

**The London School of Economics and Political Science  
MY459-Special topics in quantitative text analysis**

**Words on the Floor:  
Thematic structure of political control speeches in Congress**

**Camila Salazar Mayorga  
Candidate number: 20337**

## **Abstract**

This paper analyses the thematic structure of political control speeches in the Costa Rican Congress, where legislators request to take the floor to discuss topics related to their own policy interests or to those of their political groups. Using a Latent Dirichlet Allocation (LDA) model, this paper identifies dominant topics addressed in the 2018-2019 legislature, and spots differences between and within parties. While the ruling party pushed an agenda related to gender and the economy, the opposition devoted their attention to infrastructure, criticizing the government and talking about the coastal areas of the country. Moreover, I found differences in the way parties addressed polarizing topics.<sup>1</sup>

---

<sup>1</sup> The full code for cleaning and analyzing the data can be found here:  
<https://github.com/Camilamila/Legislative-speeches-QTA>

# 1. Introduction

Legislators have different direct and indirect mechanisms to oversight the government and exert political control over the bureaucracy in general. For instance, they can “affect presidential action and accountability through budgetary control or the defeat of presidential initiatives” (Siavelis, 2007). Also, there can be special committees devoted to investigating particular matters or even some legislations allow impeachment.

However, in some Congresses there is an additional form of political scrutiny through legislative speeches, where a space during the legislative sessions is dedicated to political control. In that period legislators can request to take the floor in order to voice out their opinion about different topics or make accusations about policies or the government. While this mechanism can be more indirect than other types of oversight, in practice it is a good indicator of the topics legislators are interested in.

The objective of this paper is to identify the thematic structure of the political control speeches in the parliamentary sessions, using the Costa Rican parliamentary sessions as a case study. In comparison to other speeches in parliament, in this part of the session congressmen have no restrictions over the topics they discuss, which allows them to talk about particular matters, that can be related to their own policy interests or to those of their political parties. In particular this research aims to identify the dominant topics and spot differences between and within parties.

The analyzed corpus comprises of 763 speeches given from May 2018 to March 2019 in the Costa Rican Parliament, that corresponds to the first legislature of the 2018-2022 Congress. To detect the thematic structure of the speeches, I use a Latent Dirichlet Allocation (LDA) probabilistic model, which is an unsupervised method that does not require previous coding of the expected categories in the corpus and automatically detects the different topics.

The results show that there are differences in the topics by political groups, where the governing party dominates topics about women and gender and the economy, and the opposition focuses on regional issues, infrastructure and on criticizing the government. Moreover, the political control in Congress seems to have a diverse daily agenda, with legislators pushing topics of their own interest, and rarely this period of the sessions is devoted to a single topic in the current news agenda. Finally, I found differences in the way parties address polarizing issues.

# 2. Motivation

The 2018 parliamentary and presidential elections in Costa Rica resulted in a particular shift of the country’s politics. The winning presidential party, Partido Acción Ciudadana (PAC), became the governing party with the strongest opposition in the Congress’s history, since they obtained only

10 seats out of 57 and came to be the third majority in Parliament. Moreover, a minority party with little or no previous political experience, and strong conservative views, increased their seats from 1 to 14 congressmen, becoming the second strongest party in Parliament. Thus, it is of particular interest to study the thematic structure of this legislature.

The main hypothesis of this paper is that since during “political control” in the sessions there is no restriction over the topics in discussion, legislators talk about issues of their own interest, regardless if they are in the current legislative or political agenda. This could help identify more clearly the political interests of individual legislators and spot similarities or differences among political parties. Blackstone and Oldmixon (2015) citing Maltzman and Sigelman (1996) and Morris (2001) consider that “speeches are a vehicle through which representatives can express their legislative policy agenda to constituents and colleagues, and they may be a vehicle for shoring up electoral goals” (p. 3).

Legislative speeches related to political control in the Costa Rican Congress have been previously studied by Arias (2007), who analyzed the speeches of the 2006-2007 legislature. Using manual methods, he identified which institutions and issues were the most common in the discourses. He found that the opposition parties devoted more time to political scrutiny and most of the discussion focused on legal and administrative issues regarding the Congress operation. Other research has analyzed speeches using manual coding or supervised methods to identify religious agendas (Blackstone & Oldmixon, 2015), to study the gender differences in the legislator’s speech participation (Pearson & Dancey, 2011) and political ideologies (Diermeier et al., 2011).

In comparison, my approach uses unsupervised machine learning methods to analyze the legislative speeches, which has proved to be a useful technique to study large corpus of political texts. For instance Quinn et al. (2010) proved topic models to be useful to identify the topics that dominated the legislative debate in the United States Congress, Curran et al. (2018) used this technique to study speeches in the New Zealand Parliament and Barberá et al. (2014) studied the distribution of topics US legislators discussed on Twitter.

