# Principal Components for European Parliaments: A case study of Ireland and Sweden

Adam Tallon[1], Astrid S. de Wijn[2], and Mikael Fremling[1,3,*]

[1]Department of Theoretical Physics, Maynooth University, Maynooth, Ireland

[2]Department of Mechanical and Industrial Engineering, , Norwegian University of Science and Technology, Trondheim, Norway

[3]Institute of Theoretical Physics, Utrecht University, , Utrecht, The Netherlands

[*]m.h.o.fremling@uu.nl

June 11, 2019

**Abstract**

In this work we study Principal Component Analysis (PCA) method as a tool for roll-call analysis. Using PCA allows us to objectively determine the number of dimensions (or factors) that are necessary to faithfully represent the voting behavior within a parliament. As test subjects we use two very different European parliaments, the Irish and the Swedish. We find that the dimensionality of a parliament most often does not reflect the number parties in a parliament, and that it can vary between different government periods. Rather it is a measure of the number of *independent* parties or blocs, because coalition parties are effectively one party when it comes to voting. We find that the Irish Parliament of 2017 was at least three-dimensional, but the Swedish parliaments of 2010-2014 and 2014-2018 where four- and six-dimensional respectively. We also find that the political map of the Irish Parliament for the year 2017 conforms to the expectations set forth earlier by Costello [Costello, 2017].

# 1 Introduction

Traditionally, the political landscape has been discussed mainly in terms of left vs. right. This distinction has its roots in the French parliament, where the royalists were seated to the right of the speaker, and the anti-royalists were seated to the left. These labels have over time transcended from labelling which authority the King should have, to focusing more on economic ideas. They have become quite efficient at drawing up the main battlegrounds of political controversy. Nevertheless, due to the simplicity of assigning people *to the left* or *to the right*, this type of labeling is only useful if the political landscape it tries to describe is highly correlated. This, however, need not always be the case, as it is not guaranteed that a politician's view of *e.g.* economy correlates with their view on *e.g.* immigration, law enforcement or education. Over the years, there have been many attempts to map out the details of the policy space at hand. Various different methods may be applied for this task. In the context of the American senate and house of representatives, the NOMINATE technique of Poole and Rosenthaal has for many years been the gold standard[Poole and Rosenthal, 1990, Poole and Rosenthal, 2001]. Although the technique has mostly been applied in the context of the American congress it has found usage also in the the European Parliament [Hix, 2001, Hix et al., 2006, Hix and Noury, 2016]. See also a variety of regression models based on the assumption of utility functions [Erickson, 1971, Bullock III and Brady, 1983, Carrubba and Gabel, 1999, Jenkins, 1999, Snyder and Groseclose, 2000, Thiem, 2006, Hansen, 2009]. Many of the above aforementioned works take as their starting point roll-call data, *e.g.* votes in parliaments or supreme-courts.

In this paper we will focus on the methods of Principal Component Analysis. This method has previously been applied on the US congress [Heckman and Snyder, 1997, Ansolabehere et al., 2001], though in a different context.

Other techniques that have also been applied includes the classification of speeches given in parliament [Proksch and Slapin, 2010] as well as text analysis of the political manifestos for various parties [Merz et al., 2016]. One may of course conduct surveys [Bakker et al., 2015] to directly probe the political landscape on a number of pre-defined issues.

An important drawback of the methods mentioned above for gauging positions of politicians, is that one is inadvertently introducing a selection bias, by already having defined which how many potential axes are to be considered. In the case of NOMINATE and utility functions, this takes the form of choosing how many factors are implemented in the analysis. This choice may be influenced by the subjective opinions and experience of the person performing the analysis. For instance, it is widely known [Poole and Rosenthal, 1990] that the US Senate can be well described by at most two variables, and often a single variable is sufficient. Analyses on such a left-right scale are *e.g.* routinely performed for all the American branches of legislature (see References [Poole, 2005, Jessee, 2009, Shor et al., 2010, Shor and McCarty, 2011, Lauderdale and Clark, 2012] for a small selection). It should be noted that the USA of course has a much more diverse policy landscape than the two parties on capitol hill might suggest. This is obscured by the pre-selection process that takes place internally in both the parties due to the first-past-the-post system that is employed in elections. For this reason, importing the methods employed to analyze the US might not be so successful in a European context, see [Hansen, 2009] for a recent summary of possible objections.

In this article we investigate a method for determining how many dimensions (or factors) are needed to faithfully describe a parliament, and how important each of these dimensions actually is. For such a purpose, the measuring device itself must be constructed to be objective, and Principal Component Analysis has this property, since the number of factors is not determined beforehand. In previous works, the call to decide how many factors to include has been been justified by considering the added explanation strength (*i.e.* how many more roll-calls can be correctly classified) by adding more factors. In this work we take a different approach, and determine the number of dimensions that are statistically relevant.

As case studies we have chosen the Irish and the Swedish parliaments. These two parliaments are interesting to compare because their election and voting processes are quite different. We use roll-call data from different years to study how the number of relevant dimensions may change over time. To be precise, we will consider the votes of each parliamentarian as a high-dimensional proxy of their relative positions in the true policy space. This can then be used to examine the relative positions of parties and the voting discipline of the MPs within parties.

We wish to emphasize that most previous work on roll-call data has focused on parliaments where parties are relatively weak (*e.g.* the US congress, the European Parliament). A lot of previous work on national parliaments in Europe has also found roll-call data to be relatively uninformative, due to the overarching government-opposition dynamic, see *e.g.* [Hix and Noury, 2016]. In this work we will show that this is not always the case, and a more diverse policy space map may emerge when there are more than two clearly definable "blocks" of parties.

This paper is organized as follows: In Section 2 we discuss the *Principal Component Analysis* (PCA) methodology and our data-sources including as data-cleaning. In Section 3 we study the eigenvalues of the PCA and draw conclusions on the respective dimensionalities of the parliaments under consideration. Section 4 is devoted to constructing spacial policy maps of the various parliaments considered and some comparisons to earlier work is discussed. A short discussion of axis labeling can be found in Section 5 before we end with a summary in Section 6. Its worth mentioning at this stage that, although highly desirable, no systematic attempt at characterizing the different axes will be attempted in this paper.

## 2 Description of Method

In this paper we are primarily interested finding a good measure for how many individual policy axes are needed to describe the roll call behavior within *e.g.* a parliament. One may envision the actual policy space as having many more dimensions where each potential policy decision has its own axis. It is more than likely that the policy distributions along several axes will be correlated, and it is therefore of importance to estimate the number of "independent" political axes that are needed in order to describe the decomposition of the parliaments in question. As a secondary objective we are also interested in how significant these axes are and how the parties are distributed within this policy space.

We employ for this purpose a technique called Principal Component Analysis (PCA). PCA is a general purpose technique that allows one to extract the number of independent variables needed to reliably describe a data-set containing a potentially much larger number of variables. Crucially, it also provides information on how important each of these independent variables actually is (*i.e.* how much information it provides). For more details on PCA see [Hastie et al., 2009]. The PCA method sometimes also goes under the name of Heckman-Snyder[Heckman and Snyder, 1997].

In this section we will first briefly outline the PCA method in the context of analyzing roll-call data. We will then discuss the advantages and disadvantages of PCA to other methods in the field. Finally we will describe the specific data cleaning and interpretations that pertains to the two parliaments used as test beds in this work.

## 2.1 The PCA Method

The PCA method is a general purpose method that can be employed to distill the most descriptive combination of characteristics from a data-set. The method is used frequently in statistical analysis to build a simplified prediction model for more complicated data.

In our case the variable space will be the set of roll-call votes of a parliament during a given time interval. We may think of the roll-call votes as a forming a high-dimensional proxy of the actually policy space, where each parliamentarian is a point in this space. The PCA acts by first defining the direction in this space along which there is the largest variance of the politicians' votes (*i.e.* the axis where there is the most political disagreement). This is the first principal axis, and the variance along this axis is the first principal component. The procedure is then repeated in the space perpendicular to the first principal axis, to produce the second principal axis, and so forth.

