Physica A 571 (2021) 125832

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

On the stability of the Brazilian presidential regime: A statistical analysis

Frederico Fetter *, Daniel Gamermann, Carolina Brito

Department of Physics, Universidade Federal do Rio Grande do Sul (UFRGS), Instituto de Física, Brazil

ARTICLE INFO

Article history:
Received 26 December 2020
Received in revised form 29 January 2021
Available online 12 February 2021

Keywords:
K-means
Shannon entropy
Political term stability

ABSTRACT

Brazil's presidential system is characterized by the existence of many political parties that are elected for the Chamber of Deputies and unite in legislative coalitions to form a majority. Since the re-democratization in 1985, Brazil has had 8 direct presidential elections, among which there were two impeachments of the elected presidents. In this work we identify clear differences between stable presidential periods and Legislative terms with an impeachment by analyzing the votes that took place in the Chamber of Deputies from 1991 to 2019. Our statistical analysis are blind to the content of the bills. We start by measuring the cohesion of the parties and the congress for each bill. We then quantify the agreement between the votes of congressmen and observe that there is a stronger polarization among congressmen during legislative periods where there was no impeachment, referred here as stable legislative periods. Using clustering algorithms, we are able to associate these polarized groups observed during the stable periods with the opposition to the government and government base. For periods with an impeachment, the data shows that the congress split up in more than two groups. To characterize the impeachment of Collor and Dilma Rousseff (in 1992 and 2016, respectively) we analyze how the agreement between congressmen and the government evolved over time and we also propose a division of the congressmen in three groups. We identified that, in periods with an impeachment, the third group aligns itself against the president.

1. Introduction

Modern day technologies allow for the easy gather, storage and distribution of data relevant in many fields, such as political sciences, which have been traditionally tackled mainly by qualitative approaches. Quantitative analysis of these data should give insights in those areas [1–3], such as in the workings of the many democratic institutions that run most of the countries in the world. The idea behind western-style representative democracy [4] is that, through elections, the people choose those that better represent their interests for these many democratic institutions and that the plurality of these institutions guarantees the checks and balances needed to prevent abuses and self interests from prevailing. Moreover, representatives with common goals and ideals gather together to form political parties which, in principle, ease the identification of the electorate's interests within these groups of representatives.

One of such institutions is the chamber of representatives or national congress, where laws are analyzed and voted by the congressmen (elected representatives of the people). Several works analyze data from elections [5,6], its financing and how inappropriate money gathering might threaten the idea that the elected candidates do represent the people's

* Corresponding author.
E-mail addresses: frederico.fetter@gmail.com (F. Fetter), carolina.brito@ufrgs.br (C. Brito).

https://doi.org/10.1016/j.physa.2021.125832
0378-4371/

Table 1
A summary of the Legislative periods analyzed in this work. The columns show respectively: the Legislature number, the abbreviation of the name of the president, the starting and ending date of the term, number of congressmen N_d, number of roll calls N_{rc} voted during the respective legislative period, the number of parties N_{par} and the effective number of parties N_{par}^{eff}.

Legis.	Presidential Term	Start date	End date	N_d	N_{rc}	N_{par}	N_{par}^{eff}
49	Collor	01/02/1991	28/09/1992	507	73	19	9.1
	Itamar	29/09/1992	31/01/1995	506	85		
50	FHC I	01/02/1995	31/01/1999	573	468	18	8.1
51	FHC II	01/02/1999	31/01/2003	563	419	17	7.1
51	Lula I	01/02/2003	31/01/2007	545	450	18	8.4
53	Lula II	01/02/2007	31/01/2011	536	611	20	9.3
54	Dilma I	01/02/2011	31/01/2015	553	430	22	10.4
55	Dilma II	01/02/2015	11/05/2016	521	330	28	13.4
	Temer	12/05/2016	31/01/2019	537	525		
56	Bolsonaro	01/02/2019	31/12/2019	517	329	30	16.4

interests [7–9]. Also the dynamics in the legislative chambers is analyzed in different contexts. Such works might take into account the contents of the issues discussed and voted by the congressmen [10–12]. Ideal point models are based on our intuition that congressmen have positions in an abstract ideological space and cast their votes in roll calls based on where the voted bills lie in this space [13,14]. Such approaches, though, usually require an assessment not only of the voting data, but also of the contents of bills and speeches, which can be biased by the readers interpretations and therefore not indisputably objective. Other approaches study only the similarities between the votes cast by the politicians [15,16].

In the present work, we follow this latter approach and analyze the voting patterns in the Brazilian Chamber of Deputies (the lower chamber of Brazil's bicameral legislative power, composed by this chamber and the senate). Our approach is blind to the contents of the bills voted in the congress. It first analyzes the cohesion of the votes results for each bill by means of an entropy related measure and then the similarities in the votes sequences of each pair of congressmen using the *k*-means clustering algorithm. The first analysis suggests that there is a correlation between the polarization of the deputies inside each party and in the congress as a whole. The objective of the clustering analysis is to identify the *de facto* groups of congressmen that vote cohesively across several roll calls and some of the dynamics behind these groups. We note that there are more than 30 political parties represented in the Brazilian congress [17] and the president's party alone does not have majority in the legislative. Our analyses show that the nominal 30 different political parties end up split into two actually interacting groups [18] during the stable legislatures, but that for the terms with an impeachment, the separation into two groups is not enough to characterize the complexity of the voting patterns. We also investigate how changes in this split might lead to an unstable government resulting in the impeachment of the president.

The data analyzed, the methods and metrics used in the study are explained in the next section. In Section 3 we present the results highlighting the differences between stable regimes and those where an impeachment happened. Finally we present a discussion, summary and our conclusions.

2. Data and methods

2.1. Legislative periods analyzed in this work

Table 1 presents information about the eight legislative terms analyzed in this work. It shows the number of the legislative period, the abbreviation of the name of the presidents (which is how we will refer to them along the paper) the start and end date of the term of each president, the total number of congressmen N_d and the total number of roll calls N_{rc} in the period. The names of the presidents are Fernando Collor de Mello (identified as Collor), who resigned as president in 1992 to prevent an impeachment. He was followed by Itamar Franco (Itamar), then Fernando Henrique Cardoso (FHC), who had two consecutives mandates (that we refer to as FHC I and FHC II). Luiz Inácio Lula da Silva (Lula) also had two mandates, followed by Dilma Rousseff (Dilma) who completed her first mandate (Dilma I) and was elected for the second term (Dilma II) but was removed from office through an impeachment process in 2016 and replaced by Michael Temer (Temer). In 2019 Jair Messias Bolsonaro (Bolsonaro) became president of Brazil. A normal legislative term in Brazil lasts for 4 years but in the case of Bolsonaro we will show the data of the first year of mandate (until Dec/2019). The table also indicates the total number of parties N_{par} elected to the Chamber of Deputies [19] and an effective number of parties N_{par}^{eff} [20], which takes into account the number of congressmen per party and gives a better idea of the fragmentation of the congress. This effective number of parties is defined as: $N_{par}^{eff} = 1/\sum_i p_i^2$, where p_i is the proportion of seats the ith party has. The N_{par}^{eff} is equal to the actual number of parties if every party has the same size and is closer to one if most congressmen belong to only one party. These numbers are calculated with the parties assignments in the moment of the elections results.

We close this section with a remark about the ending dates shown in Table 1. In the legislative terms where there was an impeachment (Collor–Itamar and Dilma–Temer), the ending date corresponds to the moment at which the

Table 2
Schema of structure of the data. Each roll call is represented by r^l and congressmen are represented by d_i. For each deputy, we know a list of his/her votes in each roll call: $\mathbf{o}_i = (o_i^1, o_i^2, \ldots, o_i^{N_{rc}})$ and o_i may be 5 different options, or the congressman may be absent, as explained in the text. For each deputy there are also other information as for example the party P_i to which he/she belongs at the moment of the roll call, reminding that we consider that the congressmen belong to the party for which they voted for the first time.

deputy	r^1	r^2	...	$r^{N_{rc}}$	party
d_1	o_1^1	o_1^2	...	$o_1^{N_{rc}}$	P_1
d_2	o_2^1	o_2^2	...	$o_2^{N_{rc}}$	P_2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	...
d_{N_d}	$o_{N_d}^1$	$o_{N_d}^2$	...	$o_{N_d}^{N_{rc}}$	P_{N_d}

president was removed form her/his position. According to Brazilian law, the president is removed from office once the impeachment process is started and the official judgment of the impeachment happens three months latter by the Senate, when she/he might be brought back if declared not guilty or be officially removed otherwise.

2.2. Data format

Brazil has excellent transparency laws which have been very well implemented and nowadays it is fairly easy to access huge amounts of data concerning the public administration and the legislative [21]. In this work, we use data available from the Brazilian Chamber of Deputies, concerning roll calls in the national congress [22,23].[1] From the application programming interface (API) developed by the I.T. personnel working in the congress, one can obtain a list of roll calls voted in a given year and the votes cast by the congressmen in each open roll call.

