

UCLA

UCLA Electronic Theses and Dissertations

Title

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

Permalink

<https://escholarship.org/uc/item/1c192923>

Author

Magyar, Zsuzsanna Blanka

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

What Makes Party Systems Different?

A Principal Component Analysis of 17 Advanced Democracies 1970-2013

A thesis submitted in partial satisfaction

of the requirements for the degree

Master of Science in Statistics

by

Zsuzsanna Blanka Magyar

2017

© Copyright by

Zsuzsanna Blanka Magyar

2017

ABSTRACT OF THE THESIS

What Makes Party Systems Different?

A Principal Component Analysis of 17 Advanced Democracies 1970-2013

by

Zsuzsanna Blanka Magyar

Master of Science in Statistics

University of California, Los Angeles, 2017

Professor Mark Stephen Handcock, Chair

Parties are the main vehicles of representation in modern, democratic societies. Party systems, that is the number and the size of all the parties within a country, can vary greatly across countries. There is an ongoing debate in the political science literature about the appropriate way to reduce the dimensionality of the cross-country party system data for comparative purposes. This thesis reviews that literature and offers a new solution: Principal Component Analysis to find the most important information in the data matrix. I use data from 17 advanced democracies from 1970-2013. I conduct analyses using various related methods (Principal Component Analysis, Principal Component Analysis on the Residuals, kernel Principal Component Analysis, Non-Linear Principal Component Analysis, Principal Component Analysis on log-ratio transformed variables and Principal component Analysis on non-centered variables). I find that the most important differences across countries are: “the size of the biggest two parties”, “competition between the two biggest parties”, “existence of a third party” and “balanced multipartism.” I argue that most of the current political science literature uses summary measures that are only correlated with the first of those four dimensions. I suggest a strategy for incorporating a measure of the second dimension that relies on indices of opposition structure.

The thesis of Zsuzsanna Blanka Magyar is approved.

Chad J. Hazlett

Nicolas Christou

Mark Stephen Handcock, Committee Chair

University of California, Los Angeles

2017

To Anya and Apa

TABLE OF CONTENTS

1	Introduction	1
2	Literature Review	4
2.1	Typologies	4
2.2	Summary of Typologies	7
2.3	Indices	9
2.3.1	Maximum Entropy	9
2.3.2	Concentration	10
2.3.3	Party Power and the Number of Parties	11
2.3.4	Full distribution	12
2.4	Measuring the Party System: Summary	13
3	Data	14
4	Principal Component Analyses	19
4.1	Simple PCA	19
4.1.1	PCA on the Residuals	26
4.2	Introducing Non-linearity to the PCA Analysis	31
4.2.1	Kernel Principal Component Analysis (kPCA)	31
4.2.2	Non-Linear Principal Component Analysis (NLPCA)	35
4.3	Issues with the Compositional Dataset	43
4.3.1	PCA on Log-ratio Transformed Variables	44
4.3.2	PCA on Non-centered Variables	49
4.4	Summary	51

5	How do the PCA Dimensions Relate to Traditional Typologies and Measures	55
5.1	Comparing the Results of the PCA Analysis to Traditional Typologies	55
5.2	Comparing the results of the PCA analysis to Party System Size Measures .	58
5.3	Measures of Opposition Concentration	58
6	Conclusion	64

LIST OF FIGURES

3.1 Party System Distribution, 17 Countries	15
3.2 Party System Distribution by year, Spain, Sweden, United Kingdom, Greece	16
3.3 Party System Distribution by year: France, the Netherlands, Norway, Portugal	17
4.1 Scree Plot, PCA	21
4.2 Loadings, PCA	22
4.3 All Observations on the Two-dimensional PCA Plane	24
4.4 Biplot, PCA	25
4.5 Screeplot, PCA on the Residuals	28
4.6 Loadings, PCA on the Residuals	29
4.7 Biplot, PCA on the Residuals	30
4.8 Scree Plot, kPCA	33
4.9 Biplot, kPCA	34
4.10 NLPCA Objectplot, All Parties	37
4.11 NLPCA, Scree Plot, All Parties	38
4.12 NLPCA, Loadings, All Parties	39
4.13 The Sensitivity of the NLPCA Results to the Change in the Number of Parties	40
4.14 The Sensitivity of the Scaled PCA Results to the Change in the Number of Parties	41
4.15 The Sensitivity of the Unscaled PCA Results to the Change in the Number of Parties	42
4.16 Loadings, NLPCA, Ten Parties	43
4.17 Scree Plot, NLPCA, 10 Parties	44
4.18 Objectplot, NLPCA, 10 Parties	45

4.19	Sensitivity of the Log-ratio Transformed Variables to the Size of an Added Value to Zeros	46
4.20	Loadings, PCA on Log- ratio Transformed Variables	47
4.21	Biplot, PCA on Log- ratio Transformed Variables	48
4.22	Screepplot, PCA on Non-centered Variables	50
4.23	Loadings, PCA on Non-centered Variables	51
4.24	Biplot, PCA on Non-centered Variables	52
5.1	Correlogram, Traditional Measures of Party System Size and PCA Principal Components	59
5.2	Correlogram, Opposition Structure and PCA Principal Components	61
5.3	Correlogram, Traditional Measures of Party System Size and Opposition Structure and PCA Principal Components	62
6.1	Austria on the PCA Dimensions	66
6.2	Belgium on the PCA Dimensions	67
6.3	Denmark on the PCA Dimensions	68
6.4	Finland on the PCA Dimensions	69
6.5	France on the PCA Dimensions	70
6.6	Germany on the PCA Dimensions	71
6.7	Greece on the PCA Dimensions	72
6.8	Iceland on the PCA Dimensions	73
6.9	Ireland on the PCA Dimensions	74
6.10	Italy on the PCA Dimensions	75
6.11	Luxembourg on the PCA Dimensions	76
6.12	the Netherlands on the PCA Dimensions	77

6.13 Norway on the PCA Dimensions	78
6.14 Portugal on the PCA Dimensions	79
6.15 Spain on the PCA Dimensions	80
6.16 Sweden on the PCA Dimensions	81
6.17 The United Kingdom on the PCA Dimensions	82
6.18 The Party Systems of 17 European Countries in 1970 on the PCA Dimensions	83
6.19 The Party Systems of 17 European Countries in 1975 on the PCA Dimension	84
6.20 The Party Systems of 17 European Countries in 1980 on the PCA Dimensions	85
6.21 The Party Systems of 17 European Countries in 1985 on the PCA Dimensions	86
6.22 The Party Systems of 17 European Countries in 1990 on the PCA Dimensions	87
6.23 The Party Systems of 17 European Countries in 1995 on the PCA Dimensions	88
6.24 The Party Systems of 17 European Countries in 2000 on the PCA Dimensions	89
6.25 The Party Systems of 17 European Countries in 2005 on the PCA Dimensions	90
6.26 The Party Systems of 17 European Countries in 2010 on the PCA Dimensions	91

LIST OF TABLES

2.1	Party System Classifications	8
4.1	Eigenvalues and Explained Variance, PCA	20
4.2	Eigenvalues and Explained Variance, PCA on the Residuals	27
4.3	Eigenvalues and Explained Variance, kPCA	32
4.4	kPCA and PCA Covariances	35

ACKNOWLEDGMENTS

First and foremost, I would like to express my gratitude to my committee. I would like to thank Professor Mark Handcock, whose social statistics class inspired me to pursue the MS degree in Statistics. He was a very dedicated and inspiring chair throughout the writing of this thesis. I would like to thank Professor Chad Hazlett who gave me the idea for this thesis and gave me careful advice on the project every step of the way. I would like to thank Professor Nicolas Christou, whose class was the first one I took at the UCLA Department of Statistics it was so interesting that I wanted to take more. I would like to thank the entire Statistics Department at UCLA for being very welcoming towards students from other departments. I would like to thank Glenda Jones, who has helped me tremendously in navigating the bureaucracy of UCLA. From the Political Science Department, I would like to thank Professor Jeffrey Lewis who I could always find with sudden questions, and Professor Michael F. Thies, whose knowledge on party system helped me through this project. I would like to thank some of my peers who have been taking this journey with me, Marika Csapo and Eoghan Stafford, with whom I spent late nights solving statistics problem sets.

Finally, I would like to thank family. I would like to thank, my mother Zsuzsa Nemeti and my father Istvan Magyar, who have been supporting me through this process. I would like to thank my father who showed me how exciting math puzzles could be. I found the same excitement in solving statistics problems. I would like to thank my late grandfather Gabor Nemeti, who I always wanted to emulate as a careful scholar, and my late grandmother Margit Oszlanczi Magyar Istvanne “Mami,” who supported me with love even when I studied at the other end of the world.

CHAPTER 1

Introduction

In the modern world, political parties are the main vehicles of political representation. However, the number of parties that win seats in the legislature varies widely across countries. This number depends on the electoral system (Duverger 1954; Lijphart 1994) and on the social cleavages in the society (Lipset and Rokkan 1967).

The number and size of the parties within countries were some of the first data that were readily available for political science. In the 1960s and 1970s, political scientists argued that the variation in the size of the party system determines how parties can interact with each other, on the whole party systems ought to be important for politics (Sartori 1976). It was not easy to analyze the data however; since party systems vary widely in size and shape (Kitschelt 2008). To determine the causes and the consequences of the differences between these party systems, political scientists hoped to find patterns across countries. They constructed typologies by grouping together parties with similar party systems. These groupings sort the countries, partly based on the number of parties, partly based on the sizes of the parties and how the parties behave in the elections (Duverger 1954; Dahl 1966; Rokkan 1970; Mair 2002). However, often the demarcation line between the groups is decided based on how the parties behave, which could already be an outcome of the party system itself.

To avoid the subjectivity, another group of political scientists tried to summarize the party system size in a single measure, based only on the legislative seat shares of the parties (Laakso and Taagepera 1979). Though it would be straightforward to characterize the political systems of the countries with the sheer number of parties in the legislature, political scientists argued, that we needed to know not only the number but the size of the parties as well, since small parties have less influence on the political processes (Blondel 1968).

Accordingly the notion of a *relevant* party was born. Political scientists, however, defined the *relevant* parties differently, and offered many different solutions as to how to weight the number and size of the parties. The most important question at this time was whether to weight big parties more (as they are relevant) or weight small parties more (so that the index stays sensitive to small changes). This debate points to a classical dimensionality reduction problem: how to reduce a high-dimensional dataset in such a way that all the important information is retained? However, even with the various measures political scientists did not find any systematic impact of the party system size on politics. Thus the project to describe, group, and measure party systems slowed down. Today, very little research concentrates on how the party system influences politics (Kitschelt 2008).

In this paper I revive the previous research agenda, that aimed to understand what makes party systems different from each other, but this time I use another set of dimensionality reduction technique to group together the countries that are most alike. Using legislative seat share data from 17 European countries, I conduct a Principal Component Analysis (PCA), a Kernel Principal Component Analysis (kPCA), and a Nonlinear Correspondence Analysis (NLCA) to find the latent dimensions that make party systems different from each other. The seat share data is compositional. This means that the sum of the seat shares equals to one.¹ This may cause issues with the PCA analysis as the dataset does not have subcompositional coherence (Aitchison 1983). For this reason I also conduct a PCA on the log-ratio transformed and the non-centered variables. I then compare the categories that I find in my analysis to the categories that the qualitative researchers established. Finally, I compare the categories to the party system indices that political scientists have created to see what features of the party system these indices measure. Through this analysis I implicitly can compare the typologies with the party system size measures. I find that while the typologies divide the countries based on the sizes of their party systems and the competition within the party system the traditional party system measures only measure one aspect, the size of the party system. I suggest that the opposition concentration measures may offer a

¹Compositional data “consists of vectors of positive values summing to a unit” or any other constant for all vectors (Aitchison and Greenacre 2002).

solution and can measure the competitiveness of the party systems.

CHAPTER 2

Literature Review

The number and relative sizes of the parties is one of the main characteristics of the political system in any developed country. Most country studies in political science include a description of the political development of the party system. However, each country's political system developed independently and the interactions of the parties may seem to be unique from this perspective. Thus, in spite of the ample information available, political scientists have struggled to understand the extent to which these interactions are determined by systematic, structural factors and the extent to which they are determined by the stochastic political behavior of the elites.

Two big schools of thought emerged in the literature. First, some scholars classified the party systems into several broad groups with qualitative methods. The aim of this classification effort was to find groups of countries in which the parties and politicians behaved similarly. At the same time, other scholars created continuous measures to quantify the size and shape of the party systems that could be used in comparative quantitative analyses. Before running the PCA analyses, I am going to first review some of the most important party system typologies and party system measures from the previous literature.

2.1 Typologies

In any parliamentary system, a majority is needed to pass legislation. Normally, this legislative majority chooses the prime minister and the government. The rest of the parties are considered to be in the opposition. The most canonical difference in party systems across countries is between two-party and multi-party systems. In a two-party system, the winning

party always holds the legislative majority by itself. In contrast, if there are many parties, no party may hold a majority by itself. If none of the parties wins a majority, some parties have to form a governing coalition.

Duverger (1954) argues that plurality electoral systems (in which only one candidate can win in a given district) lead to a two-party competition, at least on the district level. This is because the voters do not waste their vote on third party candidates, thus small parties fall out from the competition. By contrast, proportional representation electoral system (PR) leads to a multi-party party system. Under PR, several candidates can win seats within a given electoral district. The parties get seats based on their vote shares in the election (thus a party that got 15% of the votes receives roughly 15% of the seats in the legislature). Under this system, small parties can gain legislative representation. Duverger considers the two-party system the ideal type, while he thinks that multi-party systems are unstable and inchoate, as the coalition governments are less stable than single party governments. In practice, however, there are very few countries with ideal two-party systems (countries that have close to two-party systems, at least in the 1970s include Britain, the United States, Canada, New Zealand, Austria and Australia (Sartori 1976)).

The rest of the countries are multi-party countries. Within the countries with non-majoritarian electoral systems, there is a wide variety of different sized and structured party systems. One reason for this again is the electoral system: In some of the PR countries, electoral districts are relatively small, - there are electoral districts in which only a few seats get allocated. Even though *within* the electoral districts seats are allocated proportionally, the smallest parties cannot gain seats (for instance if there are only 5 seats available a party with 15% of the votes may not gain seats). However, the variation is not limited to electoral causes. Even in countries with the same electoral system, different party systems have developed, and keep evolving. To impose order in the chaos (to group *similar* countries together), political scientists classify the multi-party countries into more refined categories (Blondel 1968; Rokkan 1970; Sartori 1976).

Blondel (1968) is the first to recognize that not only the number of parties, but also their relative sizes, are important to compare party systems, as small parties are less

important than big ones.¹ Most of the typologies following Blondel (1968) sort the countries based on the number of the parties and based on how the parties compete. Depending on their approach, some authors argue that the competition style is a direct outcome of the party system size and structure, while others argue that the competition between the parties is an independent feature, a separate dimension. Rokkan (1970) classifies the countries based on whether the parties in the party system are roughly the same sizes (compared to each other) or whether there is one or more dominant parties facing small parties.² In a related paper, Laver and Benoit (2015) create a party system classification based on the government, and coalition potential of the different parties.³

Other authors consider competition a separate feature. Dahl (1966) sorts the countries into different categories based on whether the parties only compete or at the same time cooperate with each other (which happens in party systems in which parties regularly have to build coalitions). He argues that the competition style is directly influenced by the party system.⁴ On the other hand, Sartori (1976) argues that party fragmentation and the ideological distances between the parties are two separate characteristics and these two dimensions determine the type of political competition in a country.⁵ Finally, Mair classifies party systems based on whether a country has open or closed party competition, whether new parties can enter the race. Thus the party system defines the type of the country (Mair

¹He sorts the party systems into two-party systems, two and a “half” party systems, multi-party systems with a dominant party, and multi-party systems without a dominant party.

²Rokkan’s categories are named after the sizes of the parties in these groups. For example the British-German “1 vs. 1+1” system describes a two and a half party system – a dominant party facing one dominant and one small party (Rokkan 1970).

³Laver and Benoit (2015) establish categories based on how the ranked parties (biggest, second biggest etc.) could form winning coalitions (reach the 50% seat share threshold). Thus the authors classify countries based on their party seat share constellations. The authors do not explicitly show the countries that belong to each category, as they argue that the multi-party countries shift in and out of these categories quite frequently, based on small changes in the electoral results.

⁴Dahl (1966) claims that the opposition is competitive in two-party systems –in which only two parties compete– while it is cooperative-competitive in multi-party systems –in which small opposition parties have a chance to join the government coalition without changing the entire government.

⁵Sartori draws a distinction between countries in which two ideologically close party groups compete (limited or moderate pluralisms), and between countries in which the opposition is fragmented, and ideologically diverse (extreme pluralisms). In his classification, the cut off between moderate and polarized pluralism is around five or six parties (Sartori 1976, 328).

2002).⁶

2.2 Summary of Typologies

In a later part of this paper, I examine the most important features that separate party systems, and I compare these to the typologies I discussed above. *Table 2.1* summarizes the typologies created by previous literature. The table lists the countries that the authors bring up as examples for the categories. Most of the typologies were created in the 1960s and 1970s and as a result the universe of the cases that the authors discuss is primarily limited to European democracies. Often the authors are cautious about discussing the political institutions in non-democracies or newly democratized countries. Greece, Spain and Portugal are also missing for the same reason (Greece becomes a democracy in 1974, Spain in 1978 and Portugal in 1976).

As the authors are writing in the same decade (apart from Mair (2002) and Laver and Benoit (2015)) the typologies are comparable to each other. Most of the authors sort two-party systems in their own separate category.⁷

There is less consensus about countries with more parties. As the number of parties within a country increases, the consensus on the ideal category for the country decreases. Countries that have party systems close to a two-party system (Germany, Ireland), get their own separate category in most of the classifications. However, it is unclear whether the

⁶Closed party systems are those where the alternation in government is fully predictable and new parties have no chance of gaining power. In contrast, it is unclear how the next government is going to look in an open system. Mair argues that open competition emerges in transitional (inchoate) party systems, or is a sign of party system failure which is reminiscent of how Duverger characterized multi-party systems (Mair 2002).

⁷There is considerable agreement that New Zealand, the United States and Australia are within this category, and Austria (at this time) as well. Sartori (1976) argues that the consensus is that most of the anglo-saxon countries are close to the two-party system ideal (Britain, the United States, Canada, New Zealand, Australia). However, in Canada there is a clear third party, and in Australia a single party competes with a two-party coalition. Often Austria also listed as a two-party system, although it does not adhere to the “two-party competition” ideal. In Austria in the 1960s and 1970s the two biggest parties, SPÖ and ÖVP formed a coalition to keep the radical right FPÖ out of the government

Table 2.1: Party System Classifications

Author	Criteria	Typology	Countries
Duverger (1954)	Numbers of Parties	1. Two-party Systems 2. Multi-party Systems	Sartori: 1. England, United States, New Zealand, Australia, Canada, Austria 2. All else
Dahl (1966)	Competitiveness of the Opposition	1. Strictly competitive 2. Co-operative-competitive 2a. two-party 2b. multi-party 3. Coalescent-competitive 3a. Two-party 3b. Multi-party 4. Strictly coalescent	1. Britain 2a. United States 2b. France, Italy 3a. Austria, Wartime Britain 3b. (no example) 4. Colombia
Blondel (1968)	Numbers of parties Relative size of parties	1. Two-party systems 2. Two-and-a-half-party systems 3. Multi-party systems with one dominant party 4. Multi-party systems without dominant party	1. United States, New Zealand, Australia, England, Austria 2. Germany, Canada, Belgium, Ireland 3. Denmark, Norway, Sweden, Iceland, Italy 4. Netherlands Switzerland, France, Finland
Rokkan (1970)	Numbers of parties Proximity to the majority Evenness of the competition	1. The British-German "1 vs. 1+1" system 2. The Scandinavian "1 vs. 3-4" system 3. Even multi-party systems "1 vs. 1 vs. 1+ 2-3" 3a. scandinavian "split working class" systems 3b. segmented pluralism	1. Austria, Ireland, some periods Belgium 2. Sweden, Denmark, Norway 3a. Finland, Iceland 3b. Netherlands, Belgium, Luxembourg Switzerland
Sartori (1976)	Party fragmentation (number of parties) Ideological distance	1. Predominant party regimes 2. Two-party systems 3. Moderate pluralism 4. Polarized pluralism	1. Norway (or 3), Sweden (or 3), Japan, Uruguay, India Turkey 2. Canada, Australia, Austria, England, New Zealand, United States 3. Switzerland, Netherlands, Israel, Denmark, Iceland, Luxembourg, Belgium, Ireland, France (after 1958), Germany 4. Finland, Chile, France (before 1958), Italy
Mair (2002)	Type of Competition (alternation of the government, new parties: in the system, in the government)	1. Open Party System 2. Closed Party System	1. Denmark, the Netherlands, post authoritarian systems 2. United Kingdom, New Zealand (till mid 1990s), Japan (1955-93, Switzerland, Ireland (1948-89)
Laver and Benoit (2015)	Type of Competition (potential winning coalitions)	1. Single Winning Party 2. No Single Winning party 2a. Strongly dominant party ($S_2 + S_3 < W$) 2b. Top- three ($S_2 + S_3 \leq W$) 2c. Top-two 2d. Open	countries change categories

Note: The table is modified from Table 1 in Mair (2002). The countries in the different categories are the authors' own except for Duverger, where I take my information from (Sartori 1976).

Scandinavian countries are their own category or not.⁸ The most problematic countries to categorize into the typologies are Finland and France.⁹

Overall, it seems, that finding the proper categorization of multi-party countries is more difficult than dividing two-party and multi-party countries. However the typologies, in fact, are not too different from each other. Rokkan's (1970) idea to categorize parties based on how parties face each other within party system is made more precise 40 years later by Laver and Benoit (2015). Sartori's (1976) distinction between moderate and polarized pluralisms creates a very similar categorization to Mair (2002). All typologies suggest, that apart from the size and relative power of the parties, we should consider the competition within the party system to separate countries into groups.

2.3 Indices

In contrast to the authors who sort countries into party system categories, other authors summarize party systems with a single, continuous variable. Later, I will calculate some of these measures to compare them with the results of the PCA analysis.

2.3.1 Maximum Entropy

Kesselman (1966) develops an entropy-based hyperfractionalization index to characterize the shapes of party systems (Taagepera and Shugart 1989, 5). The entropy measure evaluates the probability of the i -th bin in a histogram. It counts the number of ways how we could rearrange the parties while still arriving at the same histogram (Bishop 1995).¹⁰ Kesselman

⁸While Sweden and Norway are usually in the same category, the appropriate category for Denmark and especially Iceland is less clear.

⁹France has several parties but these parties form coalitions, so depending on the author the country is categorized as either a quasi two-party system; or a party system with several, equally strong parties. Finland on the other hand gets categorized with the Netherlands (and France) by Blondel (1968) as the country has many small parties, Rokkan (1970) puts the country into the same category as Iceland (Scandinavian split working class country), while Sartori (1976) sorts the country to a category in which countries with a dominant parties are (along with Italy).

¹⁰For the i -th bin there are $N_i!$ such ways how we could arrange the objects and arrive at the same histogram. Where N is the number of objects and N_i is the number of objects in each bin. Altogether

defines his index as $I = \exp[-\sum_i^k p_i \log_e p_i]$ ¹¹ where k is the number of candidates or lists, p_i is the proportion of vote for i -th list and $\sum_i p_i = 1$ (Kesselman 1966).

Thus the hyperfractionalization indices uniquely characterize each party distribution. However, entropy-based indices are sensitive to the smallest changes in the distribution. This can make the measure unreliable, as similar party systems may end up with very different numbers (Laakso and Taagepera 1979).

2.3.2 Concentration

Next, in order to give more weight to bigger parties in the system and minimize the weight of smaller parties (to make the measures more reliable), political scientists adapt an economic measure. The basis of this family of measures is the Herfindahl-Hirschman concentration index, which is the sum of squares of the market share of each company in a given market $HH = \sum s^2$. (Where s is the market share of each company). The range of this index is 0 to 1 where a 1 means that the market is dominated by one company and 0 means that all companies are equal.¹² Rae and Taylor argue that this measure shows the probability that two randomly selected voters would vote for the same party (Molinar 1991).

Laakso and Taagepera (1979) argue that an intuitive transformation is $1/HH (\frac{1}{\sum s^2})$, which shows how many equal sized parties would be equivalent to the current party system. They call this measure the Effective Number of Parties (ENP). Currently the ENP is probably the most widely used measure of party-system concentration. However the measure

the multiplicity of the objects can be given by $W = \frac{N!}{\prod_i N_i!}$ the entropy is the negative logarithm of the multiplicity. $S = -\frac{1}{N} \{\ln N! - \sum_i \ln N_i!\}$ if we assume that $N \rightarrow \infty$ and use Stirling's approximation we find that $S = -\sum_i p_i \ln p_i$, the entropy. Consequently, a very high peaked histogram has a very low entropy (a histogram with one bin would have an entropy of 0) while a uniformly distributed one has a high entropy. (Bishop 1995).

¹¹According to Wildgen (1971) this measure measures “the voters’ tendencies to diverge or converge relative to parties or candidates.”

¹²Rae and Taylor (1970) calculate a fractionalization index by exchanging the companies’ market shares to seat shares in the formula, and changing the formula to $1/HH$ or $1/\sum s_i^2$ where the (s_1, \dots, s_n) are the legislative seat shares of the parties. This measure is in fact the Effective Number of Legislative Parties. Depending on the issue at hand, this measure can be calculated as $1/\sum v_i^2$ where (v_1, \dots, v_n) is the vote share of all the parties that ran in the elections.

has been criticized both because it insufficiently weights big parties, and because it does not show small changes in the party system. The generic formula for this family of indices is: $N_a = [\sum_1^x v_i^a]^{1/(1-a)}$ (Dunleavy and Boucek 2003). Where we raise the decimal vote shares to a power (a) add these numbers together and raise the resulting summed number to 1 divided by (1-a). We can see that the ENP is a special case of this formula where a=2 (Dunleavy and Boucek 2003).

2.3.3 Party Power and the Number of Parties

The first criticism of the ENP is that it overestimates the number of relevant parties. The critics argue that we should only consider parties to be relevant if they have a real probability of joining a governing coalition (Kline 2009) or at least of influencing the behavior of parties that do have coalition potential (Sartori 1976). Thus new measures are created to put more weight on the bigger parties if they have a higher coalitional potential or “power”.¹³

The Shapley-Shubik power index shows how many times a party would be pivotal in coalitions (Shapley and Shubik 1954).¹⁴ 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).¹⁵ Caulier and Dumont (2005), Grofman (2006), and Kline (2009) all suggest using the sum of squared power shares instead of the seat shares of the parties in the formula

¹³ In practice all power indices use the same data as the party number indices: the seat shares (or the vote shares) of the parties. The only difference is that, based on some combinatorial rules, the parties may receive bigger or smaller weight than their original seat (vote) shares.

¹⁴This measure starts from the premise that all possible coalitions are ordered as the parties join them in particular order. After listing all coalitions, in each coalition the pivotal player is identified. The pivotal player is the player that can make the coalition’s total vote share pass the threshold that is needed to win the particular vote. The index is calculated for each actor (party) and it shows how many times a player would be pivotal out of all possible permutations of party coalitions.

¹⁵Mathematically the Shapley index for a simple game of n players for party i is the following: $\Phi_i = \frac{1}{n!} \sum_{\{i\text{-swings-in-}S\}} (s-1)!(n-s)!$, ($s = |S|$). Where the sum is taken over all such coalitions S that i is in S , S is winning but $S - (i)$ is losing. With similar notation the Banzhaf index is the following: $\beta_i = \frac{1}{2^{n-1}} \sum_{\{i\text{-swings-in-}S\}} 1$ (Straffin Jr. 1988). In practice, bigger parties could get a higher Banzhaf power value than Shapley-Shubik value. This is because in an oversized coalition, a big party may be the only one whose leaving could swing the coalition from winning to losing so it would be the only party that is relevant for the calculation of the Banzhaf index. But the big party still may not be a majority party and may therefore need coalition partners, so it would not be the only party relevant for the calculation of the Shapley-Shubik value.

of the ENP in order to address the potential over-valuing of small parties. Mathematically this measure is: $LTB = 1/\sum_{i=1}^n B_i$ where B_i is the Banzhaf score of i -the party.¹⁶

Several other measures have been created to increase the weight of bigger parties.¹⁷ Dunleavy and Boucek suggest that because all of these measures are correlated with the size of the biggest party, we might as well use the latter to measure the size of the party system (2003). They suggest using $\frac{1}{V_1}$ where V_1 is the vote share of the biggest party. This is also suggested by Taagepera (1999).

In practice, studies find that there are sharper step-downs in the number of parties in the measures modified by the power of parties (Kline 2009). In fact, this modification amplifies that problem that ENP has, that some very different party configurations end up with the same index numbers.

