Measures of Populism in the CHES 2017 Dataset

By Constantin Colonescu[*]

Populism is a fuzzy concept in world politics; it can take many shapes and colors, thus evading rigorous definition. Using expert evaluations provided in Chapel Hill Expert Survey (CHES) 2017 data, we try to identify features that predict populism and to characterize various European parties on a populism scale. As a byproduct, we find that experts have often diverging opinions on a party's stance on various issues.

Introduction and Scope

Chapel Hill Expert Survey (Polk et al. 2017) is a collection of expert evaluations of party positions in several European Union countries; the survey asks questions concerning a party's position with respect to the European Union project, general left-right positioning, both political and economic, and, for the first time in the sequence of the CHES surveys, questions about populism, migration, and coordination of European Union policies. Most of the data are scores on a scale of 1 to 7 or 0 to 10. Since the scales are different for different variables, the data will be standardized for the purpose of this analysis. Using various statistical methods, such as clustering, structural equation modeling and recursive partitioning, we provide an image of how close various parties are with respect to each other when multiple aspects are considered and identify the most relevant attributes in predicting the degree of populism for all parties.

Hooghe et al. (2010) compare the 2002 and 2006 waves of the CHES survey with other sources (Comparative Manifesto, Benoit-Laver, and Rohrschneider-Whitefield) and conclude that the CHES data are "reasonably valid and reliable." Hooghe et al. (2010) use listwise deletion for missing data but are somehow cryptic about the details of their method. As opposed to many of the studies using CHES data, which share a dismissive attitude towards how they treat missing data, we try to pay more attention to this aspect.

The dataset contains 2168 observations on 31 variables, 14 countries, and 132 political parties; a total number of 228 experts provided assessments for one party or another. The CHES 2017 Codebook contains detailed information about the data (CHES 2017).

Here is a succinct guide to the symbols representing the variables in CHES data: *lrecon* measures a party's stance (left-right) on economic matters, with low values for left and high values for right; *lrgen* is the same, but on general matters, not only economic; *galtan* is a party's position from green, alternative, and libertarian at the lower end of the scale to traditional, authoritarian, and nationalist

---

[*]Associate Professor, MacEwan University, Canada.

at the highest end, a variable sometimes criticized for being difficult to measure; *position* measures a party's position toward the European Union project; the attributes *clear* and *salience* in the name of a variable refers to such qualities of a party's rhetoric on the respective issue; the attribute *dissent* refers to the degree of dissent within a party on the respective issue. Other variable names are more self-explanatory.

Descriptive Statistics and Missing Data in CHES 2017

Table 1 provides some descriptive statistics, including the percentage of missing values for each variable[1]; the variables are sorted in descending percentage of missing values. It shows that the variable *lr_econ salience* (salience of economic issues in a party's stance) has the least percentage of missing data, while *multicult_dissent* (the degree of within-party dissent concerning the integration of immigrants) has the highest percentage (about 28%).

Of the total number of 2168 observations in CHES 2017, only 960 are complete (with no missing data whatsoever), or about 44%. Thus, if listwise deletion is used in an analysis involving all variables, more than half of the dataset must be discarded. The situation is, though, less serious when the observations and the variables having the most missing data are eliminated from the analytic model, instead of applying list-wise deletion to the whole dataset. It will be useful, therefore, to identify the variables and observations containing the most missing values.

Let us count all missing values by country, to determine whether potentially eliminating all parties in a country may substantially reduce the missing values proportion in the data. Table 2 shows a country ranking by missing values; the *prMiss* variable is the percentage of missing values in all variables, by country, sorted in descending order of this percentage.

Table 2 shows that Italy stands out with the highest number and percentage of missing values. Table 3 shows the parties having the highest number of missing values in the dataset, across all variables. The parties with 75% or more missing values will be removed from the dataset. These parties are VdA(UV) and LKR.

Listwise, or pairwise deletion may be acceptable when a suitable combination of observations and variables is selected to minimize the percentage of missing values, a combination that still serves the purpose of one's investigation. For most purposes, though, there are better alternatives to listwise deletion. These are either multiple imputation, or some full-information maximum likelihood (FIML) method that can use all the information available in an incomplete dataset, without imputation.

