}{\sum_{k=1}^{N_C} n_{ik}}$ and the definition of $0 \log_2(0) = 0$ that is consistent with the $\lim_{t\to 0} t \log_2 t = 0$.

The local weight ω_{local} measures the importance of a word inside a speech. The larger ω_{local}, the larger is the importance of a word, since it is basically a measure of a frequency of a word inside a speech. On the other hand, the global weight ω_{global} measures the importance of a word in a speech when compared to its presence in all other speeches of the sample. A word that arises simultaneously in several speeches is less important than a word that arises specifically in few speeches. In NLP framework, ω_{local} and ω_{global} are often defined as functions of the term-frequency and inverse-document-frequency, respectively. We then apply Eq. (3.2) to the speeches of each pair of senators and use this result to define the strength of the link between them.

3.2.2 Voting network model

To construct a network from the roll call votes of senators, we establish a connection between them based on the frequency of their joint votes on selected sessions (votes with a maximum winning margin of 70%) debated in parliament. Thus, senators who consistently vote in agreement on these issues demonstrate a stronger linkage within the network, indicating closer alignment in their parliamentary voting behavior. So, we propose the following measure:

$$q_{k,l} = \sum_{j=1}^{N_b} \left| \frac{v_{k,j} + v_{l,j}}{2} \right|, \tag{3.6}$$

where $q_{k,l}$ is the strength of the link between senators k and l, N_b is the number of selected votes in the legislature and $v_{k,j}$ is the vote of senator k on bill j ($v_{k,j} = 1$ for a "yes" vote and $v_{k,j} = -1$ for a "no" vote).

3.3 Adjustment

Due to the frequent occurrence of party affiliation changes among senators, it is expected that the speech and voting networks for a specific legislature might not exhibit an identical number of nodes. This discrepancy often arises when a senator's tenure within a particular party, or as an independent, is too brief to result in any recorded votes or speeches during that timeframe. Furthermore, the exclusion of votes characterized by high consensus can significantly reduce the volume of votes, potentially nullifying the connections for some individuals, especially those not engaged in voting on contentious issues. Additionally, certain senators may abstain from participating in debates, leading to an absence of recorded speeches. Consequently, prior to proceeding to the phase of distance evaluation, it is crucial to align the networks by ensuring a uniform number of nodes, thereby rendering them compatible for comparative analysis.

3.4 Intra-party and inter-party connectivity analysis

Parties represent the principal organizational structure within parliamentary systems. This study evaluates the connectivity both within these groups and with members outside these groups, utilizing both the speech and voting networks for analysis. To conduct this analysis, we construct unweighted versions of the networks selecting the top 12,5% strongest connections, thereby capturing the most substantive links between senators.

The intra-party connectivity, denoted as ρ_{intra}, is measured by the network density within a party. This is calculated by considering only the nodes representing senators from the same party and their mutual connections:

$$\rho_{\text{intra}} = \frac{\text{Number of intra-party connections}}{\text{Number of possible intra-party connections}}$$

where the number of possible intra-party connections is given by $\frac{n(n-1)}{2}$, with n representing the total number of senators affiliated with the party. Conversely, the inter-party connectivity, denoted as ρ_{inter}, is assessed through a similar density calculation but extends the analysis to include first-order neighbors, that is, directly connected nodes outside the party:

$$\rho_{\text{inter}} = \frac{\text{Number of connections to senators from other parties}}{\text{Number of possible external connections}}$$

Here, the denominator is calculated as $n(N-n)$, where n is the number of senators within the party under consideration, and N is the total number of nodes in the network.

3.5 Comparing networks

For making a comparison between the networks found in the steps above we first construct unweighted versions of the networks isolating the top 12,5%, 25,0% and 37,5% strongest connections. Then we propose the evaluation of *Euclidean like distance* $\delta(\alpha, \beta)$ of (Andrade et al., 2008). It provides a method of evaluating a distance between non-isomorphic networks. Let $M_\alpha(g)$ be the adjacency matrix for order g neighborhood of a network α, i.e.,

$$M_\alpha(g)_{i,j} = \begin{cases} 1, \text{ if } d_{i,j} = g \\ 0, \text{ otherwise} \end{cases}, \tag{3.7}$$

where $d_{i,j}$ is the shortest path between nodes i and j and. Then, the neighborhood matrix is

$$\hat{\mathbf{M}}_\alpha = \sum_{g=0}^{D} g M_\alpha(g), \tag{3.8}$$

where D is the network diameter. $\hat{\mathbf{M}}_\alpha$ carries all information on the shortest path between any two nodes i and j. The proposed distance measure is achieved by evaluating $\delta^2(\alpha, \beta)$ with the following equation:

$$\delta^2(\alpha, \beta) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[\frac{(\hat{M}_\alpha)_{i,j}}{D_\alpha} - \frac{(\hat{M}_\beta)_{i,j}}{D_\beta} \right]^2. \tag{3.9}$$

Equation 3.9 scales all terms in the sum to the $[0, 1]$ interval. Appropriately, according to this distance definition, networks α and β can only be identical (isomorphic) if $\delta(\alpha, \beta)$ equals zero. Furthermore, the definition requires that the networks have exactly the same number of nodes N, which renders it suitable for our study's objective. Indeed, within a legislature, both networks from roll call voting data and from speech records, after proper adjustment processing, will have the same number of nodes N. Also the node numbering is fixed in our case because we already know who are the senators, so we can pair them in the same positions in the neighborhood matrices. Therefore, the distance evaluation is very straightforward, not needing any algorithm to find better numbering to minimize the distance.

4 Data

To understand how parliamentary dynamics evolved over time, we collect data from February 1999 to January 2023 in the Brazilian Senate, which comprises the legislatures 51st to 56th. The speech and voting data are public and were obtained at the webpage of the Federal Senate of Brazil [1]

4.1 Voting network data

The dataset comprises 124.924 votes cast by senators across 1.544 voting sessions during the 51st to 56th legislatures. It includes not only the names of the senators and their parties, but also the codes for the voting sessions and the votes themselves (indicated as "Yes" or "No"), along with details on the subjects of the votes. Table 1 reports summary statistics for the votes dataset.

Table 1 – Summary statistics for initial votes dataset: count of variables

Legislature	Period	Votes	Sessions	Senator-Party
51	1999 - 2003	22.564	279	143
52	2003 - 2007	18.142	224	136
53	2007 - 2011	14.407	178	190
54	2011 - 2015	16.846	208	136
55	2015 - 2019	20.488	253	172
56	2019 - 2023	32.477	402	173
Total		124.924	1.544	

4.2 Speech network data

The speech dataset contains 85,585 speeches delivered by Brazilian senators. This dataset not only records the senators' names, their party affiliations, and the speeches themselves but also provides summaries of the speech content. Similar to the approach taken with the voting dataset, we introduced a "Senator-Party" variable. Summary statistics for the speech dataset are presented in Table 2.

[1] All the data on senators, including their speeches and votes, was available at the following URL:<https://www12.senado.leg.br/dados-abertos>

Table 2 – Summary statistics for speech dataset: count of variables

Legislature	Period	Speeches	Senator-Party
51	1999 - 2003	7.404	143
52	2003 - 2007	18.859	136
53	2007 - 2011	17.219	190
54	2011 - 2015	17.347	136
55	2015 - 2019	12.147	172
56	2019 - 2023	12.609	173
Total		85.585	

5 Results

5.1 Voting Network

The initial phase of constructing the voting network involved selecting votes that could provide insights into the senators' organization within an underlying network based on voting patterns. Consequently, by applying a filter to include only those votes with a maximum winning margin of 70%, we retained 18,26% of the voting sessions in our dataset. This indicates that a vast majority of roll call votes in the Brazilian Senate result in a high consensus among senators. Therefore, these votes have been omitted from the dataset. Table 3 report some statistics on the resulting dataset after filtering.

Table 3 – Summary statistics votes dataset after filtering out votes with high consensus (>70% for winning margin): count of variables

Legislature	**Votes**	**Sessions**	**Senator-Party**
51	3.554	44	140
52	3.402	42	130
53	2.508	31	129
54	2.754	34	144
55	3.644	45	135
56	6.927	86	151
Total	22.789	282	

Henceforth, Equation 3.6, yielded voting networks for 51st to 56th legislatures with properties exhibited in Table 4. N denotes the number of nodes, reflecting the total count of unique "senator-party" combinations within each legislature. Density ρ indicates the proportion of actual connections relative to the possible connections, providing insight into the network's overall connectivity. Maximum weight q_{max} is defined as the highest possible weight for a connection. It equals the number of selected votes within the legislature. Q1, Q2 and Q3 represent the first, second (median), and third quartiles in a distribution, respectively. Max d and Max q are the maximum degree and weight observed in the network, respectively.