2.3.4 Full distribution

The second major criticism about the ENP measure is coming from the other direction. Some authors argue that by weighing big parties more than small parties, a lot of different party configurations end up with the same ENP value, thus the index may mask important differences among the party systems.¹⁸ Thus in recent years, some political scientists have created measures, to describe the full distribution of parties in order to measure the nuanced

¹⁶This measure in practice ends up having sharper step-downs in the number of parties than the ENP when certain thresholds (of coalitional potential) are hit. Especially around these thresholds, the measure diverges from ENP. Kline argues that we should use this measure when we are interested in outcomes related to coalitional potential such as government duration (Kline 2009, 21)

¹⁷For instance Molinar (1991) argues that we should always count the winning party as one, and then he suggests that we should calculate the ENP with all the other parties and add the two values together. Mathematically this index is the following $M = 1 + \left(\frac{1}{\sum_1^x v_i^2} * \frac{\sum_1^x v_i^2 - V_1^2}{\sum_1^x v_i^2} \right)$, where v are all the parties. and V_1 the voteshare of the opposition parties. This index is criticized by Dunleavy and Boucek (2003) as it behaves erratically under certain circumstances.

¹⁸However, the original goal of Laakso and Taagepera (1979) was to create exactly such a measure. They believed that the party systems that they characterized with the same value were indeed similar. “The effective number of parties is the number of hypothetical equal-size parties that would have the same total effect on fractionalization of the system as have the actual parties of unequal size” (Laakso and Taagepera 1979). The goal of the authors with the index was to create a measure that will not change significantly when there is an additional small party in the party system

changes in the party system. These efforts create predicted vote shares of each party by using the log-ratio transformed party vote shares. (Katz and King 1999; Rozenas 2011).¹⁹

2.4 Measuring the Party System: Summary

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 the bigger more relevant parties. The former approach yields a measure that weights smaller parties more, while the latter yields a measure that weights larger parties more. All the measures were created to reduce the dimensionality of the party system data matrix by extracting the most important information in the dataset. The debate between scholars has been over which information to keep and which information to discard. In the following section, I will use another way to reduce the dimensionality of the data– Principal Component Analysis – and I will compare the results of these classical measures. Currently, in most empirical studies that evaluate whether certain factors influence government policies, the author picks one or more controls for the party system size (which is usually the *ENP*) without sufficient attention to what the indices actually measure. This may be one of the reasons why previous studies on the influence of the size of the party system did not lead to substantive results.

¹⁹Katz and King (1999) use district level electoral data from England to calculate the changes of party vote shares within the system. With the full distribution, they predict the expected vote share for each party in the districts and can calculate whether the politicians have incumbency advantage. Rozenas (2011) uses the relative sizes of the parties similarly. The parties are not defined by their names but by their electoral results (biggest, second biggest etc.). Both of these papers use the mathematical transformation that is suggested by Aitchison (1986) for compositional data. For party J let the voteshares in the districts i ($i = 1, \dots, n$) be $V_i = (V_{i1}, \dots, V_{iJ-1})$. In addition let Y_i be the vector of $J-1$ log-ratios. $Y_{ij} = \ln(\frac{V_{ij}}{V_{iJ}})$. Then we transform the voteshares as $V_{ij} = \frac{\exp(Y_{ij})}{1 + \sum_{j=1}^{J-1} \exp(Y_{ij})}$ where Y_i is the vector of $J-1$ to get the observed voteshares Katz and King (1999).

CHAPTER 3

Data

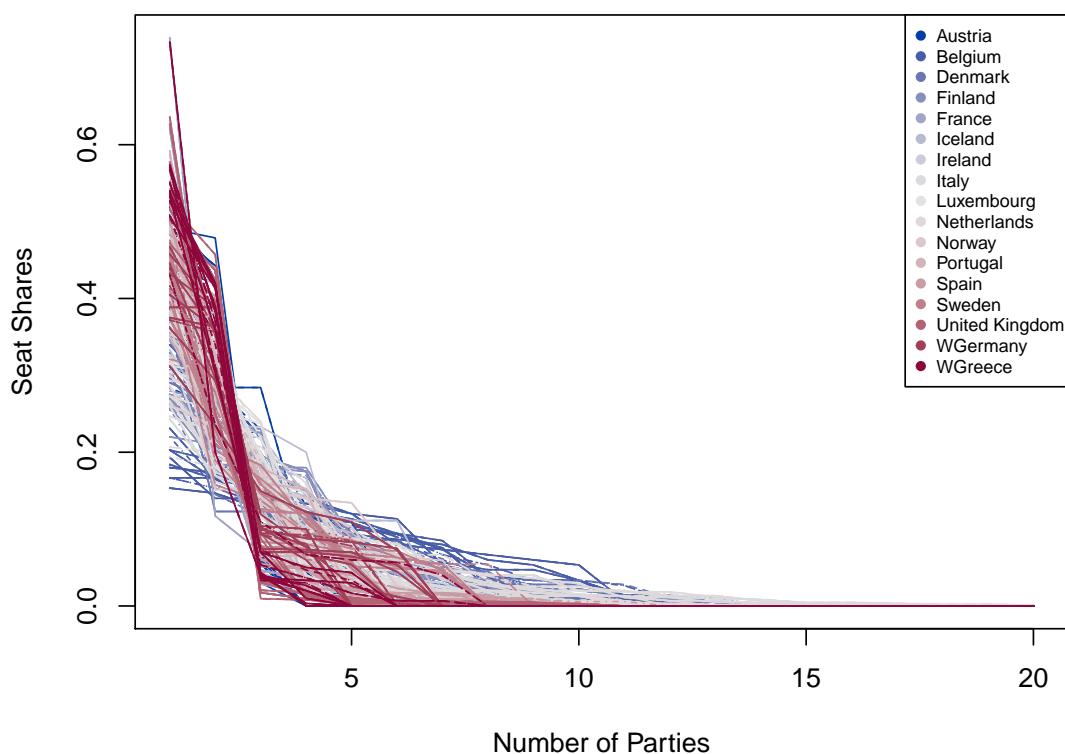
In this paper, I use dimensionality reduction techniques to explore the underlying structure of the party-system dataset. This will make it easier to understand what the party system size indices measure. The data that I am using consists of the party seat shares in the legislature of 17 European countries from 1970 to 2013.¹ In each row (country-year) of the matrix, I rank parties based on their sizes. The first variable is the seat shares of the biggest parties; the second variable is the seat shares of the second biggest parties etc. Thus, the dataset does not contain the identity of any individual party, but it allows me to compare the party systems across countries. If all the parties have been accounted for in a given country-year, the next entry in the row is a 0. The number of parties ranges from 3 in Austria from 1970-1986 and Greece in 1981-1984, to 20 in Italy in 2006 and 2007. The matrix that I create has 719 rows and 20 columns.

Figure 3.1 shows the party-system size distributions from around the world. The different colors indicate the different countries. The plot shows that party systems vary considerably across countries, and even within countries over time. Since the dataset measures the percent of legislative seat shares over the total number of seats, the dataset is a compositional dataset. The sum of seat shares within a country-year add up to 1 and each datapoint is between 0 and 1 (Aitchison n.d.).

To clarify how the party system structure changes within countries, in *Figure 3.2*

¹The countries in the dataset are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, and the UK. Data available: <http://www.parties-and-elections.eu/countries.html>. I use the legislative seat shares of the parties. Later to develop the relevant indices, I also use the identity of the government (also available above). Countries that democratized later than 1970 appear in the dataset after the first democratic elections.

Party Systems in 17 OECD Countries 1970–2013



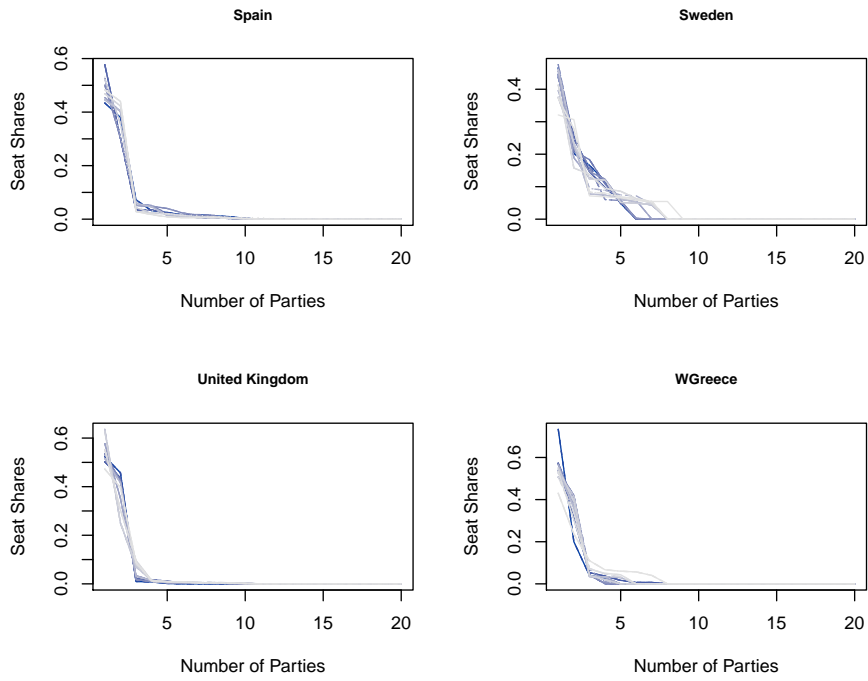
Notes: The plot shows the party system distributions in all 17 countries from 1970-2013. Different colors indicate different countries. The distribution varies across, and within countries, through the years.

Figure 3.1: Party System Distribution, 17 Countries

and in *Figure 3.3*, I present the party system distributions of selected countries. The shading shows the year of the observation. Lighter shades show the years closer to 1970, while the darker shades show years closer to 2015.

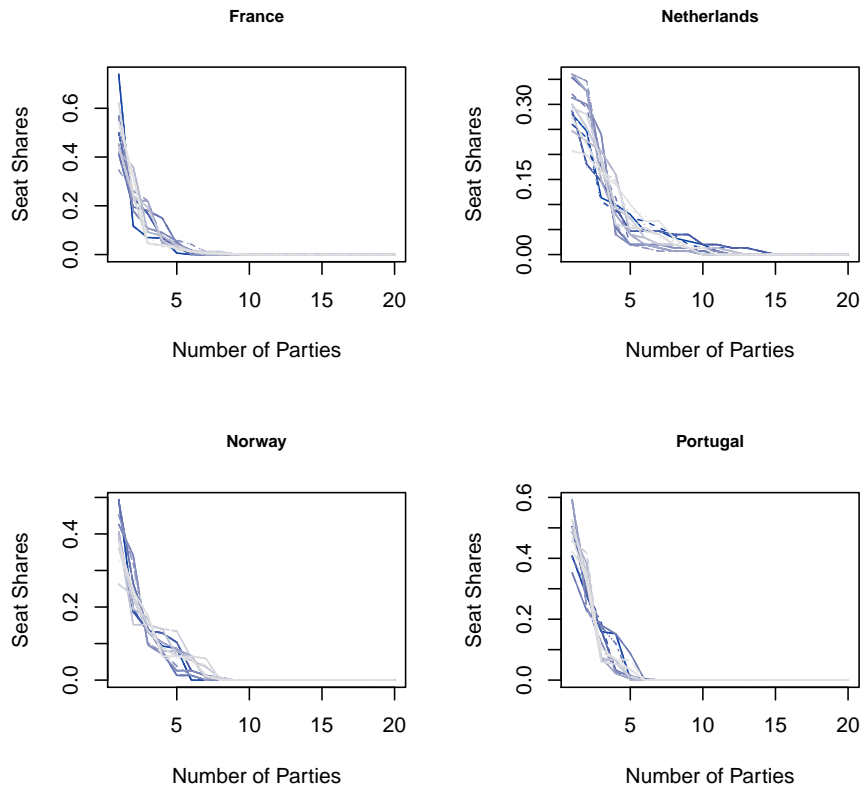
As *Figure 3.2* shows, in countries such as Spain, the UK, Greece and even Sweden, the party system is fairly stable. In these countries the party system is close to the two-party ideal, and the electoral system is close to majoritarian.² The odd country out of this group is Sweden, which has a big district PR electoral system. Still, the party system did not change

²The UK has a majoritarian (first-past-the-post) electoral system. In addition, while Spain and Greece have proportional representation (PR) electoral systems, they have small electoral districts and thus concentrated party systems.



Notes: The plot shows the party system distribution in four countries where the party system did not change throughout the years. Lighter shades show earlier years while darker shades show later years.

Figure 3.2: Party System Distribution by year, Spain, Sweden, United Kingdom, Greece considerably over the years.



Notes: The plot shows the party system distribution in four countries where the party system changed throughout the years. Lighter shades show earlier years while darker shades show later years.

Figure 3.3: Party System Distribution by year: France, the Netherlands, Norway, Portugal

On the other hand, we can see that the party system varies considerably over time in France, the Netherlands, Norway, and Portugal (*Figure 3.3*). From this group of countries Portugal is often classified together with Greece and Spain because the country has a PR system with small district size. However, as we can see, the party system has changed a lot throughout these years. In France,³ parties often split and merge due to coalitions competing. Overall, we can see that while there is some correlation between the electoral system of the countries and the stability of the countries' party systems we can see some deviation from the traditional wisdom that the party system is more stable in majoritarian countries.

³France has a two-round majority plurality electoral system. This means that voters have to vote again if a majority winner is not selected in the first round of the elections. Parties form coalitions to support the ideologically closest candidate in each district.

CHAPTER 4

Principal Component Analyses

4.1 Simple PCA

In this paper, I first analyze the results of a Principal Component Analysis (PCA) that I conducted to explore the structure of the party system dataset. PCA is a dimensionality reduction technique, in which the goal is to recover the minimum information needed to reproduce the maximum information present in the data matrix. The method projects high-dimensional data on a lower dimensional space. The lower dimensional space is determined by the directions in which the data varies most, so the least amount of information is lost during the projection. The projection is linear, so it can create an accurate summary of the data if the data is Gaussian distributed.

In practice, the PCA estimation first calculates the dimension where the data has the biggest variation. It then finds n orthogonal dimensions which explain the biggest part of the remaining variation within the data. Mathematically, we would like to map vectors x^n in a d -dimensional space onto z^n in a 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 in the following way: we are looking for such \tilde{x} 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 $\sum u_i =$

$\lambda_i u_i$, where Σ is the covariance matrix of the set of vectors, and λ are the eigenvalues (Bishop 1995, 310). Thus this is an eigen decomposition of the covariance matrix. In the result, the eigenvalues are the scale and the eigenvectors are the direction of the new reduced dimensions. We call the eigenvectors principal components. The eigenvector with the highest eigenvalue is the first principal component of the data set (the dimension in which the data has the most variation or the direction in which the data is the most dissimilar); the second principal component is orthogonal to this dimension etc.

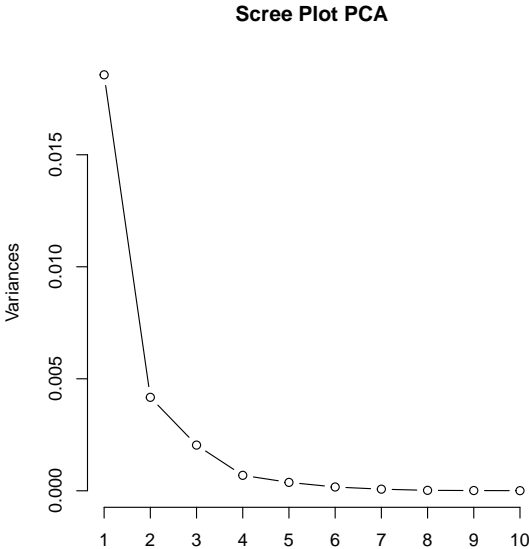
In this paper, I project the 719- dimensional party seatshare data matrix to a 20-dimensional space (the number of variables). To do this, I first mean center the data. Then, I find the eigenvalues and the eigenvectors of the covariance matrix. I do not scale the variables at this point. Transforming the variables to have unit variance would mean that the PCA is done on the correlation matrix instead of the covariance matrix. This could be useful if the variables do not have the same measurement. However, in this case the variables are all seat shares. Moreover, because the seat share data is compositional data, the scaling would change one important feature of the data— that the seat shares add to one. Thus, in this case, scaling may not be an optimal solution. In the later part of this paper, I discuss what happens to the analysis when I standardize the variables.

Table 4.1: Eigenvalues and Explained Variance, PCA

	Eigenvalues	Explained Variance	Cumulative Variance
1	0.0184119950	70.62	70.62
2	0.0043036798	16.51	87.12
3	0.0020294706	7.78	94.90
4	0.0006882216	2.64	97.54
5	0.0003666149	1.41	98.95
6	0.0001657361	0.64	99.59
7	0.0000709305	0.27	99.86
8	0.0000171304	0.07	99.92
9	0.0000108457	0.04	99.97
10	0.0000048573	0.02	99.98
11	0.0000025334	0.01	99.99
12	0.0000011470	0.00	100.00
Sum	0.0260737		

Since the purpose of the PCA analysis is to reduce the dimensionality of the matrix,

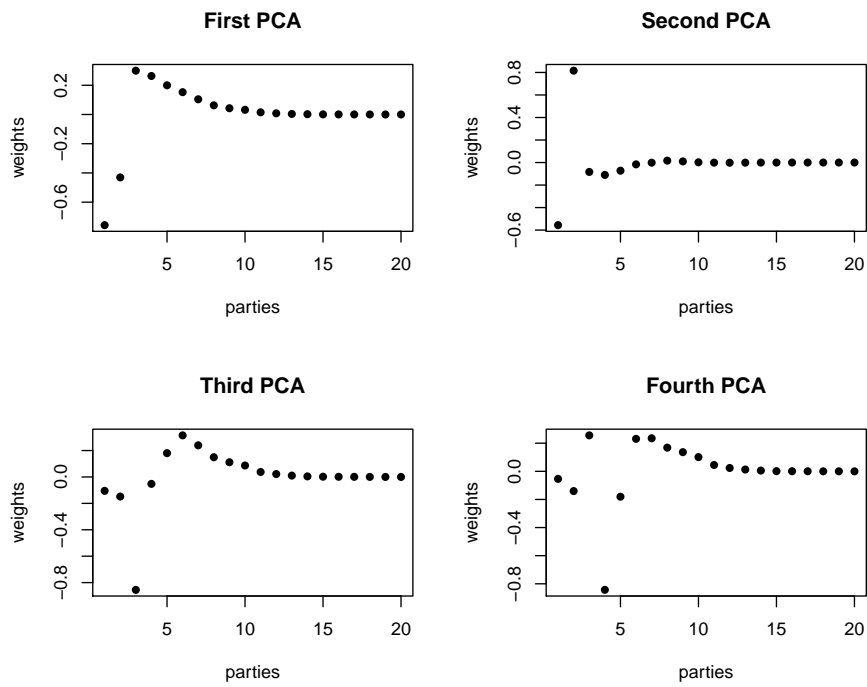
next we have to evaluate the appropriate number of dimensions to use for the analysis. I plot the screeplot, *Figure 4.1*, which shows in descending order how much of the total variance the eigenvalues explain. *Figure 4.1* shows that the first four eigenvalues account for most of the variation in the dataset. *Table 4.1* shows that the first four eigenvalues explain 97.54% of the variation in the data.



Notes: The plot shows the Scree Plot of the Simple PCA. On the x-axis are the Eigenvalues, on the y-axis the unexplained variance. We can see that the first four vectors account for most of the unexplained variance.

Figure 4.1: Scree Plot, PCA

Thus, in the following paragraphs, I focus on the meaning of these first four principal components. To analyze the source of the biggest variation in the data, next I picture the eigenvectors recovered by the PCA analysis. Here, I call the weight of the variables in the eigenvector as loadings. *Figure 4.2* shows the first four eigenvectors, or principal components, that the analysis has recovered. We can see which parties get a weight in separating the most dissimilar party systems over country-years. On the X-axes of the plots, we can see the number of parties. On the Y-axes of the plots we can see the weights that each party has in the given principal component.



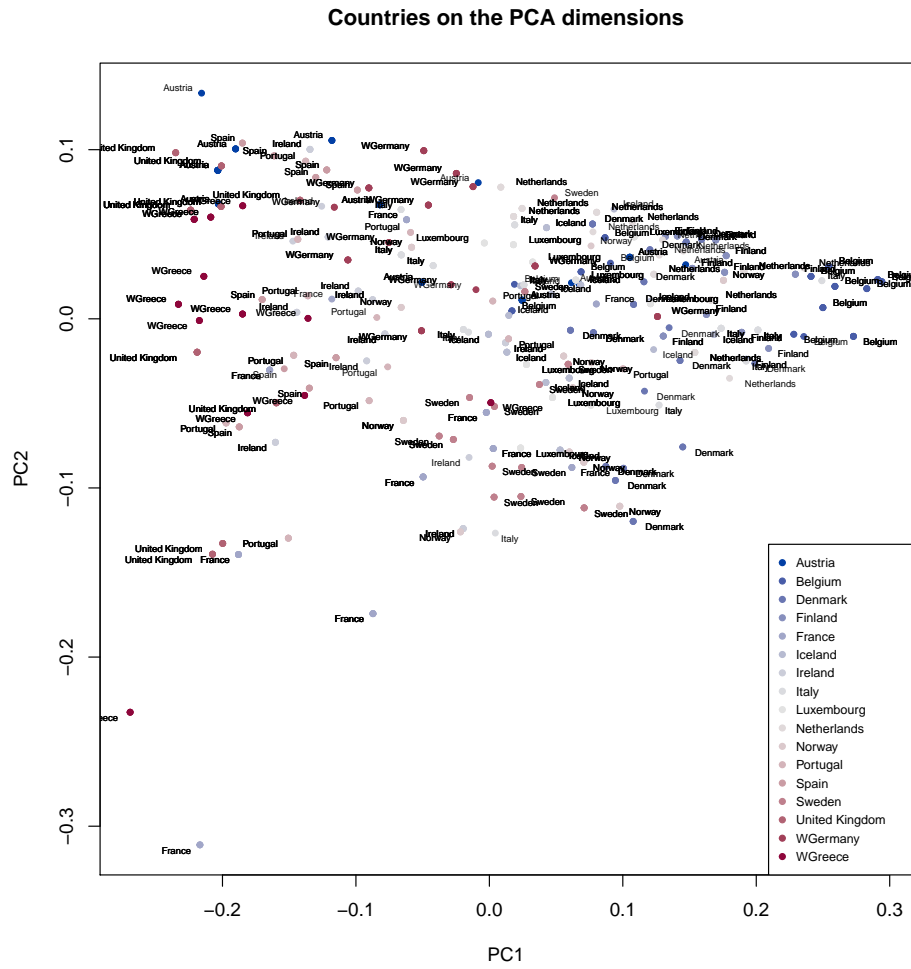
Notes: The plot shows the loadings, the weight of parties in determining the principal components in the PCA analysis.

Figure 4.2: Loadings, PCA

One of the benefits of the PCA analysis is that the lower dimensions of the data may be more easily interpretable than the complex dataset (Jolliffe 2002). The first three principal components seem to show a clear picture of what makes party systems most unlike. As *Figure 4.2* shows, the first dimension (PC1) contrasts countries where the sizes of the two biggest parties are big relative to the other parties, with countries where the sizes of the two biggest parties are small relative to the other parties. I call this dimension “Size of the Biggest Two Parties.” The second dimension (PC2) contrasts the countries where the sizes of the two biggest parties are close to each other with countries where the sizes of the two biggest parties are far from each other. This dimension contrasts countries with two party competition with countries with 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, so it contrasts countries with a big third party with countries with a small third party (we can call this dimension “Third Party”), while the fourth dimension (PC4) is somewhat unclear. This dimension seems to be defined by Parties 3-5, and, tentatively, I call it: “Multipartism.” From looking at the plot, this last dimension may separate countries that have a balanced multi-party system from countries in which the fourth and the fifth parties are non-existent or very small. In the next section of this paper, I analyze more closely what this fourth dimension might be. All of the dimensions (PC1, PC2, PC3 and PC4) are orthogonal to each other.

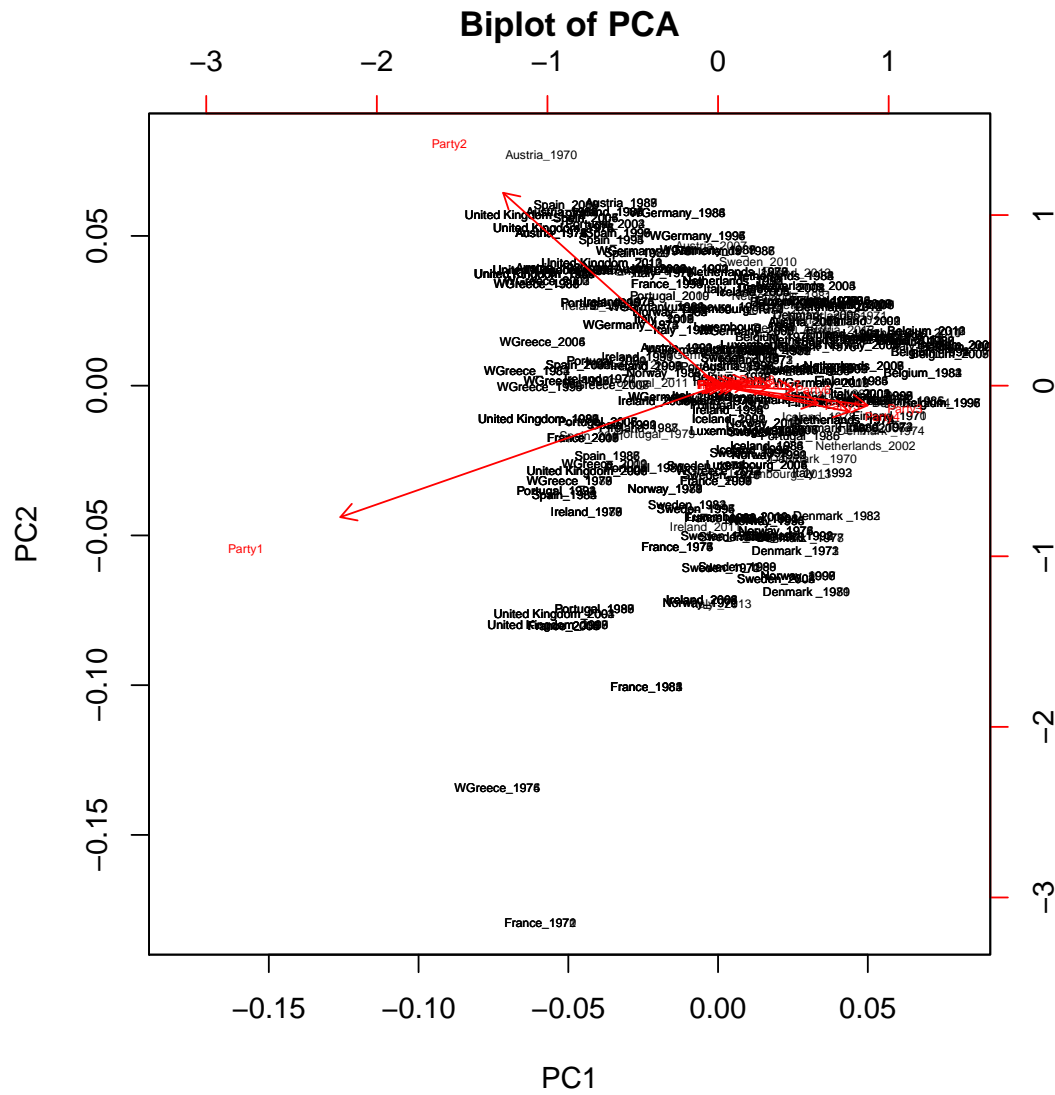
I plot the cases (which are the country-years) on the two-dimensional plane. First, *Figure 4.3* shows all the countries on the two-dimensional plane that the first two dimensions (PC1 and PC2) create. This means that the location of each country is its projected position on the PC1 and PC2 dimensions. The different colors indicate the different countries. We can see on one side of Dimension 1 the United Kingdom and Greece, while on the other end of this dimension Belgium and Finland. Dimension 2 separates Germany, Iceland, and Luxembourg from Sweden and Norway. Later, I discuss how these dimensions relate to the typologies and party system size measures.

Figure 4.4 shows the biplot. The biplot presents all the observations and the variables on a two-dimensional plane determined by PC1 and PC2. The parties (variables) are



Notes: All the cases projected onto the two-dimensional plane defined by the first (PC1) and second (PC2) dimensions recovered by the simple PCA. The colors indicate the countries.

Figure 4.3: All Observations on the Two-dimensional PCA Plane



Notes: Biplot of the PCA. The red arrows show how the variables (parties) relate to the first two dimensions of the PCA analysis. In addition, we can see the cases (country-years) on the two-dimensional plane determined by the first two dimensions.

