PA

What Makes Party Systems Different? A Principal Component Analysis of 17 Advanced Democracies 1970–2013

Zsuzsanna B. Magyar[ID]

University of Lucerne, Lucerne, Switzerland. Email: zmagyar@ucla.edu

Abstract

Party systems, that is, the number and the size of all the parties within a country, can vary greatly across countries. I conduct a principal component analysis on a party seat share dataset of 17 advanced democracies from 1970 to 2013 to reduce the dimensionality of the data. I find that the most important dimensions that differentiate party systems are: "the size of the biggest two parties" and the level of "competition between the two biggest parties." I use the results to compare the changes in electoral and legislative party systems. I also juxtapose the results to previous party system typologies and party system size measures. I find that typologies sort countries into categories based on variation along both dimensions. On the other hand, most of the current political science literature use measures (e.g., the effective number of parties) that are correlated with the first dimension. I suggest that instead of these, indices that measure the opposition structure and competition could be used to explore problems pertaining to the competitiveness of the party systems.

Keywords: principal component analysis, party system, typology, effective number of parties

1 Introduction

In the modern world, political parties are the main vehicles of political representation. The number of parties that win seats in the legislature and the number of seats that each party holds vary widely across countries and even across years within the same country. Finding the most important features that characterize party systems is a classical dimensionality reduction problem: How a high-dimensional dataset be reduced in such a way that all the important information is retained? In this paper, I use a principal component analysis (PCA) to assess the seat (and later the vote) share matrix of the legislatures of 17 European countries from 1970 to 2013 to understand what makes party systems different from each other. This analysis allows me to compare changes in legislative and electoral party systems. In addition, it allows me to examine some of the early solutions to the dimensionality reduction dilemma. In the 1960s and 1970s, political scientists created typologies that clustered countries with similar party systems together based on qualitative studies (Duverger 1954; Dahl 1966; Rokkan 1970; Sartori 1976; Mair 2002). These typologies have not been used in quantitative studies as their application requires many decisions. First, there are several different typologies. Second, as party systems evolve, a country might have to be classified as a different "type" in different years (Kitschelt 2008).

Other political scientists have created quantitative measures based on the legislative seat and vote shares of different parties to reduce the dimensionality of the data (Rae and Taylor 1970; Laakso and Taagepera 1979; Katz and King 1999; Caulier and Dumont 2005; Rozenas 2011). The most widely used of these measures is the effective number of parties (ENP). Surprisingly, the scholarship thus far has found little evidence that the ENP influences important political outcomes. These two different scholarships—qualitative and quantitative—were never in a dialogue with each other, and neither has considered the other as a different solution for the same dimensionality reduction problem.

In this paper, using PCA, I identify the features that make party systems most dissimilar (i.e., the orthogonal dimensions in which the data vary the most). I find that the most important features are the sizes of the two biggest parties and the level of competition between these two parties. Through this analysis, I compare electoral and legislative party systems. Next, I plot party systems in different years to show how close the solutions of the earlier studies are to the machine learning approach. I show that typologies sort the countries on the two main dimensions that the PCA recovers. At the same time, I show that commonly used party system indices (such as the ENP) measure only one aspect of the party system (i.e., the sizes of the two biggest parties). I also show that some of the measures that studies use—sometimes within the same statistical model (e.g., ENP and the number of parties in the government)—are highly correlated with each other. Finally, I suggest that opposition concentration indices and competition indices could measure the competitiveness of the party systems and could lead us back to a multidimensional approach to assessing the effect of party system structure on political outcomes.

2   Previous Scholarship on Party Systems

The term "party system" refers to the number and the size of the parties in a country (Sartori 1976). Comparing party systems is not straightforward: the differences between party systems can be subtle, and changes in the number and sizes of the parties may frequently occur within a country (Kitschelt 2008).

At first, political scientists created typologies that would group similar countries together and divide the countries that had big enough differences in their party systems (Duverger 1954; Dahl 1966; Rokkan 1970; Sartori 1976). The most canonical difference in party systems across countries is between two-party and multiparty systems. Early political scientists found two-party systems to be "ideal," whereas multiparty systems weak and chaotic, because coalition governments frequently formed in these systems (Duverger 1954). In practice, however, there are very few countries with two-party systems.

The rest of the countries are multiparty countries. Within the countries with nonmajoritarian electoral systems, there is a wide variety of differently sized and structured party systems. Finding a way to group these countries for analytical purposes is a multidimensionality problem that researchers face and try to solve through different methods. Blondel in 1968 classifies party systems not only by the number of parties but also by the relative sizes of the parties (Blondel 1968).[1] Most of the typologies following Blondel (1968) sort countries based on the number of the parties as well as the extent to which those parties compete or cooperate with each other (Dahl 1966; Rokkan 1970; Sartori 1976; Mair 2002; Laver and Benoit 2015). Authors focus either on the sizes of the parties that oppose each other (whether they can govern alone or need a coalition partner; Rokkan 1970; Laver and Benoit 2015), or how easy it is for a new party to enter the system (Sartori 1976; Mair 2002). While these typologies solve the multidimensionality problem, it is difficult to use them in quantitative studies. There are several different typologies, and it is not always clear which one should be used. In addition, party systems evolve, and the category that a country belongs to should be re-evaluated after every election.

Thus, in contrast to some authors who sort countries into party system types, other authors use quantitative data (the seat shares and vote shares of parties) to create measures of the party system size and structure. The first approach to this problem is to summarize the entire party-size distribution histogram with a single measure. The first measure of choice is entropy, which counts the number of ways we can rearrange the histogram of party size shares while arriving at the same histogram. Kesselman names this *hyperfractionalization index* (Kesselman 1966; Bishop 1995). Later, Katz and King (1999) and Rozenas (2011) use the entire vote share distribution of

---

1   For a detailed summary of all the typologies, see Online Appendix D.

the parties (in a log-ratio transformed way) as an outcome variable to predict election results. Thereafter, political scientists adopt a measure from the economics literature: the *Herfindahl–Hirschman* concentration index (HH).[2] Rae and Taylor (1970) calculate a party *fractionalization index* by changing the formula to 1-HH or 1- $\sum s_i^2$, where the $(s_1, ..., s_n)$ are the legislative seat shares of the parties. They change the direction of the original measure such that the more parties there are in a party system, the higher the index value of that system.

Laakso and Taagepera (1979) argue that the previous measures give too much weights to smaller parties, which means the entry or an exit of a minor party may change the measure a lot. In addition, similar party systems may end up with very different index numbers (Laakso and Taagepera 1979). To avoid these problems, Laakso and Taagepera (1979) decide to weigh bigger parties in the party system more compared to the fractionalization and the hyperfractionalization indices. Laakso and Taagepera (1979) thus transform the *HH* index in a new way as $1/HH = \left(\frac{1}{\sum s^2}\right)$. Unlike the fractionalization index, this measure does not produce a value between 0 and 1; instead, it ranges from 0 to ∞, and shows roughly how many equally sized parties there are in the party system. This index is the aforementioned *ENP*.[3] Currently, the ENP is probably the most widely used measure of party system concentration.

Some authors argue that we should weigh the seat shares of bigger parties even more heavily than prescribed by the *ENP* measure, because only big parties have any chance of entering the government (Sartori 1976; Molinar 1991; Kline 2009). Some of these approaches suggest weighing a party's seat shares more heavily if it has a higher coalitional potential or "power" or if it is pivotal to any given coalition (Caulier and Dumont 2005; Grofman 2006; Kline 2009).[4] On the extreme end of this argument, some suggest that we should use the size of only the biggest party to characterize the entire party system (Taagepera 1999; Dunleavy and Boucek 2003). All of these latter measures have been criticized, because they create the opposite problem from the entropy type indices: they are not sensitive to small changes. This means that many different configurations of party sizes (party systems) can produce the same value (Kline 2009). In practice, the more weight big parties receive, the more different party configurations are associated with the same index number.

Overall, there is a trade-off between how comprehensively we would like to describe the party system on the one hand and how much we would like to identify how many big (relevant) parties there are in a given country on the other hand. Whether we are interested primarily in the number of parties or more in the evenness of the distribution. The former approach yields a measure that weights smaller parties more, while the latter yields a measure that weights larger parties more. In most recent empirical studies that evaluate whether certain factors influence government policies, the authors choose one or more controls for the party system size (usually the *ENP*). Contrary to the typologies, this line of literature has paid little attention to which features make party systems different from each other or which features different measures quantify.

3 Data

In this paper, I use PCA as a dimensionality reduction technique to explore the underlying structure of a party system dataset. To do this, I use the party seat shares in the legislature of 17 European

---

2 Originally, this measure was developed in economics, and it is the sum of squares of the market share of each company in a given market. The range of this index is 0–1, where a 1 means that the market is dominated by one company and 0 means that all companies are equal.

3 Laakso and Taagepera (1979) show that the general form of their index is $N_a = \left[\sum_{i=1}^{N} p_i^a\right]^{1/(1-a)}$, where N is the ENPs, and p is either the seat of vote shares of the parties. The same formula is also the general form of the fractionalization and the hyperfractionalization indices. The value of the a parameter determines how sensitive the index is toward small parties or big parties. If a is close to 0, then small parties are weighted more, while if it is close to ∞, bigger parties are weighted more.