These quartiles help characterize the weight distribution, outlining the variance in linkage strength throughout the network. Meanwhile, the degree distribution illustrates the extent of a senator's connections with peers within the timeframe. These metrics collectively offer an overview of the network's structure and the distribution of connection strengths and node connectivity across different legislative periods.

Table 4 – Summary statistics for baseline voting network. N is the number of nodes; ρ is network density; q_{max} is maximum possible weight for an edge; Q1, Q2, and Q3 are quartiles thresholds of a distribution; Max d is maximum degree in the network; Max q is maximum observed weight in the network.

				Weight dist.				**Degree dist.**			
Leg.	**N**	ρ	q_{max}	**Q1**	**Q2**	**Q3**	**Max** q	**Q1**	**Q2**	**Q3**	**Max** d
51	128	0,728	44	2	5	11	39	74	99,5	116	122
52	126	0,706	42	2	5	13	37	66,75	99,5	110	119
53	125	0,677	31	2	3	8	28	61	93	107	120
54	115	0,66	34	2	4	8	30	61,5	86	95,5	107
55	113	0,759	45	3	6	12	40	76	88	103	109
56	122	0,689	86	3	10	24	76	67	90	108	117

Connectivity Analysis

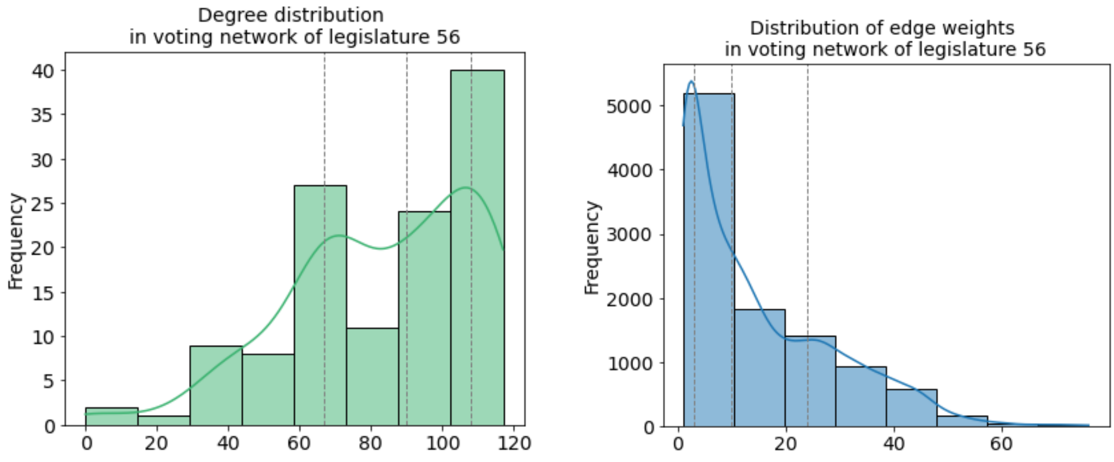
The density values, which range from 0.66 to 0.759, signify a high level of connectivity within each network. A density near 1 would indicate a fully connected network, so values in this range suggest that a significant proportion of all possible connections among senators (nodes) are realized. The 55th legislature exhibits the highest density (0.759), indicating the most interconnected network among those studied. In fact, the median node degree across legislatures range from 86 to 99,5, meaning that senators vote at least once with the great majority of its peers in controversial subjects. Given that we are working with a weighted network, the highly connected nodes do not bring very significant insights about the Senators activity if not analyzed concurrently with the distribution of weights among the connections.

Quartile Analysis

The quartile values of weight and degree distributions in Table 4 provide an overview of the variation and median tendencies in connection strength and node connectivity. While the networks demonstrate significant connectivity, it emerges that the majority of connections are of minimal weight. The analysis of degree and weight distribution plots across all legislatures reveals a right-skewed degree distribution, signifying a densely connected network, alongside a left-skewed weight distribution, indicating a network characterized by a limited number of strong connections.

Figure 1 shows the distributions for 56th legislature, which is very similar to the other legislative periods. The complete set of plots are in Appendix A. In general, in every legislature half of the senators have around 80 connections, and the majority (75%) connects with at least 60 of its peers. Conversely, 75% of these connections don't go further than 15 votes. On the other hand, in some legislatures the top 25% strongest connections can reach up to 60 votes, as in 56th.

This overall profile of connectivity must be interpreted taking into consideration the q_{max}, the maximum possible votes two senators can have together. Table 4 show that q_{max} varied from 31 to 86. So, Figure 2 show how the weight distributions quartiles as percetage of q_{max} evolved in time. The 1st, 2nd and 3rd quartile thresholds did not vary much across the period of analysis. The median of the distributions oscillated around 11% of q_{max}, and the first quartile remained close to 5%. The third quartile threshold had the biggest variance, ranging from 23,5% in the 54th legislature, to 31,0% in the 52nd.



(a) (b)

Figure 1 – Degree and edge weights distribution for baseline voting networks for 56th legislature. Dashed lines indicate quartile thresholds

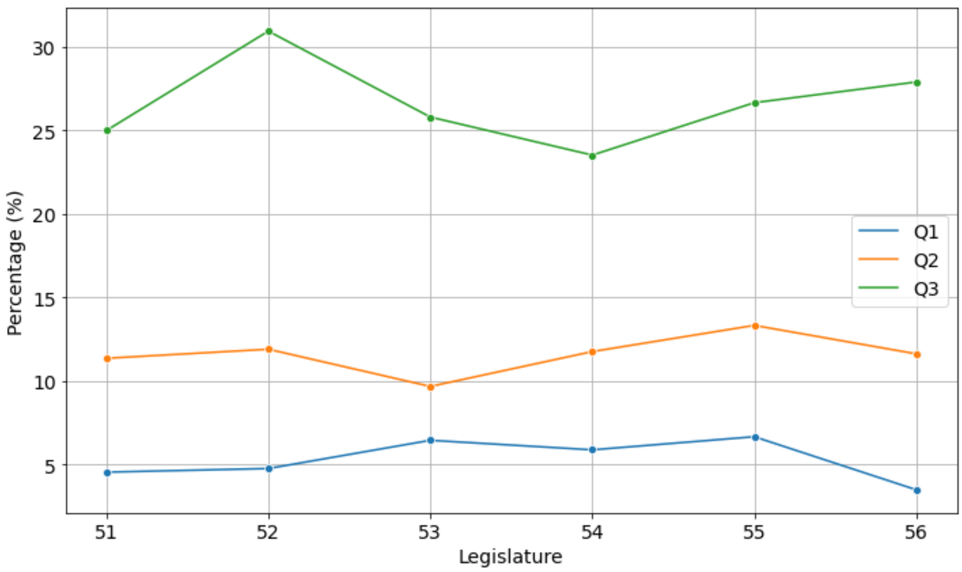


Figure 2 – Quartile thresholds of weight distributions as percentage of q_{max} of the legislature in voting networks

5.2 Speech network

The baseline speech network is a connected network, given that the edges weights are a cosine similarity evaluation i.e., it attributes a value in the interval [0,1] to the link between senators. Table 5 brings the summary of speech networks. N denotes the number of nodes, reflecting the total count of unique "senator-party" combinations within each legislature. Density ρ indicates the proportion of actual connections relative to the possible connections. For these baseline networks, density will always be 1. Maximum weight q_{max} is defined as the highest possible weight for a connection. Q1, Q2 and Q3 represent the first, second (median), and third quartiles in a distribution, respectively. Max d and Max q are the maximum degree and weight observed in the network, respectively.

Table 5 – Summary statistics for baseline speech network. N is the number of nodes; ρ is network density; q_{max} is maximum possible weight for an edge; Q1, Q2, and Q3 are quartiles thresholds of a distribution; Max d is maximum degree in the network; Max q is maximum observed weight in the network.