Figure 4.4: Biplot, PCA

represented with red arrows, the n -observations are the dots. The arrows in relation to the two-dimensional plane show how the dimensions are defined. The arrows are the least squares projections of the variables to the plane, and the lengths of the rays are proportional to the variances that the variables explain (Aitchison and Greenacre 2002). We can see in *Figure 4.4* that the first dimension PC1, is defined by *Party 1* and *Party 2*. The second dimension, PC2, on the other hand, is defined by the difference between the sizes of *Party 1* and *Party 2* (in the plot the arrows are pointing to different vertical directions). The size of the arrows indicates how much variation gets explained by the variables. The arrows of the variables: *Party 1* and *Party 2* are much longer than the rest of the arrows. This is because the two variables have high variances. One of the characteristics of the PCA as a method is that the result can get dominated by the influence of variables with high variances. As in *Figure 4.3*, in the biplot *Figure 4.4*, we can also see how the dimensions separate the cases; however, as I have all country-years in this plot, it is a little bit harder to read. Later in the paper, I plot and analyze the location of each country, and I discuss the results in greater detail.

4.1.1 PCA on the Residuals

In this section, I am going to investigate further what PC3 and PC4 mean. The analysis above shows that PC1 and PC2 are mainly influenced by the sizes of the first two parties, while PC3 is defined by whether or not there is a third party in the party system. One reason why it is difficult to interpret PC4 could be that the sizes of the first two parties (hence the variance of these variables) are so big that they mask how the other variables relate to each other. To address this problem, I control for the sizes of *Party 1* and *Party 2* on the sizes of all the other parties to clarify the meaning of the lower dimensions. With this step, I essentially normalize all the party systems by the size of the first two parties. In practice, I regress all other party sizes on *Party 1* and *Party 2* (Vitt et al. 1997).

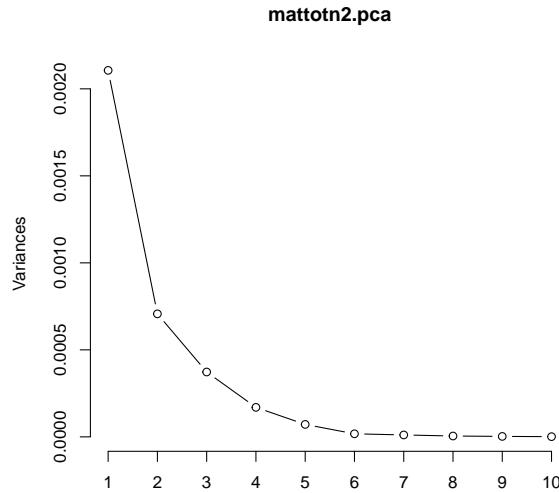
On the residuals, I again perform a PCA analysis. We can examine whether this method leads to a loss of information if we compare the variance explained by the third and

Table 4.2: Eigenvalues and Explained Variance, PCA on the Residuals

	Eigenvalues	Explained Variance	Cumulative Variance
1	0.0021061443	60.79	60.79
2	0.0007068498	20.40	81.20
3	0.0003727539	10.76	91.96
4	0.0001695409	4.89	96.85
5	0.0000714137	2.06	98.91
6	0.0000175559	0.51	99.42
7	0.0000109821	0.32	99.74
8	0.0000048574	0.14	99.88
9	0.0000025608	0.07	99.95
10	0.0000011537	0.03	99.98
11	0.0000004139	0.01	100.00
Sum	0.003464353		

fourth eigenvalues in the original PCA to the first and second eigenvalues of the PCA on the residuals. Overall, *Table 4.2* shows that the remaining variation in the data matrix is 0.003464353. Out of this, the first eigenvalue explains 60.79%, which is 0.0021 comparable to what the third eigenvalue in the original PCA (0.0020) explains (*Table 4.2*). The second eigen value in the PCA on the residuals explains 20.4% of the variation, which is 0.00070, similar to the fourth eigenvalue in the original PCA 0.0006 (*Table 4.2*). The screeplot of the PCA on the residuals *Figure 4.5* shows that the first eigenvalue explains less variation than the first eigenvector of the PCA. As previously however, most of the variation (96.85%) , is explained by the first four eigenvalues.

As earlier, I show the loadings that determine these four lower dimensions (*Figure 4.6*). The first dimension is mostly influenced by the size of the third biggest party (this is what we have seen in the original PCA as well), and somewhat by *Party 4*, *Party 5* and *Party 6*. Thus, PC3 (in the original PCA) is probably indeed “Third Party,” and separates countries with relatively small third parties from countries with relatively big third parties. This makes sense if we compare this finding to the traditional typologies, which divided countries that had two-party systems from countries that had multi party systems. The second dimension is determined by the competition between *Party 3*, *Party 4* and *Party 5*. While I called PC4 “Multipartism.” before, the correct idea is probably the balance in power

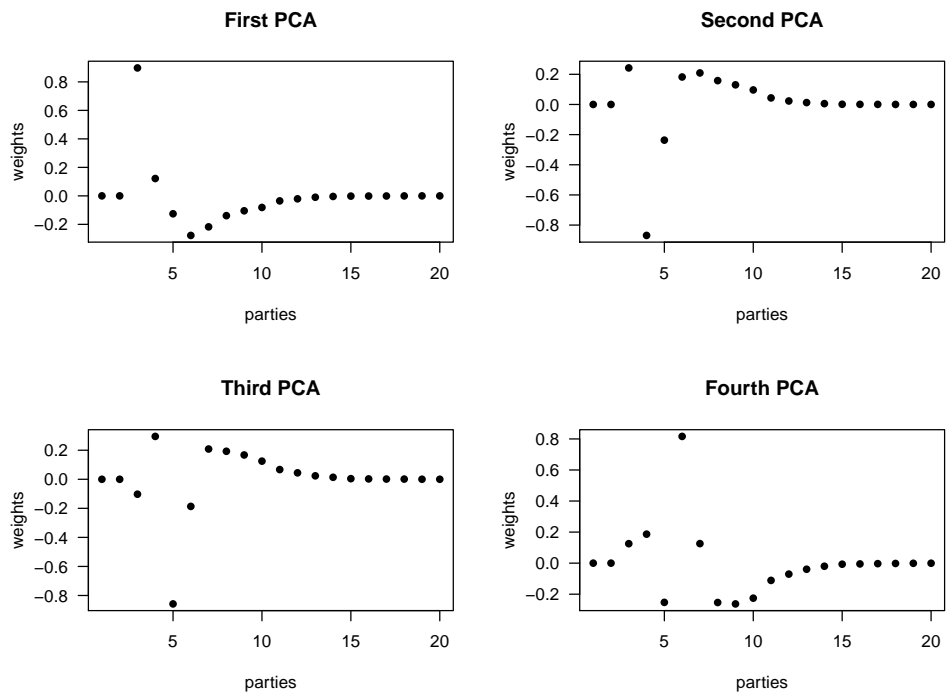


Notes: Screeplot of the PCA analysis on the residuals. The plot shows the variance explained by the eigenvalues. While the first four eigenvalues explain most of the variance as previously, the slope is less steep.

Figure 4.5: Screeplot, PCA on the Residuals

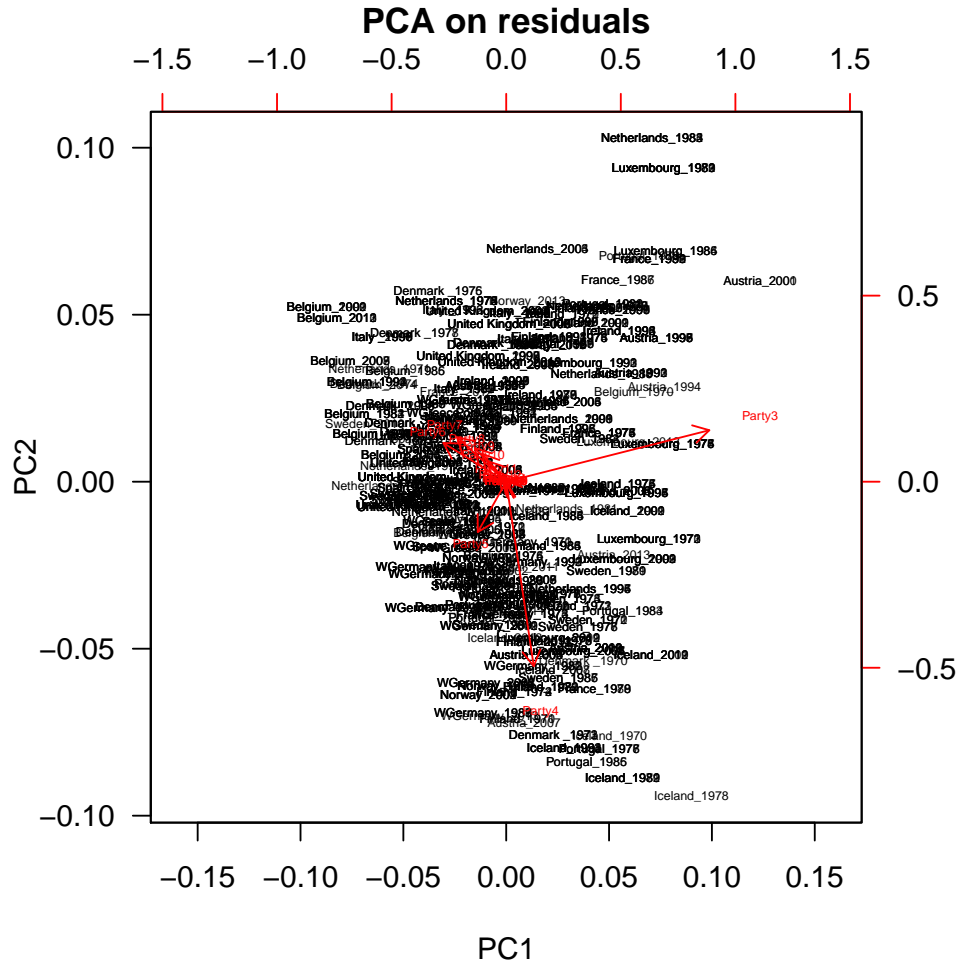
within the party system (whether the smaller parties are equal sized or not), as discussed above.

Similarly to the higher dimensional analysis, the biplot (*Figure 4.7*) plots the variables with the observations on the plane determined by the first two principal components. In this plot, we can see that the variation is smaller between the observations as it was previously as we explained some of the variation in the sizes of the parties with the sizes of the biggest two parties. No clear pattern of countries emerges in the plot. Luxembourg and Austria are on one end of the “Third Party” dimension, while Belgium is on the other side of this dimension (which means that compared to Belgium, Luxembourg and Austria have relatively small third parties). In the second dimension, (which is the original PC4, now identified as “Balance of the Party System”) Belgium is on one side of the dimension opposite to Iceland and West Germany. Overall, there seems to be a bigger within country variation on these dimensions than on the first two dimensions of the original PCA.



Notes: The plot shows the weight of parties in determining the principal components based on the PCA analysis on the residuals, after regressing the data on Party 1 and Party 2.

Figure 4.6: Loadings, PCA on the Residuals



Notes: Biplot of the PCA on the residuals. The red arrows show how the variables (parties) relate to the first two dimensions of the PCA analysis. In addition, we can see the cases (country-years) on the two-dimensional plane determined by the first two dimensions.

Figure 4.7: Biplot, PCA on the Residuals

4.2 Introducing Non-linearity to the PCA Analysis

The PCA analysis has limitations. During the PCA analysis, the data is linearly projected to the new dimensions. The old data is decomposed as a linear combination of lower level dimensions (eigenvectors) and weights (eigenvalues). This projection finds the correct solution if the data is close to Gaussian distributed; however, if the data is non-linear, we may not find the most important dimensions of the data (Bishop 1995).

In the following sections, I am going to present two ways that non-linearity can be introduced into the PCA analysis, and I am going to analyze the dataset through these methods. The first method, the kernel PCA (kPCA), non-linearly transforms the dimensions on which we are projecting the data, while the Non-linear Principal Component Analysis (NLPCA) finds the optimal, potentially non-linear, quantifications (transformations) of the data, at the same time as it projects the new data on linear lower level dimensions.

4.2.1 Kernel Principal Component Analysis (kPCA)

The Kernel Principal Component Analysis (kPCA) offers one solution to how to find the appropriate reduced dimensional space if the data is non-linear. With this method, we first map the data to a higher dimensional non-linear feature space. After this, in this non-linear subspace, we do a traditional PCA calculation (Schölkopf, Smola and Müller 1997). Thus, the result will be non-linear on the original data space. Schölkopf and Smola (2002) find that kernel PCA provides a better classification rate than does the linear PCA, and more components can be extracted with this method than with the linear PCA.

As I discussed above, the minimum of the sum of squared errors in the PCA estimation can be found when the covariance matrix is diagonalized. The kernel PCA proceeds as the regular PCA. However, the covariance matrix of the data is transformed by the kernel function. The covariance matrix of the non-transformed data is the following $\Sigma = \frac{1}{N} \sum_{i=1}^N x_i x_i^T$ (where $x_k, k = 1 \dots N, x_k \in R^N, \sum_k x_k = 0$).

In this instance, we transform the data to a feature space F by a function ϕ , which

will result in: $R^N \rightarrow F, x \rightarrow X$. Hence the data will be the following: $\phi(x_1)\dots\phi(x_N)$ and the covariance matrix of the data will look like this: $\bar{\Sigma} = \frac{1}{N} \sum_{i=1}^N \phi(x_i)\phi(x_i)^T$. After this transformation, the solution is similar to the regular PCA. To minimize the loss function we find the eigenvectors satisfying $\lambda_i u_i = \bar{\Sigma} u_i$. u_k are the directions of the space. Let's define $u_k = \sum_{j=1}^N \alpha_j^k \phi(x_j)$. The inner product space is $K = k(x, y) = \langle \phi(x), \phi(y) \rangle = \phi(x)^T \phi(y)$ by definition. We can use the kernel trick here, since for the estimation of the data matrix the PCA uses the inner products of component scores (eigenvalues) and component loadings (eigenvectors), consequently, we do not need the explicit function to calculate these. The kernel trick means that we do not explicitly use the high order function, but, instead, we directly evaluate kernel k . We use the inner product $\langle \phi(x), \phi(y) \rangle$ between the images of two data points x, y in the "feature space" (ϕ space).

While the features $\phi(x_1)\dots\phi(x_N)$ are not unique, their dot product is unique. As we do not use the explicit function, however, we cannot compute the principal components themselves, only the kernel projected data which is computationally given by: $\langle u_k, \phi(x_j) \rangle = \langle \sum_{j=1}^N \alpha_j^k \phi(x_j), \phi(x_j) \rangle = \sum_{j=1}^N \alpha_j^k \langle \phi(x_i) \phi(x_j) \rangle = \sum_{j=1}^N \alpha_j^k k(x_i, x_j) = K \vec{\alpha}$ (Schölkopf and Smola 2002).

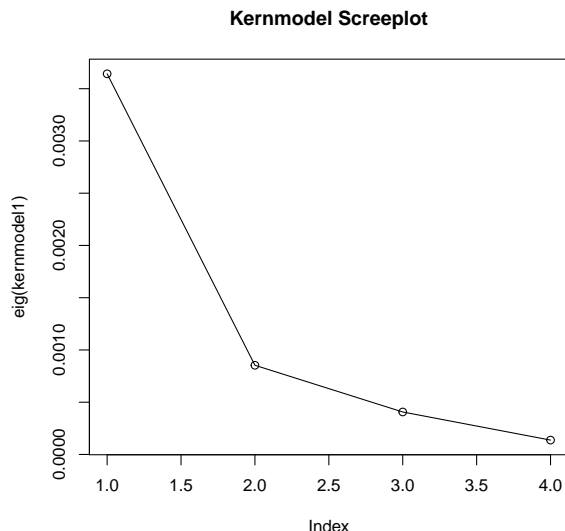
Table 4.3: Eigenvalues and Explained Variance, kPCA

	Eigenvalues	Explained Variance	Cumulative Variance
Comp.1	0.0036424550	72.27	72.27
Comp.2	0.0008534100	16.93	89.21
Comp.3	0.0004065058	8.07	97.27
Comp.4	0.0001375362	2.73	100.00
Sum	0.005039907		

For my analysis I use the Gaussian Radial Basis as my kernel function. This kernel, $k(x, x') = \exp(-\sigma ||x - x'||^2)$ is a general purpose, smooth kernel that we can use if we do have deeper knowledge about the structure of the data.¹ Below, I present the results of the kPCA analysis. *Figure 4.8* shows the eigenvalues that the kPCA recovers in the high dimensional space in descending order. The first eigenvalue explains most of the variation

¹the Gaussian kernel is "universal." It is positive definite and they are invariant under the Euclidean group. These are desirable properties if we want to estimate bounded continuous functions (Hofmann, Schölkopf and Smola 2008)

in the data. *Table 4.3* shows that in this case the first eigenvalue explains 72.72% of the variation. Overall, the four eigenvalues explain 100% of the variation.

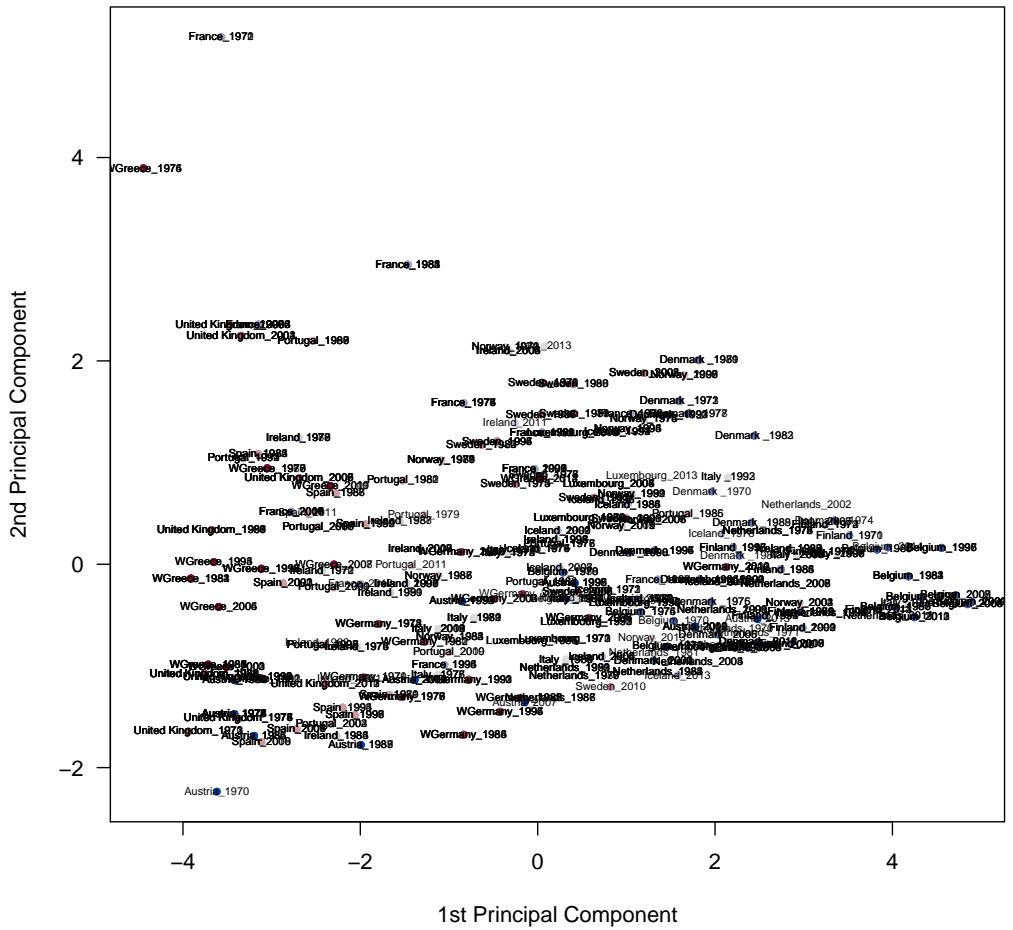


Notes: The plot shows the screeplot Kernel PCA.

Figure 4.8: Scree Plot, kPCA

Figure 4.9 shows the observations in the two-dimensional plane determined by the dimensions kPC1 and kPC2. This plot is similar to the PCA plots discussed previously (the PC2 has the loadings in the opposite direction from the PCA calculation, but since any PCA analysis is non-directional method, this does not have any impact on the analysis). Even though we cannot extract the loadings from this estimation process directly, we can see that the first two principal components are very similar to the first two principal components I obtained from the normal principal component analysis. To demonstrate this connection I created show the covariances between the transformed datasets based on the first four PCAs of the linear and the kernel PCA. (*Table 4.4*) shows that the respective principal-components are related. Since the results of the PCA analysis are easier to interpret than the results of the kPCA analysis. Because the two sets of results are reasonably similar, I will use the PCA dimensions later in this paper.

Kernel PCA



Notes: The plot shows the result of the Kernel PCA analysis

Figure 4.9: Biplot, kPCA

Table 4.4: kPCA and PCA Covariances

	kPC1	kPC2	kPC3	kPC4
PC1	0.21	-0.00	-0.00	-0.00
PC2	-0.00	-0.05	0.00	-0.00
PC3	-0.00	-0.00	-0.02	0.00
PC4	-0.00	0.00	-0.00	-0.01

4.2.2 Non-Linear Principal Component Analysis (NLPCA)

I also conducted a Non-Linear Component Analysis (NLPCA) on the data. In this section, I present the results of this analysis, and I compare them to the results of the PCA and the kPCA analyses. The NLPCA is a special case of multiple correspondence analysis or homogeneity analysis. Homogeneity analysis maximizes the correlation between variables at the same time as it does optimal scaling of the variables (optimal quantification of the variables). One generalization of this method is a non-metric principal components analysis, for which we can use not only categorical but also ordinal and ratio variables. (Michailidis and de Leeuw 1998).

Thus, contrary to the kPCA, the non-linearity of the NLPCA does not come from the transformation of the space on which we project the data, but from the potentially non-linear optimization of the data matrix. During the process, the data matrix is optimized to ensure that the variable variances are explained to the greatest degree possible. The traditional PCA analysis minimizes the loss function over the eigenvectors and eigenvalues. The NLPCA analysis also minimizes the loss over the admissible transformations of the data columns (de Leeuw 2005). This means that the PCA loss function: $E_M = \frac{1}{2} \sum_{n=1}^N \sum_{i=M+1}^d (z_i^n - b_i)^2$ is not only minimized with respect to b_i but also with respect to z_i (or x_i , since $z_i = u_i^T x$). Thus, the solution will be: $\sum u_i = \lambda_i u_i(X)$.²

²Furthermore, de Leeuw (2005) discusses that not all transformations are admissible: the first restriction is that the transformed variables must be in a convex cone K . Convex cones are defined by $x \in K$ implies $\alpha x \in K$ for all real $\alpha \leq 0$ and $x \in K$ and $y \in K$ implies $x + y \in K$. However, since α is in the cone, this means that its positive linear function: $\alpha x + \beta$ with $\alpha \leq 0$ must also be in a convex cone. Thus, this could lead to the trivial solution that all transformations will be set to zero. To avoid this, de Leeuw argues that we have to make another restriction: we can redefine the cone to only contain centered vectors, so the cone $K_j \cap S$ is going to be a convex cone of centered vectors. Because of this the optimization problem finds admissible transformations of the variables where the sum of the $n - p$ smallest eigenvectors of the

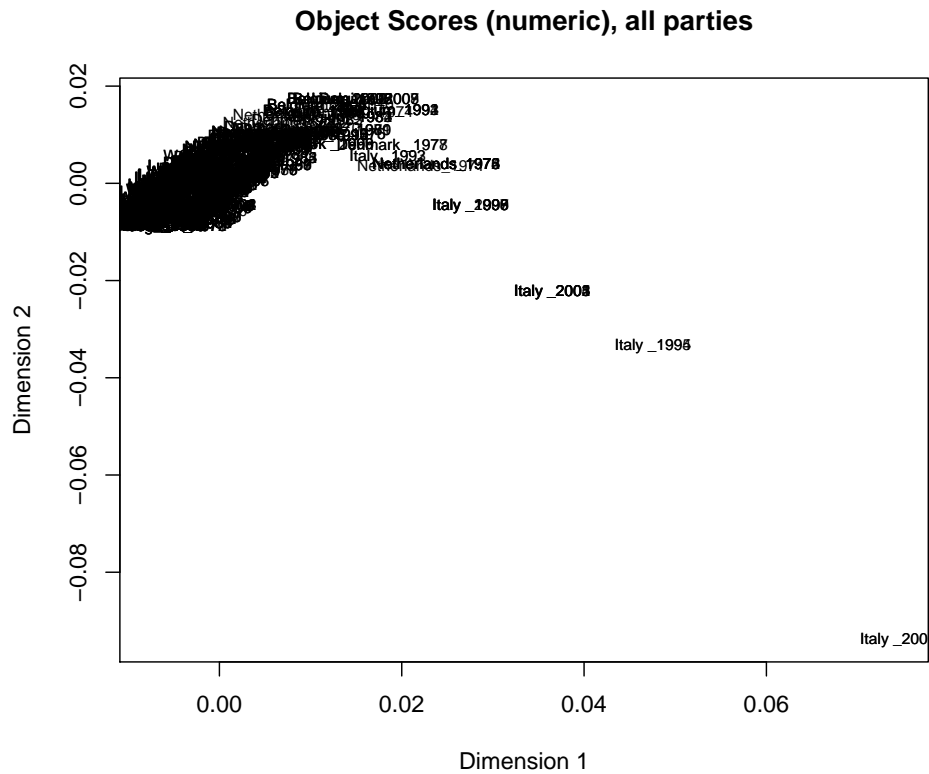
This means that the PCA is performed while the variables are also optimized. As an algorithm, the method alternates between the two processes in an iterative way, until the loss function is minimized, and the algorithm converges. At this point, neither the variable quantifications nor the PCA solution change (Linting et al. 2007). The NLPCA can be used if the data is non numerical or if it is rank ordered since this method handles the non-quantifiable distances between variables and can also clarify the results if there is non-linear relationship between the variables (de Leeuw 2005). In this analysis I use the party seat shares as numerical data, since in the dataset the parties are ordered from largest to the smallest. Even though I specify the data as numerical, the NLPCA method considers these variables as categorical. Thus, each observed numerical value becomes a category (Linting 2007).

As I discussed above, I did not standardize the data frame in the PCA and kPCA calculations. As I also discussed above the two solutions from the PCA and the kPCA analysis were similar to each other. However, the NLPCA solution is quite different from these two solutions (*Figure 4.10*). This is because the NLPCA method essentially standardizes the variables when it creates optimal quantifications. By dividing the mean centered variables with their standard deviation we can standardize the variables to have unit variance (which equals to performing the PCA on the correlation matrix). As we have seen above, in the party system dataset the variances of the two first variables are big, and thus they influence the solution the most. Through standardization, we give all the variables equal weight. This diminishes the influence of the variances of the biggest two parties and could lead to a solution which shows the structure determined by the sizes of the other parties.³

As the biplot of the NLPCA shows (*Figure 4.10*), Italy is separated from the rest of the countries on the first dimension. Italy is a unique case because the country had the

correlation matrix is minimized (or the sum of the p largest eigenvectors is maximized). Thus the final form of the NLPCA is $\max_{x_j \in K_j \cap S} \sum_{s=1}^p \lambda_s(R(X))$, where the real valued function ϕ is defined as the sum of the p largest eigenvalues of the correlation matrix $R(X)$ (de Leeuw 2005).

³Thus the variation of the bigger parties (Party 1 and Party 2) becomes smaller and the variation of the smaller parties (Parties 8-20) becomes bigger. Thus after the standardization Dimensions 1(NLPC1) and 2 (NLPC2) are influenced more by smaller parties than the first and second dimensions I recovered with PCA and kPCA.



Notes: The plot shows the location of the 17 countries on the two- dimensional plane determined by the first two dimensions that the NLPCA analysis found (with all the parties in the dataset).

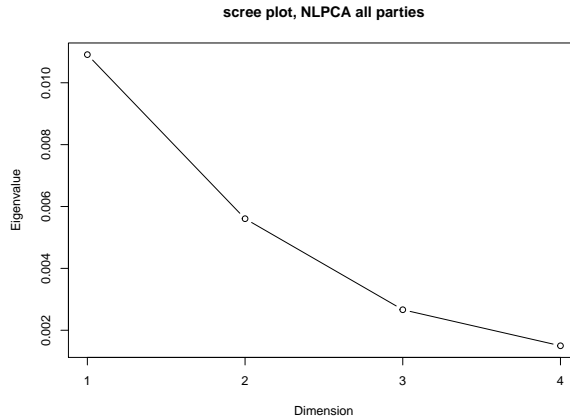
Figure 4.10: NLPCA Objectplot, All Parties

most parties in the legislature out of all countries (20 in 2006 and 2007). Neither the non-standardized kPCA nor the PCA analysis revealed that Italy is a special case previously.⁴

The screeplot (*Figure 4.11*) shows that the first eigenvalue that the NLPCA extracts, explains less variation in the data, compared to the the first eigenvalue that the PCA and the kPCA methods have found.⁵ This is because most of the variation in the data has been generated by the variation in the sizes of the first two parties.

⁴If we remove Italy from the dataset, the first dimension separates Belgium from the rest of the countries (as we have seen, the kPCA and PCA solutions also put Belgium at the far end of the dimension that separated countries with two big parties from the rest of the countries).