4 Caulier and Dumont (2005), Grofman (2006), and Kline (2009) suggest using the sum of squared power indices instead of the seat shares of the parties in the *ENP*'s formula to prevent overvaluing small parties. The Shapley–Shubik power index shows how many times a party would be pivotal in coalitions (Shapley and Shubik 1954), while the Banzhaf index measures how many times a coalition would shift from winning to losing if a particular actor were to change their vote (Banzhaf 1965).

countries from 1970 to 2013.[5] In each row (country-year) of the matrix, I rank each party based on its size. The first variable is the number of seat shares held by the biggest parties; the second variable is the number of seat shares held by the second-biggest parties, and so on. Thus, the dataset does not contain the identity of any individual party, but it still allows me to compare the structure of the party systems across countries. If all the parties have been accounted for in a given country-year, a value of 0 is entered into the next column. The number of parties ranges from 3 (in Austria from 1970 to 1986 and in Greece from 1981 to 1984) to 20 (in Italy in 2006 and 2007). The party seat share data comprise a compositional dataset. The party seat shares add up to one ($\sum s_i = 1$). This means that within any given case, the size of each data point depends on the size of the others (Aitchison 1983). The matrix that I create has 675 rows and 20 columns. I use party seat shares foremost, so that I can compare the results to typologies and seat-share-based indices in the second part of the paper.

4 Principal Component Analysis

4.1 Simple PCA

First, I conduct a PCA to explore the structure of the party system dataset. This type of analysis is a popular statistical approach that can help with the analysis of complex datasets (Dennis 1988; Downes 1988; Ben-Ur and Newman 2002; Kestilä 2006; Zoco 2006; Guzmán and Sierra 2009; Jakulin *et al.* 2009; Falcó-Gimeno and Vallbé 2013; Akkerman, Mudde, and Zaslove 2014; Savoy 2015). The goal of the PCA is to recover the minimum amount of information needed to reproduce the maximum information present in the data matrix. This method projects high-dimensional data onto a lower-dimensional space. The lower-dimensional space is determined by the directions in which the data vary most. As such, the least amount of information possible is lost during projection. The projection is linear, meaning that it can create an accurate summary of the data if the data are Gaussian-distributed.

In practice, PCA estimations begin with a calculation of the dimension in which the data have the most variation. It then finds n orthogonal dimensions that explain the most significant part of the remaining variation within the data.[6] Within the result, the eigenvalues are the scale, and the eigenvectors are the direction of the new (reduced) dimensions. The eigenvectors are the principal components. The eigenvector with the highest eigenvalue is the first principal component of the dataset (i.e., the dimension in which the data have the most variation or the direction in which the data are the most dissimilar). The second principal component is the second eigenvector, and it is orthogonal to the first dimension, and so on.

I analyze the matrix by row. I project the 675-dimensional party seat share data matrix to a 20-dimensional space (the number of variables). To do this, I first mean center the data. Then, I find the eigenvalues and the eigenvectors of the covariance matrix. I do not scale the variables, because all the variables are seat shares and such commensurable variables do not require scaling (Jolliffe 2002). Moreover, because the seat share data are compositional, scaling would change the crucial

---

5   The countries in the dataset are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, and the United Kingdom. Data are available at https://doi.org/10.7910/DVN/Q0OIMY. Countries that democratized later than 1970 appear in the dataset after the first democratic elections. These dates are 1975 for Portugal and 1977 for Spain. In addition, the German data refer to West Germany before 1990, and the Unified Germany after 1990.

6   Mathematically, we would like to map vectors x^n in a d-dimensional space onto z^n in an M-dimensional space ($M < d$). This means that we need a transformation matrix, a set of vectors that can help with this mapping. If we represent x as a linear combination of orthonormal vectors such that $x = \sum_{i=1}^{d} z_i u_i$, we can write this problem in the following way: we are looking for such and x where $\tilde{x} = \sum_{i=1}^{M} z_i u_i + \sum_{i=M+1}^{d} b_i u_i$ (Bishop 1995). To arrive to the optimal solution, the PCA (similarly to the regression) reduces the sum of squared errors of this approximation. The error can be written as $x^n - \tilde{x}^n = \sum_{i=M+1}^{d} (z_i^n - b_i) u_i$; as a consequence, we minimize $E_M = \frac{1}{2} \sum_{n=1}^{N} ||x^n - \tilde{x}^n||^2 = \frac{1}{2} \sum_{n=1}^{N} \sum_{i=M+1}^{d} (z_i^n - b_i)^2$. If we calculate this optimization, we can show that this minimum occurs when $\Sigma u_i = \lambda_i u_i$, where Σ is the covariance matrix of the set of vectors, and λ are the eigenvalues (Bishop 1995). Thus, this is an eigen decomposition of the covariance matrix.
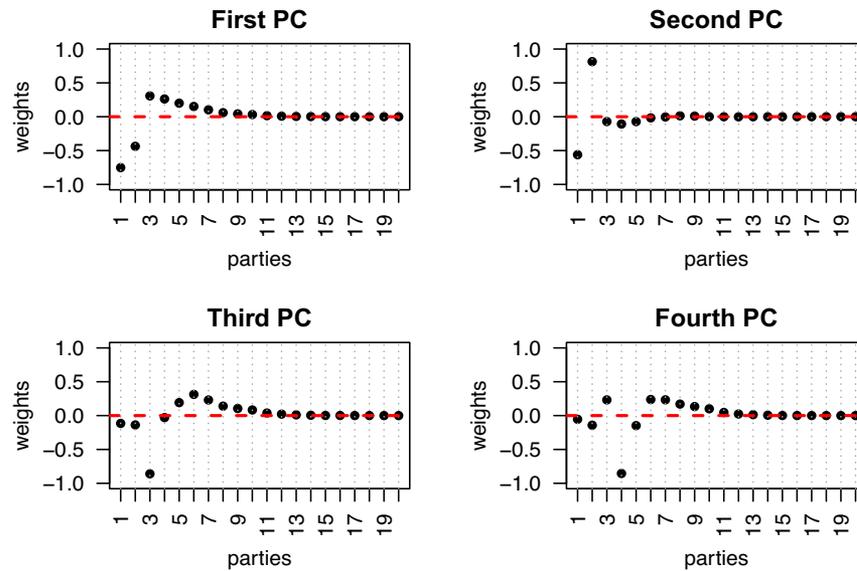

Figure 1. Loadings, PCA. *Notes:* The plot shows the loadings, the weight of parties in determining the principal components in the PCA.

condition that the seat shares add up to 1. Thus, in this case, scaling may not be an optimal solution (Online Appendix B).

One of the benefits of PCA is that the lower dimensions of the data are more easily interpretable than they are in a complex dataset (Jolliffe 2002). The first four principal components illustrate what makes party systems most unalike. As Figure 1 shows, the first dimension (PC1) contrasts countries where the two biggest parties are much larger than the other parties with countries where the two biggest parties are not much larger than the other parties. I call this dimension "Size of the Biggest Two Parties." The second dimension (PC2) contrasts countries where the two biggest parties are similar in size with countries where the biggest party is significantly larger than the second-biggest party. This dimension contrasts countries that have two-party competition with countries that have one dominant party. We can understand this dimension as the "Competition Between the Biggest Two Parties." The third dimension (PC3) is most heavily influenced by the size of the third party, and so it contrasts countries that have a big third party with countries that have a small third party (we call this dimension "Third Party"). The fourth dimension (PC4) is somewhat unclear. This dimension is influenced by the third, fourth, and fifth parties; tentatively, I call it "Multipartism." Based on the plot, this last dimension may separate countries that have balanced multiparty systems from countries in which the third to the fifth parties are nonexistent or very small. All dimensions (PC1, PC2, PC3, and PC4) are orthogonal to each other. The purpose of the PCA is to reduce the dimensionality of the matrix. Therefore, the next step is to evaluate the appropriate number of dimensions to use for the analysis.

Previous studies suggest different rules for choosing which PCA dimensions to analyze (Jolliffe 2002; Peres-Neto, Jackson, and Somers 2005). According to Jolliffe (2002), a good rule of thumb is to use PC dimensions that explain between 70% to 90% of the (cumulative) percentage of the total variation in the data. In the following sections, I will focus on the first two dimensions, because as Table 1 shows the first two eigenvalues explain 87.12% of the variation in the data. Figure 2 shows the scree plot of the PCA.[7]

---

7 There is a significant scientific debate in the literature about the number of PCA dimensions to retain. While the broken stick rule examination of the scree plot suggests that I retain two eigen values, other methods such as the Kaiser criterion suggest that I should retain three. In Online Appendix G, I examine this dimension. I find no connection of it to previous party system typologies or indices.

Table 1. Eigenvalues and the explained variance, PCA seat shares.