To be concrete, assume a parliament contains $N_p$ parliamentarians labeled $i = 1, \ldots, N_p$. During some time interval the parliament votes in $N_r$ roll-calls, labeled $\alpha = 1, \ldots, N_r$. As an example, consider the case of Ireland. Here the $N_r = 200$ roll-calls of the Irish parliament form together a 200 dimensional space and each of the $N_p = 157$ parliamentarian is one point in this space. Over the year in question MP $i$ will vote "Yae", "Nay" or "Abstain" on each of the 200 votes. Natural values of 1, -1, or 0 can be assigned to these votes. We let $R_{\alpha,i}$ denote the numerical value cast by the $i$-th parliamentarian in the $\alpha$-th roll-call. The total number of parliamentarians and votes is shown in Table 1.

Next we will construct an object known as the Covariance Matrix, $C_{\alpha,\beta}$. The covaraince matrix $C_{\alpha,\beta}$ is a $N_r \times N_r$ matrix. Its components are given by the following formula

$$C_{\alpha,\beta} = \frac{1}{N} \sum_{i=1}^{N_p} \tilde{R}_{\alpha,i} \tilde{R}_{\beta,i}. \tag{1}$$

Here $\tilde{R}_{\alpha,i} = R_{\alpha,i} - \frac{1}{N} \sum_{k=1}^{N_p} R_{\alpha,k}$ are normalized such that the average numerical vote for roll-call $\alpha$ is centered at zero. By identifying $\langle R_\alpha \rangle = \frac{1}{N} \sum_{k=1}^{N_p} R_{\alpha,k}$ and $\langle R_\alpha R_\beta \rangle = \frac{1}{N} \sum_{k=1}^{N_p} R_{\alpha,k} R_{\beta,k}$ we can see that equation (1) is simply $C_{\alpha,\beta} = \langle R_\alpha R_\beta \rangle - \langle R_\alpha \rangle \langle R_\beta \rangle$ in a different guise.

Armed with the covariance matrix $C_{\alpha,\beta}$, we diagonalize it by numerically solving the eigenvalue equation.

$$\sum_{\alpha=1}^{N_r} V_\alpha^{(k)} C_{\alpha,\beta} = \lambda_k C_{\alpha,\beta}. \tag{2}$$

This equation has $N_r$ solutions, and $k$ labels those solutions. The eigenvalues $\lambda_k$ are called the principal values, and the vectors $V_\alpha^{(k)}$ are called the principal axes. On a standard laptop this equation can easily be solved for matrices including up to a few thousand rows and columns, and is a standard routine in most statistical software including R and Python. The largest value of $\lambda$ is $\lambda_1$ which represent the largest possible variance in any data direction, while the corresponding vector $V_\alpha^{(1)}$ is pointing in that direction. Since $C_{\alpha,\beta}$ is symmetric ($C_{\alpha,\beta} = C_{\beta,\alpha}$), all the principal axes $V_\alpha^{(k)}$ are guaranteed to be perpendicular to one another. We may then interpret the remaining $\lambda_k$ as the variances in directions perpendicular to $V_\alpha^{(1)}$ (and to each other). The statistical significance of the dimensions can be obtained by comparing to a new data set where the roll-call votes have been randomly redistributed. This shows the noise level. If the variance in the $k$-th dimension of the original parliament exceeds that of the random parliament, it is statistically significant.

The final step is to project the votes of all the individual parlamentarians on the principal axes corresponding to a few of the largest values of $\lambda_k$. The projected position along the principal axes $k$ are given by

$$P_{k,i} = \sum_{\alpha=1}^{N_r} V_\alpha^{(k)} R_{\alpha,i}, \tag{3}$$

provided we first make sure that $\sum_{\alpha=1}^{N_r} |V_\alpha^{(k)}|^2 = 1$.

For a discussion of labeling of the principal axes, we defer the discussion to Section 5. We wish to stress, however, that even without having knowledge of the policy labels pertaining to each individual axis, equation (3) is still able to produce spatial maps of the positions in policy space. The special maps are useful already, as we will see in *e.g.* Figure 2, to gain information about the relations between the different parties.

## 2.2 Relation to other methods

The method of PCA is only one of a number of methods that can be applied to the model *e.g.* the roll-call behavior in parliaments. The two other major frameworks for this type of analysis is the NOMINATE[Poole and Rosenthal, 1997] and IDEAL[Clinton et al., 2004]. It should be clear to the reader that both of these methods are quite more sophisticated than PCA. They both assume at the core (in their simplest versions) that the voting outcome of individual legislators is given by the (quadratic) utility functions $Y_i(\zeta_j) = -\|\zeta_j - \xi_j\|^2 + \eta_{ij}$ and $N_i(\psi_i) = -\|\psi_i - \xi_j\|^2 + \nu_{ij}$. Here $\xi_j \in \mathcal{R}^d$ is the ideal point of legislator $j$ in a $d$-dimensional policy space; $\zeta_j$, $\psi_j$ are the legislative outcome of "yes" or "no" being adopted and $\eta_{ij}$, $\nu_{ij}$ are random fluctuations in the decision process. The legislator is then assumed to vote "yes" if $Y_i(\zeta_j) > N_i(\psi_j)$ and "no" if the reverse is the case. The task of both NOMINATE and IDEAL is then to estimate the ideal positions $\xi_j$ given the roll-call data at hand. See references [Carroll et al., 2009] and [Clinton and Jackman, 2009] for an in-depth comparison between the two methods.

A subtle, but important point, is that the best estimators of $\xi_j$ will depend on the dimension $d$ of the model, much in the same way that regression coefficients will change depending on whether one fits to a linear or quadratic function. The major advantage of PCA is that it does not require a predetermined dimensionality, such as in the case of the above mentioned, otherwise more sophisticated methods, for mapping out policy spaces. For this reason we can use PCA to extract the dimensionality of the political landscape without human-induced bias. Unlike NOMINATE and IDEAL, however, PCA does not resolve the non-linearities at the far extremes[Clinton et al., 2004]. PCA also has the advantage (in our opinion) that it is easy to implement, as it only requires linear algebra. Moreover, it is easy to visualize which issues contribute the most to a certain direction in policy space, by considering the decomposition of the eigenvectors in equation (3).

## 2.3 Data cleaning

In this work we will consider official roll-call data from the Irish parliament[IRL, 2018] in 2017 and from the Swedish parliament[SWE, 2018] in the periods 2010-2014 and 2014-2018. We choose both the periods 2010-2014 and 2014-2018 since there has been a change of coalition government between these two periods. The Swedish parliament votes on an average of 600 issues per year, and the Irish votes on somewhat fewer, ca 200. All of these data-sets are publicly available online from the respective parliaments websites[IRL, 2018, SWE, 2018]. The idea is that since many of these votes will be on similar issues, economic, social, etc., the sheer number of votes will paint a faithful picture of the relevant policy space of said parliament.

In the data for Ireland there are votes recorded for all the MPs. Sometimes, however, an MP will be unable to take part in a roll-call, *e.g.* because of sickness or travel. In these situations, the vote is recorded as an "Absent". In the case of "Absent" we choose to treat the absence of a vote numerically the same as an abstain, *i.e.* recorded as "0". Although technically the "Abstain" and "Absent" have different origins, their impact on the amount of information in the vote is the same (*i.e.* no preference in either direction) and it is therefore justified to threat them in a similar manner.

In the data for Sweden, each member of parliament (MP) is given a unique ID-number that we use as the main key. Here when an MP is unable to take part in a roll-call there are two options. The vote is either 1) recorded as an "Absent"or 2) cast by another – substitute – member of the parliament. The case 1) is treated the same as an Irish "Absent", namely with the numerical value of 0. The case 2) causes some problems in data interpretation, since sometimes the same substitute can act as substitute for more than one MP, and sometimes an MP seems to have had more than one substitute. Ideally we would like to match each vote cast by a substitute with the intended primary MP, but this is not always possible.