In Table 2 we show schematically the type of data we obtain from this database. Each roll call is represented in Table 2 by r^l, where $l = 1, \ldots, N_{rc}$ where N_{rc} is the total number of roll calls voted during a given legislative period. The congressmen are represented by d_i, where $i = 1, \ldots, N_d$ where N_d is the total number of congressmen. For each congressman d_i one has a sequence of votes which can be represented by a vector of options $\mathbf{o}_i = (o_i^1, o_i^2, \ldots, o_i^{N_{rc}})$ where each $o_i^l = v_j$ can assume 5 different values: v_1 =Yes, v_2 =No, v_3 =Abstention, v_4 =Obstruction and v_5 =Art.17[2] (a given congressman may also be absent from a given roll call, an issue which will be discussed later).

For each congressman, there is the information about his party affiliation and the federal unity she/he represents. Some roll calls also contain information about how the government, coalitions and the parties oriented their congressmen to vote.

It is important to point out some characteristics of these data to understand the limitations of our study. The roll calls whose votes are registered in the database [22,24] are called "nominal" or open. These are the ones we evaluate and they represent less than 20% of the total vote sessions in the congress; more than 80% of the votes in the plenary are secret and only the results are made public.[3] Moreover most of the bills are never put to vote and this is related to the fact that mostly, it is the President of the Chamber of Deputies who decides which bills are voted and in which order. This choice is clearly not random, but is subject to political calculation. This introduces a bias in our analysis that is a common feature of many roll call analysis studies [25].

We end this section with three considerations about our choice of analyses. In this work we consider all roll calls of the database as equally important, without giving weight to different types of projects or the subject tackled in the bills and roll calls. Considering the number of congressmen, there are 513 congressmen elected for each legislative term. Some of them are nominated as ministers or any other functions and are replaced by an alternate congressman. Therefore, in practice, there are more than 513 congressmen who vote along the four years of mandate, as Table 1 indicates. Some congressmen participate in only a few roll calls and were excluded from our analysis with a criterion that is explained in the Appendix A. This question of congressmen migration and participation in the roll calls is studied in detail in [26]. Also, congressmen can change parties during a legislative period. In this work we assume that a deputy belongs to the party for which he voted differently from "Absent" for the first time.

2.3. Measurements

In this subsection we define the quantities used in this work to analyze the legislative data.

[1] From [23], one may use the APIRESTful or click in "*arquivos*" (files) and directly download in "*Votações*" (votes) a list of all votes done in a given year, then in "*Voto de cada parlamentar*" (vote for each congressman) obtain the vote cast by each politician in each session for every year.

[2] This last option is reserved for a small fraction of congressmen that compose the presiding table at the plenary.

[3] From all roll calls voted per year that one can download from [23], only in around 20% of the cases one is able to obtain the list of votes by congressman. The real number of secret and open roll calls might be different since the database itself claims not to be completely up to date.

2.3.1. Cohesion in the roll calls

As mentioned above, there are 5 different ways a congressman may vote besides being absent in a roll call. These different alternatives do point to different strategies the government, opposition and the different parties may be adopting given the bill at hand. There are of course the yes and no alternatives indicating support to approve or dismiss a given bill, but other alternatives like obstruction or abstention may indicate a push to either postpone or alter the bill.

In order to evaluate the cohesion of a given party in a roll call, and therefore to assess whether the party is cohesively following a given strategy, we adopt a concept from information theory: the Shannon entropy [27], which measures the uncertainty in a distribution. Given that a fraction q_i of the congressman from a given party P_j voted option v_i in a given roll call, one can evaluate

$$H_j = -\sum_{i=1}^{N_V} q_i \ln q_i \tag{1}$$

$$C_j = 1 - \frac{H_j}{\ln N_V}, \tag{2}$$

where H_j is the Shannon entropy for the distribution of q_i's, which assumes values between 0, when all congressmen in the party P_j voted the same option (zero uncertainty), and $\ln N_V$ when the congressmen voted evenly among the N_V different options (in a given roll-call $1 \leq N_V \leq 5$) which received at least one vote by congressmen of party P_j (maximum uncertainty). Note that the value of C_j, which we will be calling cohesion, will be between 1 (cohesive strategy adopted) and 0 (party most divided).

Given a roll call, one can evaluate two global cohesions associated to it, which are going to be called effective cohesion C^{Eff} and party cohesion C^{Party}. The effective cohesion is evaluated using the q_i distribution in a given roll call without regard to the parties, just evaluating the total number of congressmen that voted each option. For evaluating the party cohesion, given a roll call, first for each party the particular party cohesion C_j is evaluated, and the total roll call party cohesion is then the weighted average of the parties cohesions:

$$C^{\text{Party}} = \frac{\sum_j C_j n_j}{\sum_j n_j}, \tag{3}$$

where the sum runs over all parties that participated in the roll call, C_j is the cohesion of party P_j in the roll call and n_j is the number of congressmen from party P_j that participated in that roll call.

2.3.2. Agreement between congressmen

To quantify how similar are the sequences of votes between two congressmen $\mathrm{d_i}$ and $\mathrm{d_j}$, we define the agreement as:

$$A_{\mathrm{d_i},\mathrm{d_j}} = \frac{1}{N_{\mathrm{d_i},\mathrm{d_j}}} \sum_{l}^{N_{rc}} \delta^{l}_{\mathrm{d_i},\mathrm{d_j}} \tag{4}$$

where l is the index of the roll call, $\delta^{l}_{\mathrm{d_i},\mathrm{d_j}}$ is a Kronecker delta such that, $\delta^{l}_{\mathrm{d_i},\mathrm{d_j}} = 1$ if congressmen $\mathrm{d_i}$ and $\mathrm{d_j}$ vote with the same option in the roll call l and $\delta^{l}_{\mathrm{d_i},\mathrm{d_j}} = 0$ otherwise, $N_{\mathrm{d_i},\mathrm{d_j}}$ is the total number of roll calls that both congressmen $\mathrm{d_i}$ and $\mathrm{d_j}$ voted differently than Absent. With this definition, $A_{\mathrm{d_i},\mathrm{d_j}} = 0$ if both congressmen voted completely different or if they are never present simultaneously in any roll call and $A_{\mathrm{d_i},\mathrm{d_j}} = 1$ if they have exactly the same sequence of votes in all roll calls they both participated in a given legislative period.

2.3.3. K-Means to identify groups in the Chamber of Deputies

In order to identify groups of congressmen who voted in a similar way, we use the k-means clustering algorithm [28]. Below we succinctly explain how this method works.

Assume we have m observations lying in an N-dimensional space

$$obs_i = (obs_{i,1}, obs_{i,2}, \ldots, obs_{i,N}) \tag{5}$$

where each component $obs_{i,j}$ is the jth feature of the ith observation.

The point of the algorithm is to cluster observations into k clusters and it proceeds as follows:

1. randomly select k observations. These will be called the centroids μ_n $(n = 1, \ldots, k)$ of the clusters G^k_n
2. assign each observation to the cluster whose centroid is closest to it.
3. update the centroids as the mean point of the observations belonging to the cluster.

Repeat the steps 2 and 3 until the change in the centroids positions is below a given precision. This process will cluster the points which are closer together by minimizing the intra-cluster sum of squared-distances (also called inertia) [28],

$$I = \sum_{i=0}^{m} \min_{\mu_n} \left[(obs_i - \mu_n)^2 \right]. \tag{6}$$

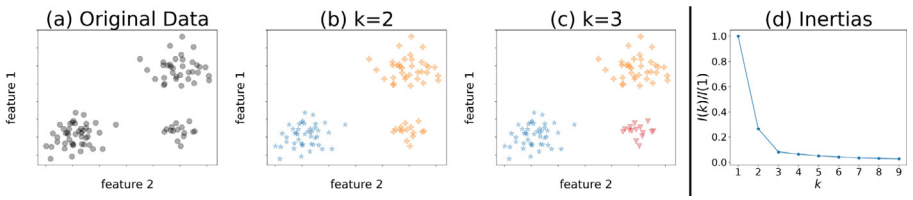

Fig. 1. Example of the k-means algorithm applied to 2-dimensional data. Each point represents one observation, with each axis representing one feature, as defined in Eq. (5). In **(a)** it is shown the original data, in **(b)** points are clustered in $k = 2$ groups and in **(c)** the algorithm is applied for $k = 3$. In (b) and (c) different groups identified by the algorithm are identified by different colors. **(d)** presents the relative inertia $I(k)/I(1)$ for different values of k. There is no significant decrease in inertia when we go from $k = 3$ to $k = 4$, which indicates, by the elbow method, that 3 is the ideal number of clusters as would be intuitively expected by visual inspection of (a).

In our case, each observations will be the vector

$$A_i = (A_{\mathrm{d_i d_1}}, A_{\mathrm{d_i d_2}}, \ldots, A_{\mathrm{d_i d_{N_d}}})$$

which characterizes the agreement of the $\mathrm{d}_\mathrm{i}^{th}$ congressman to all others, such that $m = N_\mathrm{d}$ and also $N = N_\mathrm{d}$ (that is, the N_d observations lie in a N_d-dimensional space) and the agreement to another congressmen d_j is treated as a feature. The absolute centroid position of each group is not relevant here but, after running the algorithm, each congressman has a label indicating to which group he/she belongs i.e. to which other congressmen he/she is more afine. In Fig. 1 there is a pictorial representation of the results of the algorithm for a set of random points generated for two arbitrary features ($obs_i = (obs_{i,1}, obs_{i,2})$).