	Eigenval.	Expl. variance	Cum. variance
1	0.0194012464	71.04	71.04
2	0.0045310786	16.59	87.63
3	0.0020397785	7.47	95.10
4	0.0006762824	2.48	97.57
5	0.0003635705	1.33	98.91
6	0.0001867591	0.68	99.59
7	0.0000764051	0.28	99.87
8	0.0000170232	0.06	99.93
9	0.0000113822	0.04	99.97
10	0.0000036437	0.01	99.99
11	0.0000023248	0.01	100.00
12	0.0000009147	0.00	100.00
13	0.0000003156	0.00	100.00
14	0.0000000699	0.00	100.00
15	0.0000000216	0.00	100.00
16	0.0000000152	0.00	100.00
17	0.0000000039	0.00	100.00
18	0.0000000007	0.00	100.00
Sum	0.0260737		

Notes: The table shows the eigen values, the explained variance, and the cumulative variance explained by the principal components.
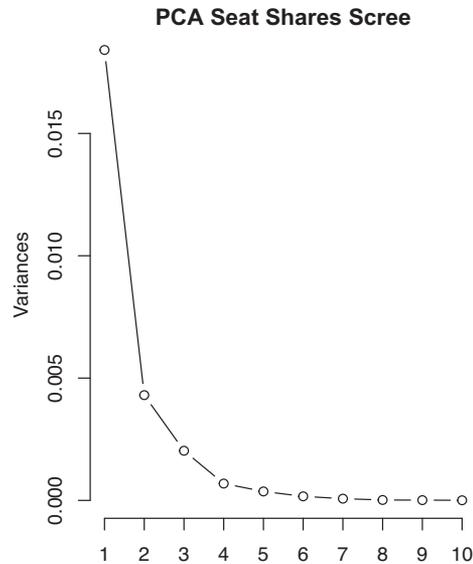

PCA Seat Shares Scree

Figure 2. Scree plot, PCA seat shares. *Notes:* The plot shows the scree plot of the PCA on seat shares. On the *x*-axis are the numbers of the eigenvalues, and on the *y*-axis are the unexplained variance.

Figure 3 shows how some selected countries' legislative party systems have changed from 1970 to 2013 on the two first dimensions. The advantage of the PCA that it can handle the dynamics of party system changes. The PCA shows that neither competition nor the weights of the two biggest parties change in some countries, such as in Iceland and Finland between 1970 and 2013 (Online Appendix A). The same is true for Western Germany (Figure 3c). On the other hand, the United
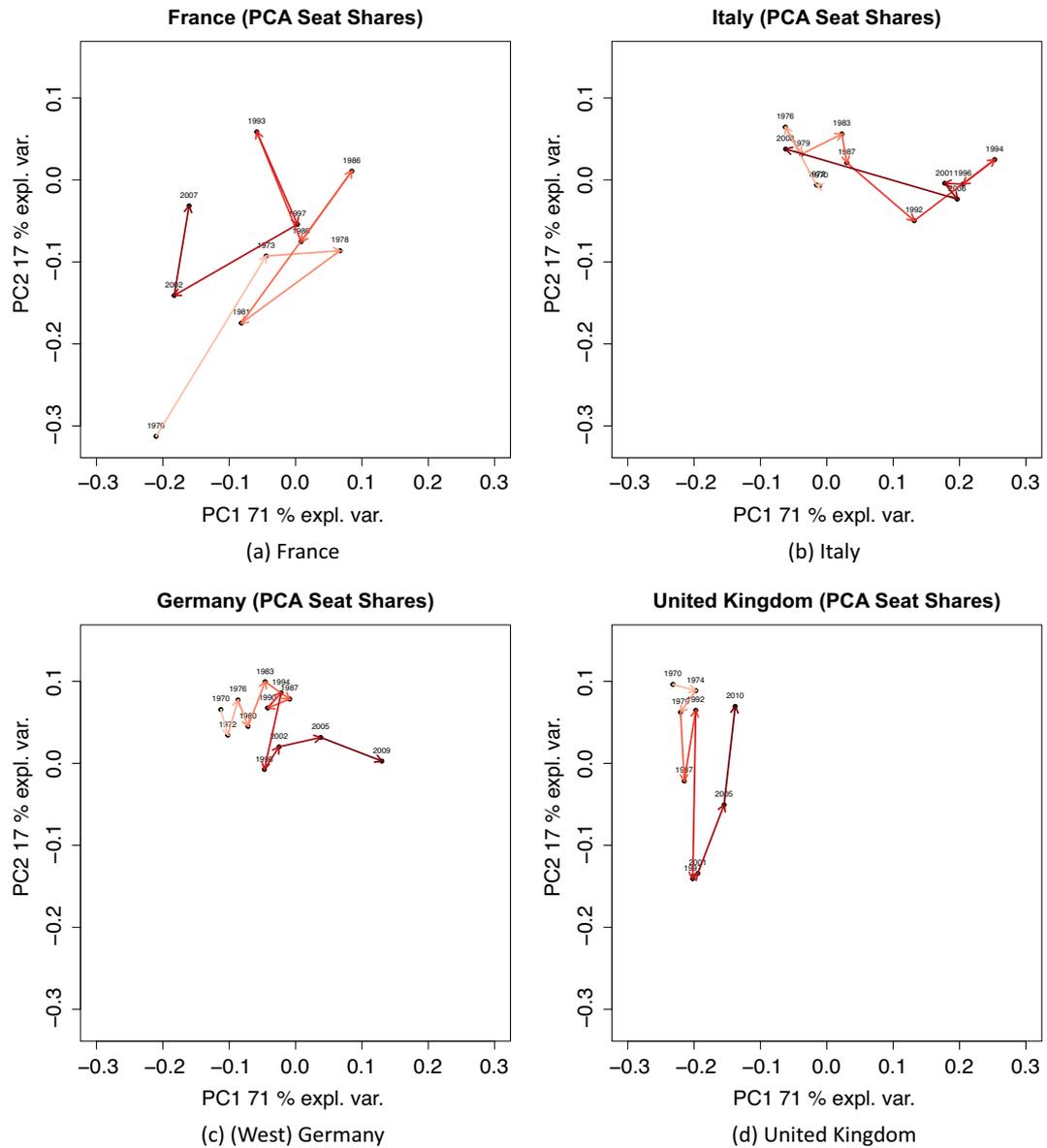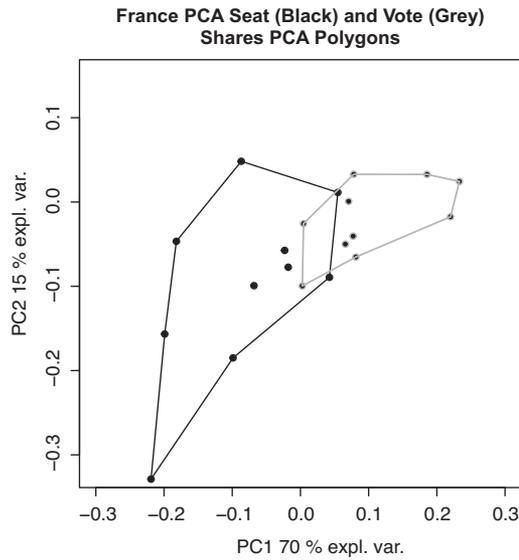
Figure 3. Individual countries in the PCA two-dimensional plane. *Notes:* The plots show how the countries' party systems change along the PC1 (size) and PC2 (competition) dimensions. The shading of the arrows shows the time progression, with darker colors signifying times closer to the present day.

Kingdom has shifted noticeably in the second dimension (Figure 3d). This means that while the country stayed close to a two-party system, the relative sizes of the two big parties have changed.

Some countries stay in the same place in *Dimension 2* (i.e., the competition between the two biggest parties stay the same) but move within *Dimension 1* (sometimes the two biggest parties have a comfortable advantage, while other times the other parties grow), such as in Austria, the Netherlands, Belgium, and Italy (Online Appendix A). Figure 3b shows how the 1993 electoral reform in Italy increased the fragmentation of the party system at first, but then by the 2010s, the size and the competitiveness of the party system returned to the preform era.[8] In addition, some countries—notably France, Portugal, and Norway—move in both directions. Figure 3a shows the immense changes that the French party system went through over the past decade, starting
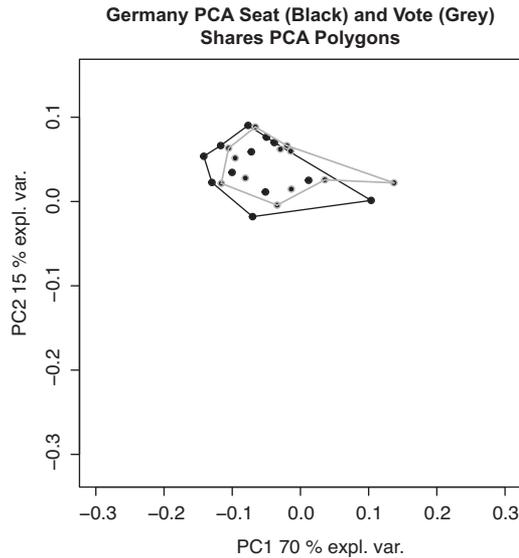
---