We clean the data by considering the set of all votes cast by MPs $A$ and $B$ (we ignore MPs that have participated in all the votes). If the votes that where not cast by $A$ precisely coincides with the ones that were cast by $B$ – and $A$ and $B$ have the same party affiliation – then we conclude that $B$ is the substitute of $A$. In some situations we have also identified three MPs $A$, $B$ and $C$ that share the same party affiliation, have no mutual votes, but together vote in all the roll-calls. In this case we identify $B$ and $C$ as substitutes for $A$. Any MP who does not fulfill this criteria is removed from the analysis. As an observation we note that the coalition MPs are typically over-represented in removed MPs during their coalition years.

There is a possibility of seasonal bias. During different periods in the year, different types of issues may be discussed or voted on. To prevent this from distorting our results, we consider votes only for the full calendar year that has so far been completed. For the Irish parliament, this means we have to restrict ourselves to 2017. For the Swedish

| | Ireland 2017 | Sweden 2010-2014 | Sweden 2014-2018 |
|---|---|---|---|
| Total MPs | 157 | 349 | 349 |
| Usable MPs | 157 | 336 | 323 |
| Roll-calls | 175 | 2699 | 2601 |

Table 1: Summary of the data used for the parliaments under consideration. For Sweden the total number of members of parliament (MP)s and the number used in the analysis is not the same. This is because the Swedish MPs use substitutes when they are not able to vote, and sometimes identifying the correct substitute is not possible.

data we simply use the entire set of roll-calls for both periods since they are exactly four years long.

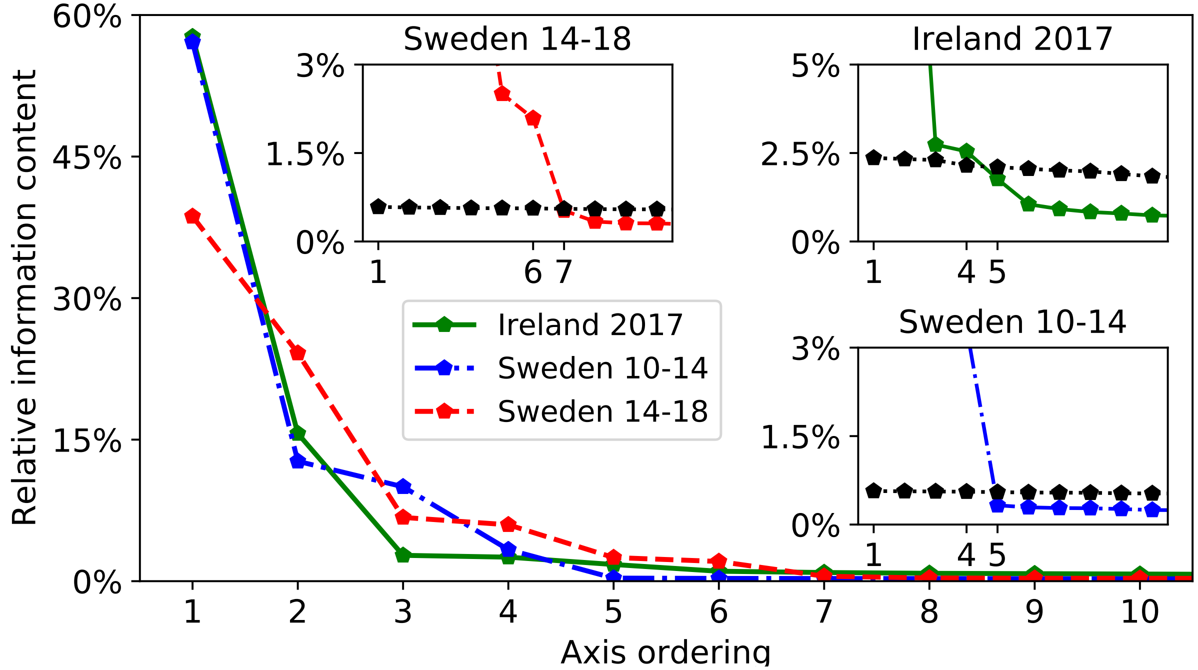


Figure 1: (Main panel) The relative information content in each of the principal axes for the different parliaments under consideration. The PCA procedure selects the directions in policy space which shows the largest variance, corresponding to maximizing the relative information content of that axis. It is clearly visible that the first principal axis is enough to capture roughly 60% and 40% for the different parliaments respectively. (Insets) Each inset shows a comparison between the PCA for the actual parliament and a “random parliament” where the votes have been shuffled, roll-call by roll-cal. When the information content in the actual parliament drops below the information content of the random parliament we conclude that these axes are irrelevant. Using this measure we can see that Ireland 2017 and Sweden 2010-14 is 4 dimensional, whereas Sweden 2014-18 is 6 dimensional.

# 3 Dimensionality at a Glance

To obtain the effective dimensionality we extract the principal components of the covaraince matrix $C_{\alpha,\beta}$ by solving equation (2) numerically. The relative strength of the principal components for the different parliaments are summarized in the main panel of Fig. 1. To be able to compare the different parliaments, which have different number of parliamentarians as well as number of roll-calls, we use the normalized principal components $\tilde{\lambda}_j = \lambda_j/(\sum_k \lambda_k)$. The quantity $\tilde{\lambda}_j$ has the interpretation that is measures how much of the total variance of the data $(\sum_k \lambda_k)$ that is captured by axis number $j$.

We see in the plot that the first axis for both Ireland 2017 and Sweden 2010-14 captures roughly 60% of the variance within the data. For Sweden 2014-18 the first direction only captures 40% of the variance. The subsequent axes capture less and less information and by the 7-th axis, we can see that each axis now explain less than one percent of the total variance.

We now estimate how many of these axes are relevant, *i.e.* how many axes contain more information than random noise. To this end we randomly scramble the votes of all the MPs, roll-call by roll-call. This produces a "random parliament" with the same roll-call distributions as the actual parliament, but without any correlations between the different roll-calls. We then compare the information content of each principal axis of the random parliament with the information content of the actual parliament. If the actual information drops below the "random" information, we conclude that this axis is not necessary to describe the dimensionality of the parliament. See the insets of Fig. 1.

Using this measure we conclude that for Ireland 2017 and Sweden 2010-14 there are 4 axes that carry actual information, whereas for Sweden 2014-18 there is relevant information up to and including axis 6. Note however that the transition from relevant to irrelevant axes in the Irish parliament happens less abruptly than in the Swedish parliament. See *e.g.* the sub-panels for Sweden and Ireland, where in the former cases, there is a sharp drop in information content between axis 6 (or 4) and axis 7 (or 5) respectively. In the Irish system on the other hand, the drop in information happens more gradually over axis 3, 4 and 5. We speculate that this tail is related to the Irish parliament having a significant number of independents and smaller parties, and we will give some evidence to support this in Section 4.1.

One would on general grounds expect that a parliament consisting of $n$ parties with strong MP cohesion, would not show a dimensionality that is larger than $n-1$. This does not exclude the possibility of being lower than $n-1$ dimensions. It is therefore interesting to note that the effective dimensionality for the Swedish parliament varies between the 2010-14 and 2014-18 terms. The main reasons for this, which we will return to later, is the change of government between the two terms. This affects the number of parties taking part in the governing coalition, and indirectly the effective dimensionality. What happens is that when several parties enter into a coalition, any intra-party differences will (often) have been ironed out in internal negotiations, making the coalition effectively appear as one party. As a simple counting exercise we note that for 2010-14 and 2014-18 respectively there where eight parties with a government coalition consisting of four or two parties respectively. The effective number of parties is thus five (or seven), giving rise to an upper bound of four (or six) dimensions.