In addition, Grimmer (2010) used the expressed agenda model and hierarchical topic models, to analyze the attention senators in the United States allocated to issues in their press releases. “It is an expressed agenda because the attention dedicated to issues in communication are the issues that senators express as their priorities, not necessarily those issues that receive the most attention from senators while in Washington or from their staff” (p. 3). In our case instead of press releases we use the floor speeches as an indicator of the attention legislators dedicate to specific issues. This can be a good proxy to measure topics of interest and positions, first because legislators have to ask to take the floor, so only those legislators who have an expressed interest in voicing out their opinions will speak during the political control period. Moreover, as stated before, there are no

restrictions over the topics they can discuss, as in other parts of the session, when congressmen who request to speak can only refer to the bills in the agenda.

### 3. Corpus

According to the parliamentary regulations<sup>2</sup>, in the Costa Rican Congress, the first part of the ordinary legislative sessions each Wednesday is dedicated to the discussion of issues related to political control, scrutiny and oversight. However, in practice legislative authorities have implemented a thirty minute period in sessions from Monday to Thursday for these matters (Arias, 2008). During this time, legislators can request to take the floor and voice out their opinions. For this paper, I focus on the speeches given during this part of the session.

The minutes of each Congressional ordinary session were downloaded from the Congress webpage<sup>3</sup>, and then the speeches given in each session were extracted. The final corpus comprises 763 speeches (documents) given during 101 legislative sessions, between the 1<sup>st</sup> of May 2018 until the 31 of March 2019. This corresponds to the first legislative period of the 2018-2022 Congress, excluding the last month of the legislature, from the 1<sup>st</sup> to the 30 of April, that had a total of 11 sessions, since the minutes were not available online. While this could be a limitation of the study, the methods could be applied to the full corpus once the data becomes available.

In total, the Costa Rican Congress has 57 legislators and currently has representation of seven political parties: PLN (19), PRN (14), PAC<sup>4</sup> (10), PUSC (9), PIN (4), PRSC (2) and FA (1).

To clean the documents, I removed all punctuation and stop words in Spanish and placed all the words into lowercase. Moreover, I removed words that appeared in fewer than 2 documents and very common words used by most of the legislators that were uninformative<sup>5</sup>. Moreover, I grouped all the interventions each legislator had in the same session, in a single speech. Finally, I removed documents with less than 100 words that were related to short interventions and not speeches. Figure 1 shows the distribution of words per document. In general, speeches have a median of 708 words.

---

<sup>2</sup> Reglamento de la Asamblea Legislativa

<sup>3</sup> <http://www.asamblea.go.cr>

<sup>4</sup> Governing party

<sup>5</sup> Very common uninformative words: diputad[ao]s, diputad[ao], señora, señor, compañer[ao], compañer[ao]s, gracias, don, doña, presidenta

Figure 1. Distribution of the word count in legislative speeches

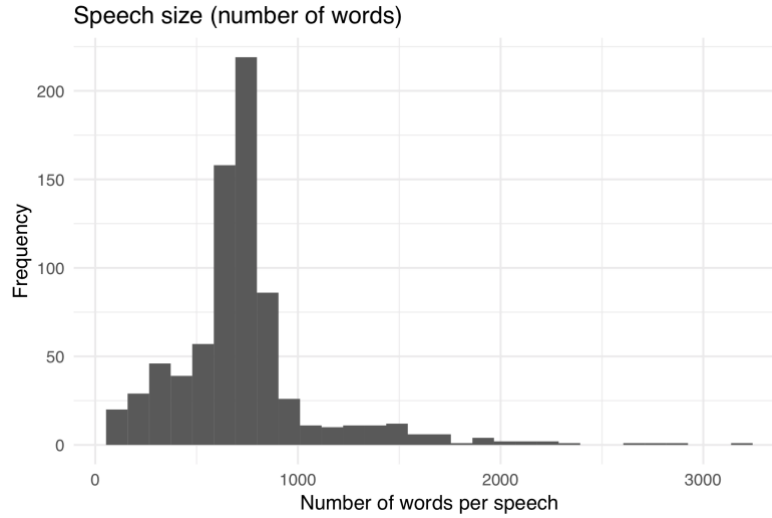


Table 1 shows the distribution of speeches per party. Two opposition parties, PLN and PRN, who have the highest number of seats in Parliament (17 and 14 seats respectively) are the ones who dominate the floor, while the ruling party PAC is in the third place. This shows that opposition parties talk more in Congress, which could relate to the fact that these legislators want to draw attention to particular topics or criticize the government.

Table 1. Words and speeches by party

Party	Speeches	Words	Number of legislators
PLN	200	154,962	17
PRN	191	126,902	14
PAC	124	101,736	10
PUSC	126	89,267	9
PIN	54	38,998	4
FA	37	26,970	1
PRSC	31	21,051	2

## 4. Methods

To identify the different topics in the parliamentary speeches, I use the Latent Dirichlet Allocation (LDA) probabilistic model, where “the basic idea is that documents are represented as random mixtures over latent topics, where each topic is characterized by a distribution over words” (Blei