A caveat of this method is that the number of clusters is an input, not an output. So, for the same data, one may run the algorithm with different values of k obtaining different results. A criterion to assess the best number of clusters is the elbow method which seeks to identify the value of k for which adding a new cluster does not significantly reduces the inertia (variance) I for the distribution of the data points around their assigned centroids [29].

2.3.4. Alignment of the clusters with the president's party

Once we define a group of congressmen G_n^k using the k-means algorithm, we can quantify the degree of alignment or support of these congressmen with the president's party P_p by defining the following quantity:

$$\mathcal{S}_n^k = \frac{1}{|G_n^k||P_p|} \sum_{\mathrm{d_j} \in G_n^k} \sum_{\mathrm{d_i} \in P_p} A_{\mathrm{d_i,d_j}}, \tag{7}$$

where $|G_n^k|$ and $|P_p|$ are the total number of congressmen in the group G_n^k and in the president's party P_p, respectively. To avoid considering dissident congressmen as proper representatives of the president, we consider as belonging to the president's party only those congressmen who stayed in this party during the whole presidential term. Note that it is just an average of the agreement over pairs of congressmen, one belonging to a cluster G_n^k and the other to the presidents party P_p. This quantity is expected to be closer to 1 if the group G_n^k has parties aligned politically with the government and is expected to be smaller if G_n^k has congressmen who belong to the opposition to the government.

The $\mathcal{S}_n^k$ can be evaluated for different intervals of time in a given legislative term in order to study its change and behavior along the time. The dynamics of this measure over time is a way to identify possible instabilities of a presidential term.

3. Analysis of the legislative activity data

In this section we present the results of the analyses using the data for the roll calls in the Chamber of Deputies.

3.1. The roll calls in cohesion space

As explained in Section 2.3.1, one can associate to each roll call two measures of cohesion, the effective cohesion C^{Eff} and the party cohesion C^{Party}. These cohesion parameters define, therefore, a two dimensional phase-space where one can locate each roll call.

In Fig. 2 it is shown the scatter plot of the roll calls for all the different legislative terms in this cohesion phase space. In this plot, each roll call is a point in the $C^{\mathrm{Eff}} \times C^{\mathrm{Party}}$ space. Points on the diagonal line indicate that the average cohesion inside the parties is the same as the cohesion on the whole National Congress for a given roll call. Points below the diagonal indicate the roll calls that have higher cohesion inside of the parties than in the whole congress, indicating that these are the votes polarizing the parties in opposite directions. From the figure it is possible to observe that effective and party cohesion seem to be correlated, one tends to be high when the other is high as well. We should note here, that

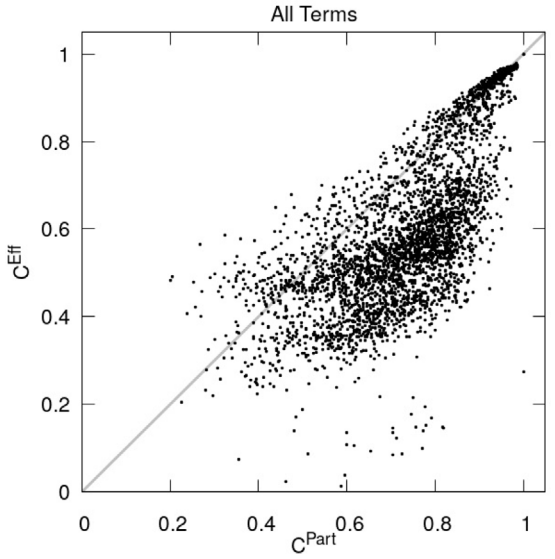

Fig. 2. Roll calls in cohesion phase space for all periods superposed. The main diagonal line is traced in light gray color.

this correlation is not evident from the definitions of these parameters. One is measuring the parties intrinsic cohesion and the other the overall result of a roll call. Were the congress evenly divided into two strongly opposed parties, a high party cohesion would imply a low effective cohesion for one party would systematically vote the opposite option than the other, meaning that the party cohesion would be high, but the effective cohesion low because the overall result of the vote would be divided. Therefore, it is interesting to observe here that very consensual issues in the congress tend not only to unite intrinsically the parties, but also the parties among them and dividing issues, tend not necessarily to oppose parties against each other, but rather to divide the parties intrinsically.

In the Appendix B we show these results individually for each term where one can see that the same pattern repeats in all legislative periods, as we explain now. A dense group of points lay on the diagonal line on the top right in the scatter plots. These are roll calls for which there is a very high cohesion inside of the parties and in the congress itself. These roll calls correspond to subjects that are consensual and therefore are not useful to distinguish different ideologies among parties. On the other hand, more diffuse points all over this cohesion space represent roll calls for which there is less cohesion. Because these points appear for smaller values of C, they indicate that, both deputies inside a party or of the whole congress vote in a less cohesive way. These can be seen as more controversial roll calls both among congressmen of a given party and among all the congressmen of the National Congress.

3.2. Distribution of the agreement between congressmen

Fig. 3 shows the distribution of A_{d_i,d_j}, as defined in Section 2.3.2, during the mandates of the ten presidents considered in this work. By inspecting the distributions we identify that most of the terms present a bimodal distribution, the exception being the periods governed by Collor, Itamar, Dilma I and Dilma II. The periods Collor–Itamar and Dilma–Temer are referred to as *politically unstable* because in both there was an impeachment of the president.[4]

We will see that this observed bimodality during stable legislative periods suggests that a polarization into two opposing blocks seems correlated to the stability of the Brazilian political system. This is a curious result because Brazil has one of the most fragmented political systems in the world, having many political parties, as shown in Table 1. We point out that each separate peak in these distributions is not directly referring to the opposing blocks (government and the opposition), but the peak with high agreement contains the values of both groups together and the peak with low agreement the cross-terms (agreements between congressmen from different groups). This distinction will become clearer in Fig. 4-d, where these distributions are colored according to groups divided by the k-means.

A similar pattern was observed using a different type of measure of agreement in [16], where the authors fit a bimodal function to the distributions they find and associate the relative distance between the peaks to an indicative of instability.

To better characterize these periods of stability and instabilities, in the next sections we cluster congressmen in groups and analyze their behavior over time.

[4] Officially, Collor resigned before being impeached to avoid losing his political rights (a move similar as done by Nixon).

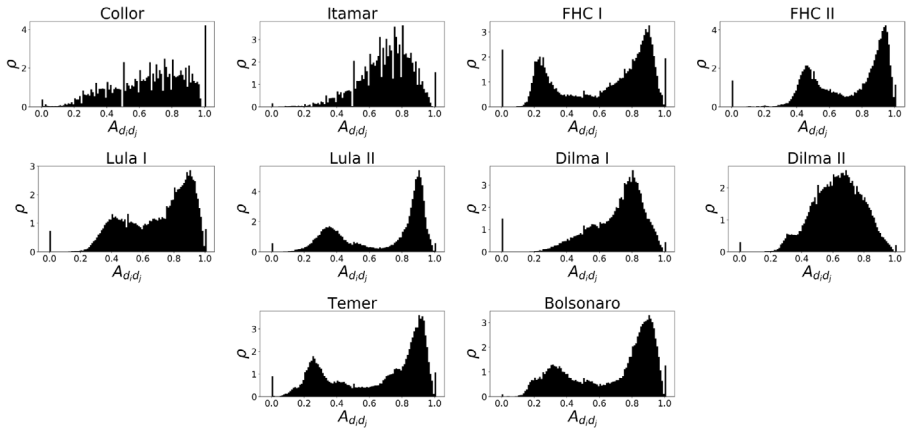

Fig. 3. Normalized distribution of the agreement matrix elements for each analyzed period.

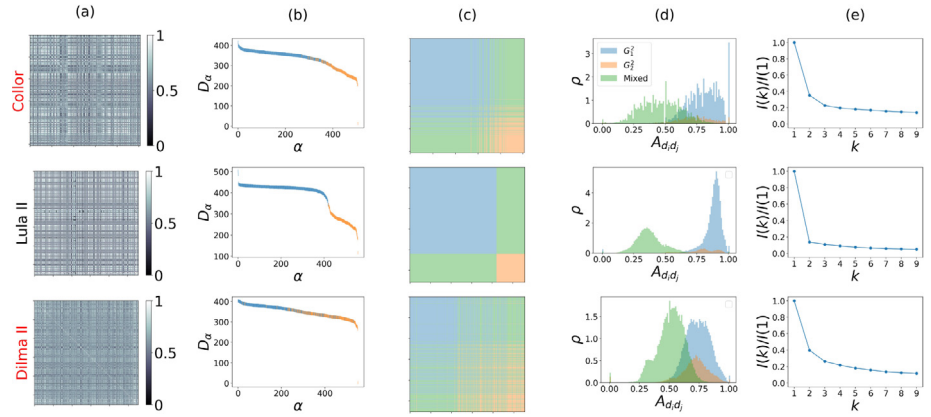

Fig. 4. For Collor, Lula II and Dilma II periods we show (a) the agreement matrix with the agreement according to the color-bar to the right, (b) ranked degrees centrality $\mathcal{D}_\alpha$ of the congressmen colored by which cluster they were assigned to (blue for G_1^2 and orange for G_2^2), (c) ordered agreement matrix with congressmen ranked by degree centrality and colored blue if both belong to G_1^2, orange if both belong to G_2^2 and green if one belongs to G_1^2 and the other to G_2^2, (d) distribution of agreement matrix separated by the same coloring criterion of (c), and (e) the relative inertia $I(k)/I(1)$ calculated for different k.