⁵The scree plot of the NLPCA *Figure 4.11* is less steep than the ones we have seen before: *Figure 4.1*, *Figure 4.5*.



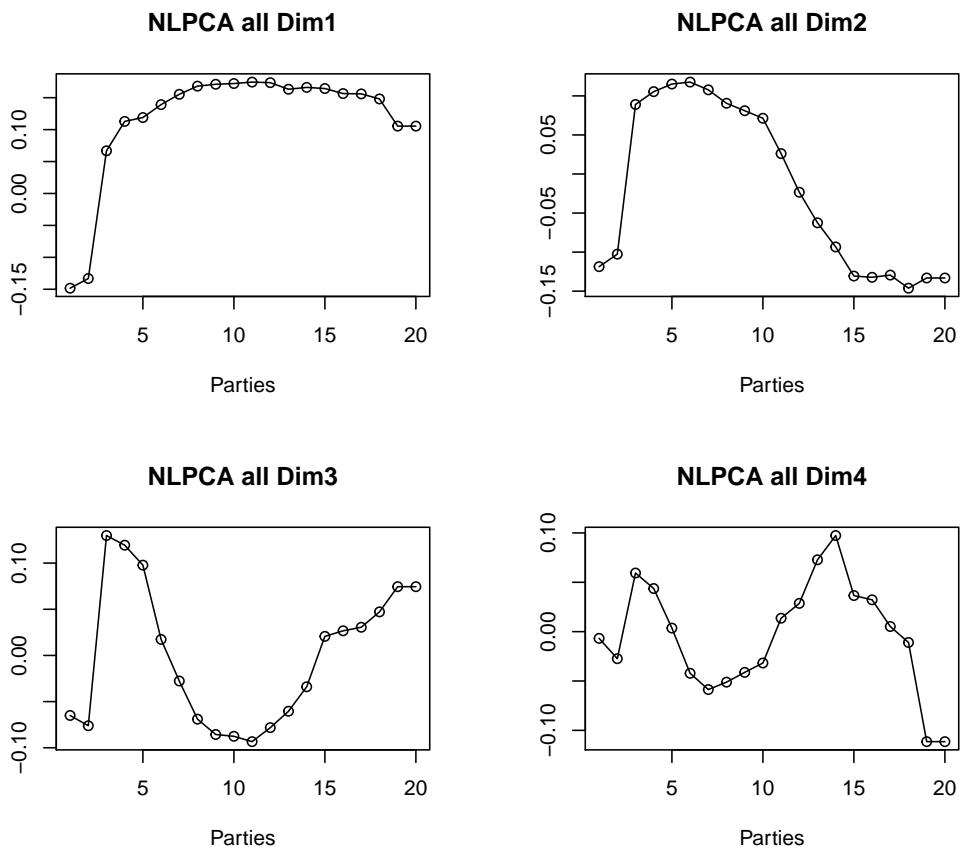
Notes: The plot shows the scree plot of the NLPCA analysis, when all parties are in the data matrix.

Figure 4.11: NLPCA, Scree Plot, All Parties

Because I reduced the variance of the variables the first and second principal components that the NLPCA analysis finds are influenced less by the sizes of the two biggest parties. Because all parties get equal weights in determining the dimensions, NLPC1 and NLPC2 separate the countries based on the actual number of parties. At the same time, NLPC3 separates moderate party systems (party systems up to 5 parties) from the very large party systems (*Figure 4.12*).⁶ In order to be able to compare the results of the NLPCA analysis to the kPCA and to the PCA results, in the following section, I reduce the number of parties to ten. This way I can avoid that the countries with fragmented party systems would define the Dimension 1 of the NPLCA.

Before reducing the number of the variables however, I examine the impact of this change on the NLPCA and the PCA results. *Figure 4.13* shows how the loadings change if we change the number of parties in the NLPCA analysis. The colors show the number of parties in the analysis. The number of parties start at 4 and go up to 20. *Figure 4.13* shows

⁶*Figure 4.12* shows that even though the sizes of the first two parties are in the opposite direction from the rest of the parties, the sizes of the smaller parties weigh almost the same as the sizes of the bigger parties. As Italy has many small parties that have a high weight in this analysis, the country gets separated from the other countries. In NLPC1 apart from Party 1 and Party 2, Parties 7-16 have the highest loadings. NLPC2 is determined by the sizes of Party 1 and 2 and also it is influenced by Parties 3-11. NLPC3 is influenced by Party 1 and Party 2 and Parties 3-5, while Dimension 4 has high loadings from the smaller parties, Parties 15-20.



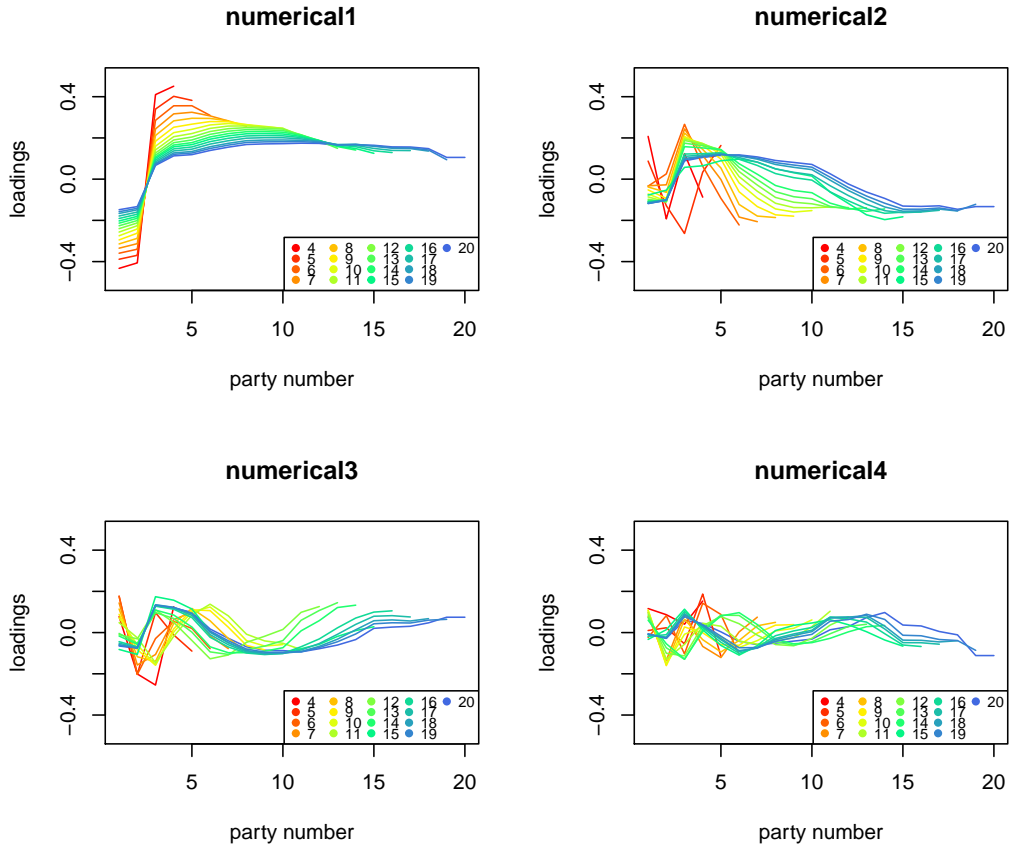
Notes: The plot shows the weight of parties in determining the principal components based on the NLPCA analysis.

Figure 4.12: NLPCA, Loadings, All Parties

that the NLPCA does not find exactly the same solution when we increase the number of parties. When we increase the number of parties the first parties get less weight. However, the solutions are similar in their underlying structure. We can contrast the NLPCA solution (*Figure 4.13*) with the scaled (*Figure 4.14*) and the unscaled (*Figure 4.15*) PCA solutions.

Figure 4.15 shows that even when we limit the number of parties radically, the unscaled PCA analysis finds the same first two dimensions, and while the sign of the loadings might change, the PC3 and PC4 remain very similar as well. This is because the small parties get less weight in this analysis than the bigger parties, and the parties after the fourth party tend to be small. This is not the case when we scale the variables *Figure 4.14*.

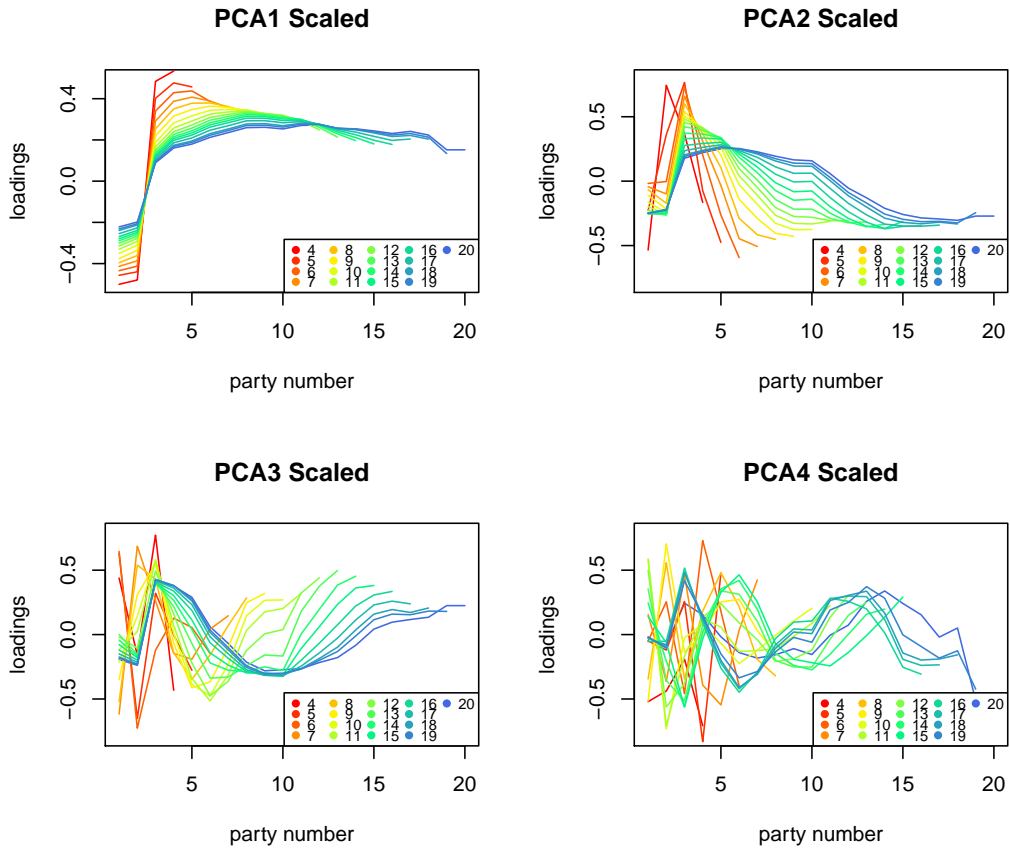
A comparison between the scaled PCA *Figure 4.14* and the NLPCA *Figure 4.13*



Notes: The plot shows how the NLPCA changes when we run the analysis on fewer and fewer parties. The color of the lines indicates the number of parties in the analysis.

Figure 4.13: The Sensitivity of the NLPCA Results to the Change in the Number of Parties

reveals that the NLPCA solution is not the same as the scaled PCA solution. While eventually the dimensions that the two methods find seem to be similar (apart from the fact that Dimension 4 is still unclear) the loadings change more when we change the number of parties in case of the scaled PCA. Overall, it seems that small changes in the party system can influence the dimensions that the scaled PCA recovers more than it can influence the dimensions that the NLPCA recovers. In contrast, the unscaled PCA solution remains pretty steady when we include the smaller parties *Figure 4.15* This may indicate that we have to consider a trade-off: the NLPCA analysis may be more suitable if we want to explore party system changes when small disturbances happen within a single country, while the PCA analysis may be more suitable for cross-country, cross-era comparison.

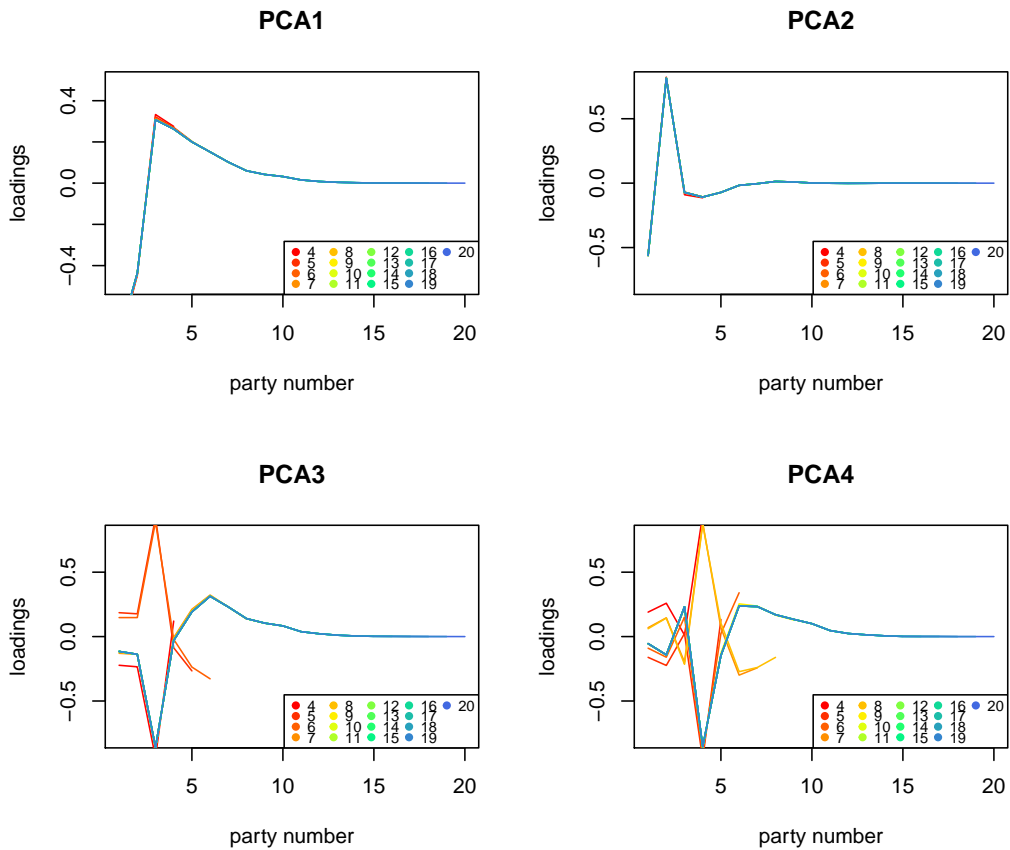


Notes: The plot shows how the scaled PCA changes when we run the analysis on fewer and fewer parties. The color of the lines indicates the number of parties in the analysis.

Figure 4.14: The Sensitivity of the Scaled PCA Results to the Change in the Number of Parties

Next, I analyze the results of the NLPCA results that I get when I limit the parties to the 10 biggest parties in the legislature. As the following plots show, the dimensions that the NLPCA recovers from the limited data, are similar to the ones that the methods finds with the full dataset- although there are some differences.

Figure 4.17 shows that when there are fewer parties, the first eigenvalue that the NLPCA finds explains more variation of the data compared to the rest of the eigenvalues than when all the parties are included. Again, this happens because if there are only the biggest 10 parties included, the first two parties get more weight than if all parties are included. *Figure 4.18* shows the two-dimensional plane that the NLPCA (with 10 parties)

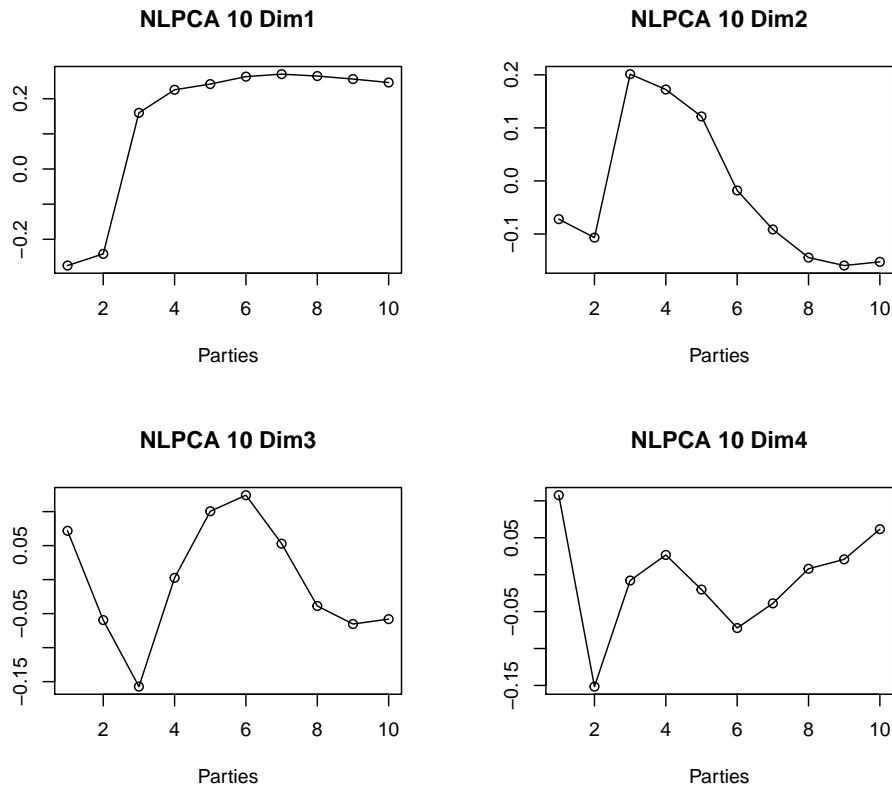


Notes: The plot shows how the unscaled PCA changes when we run the analysis on fewer and fewer parties. The color of the lines indicates the number of parties in the analysis.

Figure 4.15: The Sensitivity of the Unscaled PCA Results to the Change in the Number of Parties

finds, and the object scores of the countries on these dimensions.

The object plot (*Figure 4.18*) shows that Dimension 1 that separates the parties (similarly to the PCA and the kPCA) is based on the size of the party system. In one end of the dimension we can see Greece, Spain, the United Kingdom (countries that have concentrated party systems) while at the other end of the dimension we can see Belgium, the Netherlands and Finland (countries that have fragmented party systems). On one side of the second dimension we can see Luxembourg, Iceland, while on the other side Italy, Denmark, Belgium and Spain. Again, in this dimension the countries seem to change their positions throughout the years.



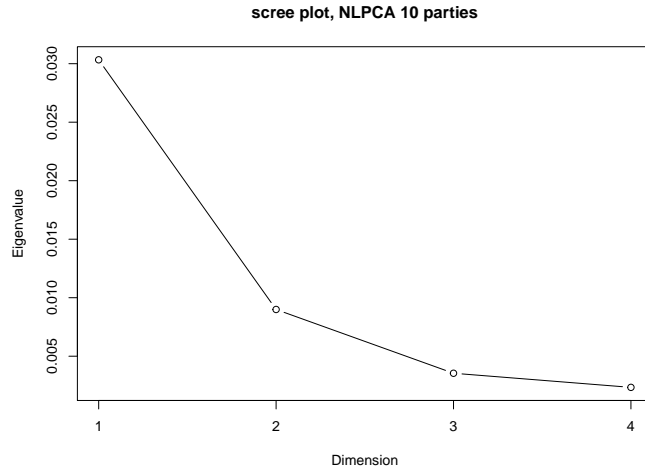
Notes: The plot shows the weight of parties in determining the principal components based on the NLPCA analysis on the 10 biggest parties in each country.

Figure 4.16: Loadings, NLPCA, Ten Parties

4.3 Issues with the Compositional Dataset

As we have seen, the result of the PCA analysis depends on whether we have the full dataset or just part of the data. The cause of this problem is the structure of the data. My data is a compositional dataset, the party seat shares add up to one $\sum s_i = 1$. The size of each individual data point depends on the size of the others within a case (Aitchison 1983). The potential issue with this type of data is that the correlations between the variables might have a negative bias (Jolliffe 2002).⁷ Also Aitchison (1983) notes that one of the issues with

⁷If we have a D part composition $[x_1, \dots, x_D]$ where $\sum_{i=1}^D x_i = 1$ the covariance is going to be $cov(x_1, x_1 + \dots + x_D) = 0$ thus $cov(x_1, x_2) + \dots + cov(x_1, x_D) = -var(x_1)$. According to Aitchison (n.d.) this means that there will be at least one negative element per row in the covariance matrix.



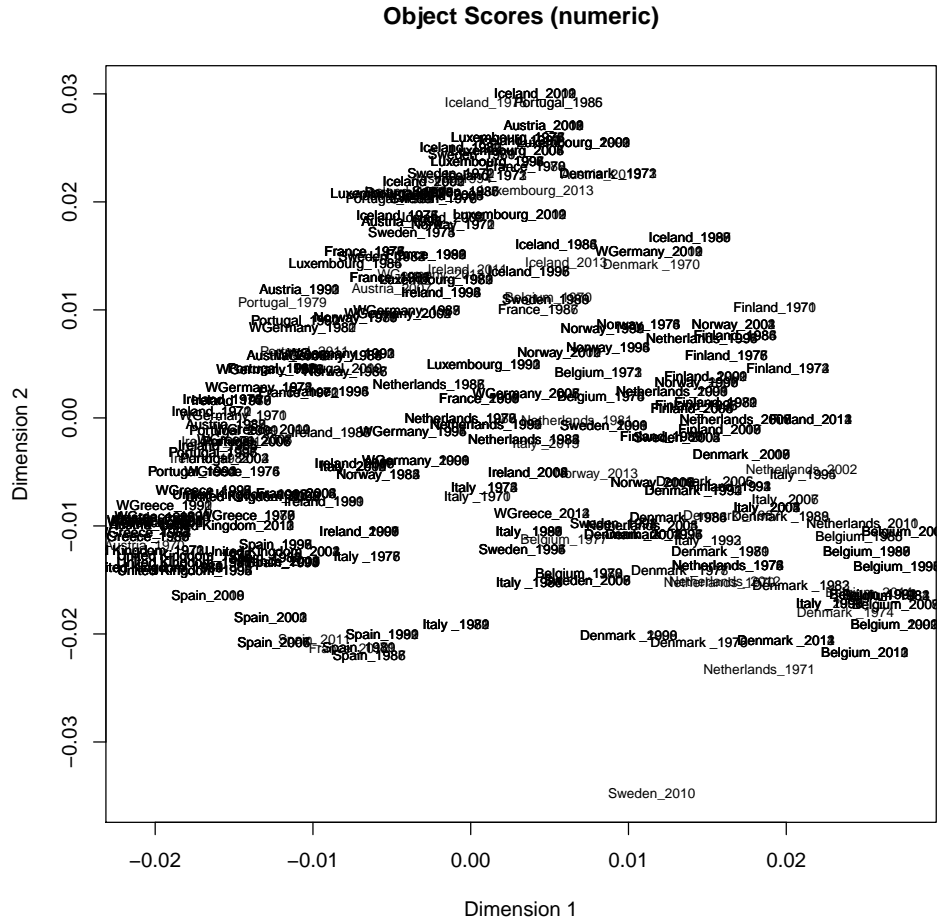
Notes: The plot shows the scree plot of the NLPCA analysis on the 10 biggest parties in each country.

Figure 4.17: Scree Plot, NLPCA, 10 Parties

the compositional dataset is that the dataset does not have *subcompositional coherence*. This means that if we have only a subset of the data, the PCA analysis on the covariance of this subset will lead to a different result from an analysis on the entire dataset. It has been widely debated in the literature how to run a PCA analysis on a compositional data set. One solution would be to leave one party out of all the party systems and calculate the PCA on the remaining data. However, in my dataset the party systems vary widely in size. This means that leaving out one party from all party systems could change the analysis considerably (as the smallest parties in some countries are relatively big compared to other countries). Below, I apply some of the techniques that previous authors suggested to analyze compositional data. First, I log-transform the variables, second, I perform a PCA on non-centered variables.

4.3.1 PCA on Log-ratio Transformed Variables

One recommendation about how to perform any calculation on a compositional dataset comes from Aitchison (1986), who suggests the log-ratio transformation of the data. He argues that this transformation makes the observations uncorrelated, and solves the issue

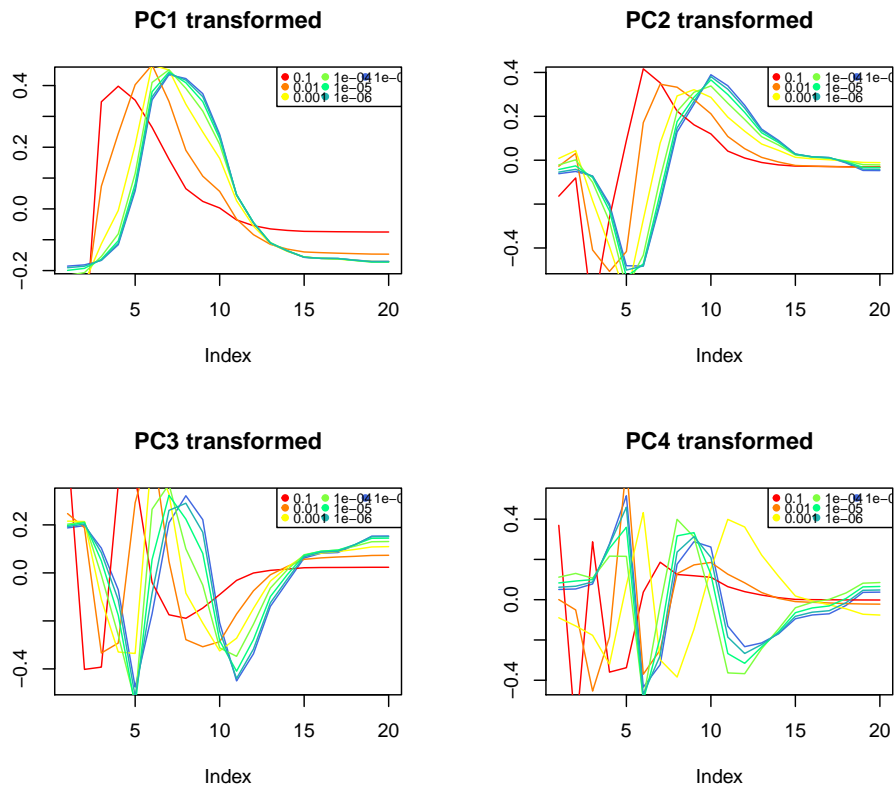


Notes: The plot shows the location of the countries based on the NLPCA analysis on the 10 biggest parties in each country.

Figure 4.18: Objectplot, NLPCA, 10 Parties

of subcompositional coherence (Aitchison 1983).⁸ Aitchison (1986) is aware of the problem that some datasets may have zeros in them that we cannot transform with the log-ratio transformation. He suggests adding a small number to the zeros so that the transformation can be done. In this analysis, first I add 1^{-5} to each zero in the dataset, and then transform the variables with centered log-ratio (clr) transformation. Aitchison (1986, 156) suggests

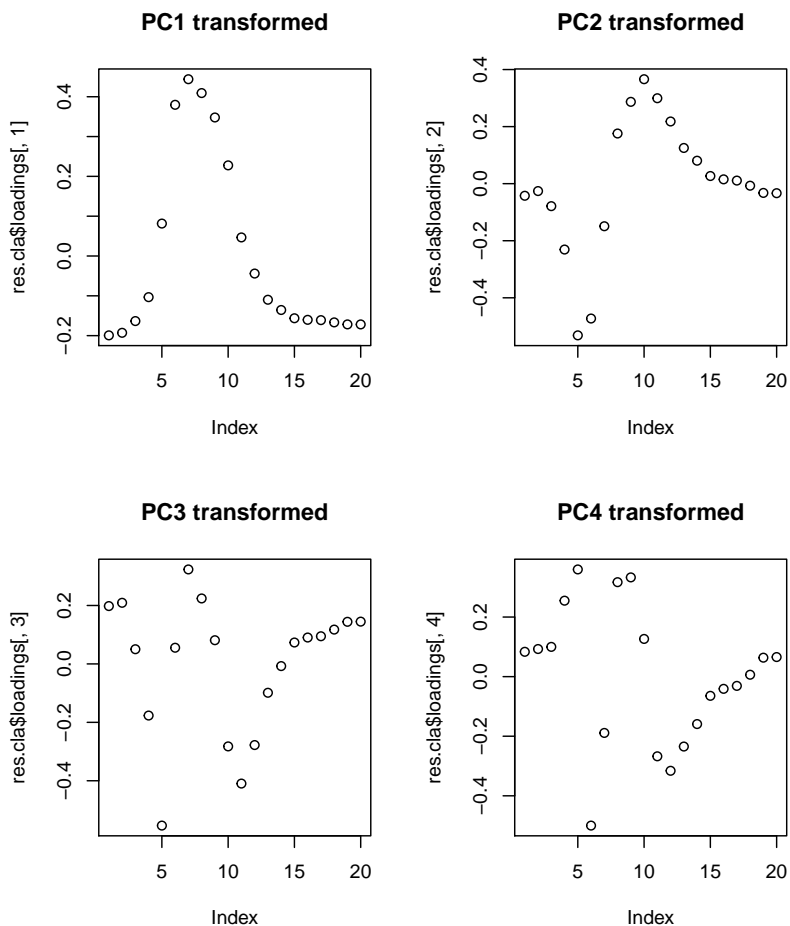
⁸Aitchison suggests the transformation of the data in such a way that the new data is going to be: $v = \log[x/g(x)]$ where $g(x) = (\prod_{i=1}^p x_i)^{\frac{1}{p}}$. This means that we divide each variable with its geometric mean and do a logarithmic transformation the following way: $v_j = \log x_j - \frac{1}{p} \sum_{i=1}^p \log x_i, j = 1, 2, \dots, p$ (Jolliffe 2002).