				Degree dist.				Weight dist.			
Leg.	**N**	ρ	q_{max}	**Q1**	**Q2**	**Q3**	**Max** q	Q1	Q2	Q3	Max d
51	128	1	1	0,173	0,228	0,291	0,529	127	127	127	127
52	126	1	1	0,229	0,304	0,371	0,538	125	125	125	125
53	125	1	1	0,283	0,370	0,438	0,730	124	124	124	124
54	115	1	1	0,246	0,339	0,424	0,608	114	114	114	114
55	113	1	1	0,213	0,284	0,361	0,634	112	112	112	112
56	122	1	1	0,162	0,242	0,311	0,527	121	121	121	121

Figure 3 presents the evolution of the quartiles thresholds for the speech networks across legislatures. In a initial phase, similarity between senators discourse grew from 51st and 53rd legislatures, yielding networks with stronger bonds between senators. After reaching a peak in 53rd legislature, the weights started falling, reaching, in legislature 56th a level similar from that observed in the 51st. For a more detailed look in the distributions, Appendix B contains the complete set of plots.

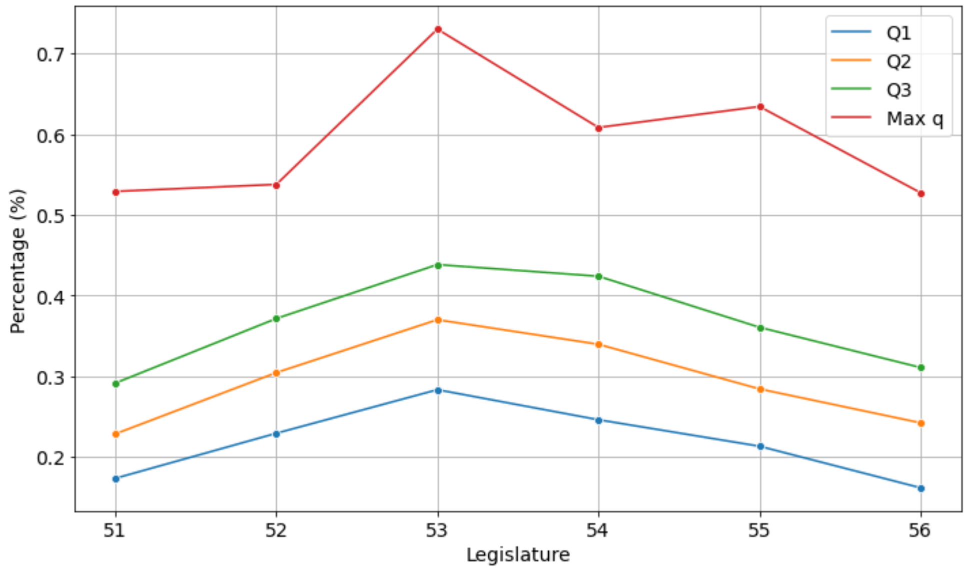


Figure 3 – Quartile thresholds of weight distributions in speech networks

5.3 Intra-party and inter-party connectivity analysis

The intra and inter-party connectivity assessment methodology starts from an unweighted network configuration. However, our baseline speech and voting networks are initially weighted and exhibit significant connectivity. To align them with our methodology's requirements, it becomes essential to derive their unweighted counterparts. The analysis of distributions, conducted earlier, facilitates the identification of parameters necessary for generating these new networks.

For the speech network, our examination revealed a predominantly left-skewed distribution, characterized by a predominance of weak connections across all legislatures. Consequently, our objective is to retain only the most substantial connections in the unweighted network. Such an approach ensures that the presence of an edge between two nodes distinctly signifies a cohesive voting pattern. To this end, we have developed variants of the voting and speech networks, where edges are selectively preserved based on their strength, specifically targeting the heaviest 12.5% connections. The complete set of tables containing the results of this analysis are displayed in Appendix E.

Voting Networks

The examination of intra-party connectivity within the voting network highlighted the Partido dos Trabalhadores (PT) as exhibiting the highest degree of internal interconnectedness among its members, consistently achieving top rankings across all examined legislatures. Notably, in the 52nd, 55th, and 56th legislatures, PT demonstrated the highest intra-party connectivity measure, ρ_{intra}, for these periods. Other parties, such as PSB and PSDB, also displayed significant internal cohesion throughout the legislatures, as evidenced by voting patterns.

Conversely, the analysis of inter-party connectivity identified PT as persistently the most influential party throughout the legislative terms, consistently attaining top rankings of ρ_{inter}. Other parties that emerged as noteworthy in this context encompass PSDB, PMDB/MDB, PSB, and PFL.

Speech Networks

The analysis of intra-party connectivity within speech networks revealed an even greater alignment among Partido dos Trabalhadores (PT) members, based on their speech patterns. This reflects high similarity values in their connections, positioning PT at the forefront rankings across all legislatures examined. Similarly, PSDB and PSB were other parties that achieved high rankings for this metric across the legislative periods.

In terms of inter-party connectivity, PT was also identified as highly interconnected with members of other parties, establishing itself as an influential entity not only in parliamentary voting activities but also in speech. This suggests PT's potential role as a leader in guiding thematic and ideological debates within the parliament. Other parties that demonstrated significant connectivity included PSB, PSDB, and PCdoB.

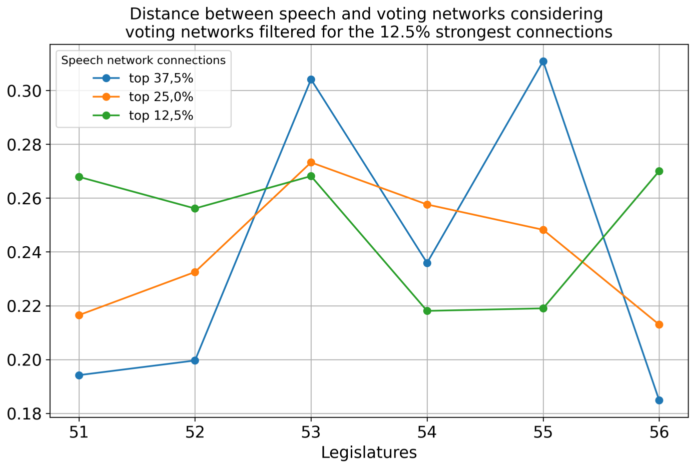
5.4 Network Comparison

To construct unweighted versions of our baseline networks, analogous to the approach outlined in the preceding section, we have formulated three variants for both the speech and voting networks. In these variants, edges are selectively retained according to their strength, focusing specifically on the most substantial 12.5%, 25.0%, and 37.5% of connections. Figure 4 illustrates the temporal progression of distance measures as influenced by these network configurations.

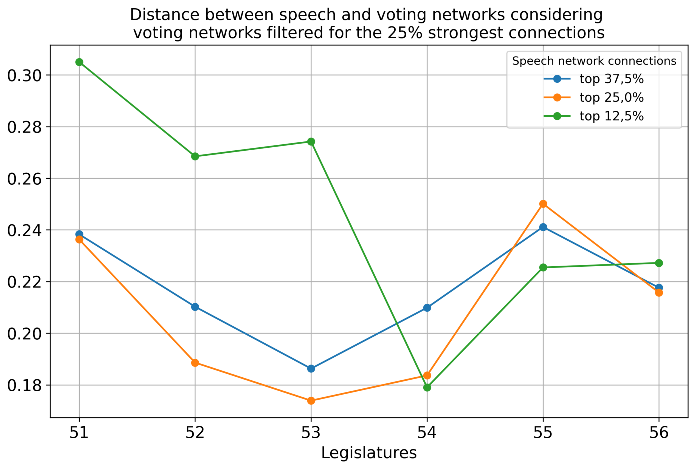
The application of diverse filters to both speech and voting networks results in significantly varied, and occasionally contradictory, tendencies. Such variability stems from the distinct connections that emerge when various dynamics are excluded through the filtering process. Specifically, in speech networks, certain connections might materialize due to the coincidental recurrence of common words across the speeches of different senators, rather than reflecting genuine thematic or ideological alignment.

To elucidate the reasons behind the marked sensitivity of distance measures to the applied filters, a comprehensive qualitative analysis of both speech and voting networks is imperative. This examination aims to decipher the underlying factors contributing to the pronounced fluctuations in distance measures observed across different filter configurations. Through this analytical approach, insights into the specific attributes of speech and voting patterns, which render them susceptible to the influence of filtering, can be gained, providing a deeper understanding of the networks' dynamics.