3.3. Discriminating groups in the Chamber of Deputies

In principle, a direct way to analyze the dynamics of groups in the congress is by dividing congressmen in their own political parties. For each party it is possible to quantify, for example, how aligned they are with the government and how faithful the congressmen are to their own parties. These quantities can be evaluated for different roll calls and give us a notion of how groups behave over time. In spite of being a natural choice of groups, this approach might be problematic for some reasons. First because there are many parties in the Chamber of Deputies, as shown in Table 1, and they have different sizes and influence in the political scenario. Second, because parties break up and congressmen change parties along the same legislative periods, an issue which is more thoroughly discussed in [26]. It can be shown that most of the movement of the congressmen happens between the parties of the opposition or between the parties of the ruling coalition, categories that, for our purposes, will be properly defined in the next section. Here also one might resort to the conclusions drawn from Fig. 2 and realize that those issues dividing the congress are, as well those that divide the parties. To avoid these problems, in the following we use the sequence of agreements between the votes of the congressmen as entries in the k-means clustering algorithm to analyze how congressmen effectively aggregate. It allows us to identify the important groups in the congress and how they evolve in time.

3.3.1. Identifying an effective ruling coalition and opposition for each presidential term

We use the k-means algorithm to divide the congressmen in groups. As explained in Section 2.3.3, this type of algorithm requires a pre-definition of the number k of groups into which one wants to partition a given sequence of observations.

Table 3
Absolute size of the clusters G_n^2 in terms of number of deputies (relative sizes G_n^2/N_d) and their support $\mathcal{S}_n^2$ to the president's party in the correspondent period.

President	G_1^2	G_2^2	$\mathcal{S}_1^2$	$\mathcal{S}_2^2$
Collor	356 (70%)	151 (30%)	0.84	0.4
Itamar	400 (79%)	106 (21%)	0.77	0.61
FHC I	440 (77%)	133 (23%)	0.81	0.31
FHC II	402 (71%)	161 (29%)	0.91	0.49
Lula I	412 (76%)	133 (24%)	0.87	0.4
Lula II	403 (75%)	133 (25%)	0.89	0.34
Dilma I	438 (79%)	115 (21%)	0.80	0.45
Dilma II	326 (63%)	195 (37%)	0.73	0.4
Temer	417 (78%)	120 (22%)	0.86	0.28
Bolsonaro	384 (74%)	133 (26%)	0.86	0.29

Table 4
Lula II — Crosstabulation of identified clusters G_n^2 and self declared (sd) party affiliation at the second round of the election.

	G_1^2	G_2^2	Total
B^{sd}	232 (98.7%)	3 (1.3%)	235
O^{sd}	22 (15.3%)	122 (84.7%)	144
Non Declared	149 (94.9%)	8 (5.1%)	157
Total	403 (75.2%)	133 (24.8%)	536

Table 5
Dilma II — Crosstabulation of identified clusters G_n^2 and self declared party affiliation at the second round of the election.

	G_1^2	G_2^2	Total
B^{sd}	275 (89.3%)	33 (10.7%)	308
O^{sd}	51 (24.5%)	157 (75.5%)	208
Non Declared	0 (0%)	5 (100%)	5
Total	326 (62.6%)	195 (37.4%)	521

For each presidential term, we consider $k = 2$ and divide all N_d congressmen in two groups,[5] called G_1^2 and G_2^2. For each of these groups, we quantify their support to the president's party by using the quantity $\mathcal{S}_n^k$ defined in Section 2.3.4. Table 3 shows the number of congressmen identified in each group, the proportion of these groups in terms of the total number of congressmen N_d and how each group aligns with the president's party.

We observe in Table 3 that the group G_1^2 corresponds at least to 70% of the total number of congressmen in all terms, except in the case of Dilma II. Moreover, the support quantity of this group $\mathcal{S}_1^2$ is around 0.8 in all terms, which indicates that this group is aligned with the president's party. This suggests that this majority group can be associated to a *ruling coalition* with the president. The minority group G_2^2 has alignment with the president's party $\mathcal{S}_2^2$ much smaller than $\mathcal{S}_1^2$. We associate this minority group with an *effective opposition* to the government.

To check how much these effective ruling coalition G_1^2 and effective opposition G_2^2 coincide with an officially declared base and opposition to the government, we identified the legislative terms for which the election occurred in two rounds. When this is the case, in the second round, many parties organize themselves into a coalition, declaring publicly support to one or the other candidate. The parties which belong to the coalition that won the election we identify as *self-declared ruling coalition* B^{sd}, while parties in the coalition that lost the election are identified as *self-declared opposition*, O^{sd}. Some parties do not declare any position before the election and were identified as *non-declared*. Tables 4 and 5 show, for two legislative terms, the fraction of congressmen that belong to a party of the B^{sd} that is identified as well as either an effective basis (G_1^2/B^{sd}) or effective opposition, (G_2^2/B^{sd}), and the same for the self declared opposition (G_n^2/O^{sd}) and non-declared position. Observing the diagonal of these tables one can see that G_1^2 corresponds to 98.7% for Lula II and 89.3% for Dilma II of the self-declared basis B^{sd}. G_2^2 coincides with 84.7% in the case of Lula II and with 75.5% for Dilma II of the self-declared opposition. These high values occur for all the legislative terms for which we could do this analyses and they justify our interpretation of the groups encountered by the k-means algorithm as ruling coalition and opposition. We remark that the boundary between clusters generally respect the boundaries between parties (that is, parties are usually contained inside clusters).

[5] The choice of $k = 2$ is made according to the criterion based on the measurement of the inertia for different values of k and choosing the one for which the addition of new clusters does not significantly reduce the inertia (elbow rule).

3.3.2. Interpretation of the agreement between congressmen

In this subsection we use the separation in two groups discussed in the previous subsection to interpret the peaks in the distribution of agreements and how good this separation is to identify different behaviors in the Congress. From this point forward we will focus on Collor, Lula II, and Dilma II terms. Collor and Dilma II were the periods in which there were impeachment processes and Lula II will be taken as a typical stable period. Fig. 4 shows the data for these three legislative periods and the equivalent results for the other periods can be seen in Appendix B. We remind the reader that Dilma II is defined as the period before the beginning of impeachment process in the legislative period called Dilma–Temer while Collor is the one before the impeachment in Collor–Itamar. The exact dates used to separate the two periods are shown in Table 1.

Fig. 4-a shows the matrix of entries A_{d_i,d_j} as a function of the indexes of the deputy d_i and d_j. We then measure the degree centrality of the d_i^{th} congressman, defined as $\mathcal{D}_{d_i} = \sum_{d_j} A_{d_i,d_j}$ and introduce a label α to rank the congressmen in descending order of $\mathcal{D}_\alpha$, $\mathcal{D}_1 > \mathcal{D}_2 > \cdots > \mathcal{D}_{N_d}$. This ordered quantity $\mathcal{D}_\alpha$ as a function of α is shown in Fig. 4-b. The blue colors shown in this figure represents the congressmen of the group G_1^2 and orange represents the congressmen in G_2^2. We use this rank defined by D_α to reorder the matrix of agreements presented in Fig. 4-c. The colors in this reordered matrix are defined using the following criterion: if two congressmen d_i, d_j belong to the same cluster G_1^2, A_{d_i,d_j} is colored blue; if both congressmen belong to cluster G_2^2, they are colored orange; and if each one belongs to a different group, they are colored green and referred to as "mixed term". This same notation of colors is used to represent the distribution of A_{d_i,d_j} for these tree cases in Fig. 4-d. In this case the distributions are normalized in such a way that the sum of the areas under all the curves together is one, but observing that the area below each colored distribution respect the size (proportion) of its respective group. Our sorting strategy could be contrasted with the one adopted in [16] which is to use the Prim algorithm to determine the minimum spanning tree for their set of distances (agreements). This strategy considers only the strongest agreements between congressmen, while ours is taking into account all information (all agreements) besides being much simpler and yet capable of achieving similar results. In Fig. 4-e we show the relative inertia $I(k)/I(1)$, as calculated by Eq. (6).