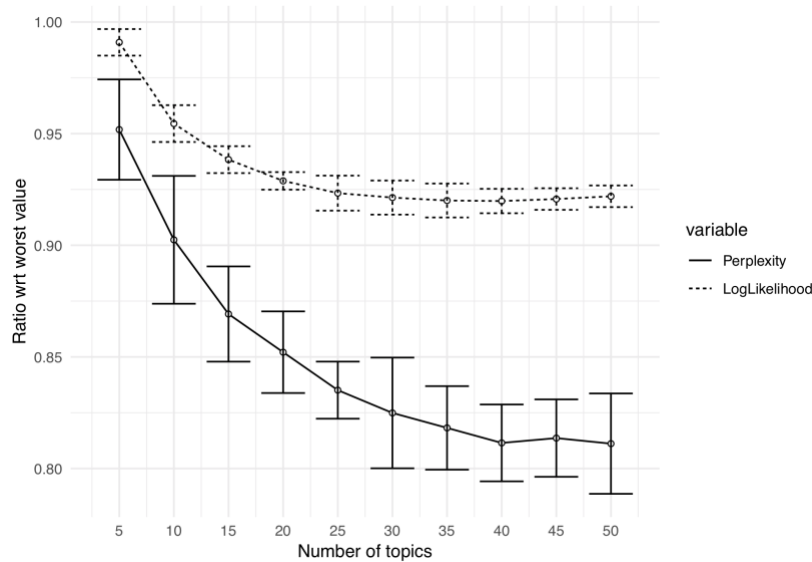
et al., 2003, p. 996). The details of the degenerative model can be found in Blei et al. (2003), p. 996).

This unsupervised machine learning method is useful for our application since it “does not require a researcher to know the underlying taxonomy of categories with certainty. Instead, the division of topic and keywords that identify each topic are estimated from the text” (Quinn et al., 2010). In comparison, other supervised methods require previous classification tasks that can introduce bias to the estimations. Moreover, while we could try to use clustering methods to classify the documents, most of these models “restricts a document to being associated with a single topic” (Blei et al., 2003, p. 997), and the LDA allows the documents to belong to different categories with different probabilities.

Each document  $D$  in the corpus will be on a finite number of  $K$  topics, and I assume that topic distributions are independent over time. In the model, there are two parameters of interest:  $\beta$ , a matrix of dimensions  $K$  (topics)  $\times$   $M$  (words), where  $\beta_{km}$  corresponds to the probability that word  $m$  belongs to topic  $k$ . And  $\theta$ , a matrix of dimensions  $D$  (documents)  $\times$   $K$ , where  $\theta_{dk}$  is the probability that document  $d$  belongs to topic  $k$ .

In this model, it is necessary to fix the number of topics ( $K$ ). For this I ran the model with different values of  $K$  (5, 10, 15, 20, 25, 30, 35, 40, 45, 50) using Gibbs sampling methods, and chose the model with a low log-likelihood and perplexity, avoiding overfitting. I found that a value of  $K=35$  fits the model best, as seen in Figure 2.

Figure 2. LDA model fit



Finally, to validate the results of the model I checked the semantic and predictive validity. The first one refers to “the extent to which each category or document has a coherent meaning and the

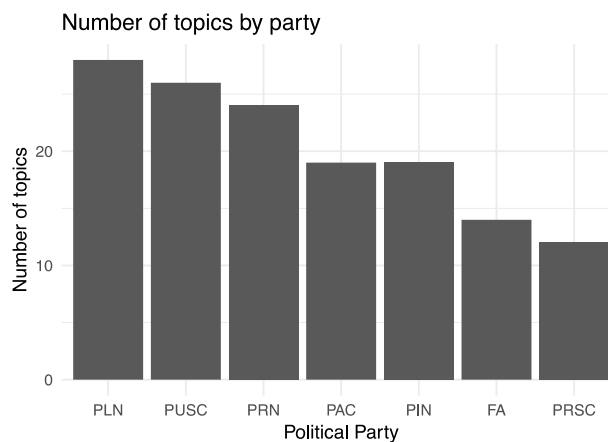
extent to which the categories are related to one another in a different way” (Quinn et al., 2010, p. 216), while the predictive validity is used to see if the identified topics in the documents relate to exogenous events in the news and political agenda.

## 5. Results

The top scoring words of each of the 35 topics detected by the model, show that 28 of the topics had a coherent meaning and proved to be informative. These topics account for 80% of the speeches given during the analyzed period. They range from economic issues such as budget, taxes and investment, to specific events or social matters. In general, the agenda during the first legislature of this Parliament is diverse and was not dominated by a specific matter. Table A.1 in the appendix shows the top words associated with each topic, the label given to the topic, the proportion of the total speeches it appears in, and the party that speaks most about each topic.

I found that there is a division of topics among parties, which can be seen clearly in Figures 3 and 4. PLN spoke, at least once, about all the 28 identified topics, while the ruling party (PAC) covered only 19. Other opposition parties, such as the PRN and PUSC had also more topic diversity, with 24 and 26 topics, respectively. However, the attention legislators devote to specific matters varies among the different political groups.