8 At this time, the country changed its electoral system from proportional representation (PR) to a mixed-member majoritarian electoral system.
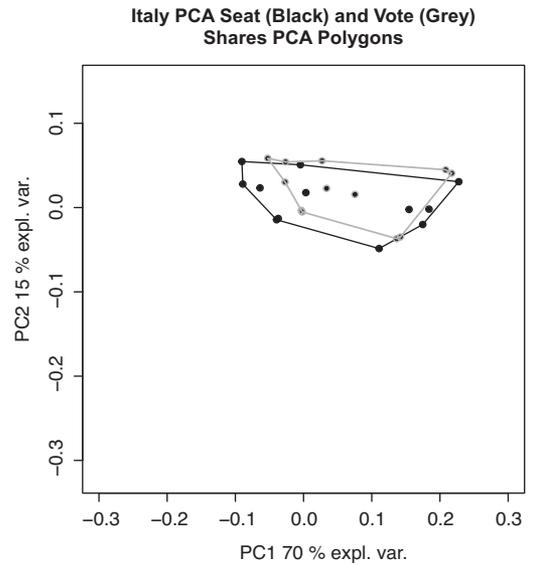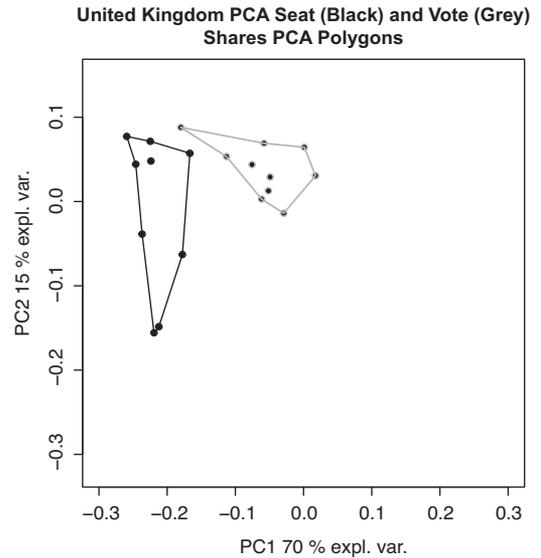
Figure 4. Electoral versus legislative party system changes 1970–2013. The party system changes throughout the years are plotted in the two-dimensional plane determined by a PCA on both seat and vote shares. In black, I plot the outer polygon surrounding all the positions of a given country based on seat share PCA, whereas in dark grey, based on vote share PCA.

from the dominance of the UDR (Union for the Defense of the Republic) in 1973 before increases took place in terms of party fragmentation and the overall competitiveness of the party system. Eventually, the country had an almost British-style balanced two-party system by the 2010s.

We can quantify these changes. Each country's party system changes can be superimposed by a polygon (Kollár-Hunek *et al.* 2008). The area of the polygon represents how much the country's party system changed between 1970 and 2013 (Figure 4a–d). In addition, we can calculate the standard deviation of all the data points for each country on the PC1 and on the PC2 dimensions ($Sd_{diffxy} = Sd_x - Sd_y$). This measure is positive when in a country the sizes of the two biggest

parties relative to other parties changed more throughout the years, while it is negative when the competition between the two biggest parties changed more. It is close to 0 if the changes are roughly the same. I also calculate the area–perimeter ratio of the polygons A/P. This measure is high if the polygon is elongated in any direction and low if the party system of the countries changes on both dimensions roughly the same way over the years. According to Figure 4a, France has the biggest polygon area (0.038). The United Kingdom, Italy, and Germany have polygon s that are roughly one-third the size of France's, which means they have experienced much less change in their party systems between 1970 and 2013. In the case of France, the Sd_{diffxy} is close to 0 (−0.0012) while the party systems of Italy and Germany change more in terms of the sizes of the biggest two parties (0.0824 and 0.0321), whereas the United Kingdom's party system changes more in terms of the competitiveness of the biggest two parties (−0.0638) in the same period.

4.2 Difference Between the Electoral and the Legislative Party Systems

While many parties can get (some) votes in elections, not all of these votes translate to seats in the legislature. How the vote shares translate to seat shares depends on the permissiveness of the electoral system, the threshold to enter the legislature, and the geographical concentration of voters who prefer the same party. I call the number and sizes of parties that get votes in elections: electoral party system. A legislative party system, on the other hand, is the number and sizes of the parties in the legislature.[9] Next, I use the PCA results to compare the changes in the electoral and in the legislative party systems of the counties throughout the years.

First, I conduct a PCA on the vote shares of the parties in the previous 17 European countries from 1970 to 2013. The first two principal components that I find through the analysis of the vote shares are highly correlated with the first two principal components that I find through the analysis of seat shares (both have a correlation of $cor = 0.98$). To be able to compare the legislative and the electoral party systems, I conduct a joint PCA of the seat shares and the vote shares. I create a matrix with both the vote shares and the seat shares of the parties from 1970 to 2013 in the 17 European countries I have in my sample. The matrix consists of 1,350 rows and 20 columns. Each country-year is represented as two rows—one with the seat shares of all parties, and one with the vote shares of all parties. Similarly to the previous analysis, I center the variables, but I do not scale them, and I analyze the matrix by row.

Next, I plot the legislative and the electoral party system changes of each country on Dimensions 1 and 2 (Online Appendix F). Around the points that represent the positions of the countries in each year, I superimpose outer polygons. In most of the countries with PR electoral systems, the electoral and the legislative party system changes are almost identical between 1970 and 2013 (Austria, Belgium, Denmark, Finland, Iceland, the Netherlands, and Sweden). In Figure 4a–d, I present the same countries as in Figure 3. These countries either have single member district (SMD) electoral systems (the United Kingdom and France) or mixed-member electoral systems (Germany and Italy), which means that vote shares do not directly translate to seat shares. In the plots, the polygons represent the legislative party system change (black) and the electoral party system change (grey) of the country. Figure 4a–d shows that even though Italy and Germany have

---

9 Among the 17 countries that I have in my dataset, there are countries with highly proportional representation (PR) electoral systems: Austria, Belgium, Denmark, Finland, Iceland, Greece (before 1989), France (in 1986), Italy (before 1993), Luxembourg, the Netherlands, Norway (after 1989), and Sweden. These electoral systems are considered to be permissive. Some countries use PR electoral system, but each district has few representatives, so the allocation of seats is less proportional: Spain, Portugal, and Greece (after 1989; the winner party here gets 50 extra seats). Germany and Italy have mixed-member electoral systems. Under this system, the voters have both a PR list vote and a single member district (SMD) vote. Finally, the United Kingdom, Ireland, and France have single-member district electoral systems (SMDs). Ireland has a Single Transferable Vote electoral system, in which the vote is transferred to another candidate if the preferred candidate of the voter is eliminated. France has a two-round majority-plurality runoff electoral system. The SMD electoral systems are the least permissive, and the seat shares of the parties can diverge significantly from the vote shares of the parties (Powell 2000).

mixed-electoral systems, the countries experienced very similar changes in their electoral and in their legislative party systems.[10] In case of France and the United Kingdom (both with single-member districts), the changes in the legislative party systems are much more radical than the changes in the electoral party systems. I calculate the differences in the area between the electoral and legislative party system change polygons. I also calculate the area–perimeter ratio of the polygons *A/P*. In addition, again we can use the standard deviation differences of the *x*-axes and *y*-axes to quantify in which direction do electoral systems distort the party systems. France's legislative party system change polygon is three times as big as its electoral party system change polygon. The area–perimeter ratio suggests that the competitiveness in the legislature changes almost the same amount as the sizes of the two biggest parties, whereas the electoral party system mainly changes in one dimension. In the United Kingdom, this measure is almost the same for the legislative and the electoral party systems. In France, the standard deviation difference on *Dimension 2* is twice as big as on *Dimension 1* (Figure 4a). The results are similar in case of the United Kingdom (Figure 4d) in which case the impact of the electoral system is almost limited to *Dimension 2*.

Overall, with the separate vote and seat share PCA, I find that the same features are the most important two dimensions separating the electoral and legislative party systems. A joint PCA and a comparison where the countries fall on the first two dimensions shows the effect of the different electoral systems: the legislative party systems change more (especially on the competition dimension) in countries with more majoritarian electoral systems, while the electoral and the legislative party system changes are almost identical in countries with PR electoral systems.

4.3 Issues with the Data Structure

The data are compositional and sparse. I address these issues with several further analyses. The party seat share data are a compositional dataset, as the party seat shares add up to one. Some technical issues arise when we conduct a PCA on a compositional dataset, as the data points in each row are not independent from each other (Aitchison 1983). Therefore, the correlations between the variables might have a negative bias (Jolliffe 2002).[11]

I conduct a principal component factor analysis on the data to evaluate the common and the specific variance of the variables. Similarly to the PCA, a factor analysis reduces the original dataset into a smaller number of dimensions or factors. The assumption behind this method is that there are latent variables (factors) that drive variation in a set of observed variables. These factors are linear combinations of the latent variables. Thus, the variation in the variables can be decomposed into common variation (also known as communalities) and variable-specific variation. I anchor the factors to the principal component dimensions. The communalities of the two biggest parties (0.0118 and 0.0064) reveal that the biggest parties vary the most with the other parties. The third party however, has the highest specific variance.