We can now critically analyze Fig. 4 to observe differences between these legislative terms. In Figs. 4-b, we observe that in Lula II period there is a step in the quantity $\mathcal{D}_\alpha$, and the two levels of the step correspond to the two groups classified by the clustering algorithm, as can be seen by their colors. On the other hand, in Collor, and Dilma II there is a continuous decay of $\mathcal{D}_\alpha$, where these two groups mix together. This presence or absence of a step in this quantity is a manifestation of how good a division in two clusters is. Ordered matrices represented in Fig. 4-c show a very well defined division in two groups during Lula II, but in Collor and Dilma II, the division in two groups is not enough to capture the complexity of the groups' organization in the congress. Fig. 4-d shows that, during Lula II period, the blue color that correspond to the effective ruling coalition, is large, while the orange group, which was shown to be an effective opposition, is very small. The green group, which are the mixed terms, has an intermediate size in this period, while is the most important distribution in Dilma II and almost as important in Collor. This mixed group has a peak at smaller values of A_{d_i,d_j} and this is just saying that congressmen from two different groups have smaller agreement in their sequence of votes, while for the effective ruling coalition and opposition the average agreement is high inside of each group, which indicates strong cohesion among them. One can see in Fig. 4-e, the inertia curves, a distinctive difference when comparing Lula II, a stable period, with Collor and Dilma II, unstable periods. In Lula II, one sees a sharp edge at $k = 2$ after which it goes down much slower. On the other hand, in Collor and Dilma II, the curves are smoother and there is no obvious breaking point. As prescribed by the elbow criterion, this indicates that $k = 2$ is a good choice for Lula II period, but not as good for Collor or Dilma II, which might need more clusters to be properly characterized.

To summarize, our analyses have shown that there are two standard legislative periods with their own characteristics: a stable period and an unstable period. During stable political periods, the government formed an effective ruling coalition which contained at least 70% of the congressmen and this group presented a strong support to the president's party. In unstable legislative terms, characterized by an impeachment, the separation into two effective groups was not enough to discriminate the matrix of agreement between congressmen completely. In the next section we propose to analyze how these two groups (G_1^2 and G_2^2) evolve along time in the different legislative terms and also a further analysis in terms of considering more than two groups to capture what happens in unstable periods.

3.4. Time evolution of the groups in the Chamber of Deputies

An average measure over the whole legislative terms might not allow one to understand what happens before an impeachment or how this is compared to a politically stable legislative term. To verify how the support to the government of different groups behaves along time within a legislative term, we use the following procedure. We first cluster groups of congressmen using k-means. Once we have these groups defined, we separate the roll calls in windows of four months and, for each period, we measure the support S_n^k of each group to the president's party as defined in Eq. (7). The result of this measure in time is shown in Fig. 5 for the Collor–Itamar, Lula II and Dilma–Temer periods (these graphics for all the other periods can be found in Fig. B.10 at Appendix B). Figures on the left show the cases where the congressmen are segregated in two effective groups. We emphasize that, for Lula II term, groups were defined using the whole presidential term, while for the Collor–Itamar and Dilma–Temer, groups were defined using the period before the impeachment (the exact dates are shown in Table 1). Also, for these unstable periods, it is represented the alignment with the party of first

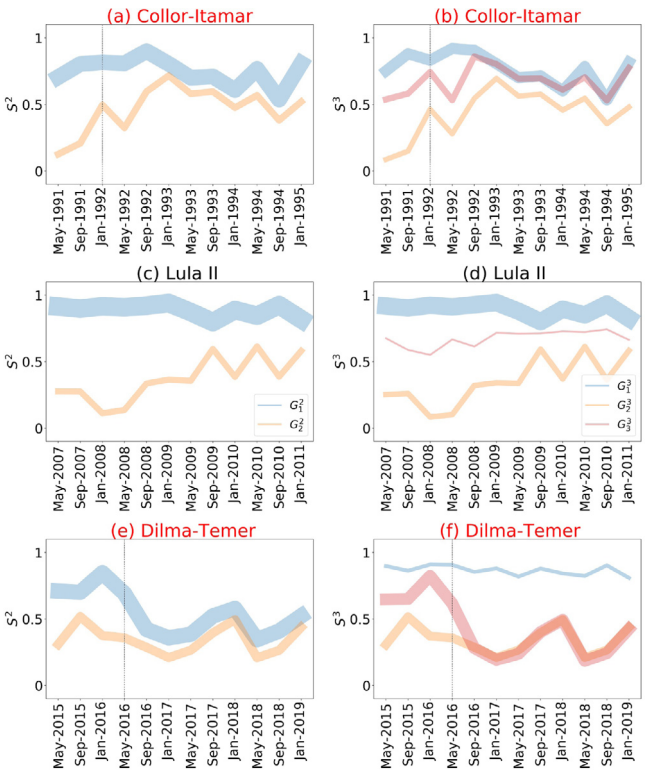

Fig. 5. Evolution of clusters support to the president's party. The clusters were found with k-means using the whole period except for the Collor–Itamar and Dilma–Temer (highlighted in red) in which they were found using rollcalls from the beginning of the period up to the beginning of his impeachment process, indicated by the vertical line. The width of the curves are proportional to the cluster sizes. The legend in Lula II periods indicates the group name for all the periods.

president of the term, namely Collor and Dilma. In these figures, the line width of each group is proportional to their size and the notation of colors is the same as in the previous section: blue denotes effective ruling coalition, G_1^2, and orange the effective opposition, G_2^2. Vertical lines in Fig. 5-a, b, e and f indicates the end of Collor's and Dilma Rousseff's mandate (i.e. the moment they were removed from government because the impeachment process had begun).

One can observe that, during the whole period Lula II, which is typical of a stable mandate, the effective ruling coalition keeps its support to the president's party and that the effective opposition has a smaller support value that increases at the end of his mandate. This behavior is observed in the other politically stable legislative terms, as well, and is in line with the observation in a recent work that the polarization decreases as the end of mandate is approached [16]. In the Dilma–Temer period, the situation is very different: i) the support and the size of the effective ruling coalition is smaller than in the Lula II period and ii) the group associated to an effective ruling coalition in the Dilma's government decreases its support S_1^2 to a value comparable to the one of the effective opposition S_2^2 after the impeachment. In Collor–Itamar, the ruling coalition keeps aligned to the president's party for the whole period. One may wonder why these two impeachments look like so different and the reason for it is the fact that in Collor's case its own party voted against him (21 out of 29 congressmen belonging to his party voted favorable to his impeachment), while Dilma's party stood by her (none voted favorable to her impeachment). This means that the ruling coalition kept aligned to the president's party, but the party itself was against the president, in Collor's case.

To better understand this behavior we repeat the previous analysis, but this time using k-means with $k = 3$ to separate the congressmen in clusters G_1^3, G_2^3 and G_3^3. For each of these groups, we measure their support to the president's party as a function of time and show the results in the right column of Fig. 5. In these figures, red is the third group identified in this process. As previously, the line width is proportional to the group sizes. In the period Lula II, it is very small and its support S_3^3 stays in between the ruling coalition and opposition. This pattern is similar in all stable periods as can be seen in the Appendix, Fig. B.10. However, for the two terms with an impeachment, this third group G_3^3 is bigger than or at least comparable with the effective ruling coalition and it aligns with one of the other groups. In the case of the Collor term, it aligns itself with the ruling coalition, and in the case of Dilma, with the opposition. As mentioned above, we emphasize that Collor lost the support of his own party. Then, the similarity of both impeachments is that the third group aligned itself with the group against the president.

Table 6 summarizes the size of each group G_n^k and its alignment to the president's party S_n^k for the eight presidential terms (ten presidents). We note that both the group size *and* its alignment needs to be relatively high to guarantee the

Table 6
Absolute sizes of clusters G_n^3 (their proportion) and their support $\mathcal{S}_i^2$ to the president party in the correspondent period.

President	G_1^3	G_2^3	G_3^3	S_1^3	S_2^3	S_3^3
Collor	250(49%)	118(23%)	139(27%)	0.89	0.35	0.67
Itamar	312(62%)	100(20%)	94(19%)	0.8	0.61	0.65
FHC I	354(62%)	117(20%)	102(18%)	0.84	0.28	0.66
FHC II	387(69%)	134(24%)	42(7%)	0.92	0.46	0.66
Lula I	320(59%)	121(22%)	104(19%)	0.9	0.38	0.72
Lula II	376(70%)	122(23%)	38(7%)	0.9	0.32	0.7
Dilma I	138(25%)	115(21%)	300(54%)	0.9	0.45	0.76
Dilma II	82(16%)	191(37%)	248(48%)	0.89	0.39	0.67
Temer	363(68%)	107(20%)	67(12%)	0.89	0.26	0.66
Bolsonaro	371(72%)	101(20%)	45(9%)	0.87	0.23	0.54

president stability. One can see that Dilma I and Dilma II periods stand out by the small size of G_1^3 and large size of G_3^3 (though G_2^3 is also large in Dilma II). In this table we do not yet observe a small value of S_3^3 for Dilma II period because the support of the G_3^3 only decays significantly very close to the impeachment process, as shown in Fig. 5.

4. Summary and discussion

Using data from all nominal roll calls occurred in the Brazilian Chamber of Deputies from 1991 to 2019, after locating the roll calls in a cohesion phase space, we computed how similar is the sequence of votes between two congressmen d_i and d_j and refer to it as "agreement between congressmen", A_{d_i,d_j}. Using this quantity as input data of the clustering algorithm k-means, we separate the congressmen in k groups identified as G_n^k. For each of these groups, we defined S_n^k that quantifies how strong is each group's support to the president's party. We then measure how this quantity evolves in time for each legislative term.