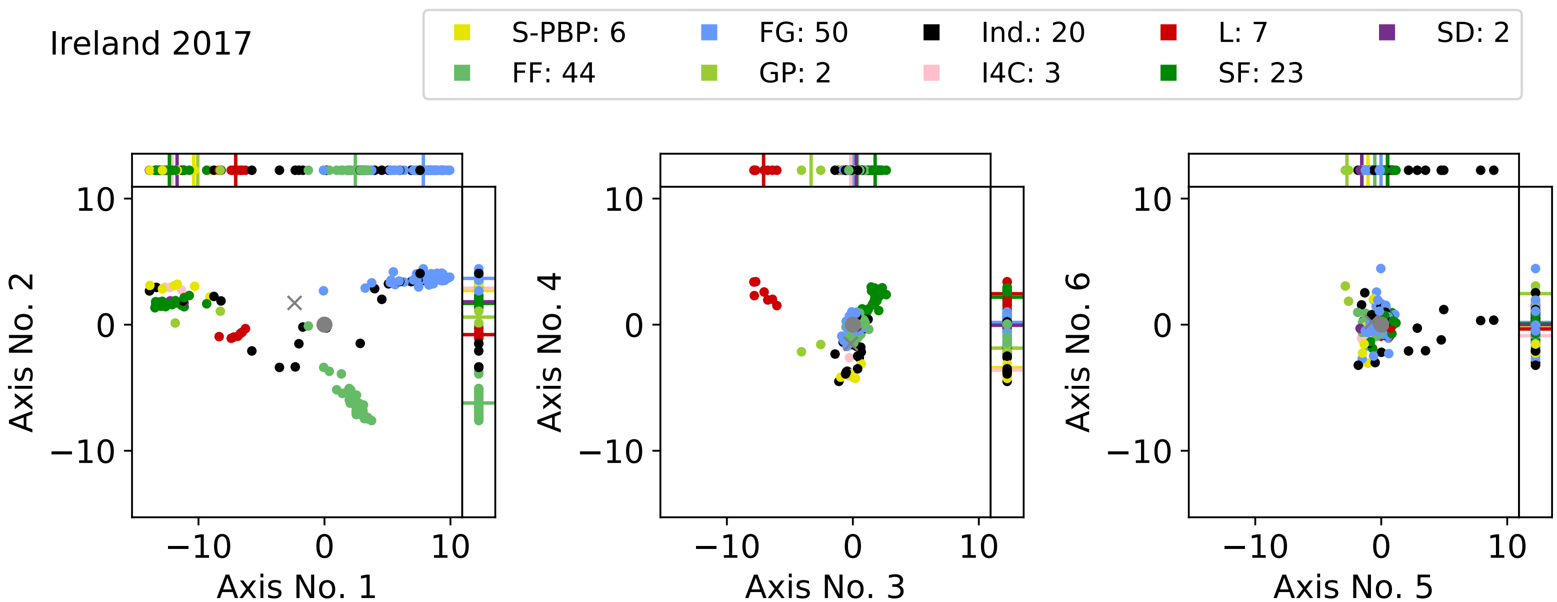


Figure 2: Projection onto the principal axes, Ireland 2017. On the first axis we see FG and FF to the right, while SF, Labour and many of the smaller parties are to the left. The scale in the different plots are chosen to be the same, for all axes, to make it easier for the reader to compare the projections on the different axes. The average position of each party (independents excluded) is marked with a solid line. A legend of the party names and affiliations can be found in Table 3.

# 4 Mapping the Political Space

In this section we show the spatial maps of the relevant policy space for our three parliaments in question. We construct these maps by projecting onto the most relevant axes that the PCA gives us.

We begin with a general remark on the distribution of MPs within the parties themselves. In many of the maps to be presented in Figure 2, 4 and 5, the distributions of MPs for a given party will take on a elongated shape. This should be contrasted with the expectation that all MPs should be on top of each other if the party voted coherently on all issues. The spreading out is to a certain extent an artifact of the more or less "random" absentees that adds a widening to the distribution of MPs. It also explains why the elongated distributions are radially directed, since The MPs that vote less often will be pulled towards the origin.

We would also like to mention that the maps presented are using the normalized roll-call values $\tilde{R}_{i,\alpha}$. To make this clear we mark the average of all the votes with a "X" and the abstaining on all the votes with a bullet "•".

## 4.1 Ireland

We begin by taking a closer look at the Irish parliament. Roll call analysis for Ireland has previously been performed using different methods in for example [Hansen, 2009]. The Irish MPs are elected from 40 constituencies each comprising of three to five seats. The MPs are elected by a single transferable vote system, and run independently for office.

The distinction between Government and opposition in the Irish parliament is obscured slightly by the "Confidence and Supply agreement" between the two largest parties, Fine Gael (FG) and Fianna Fáil (FG). Following the General election in 2016, neither FG or FF were willing to enter coalition with the third largest party Sinn Féin (SF) and, forming a viable coalition became impossible. The aforementioned "Confidence and Supply agreement" was the solution, where FG would be facilitated by FF to form a minority government. In practical terms this means that FF would abstain from the election of Prime Minister and his cabinet, vote against all votes of no-confidence in the government and support budgets put forward by FG (dependent on their adherence to a set of agreed red lines). FF remains the lead opposition party and are acknowledged as such by the confidence and supply agreement. They are free to propose their own legislation and vote according to their own party message on all votes not specifically outlined in the confidence and supply agreement.

We now consider Figure 2 (left panel). We first notice that the parliamentarians from the same parties group together, and one can visually distinguish different grouping even without the color coding. Focusing on the projection onto the first axis, we find Sinn Féin (SF) and Fine Gael (FG) at opposite ends of the spectrum with Labour (L) and Fianna Fáil (FF) between them (as well as a smattering of smaller parties and independents throughout).

On the second axis we can see how FF separates further from FG such that the three groups are clearly visible, comprising of FG, FF and SF plus independents.

The third and fourth axes Figure 2 (middle panel) show the lines of disagreement between Labour, Sinn Féin and the smaller liberal leaning parties (such as Solidarity - People Before Profit and Independents for Change). On the two most principal axes these appear to be quite close, but here they start to distinguish themselves.

**Ireland 2017**

| | FG | FF | L | S-PBP | GP | SF | SD |
|---|---|---|---|---|---|---|---|
| I4C | 100 | 87 | 58 | 33 | 45 | 37 | 41 |
| SD | 99 | 83 | 54 | 38 | 44 | 33 | |
| SF | 98 | 82 | 51 | 37 | 47 | | |
| GP | 95 | 78 | 49 | 45 | | | |
| S-PBP | 92 | 80 | 55 | | | | |
| L | 83 | 65 | | | | | |
| FF | 54 | | | | | | |

**Sweden 2010-2014**

| | SD | C | L | KD | M | MP | S |
|---|---|---|---|---|---|---|---|
| V | 100 | 85 | 86 | 88 | 90 | 52 | 60 |
| S | 94 | 77 | 77 | 79 | 81 | 65 | |
| MP | 96 | 76 | 77 | 79 | 80 | | |
| M | 72 | 8 | 7 | 6 | | | |
| KD | 70 | 8 | 8 | | | | |
| L | 69 | 6 | | | | | |
| C | 69 | | | | | | |

**Sweden 2014-2018**

| | SD | C | L | KD | M | MP | S |
|---|---|---|---|---|---|---|---|
| V | 98 | 92 | 89 | 91 | 97 | 60 | 61 |
| S | 91 | 81 | 77 | 79 | 85 | 6 | |
| MP | 91 | 81 | 76 | 78 | 85 | | |
| M | 100 | 58 | 59 | 58 | | | |
| KD | 97 | 52 | 53 | | | | |
| L | 96 | 54 | | | | | |
| C | 100 | | | | | | |

Table 2: The euclidean distances between the different parties based on the average position of the MPs of the parties. The distances are normalized such that the two parties that are furthest apart are at a distance of 100(%). In the case of Ireland one can see that Fine Gael (FG) and Independents for Change (I4C) are furthest away from each other, with Social Democrats (SD) and Sinn Féin (SF) at roughly the same distance. In Sweden 2010-2014 one finds Sverigedemokraterna (SD) and Vänseterpartiet (V) at different ends of the policy space, whereas for 2014-2018 SD is almost equally far from all other parties. A legend of the party names and affiliations can be found in Table 3 and Table 4.