Notes: The plot shows how the PCA on the log-ratio transformed analysis changes when we add different small numbers to the 0 seat shares of the parties.

Figure 4.19: Sensitivity of the Log-ratio Transformed Variables to the Size of an Added Value to Zeros

that if we add a small number to the zeros, next we should do a sensitivity analysis to check how much this manipulation changes the results of the PCA analysis. In *Figure 4.19* I present the results of this sensitivity analysis. We can see in this plot that if we add a number smaller than 10^{-4} to the zeros we will arrive to a stable solution.

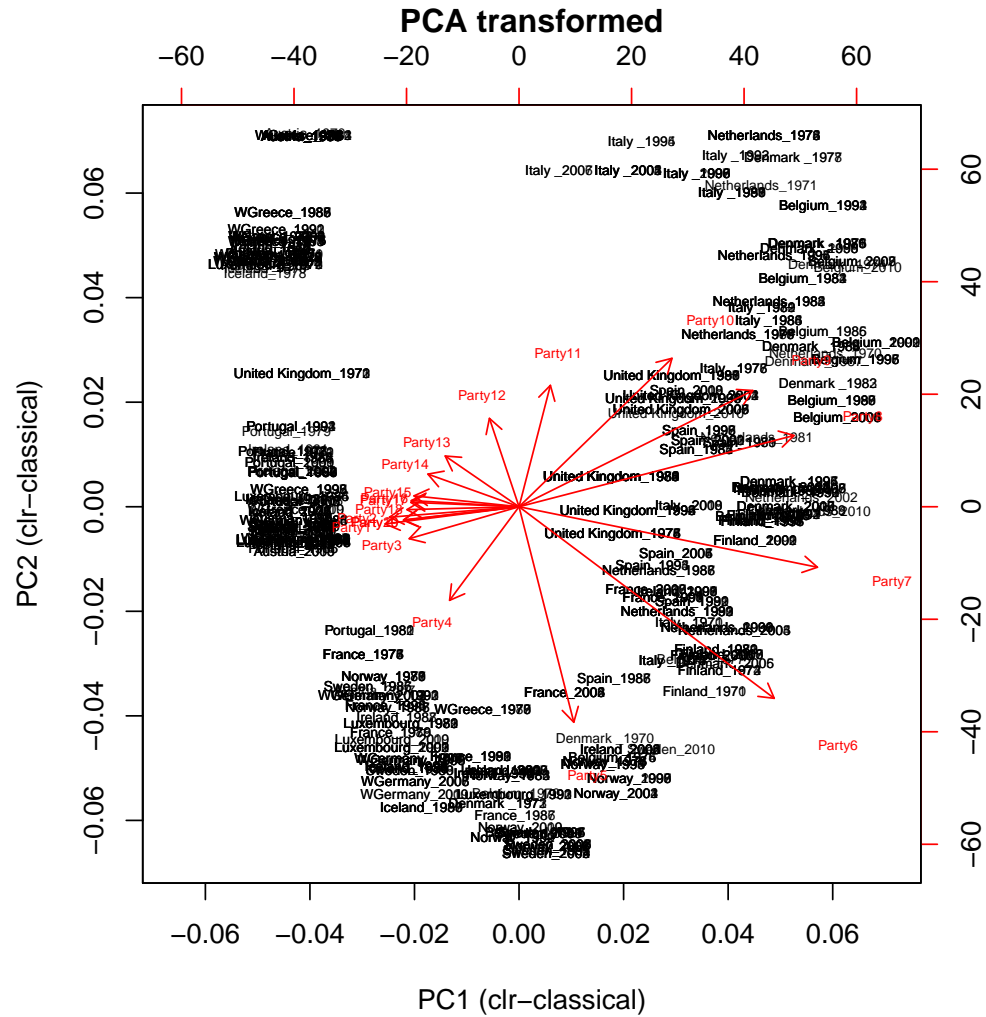


Notes: The parties defining the principal components when the variables are log-ratio transformed.

Figure 4.20: Loadings, PCA on Log- ratio Transformed Variables

Figure 4.20 shows that Dimension 1 of the PCA on the log-ratio transformed variables separates the small party systems (and extremely big party systems) from the bigger ones. *Party 5* to *Party 10* have a big influence on this dimension. Dimension 2 separates the moderately big party systems from the very big party systems: *Party 5* and *Party 10*

have opposite loadings in this dimension.



Notes: Biplot of the PCA analysis when the variables are log-ratio transformed.

Figure 4.21: Biplot, PCA on Log-ratio Transformed Variables

Figure 4.21 shows the biplot of this PCA analysis. An advantage of this method is that the countries are separated in quite clear groups. In Figure 4.21, we can see that the PCA on the transformed variable sorts the countries in groups based on the number of parties in the legislature. Thus, while we can see that the PCA on transformed variables recovered an important feature of the party system (the number of parties) the result is not

very informative. In line with this conclusion, Baxter argues that if there are a lot of zeros and small values in the dataset, performing the PCA on the original dataset is potentially 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).

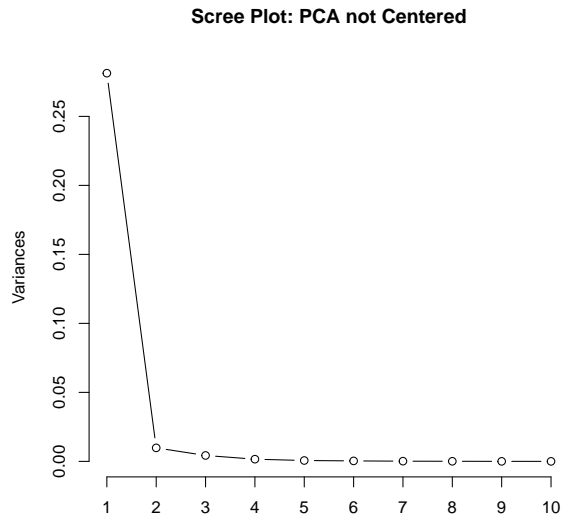
4.3.2 PCA on Non-centered Variables

Another approach to reduce the dimensionality of a compositional dataset is to conduct a PCA analysis on the non-centered variables (ter Braak 1983). The reason follows from the geometrical properties of the compositional dataset. As Aitchison (1983) discusses we can understand each observation in a compositional dataset, as a point on an n -dimensional simplex. This means that each country-year could be represented as a point or vector on a 20 dimensional space, where the 20 coordinates are the seat shares of the 20 parties in the dataset. In case of a compositional dataset thus it is informative to find the space going through the origin of the data as this defines the simplex. This projection can show us the locations of the points on the simplex.⁹

A non-centered PCA does exactly this, it projects the data to the best fitting plane through the true origin and not the center of the data. The data and the direction are projected to this plane (ter Braak 1983). Thus while we get a different projection of the data, this can be useful if we want to find within group variance as opposed to simply between group variance (ter Braak 1983). The result is an ordination plot. On this plot the countries that have unstable party systems will be far from the origin, and countries that have stable party systems will be close to the origin. In addition, country-years that have similar party systems will be grouped together (ter Braak 1983).

The drawback of this technique is that if the observations are a long way away from the origin, the first Principal Component that the analysis finds is the center of the data. This is what we can observe here too, if we conduct this analysis. As *Figure 4.23* shows the

⁹If we denote the seat shares of Party k as s_{ki} in year i so that $\sum_{k=1}^{20} s_{ki} = 1$, each of the individual country-year can be represented as $s_i = (s_{1i}, \dots, s_{20i})$. The space that these points are on is determined by the basis vectors that are length 1 orthogonal in each of the 20 directions.

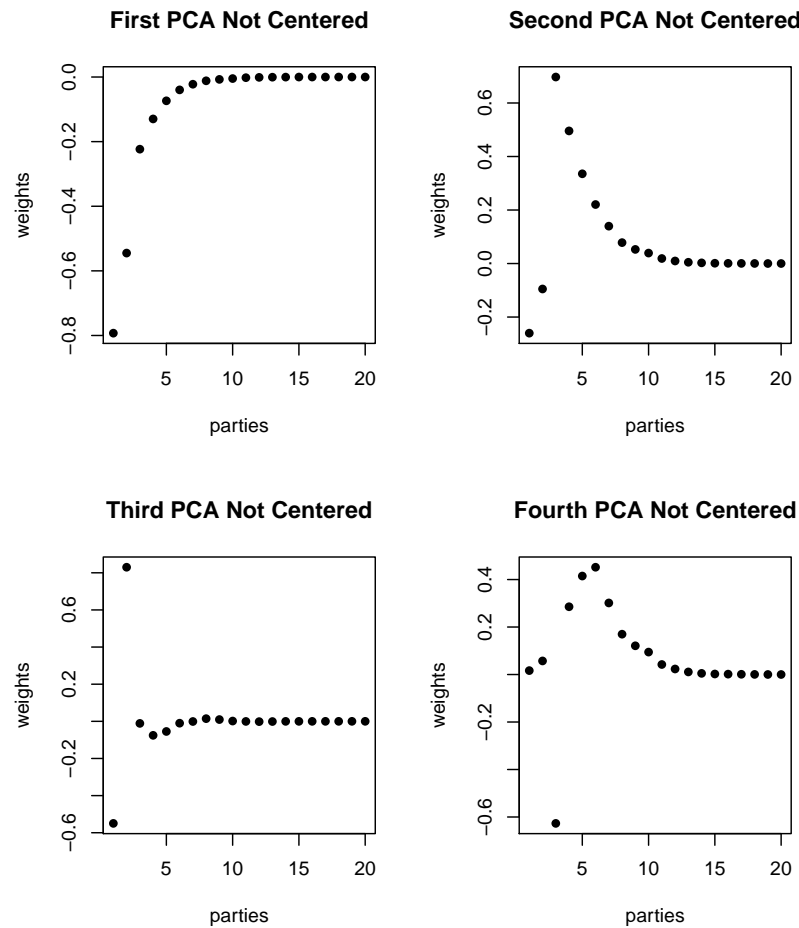


Notes: Screeplot of the PCA analysis on the not centered variables

Figure 4.22: Screeplot, PCA on Non-centered Variables

first two dimensions absorb the influence of variation in the size of the biggest two parties. The two later principal components (here PC3 and PC4) are exactly the same as PC2 and PC3 were in the original PCA.

The biplot (*Figure 4.24*) of this analysis shows that the countries line up mostly based on what extent is their party system concentrated. Indeed, countries that were identified to have smaller party systems (Britain, Germany, France, United Kingdom) are on one end of the dimension, while party systems that are generally considered fragmented (Belgium Finland, the Netherlands) are on the other end. The order is similar to what we have seen on Dimension 1 in the normal PCA analysis. Only three distinct groups arise: in one of these groups we find Greece, Britain and France in certain years, in the second, we can see all the multi-party countries (in the middle of the plot) and in the third group Belgium in the 2000s is its own category.

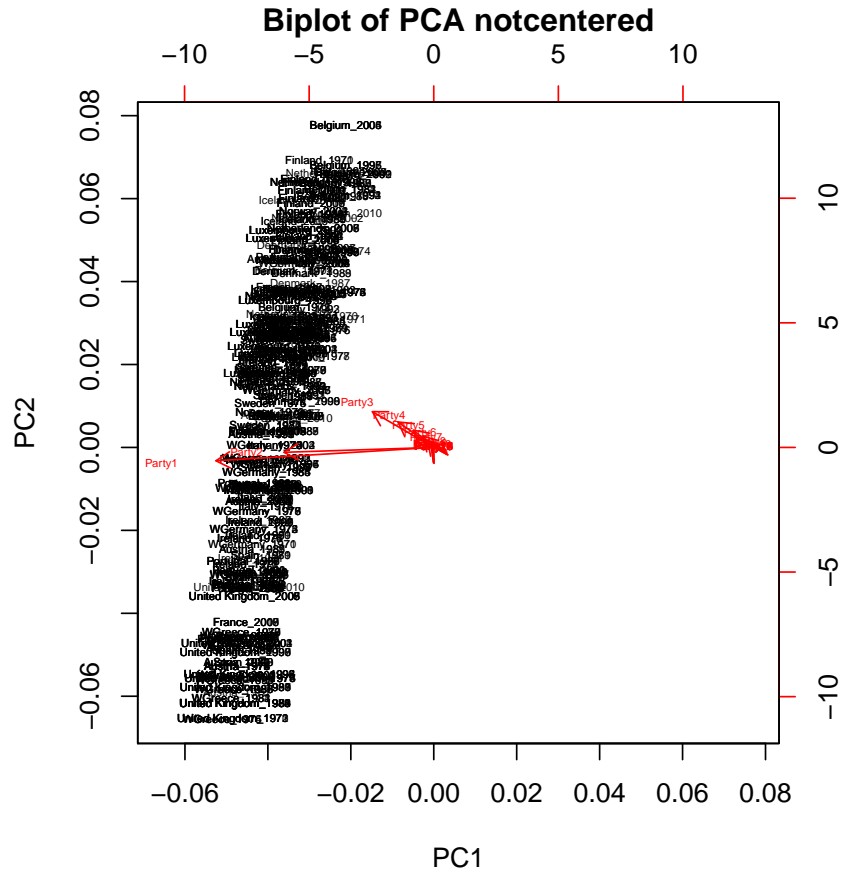


Notes: The parties defining the principal components when the variables are not centered.

Figure 4.23: Loadings, PCA on Non-centered Variables

4.4 Summary

In this paper I analyzed a dataset that contains the legislative seat shares of parties in 17 European countries from 1970-2013. The dataset is ordered so that the variables represent parties with decreasing sizes. As discussed above, some political science literature classified party systems with qualitative methods, while other studies created summary measures to characterize the size and shape of the party systems. In this paper, I explored the underlying structure of the data, so that we could better understand what these previously created indices measure.



Notes: Biplot of the PCA analysis on the Non-centered Variables.

Figure 4.24: Biplot, PCA on Non-centered Variables

In this paper I explored the dataset with a principal component analysis (PCA). Next, I relaxed the linearity assumption of the PCA analysis and I ran a kPCA and a NLPCA analysis. Overall, all solutions showed that the variations of the two biggest parties were the most important features of the dataset. I found with the PCA analysis on the unscaled variables, that the first two principal components were related to the absolute and the relative sizes of the two biggest parties. In addition, the PCA analysis on the non-centered variables separated countries with small and with large party systems. The kPCA analysis showed that the untransformed data is close to a Gaussian distributed, so the kPCA solution was very similar to the simple PCA solution.

I argued that because the variation in the sizes of the two biggest parties were the biggest, these might mask important features in the lower dimensions of the data. To understand the deeper structure of the data, first I controlled for the sizes of the biggest and the second biggest parties, and I conducted a PCA analysis on the residuals. This analysis showed that the lower dimensions were also defined by the absolute sizes of the parties, and their relative sizes compared to each other. The party systems were separated into two groups: countries in which the parties were equal in size (competitive party systems) and countries in which some parties were dominant (less competitive party systems).

Another way to decrease the influence of the variables with high variances is to scale them. Scaling means that we set the variance of the variables to unit, thus we equalize their impact on the analysis. This transformation, however, may not be the most optimal one. Potentially, a better way to reduce the variation in the data, is to use Non-Linear Principal Component Analysis. The NLPCA methodology created optimal quantifications of the variables at the same time as it optimized the PCA loss function. Thus, the variance of the data matrix was reduced to the minimum. With the NLPCA analysis I found that when the variables were optimized this way, the most important feature that the PCA extracted was the raw number of parties. Because of this countries that had very big party systems in certain times became their own category (like Italy). The NLPCA analysis that I conducted on the 10 biggest parties in each party system had a solution similar to the PCA and kPCA solutions. This makes sense as PCA and kPCA solutions were not influenced by the variation of the smallest parties as much.

The sensitivity analysis of the NLPCA analysis showed that the data did not have subcompositional coherence. This was because the dataset I was using was a compositional dataset: the variables were proportions that added up to one. To remedy the bias that this might have caused, I ran one PCA analysis on the log-ratio transformed variables and another one on the non-centered variables as well. While the results were clear on the log-ratio transformed variables (the countries with the same number of parties were grouped together) it was less informative on the non-centered variables (it mostly separated country- years in which countries had concentrated party systems from country-years in which countries had

fragmented party systems). Overall, it seems that in the case of the party system size data, the original PCA analysis is the most informative.

CHAPTER 5

How do the PCA Dimensions Relate to Traditional Typologies and Measures

5.1 Comparing the Results of the PCA Analysis to Traditional Typologies

In this section I compare the results of the PCA analysis to the traditional typologies. As we have seen in *Table 2.1*, the traditional typologies separate party systems by the number of the parties and by the competition within the party system. Below, I plot the position of each country in the lower dimensional plane that the PCA analysis have found. In addition, I plot the position of all the countries, every five years on the same plane. With these plots we can examine how the party systems have changed within and across countries as well. Previously, I identified the two PCA dimensions as PC1: *Size of the Biggest two Parties*, and PC2: *Competition between the Biggest two Parties*.

As we can see from these plots, in some countries the party system is relatively stable. These are countries such as Iceland (*Figure 6.8*) or Finland (*Figure 6.4*). Other countries stay the same place on Dimension 1, like the United Kingdom (*Figure 6.17*), Luxembourg *Figure (6.11)* and Sweden (*Figure 6.16*). There are countries that stay in the same place on Dimension 2 but move on Dimension 1 like Austria (*Figure 6.1*), the Netherlands (*Figure 6.12*), Belgium (*Figure 6.2*) and Italy (*Figure 6.10*). Also there are countries which move on both directions, notably France (*Figure 6.5*), Portugal: (*Figure 6.14*) and Norway (*Figure: 6.13*).¹

¹Some of these changes are the results of institutional changes in these countries. For instance, in Belgium the national parties split into Wallonian and the Flemish regional parties in the 1960s and 1970s.

It is probably more fair to evaluate the previous party system typologies based on how the party systems looked like in the years the studies were written. I have yearly data but here I only include plots for every five years, since the party system does not change very rapidly. The two most relevant years for this comparison are 1970 (*Figure 6.18*) and 1975 (*Figure 6.19*). As I discussed above, some of the first party system typologies separated two-party systems from multi-party systems. In the data, there are no pure two-party systems (in the legislatures there are at least three parties), however, in some countries the first two parties are much bigger than the rest of the parties. All typologies create different categories for two-party and for multi-party countries (Duverger 1954; Blondel 1968; Rokkan 1970; Sartori 1976; Mair 2002).

The PCA analysis recovers the same difference between the (close to) two-party systems and the multi-party systems (with PC1). The countries that are two-party systems are on one side of this dimension: The United Kingdom is always close to this two-party ideal (*Figure 6.17*) and Austria in the 1970s indeed seems to be close to this ideal as well. Later Spain, Greece and Portugal joined this group (*Figure 6.15*, *Figure 6.7*, *Figure 6.14*). On the other side of Dimension 1, we find the fragmented party systems without dominant parties: the Netherlands, Belgium, and the Scandinavian countries. Overall if we project the countries down to this Dimension 1, we can see that the order is similar to the categorization

The Flemish/Wallonian pairs of Christian Democratic Parties split in 1968, Socialist Parties split in 1971, and Liberal Parties split in 1978 in Belgium.

In Italy the electoral system changed from PR to a mixed-member electoral system in 1994. The mixed-member electoral system is often associated with fewer parties, however in Italy's case many small parties formed and then competed in two big coalitions. Under the mixed-member system the electoral competition consists of two separate competitions at once. The people vote for parties and at the same time with a separate ballot they vote for a district candidate. Parties get legislative seat shares based on their vote shares, while individuals get seats if they win electoral districts. While small parties can survive in the proportional tier, in the Single Member District (SMD) tier usually the two biggest competitors remain.

The electoral system did not change drastically in Portugal, however it changed incrementally after the country became democratic. In *Figure 6.14* we can see that the party system at first is fragmented, and it becomes more competitive and less fragmented throughout the years. The small parties gradually form coalitions, first, to compete in the elections, however these coalitions merge. Norway *Figure 6.13*, moves to the other direction, while between 1970 and 1990 the country had a pretty stable and closed party system, in the 2000s new parties enter into the electoral competition and the non-left parties become stronger. In the plots we can also see why Finland and France was hard to categorize by the scholars: *Figure 6.5* shows that the French electoral system changes the most throughout the years. In contrast, Finland remains pretty much at the same place- on the very edge of Dimension 1 (at least in the early years) but it is hard to either classify it as a competitive or an uncompetitive system based on its position on Dimension 2 (*Figure 6.4*).

of Blondel (1968).²

Rokkan (1970)'s classification is more complex, and it is similar to how both Dimension 1 and Dimension 2 separate the countries. Rokkan's first category (the British-German 1vs 1+1 system) engulfs the countries that are separated from the rest of the countries on Dimension 1, while his 2, 3a, 3b categories include the countries that are separated from each other by Dimension 2 *Figure 6.18*. Rokkan's categorization is based on the numbers of parties and the evenness of the competition. This gives qualitative support to the notion that Dimension 2 separates countries based on the competition between the two biggest parties. Sartori's (1976) classification, similarly to Rokkan's, shows that Dimension 2 captures the competitiveness in the party system.³

While the qualitative typologies were useful to draw attention to that "competitiveness" can be considered a separate dimension from the party system size, the countries change their position more frequently on Dimension 2, so the classification of the countries in distinctive groups may be difficult. On Dimension 1 changes are slow,⁴ on Dimension 2, however the changes are more rapid.⁵

Next, I discuss how the PCA dimensions relate to the party system size indices that I introduced before. After that, I show some indices that may be useful to measure Dimension 2, the competitiveness dimension.

²The order of in which he classifies the countries is: England, Germany, Ireland, Denmark, Sweden, Norway, Iceland, Italy, Netherlands, and Finland, is close to the order of the countries that the PCA analysis finds on Dimension 1. The notable exception is France (and the country's position is moving around throughout the years) and Denmark, which seems to be closer to the Netherlands and Finland than to the rest of the Scandinavian countries.

³Sartori's classification seems to reflect the 1970s as opposed to 1975. In 1970 (*Figure 6.18*) according to the PCA analysis Norway and Sweden could indeed constitute their own group. This is less so by 1975 (*Figure 6.19*). While in 1970, Finland and Italy (which are categorized as polarized pluralisms) are on the same level on PC2 and a bit below the other countries, Denmark should potentially be in the same category.

⁴Although these happen too. *Figure 6.26* shows that by 2010 the party systems of Spain, the UK Portugal, Italy, Ireland, Greece and France all became more concentrated. On the other hand, Austria and Germany have a more fragmented party system than they had before.

⁵Just by comparing 2005 (*Figure 6.25*) and 2010 (*Figure 6.26*) we can see that Sweden moves from one end the other.

5.2 Comparing the results of the PCA analysis to Party System Size Measures

In this section, I compare the PCA results to the indices that previous scholarship has identified as useful measures of the party system size. I calculated the measures: *the Effective Number of Parties in the Legislature (ENP)*, *the Biggest Party in the Legislature (BigP)*, *the Raw Number of Parties in the Government (GPs)*, *the Shapley ENP*⁶

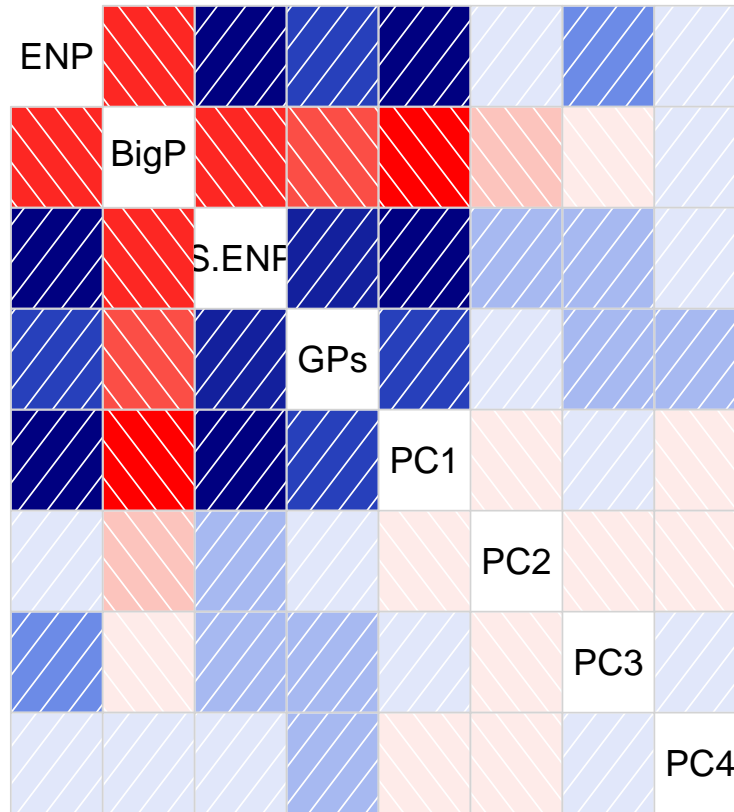
I represent the correlations between these measures and the traditional measures on a correlogram in *Figure 5.1*. A correlogram is a graphical representation of the correlations between the variables in a given dataset. Along the main diagonal of the big square we can see the variables. The small squares on the two side of the diagonals show the direction of the correlations. Upward slopes mean that there is a positive correlation between two variables, while downward slopes indicate a negative correlation between the two variables. The darkness of the shading indicates the strength of the correlations. We can see from the correlogram that all of these measures are highly correlated with each other and with Dimension 1 (PC1) which is the dimension defined by the size of the biggest and the second biggest parties. As we can see through this method we can implicitly compare the typologies with the party system size measures. And we can see that the party system size indices relate closely to Dimension 1, and thus to the typologies of Duverger (1954); Blondel (1968).

5.3 Measures of Opposition Concentration

As we can see most of the party system size indices measure only one aspect of the party system: the size of the biggest two parties (PC1). However, the concentration of the opposi-

⁶This is an ENP like measure, in which I replace the parties' seat shares with their Shapley-Shubik indices (S.ENP). I calculated a composite measure of the Effective Number of Parties including the Shapley-Shubik power index. Grofman and Kline (2011) have used the Banzhaf index in their calculations. I am using the Shapley-Shubik index as opposed to the Banzhaf index because I examine how coalitions are formed, thus the sequential approach to the coalition formation seems to be more appropriate. In addition, mathematically, the Banzhaf index puts additional weight on the biggest party. As I am interested in the opposition structure, I chose a measure that evaluates the opposition power more precisely.

Correlogram with PCA



Notes: PC1: “Size of the two biggest parties”, PC2: “Competition”, PC3: “3rd party”, PC4: “Multipartism”, ENP: Effective Number of Parties, BigP: Size of the Biggest Party over the size of the Legislature, S.ENP: Effective Number of Parties (Shapley), GPs: Parties in Government.

Figure 5.1: Correlogram, Traditional Measures of Party System Size and PCA Principal Components

tion could be measured similarly to the party system size. This notion has been discussed in a few studies previously (Maeda 2010, 2015). However, very few studies in political science have used these measures. In this part I examine whether these indices measure the same dimension of the party systems as the indices discussed above.

The *Effective Number of Opposition Parties ENOP* is calculated the same way as the *Effective Number of Parties ENP* suggested by Laakso and Taagepera (1979), except I calculate the measure only for the opposition parties.⁷ *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.⁸ Finally the *Size of the Biggest Opposition Party (BOPP)* tellingly measures the size of the biggest party in the opposition.⁹

In *Figure 5.2* and *Figure 5.3* we can see that the opposition measures relate more closely to *PC2* while the party system size measures are more closely related to *PC1*. The only exception is the size of the Biggest Party (*BigP*) 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.