From the assessment of the cohesion of the vote results in the roll calls, we observe that effective and party cohesion's seem to be correlated. This means that issues that unite the congressman from a same party are those that also unite the parties around the same cause and issues that divide the congressman are not pitching only parties with opposing ideas against each other, but also causing some kind of dissension within the parties themselves. This observation calls into question whether parties are an effective means for the electorate to associate its interests in a cohesive group of representatives. One does observe from the data, though, some kind of coherent behavior in the congress in respect to the mater of forming a ruling coalition.

Since 1991, Brazil has had eight direct presidential elections and in two of them there was an impeachment of the elected presidents. The periods without an impeachment are referred to as stable political terms in contrast with the terms with an impeachment, referred to here as unstable. Our analyses show some differences between these two types of periods as we now discuss and summarize in Fig. 6.

Stable political periods can be characterized as follows. The distribution of A_{d_i,d_j} presents two peaks, while a continuous, unimodal distribution is observed in unstable ones, legislative periods that have had an impeachment (Fig. 6-a). The exceptions are Dilma I in the term before her impeachment, where we observe unimodality in this distribution and Itamar, the term that followed Collor's impeachment (both terms are shown in Appendix B) . Although Dilma was not impeached in her first mandate, this unimodality there might indicate that the instability that lead to her impeachment in the second term had already began in her first term. When analyzed in terms of inertia I, stable terms are also characterized by a strong prescription for a $k = 2$ number of clusters, while unstable periods are less clear by that criterion, Fig. 6-c. This coincides with the ranked degree centrality, which shows two steps, characteristic of each cluster, and makes the groups explicit when used to order the agreement matrix, Fig. 6-b. When congressmen are divided in $k = 2$ groups called G_1^2 and G_2^2, it is observed that G_1^2 is composed by at least 70% of the congressmen (Table 3), Colored distributions shown in Fig. 6-a allow us to identify the origin of the peaks in the distribution of A_{d_i,d_j}: one of the peaks comes mostly from the terms relating two congressmen from G_1^2 with a small contribution from terms relating the congressmen from group G_2^2 among themselves, while the other comes from the mixed terms, which relate congressmen from G_1^2 to G_2^2, and not one peak for each cluster, as a naive guess might lead.

Fig. 6-right summarizes the support to the president's party S_n^k over time, allowing us to observe differences between the two types of legislative terms. When looking at the evolution of S_n^k it is observed that i) G_1^2 presents a high support for the president's party (Table 3) and ii) for stable periods, this support is high along the whole legislative term. This biggest group with a high support to the president's party is associated with an effective ruling coalition, which correlates well with a self-declared coalition in cases where this comparison is possible (Tables 3 and 4). All this suggests that the stability of the presidents requires a big and cohesive effective coalition during the whole mandate. Our data show that "big" means at least 70% of the congressmen, although we do not have a theory to assume that this is a minimum value. Moreover, we analyzed the stability of this division in two groups by dividing them in $k = 3$ (Fig. 5-right). We observe that, for stable legislative terms, the third group is very small (from 7% to 20% of the congressmen - Table 6) and it

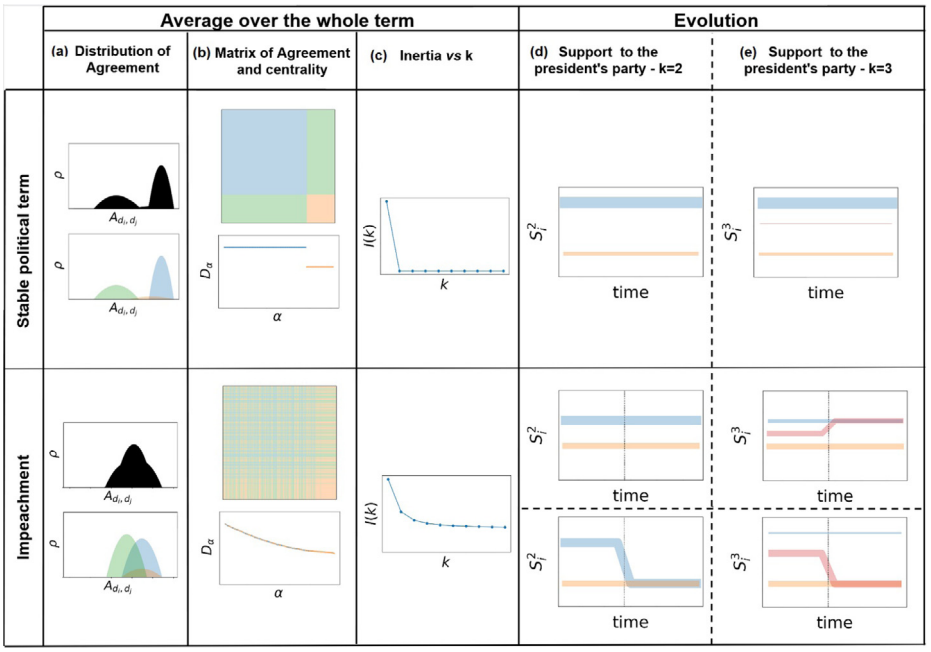

Fig. 6. Schematics of the two types of regimes with their main characteristics. One type of regime is characterized by a political stability and other by an impeachment of the elected president. On the left, there are measures of the average over the whole legislative terms, which show that two clusters characterize well the Chamber of Deputies only for stable periods. On the right it is shown how the support to the president's party evolve in time, where we see that stable periods are largely homogeneous in time, while there is a sudden change in unstable one around the impeachment. For stable and unstable periods we show **(a)** the distribution of agreement before (upper, in black) and after being separated by the k-means with $k = 2$ (lower), **(b)** the agreement matrix ordered by degree centrality (shown below the matrix) and both colored according to the clustering criterion described in the text, **(c)** how inertia varies for different values of k and how the clusters' support to the president's party evolve for **(d)** $k = 2$ and **(e)** $k = 3$.

is composed by congressmen from the two other groups G_1^2 and G_2^2. Also, this group presents an intermediary value of support to the president's party slightly in favor of it, but in unstable term, this third group is bigger and aligns itself against the president during and after the impeachment process.

The unstable terms, from the perspective of the evolution of the support to the president's party, presents a relatively bigger G_3^3 that eventually aligns itself with one of the other groups. Dilma's impeachment in her second term is characterized by the decay of the S_1^2 along the legislative term. When the congress is separated into 3 groups, one clearly observes that this coalition represented by the group G_1^2 which is composed by 63% of congressmen splits into two groups: a very small group G_1^3 with around 16% of the congressmen that keeps its support and the other part that aligns with the opposition, the group G_3^3, with 48% of the congressman (Fig. 5). These 16% of congressmen corresponds to her own party and another one; all other parties migrate to an opposition. Collor was also impeached and does not present a decay of S_1^2, but this is because the support we are able to measure is for the president's party and not to the president itself and Collor lost the support of his own party. To exemplify this statement, it is easy to verify that more than 70% of the congressmen of his own party voted in favor of his impeachment.

Our methods should be readily adaptable to other regimes as long as there are nominal roll call data available and a head of government such as a president or prime minister, with a subset of congressmen that can be interpreted as his or her main representatives. In that regard, one advantage of our methods is that they are independent of the number of voting options available to the congressmen and their interpretation. It remains to be seen how different party structures and congress dynamics emerge in other parliaments (other countries).

5. Conclusion

Thirty years ago the concept of coalitional presidentialism was proposed [30] to understand the stability of presidential regimes with many parties. The idea is simple: if, in a multiparty presidential regime, a directly elected minority president cannot form a majority in the Parliament with only his/her own party, a coalition between many parties is necessary to build a durable government. [31–34]. Moreover, it was also shown that the projects proposed by the president's party have a very high rate of approval in the Brazilian congress [35]. These results are in line with our observations during stable periods. We identified that we can classify the presidential terms in two categories that we called stable and unstable based on their characteristics such as uni- or bimodality of the distribution of agreement between congressmen, shape of

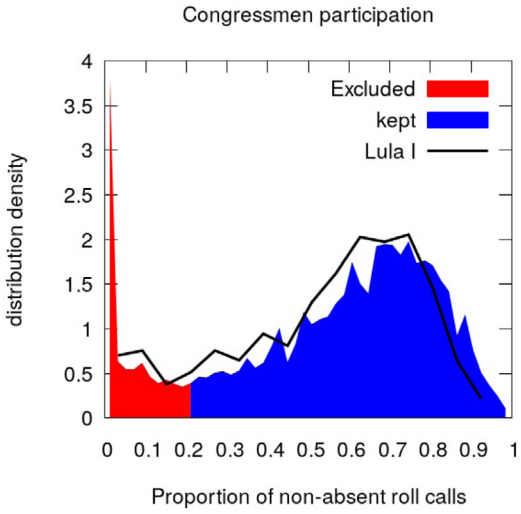

Fig. A.7. Distribution of the number of non absent votes for the congressmen in all periods. In red are the excluded congressmen, in blue the kept ones. Lula I mandate is shown individually for comparison.