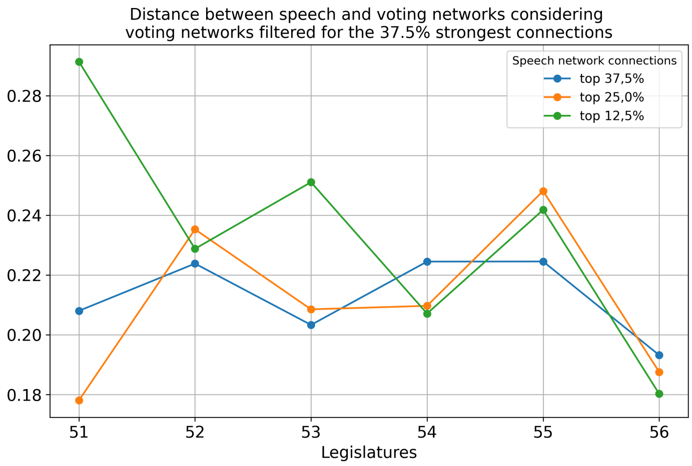
Appendix C presents a complete set of distance measures of the networks, refined through the application of octile filters to both speech and voting connections. Notably, discernible patterns in the speech-voting network distances begin to emerge when the analysis filters for the strongest half of the voting connections.



(a)



(b)



(c)

Figure 4 – Distance measures between speech and voting networks with voting edges filtered for the (a) 12.5%, (b) 25,0% and (c) 37,5% strongest connections.

6 Summary and conclusion

Our the primary objective was to conduct a network analysis of the Brazilian Senate, utilizing data from speeches and roll call votes across the 51st to the 56th legislatures. This research employed an approach that examines connectivity both within and between parties and measuring the distance between speech and voting networks for each legislature.

The results highlight that while voting networks are highly connected, they predominantly feature weak connections. The median connection strength in voting networks hovered around 11% of contentious votes, with the strongest 25% of connections involving joint voting on contentious issues varying from 23.5% to 31%. Furthermore, the speech network analysis showed a strengthening of connections between senators, particularly between the 51st and 53rd legislatures, indicating periods of increased rhetorical alignment. However, this trend reversed from 53rd legislature onward, so that by the 56th legislature it returned to levels of connectivity observed at the beginning of the study period.

Intra-party network analyses revealed the Partido dos Trabalhadores (PT) as exhibiting the highest internal cohesion in both voting and speech criteria, consistently ranking high across all legislatures. The PSB and PSDB also emerged as noteworthy for their internal connectivity.

Looking at inter-party network dynamics, PT also emerges as the most well-connected party across legislatures, being established as an influential entity not only in parliamentary voting activities but also in speech. PSB, PSDB also demonstrated significant connectivity with other parties members.

However, the attempt to compare speech and voting networks did not yield conclusive results, pointing to the need for a more detailed exploration to understand the sensitivity of network metrics to applied weight filters. This aspect underscores the complexity of translating quantitative network measures into qualitative insights about political behavior.

This dissertation contributes to the understanding of parliamentary dynamics through the lens of complex networks, offering a innovative methodological framework for analyzing the web of interactions that define legislative behavior. By unraveling the patterns of connectivity and division within and between parties, this work underscores the value of network analysis in capturing the subtleties of political discourse and alliance formation. As political landscapes evolve, such methodologies become increasingly crucial in providing insights into the underlying mechanisms of parliamentary functioning, offering a foundation for future research aimed at understanding the dynamics of political institutions and their impact on governance and policy-making.

A Distribution Plots for Voting Networks

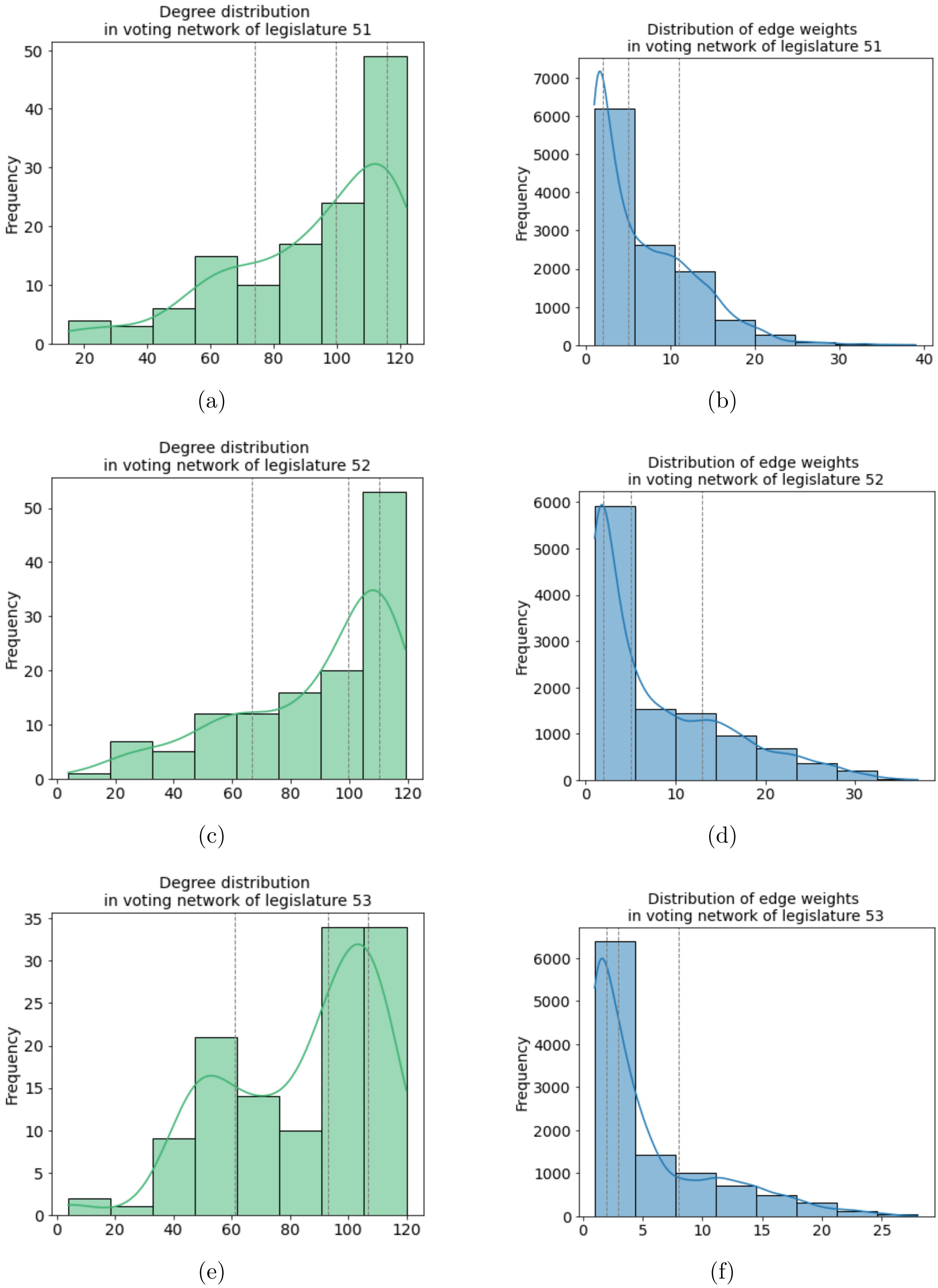
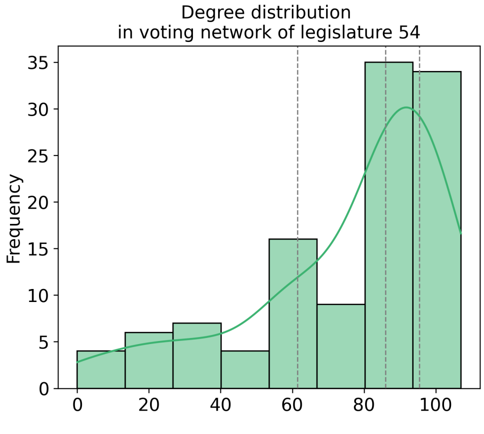
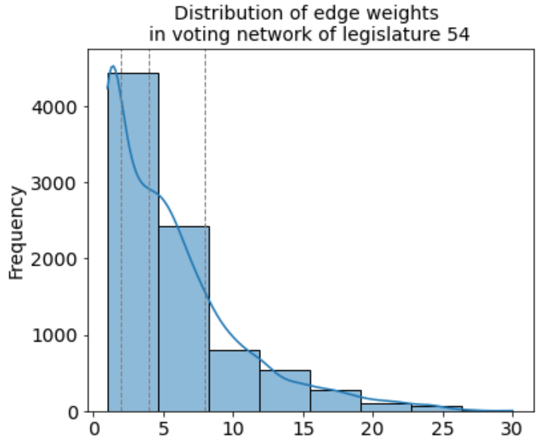