Considering also the fifth and sixth axis in (fig. 2 right panel) shows how the independents contribute to the slow decrease in axis information that was observed in Section 3. See especially the insets in Fig. 1. We thus see that a four-dimensional space is required to accurately represent the positions of the eight parties in the Irish parliament, but if one wished to also model the various independents, then at least one extra axis is needed.

We mention here that using a linear space clearly works if one is interested the party's relative positions. However, if one wishes to resolve the parties themselves, then other methods will be necessary.

We may compare our results to the findings that where presented by Costello[Costello, 2017]. Their survey compared the self-reported policy-preferences of politicians and voters, prior to the general election of 2016. In that work it was concluded that four (4) policy axes was enough to accurately capture the space of Irish politics: Economic, Austerity, Religious and Cultural. We focus on the placements of the political parties on these four axes as reported in that work, and reconstruct a parliament with only 4 votes on these axes with Costello's continuous values between one and three. To be precise, we construct the table of roll-calls $R_{\alpha,i}$ with fictitious MPs (equal to the amount in the current parliament) and perform PCA on that. We can then compare with the actual behavior in parliament. See Figure 3. The area of the circles representing each party is proportional to their number of MPs, as well as representation in the real parliament.

We first note that our PCA analysis concludes that only two axes are necessary to represent the Costello result (right panel). Here we do not attempt a comparison with a "random parliament" as in Section 3, as the policy space is only made up of four axes. We can however construct a policy pallette by following the steps that will be outlined in Section 5. We find that the policy palette that makes up the axes is found to be for axis 1: 25% economic, 2% cultural, 30% religious, 43% austerity and on axis 2: 6% economic, 31% cultural, 20% religious, 43% austerity. We find that the general positioning of the three biggest parties FG, FF and SF on the first two axes is qualitatively similar to what we observed earlier in Figure 2. We find a discrepancy in that L and GP show up diametrically opposed to FF, where they in the parliamentary analysis are somewhere between SF and FF. While potentially this represents some incongruity between these parties pre- and post election political positions, this difference may however be accounted for if one considers that L and GP on the third axis of Figure 2 accounts almost single-handily for a dimension. It is not unreasonable that this difference can be attributed to the already truncated policy-space that is used in the Costello paper.

We finish this section by constructing the euclidean distance between the different parties in the Irish parliament. For this we begin computing the mean positions $\mathcal{R}_{\alpha,i}$ of all the parties in the 175 dimensional space of roll-calls. Here $\alpha$ enumerates the roll-call and $i$ enumerates the party. Following that we compute the euclidean distance $d_{i,j}$ between two parties $i$ and $j$ as $d^2_{i,j} = \sum_\alpha (\mathcal{R}_{\alpha,i} - \mathcal{R}_{\alpha,j})^2$. The results can be found in Table 2, where the distances $d_{i,j}$ have been normalized such that the largest distance is 100(%). In the table we can see that Fine Gael (FG) is significantly closer to Fianna Fáil (FF) than to any other party (54%), and furthest way from Independents for Change (I4C). It is interesting to note (as mentioned earlier in the section) that the difference between SF and the smaller parties like I4C and S-PBP and SD, which seems to lie very close when only the first and second principal axis are considered, separate considerably when the full data-set is taken into account, and end up at a distance of roughly 35%. We stress again that the euclidean distance measure reported here is not able to factor in how important each roll-call is perceived to be.

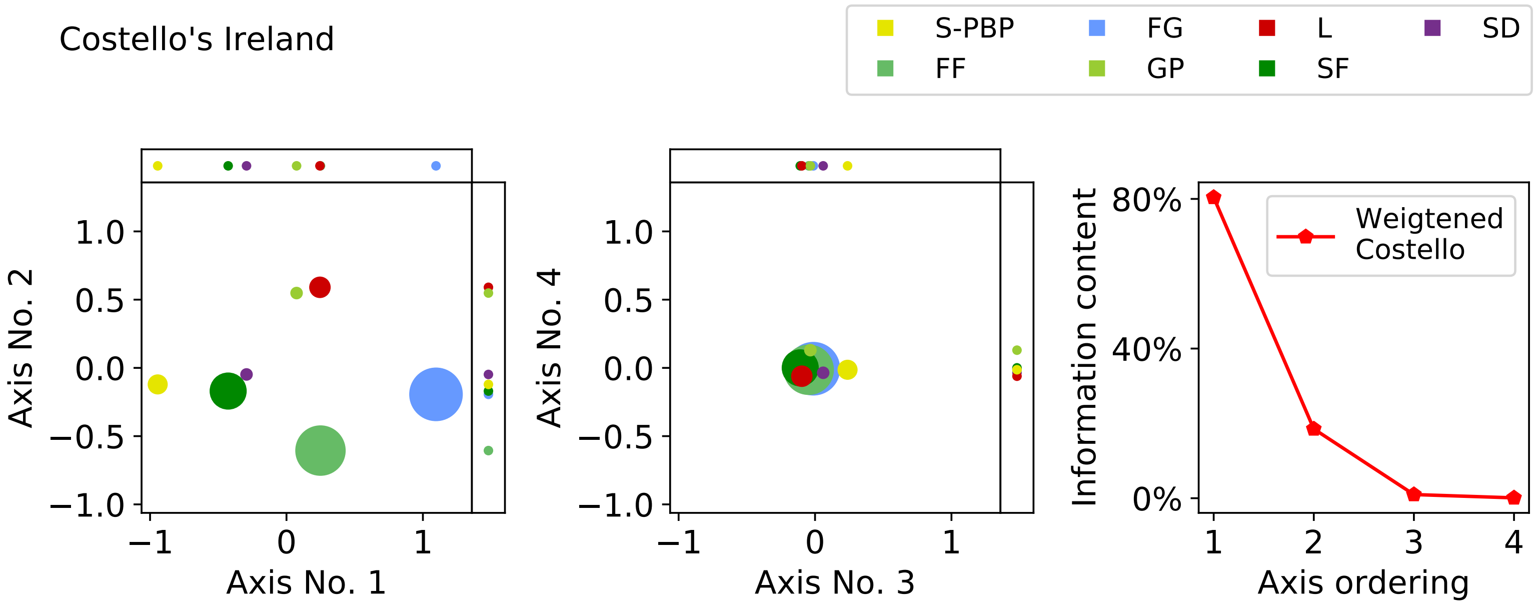


Figure 3: Policy space map of Ireland constructed using the parties positions issues of economy, religion, austerity and culture as reported by Costello in Ref. [Costello, 2017]. A fictitious parliament with MPs distributed proportional to the post-2016 parliament was created and analyzed identically to that in Figure 2. The overall positions of the parties agrees well with the map presented in the same figure. Since only four "votes" are considered, the random parliament analysis cannot be used. It is however evident that third and fourth axis contain next to no information. The middle panel shows very little structure, as is expected, given the very little information content in the third and fourth axis. A legend of the party names and affiliations can be found in Table 3.