*Figure 3. Number of topics by political party*



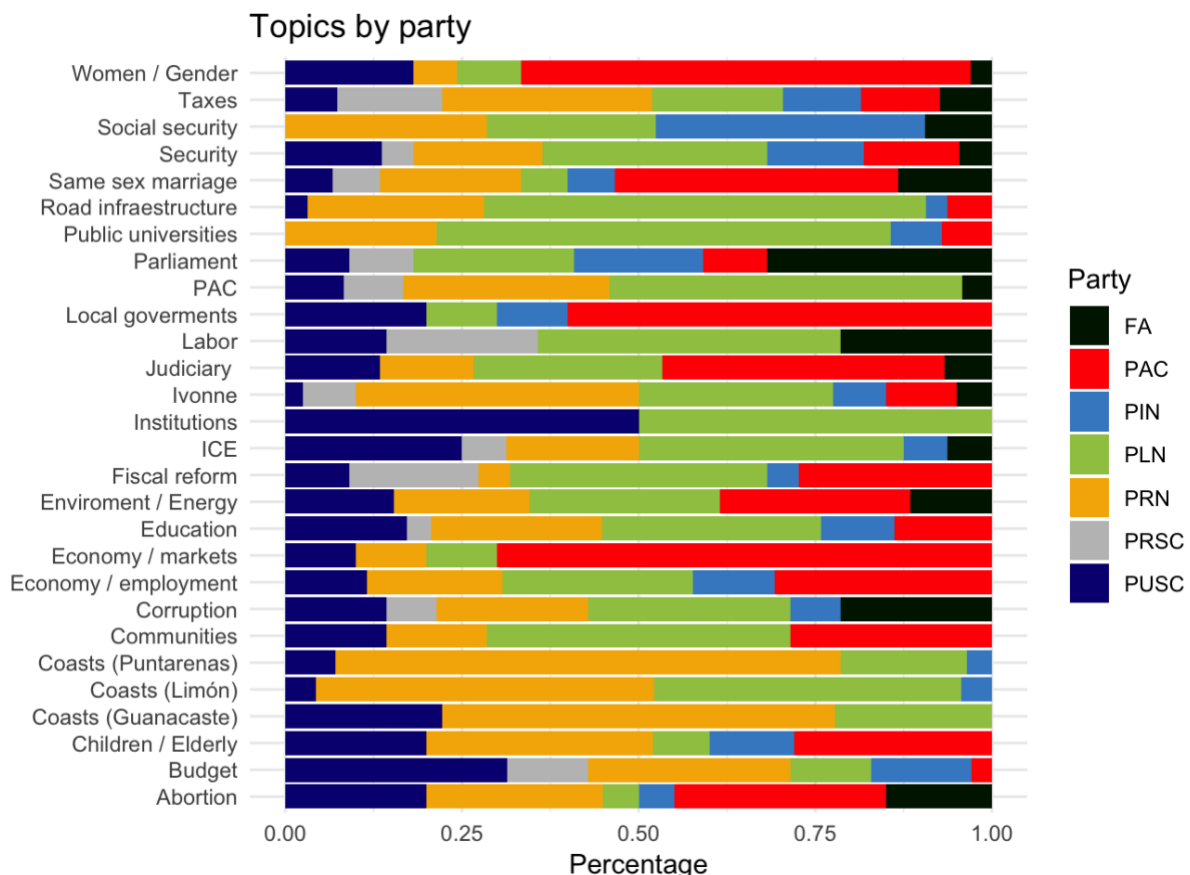
Speeches related to women and gender (topic 10) and to the economy and markets (topic 15) came mainly from the governing party PAC, while those related to the budget of public universities and road infrastructure were mainly discussed by the PLN, the biggest opposition party in Congress. Moreover, the PRN, the second majority, dominated topics related to the coastal areas of the country, which makes sense since most of the legislators from this party were elected in these territories. These two parties are also the ones who talked more about the government (topic PAC) and corruption, which demonstrates the opposition uses this part of the sessions to play an active



role in scrutinizing the current administration. Additionally, dedicating more speeches to a single topic, could be more effective in terms of positioning a subject on the political agenda.