In addition, a compositional dataset does not have *subcompositional coherence* (Aitchison 1983). This means that if we have only a subset of the data, the PCA on the covariance of this subset will lead to a different result from an analysis on the entire dataset. To address these concerns, I conduct the PCA on the log-ratio transformed data (Aitchison 1986; see the analysis in Online Appendix B).[12] The PCA on the transformed variables sorts the countries in groups based on the

---

10   Germany has a mixed-member proportional electoral system which explains this. The Italian electoral system changed in 1993 from PR to mixed-member majoritarian electoral system. However, the analysis shows that this change did not impact the conversion between votes and seats significantly.

11   If we have a *D* part composition [$x_1, \ldots x_D$], where $\sum_{i=1}^{D} = 1$, the covariance is going to be $cov(x_1 x_1 + \cdots + x_D) = 0$; thus, $cov(x_1, x_2) + \cdots + cov(x_1, x_D) = -var(x_i)$. This means that there will be at least one negative element per row in the covariance matrix (Aitchison n.d.).

12   This transformation makes the observations uncorrelated, and solves the issue of subcompositional coherence (Aitchison 1983).

---

number of parties in the legislature. Thus, the result is not very informative. This is because party systems have different sizes and there are a lot of zeros in the matrix. If there are a lot of zeros and small values in the dataset, performing the PCA on the original dataset can be more informative than any of the other approaches, as the absolute variation in the variables may be an important feature of the data (Baxter and Freestone 2006).

I also relax the linearity assumption of the PCA. I use kernel principal component analysis (kPCA) and nonlinear component analysis (NLPCA) to analyze the data. The kPCA offers one solution to how to find the appropriate reduced dimensional space if the data are nonlinear. With this method, we first map the data to a higher-dimensional nonlinear feature space, and after that, we run a normal PCA on the transformed data (Schölkopf, Smola, and Müller 1997). Even though we cannot extract the loadings from this estimation process directly, the biplots of the analyses show that that the first two principal components are very similar to the first two principal components I obtained from the simple PCA. The kPCA thus shows that the untransformed data are close to a Gaussian distribution (Online Appendix C).

I also conduct a NLPCA on the data. Contrary to the kPCA, the nonlinearity of the NLPCA does not come from the transformation of the space on which we project the data, but from the potentially nonlinear optimization of the data matrix. The traditional PCA minimizes the loss function over the eigenvectors and eigenvalues. The NLPCA also minimizes the loss over the admissible transformations of the data columns (de Leeuw 2005). The dimensions that the NLPCA recovers are also similar to the ones that the simple PCA finds. These analyses reveal that the result we got with the simple PCA does not depend on the compositional nature of the data (Online Appendix C).

5   How Do the PCA Dimensions Relate to Traditional Typologies and Party System Size Measures
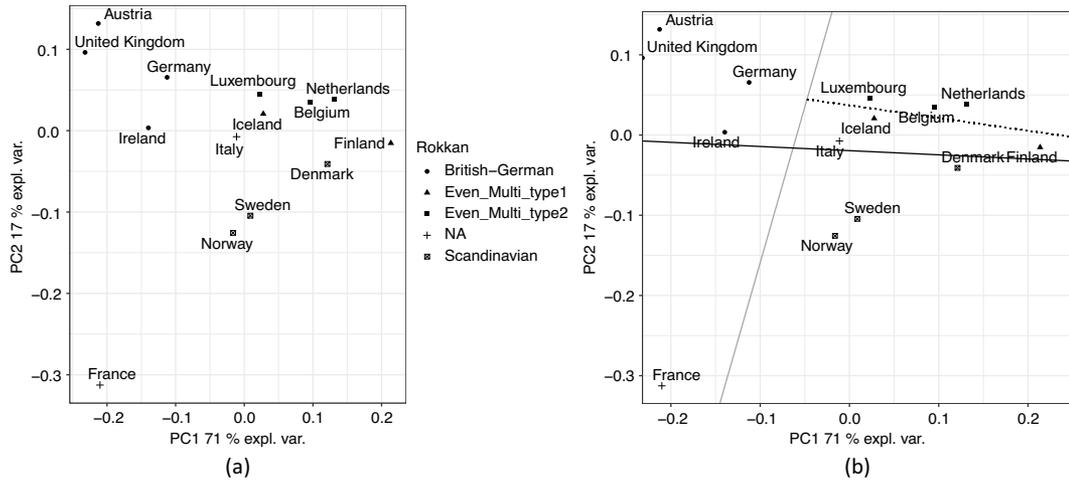
5.1   Typologies

Throughout the years, several authors have developed new party system typologies. In theory, the typologies should be static, and countries could exit and enter a group. However, the groups themselves should not change. Despite this, the authors find slightly different ways of splitting the countries. Through the PCA, one can understand better why periodically new typologies emerged, as well as why few typologies have been developed in recent years. The PCA shows that at different times, different groups of countries can become similar to each other on the two-dimensional plane. This means that all typologies could have been correct at a certain point in time and may have lost their relevance a decade later.

It is probably fair to evaluate the previous party system typologies based on how the party systems were situated on the two-dimensional plane of the PCA *vis-a-vis* each other when the relevant studies were written. In this paper, I show this with the typology of Rokkan (1970) and compare his typology to the dimensions that the PCA finds in 1970.[13]

Rokkan (1970) classifies countries based on whether the parties are roughly the same size or whether there are one or more dominant parties opposed by small parties. Rokkan constructs three categories: (1) the British–German system, in which one party faces another or another big party (and maybe a small one); (2) the Scandinavian system, in which a big party faces three to four small parties; and (3) even multiparty systems, in which there are two to five parties of about the same size. Rokkan splits the final group into two subcategories: (3a) Scandinavian split working-class systems, in which workers do not back the same parties, and (3b) segmented pluralisms, in which different identity groups create roughly equal-sized parties (Rokkan 1970). I show the

---

13   In Online Appendix E, I show the same analysis with Sartori's classification (Sartori 1976) and compare it to the PCA in 1976. I also present in Online Appendix E all the plots with the classifications of Sartori and Rokkan.

Figure 5. Comparison of Rokkan's (1970) typology to PCA results. *Notes:* Figure 5a shows the first two dimensions of the PCA and the classification of Rokkan. Figure 5b shows the demarcation lines between groups. Top-left: Category (1). Top-right: Category (3)—Top(3b) and Bottom(3a). Bottom-right: Category (2).

classification of Rokkan in 1970 along with the countries on the two-dimensional plane in 1970 in Figure 5. In Figure 5a, I designate different symbols to different categories. The PCA quite clearly separates the different groups.

In Figure 5b, I draw the demarcation lines that separate these groups. I estimate these lines through logistics models.[14] With the help of this plot, I can evaluate the original typology of Rokkan. In addition, the plot helps in sorting the countries in other years and/or countries that are not in the original typology into types. First, Figure 5b shows that Rokkan does not create one category that logically should exist. He does not include a category for countries that have few parties, one of which is dominant. In the plot, these countries should be placed in the bottom-left quadrant. It could be the case that no such country existed in 1970, but the plot shows that there is a country that Rokkan does not classify—France—which could fall into this category. Indeed, in 1968, in the French legislative elections, the UDR (Union of Democrats for the Republic) won 354 of the 458 legislative seats of the French National Assembly, while the FGDS (Federation of the Democratic and Socialist Left) ended up a distant second with 57 seats. Rokkan also did not classify the Italian party system in 1970. With the help of Figure 5b, we can infer that Rokkan probably would have classified Italy as a "Scandinavian split working class" party system along with Finland and Iceland (Category [3a]) at this time.