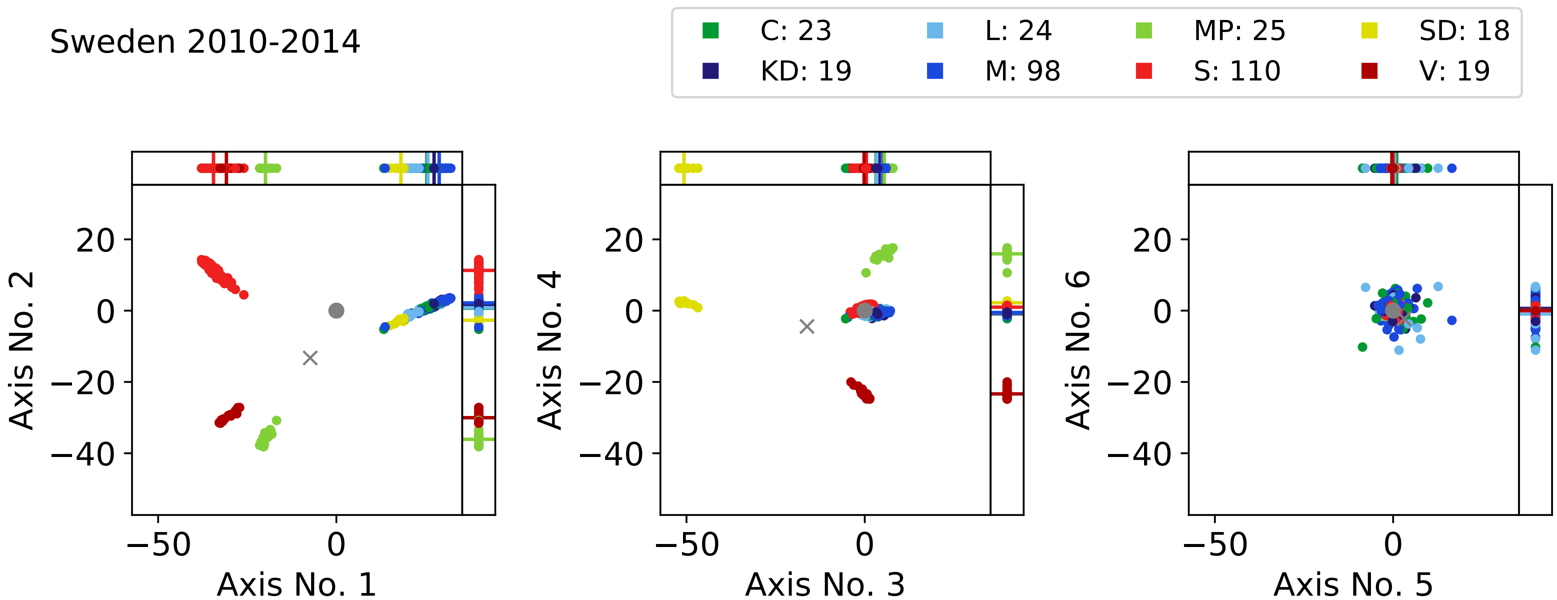


Figure 4: Policy-space map of Sweden 2010-2014 constructed by projection onto the principal axes for that parliament. The first principal axis shows a strong economic component with the conservative alliance to the right and the more left leaning parties S, V, MP to the left. See Table 4 for a list of all party names and acronyms.

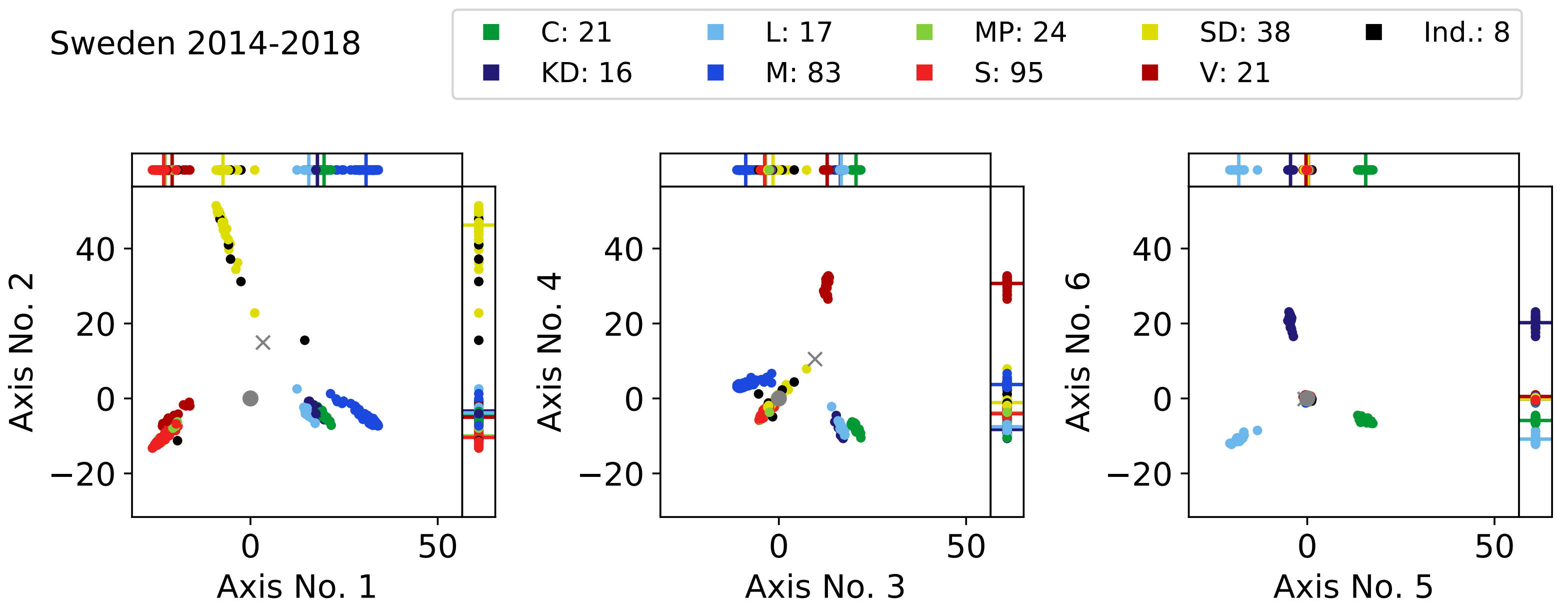


Figure 5: Policy-space map of Sweden 2014-2018 constructed by projection onto the principal axes for that parliament. Just as in Figure 4 the first principal axis shows a strong economic component with the conservatives M,L,C,KD to the right and the red-green coalition government to the left. An important difference to the previous period is that SD is now more clearly forming a third leg in the parliament. One can also see that SD is the party that is plagued the most by defecting MPs (labeled independents in the plot). See Table 4 for a list of all party names and acronyms.

## 4.2 Sweden

We now turn our attention to Sweden. MPs are elected from 29 constituencies of varying sizes, where the smallest is nominating 2 MPs and the biggest is nominating 39. The MPs do not run for office individually, but as part of a party, and the seats of the parliament are distributed proportionally to the national election result. The Swedish voting system does in practice not allow for independents, as there is a national 4% bar for a party to get into parliament [1].

### 4.2.1 Parliament of 2010-14

We begin by looking at the earlier parliament 2010-2014, as that seems to have a lower dimensional policy space compared to 2014-2018. During these years (just as 2006-2010) Sweden had a coalition government including Moderaterna (M), Liberalerna (L), Centerpartiet (C) and Kristdemokraterna (KD). See Table 4 for a list of all party names and acronyms. This coalition was popularly called Alliansen (The Alliance) and during the preceding election the members of Alliansen were campaigning together as one predetermined alternative, but still as separate parties.

The very strong cohesion that signified Alliansen becomes evident from the relevant policy space of Sweden during these years. See Fig. 4. One can clearly see that there is no separation of Alliansen parties at all in the four most relevant axes. Axis five and six – which where discarded as irrelevant based on the criteria described in Section 3 – are included to show that they really are irrelevant for the point of view of party politics.

During this period, there were four parties in the parliament that were not in government, Socialdemokraterna (S), Vänsterpartiet (V), Miljöpartiet (MP) and Sverigedemokra-

[1] There also exists a second mechanism to allow seats to regionally strong parties. If a party gets at least 12% in a constituency, they are entitled to their relative share of the seats from that particular constituency.

terna (SD). The three former are economically of varying degree to the left, which can be seen in the (left panel) of Fig. 4. At this point we should be cautious to give interpretations to the axes, but the first axes does not exclude a large left-right economic component, given that the traditionally economically right and left leaning parties are clearly separated.