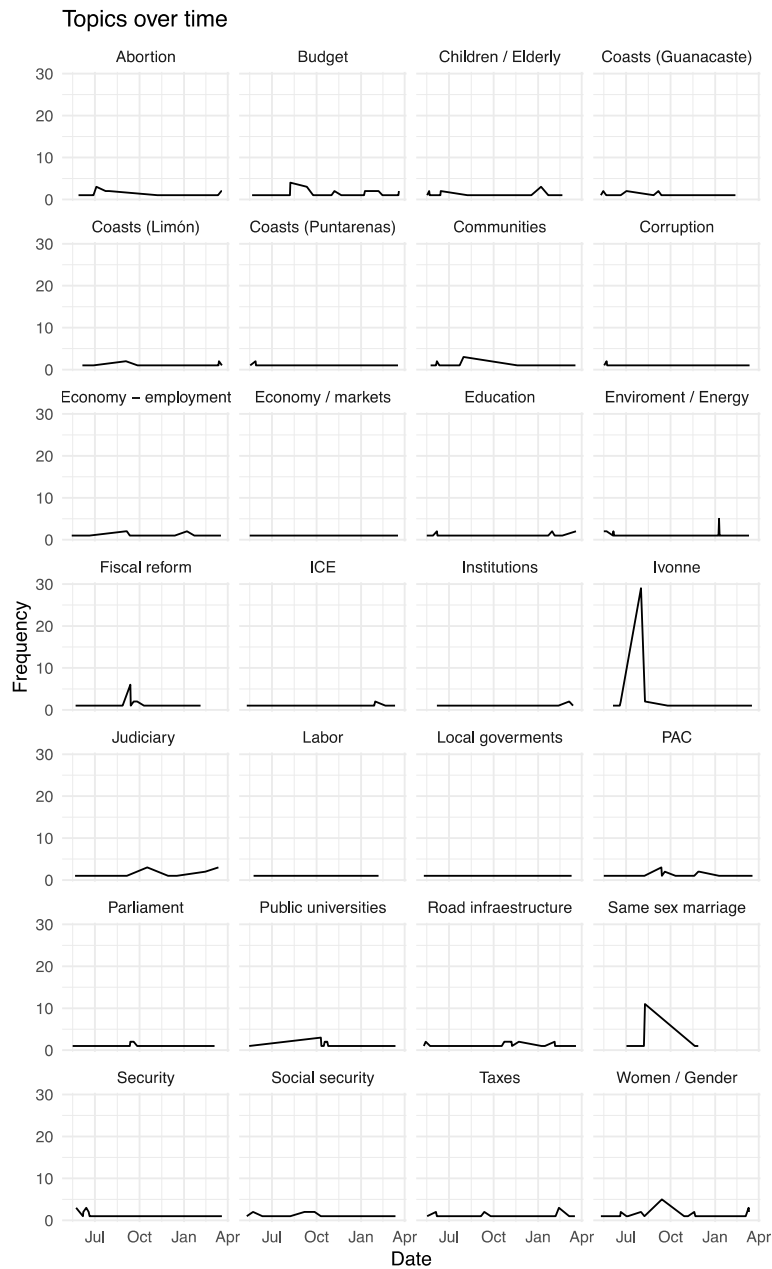
On the contrary, topics related to security, taxes, the fiscal reform and education concentrated the attention of different parties. This confirms the hypothesis that there are partisan and non-partisan topics in the congressional speeches.

Figure 4. Distribution of topics by party



When analyzing the frequency of the topics through time, there are only two topics that monopolized the agenda of specific sessions and are related to precise news events (predictive validity). The first one relates to a sentence by the Costa Rican Supreme Court on the 8<sup>th</sup> of August 2018, that lifted the ban on same-sex marriage and gave the Parliament 18 months to legalize it. The following day, all of the speeches in Congress referred to that topic. In addition, the speeches on the 1<sup>st</sup> of August, are all about the topic “Ivonne”, were the legislators condemned a social media offensive publication of a renowned doctor, against the legislator Ivonne Acuña (PRN). The rest of the topics, had an evenly distribution over time. This shows that in the same sessions, legislators rarely discuss about the same issues and the daily agendas are diverse.

Figure 5. Distribution of topics over time

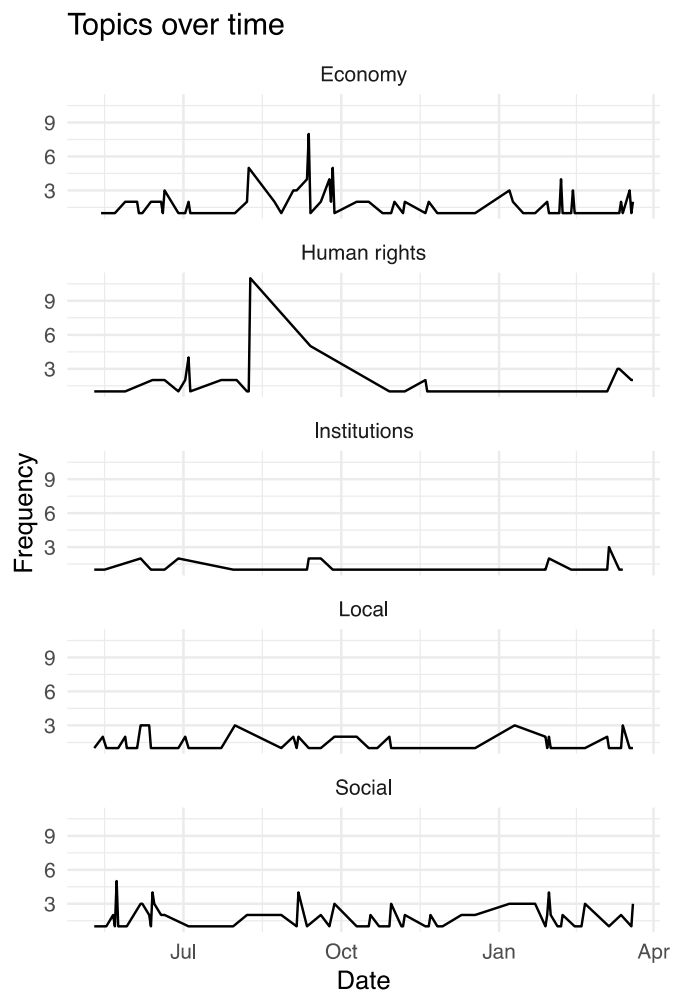


Moreover, to see the evolution of topics more clearly, I regrouped the topics into 6 clusters<sup>6</sup>: Economy (Budget, Economy / employment, Economy / markets, Fiscal reform, Taxes), Social (Children / Elderly, Education, Labor, Social security, Security), Human rights (Abortion, Women / Gender, Same sex marriage), Local (Coasts, Communities, Local governments), Institutions

<sup>6</sup> I excluded the Topic Ivonne since it was unrelated to other topics

(ICE, Institutions, Judiciary, Parliament, Environment / Energy) and Other. In Figure 6 we see that topics related to the economy and social and local issues are constantly in the legislative agenda, while speeches about human rights or related to institutions are discussed less, and only in particular moments of time (as in the case of same sex marriage).