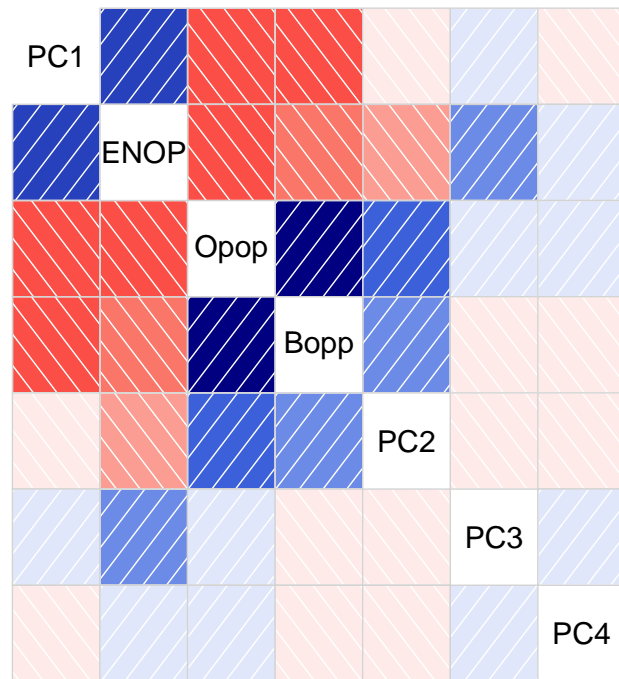
Most political science studies use one of the party system size measures when they

⁷This measure 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. The variable ENOP is utilized by Maeda (2010) and Maeda (2015) although contrary to my calculation Maeda does not include changes in the government between two elections. In addition, Falcó-Gimeno and Jurado (2011) uses the Herfindahl-Hirschman index to measure the opposition concentration which is a reciprocal transformation of ENOP as discussed above.

⁸This index is: $OPOP = \frac{S_{Opp_1} - S_{Opp_2}}{n}$ 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.

⁹Several works argue that if we want to measure the power structure of the entire party system we can look at the size of the biggest party (Taagepera 1999; Dunleavy and Boucek 2003). I calculate this measure as the seat share in the legislature of the biggest opposition party. Mathematically $BOPP = \frac{S_{Opp_1}}{n}$ 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.

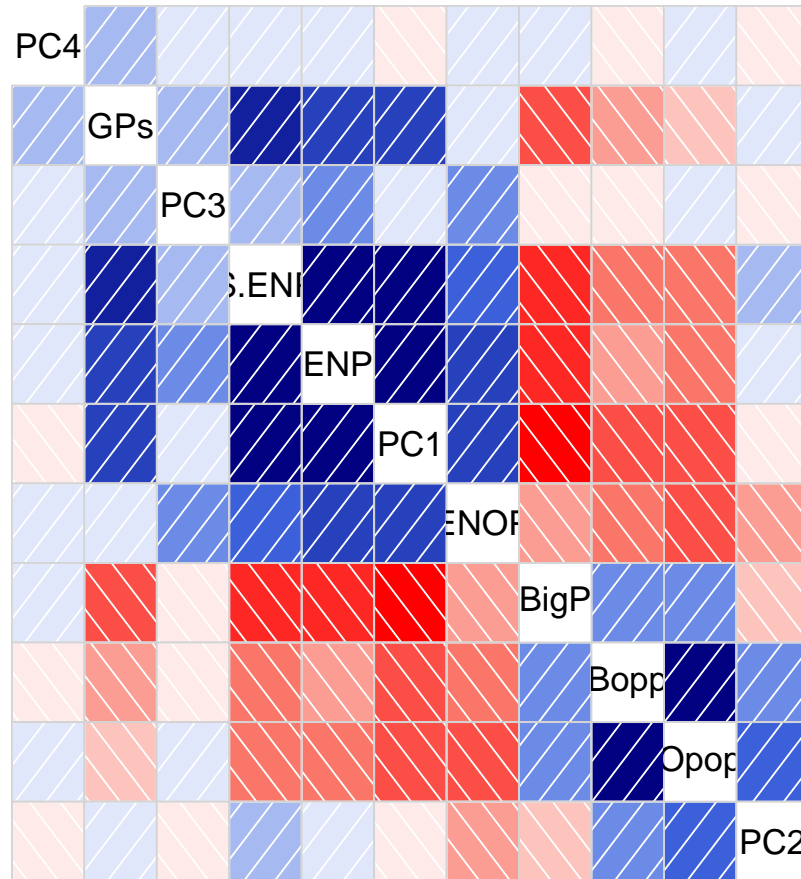
Correlogram with PCA



Notes: PC1: “Size of the two biggest parties”, PC2: “Competition”, PC3: “3rd party”, PC4: “Multipartism”. *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.

Figure 5.2: Correlogram, Opposition Structure and PCA Principal Components

Correlogram with PCA



Notes: PC1: “Size of the two biggest parties”, PC2: “Competition”, PC3: 3rd party, PC4: Multipartism. *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. *ENP*: Effective Number of Parties, *BigP*: Size of the biggest party over the size of the legislature, *S.ENP*: Effective Number of Parties(Shapley), *GPs*: Parties in Government.

Figure 5.3: Correlogram, Traditional Measures of Party System Size and Opposition Structure and PCA Principal Components

want to measure the competitiveness of the party system. However, the analysis in this section shows that these measures capture the size of the biggest two parties, and not necessarily show how competitive the party system is. While most traditional party system size indices measure the size of the party system (PC1), the opposition structure measures relate to a different dimension: competitiveness of the party system (PC2). These measures so far have been infrequently utilized by the quantitative analyses. I suggest that when a political science theory relates to the competitiveness of the party system, we should consider using one of the opposition concentration indices instead of the party system size measures.

CHAPTER 6

Conclusion

The party system is an important part of the political system in any country. The way how the party system has evolved and how parties interact with each other are almost always discussed by country studies. However, since there is a considerable variation in party systems across the countries, and the party systems evolve within countries, political scientists have had debates on how accurately measure the size of the party systems. In this paper, I introduced two approaches that political scientists took to create meaningful comparisons. One group of scholars sorted the countries into categories based on the characteristics of the party systems. I argued that most of the party system typologies divided the countries based on the number of parties, and the competition between the parties. Another group of scholars created summary indices that characterized the party systems. I argued that these scholars faced a classical dimensionality reduction dilemma: they tried to represent the most information possible with a single measure. The debate was about which information was important to keep, and which could be discarded: political scientists weighted either the bigger parties or smaller parties more heavily in their calculations.

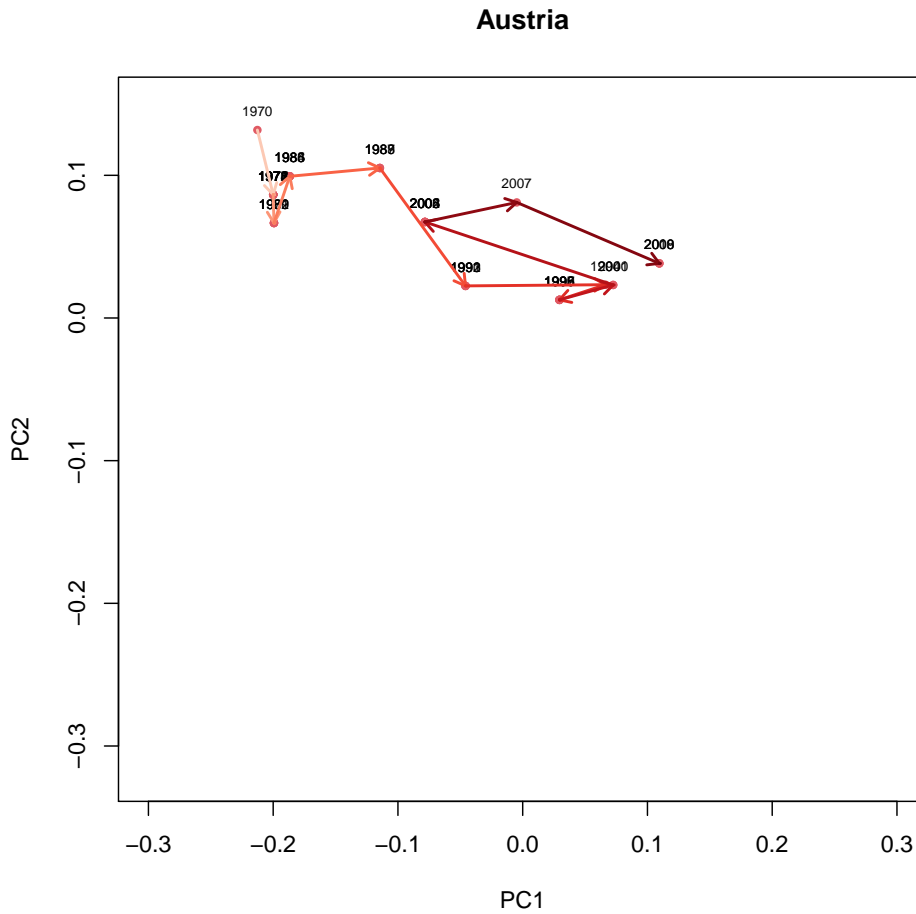
In this paper, I compared these two approaches to the underlying structure of the party system data with yet another dimensionality reduction technique. I conducted a Principal Component Analysis on data on party systems from 17 countries from 1970-2013. In addition, to verify the results, I allowed for non-linearity of the data with a kernel Principal Component Analysis, and a Non-Linear Principal Component Analysis. I also examined whether the analysis was influenced by the structure of the data (as the data was compositional) and ran a Principal Component Analysis on the log-ratio transformed and the non-centered variables. The PCA analysis showed that the absolute sizes of the biggest

two parties, the relative sizes of the biggest two parties (the competition between the two biggest parties), the size of the third party, and the relative sizes of parties 5-6 (the balance in the party system) were the important features separating the countries from each other. I also found that the two most important dimensions that separate countries according to the PCA analysis were similar to the dimensions that the typologies identified. One of these was the number of parties within the party system (Dimension 1), and the other one was the competitiveness of the two biggest parties (Dimension 2).

In the last part of the paper, I compared the results of the PCA analysis to the party system size measures. I showed that the traditional party system size indices were correlated mostly with Dimension 1, most of these indices show whether the sizes of the biggest two parties were big or not relative to the other parties. However, none of these indices were correlated with the other PCA dimensions. Then, I discussed some indices that measure the opposition concentration, and I suggested that these indices could measure the competitiveness of the party systems. I showed that indeed, the opposition size measures relate partially to Dimension 2 that the PCA recovered.

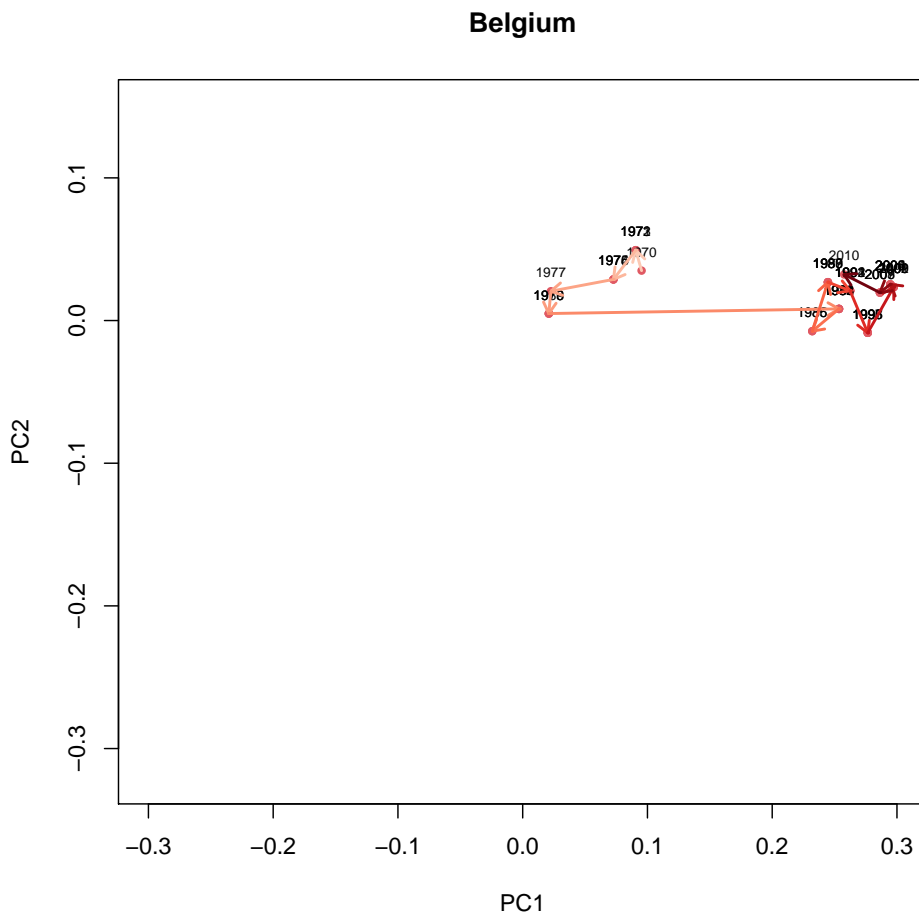
Overall, this paper shows that we should consider the structure of the dataset more carefully when we decide how to operationalize our key variables. As I discussed, for a long time researchers in political science did not find any evidence that the party system influences political outcomes. However, most of these studies use only one measure of party system fragmentation: the ENP. Even though qualitative studies have noticed that the competitiveness of the party systems matter, until recently there have been no attempts to quantify competitiveness. For this reason, in a lot of studies that required some measure of competitiveness, political scientists in reality controlled for the size of the party system and not its competitiveness. As a consequence, maybe it is not surprising that the quantitative studies have found little evidence supporting that the party system influences political outcomes. The results in this paper show that the size and the competitiveness of the party system are two different features of the party systems. I suggest that studies that require measures of competitiveness should consider using one of the opposition concentration indices.

APPENDIX: COUNTRY AND YEAR FIGURES



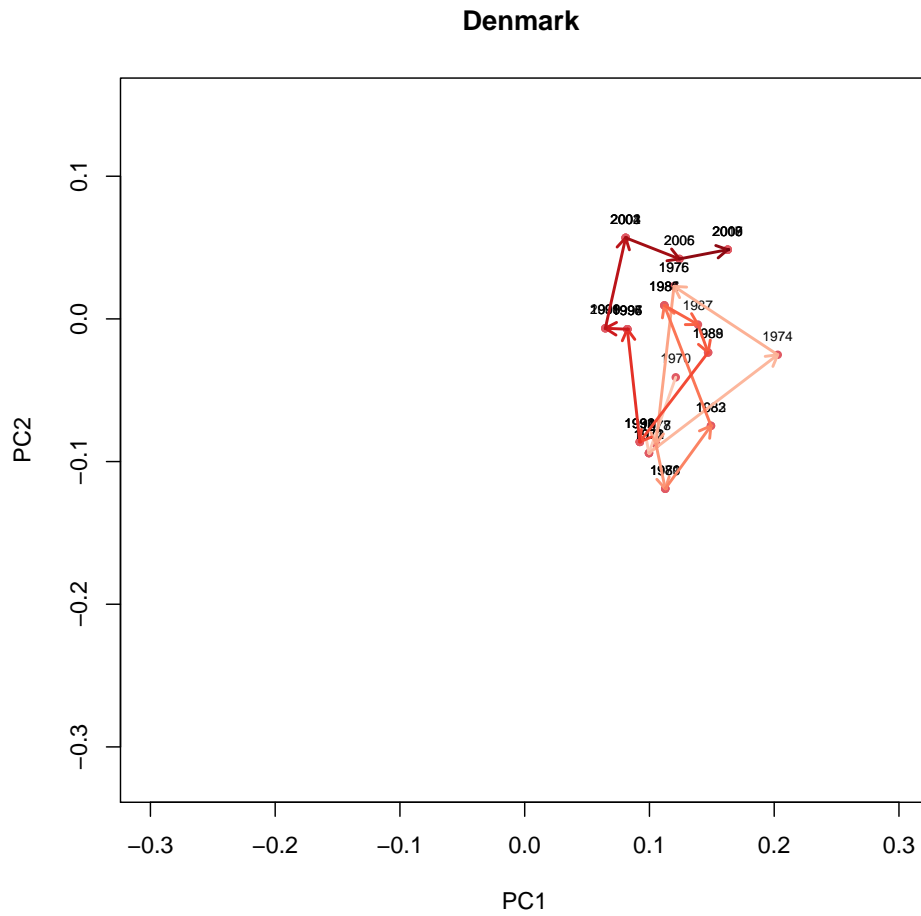
Notes: The plot shows the Austrian party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

Figure 6.1: Austria on the PCA Dimensions



Notes: The plot shows the Belgian party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

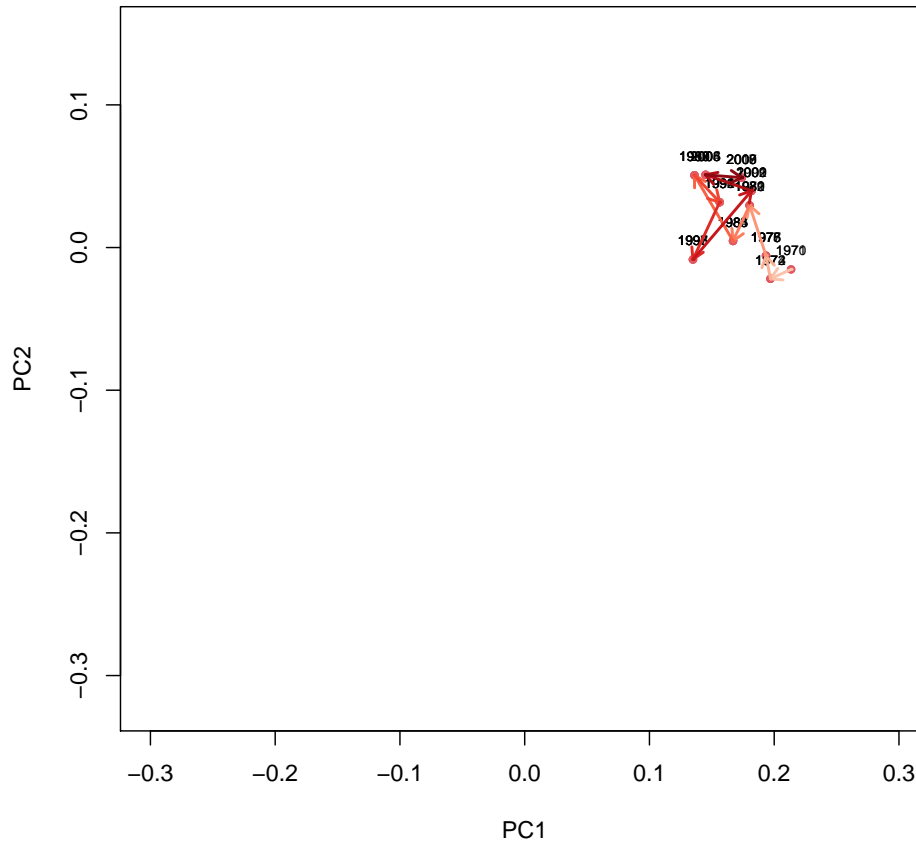
Figure 6.2: Belgium on the PCA Dimensions



Notes: The plot shows the Danish party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

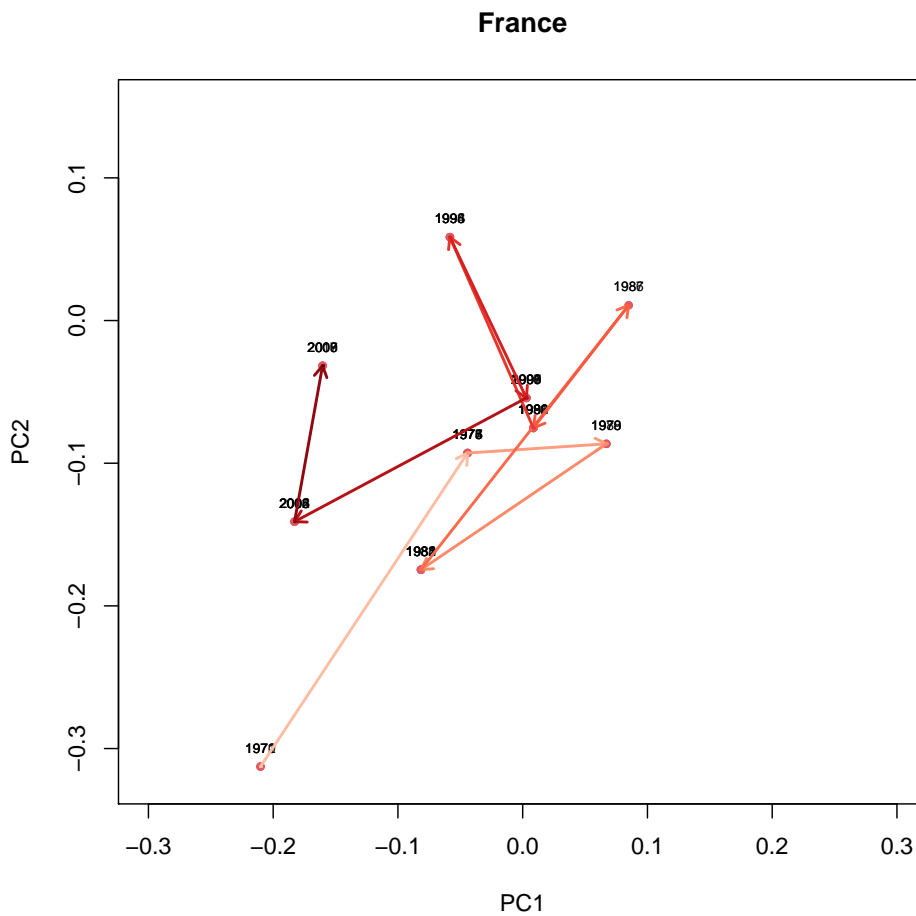
Figure 6.3: Denmark on the PCA Dimensions

Finland



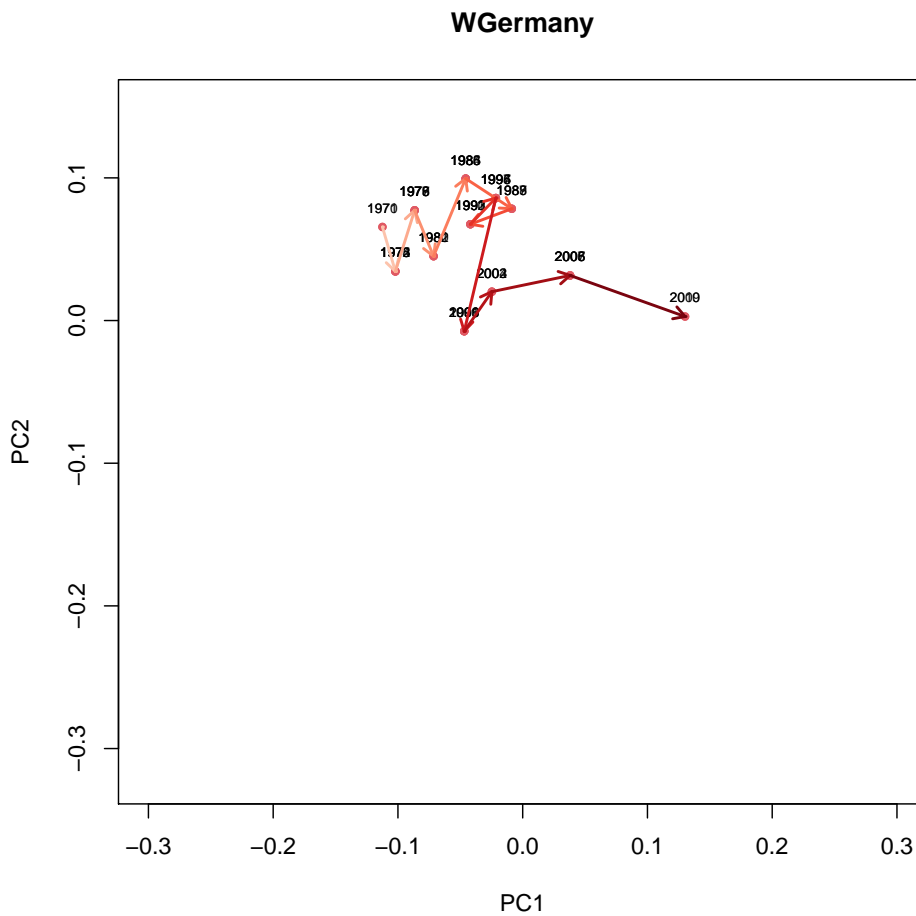
Notes: The plot shows the Finnish party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

Figure 6.4: Finland on the PCA Dimensions



Notes: The plot shows the French party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

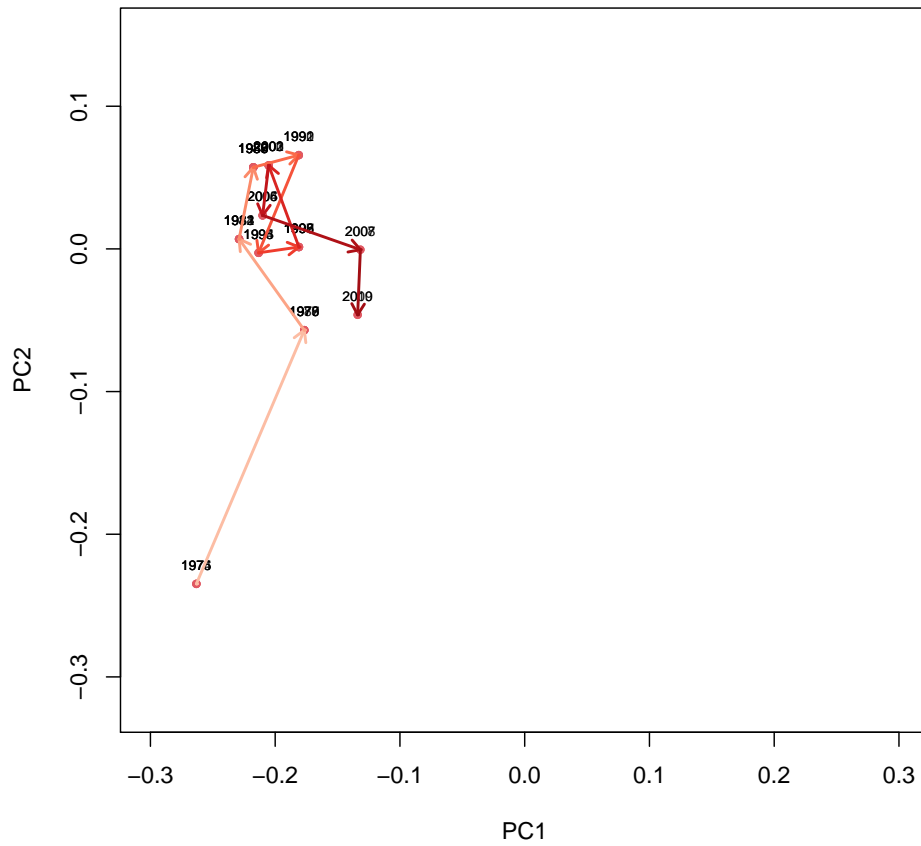
Figure 6.5: France on the PCA Dimensions



Notes: The plot shows the German party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

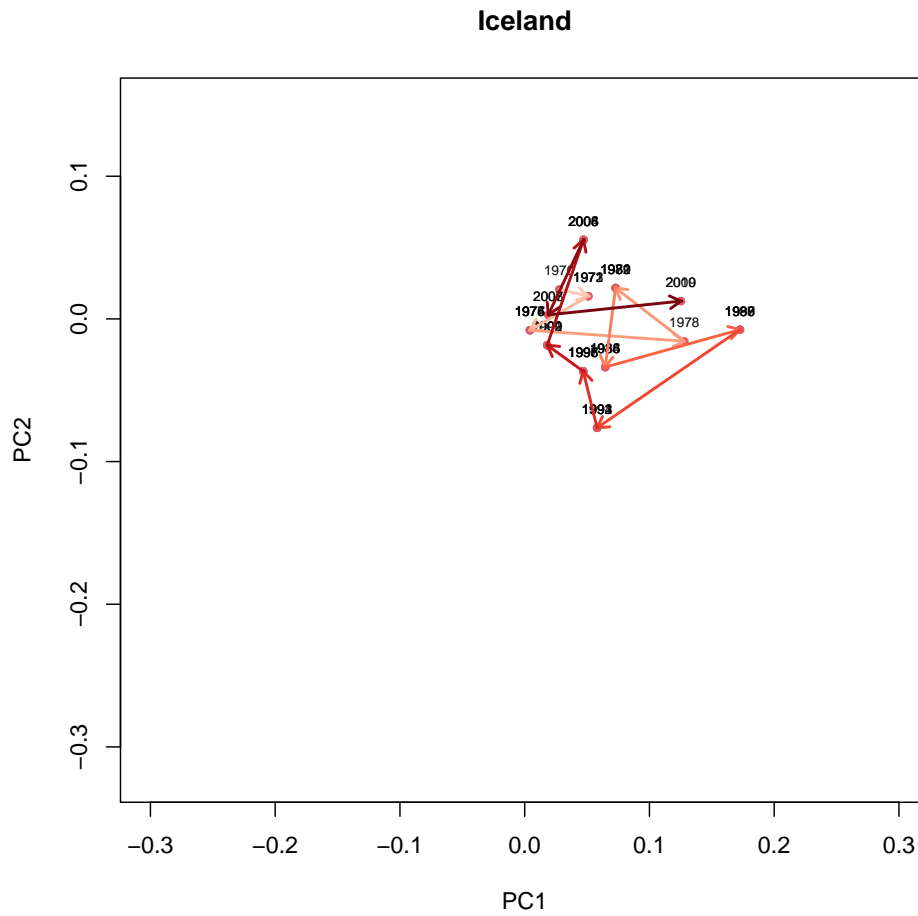
Figure 6.6: Germany on the PCA Dimensions