#### 4.2.2 Parliament of 2014-18

In the 2014-18 government we see a change of coalition. After the election, by virtue of being the bigger coalition, S and MP entered the government with V as support to form a minority government. The first months of the new government saw much drama when SD voted for the budget put forth by Alliansen, causing a near re-election. The situation was resolved by the "December agreement" of 2014, in which Alliansen pledged not to vote for their own budget, and thus allow for a minority government, on the premise that if the roles were reversed after the 2018 election, then S and MP would honor the agreement. The agreement was later scrapped by Alliansen, but remained *de facto* in place until the end of the parliamentary period.

That S and MP are in the same government can be seen in Figure 5 by noting that these two parties are not separable on any of the six relevant policy axes. Alliansen, which formed the previous government, is still very much in existence, but the parties are not as tightly bound. It is therefore possible to see a distinction between the four Alliansen parties already on the first and second principal axes.

Another important difference between this parliament and the previous is that SD has grown significantly. From having 5.7% in the 2010 general election, they have 12.9% in 2014. SD, which one could classify as a part of the alt-right movement is very focused on immigration, and has a decidedly anti-establishment (or as they would put it anti-globalist) agenda. This party does not fit naturally into the left-right paradigm that usually has dominated Swedish politics, and can hence be seen as a third power in the parliament. This is also made visually quite clear by looking at the first and second axes, in which SD forms the third leg of a triangle, with Alliansen and "red-green" on the other two legs. A similar tendency can be seen already in the third axis of 2010-14, which now corresponds to the second axis in 2014-18. The shift in axes order is likely mostly attributed to the fact that SD is roughly twice as large in the later period.

Due to the less cohesive alliance, there are extra relevant policy axes, where these parties can now distinguish themselves. We note that if we treat the parties in a coalition government as one effective party, the effective dimensionality of the Swedish parliament, both for the years 2010-14 and 2014-18, is one less than the effective number of parties. The reason for this is linked with the strong cohesion that can be found inside all of the Swedish parliamentary parties themselves, together with the lack of independents.

It is reassuring to see that the simple dimensionality analysis that was performed in section 3, coincides almost perfectly with the visual inspection of the axis projections that was carried out in this section. Of course, by looking at the axis projection, much more information can be deduced than by simply considering the information content within each axis. Given the strong cohesion that we see in the Swedish parliament, we are not surprised that the dimensionality turns out to be one less than the effective number of independent parties. We wish to stress however, that since these axes are not of equal weight, some policy decisions will have had more controversy linked with them than others.

Just as for Ireland, we may construct the euclidean distances between the different Swedish parties. Please refer to Table 2. Starting with the parliament of 2010-2014 we can again clearly see the cohesion that followed from the coalition government Alliansen. The distances between these parties was in the range 6%-8%. The party

SD is, maybe not surprisingly, furthest away from V, but is almost equally far from S (94%) and MP (96%), which is a bit noteworthy given that these parties are a distance 52%-65% apart. As a comparison we may note that the distance from SD to Alliansen is 69%-72%, (which is a much tighter constellation) has only a little bit less variance than the distance to S, V and MP.

Now turning our attention to the period 2014-2018 we can here clearly see the coalition between S and MP in that their distance is only 6%. The party V, which is supposed to support the coalition government, following the "December agreement" is however at a distance from the government of 60%-61%. This is similar to the distance during the period 2010-2014 which was 52%-60%. We can see how Alliansen is not as tightly bound as the previous period when they were in government, and actually has an inter-party distance in the range of 55%, most of which is picked up on axes three, five, and six. A final interesting observation is that the distance between SD and any other party is above 90%, and is remarkably largest to M and C (100% each) and smallest to S and MP (91% each). This indicates as shift (or realignment) in the relatives positions of parties during this period. This again shows how SD in the 2014-2018 did not fit in naturally into any of the two major constellations.

A very important question that we have not really addressed so far, is what political dimensions the different axes actually do correspond to. We will discuss this in the next Section.

# 5 Labelling of the Axes

In the preceding discussion of how the parties place themselves along the most principal axes, we have been careful not to try and read to much into the underlying policies that these axes represent. This choice is deliberate, as the PCA method has no intrinsic knowledge of what the different principal axes represent. That does however not mean that a characterization of the axes is not possible, only that one needs to exercise care, as several policy stances that a human observer would consider different could coexist along the same axis. It would be naïve to expect that any single axis will be uniquely related to any particular policy, rather we should expect to see distinct policy palettes (distributions) for the different axes. In the context of two dimensional policy maps, the observation of intertwined policy axes is referred to as cutting angles, see [Jenkins, 1999] for an example.

In principle, the route to establishing labels for the various principal axes is straightforward, albeit tedious. One proceeds by assigning a label $L_\alpha$ to all of the roll-calls that make up the PCA analysis. Since we require that $\sum_{\alpha=1}^{N_r} |V_\alpha^{(k)}|^2 = 1$, we can think of $|V_\alpha^{(k)}|^2$ as the percentage with which axis $\alpha$ contributes to the $k$:th principal axis. The policy pallet is now simply constructed by summing $|V_\alpha^{(k)}|^2$ over all axes with a common label. Formally the percentage with which label $\mathcal{L}$ contributes to axis $k$ would be given by

$$p_{\mathcal{L}}^{(k)} = \sum_{\alpha \in A_{\mathcal{L}}} |V_\alpha^{(k)}|^2,$$

where $A_{\mathcal{L}}$ is the set of axes where $L_\alpha = \mathcal{L}$. See the discussion of Costello's results in Section 4.1 and Fig. 3 for an example of what this looks like.

For this particular work, the proper labeling $L_\alpha$ of all the roll-calls in the considered parliaments has not been the main focus, even though it can be done, and thus lies outside our scope. Also, it would require (a lot of) subjective interpretation of the policy labels relating to the various roll-calls, which lies beyond the skill-set of the authors.

# 6 Summary and Discussion

In this paper we have investigated the technique of Principal Component Analysis (PCA) for determining the dimensionality of three incarnations of two European parliaments, the Irish and the Swedish. PCA allows us to avoid introducing any subjective bias as to the dimensionality itself, or the nature and significance of each dimension. It can thus be used to objectively quantify political positions.

We find that the dimensionality of a parliament most often does not reflect the number parties in a Parliament. Rather it measures the number of *independent* parties (minus 1), because coalition parties are effectively one party when it comes to voting. This was clearly seen in the case of Sweden, where party discipline is high and the effective dimensionality varied between two consecutive elections as a result of different government coalitions.

In this work we found that the Irish Parliament of 2017 was at least 3 dimensional, but that up to 5 dimensions could be necessary if one also wishes to account for independents. We then compared this with the work by Costello in Ref. [Costello, 2017] which showed that there is a qualitative agreement between the two approaches to gauge the Irish political landscape.

It was noted that the parties, both in Ireland and Sweden, show a large amount of cohesion. However, the Irish system does, opposite to the Swedish, allow for independents to enter the parliament. Consequently, there is more small-scale structure in the Irish parliament than the Swedish, and potentially explains the gradual decrease in information content that can be seen between axis 3 and 5 in the Irish Parliament.

A major remaining challenge is surely the labelling of the different policy axes. Given the geometric interpretation of the PCA method in terms of a high dimensional policy-space, the construction of policy-palettes, as discussed in the main text, could be achieved by appropriately labeling the individual roll-calls that go in the analysis. However, this labeling is non-trivial, labor-intensive, and subjective. As such it lies beyond the scope of this work.