*Figure 6. Distribution of grouped topics over time*

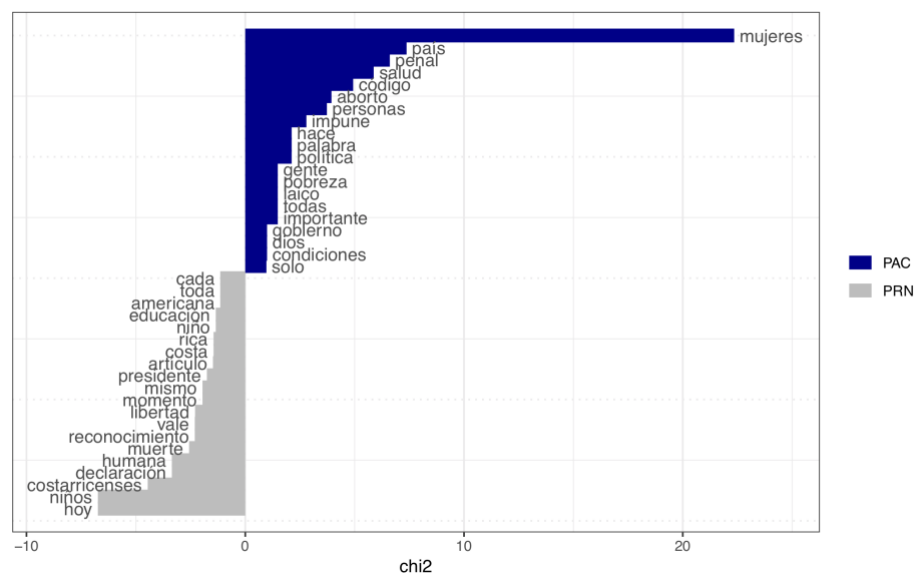


Besides studying the topics by party, it is relevant to analyze if the most participative legislators have diverse agendas or if they focus on specific topics. To answer this question, I analyzed the legislators with the highest number of speeches (top quartile). Out of these 15 deputies, Welmer Ramos (PAC) and Melvin Núñez (PRN) are the most monothematic congressmen, since 85% of the speeches they gave are about the same topic. Núñez spoke mainly about the coastal area of Puntarenas, where he is originally from, reinforcing the finding that some legislators have a regional instead of a national focus, and address issues that directly concern their constituents.

Ramos, who was the Minister of Economy in the previous administration and a presidential pre-candidate, focused on the economy and markets, which is consistent with his political agenda. Other legislators that dedicated more than 40% of their speeches to a single topic are David Gourzong (PLN) and Eduardo Cruickshank (PRN), who talked about their home province Limón, Wagner Jiménez (PLN, Education), Nielsen Pérez (PAC, Women & Gender), Wálter Muñoz (PIN, Social Security) and Paola Vega (PAC, Environment). On the contrary, José María Villalta, who is the Congressman with the higher number of speeches, divided his attention in 14 different topics instead of focusing on a single matter.

Finally, I explore the differences between the governing party (PAC) and the PRN, the two parties that advanced to the second round of the 2018 presidential elections, when talking about polarizing topics such as abortion and same sex marriage, since these two parties, held opposing views about these issues during the political campaign. For this I use keyness analysis to identify which words are more commonly used by each political group. The most common words used in the PAC speeches about abortion (Figure 5) are “mujeres” (women), “penal” and “código” (referring to the Penal Code) and “salud” (health). On the contrary, PRN legislators use words like “niños” (children) and “muerte” (death). This shows the different focus these parties have towards the topic: the PAC (in favor of abortion) considers the topic as a health issue for women, while the PRN (against) opposes abortion advocating for the unborn children and considering the practice as deathly.

Figure 7. Keyness analysis of speeches about abortion



In the case of sex same marriage (Figure A1 in the Appendix), the results are less clear, since the most discriminant word used by the PAC, “ser” (being) is uninformative, but they use other words such as “parejas” (couples), “derechos” (rights) and “discriminación” (discrimination). While the PRN uses “patrimoniales” (patrimonial rights) and “nunca” (never). Thus, while there seems to

be differences in the words used by the parties, they are not as informative as in the abortion speeches.

## 6. Conclusions

Political control speeches in the 2018-2019 legislature of the Costa Rican Congress were dominated by the biggest opposition parties (PLN and PRN) and addressed diverse topics that ranged to the economy, fiscal reform and education, to gender and same sex marriage. I found evidence of differences in the attention political parties and individual legislators dedicate to specific issues.