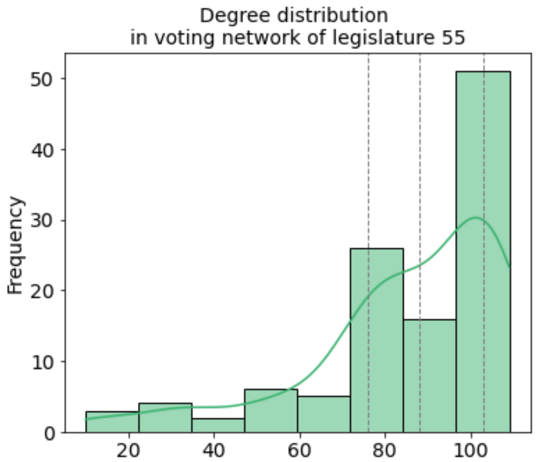
Figure 5 – Degree and edge weights distribution for baseline voting networks across legislatures 51st to 53rd. Dashed lines indicate quartile thresholds



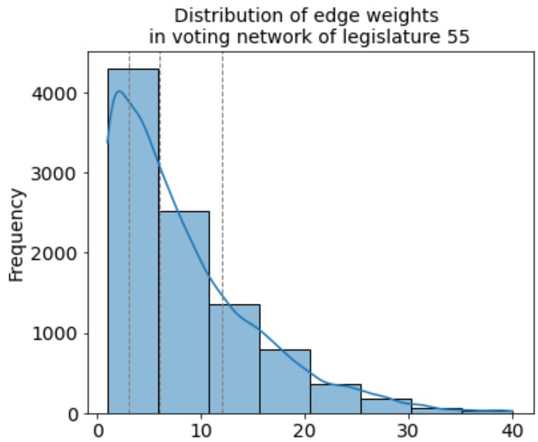
(a)



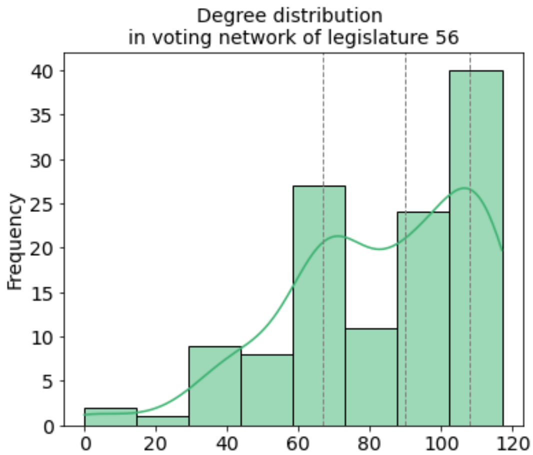
(b)



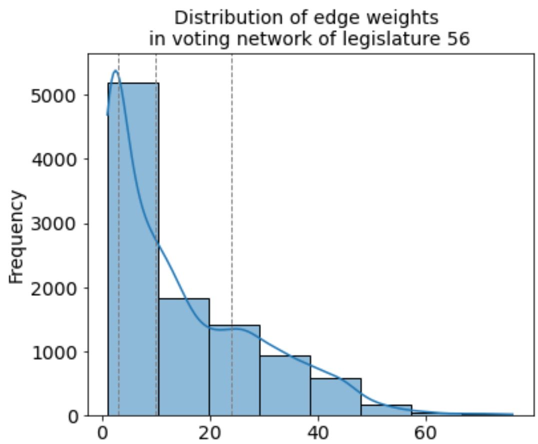
(c)



(d)



(e)



(f)

Figure 6 – Degree and edge weights distribution for baseline voting networks across legislatures 54th to 56th. Dashed lines indicate quartile thresholds

B Distribution plots for Speech Networks

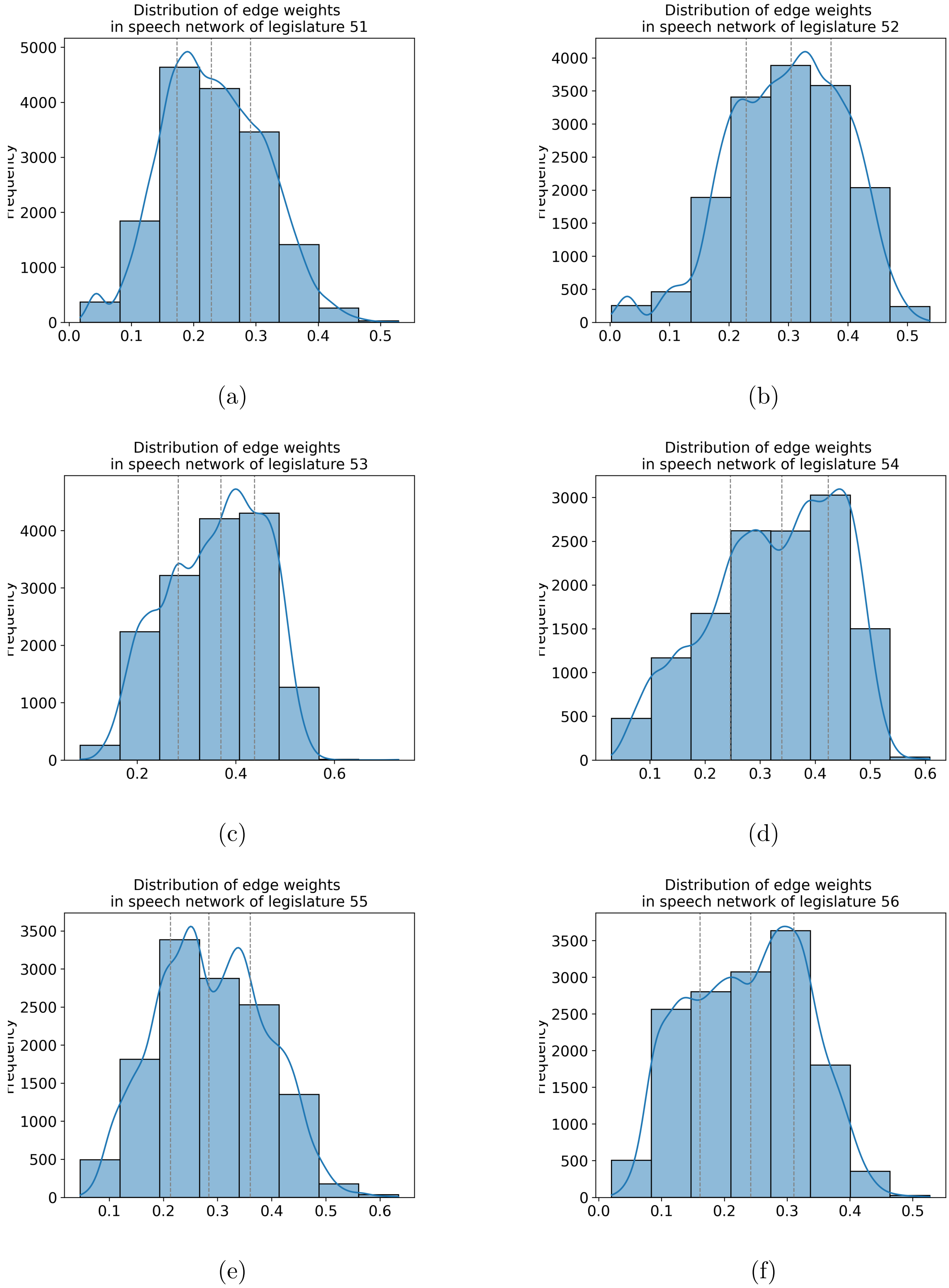
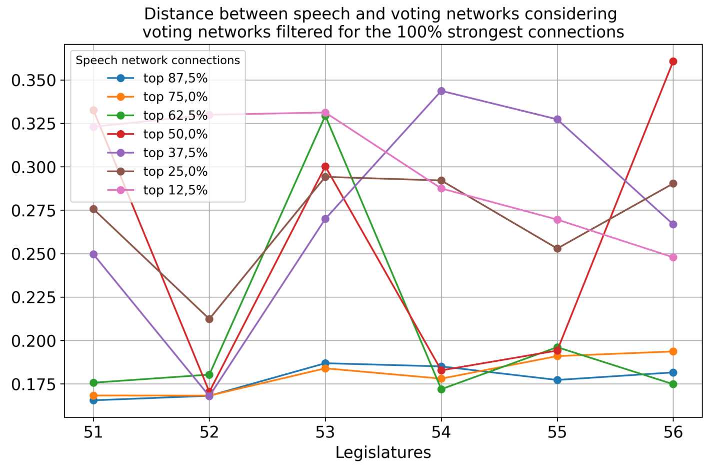
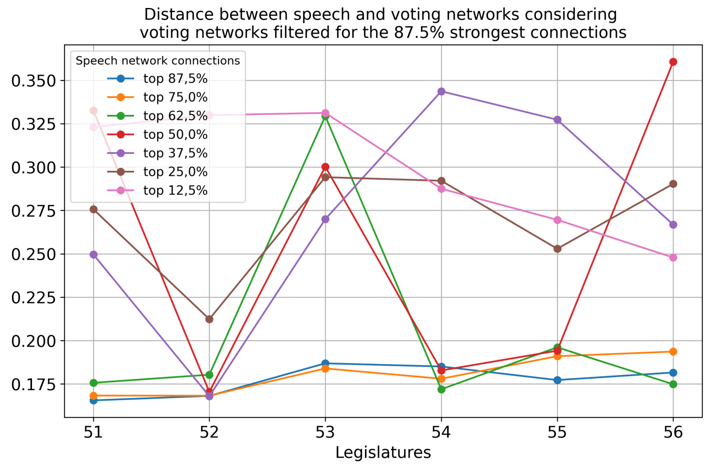