The method used here is not without its limitations. One major drawback is in the limitations of the data. It has not been recorded how long or heated any pre-roll-call discussions were, and thus we cannot distinguish the important roll-calls from the less-important ones. Sometimes, for instance, one party might disagree with the rest on a less-important sounding issue, and this could become a distinguishing feature, defining an axis. The PCA approach can easily be adapted to take "importance" of votes into account by rescaling the numerical values of "Yes" and "No" to smaller or larger values than one, but this would need to be handled with care as it opens up for interpretational bias.

We end this summary by facing the future. In our analysis of Ireland, we included only one complete year of roll-calls, so it would naturally be interesting to see if the current parliament will continue to operate with the same political map. This was the case for both the Swedish parliaments considered, where the political maps of any year showed little difference with the complete map of the whole parliamentary period.

The future of the Swedish parliament is interesting as well. After the general election in September of 2018, Sweden settled on a coalition government through "Januariavtalet" in – January – 2019. The agreement features L and C supporting a minority government consisting of S and MP through budget negotiations and comes with a list of 73 items that should be implemented during the next four years. The agreement also decisively puts and end to "Alliansen" as a unified opposition. With this in mind it will be interesting to see how future spatial maps of the Swedish parliament develop.

## Acknowledgments

The authors wish to acknowledge Rory Costello, Masud Haque, Joost Slingerland, Katarina Domijan, Getrud Fremling, Maxim Lott and Matthew Tallon for support an encouragement as well as fruitful discussions and comments on the manuscript. We gratefully acknowledge financial support from Science Foundation Ireland Principal Investigator Award 12/IA/1697 and from Netherlands Organisation for Scientific Research (NWO).

## References

[SWE, 2018] (2018). Sveriges riksdag: Riksdagens öppna data - voteringar. Web-access: http://data.riksdagen.se/data/voteringar/.

[IRL, 2018] (2018). Tithe an oireachtais / houses of the oireachtas: Dáil Éireann votes. Web-access: https://www.oireachtas.ie/en/debates/votes/.

[Ansolabehere et al., 2001] Ansolabehere, S., Snyder Jr, J. M., and Stewart III, C. (2001). The effects of party and preferences on congressional roll-call voting. *Legislative Studies Quarterly*, pages 533–572.

[Bakker et al., 2015] Bakker, R., De Vries, C., Edwards, E., Hooghe, L., Jolly, S., Marks, G., Polk, J., Rovny, J., Steenbergen, M., and Vachudova, M. A. (2015). Measuring party positions in europe: The chapel hill expert survey trend file, 1999–2010. *Party Politics*, 21(1):143–152.

[Bullock III and Brady, 1983] Bullock III, C. S. and Brady, D. W. (1983). Party, constituency, and roll-call voting in the us senate. *Legislative Studies Quarterly*, pages 29–43.

[Carroll et al., 2009] Carroll, R., Lewis, J. B., Lo, J., Poole, K. T., and Rosenthal, H. (2009). Comparing nominate and ideal: Points of difference and monte carlo tests. *Legislative Studies Quarterly*, 34(4):555–591.

[Carrubba and Gabel, 1999] Carrubba, C. and Gabel, M. (1999). Roll-call votes and party discipline in the european parliament: Reconsidering mep voting behavior.

[Clinton et al., 2004] Clinton, J., Jackman, S., and Rivers, D. (2004). The statistical analysis of roll call data. *American Political Science Review*, 98(2):355–370.

[Clinton and Jackman, 2009] Clinton, J. D. and Jackman, S. (2009). To simulate or nominate? *Legislative Studies Quarterly*, 34(4):593–621.

[Costello, 2017] Costello, R. (2017). The ideological space in irish politics: comparing voters and parties. *Irish Political Studies*, 32(3):404–431.

[Erickson, 1971] Erickson, R. S. (1971). The electoral impact of congressional roll call voting. *American Political Science Review*, 65(4):1018–1032.

[Hansen, 2009] Hansen, M. E. (2009). The positions of irish parliamentary parties 1937–2006. *Irish Political Studies*, 24(1):29–44.

[Hastie et al., 2009] Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning – Data Mining, Inference, and Prediction.* Springer, second edition.

[Heckman and Snyder, 1997] Heckman, J. J. and Snyder, J. M. (1997). Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. *The RAND Journal of Economics*, 28:S142–S189.

[Hix, 2001] Hix, S. (2001). Legislative behaviour and party competition in the european parliament: An application of nominate to the eu. *JCMS: Journal of Common Market Studies*, 39(4):663–688.

[Hix and Noury, 2016] Hix, S. and Noury, A. (2016). Government-opposition or left-right? the institutional determinants of voting in legislatures. *Political Science Research and Methods*, 4(2):249–273.

[Hix et al., 2006] Hix, S., Noury, A., and Roland, G. (2006). Dimensions of politics in the european parliament. *American Journal of Political Science*, 50(2):494–520.

[Jenkins, 1999] Jenkins, J. A. (1999). Examining the bonding effects of party: A comparative analysis of roll-call voting in the us and confederate houses. *American Journal of Political Science*, pages 1144–1165.

[Jessee, 2009] Jessee, S. (2009). Spatial voting in the 2004 presidential election. *American Political Science Review*, 103:59–81.

[Lauderdale and Clark, 2012] Lauderdale, B. E. and Clark, T. S. (2012). The supreme court's many median justices. *American Political Science Review*, 106(4):847–866.

[Merz et al., 2016] Merz, N., Regel, S., and Lewandowski, J. (2016). The manifesto corpus: A new resource for research on political parties and quantitative text analysis. *Research & Politics*, 3(2):2053168016643346.

[Poole, 2005] Poole, K. T. (2005). *Spatial models of parliamentary voting*. Cambridge University Press.

[Poole and Rosenthal, 1990] Poole, K. T. and Rosenthal, H. (1990). Patterns of congressional voting.

[Poole and Rosenthal, 1997] Poole, K. T. and Rosenthal, H. (1997). Congress: A political-economic history of roll call voting. In *Congress & the Presidency*, volume 24, page 230. Taylor & Francis Ltd.

[Poole and Rosenthal, 2001] Poole, K. T. and Rosenthal, H. (2001). D-nominate after 10 years: A comparative update to congress: A political-economic history of roll-call voting. *Legislative Studies Quarterly*, pages 5–29.

[Proksch and Slapin, 2010] Proksch, S.-O. and Slapin, J. B. (2010). Position taking in european parliament speeches. *British Journal of Political Science*, 40(3):587–611.

[Shor et al., 2010] Shor, B., Berry, C., and McCarty, N. (2010). A bridge to somewhere: Mapping state and congressional ideology on a cross-institutional common space. *Legislative Studies Quarterly*, 35(3):417–448.

[Shor and McCarty, 2011] Shor, B. and McCarty, N. (2011). The ideological mapping of american legislatures. *American Political Science Review*, 105(3):530–551.

[Snyder and Groseclose, 2000] Snyder, J. M. and Groseclose, T. (2000). Estimating party influence in congressional roll-call voting. *American Journal of Political Science*, 44(2):193–211.

[Thiem, 2006] Thiem, J. (2006). Explaining roll call vote request in the european parliament.

| Acronym | Party Name |
|---|---|
| S-PBP | Solidarity-People Before Profit |
| FF | Fianna Fáil |
| FG | Fine Gael |
| GP | Green Party |
| I4C | Independents for Change |
| L | Labour |
| SF | Sinn Féin |
| SD | Social Democrats |
| Ind. | Independent |

Table 3: Party names and acronyms of the Irish Parliament

| Acronym | Party Name |
|---|---|
| V | Vänsterpartiet |
| S | Socialdemokraterna |
| MP | Miljöpartiet |
| C | Centerpartiet |
| L | Liberalerna |
| | (before 2015: Folkpartiet) |
| M | Moderaterna |
| KD | Kristdemokraterna |
| SD | Sverigedemokraterna |
| Ind. | Independent |

Table 4: Party names and acronyms of the Swedish Parliament. Note that independents are not voted in as such, but have later defected from their original party.