The ruling party (PAC) dominated speeches about gender and the economy, while the opposition focused more on road infrastructure, public universities spending and issues related to the coastal areas of the country, where the PAC won no seats. In general, discussions about economic, social and local topics were frequent throughout the legislature and were constantly discussed by legislators, while human rights and speeches related to institutions only appeared in particular moments of the legislature. In addition, polarizing issues in the 2018 political campaign, like abortion and same sex marriage, were also discussed in Congress, and the keyness analysis of the speeches helped identify clear differences in the way the two rival parties addressed these matters.

This is a first approach to using unsupervised methods to analyze congressional speeches in the Costa Rican Congress. The results are consistent and the analysis could be expanded to a larger corpus in order to study the evolution of certain topics in Congress, spot the changes in the issues political parties discuss through time and analyze the variations in the approaches. Moreover, further research could compare the thematic structure of the political control speeches, to speeches legislators give in other spaces of political discussion in Parliament.

## 7. References

- Arias, B. (2007). Control político en el Plenario de la Asamblea Legislativa. Primera legislatura de la administración Arias Sánchez 2006-2007. *Ponencia Preparada Para El Decimotercer Informe Estado de La Nación*, 1–19.
- Arias, B. (2008). Estudio comparativo de normativa parlamentaria. *Ponencia Preparada Para El Decimocuarto Informe Estado de La Nación*.
- Barberá, P., Bonneau, R., Egan, P., Jost, J. T., Nagler, J., & Tucker, J. A. (2014). Leaders or Followers? Measuring Political Responsiveness in the U.S. Congress Using Social Media Data. *APSA Meeting Paper*.
- Blackstone, B., & Oldmixon, E. A. (2015). Discourse and dissonance: Religious agendas in the 104th congress. *Research and Politics*, 2(3). <https://doi.org/10.1177/2053168015589215>
- Blei, D. M., Edu, B. B., Ng, A. Y., Edu, A. S., Jordan, M. I., & Edu, J. B. (2003). Latent

- Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993–1022.  
<https://doi.org/10.1162/jmlr.2003.3.4-5.993>
- Curran, B., Higham, K., Ortiz, E., & Filho, D. V. (2018). Look who’s talking: Two-mode networks as representations of a topic model of New Zealand parliamentary speeches. *PLoS ONE*, 13(6), 1–16. <https://doi.org/10.1371/journal.pone.0199072>
- Diermeier, D., Godbout, J., Yu, B., & Kaufmann, S. (2011). Language and Ideology in Congress. *British Journal of Political Science*, 42(1), 31–55.
- Grimmer, J. (2010). A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in senate press releases. *Political Analysis*, 18(1), 1–35.  
<https://doi.org/10.1093/pan/mpp034>
- Pearson, K., & Dancey, L. (2011). Elevating women’s voices in congress: Speech participation in the house of representatives. *Political Research Quarterly*, 64(4), 910–923.  
<https://doi.org/10.1177/1065912910388190>
- Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H., Radev, D. R., & Quinn, K. M. (2010). How to Analyze Political Attention with Minimal Assumptions and Costs. *American Journal of Political Science*, 54(1), 209–228.
- Siavelis, P. M. (2007). Disconnected Fire Alarms and Ineffective Police Patrols: Legislative Oversight in Postauthoritarian Chile. *Journal of Interamerican Studies and World Affairs*, 42(1), v. <https://doi.org/10.2307/166466>

## 8. Appendix

### 8.1. Additional Tables and Figures

Table A1. Top features of topics (LDA model with  $k=35$ )

Topic	Label	Terms	Prop	Party
5	Ivonne	ser, usted, redes, sociales, muchas, respeto, dios, puede, ivonne, personas, siempre, creo, persona, familia, nunca, mujer, costarricense, todas, valores, hoy	5.2	PRN
6	Budget	millones, mil, colones, dólares, pensiones, presupuesto, gobierno, banco, plata, año, dinero, pagar, dos, cinco, treinta, deuda, quinientos, hacienda, recursos, lujo	4.6	PRN
10	Women / Gender	mujeres, violencia, mujer, sociedad, política, día, años, hombres, derechos, muchas, igualdad, vida, todas, género, forma, participación, sexual, quiero, políticos, espacio	4.3	PAC
7	Road infrastructure	san, obras, zona, ruta, carretera, años, obra, infraestructura, josé, muchas, ministro, construcción, norte, proyecto, carlos, conavi, puente, cantón, río, buenas	4.2	PLN
17	Education	educación, estudiantes, niños, jóvenes, centros, ministerio, escuela, ministro, bachillerato, educativos, escuelas, infraestructura, pública, educadores, colegios, semana, dual, tres, educativa, educativo	3.8	PRN
16	Coasts (Puntarenas)	puntarenas, trabajo, gente, cómo, hoy, quiero, lugar, trabajar, muchas, hace, gobierno, personas, hablar, sino, ahí, da, cosas, casa, calle, mismo	3.7	PRN