---

14 In the logistics regression models, I regress positive and negative cases of the different categories on the PC1 and PC2 dimensions. The only exception is the dividing line between Rokkan's two groups in Category (3)—(3a) and (3b)—in which case the negative and the positive cases only belong to the two subcategories. I use the coefficients of the logistics regressions to draw the separating lines.
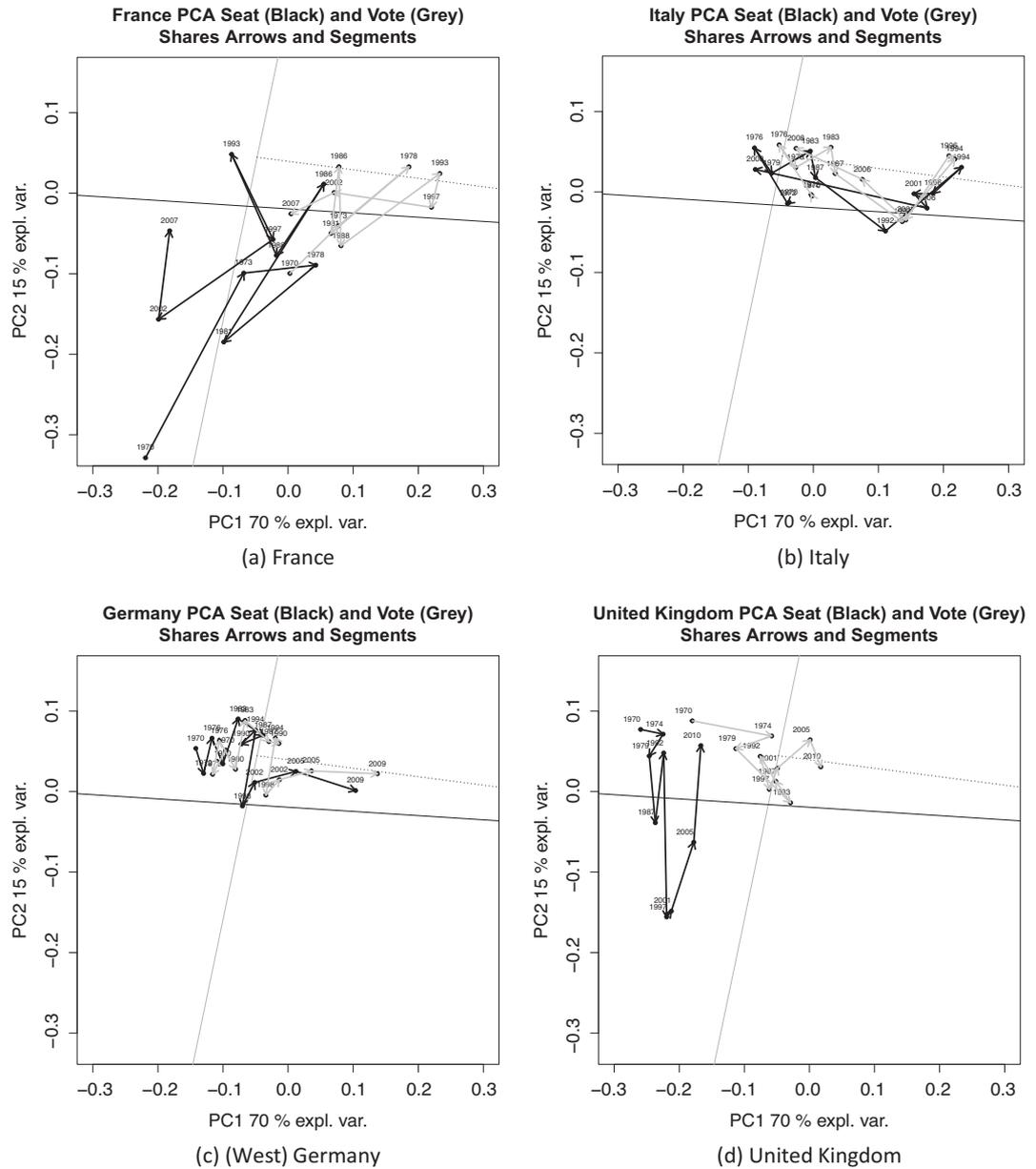
Figure 6. Individual countries in the two-dimensional plane. *Notes:* The figures show how the countries' party systems change along the PC1 (size) and PC2 (competition) dimensions through time. The lines represent the demarcation lines of different categories that Rokkan (1970) established. Top-left: 1. "British–German" party systems. Top-right: 3. Even multiparty systems—3a (bottom): Scandinavian "split working class" systems; 3b (top): Segmented pluralism. Bottom-right: "Scandinavian" party systems.

The classification of all the 17 countries between 1970 and 2013 can also be inferred based on the PCA. In Figure 6, I plot the paths through which the legislative and the electoral party systems evolved from 1970 to 2013 in France, Italy, Germany, and the United Kingdom superimposed with the demarcation lines that the analysis recovered. In this way, we can identify the times when individual countries crossed key thresholds in the typology of Rokkan (1970). The electoral party system changes are very similar to the legislative party system changes in most countries. Italy has an even multi(legislative and electoral)-party system in most of these years, but it is perhaps closer to a two-party regime in 1976 and 2002. Germany is a two-party regime until the unification of the country in 1990, at which point it becomes an even multiparty system. The trajectories

of the legislative and the electoral party system changes are different in case of countries with majoritarian electoral systems, in case of the United Kingdom and France. As for the legislative party system, France is probably a dominant party regime after the elections of 1970, 2002, and 2007; a "Scandinavian-type" party system in 1973–1981, 1988, and 1997; an even multiparty system in 1986 (the only year it has PR elections); and close to a two-party system in 1993. In 1968, which is the only PR elections in the country, the legislative and electoral party systems of France are very close to each other. Based on the electoral party system, the country starts as a Scandinavian system (2) in the 1970s and 1980s and becomes Scandinavian "spit working class" system (3a) in the late 1990s and early 2000s, and it is a segmented pluralism in 1978, 1986, and 1993. The United Kingdom's legislative party system is effectively a two-party regime in most years, with the Labour Party dominating in the late 1990s and early 2000s. The electoral party system also starts out as a two-party system, but the country moves toward multiparty system in latter years. The plot shows that in the late 1990s and early 2000s when there are more parties in the electoral party system, one of the parties gains a significant majority in the legislative party system. This shows the spoiler effect of the small(er) parties in a majoritarian electoral system.

5.2   Indices and PCA

We can also compare the PCA results to the quantitative party system size indices that previous scholarship has been identified. Based on the literature review, below I compare the PCA results to several measures. From indices that are more sensitive to the sizes of small parties to indices that are less sensitive to the sizes of small parties—*HH, Fractionalization Index* (Rae 1967), *Entropy* (Kesselman 1966)—*ENP* (Laakso and Taagepera 1979), *Shapley ENP* (ENP in which I replace the parties' seat shares with their Shapley–Shubik indices; Grofman and Kline 2011), and the size of the Biggest Party in the legislature normalized with the number of legislative seats (Taagepera 1999; Dunleavy and Boucek 2003). In addition, I include the Number of Parties in the Government, as this measure became popular for finding political outcomes (Bawn and Rosenbluth 2006). In addition, I include the *Number of Parties in the Government*, as this measure became popular for finding political outcomes (Bawn and Rosenbluth 2006).

I represent the correlations between these measures and the first two dimensions of the PCA in the first seven columns and bottom two rows of Table 2. This correlation matrix shows the Pearson correlation coefficients between the variables and their significance levels. All of these measures are highly correlated (0.7–.98) with *Dimension 1* (PC1) and with each other. Through this method, we can implicitly compare the typologies with the party system size measures. As we can see, these indices relate closely to *Dimension 1* and, thus, to the typologies of Duverger (1954) and Blondel (1968). Overall, this analysis shows that ENP distinguishes multiparty countries from two-party systems. However, this variable may not be the best solution when our theory calls for the measurement or control of interparty competition.

To measure the competition within the party system, we might calculate indices that measure the structure of the opposition. The concentration of the opposition could be measured similarly to the party system size. This notion has been discussed previously in a few studies (Maeda 2010, 2015). However, very few political science research have used these measures. The *Effective Number of Opposition Parties (ENOP)* is calculated the same way as the *ENP* suggested by Laakso and Taagepera (1979), except I calculate the measure only for the opposition parties. *The Difference Between the Biggest and the Second Biggest Opposition Parties (OPOP)* measures the competition between the biggest and the second biggest opposition parties. I normalize the difference between the seat shares of the two biggest opposition parties by the number of total available legislative seats within a country. The *Size of the Biggest Opposition Party (BOPP)* measures the size of the biggest party in the opposition also normalized with the number of the legislative seats. Finally, I calculate the *Competition*, which is the seat share difference between the biggest and the second

Table 2. Correlation table, traditional measures of party system size and opposition structure and principal components.