Figure 7 – Edge weights distribution for baseline speech networks across legislatures 51st to 56th. Dashed lines indicate quartile thresholds. Degree distributions were omitted because these networks are fully connected

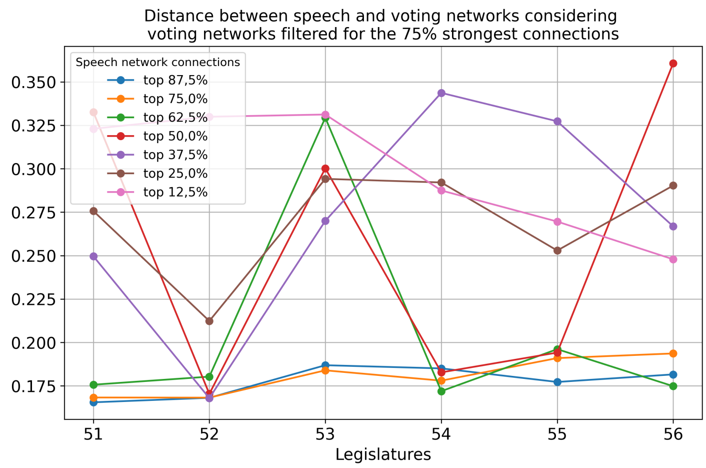
C Distance Measures Between Networks



(a)

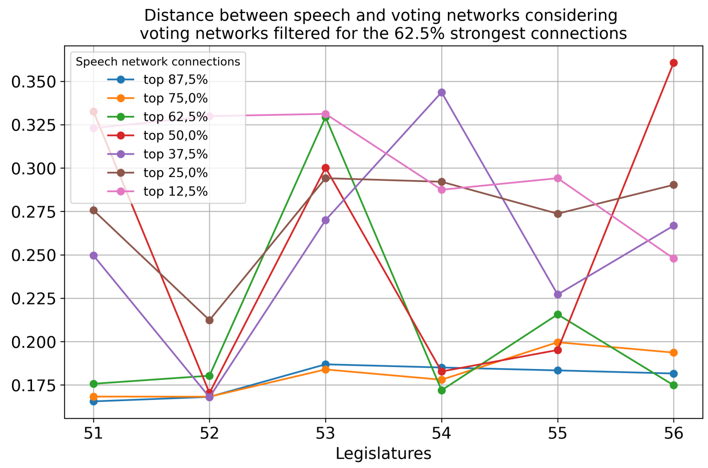


(b)

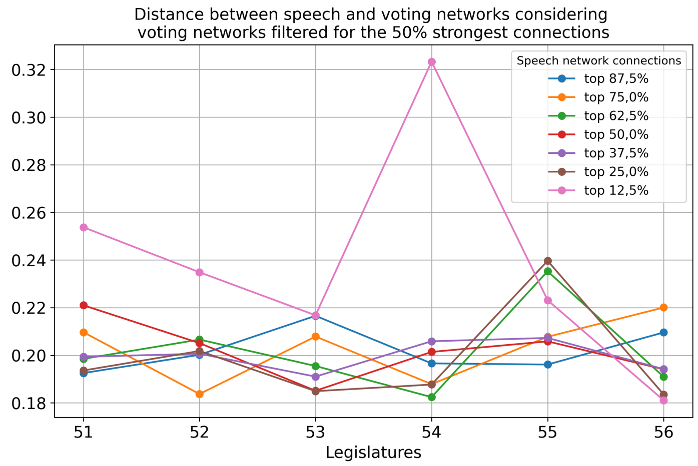


(c)

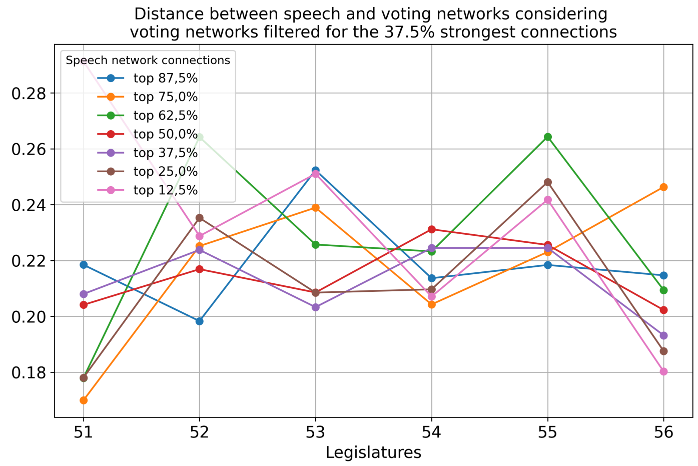
Figure 8 – Distance measures (1). Speech and voting networks are obtained applying different filters to the edges distribution of baseline models, keeping only the top x% strongest connections



(a)

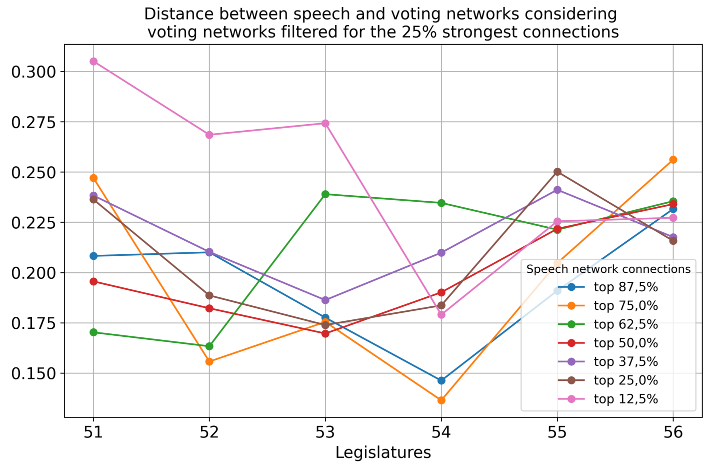


(b)