12	Taxes	fiscal, plan, impuestos, hacienda, básica, salarios, canasta, gasto, evasión, impuesto, gobierno, sector, ministra, ingresos, público, medidas, presidente, dos, gastos, crisis	3.5	PRN
1	Environment / Energy	agua, país, ambiente, crucitas, ambiental, energía, tema, años, aguas, día, ley, materia, protección, comunidades, todas, fracción, oro, ministro, problemas, recurso	3.4	PAC / PLN
32	Economy - employment	país, desarrollo, nacional, inversión, económica, sector, economía, reactivación, acciones, empleo, debemos, pública, requiere, gobierno, público, crecimiento, proyectos, económico, empresas, debe	3.4	PAC
24	Children / Elderly	personas, niños, mayores, población, mayor, atención, día, tardes, buenas, falta, acceso, sociedad, menores, nacional, niñas, adultos, edad, social, todas, servicios	3.3	PRN
28	PAC	gobierno, presidente, partido, república, carlos, pac, ministro, alvarado, fracción, acción, solís, ciudadana, nacional, ejecutivo, usted, administración, luis, comunicación, hizo, presidencia	3.1	PLN
18	Coasts (Limón)	limón, provincia, turismo, servicios, proyectos, presidente, oportunidades, japdeva, recursos, familias, ministro, área, atención, república, va, empleo, gran, trabajo, limonenses, turistas	3	PRN
8	Fiscal reform	país, responsabilidad, decisiones, situación, fiscal, tomar, crisis, decisión, ser, aquí, reforma, hoy, responsable, pueblo, muchas, seguir, forma, sectores, responsables, llamado	2.9	PLN
35	Parliament	proyecto, asamblea, legislativa, ley, comisión, plenario, reforma, acuerdo, hoy, fracción, fracciones, diálogo, legislativo, expediente, si, muchas, presentado, discusión, mociones, primer	2.9	FA
25	Security	seguridad, armas, tema, país, nicaragua, dos, hoy, fuerza, nacional, arma, medidas, tres, violencia, cuentas, ilegal, narcotráfico, llamado, nadie, crimen, puede	2.9	PLN
14	Social security	salud, caja, seguro, social, problema, hospital, régimen, costarricense, médicos, institución, médico, ver, años, atención, realmente, tema, habla, aún, relación, posibilidad	2.8	PIN
34	Abortion	vida, derechos, humanos, derecho, persona, indígenas, personas, aborto, libertad, muerte, pueblos, humano, condición, niño, ser, código, convención, respeto, día, paz	2.6	PAC
15	Economy / markets	costa, rica, mercado, si, países, tasas, mercados, ciento, ley, demás, tema, poder, competencia, ejemplo, hace, crédito, costo, podría, economía, internacional	2.6	PAC
22	Coasts (Guanacaste)	guanacaste, agua, proyecto, necesitamos, va, personas, provincia, vamos, sector, ser, día, apoyo, tener, productores, tarde, puedan, buenas, así, ustedes, llevar	2.4	PRN
19	ICE	informe, general, institución, costarricense, ice, hecho, contraloría, venido, debe, instituto, manera, costarricenses, municipalidades, muchas, control, servicio, república, gobiernos, información, locales	2.1	PUSC
21	Judiciary	poder, justicia, judicial, ser, debe, comisión, parlamento, transparencia, solo, legislativo, corte, hace, magistrados, vez, cada, poderes, votos, nombramientos, políticos, plena	2	PAC

4	Same sex marriage	sala, derechos, matrimonio, personas, constitucional, humanos, asamblea, si, corte, meses, legislativa, resolución, mismo, dieciocho, opinión, familia, derecho, igualitario, interamericana, dice	2	PAC
27	Communities	atención, gobierno, comunidades, comisión, comunidad, cantón, situación, emergencias, vez, república, atender, presidente, solo, zona, dar, instituciones, emergencia, sino, grandes, compromiso	1.8	PLN
3	Corruption	caso, comisión, denuncia, público, denuncias, comunicación, mayo, medios, casos, corrupción, últimos, sistema, primera, ayer, después, cualquier, elecciones, medio, supremo, soda	1.8	PLN
11	Labor reform	social, huelga, trabajo, política, trabajadores, derecho, hoy, mora, partido, hacer, laboral, reforma, años, duda, código, acuerdo, cualquier, sociales, costarricense, manuel	1.8	PLN
33	Public universities	nacional, universidad, universidades, presupuesto, educación, costarricenses, recursos, gasto, alberto, públicas, especial, superior, fondo, ayer, pública, catorce, sabemos, construir, cinco, construcción	1.8	PLN
26	Institutions	ministerio, empresa, si, hace, empresas, aquí, tipo, instituciones, información, ejecutivo, hago, servicios, administrativa, proyecto, públicos, contratación, atención, llamado, puede, días	1.3	PLN / PUSC
30	Local governments	público, además, así, diferentes, ciudadanía, municipal, interés, reglamento, sino, actividades, municipalidad, forma, alcalde, puestos, acciones, investigación, problemas, cartago, preocupación, falta	1.3	PAC

Figure A1. Keyness Analysis of same sex marriage speeches