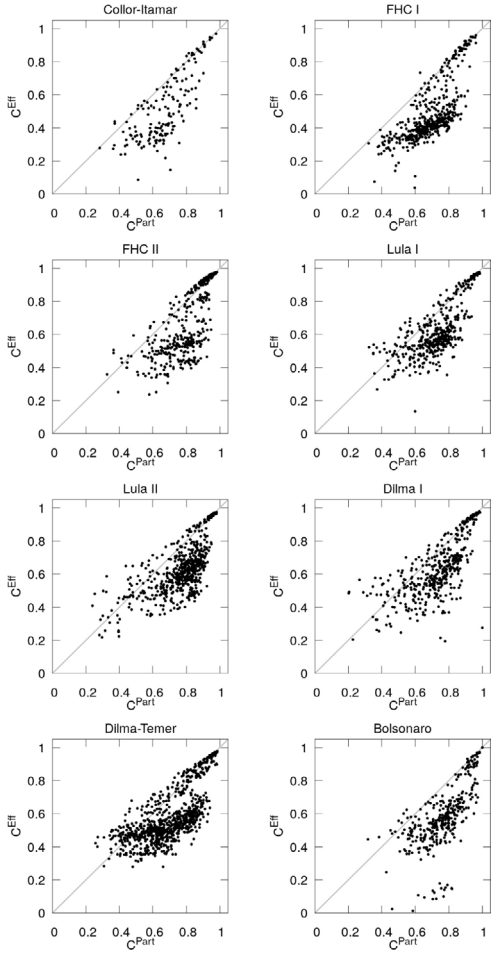

Fig. B.8. Roll calls in cohesion phase space. The plots in the two columns show each legislative period separately and the last one below has all roll calls for all periods superposed. In all plots the main diagonal line is traced in light gray color.

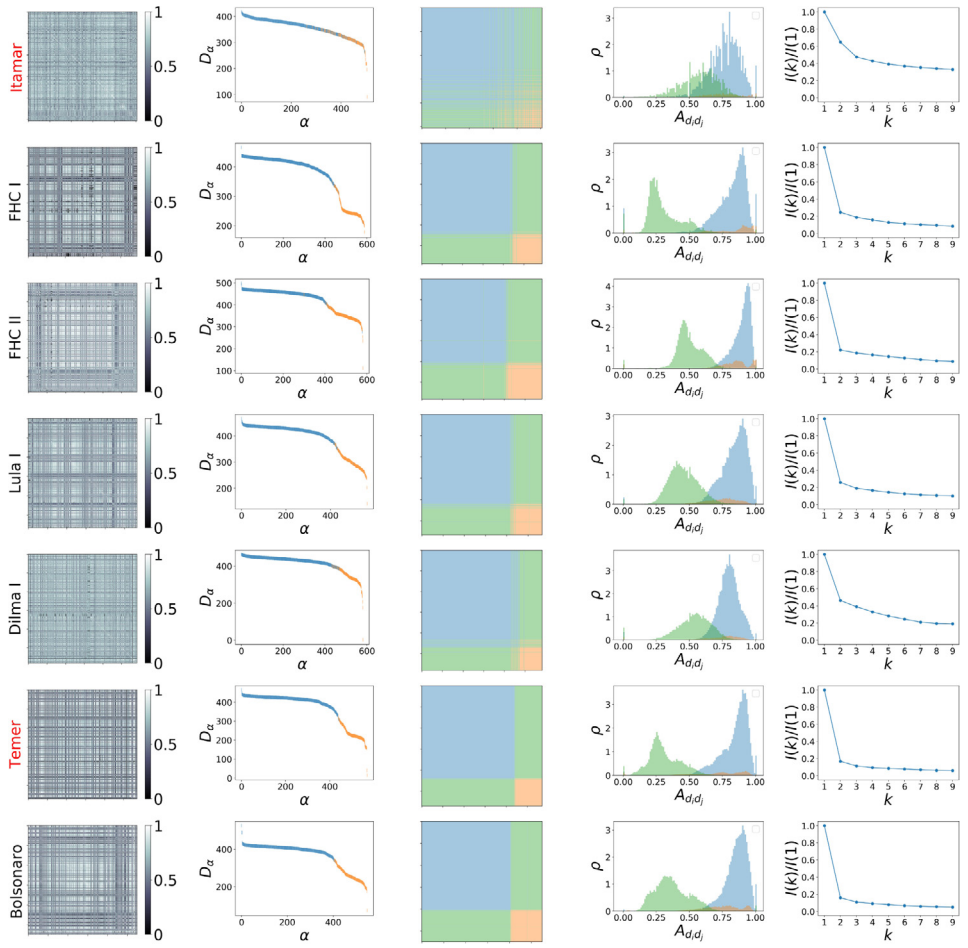

Fig. B.9. For every period we show (a) the agreement matrix with the agreement according to the color-bar to the right, (b) ranked degrees centrality $\mathcal{D}_\alpha$ of the congressmen colored by which cluster they were assigned to (blue for G_1^2 and orange for G_2^2), (c) ordered agreement matrix with congressmen ranked by degree centrality and colored blue if both belong to G_1^2, orange if both belong to G_2^2 and green if one belongs to G_1^2 and the other to G_2^2, (d) distribution of agreement matrix separated by the same coloring criterion of (c), and (e) the relative inertia $I(k)/I(1)$ calculated for different k.

the ranked degree centrality and natural number and sizes of clusters of congressmen and their support to the president and how the size and behavior of a third cluster might signal instability. Indeed we observe that stable periods have an important proportion of the congressmen in the ruling coalition with the president and the support of this group is very high. Our data also show that, once this coalition breaks down, the president's mandate does not survive. We are not able to say anything about why this happens but it can perhaps help to predict if this instability will happen with a given president. Dilma Rousseff's first mandate was signaling this instability as can be seen in some of our quantities as summarized in Table 6.

It would be important for next studies to build statistical models based on some of the characteristics of the congress that could predict some properties of each term that could lead to an instability and also to compare with results from roll calls in other countries [36]. For example, the minimum coalition size and the minimal support to maintain the mandate [37]. In this respect, we observe that the measure of cohesion proposed in this work, Fig. 2, shows that it is possible to differentiate very cohesive roll calls that align all parties together and, therefore, do not seem to bring relevant information to discriminate groups from another where the dissension and dynamics pitching the congressmen with different opposing ideas or strategies is happening. This observation also calls in question the role of the exaggerated number of parties in the Brazilian political scene [25,38]. The above mentioned measure of cohesion can be used as a filter of the roll calls that are relevant to separate parties with real different ideologies. It could then help building an objective measure of how many parties would be necessary to represent all the ideologies present in the Brazilian national Congress [39].

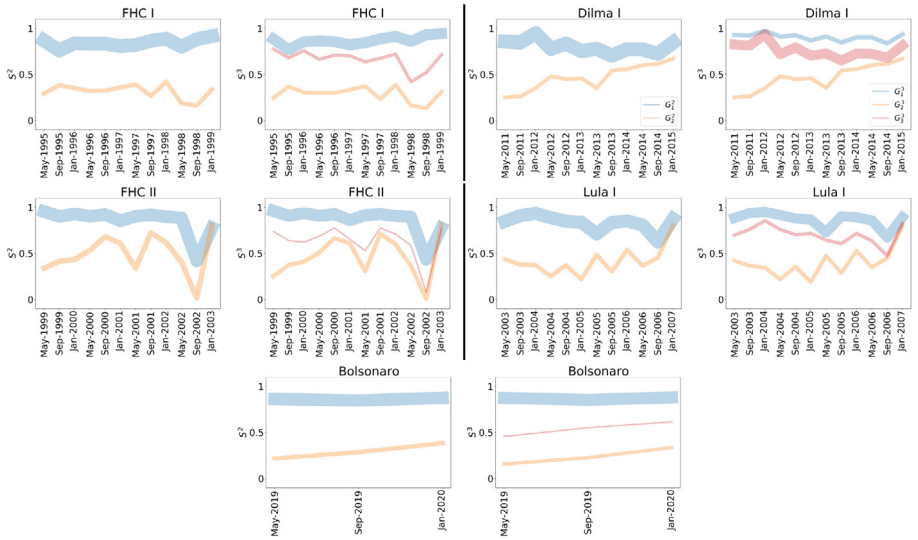

Fig. B.10. Evolution of clusters support to the president's party. The clusters were found with k-means using the whole period. The width of the line is proportional to the cluster size. The legend in Dilma I periods indicates the group name for all the periods.

CRediT authorship contribution statement

Frederico Fetter: Conceptualization, Methodology, Software, Formal analysis, Visualization, Writing - original draft. **Daniel Gamermann:** Conceptualization, Methodology, Software, Validation, Supervision, Writing - review editing. **Carolina Brito:** Conceptualization, Methodology, Resources, Supervision, Writing - review editing.

Declaration of competing interest

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

Acknowledgments

Research partially supported by the Brazilian agency *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq), Brazil.

Appendix A. Criterion of exclusion of congressmen with low legislative activity

There are congressmen that, for a variety of reasons, such as becoming ministers or taking another office appointed by the president or governors, have very few roll calls where they voted differently than Absent. Those congressmen usually have some very high (or very low) agreements with other congressmen with very different (or very similar) voting records, which introduces noise in the analysis.