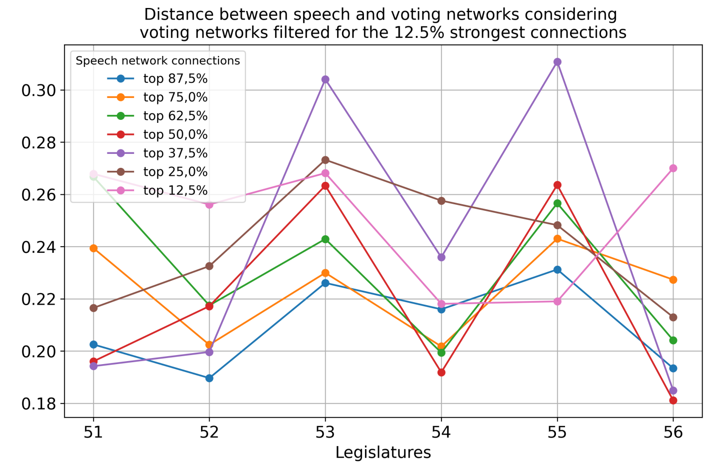


(c)

Figure 9 – Distance measures (2). Speech and voting networks are obtained applying different filters to the edges distribution of baseline models, keeping only the top x% strongest connections



(a)



(b)

Figure 10 – Distance measures (3). Speech and voting networks are obtained applying different filters to the edges distribution of baseline models, keeping only the top x% strongest connections

D Intra-party and Inter-party Connectivity of Voting Networks

Table 6 – Intra-party connectivity of party voting networks. Legislature 51st

Party	ρ_{intra}
PSB	0,500
PT	0,500
PPS	0,333
PFL	0,288
PSDB	0,233
PMDB	0,198
PDT	0,000
PTB	0,000
SPARTIDO	0,000
PPB	0,000
PL	0,000

Table 7 – Intra-party connectivity of party voting networks. Legislature 52nd

Party	ρ_{intra}
PT	0,626
PFL	0,386
PSB	0,300
PDT	0,267
PSDB	0,121
PMDB	0,116
PTB	0,067
PL	0,000
PSOL	0,000
SPARTIDO	0,000
PPS	0,000
PMR	0,000
PRB	0,000
PP	0,000

Table 8 – Intra-party connectivity of party voting networks. Legislature 53rd

Party	ρ_{intra}
PSDB	0,705
PT	0,577
PSB	0,333
PR	0,300
DEM	0,168
PMDB	0,163
PTB	0,089
PDT	0,071
PFL	0,000
PRB	0,000
SPARTIDO	0,000
PP	0,000
PSOL	0,000
PV	0,000
PSC	0,000

Table 9 – Intra-party connectivity of party voting networks. Legislature 54th

Party	ρ_{intra}
PCdoB	1,000
PSB	1,000
PT	0,562
PSDB	0,495
PP	0,381
PTB	0,143
PDT	0,095
DEM	0,095
PMDB	0,014
SPARTIDO	0,000
PR	0,000
PSC	0,000
PRB	0,000
PROS	0,000
PPS	0,000
MDB	0,000
PPL	0,000
PSD	0,000
PSOL	0,000
PV	0,000
PMN	0,000

Table 10 – Intra-party connectivity of party voting networks. Legislature 55th

Party	ρ_{intra}
PT	0,703
PSDB	0,410
MDB	0,340
PSB	0,286
PP	0,278
DEM	0,200
PSD	0,095
PDT	0,028
PCdoB	0,000
REDE	0,000
PSOL	0,000
PRTB	0,000
PRB	0,000
SPARTIDO	0,000
PTC	0,000
PROS	0,000
PMB	0,000
DC	0,000
PSC	0,000
CIDADANIA	0,000
PTB	0,000
PV	0,000
PL	0,000

Table 11 – Intra-party connectivity of party voting networks. Legislature 56th

Party	ρ_{intra}
PT	0,667
REDE	0,333
PSD	0,275
PSDB	0,267
DEM	0,143
MDB	0,133
PP	0,128
PL	0,111
CIDADANIA	0,100
PDT	0,048
PSB	0,000
PROS	0,000
PTB	0,000
PATRIOTA	0,000
PSL	0,000
SPARTIDO	0,000
PSC	0,000
REPUBLICANOS	0,000

Table 12 – Inter-party connectivity of party voting networks. Legislature 51st

Party	ρ_{inter}
PFL	0,136
PMDB	0,136
PSDB	0,125
PT	0,117
PSB	0,109
PPS	0,063
PPB	0,043
PDT	0,039
PTB	0,039
SPARTIDO	0,001
PL	0,000

Table 13 – Inter-party connectivity of party voting networks. Legislature 52nd

Party	ρ_{inter}
PT	0,168
PSB	0,114
PFL	0,110
PMDB	0,102
PTB	0,092
PPS	0,089
PSDB	0,069
PL	0,061
PDT	0,049
PSOL	0,000
SPARTIDO	0,000
PMR	0,000
PRB	0,000
PP	0,000

Table 14 – Inter-party connectivity of party voting networks. Legislature 53rd

Party	ρ_{inter}
PP	0,258
PSB	0,180
PT	0,170
PMDB	0,118
PSDB	0,114
PR	0,110
PTB	0,071
PRB	0,068
PDT	0,066
DEM	0,060
PSOL	0,016
PFL	0,000
SPARTIDO	0,000
PV	0,000
PSC	0,000

Table 15 – Inter-party connectivity of party voting networks. Legislature 54th

Party	ρ_{inter}
PV	0,289
PSB	0,257
PCdoB	0,226
PSC	0,193
PT	0,173
PP	0,164
PSOL	0,133
PSDB	0,114
PTB	0,110
PDT	0,093
PRB	0,062
DEM	0,050
MDB	0,039
PSD	0,031
PMDB	0,011
PR	0,000
SPARTIDO	0,000
PPS	0,000
PPL	0,000
PROS	0,000
PMN	0,000

Table 16 – Inter-party connectivity of party voting networks. Legislature 55th

Party	ρ_{inter}
PCdoB	0,223
PL	0,188
MDB	0,135
PP	0,134
PT	0,123
PSDB	0,121
DEM	0,111
REDE	0,098
PSD	0,094
PSB	0,094
PDT	0,047
PSC	0,027
CIDADANIA	0,014
PTB	0,012
SPARTIDO	0,006
DC	0,000
PMB	0,000
PROS	0,000
PRB	0,000
PRTB	0,000
PSOL	0,000
PV	0,000
PTC	0,000

Table 17 – Inter-party connectivity of party voting networks. Legislature 56th

Party	ρ_{inter}
PT	0,159
REPUBLICANOS	0,150
PSD	0,132
PSC	0,113
MDB	0,107
REDE	0,095
DEM	0,089
PP	0,083
PSDB	0,080
PROS	0,078
PDT	0,063
PL	0,061
CIDADANIA	0,050
PSB	0,021
PSL	0,021
PTB	0,000
PATRIOTA	0,000
SPARTIDO	0,000

E Intra-party and Inter-party Connectivity of Speech Networks

Table 18 – Intra-party connectivity of party speech networks. Legislature 51st

Party	ρ_{intra}
PSB	0,833
PT	0,500
PPS	0,333
PFL	0,145
PMDB	0,107
PDT	0,100
PSDB	0,100
PTB	0,100
SPARTIDO	0,000
PPB	0,000
PL	0,000

Table 19 – Intra-party connectivity of party speech networks. Legislature 52nd

Party	ρ_{intra}
PT	0,407
PSDB	0,263
PFL	0,229
PSB	0,100
PMDB	0,099
PDT	0,067
PL	0,000
PTB	0,000
PSOL	0,000
SPARTIDO	0,000
PPS	0,000
PMR	0,000
PRB	0,000
PP	0,000

Table 20 – Intra-party connectivity of party speech networks. Legislature 53rd

Party	ρ_{intra}
PT	0,423
PSDB	0,333
PSB	0,333
PMDB	0,137
PDT	0,107
DEM	0,100
PR	0,100
PTB	0,067
PFL	0,000
PRB	0,000
SPARTIDO	0,000
PP	0,000
PSOL	0,000
PV	0,000
PSC	0,000

Table 21 – Intra-party connectivity of party speech networks. Legislature 54th

Party	ρ_{intra}
PCdoB	1,000
PSB	0,667
PT	0,610
PMDB	0,188
PSDB	0,171
PP	0,048
PR	0,028
PDT	0,000
PPS	0,000
PV	0,000
PSOL	0,000
PSD	0,000
PPL	0,000
MDB	0,000
DEM	0,000
PRB	0,000
PSC	0,000
SPARTIDO	0,000
PROS	0,000
PTB	0,000
PMN	0,000