WGreece



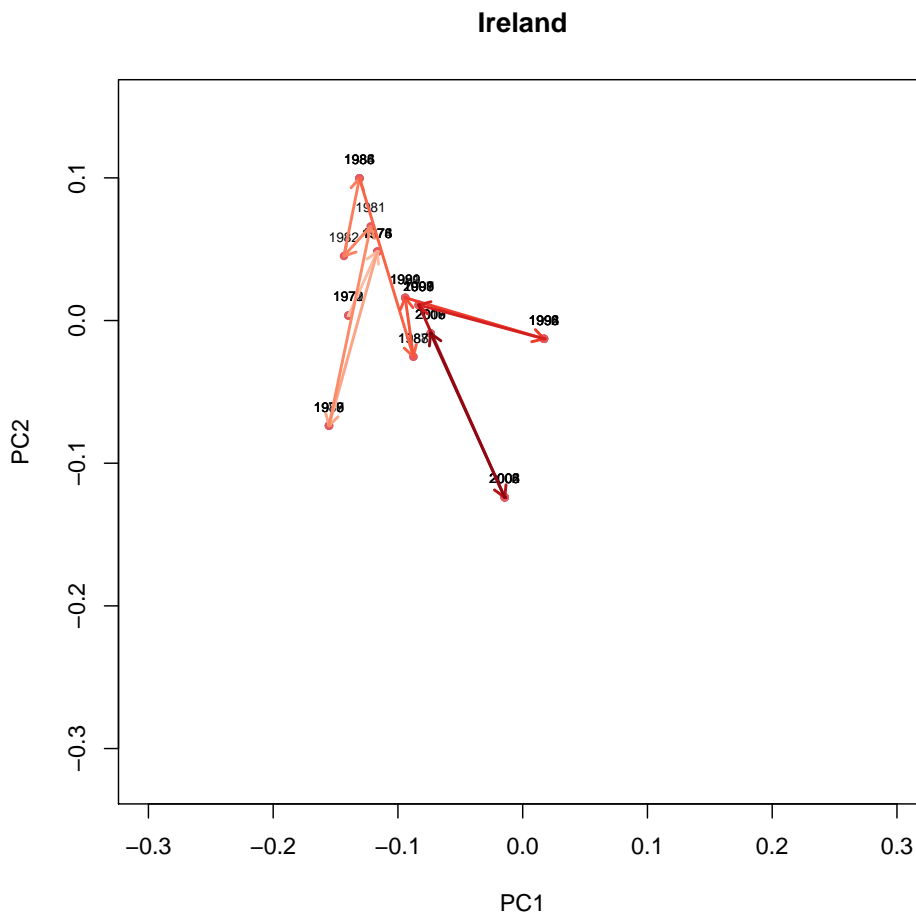
Notes: The plot shows the Greek party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

Figure 6.7: Greece on the PCA Dimensions



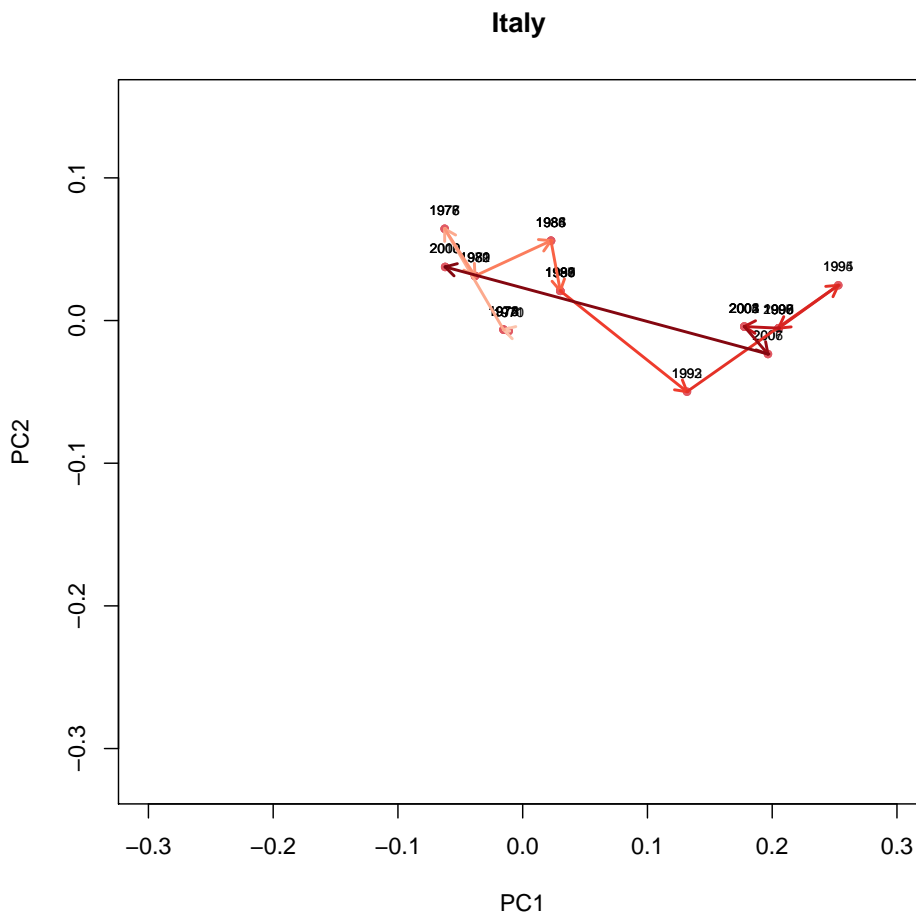
Notes: The plot shows the Icelandic party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

Figure 6.8: Iceland on the PCA Dimensions



Notes: The plot shows the Irish party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

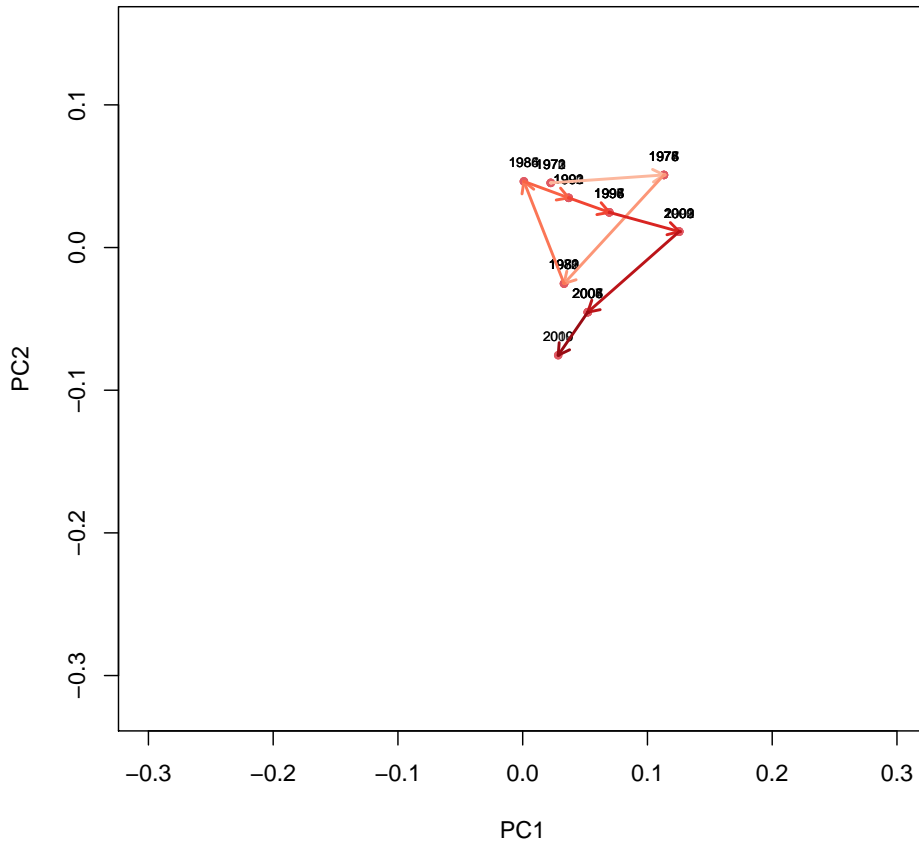
Figure 6.9: Ireland on the PCA Dimensions



Notes: The plot shows the Italian party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

Figure 6.10: Italy on the PCA Dimensions

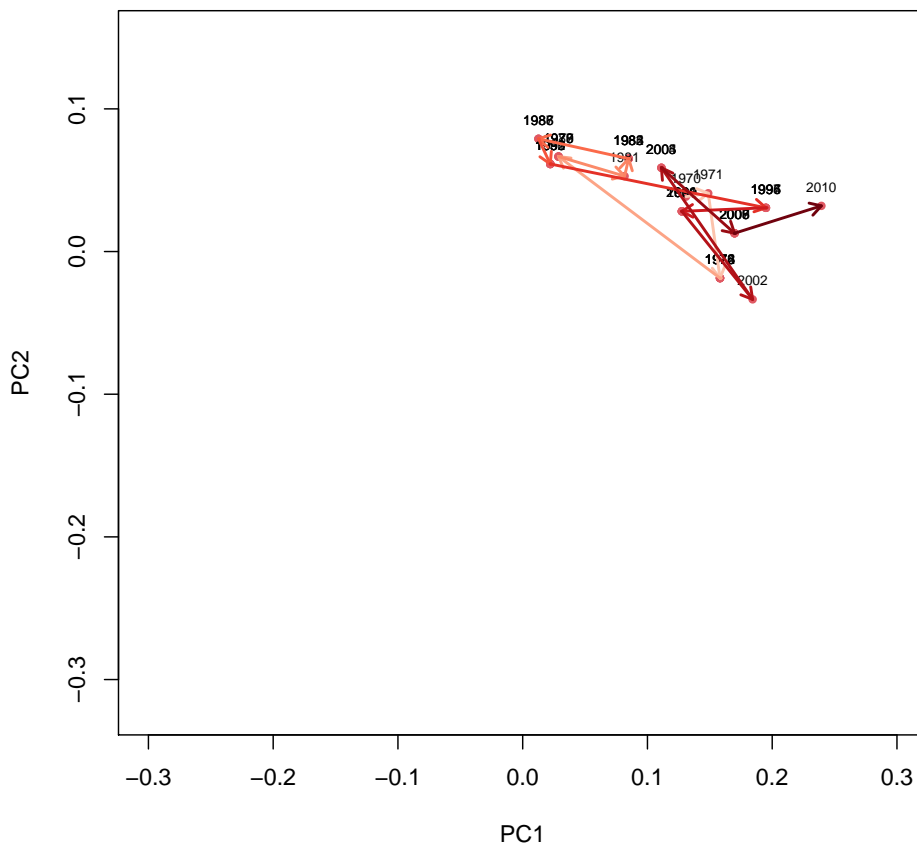
Luxembourg



Notes: The plot shows the Luxembourgish party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

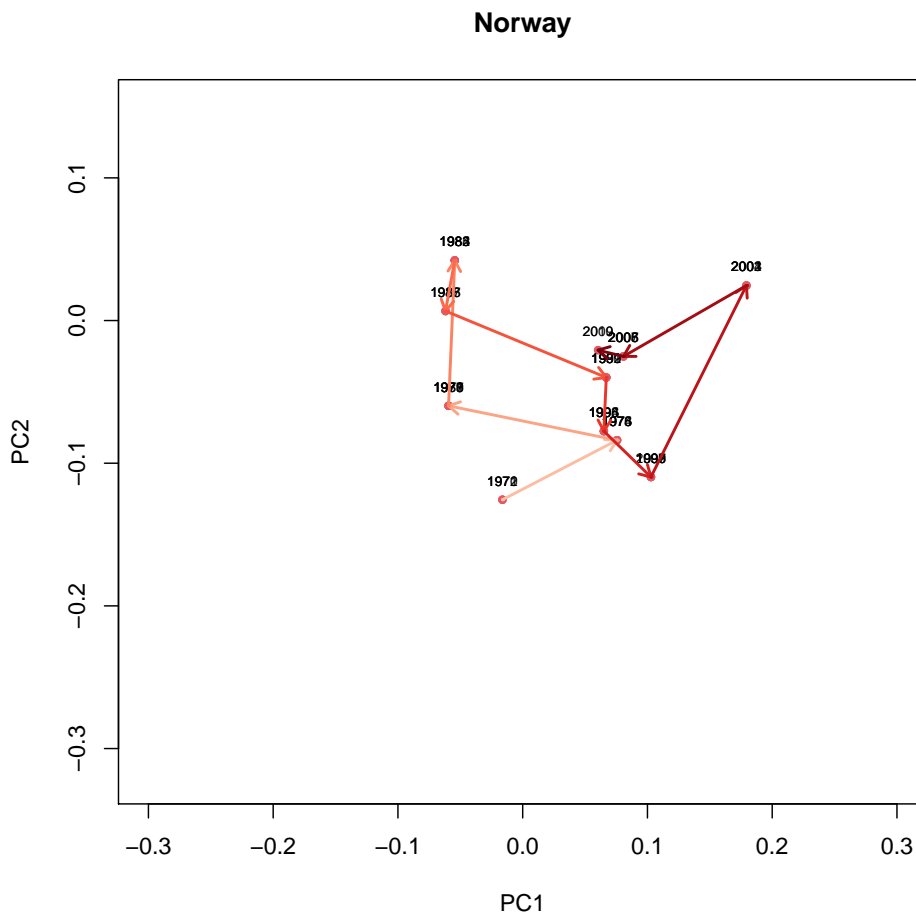
Figure 6.11: Luxembourg on the PCA Dimensions

Netherlands



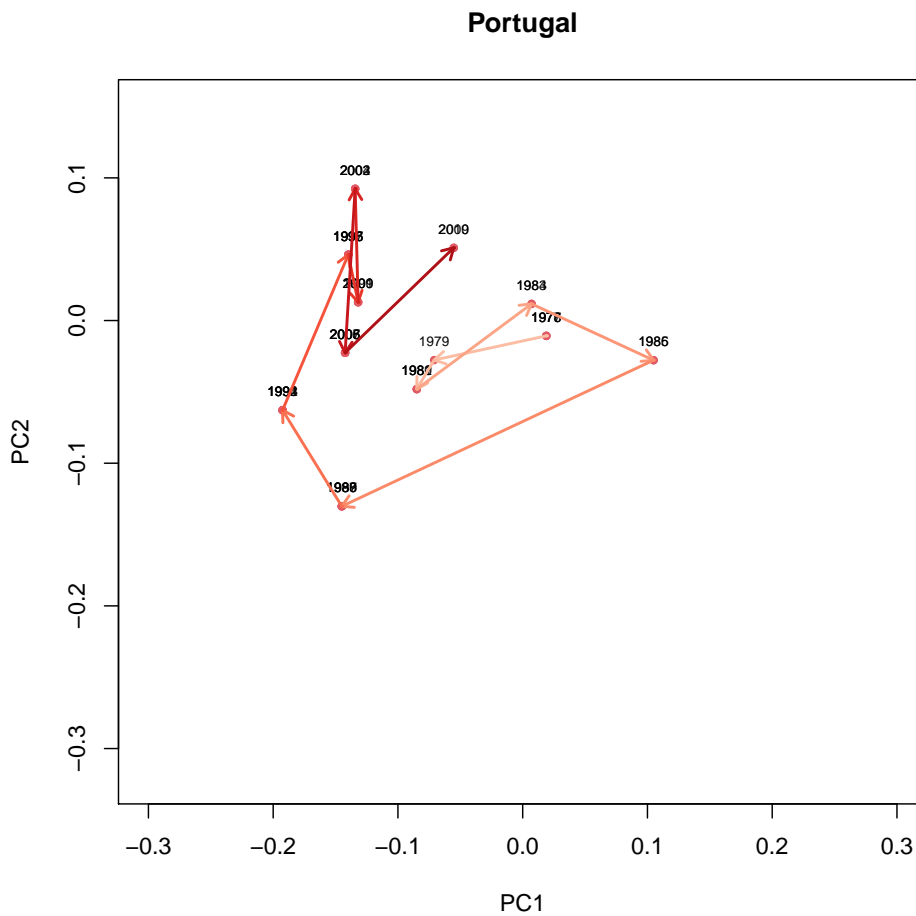
Notes: The plot shows the Dutch party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

Figure 6.12: the Netherlands on the PCA Dimensions



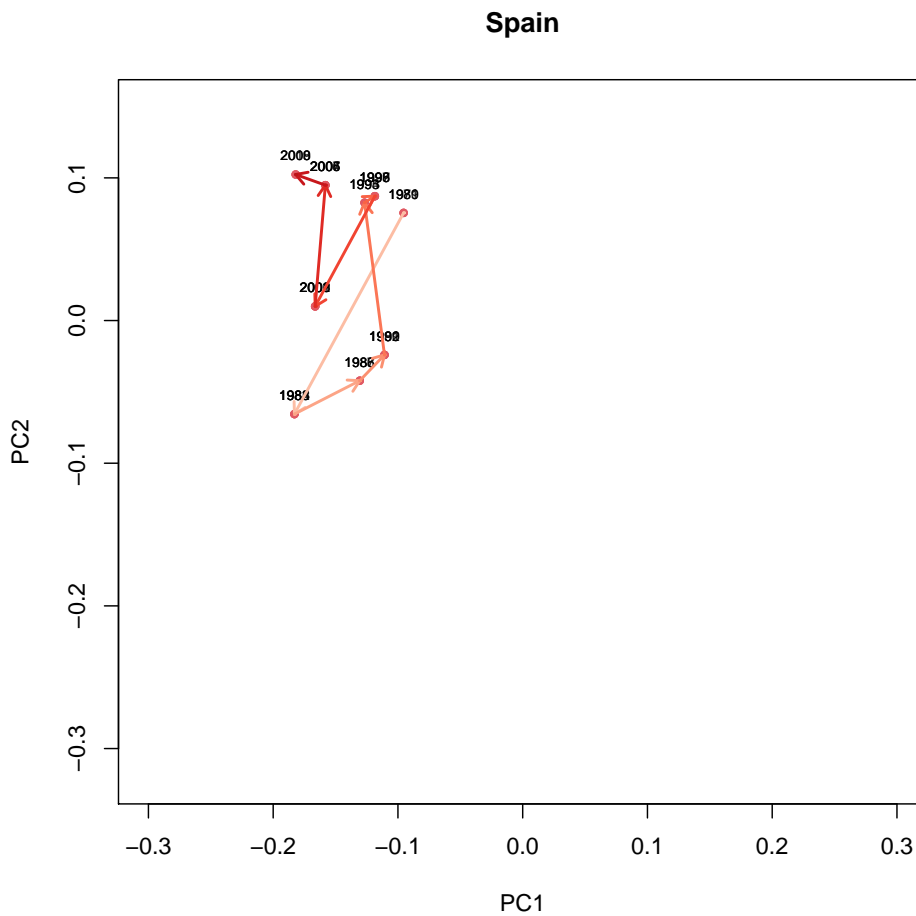
Notes: The plot shows the Norwegian party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

Figure 6.13: Norway on the PCA Dimensions



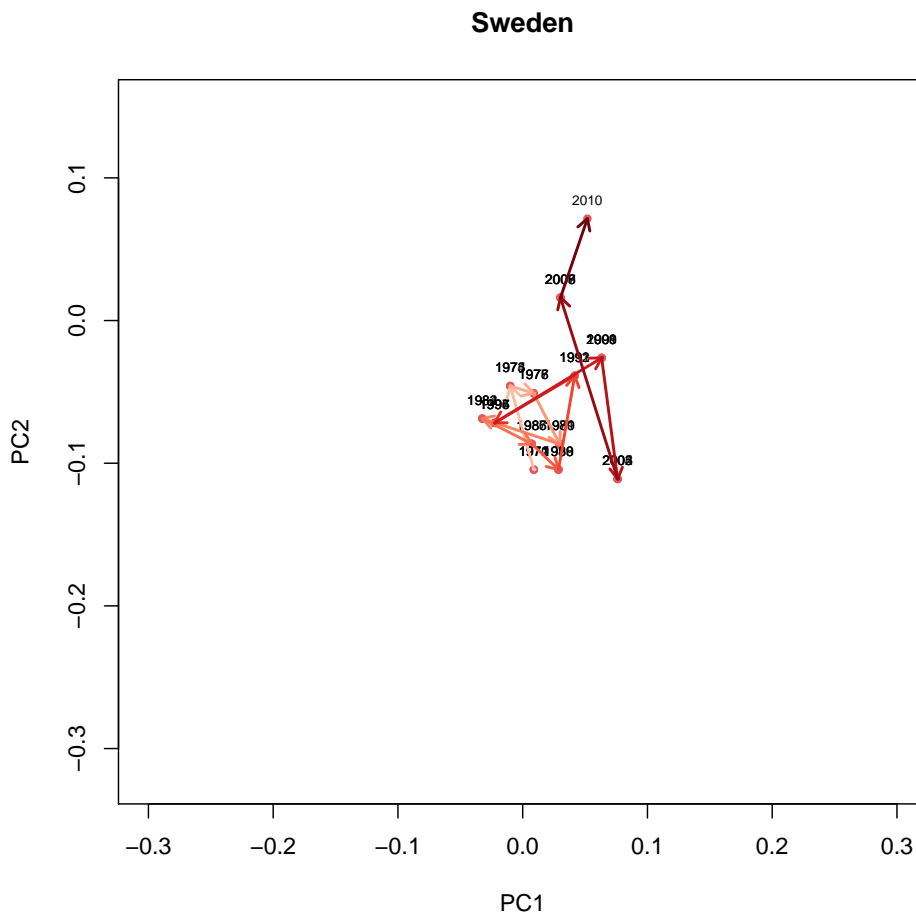
Notes: The plot shows the Portuguese party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

Figure 6.14: Portugal on the PCA Dimensions



Notes: The plot shows the Spanish party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

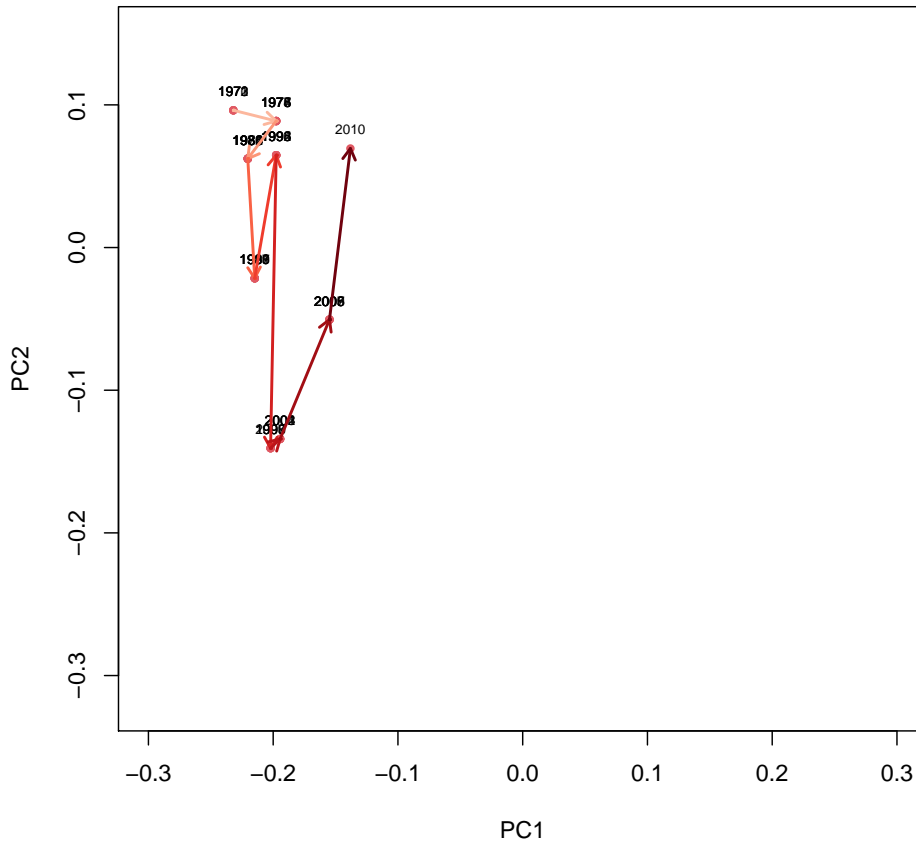
Figure 6.15: Spain on the PCA Dimensions



Notes: The plot shows the Swedish party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

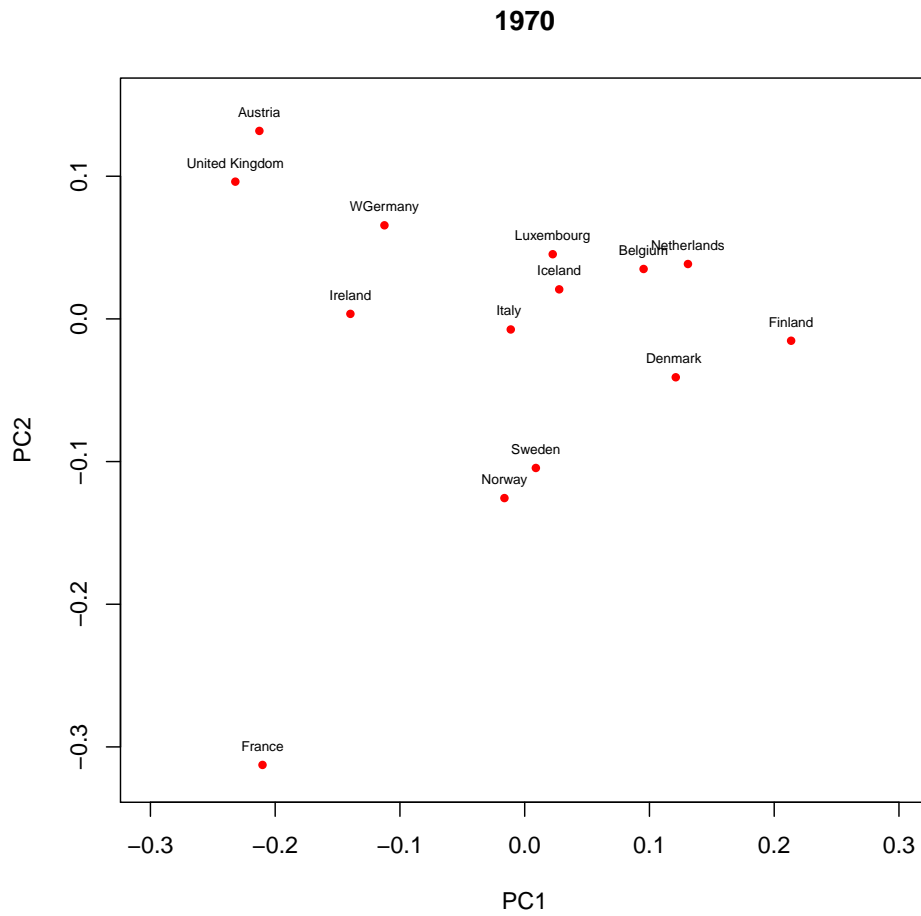
Figure 6.16: Sweden on the PCA Dimensions

United Kingdom



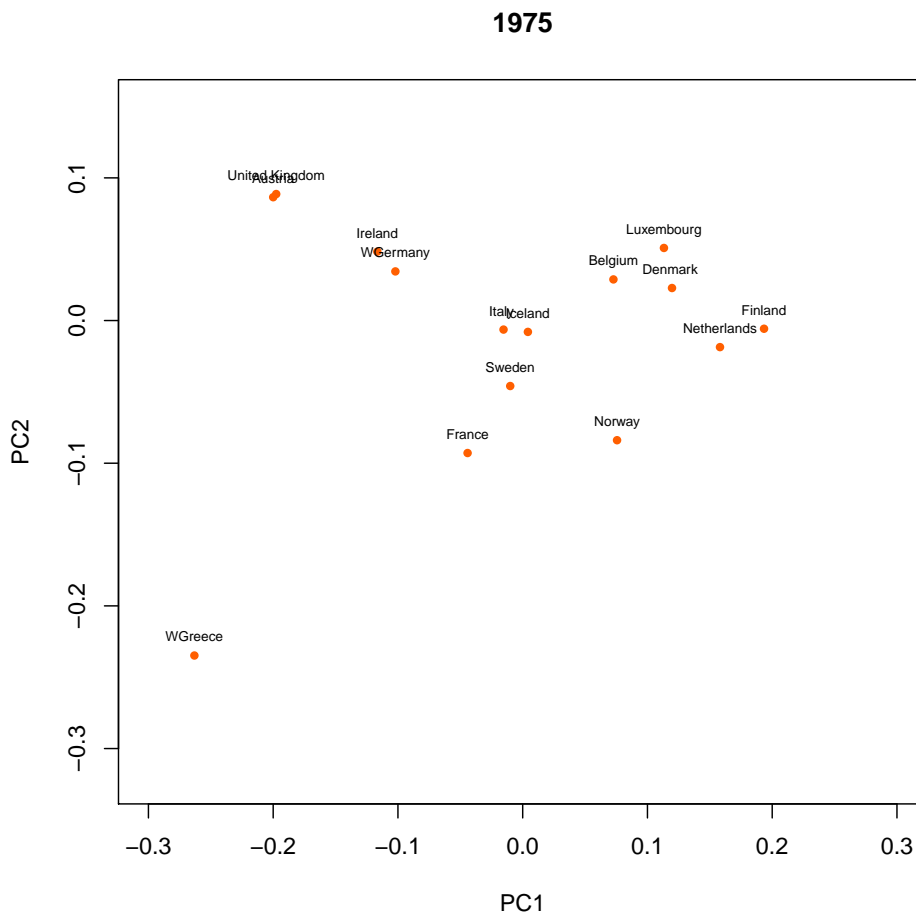
Notes: The plot shows the British party system in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis. The arrows progress from earlier to later years. The shading indicates the direction of the progress; the shading of the arrows becomes darker in later years.