	Fraction.	HH	Entropy	ENP	P.in.Gov	Shap.ENP	Big.Party	ENOP	OPOP	Bopp	Competition
HH	−1.00										
	(0.00)										
Entropy	0.96	−0.96									
	(0.01)	(0.01)									
ENP	0.91	−0.91	0.95								
	(0.02)	(0.02)	(0.01)								
P.in.Gov	0.70	−0.70	0.73	0.75							
	(0.03)	(0.03)	(0.03)	(0.03)							
Shap.ENP	0.87	−0.87	0.87	0.91	0.76						
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)						
Big.Party	−0.94	0.94	−0.89	−0.90	−0.71	−0.91					
	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)					
ENOP	0.66	−0.66	0.73	0.68	0.21	0.52	−0.50				
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)				
OPOP	−0.62	0.62	−0.63	−0.58	−0.33	−0.46	0.46	−0.74			
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)			
Bopp	−0.58	0.58	−0.61	−0.58	−0.38	−0.47	0.45	−0.68	0.94		
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.01)		
Competition	−0.47	0.47	−0.36	−0.43	−0.40	−0.58	0.71	(0.03)	−0.13	−0.09	
	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	
PC1	0.98	−0.98	0.94	0.93	0.70	0.87	−0.94	0.64	−0.66	−0.62	−0.43
	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	0.02	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)
PC2	0.06	−0.06	−0.04	0.04	0.12	0.24	−0.34	−0.33	0.46	0.39	−0.90
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.02)

Note: Computed correlation using Pearson method with pairwise deletion. Standard errors are in parentheses. *Abbreviations:Fraction*, Fractionalization Index; *HH*, Herfindahl–Hirschman Index; *Entropy*, entropies; *ENP*, effective number of parties; *P.in Gov*, number of parties in the government; *Shap.ENP*, effective number of parties (Shapley); *Big Party*, size of the biggest party over the size of the legislature; *ENOP*, effective number of opposition parties; *OPOP*, the difference between the first and the second biggest opposition parties over the size of the legislature; *BOPP*, size of the biggest opposition party over the size of the legislature; *Competition*, size difference between the biggest and the second biggest parties over the size of the legislature. *PC1*: Dimension 1, Sizes of the two biggest parties. *PC2*: Dimension 2, Competition between the biggest and the second biggest parties.
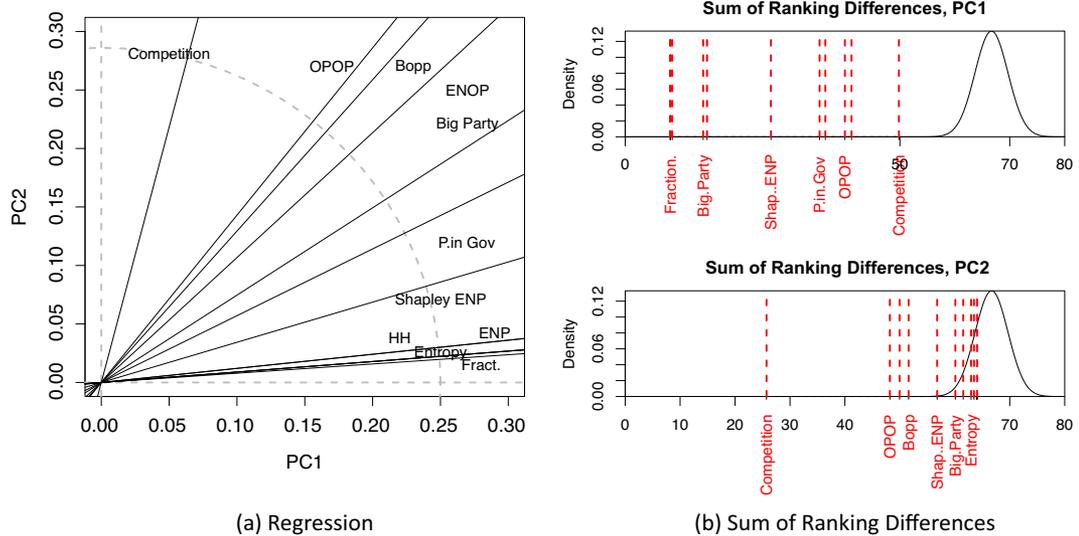
(a) Regression

(b) Sum of Ranking Differences

Figure 7. Measures on the PC dimensions. *Notes:* The plot shows how party system size measures relate to each other and the two first dimensions of the PCA. *PC1*: Dimension 1, Sizes of the two biggest parties. *PC2*: Dimension 2, Competition between the biggest and the second biggest parties. *Abbreviations: HH*, Herfindahl–Hirschman Index; *Fract*, Fractionalization Index; *Entropy*, entropies; *ENP*, effective number of parties; *P.in Gov*, number of parties in the government; *Shap.ENP*, effective number of parties (Shapley); *Big Party*, size of the biggest party over the size of the legislature; *ENOP,* effective number of opposition parties; *OPOP*, the difference between the first and the second biggest opposition parties over the size of the legislature; *BOPP*, size of the biggest opposition party over the size of the legislature; *Competition*, size difference between the biggest and the second biggest parties over the size of the legislature.

biggest parties normalized with the number of legislative seats.[15] The last two rows of Table 2 show that the opposition measures and the competition measure relate both to *PC2* and to *PC1* while the party system size measures are closely related to only *PC1*. The only exception is the size of the Biggest Party (*Big Party*), which seems to be somewhere in between the opposition measures and the party system size measures. This makes sense because it is possible that the biggest party in the party system is not a government party but an opposition party.

I also plot the indices on PC *Dimensions 1 and 2* to show how the indices relate to these dimensions (Figure 7). First, I regress each index on the PC1 and PC2 dimensions and plot the coefficients.[16] Figure 7a also shows that most of the indices that the current scholarship uses are closer to *Dimension 1* while opposition concentration measures are closer to *Dimension 2*. Figure 7a offers a way in which we can decide between measures with multicriteria decision-making process, as the curve can serve as the basis of a pareto-optimization. The more we would like to take into account competition between the parties with our measure and less the number of parties, the closer index we should choose to *Dimension 2*. Alternatively, if our theory calls for the size of the party system, we can choose one of the variables close to *Dimension 1*. Similarly, we can evaluate how the measures relate to the PC dimensions by using the *Sum of Ranking Differences*

---

15 Mathematically, ENOP is the reciprocal of the sum of squared seat shares across all parties in opposition in the legislature in a given year, $ENOP = \frac{1}{\sum s_{Opp_i}^2}$, where s_{Opp_i} is the seat share of each opposition party Maeda (2010) and Maeda (2015).

$OPOP = \frac{s_{Opp_1} - s_{Opp_2}}{n}$, where s_{Opp_1} is the number of seats that the biggest opposition party has in the legislature, while s_{Opp_2} is the number of seats that the second biggest opposition party has in the legislature, and n is the number of total legislative seats. The $BOPP = \frac{s_{Opp1}}{n}$, where s_{Opp_1} is the number of seats that the biggest opposition party has in the legislature, and n is the number of total legislative seats. *Competition* is measured as $Comp = \frac{s_{BP_1} - s_{BP_2}}{n}$, where s_{BP_1} are the number of seats that the biggest party won and s_{BP_2} are the number of seats that the second biggest party won.

16 Since the dimensions are not correlated, I can use the regression coefficients directly to construct the lines that show how the measures relate to the dimensions. Note that I am using the absolute value of the coefficients in all cases, since I do not care about the directionality of the relationship between the indices and the dimensions.

method. First, I calculate how the country-years are ranked on the PC1 and PC2 dimensions. Then, I calculate how the various country-years are ranked by the indices. Finally, I deduct the PC1 and PC2 rankings from the rankings of the measures. The sum of the differences is the *Sum of Ranking Differences*. The low values indicate that the index is close to the dimension, and higher values indicate that the index is far. Since the PC method is nondirectional, I repeat the process with inverse ranking as well to find the measure that is the closest to the PCA solution. I find that the *HH* followed by the *Fractionalization* and *ENP* (identical in their ranking) are the closest to the *Dimension 1*'s ranking of country-years. This result shows that neither transformation of the *HH* changes the measure in a meaningful way. At the same time, the measure *Competition* is the closest to *Dimension 2* followed by *OPOP*, *ENOP*, and *BOPP* (Héberger and Kollár-Hunek 2011; Bajusz, Rácz, and Héberger 2015). I plot the normalized results against the theoretical distribution of random rankings in Figure 7b.

6 Conclusion

In this paper, I conducted a PCA of data comprising party seat shares of 17 European countries from 1970 to 2013. I argued that although the data were compositional, the best choice for understanding the underlying structure of the party system data was to examine the untransformed dataset. The PCA revealed that the absolute sizes of the biggest parties and the relative sizes of the two biggest parties (i.e., the level of competition between the two biggest parties) were the two most important features that differentiated party systems from each other in Europe during the study period.

I suggested several applications of this method to improve the general understanding of party systems and how they can be measured. The first two dimensions that the method recovered could be used to track party system changes within countries. I also conducted a PCA on party vote shares and demonstrated how electoral party system changes could be compared to legislative party system changes with the proposed method.

Second, PCA can help users of party system typologies and indices to understand the measures that they use. A comparison of the PCA results with the results produced using Rokkan's (1970) typology showed that the PCA grouped countries together that Rokkan had classified as belonging to the same groups. With the help of the PCA, I showed how the typologies could be used in quantitative studies. I classified countries that Rokkan did not classify and showed that the PCA method can automatically sort each country by year in the dataset into a "type."

Finally, I compared the results of the PCA to some traditional party system size indices that many researchers currently use. This comparison revealed that most of these indices were highly correlated with each other and with Dimension 1 (i.e., size of the biggest two parties). Empirical analysts should be aware that the indices they use frequently are highly correlated with other indices that could be conceptually different yet still produce the outcomes that they are interested in. In addition, if possible, they should avoid using more than one of these indices within the same model. Furthermore, I suggested that we might want use other indices if we wanted to measure the competitiveness of a party system. I calculated some indices that measure opposition concentration and competition, and showed that these relate to *Dimension 2*. I proposed that these may be better suited for measuring the competitiveness of party systems than the traditional party system size indices.

Acknowledgments

Data Availability Statement

For Dataverse replication materials, see Magyar (2021).

Supplementary Material

For supplementary material accompanying this paper, please visit https://doi.org/10.1017/pan.2021.21.