To reduce this source of noise, we analyze the number of times that each congressman voted in the roll calls in a given legislative term. The distribution of it is shown in Fig. A.7 for all legislative terms together and for one particular term, showing that there is a regular pattern in this measure. We observe a peak close to zero that correspond to congressman who rarely vote and a second peak at higher values in this distribution. We then introduce a cutoff to exclude congressman who voted in less than 20% of the roll calls. We have played with this cutoff value and our results are robust against these small changes. In red are the excluded congressmen, in blue the kept ones. By choosing a 20% cutoff the largest number of excluded congressmen happens in the Itamar term from which 15.2% of them were excluded.

As there are substitutes for congressmen that for the many reasons leave the mandate and we consider each congressman individually, the total number of congressmen in our analyses for a given term can be greater than 513, which is the actual number of seats in the congress.

Appendix B. Results for all legislative terms

Through the paper, we analyzed mostly Collor, Lula II, and Dilma II terms, the two unstable terms and one as representative of a typical stable period. In this Appendix we reproduce the results we obtained for all the other legislative terms analyzed. In Fig. B.8 we show the cohesion phase space for every term. In Fig. B.9 we show some of metrics we used to analyze the different legislatures and in Fig. B.10 we show the evolution of the support S_n^k of each cluster G_n^k to the president's party.

References

[1] R. da Silva, L.C. Lamb, M.C. Barbosa, Universality, correlations, and rankings in the Brazilian universities national admission examinations, Physica A 457 (2016) 295–306, http://dx.doi.org/10.1016/j.physa.2016.03.014.
[2] M. Ausloos, Statistical physics in foreign exchange currency and stock markets, Physica A 285 (1–2) (2000) 48–65, http://dx.doi.org/10.1016/s0378-4371(00)00271-5.
[3] M. Kaufman, H.T. Diep, S. Kaufman, Sociophysics of intractable conflicts: Three-group dynamics, Physica A 517 (2019) 175–187, http://dx.doi.org/10.1016/j.physa.2018.11.003.
[4] N. Urbinati, Representative Democracy : Principles and Genealogy, University of Chicago Press, Chicago, 2006.
[5] B.F. Roukema, A first-digit anomaly in the 2009 Iranian presidential election, J. Appl. Stat. 41 (1) (2014) 164–199, URL http://EconPapers.repec.org/RePEc:taf:japsta:v:41:y:2014:i:1:p:164-199.
[6] A.M. Calvão, N. Crokidakis, C. Anteneodo, Stylized facts in Brazilian vote distributions, in: C. Sueur (Ed.), PLoS One 10 (9) (2015) e0137732, http://dx.doi.org/10.1371/journal.pone.0137732.
[7] W.K. Tam Cho, B.J. Gaines, Breaking the (Benford) law: Statistical fraud detection in campaign finance, Am. Statist. 61 (3) (2007) 218–223.
[8] T.C. Boas, F.D. Hidalgo, N.P. Richardson, The spoils of victory: Campaign donations and government contracts in Brazil, J. Polit. 76 (2) (2014) 415–429, http://dx.doi.org/10.1017/s002238161300145x.
[9] D. Gamermann, F.L. Antunes, Statistical analysis of Brazilian electoral campaigns via Benford's law, Physica A 496 (C) (2018) 171–188.
[10] Y. Sim, B.D.L. Acree, J.H. Gross, N.A. Smith, Measuring ideological proportions in political speeches, in: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Seattle, Washington, USA, 2013, pp. 91–101, URL https://www.aclweb.org/anthology/D13-1010.
[11] A. Fader, D.R. Radev, M.H. Crespin, B.L. Monroe, K.M. Quinn, M. Colaresi, MavenRank: Identifying influential members of the US senate using lexical centrality, in: Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, EMNLP-CONLL, Association for Computational Linguistics, Prague, Czech Republic, 2007, pp. 658–666, URL https://www.aclweb.org/anthology/D07-1069.
[12] S. Gerrish, D.M. Blei, How they vote: Issue-adjusted models of legislative behavior, in: F. Pereira, C.J.C. Burges, L. Bottou, K.Q. Weinberger (Eds.), Advances in Neural Information Processing Systems, Vol. 25, Curran Associates, Inc., 2012, pp. 2753–2761, URL http://papers.nips.cc/paper/4715-how-they-vote-iss-models-of-legislative-behavior.pdf.
[13] K.T. Poole, H. Rosenthal, Patterns of congressional voting, Am. J. Polit. Sci. 35 (1) (1991) 228, http://dx.doi.org/10.2307/2111445.
[14] J. Clinton, S. Jackman, D. Rivers, The statistical analysis of roll call data, Am. Polit. Sci. Rev. 98 (2) (2004) 355–370, http://dx.doi.org/10.1017/s0003055404001194.
[15] A. Jakulin, W. Buntine, T.M.L. Pira, H. Brasher, Analyzing the U.S. senate in 2003: Similarities, clusters, and blocs, Polit. Anal. 17 (3) (2009) 291–310, http://dx.doi.org/10.1093/pan/mpp006.
[16] L. Marenco, H.A. Carmona, F.M. Cardoso, J.S. Andrade Jr., C. Lenz Cesar, Time evolution of the behaviour of Brazilian legislative Representatives using a complex network approach, PLoS One 15 (2) (2020) 1–13, http://dx.doi.org/10.1371/journal.pone.0226504.
[17] Partidos políticos registrados no TSE, URL http://www.tse.jus.br/partidos/partidos-politicos.
[18] M. Laakso, R. Taagepera, "Effective" number of parties, Comp. Polit. Stud. 12 (1) (1979) 3–27, http://dx.doi.org/10.1177/001041407901200101.
[19] Bancadas da câmara na posse, 2020, URL https://www.camara.leg.br/deputados/bancada-na-posse. (Acessado em 14 April 2020).
[20] M. Laakso, R. Taagepera, "Effective" number of parties, Comp. Polit. Stud. 12 (1) (1979) 3–27, http://dx.doi.org/10.1177/001041407901200101.
[21] Portal Brasileiro de dados abertos, http://www.dados.gov.br/.
[22] Dados abertos - legislativo, https://www2.camara.leg.br/transparencia/dados-abertos/dados-abertos-legislativo.
[23] Dados abertos da câmara dos deputados, https://dadosabertos.camara.leg.br/swagger/api.html#api.
[24] Câmara dos deputados - assessoria de imprensa, 2018, URL http://www2.camara.leg.br/comunicacao/assessoria-de-imprensa/votacao. (Acessado em 14 November 2018).
[25] A.C. Figueiredo, F. Limongi, Presidential power, legislative organization, and party behavior in Brazil, Comp. Polit. 32 (2) (2000) 151, http://dx.doi.org/10.2307/422395.
[26] D.S. Vieira, J.M.E. Riveros, M. Jauregui, R.S. Mendes, Anomalous diffusion behavior in parliamentary presence, Phys. Rev. E 99 (4) (2019) http://dx.doi.org/10.1103/physreve.99.042141.
[27] C. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (3) (1948) 379–423, http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x.
[28] T. Hastie, R. Tibshirani, J.H. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction Second Edition, second ed., in: Springer Series in Statistics, Springer-Verlag New York, 2009.
[29] R.L. Thorndike, Who belongs in the family? Psychometrika 18 (4) (1953) 267–276, http://dx.doi.org/10.1007/bf02289263.
[30] S. Abraches, Presidencialismo de coalizão: o Dilema institucional Brasileiro, Rev. Ciências Soc. 31 (1) (1988) 5–39.
[31] P. Chaisty, N. Cheeseman, T. Power, Rethinking the 'presidentialism debate': conceptualizing coalitional politics in cross-regional perspective, Democratization 21 (1) (2012) 72–94, http://dx.doi.org/10.1080/13510347.2012.710604.
[32] P. Chaisty, S. Chernykh, Coalitional presidentialism and legislative control in post-Soviet Ukraine, Post-Sov. Aff. 31 (3) (2015) 177–200, http://dx.doi.org/10.1080/1060586x.2014.994941.
[33] S. Abranches, Presidencialismo de coalizão: Raízes e evolução do modelo político brasileiro, Companhia ds Letras, 2018.
[34] O. Neto Amorim, Presidential cabinets, electoral cycles, and coalition discipline in Brazil, Dados - Rev. Ciências Soc. (2000) URL https://www.redalyc.org/articulo.oa?id=21843303.
[35] F. Limongi, A democracia no Brasil: presidencialismo, coalizão partidária e processo decisório, Novos Estudos - CEBRAP (76) (2006) 17–41, http://dx.doi.org/10.1590/s0101-33002006000300002.
[36] A. Amelio, C. Pizzuti, Analyzing voting behavior in Italian parliament: Group cohesion and evolution, in: 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, IEEE, 2012, http://dx.doi.org/10.1109/asonam.2012.33.
[37] M. Duverger, Political Parties, their Organization and Activity in the Modern State, Librairie Armand Colin ed., 1969.

[38] S. Mainwaring, Politicians, parties, and electoral systems: Brazil in comparative perspective, Comp. Polit. 24 (1) (1991) 21, http://dx.doi.org/10.2307/422200.
[39] P.O.S. Vaz de Melo, How many political parties should Brazil have? A data-driven method to assess and reduce fragmentation in multi-party political systems, PLoS One 10 (10) (2015) e0140217, http://dx.doi.org/10.1371/journal.pone.0140217, URL https://pubmed.ncbi.nlm.nih.gov/26466365, 26466365[pmid].