Table 22 – Intra-party connectivity of party speech networks. Legislature 55th

Party	ρ_{intra}
PT	0,593
PSB	0,321
PSDB	0,267
PDT	0,250
PP	0,083
DEM	0,067
PSD	0,048
PROS	0,000
PCdoB	0,000
REDE	0,000
PSOL	0,000
PRTB	0,000
PRB	0,000
SPARTIDO	0,000
PTC	0,000
PMB	0,000
DC	0,000
MDB	0,000
PSC	0,000
CIDADANIA	0,000
PTB	0,000
PV	0,000
PL	0,000

Table 23 – Intra-party connectivity of party speech networks. Legislature 56th

Party	ρ_{intra}
PT	0,667
DEM	0,476
REDE	0,333
CIDADANIA	0,300
PROS	0,167
MDB	0,152
PSD	0,085
PSDB	0,067
PDT	0,048
PP	0,038
PL	0,028
PSB	0,000
PTB	0,000
PATRIOTA	0,000
PSL	0,000
SPARTIDO	0,000
PSC	0,000
REPUBLICANOS	0,000

Table 24 – Inter-party connectivity of party speech networks. Legislature 51st

Party	ρ_{inter}
PSB	0,310
PT	0,211
PPS	0,159
PMDB	0,147
PFL	0,145
PTB	0,130
PSDB	0,125
PDT	0,111
PPB	0,086
SPARTIDO	0,011
PL	0,000

Table 25 – Inter-party connectivity of party speech networks. Legislature 52nd

Party	ρ_{inter}
PT	0,214
PFL	0,176
PSDB	0,163
PDT	0,133
PMDB	0,132
PSB	0,121
PPS	0,103
PL	0,043
PTB	0,038
PSOL	0,008
SPARTIDO	0,004
PMR	0,000
PRB	0,000
PP	0,000

Table 26 – Inter-party connectivity of party speech networks. Legislature 53rd

Party	ρ_{inter}
PSOL	0,347
PSB	0,292
PT	0,224
PSDB	0,207
PMDB	0,148
PDT	0,125
DEM	0,122
PTB	0,106
PRB	0,104
PR	0,097
PP	0,065
PSC	0,012
PV	0,008
PFL	0,005
SPARTIDO	0,000

Table 27 – Inter-party connectivity of party speech networks. Legislature 54th

Party	ρ_{inter}
PT	0,263
PSB	0,259
PCdoB	0,248
PMDB	0,177
PSDB	0,132
PP	0,124
PSOL	0,102
PRB	0,097
PDT	0,082
PSC	0,079
PV	0,070
DEM	0,057
PTB	0,053
PR	0,043
PSD	0,009
PMN	0,009
SPARTIDO	0,000
PPS	0,000
MDB	0,000
PPL	0,000
PROS	0,000

Table 28 – Inter-party connectivity of party speech networks. Legislature 55th

Party	ρ_{inter}
PCdoB	0,366
REDE	0,348
PT	0,262
PROS	0,250
PV	0,232
PSB	0,220
PDT	0,198
PSC	0,180
PSDB	0,169
PP	0,129
PSD	0,123
PTB	0,084
DEM	0,070
PMB	0,018
PSOL	0,018
SPARTIDO	0,008
MDB	0,001
PRTB	0,000
PTC	0,000
PRB	0,000
DC	0,000
CIDADANIA	0,000
PL	0,000

Table 29 – Inter-party connectivity of party speech networks. Legislature 56th

Party	ρ_{inter}
DEM	0,276
PT	0,251
CIDADANIA	0,236
REDE	0,174
MDB	0,160
PROS	0,157
PSB	0,115
PSD	0,113
PSDB	0,106
PDT	0,093
PL	0,090
PP	0,086
PSL	0,085
PSC	0,050
REPUBLICANOS	0,033
PTB	0,004
PATRIOTA	0,000
SPARTIDO	0,000

Bibliography

Andrade, R. F. S. et al. Measuring distances between complex networks. *Physics Letters A*, v. 372, n. 32, p. 5265–5269, ago. 2008. ISSN 0375-9601. Cited 2 times on pages 15 and 21.

Arnold, C.; Doyle, D.; Wiesehomeier, N. Presidents, Policy Compromise, and Legislative Success. *The Journal of Politics*, The University of Chicago Press, v. 79, n. 2, p. 380–395, abr. 2017. ISSN 0022-3816. Cited on page 14.

Cajueiro, D. O. et al. A model of indirect contagion based on a news similarity network. *Journal of Complex Networks*, v. 9, n. 5, p. cnab035, out. 2021. ISSN 2051-1329. Cited on page 15.

Diermeier, D. et al. Language and Ideology in Congress. *British Journal of Political Science*, Cambridge University Press, v. 42, n. 1, p. 31–55, jan. 2012. ISSN 1469-2112, 0007-1234. Cited on page 14.

Izumi, M. Y.; Medeiros, D. B. Government and Opposition in Legislative Speechmaking: Using Text-As-Data to Estimate Brazilian Political Parties' Policy Positions. *Latin American Politics and Society*, Cambridge University Press, v. 63, n. 1, p. 145–164, fev. 2021. ISSN 1531-426X, 1548-2456. Cited on page 14.

Lauderdale, B. E.; Herzog, A. Measuring Political Positions from Legislative Speech. *Political Analysis*, Cambridge University Press, v. 24, n. 3, p. 374–394, 2016. ISSN 1047-1987, 1476-4989. Cited on page 14.

Laver, M.; Benoit, K.; Garry, J. Extracting Policy Positions from Political Texts Using Words as Data. *American Political Science Review*, Cambridge University Press, v. 97, n. 2, p. 311–331, maio 2003. ISSN 1537-5943, 0003-0554. Cited on page 14.

Leoni, E. Ideologia, democracia e comportamento parlamentar: a Câmara dos Deputados (1991-1998). *Dados*, Instituto de Estudos Sociais e Políticos (IESP) da Universidade do Estado do Rio de Janeiro (UERJ), v. 45, p. 361–386, 2002. ISSN 0011-5258, 1678-4588. Cited on page 12.

Moreira, D. Com a Palavra os Nobres Deputados: Ênfase Temática dos Discursos dos Parlamentares Brasileiros. *Dados*, Instituto de Estudos Sociais e Políticos (IESP) da Universidade do Estado do Rio de Janeiro (UERJ), v. 63, p. e20180176, maio 2020. ISSN 0011-5258, 1678-4588. Cited on page 14.

Morgenstern, S. *Patterns of Legislative Politics: Roll Call Voting in Latin America and the United States*. Cambridge, U.K. ; New York: Cambridge University Press, 2004. ISBN 978-0-521-82056-1. Cited on page 12.

Nery, P. F.; Mueller, B. Co-sponsorship networks in the brazilian congress: An exploratory analysis of caucus influence. *Estudos Econômicos (São Paulo)*, Departamento de Economia; Faculdade de Economia, Administração, Contabilidade e Atuária da Universidade de São Paulo (FEA-USP), v. 52, n. 1, p. 83–111, Jan 2022. ISSN 0101-4161. Available on: <https://doi.org/10.1590/1980-53575213pnbm>. Cited on page 14.

Pincombe, B. Comparison of Human and Latent Semantic Analysis (LSA) Judgements of Pairwise Document Similarities for a News Corpus. jan. 2004. Cited on page 19.

Poole, K. T. Changing Minds? Not in Congress! *Public Choice*, Springer, v. 131, n. 3/4, p. 435–451, 2007. ISSN 0048-5829. Cited on page 12.

Poole, K. T.; Rosenthal, H. A Spatial Model for Legislative Roll Call Analysis. *American Journal of Political Science*, [Midwest Political Science Association, Wiley], v. 29, n. 2, p. 357–384, 1985. ISSN 0092-5853. Cited on page 14.

Power, T. J.; Zucco, C. Estimating Ideology of Brazilian Legislative Parties, 1990-2005: A Research Communication. *Latin American Research Review*, The Latin American Studies Association, v. 44, n. 1, p. 218–246, 2009. ISSN 0023-8791. Cited 2 times on pages 12 and 14.

Schwarz, D.; Traber, D.; Benoit, K. Estimating Intra-Party Preferences: Comparing Speeches to Votes*. *Political Science Research and Methods*, Cambridge University Press, v. 5, n. 2, p. 379–396, abr. 2017. ISSN 2049-8470, 2049-8489. Cited on page 14.