References

Aitchison, J. 1983. "Principal Component Analysis of Compositional Data." *Biometrika* 70(1):57–65. https://doi.org/10.1093/biomet/70.1.57.

Aitchison, J. 1986. *The Statistical Analysis of Compositional Data*. London: Chapman and Hall.

Aitchison, J. n.d. "A Concise Guide to Compositional Data Analysis." Manuscript, University of Glasgow.

Akkerman, A., C. Mudde, and A. Zaslove. 2014. "How Populist Are the People? Measuring Populist Attitudes in Voters." *Comparative Political Studies* 47(9):1324–1353.

Bajusz, D., A. Rácz, and K. Héberger. 2015. "Why is Tanimoto Index an Appropriate Choice for Fingerprint-Based Similarity Calculations?" *Journal of Cheminformatics* 7(1):20.

Banzhaf, J. F. 1965. "Weighted Voting Doesn't Work: A Mathematical Analysis." *Rutgers Law Review* 19(2):317–343.

Bawn, K., and F. Rosenbluth. 2006. "Short Versus Long Coalitions: Electoral Accountability and the Size of the Public Sector." *American Journal of Political Science* 50(2):251–265. https://doi.org/10.1111/j.1540-5907.2006.00182.x.

Baxter, M. J., and I. C. Freestone. 2006. "Log-Ratio Compositional Data Analysis in Archaeometry." *Archaeometry* 48(3):511–531.

Ben-Ur, J., and B. I. Newman. 2002. "Motives, Perceptions and Voting Intention of Voters in the 2000 US Presidential Election." *Psychology & Marketing* 19(12):1047–1065.

Bishop, C. M. 1995. *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press.

Blondel, J. 1968. "Party Systems and Patterns of Government in Western Democracies." *Canadian Journal of Political Science* 1(2):180. https://doi.org/10.1017/s0008423900036507.

Caulier, J.-F., and P. Dumont. 2005. "The Effective Number of Relevant Parties: How Voting Power Improves Laakso–Taagepera's Index."

Dahl, R. A. 1966. *Political Oppositions in Western Democracies*. New Haven, CT: Yale University Press.

de Leeuw, J. 2005. *Nonlinear Principal Component Analysis*. Department of Statistics, UCLA. Retrieved from https://escholarship.org/uc/item/8bj075gv.

Dennis, J. 1988. "Political Independence in America, Part II: Towards a Theory." *British Journal of Political Science* 18(2):197–219.

Downes, A. S. 1988. "On the Statistical Measurement of Smallness: A Principal Component Measure of Country Size." *Social and Economic Studies* 37(3):75–96.

Dunleavy, P., and F. Boucek. 2003. "Constructing the Number of Parties." *Party Politics* 9(3):291–315. https://doi.org/10.1177/1354068803009003002.

Duverger, M. 1954. *Political Parties, Their Organization and Activity in the Modern State*. Translated by Barbara and Robert North. New York: Wiley.

Falcó-Gimeno, A., and J.-J. Vallbé. 2013. "Coalition Agreements and Party Preferences: A Principal Components Analysis Approach." 3rd Annual General Conference Of The European Political Science Association, Barcelona, Spain.

Grofman, B. 2006. "The Impact of Electoral Laws on Political Parties." In *Oxford Handbook of Political Economy*, edited by B. R.Weingast and D. Wittman, 102–118. Oxford: Oxford University Press.

Grofman, B., and R. Kline. 2011. "How Many Political Parties Are There, Really? A New Measure of the Ideologically Cognizable Number of Parties/Party Groupings." *Party Politics* 18(4):523–544. https://doi.org/10.1177/1354068810386838.

Guzmán, F., and V. Sierra. 2009. "A Political Candidate's Brand Image Scale: Are Political Candidates Brands?" *Journal of Brand Management* 17(3):207–217.

Héberger, K., and K. Kollár-Hunek. 2011. "Sum of Ranking Differences for Method Discrimination and Its Validation: Comparison of Ranks with Random Numbers." *Journal of Chemometrics* 25(4):151–158.

Jakulin, A., W. Buntine, T. M. La Pira, and H. Brasher. 2009. "Analyzing the U. S. Senate in 2003: Similarities, Clusters, and Blocs." *Political Analysis* 17(3):291–310.

Jolliffe, I. T. 2002. *Principal Component Analysis*. New York: Springer.

Katz, J. N., and G. King. 1999. "A Statistical Model for Multiparty Electoral Data." *American Political Science Review* 93(1):15–32. https://doi.org/10.2307/2585758.

Kesselman, M. 1966. "French Local Politics: A Statistical Examination of Grass Roots Consensus." *American Political Science Review* 60(4):963–973. https://doi.org/10.2307/1953769.

Kestilä, E. 2006. "Is There Demand for Radical Right Populism in the Finnish Electorate?" *Scandinavian Political Studies* 29(3):169–191.

Kitschelt, H. 2008. "Chapter 22: Party Systems." In *The Oxford Handbook of Comparative Politics*, edited by C. Boix, and S. C. Stokes, 522–554. Oxford: Oxford University Press.

Kline, R. 2009. "How We Count Counts: The Empirical Effects of Using Coalitional Potential to Measure the Effective Number of Parties." *Electoral Studies* 28(2):261–269. https://doi.org/10.1016/j.electstud.2009.01.008.

Kollár-Hunek, K., J. Heszberger, Z. Kókai, M. Láng-Lázi, and E. Papp. 2008. "Testing Panel Consistency with GCAP Method in Food Profile Analysis." *Journal of Chemometrics: A Journal of the Chemometrics Society* 22(3–4):218–226.

Laakso, M., and R. Taagepera. 1979. "'Effective' Number of Parties: A Measure with Application to West Europe." *Comparative Political Studies* 12(1):3–27. https://doi.org/10.1177/001041407901200101.

Laver, M., and K. Benoit. 2015. "The Basic Arithmetic of Legislative Decisions." *American Journal of Political Science* 59(2):275–291. https://doi.org/10.1111/ajps.12111.

Maeda, K. 2010. "Divided We Fall: Opposition Fragmentation and the Electoral Fortunes of Governing Parties." *British Journal of Political Science* 40(2):419–434.

Maeda, K. 2015. "Determinants of Opposition Fragmentation Parliamentary Rules and Opposition Strategies." *Party Politics* 21(5):763–774.

Magyar, Z. B. 2021. "Replication Data for What Makes Party Systems Different? A Principal Component Analysis of 17 Advanced Democracies 1970–2013." https://doi.org/10.7910/DVN/Q0OIMY, Harvard Dataverse, V1.

Mair, P. 2002. "Chapter 4: Comparing Party Systems." In *Comparing Democracies 2: New Challenges in the Study of Elections and Voting*, edited by L. LeDuc, R. G. Nciemi, and P. Norris, 88–108. London: Sage.

Molinar, J. 1991. "Counting the Number of Parties: An Alternative Index." *The American Political Science Review* 85(4):1383. https://doi.org/10.2307/1963951.

Peres-Neto, P. R., D. A. Jackson, and K. M. Somers. 2005. "How Many Principal Components? Stopping Rules for Determining the Number of Non-Trivial Axes Revisited." *Computational Statistics & Data Analysis* 49(4):974–997.

Powell, G. B. 2000. *Elections as Instruments of Democracy: Majoritarian and Proportional Visions*. New Haven, CT: Yale University Press.

Rae, D. W. 1967. *The Political Consequences of Electoral Laws*. New Haven, CT: Yale University Press.

Rae, D. W., and M. Taylor. 1970. *The Analysis of Political Cleavages*. New Haven, CT: Yale University Press.

Rokkan, S. 1970. *Citizens, Elections, Parties: Approaches to the Comparative Study of the Processes of Development*. New York: Scandinavian University Press.

Rozenas, A. 2011. "A Statistical Model for Party-Systems Analysis." *Political Analysis* 20(2):235–247. https://doi.org/10.1093/pan/mpr041.

Sartori, G. 1976. *Parties and Party Systems: A Framework for Analysis*. Cambridge: Cambridge University Press.

Savoy, J. 2015. "Text Clustering: An Application with the State of the Union Addresses." *Journal of the Association for Information Science and Technology* 66(8):1645–1654.

Schölkopf, B., A. Smola, and K.-R. Müller. 1997. "Kernel Principal Component Analysis." In *International Conference on Artificial Neural Networks*, edited by W. Gerstner, A. Germond, M. Hasler, and J.-D. Nicoud, 583–588. New York: Springer.

Shapley, L. S., and M. Shubik. 1954. "A Method for Evaluating the Distribution of Power in a Committee System." *American Political Science Review* 48(3):787–792. https://doi.org/10.2307/1951053.

Taagepera, R. 1999. "The Number of Parties as a Function of Heterogeneity and Electoral System." *Comparative Political Studies* 32(5):531–548. https://doi.org/10.1177/0010414099032005001.

Zoco, E. 2006. "Legislators' Positions and Party System Competition in Central America: A Comparative Analysis." *Party Politics* 12(2):257–280.