Figure 6.17: The United Kingdom on the PCA Dimensions



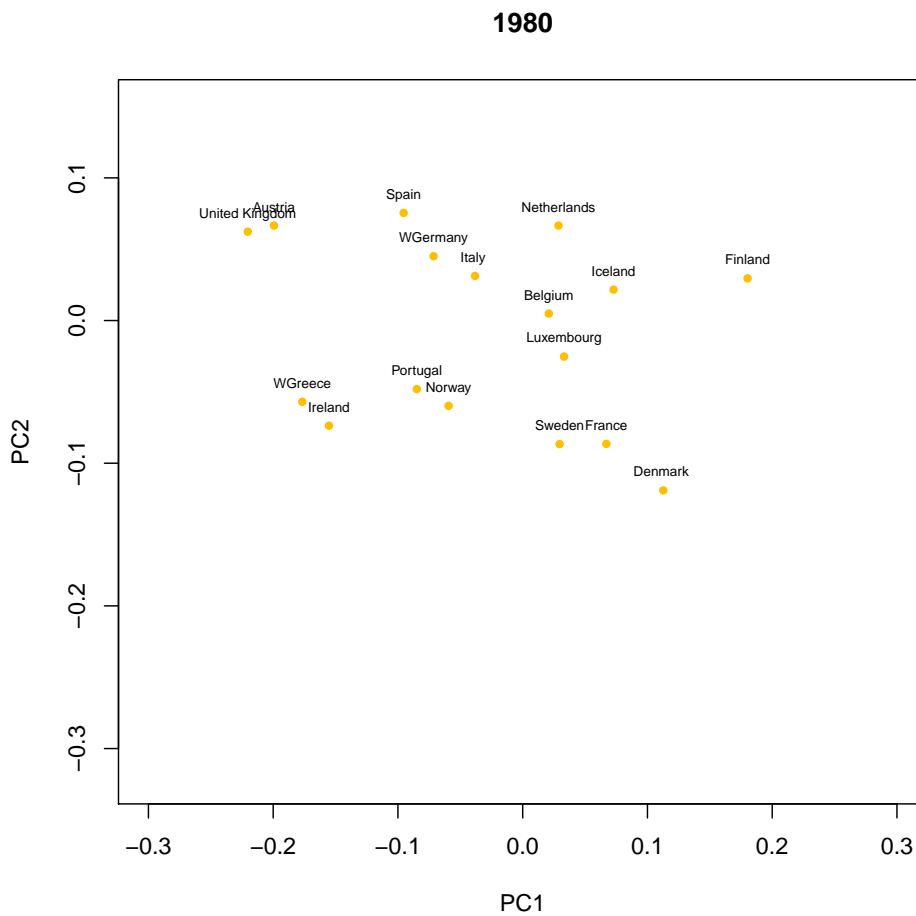
Notes: The plot shows the position of 17 countries in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis in 1970. The countries are: Austria, Belgium, Denmark, Finland, France Germany, Greece, Iceland, Italy, Ireland, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, the United Kingdom.

Figure 6.18: The Party Systems of 17 European Countries in 1970 on the PCA Dimensions



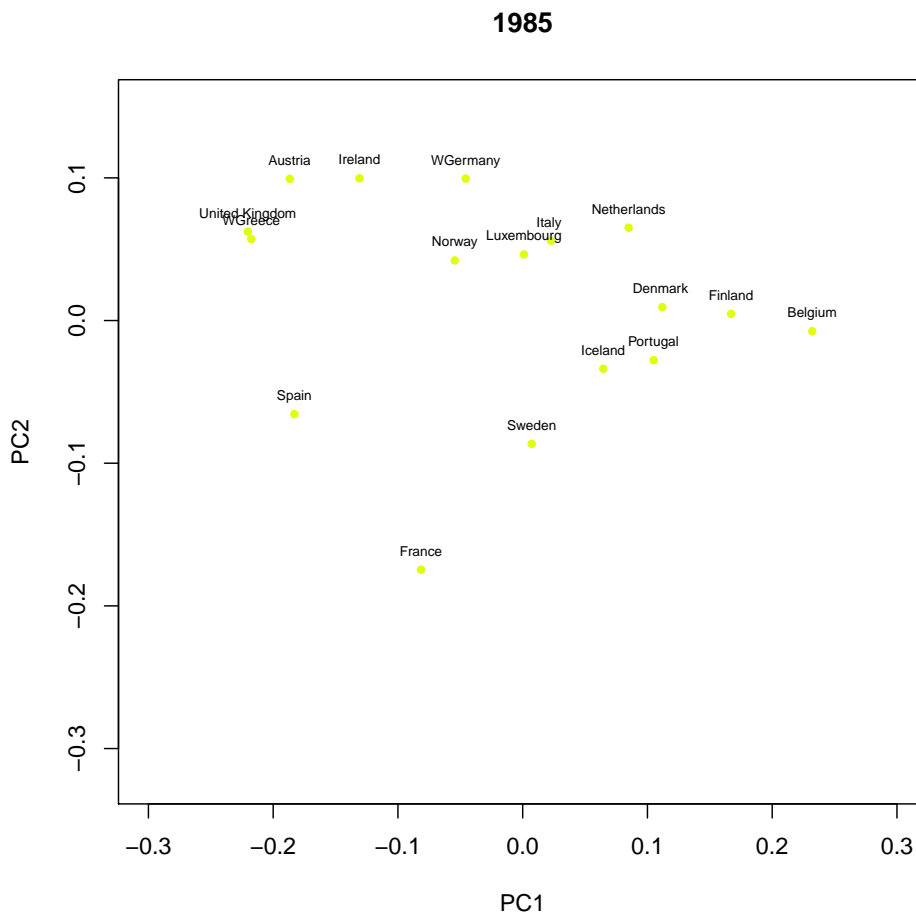
Notes: The plot shows the position of 17 countries in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis in 1975. The countries are: Austria, Belgium, Denmark, Finland, France Germany, Greece, Iceland, Italy, Ireland, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, the United Kingdom.

Figure 6.19: The Party Systems of 17 European Countries in 1975 on the PCA Dimension



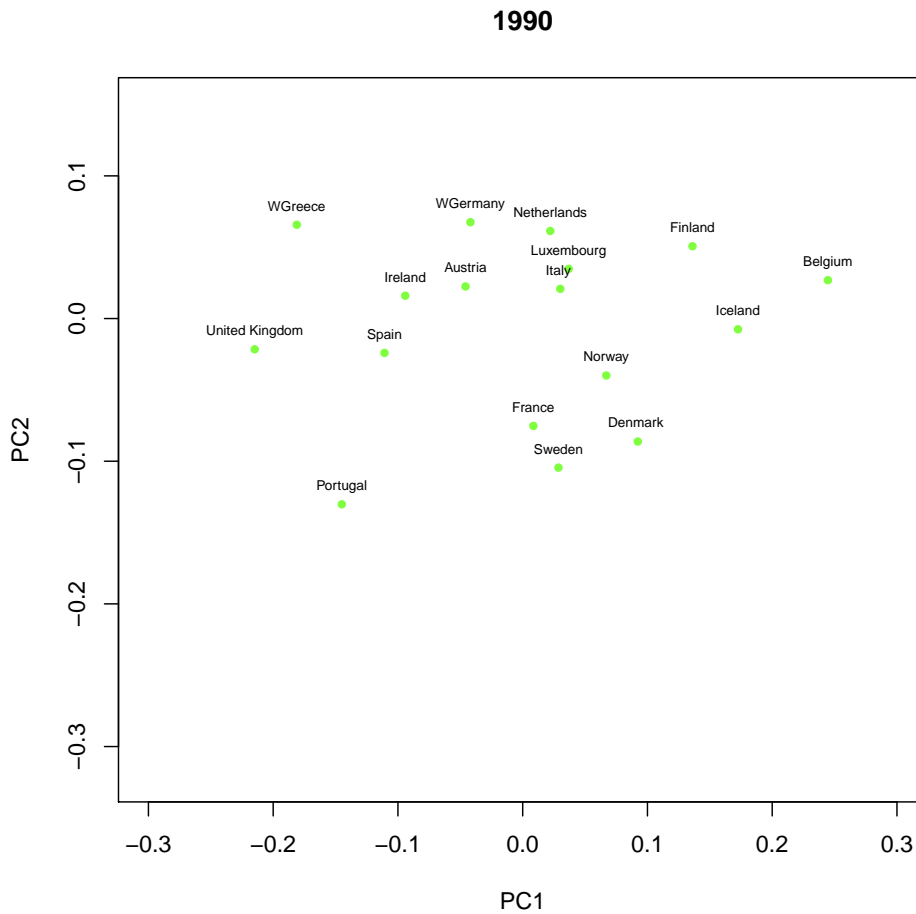
Notes: The plot shows the position of 17 countries in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis in 1980. The countries are: Austria, Belgium, Denmark, Finland, France Germany, Greece, Iceland, Italy, Ireland, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, the United Kingdom.

Figure 6.20: The Party Systems of 17 European Countries in 1980 on the PCA Dimensions



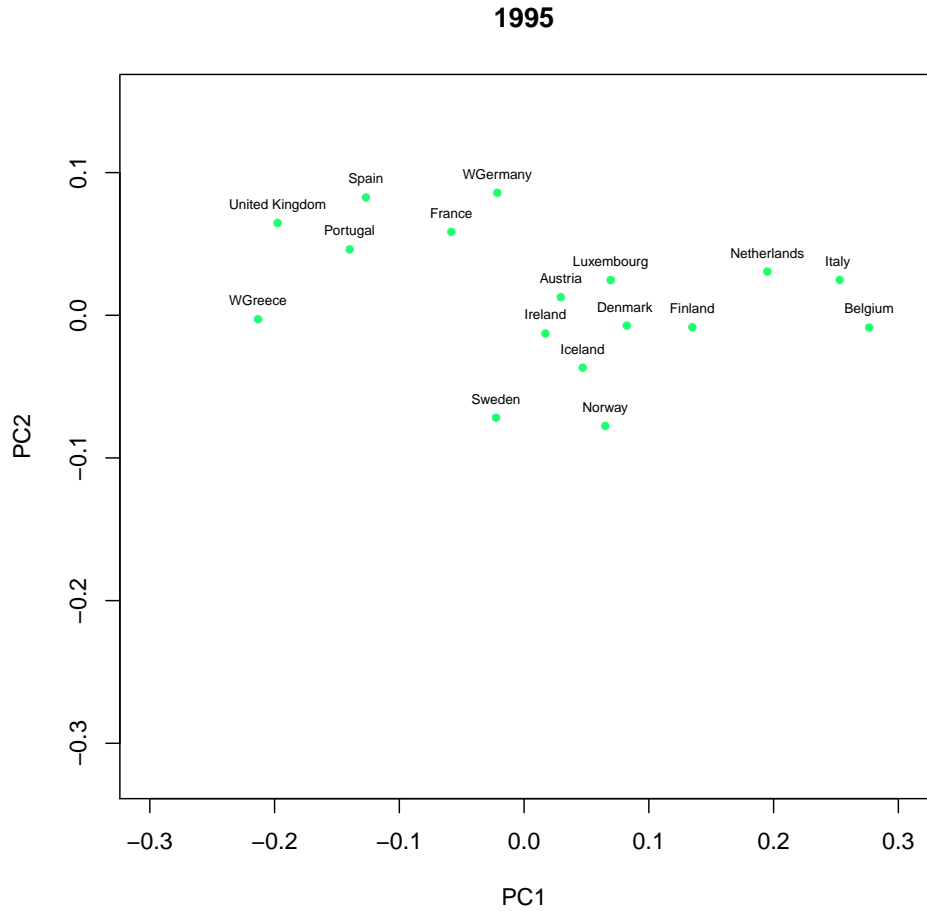
Notes: The plot shows the position of 17 countries in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis in 1985. The countries are: Austria, Belgium, Denmark, Finland, France Germany, Greece, Iceland, Italy, Ireland, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, the United Kingdom.

Figure 6.21: The Party Systems of 17 European Countries in 1985 on the PCA Dimensions



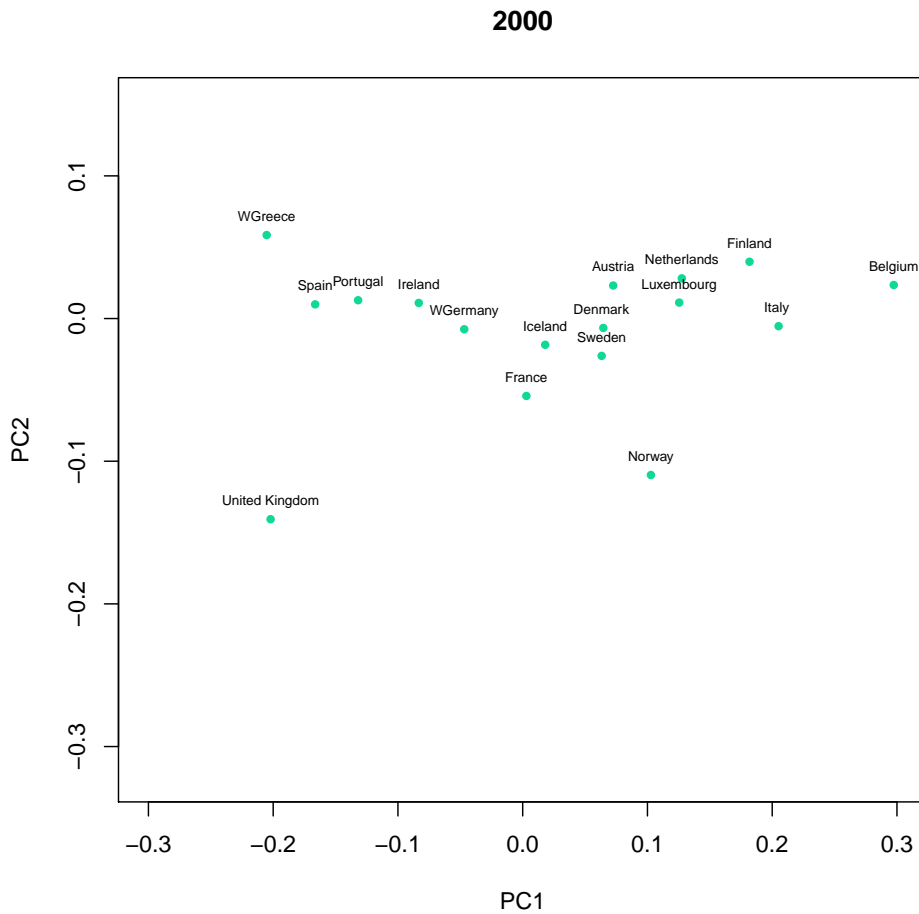
Notes: The plot shows the position of 17 countries in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis in 1990. The countries are: Austria, Belgium, Denmark, Finland, France Germany, Greece, Iceland, Italy, Ireland, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, the United Kingdom.

Figure 6.22: The Party Systems of 17 European Countries in 1990 on the PCA Dimensions



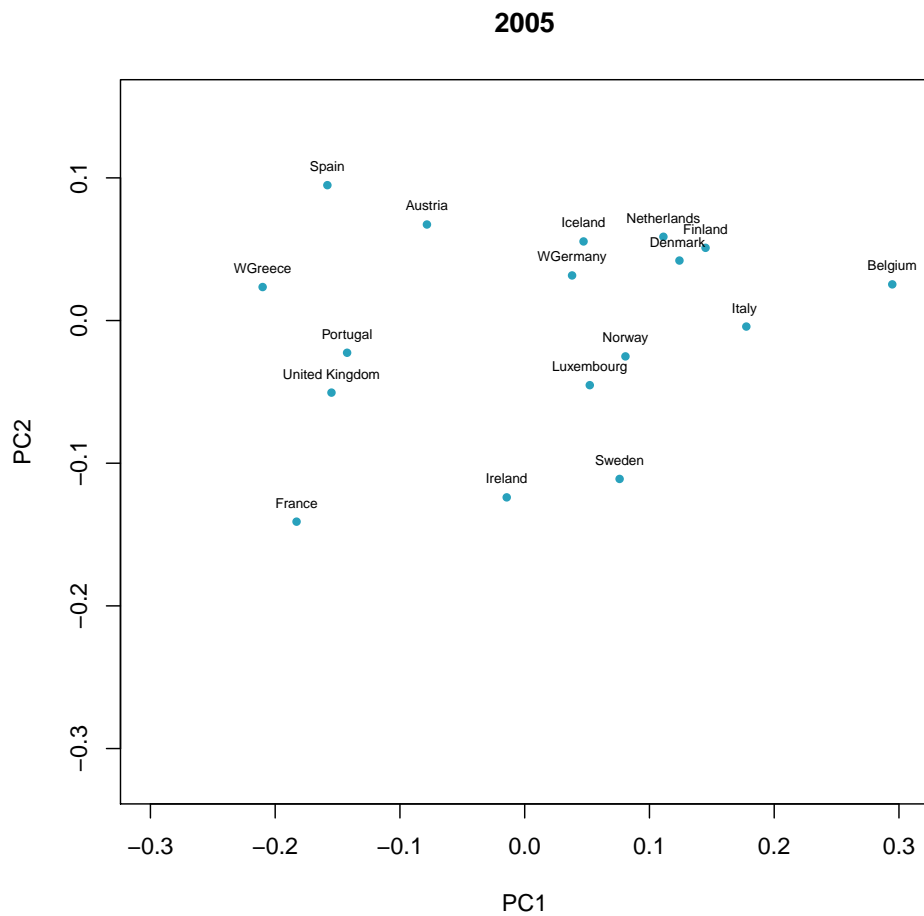
Notes: The plot shows the position of 17 countries in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis in 1995. The countries are: Austria, Belgium, Denmark, Finland, France Germany, Greece, Iceland, Italy, Ireland, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, the United Kingdom.

Figure 6.23: The Party Systems of 17 European Countries in 1995 on the PCA Dimensions



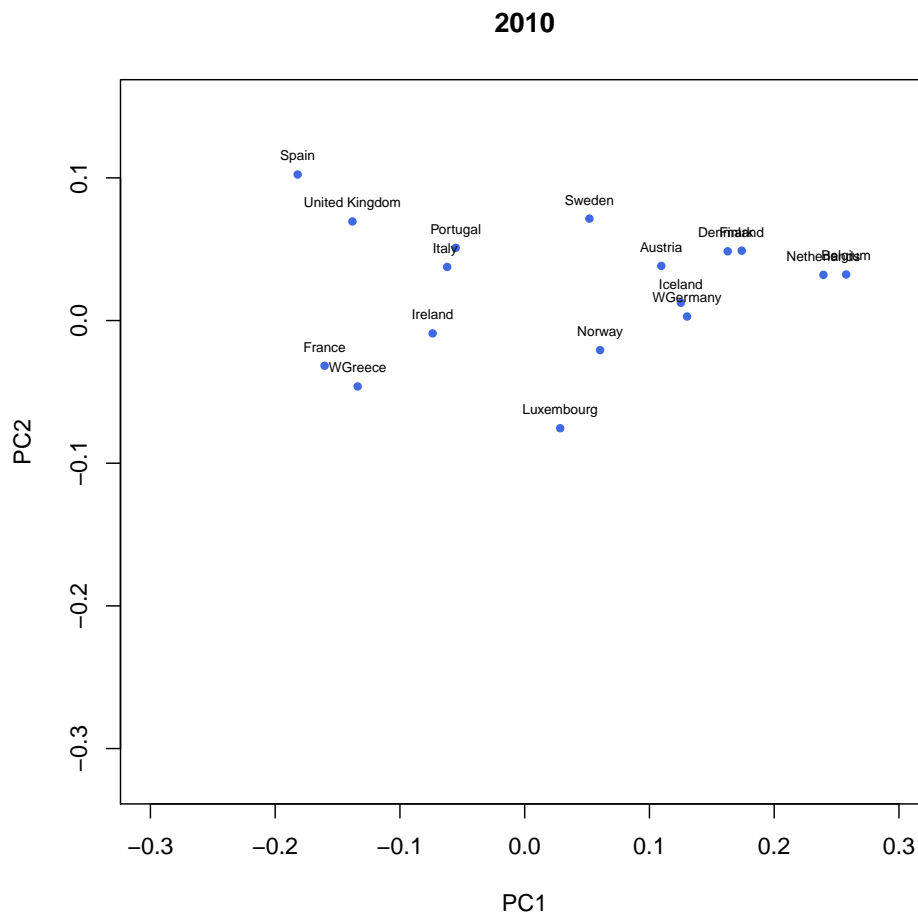
Notes: The plot shows the position of 17 countries in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis in 2000. The countries are: Austria, Belgium, Denmark, Finland, France Germany, Greece, Iceland, Italy, Ireland, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, the United Kingdom.

Figure 6.24: The Party Systems of 17 European Countries in 2000 on the PCA Dimensions



Notes: The plot shows the position of 17 countries in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis in 2005. The countries are: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Italy, Ireland, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, the United Kingdom.

Figure 6.25: The Party Systems of 17 European Countries in 2005 on the PCA Dimensions



Notes: The plot shows the position of 17 countries in the two dimensional plane defined by the first, and the second dimensions of the Principal Component Analysis in 2010. The countries are: Austria, Belgium, Denmark, Finland, France Germany, Greece, Iceland, Italy, Ireland, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, the United Kingdom.

Figure 6.26: The Party Systems of 17 European Countries in 2010 on the PCA Dimensions

BIBLIOGRAPHY

- Aitchison, J. 1983. "Principal Component Analysis of Compositional Data." *Biometrika* 70(1):57–65.
- Aitchison, John. 1986. "The Statistical Analysis of Compositional Data."
- Aitchison, John. n.d. A Concise Guide to Compositional Data Analysis. Manuscript University of Glasgow.
- Aitchison, John and Michael Greenacre. 2002. "Biplots of Compositional Data." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 51(4):375–392.
- Banzhaf, John F. 1965. "Weighted Voting doesn't Work: A Mathematical Analysis." *Rutgers Law Review*, 19(02):317–343.
- Baxter, Michael J. and Ian C Freestone. 2006. "Log-ratio Compositional Data Analysis in Archaeometry." *Archaeometry* 48(3):511–531.
- Bishop, Christopher M. 1995. *Neural Networks for Pattern Recognition*. Clarendon Press.
- Blondel, J. 1968. "Party Systems and Patterns of Government in Western Democracies." *Canadian Journal of Political Science* 1(02):180.
- Caulier, Jean-François and Patrick Dumont. 2005. "The Effective Number of Relevant Parties: How Voting Power Improves Laakso-Taagepera's Index."
- Dahl, Robert A. 1966. *Political oppositions in Western democracies*. Yale University Press.
- de Leeuw, Jan. 2005. "Nonlinear Principal Component Analysis." *Department of Statistics, UCLA* .
- Dunleavy, Patrick and Françoise Boucek. 2003. "Constructing the Number of Parties." *Party Politics* 9(3):291–315.
- Duverger, Maurice. 1954. *Political Parties, their Organization and Activity in the Modern State*. Methuen.

- Falcó-Gimeno, Albert and Ignacio Jurado. 2011. "Minority Governments and Budget Deficits: The Role of the Opposition." *European Journal of Political Economy* 27(3):554–565.
- 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.
- Grofman, Bernard. 2006. "The Impact of Electoral Laws on Political Parties." *The Oxford Handbook of Political Economy* pp. 102–18.
- Hofmann, Thomas, Bernhard Schölkopf and Alexander J Smola. 2008. "Kernel Methods in Machine Learning." *The Annals of Statistics* pp. 1171–1220.
- Jolliffe, Ian T. 2002. *Principal Component Analysis*. Springer.
- Katz, Jonathan N. and Gary King. 1999. "A Statistical Model for Multiparty Electoral Data." *American Political Science Review* 93(01):15–32.
- Kesselman, Mark. 1966. "French Local Politics: A Statistical Examination of Grass Roots Consensus." *American Political Science Review* 60(04):963–973.
- Kitschelt, Herbert. 2008. Party Systems. In *The Oxford Handbook of Comparative Politics*, ed. Carles Boix and Susan Carol Stokes. Oxford: Oxford University Press chapter 22, pp. 522–554.
- Kline, Reuben. 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.
- Laakso, Markku and Rein Taagepera. 1979. "'Effective' Number of Parties: A Measure with Application to West Europe." *Comparative Political Studies* 12(1):3–27.
- Laver, Michael and Kenneth Benoit. 2015. "The Basic Arithmetic of Legislative Decisions." *American Journal of Political Science* 59(2):275–291.

- Lijphart, Arend. 1994. *Electoral Systems and Party Systems: a Study of Twenty-seven Democracies, 1945-1990*. Oxford University Press.
- Linting, Marielle. 2007. *Nonlinear Principal Components Analysis: Introduction and Application: Chapter 2*. Manuscript.
- Linting, Marielle, Jacqueline J. Meulman, Patrick JF Groenen and Anita J van der Kooij. 2007. "Nonlinear Principal Components Analysis: Introduction and Application." *Psychological methods* 12(3):336.
- Lipset, Seymour Martin and Stein Rokkan. 1967. *Cleavage Structures, Party Systems, and Voter Alignments: an Introduction*. The Free Press.
- Maeda, Ko. 2010. "Divided we Fall: Opposition Fragmentation and the Electoral Fortunes of Governing Parties." *British Journal of Political Science* 40(02):419–434.
- Maeda, Ko. 2015. "Determinants of Opposition Fragmentation Parliamentary Rules and Opposition Strategies." *Party Politics* 21(5):763–774.
- Mair, Peter. 2002. Comparing Party Systems. In *Comparing Democracies 2: New Challenges in the Study of Elections and Voting*, ed. Lawrence LeDuc, Richard G. Niemi and Pippa Norris. London: Sage chapter 4, pp. 88–108.
- Michailidis, George and Jan de Leeuw. 1998. "The Gifi System of Descriptive Multivariate Analysis." *Statistical Science* 13(4):307–336.
- Molinar, Juan. 1991. "Counting the Number of Parties: An Alternative Index." *The American Political Science Review* 85(4):1383.
- Rae, Douglas W. and Michael Taylor. 1970. *The Analysis of Political Cleavages*. Yale University Press.
- Rokkan, Stein. 1970. *Citizens, Elections, Parties: Approaches to the Comparative Study of the Processes of Development*. Scandinavian University Press.

- Rozenas, Arturas. 2011. "A Statistical Model for Party-Systems Analysis." *Political Analysis* 20(2):235–247.
- Sartori, Giovanni. 1976. *Parties and Party Systems: a Framework for Analysis*. Cambridge University Press.
- Schölkopf, Bernhard and Alexander J Smola. 2002. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT press.
- Schölkopf, Bernhard, Alexander Smola and Klaus-Robert Müller. 1997. Kernel Principal Component Analysis. In *International Conference on Artificial Neural Networks*. Springer pp. 583–588.
- Shapley, Lloyd S. and Martin Shubik. 1954. "A Method for Evaluating the Distribution of Power in a Committee System." *American Political Science Review* 48(03):787–792.
- Straffin Jr., Philip D. 1988. The Shapley-Shubik and Banzhaf Power Indices as Probabilities. In *The Shapley Value: Essays in Honor of Lloyd S. Shapley*, ed. Alvin E Roth. Cambridge: Cambridge University Press chapter 5, pp. 71–82.
- Taagepera, Rein. 1999. "The Number of Parties as a Function of Heterogeneity and Electoral System." *Comparative Political Studies* 32(5):531–548.
- Taagepera, Rein and Matthew Soberg Shugart. 1989. *Seats and Votes: the Effects and Determinants of Electoral Systems*. Yale University Press.
- ter Braak, Cajo JF. 1983. "Principal Components Biplots and Alpha and Beta Diversity." *Ecology* 64(3):454–462.
- Vitt, L. J., J. P. Caldwell, P. A. Zani and T. A. Titus. 1997. "The Role of Habitat Shift in the Evolution of Lizard Morphology: Evidence from Tropical *Tropidurus*." *Proceedings of the National Academy of Sciences* 94(8):3828–3832.
- Wildgen, J. K. 1971. "The Measurement of Hyperfractionalization." *Comparative Political Studies* 4(2):233–243.