An important step in dealing with missing data is to determine what kind of missingness we have: missing completely at random (MCAR), when missingness has no relation to observed or unobserved data, missing at random (MAR), when missingness has some relationship with observed, but not unobserved data, or not

---

[1]All tables and figures are available at the end of the paper after the References section.

at random (MNAR), when missingness is more likely to happen for some sub-groups of observed or unobserved variables; missingness of MNAR data is directly related to the values of the missing data. In other words, MNAR data is missing precisely because of their values. MCAR values cannot be predicted with existing data, but dropping them from analysis is not likely to substantially alter the results; MAR can, and should be predicted using the available information; and, finally, MNAR can be neither predicted, nor ignored; this type of missingness is, therefore, the most problematic. Even detecting MNAR data is difficult and requires good knowledge of the data being analyzed.

The next passages try to determine the type of missingness in the CHASE dataset. To do so, a missingness matrix is calculated; each entry in this matrix is binary, with 0 for non-missing and 1 for missing. Then, we try to detect relationships between each variable in the dataset and missingness in the other variables. This part uses the R package *naniar* (Tierney et al. 2018) and the *MissMech* package (Mohan and Pearl 2014).

First, a test based on Jamshidian and Jalal (2010), using the *TestMCARNormality* function in the *MissMech* package (Mohan and Pearl 2014) rejects the hypothesis that the data is MCAR, leaving us only the MAR and MNAR possibilities. Now, we should remember that CHES data are scores, or rankings provided by independent experts, who should have no personal reason for refusing to answer certain questions; therefore, there is no conceivable reason to believe that missing data would be not at random (MNAR), leaving us with the only alternative that the CHES data is MAR, which makes it suitable for multiple imputation. After having removed the parties with more than 75% missing data, we complete the dataset by multiple imputation, using the *Amelia* package (Honaker et al. 2011); the country and party identifier variables are not used for imputation.

Factor Analysis in CHES 2017

The purpose of factor analysis is to get a sense of the similarities existing among the 26 variables in the dataset and, if possible, to identify a dimension – a linear combination of the variables – that best represents some measure of populism. The dataset will be scaled before performing factor analysis to account for differences in the scales of different variables.

Therefore, in some of the following graphs the values are in standard deviations from the mean; in other words, negative values correspond to lower values of a variable or factor. The scores provided by several experts for one party have been aggregated using the median.

Figure 1 displays a scree plot where each bar indicates the percentage of variance covered by a factor (dimension). Such diagrams allow determining an "optimal" number of factors; an extra factor shall only be included if it adds a significant share to the variance. We retain only three dimensions, those that individually contain more than 10% of the variation in the data. The fact that each dimension or group of dimensions captures a rather small amount of the total

variance reflects low correlation among the measured variables. This characteristic of the CHES data makes the data not too useful beyond the usefulness of each variable taken separately.

In Figure 2, each row of graphs contains the contributions of variables and individual parties to the first three dimensions. Again, each dimension appears to be made of small contributions by many primary variables and by many individual parties, indicating that most variables are but weakly distinct and most parties have rather indistinct profiles.

Exploring Inter-Variable Relationships

Grouping (clustering) the variables in a dataset is a simpler alternative to factor analysis, based on the assumption that each variable belongs to only one group (hierarchical cluster analysis, Revelle 1979). Figure 3 shows such a variable grouping, assigning the variables to 5 high-level groups and several sub-groups. As in a regular, by-observation clustering, one can decide how many groups to use for clustering. Such variable grouping could be informative when designing a structural equation model, as we do in the next section. More generally, this method could help constructing composite scales for measuring features that are not directly observable; according to Revelle (2018), the criteria of selecting certain variables to form a construct could be "rational, empirical, or factorial."

In Figure 3, high values for the alpha and beta statistics indicate good grouping; two groups will merge into a higher-level group if alpha and beta are increased by the merger. After experimenting with different numbers of groups, we settle on an eight-group model (see Figure 3); some of the groups, such as the ones labeled C17 and C18 turn out to be quite stable when the number of groups is changed; others, such as C14, only appears when the number of groups becomes equal to eight. Higher values of α and β for the C18 and C14 groups also indicate good reliability of these groups. The composition of most of the smaller groups does not seem to provide too much insight; therefore, we focus on the three larger groups.

The group labeled C17 in Figure 3, probably not incidentally, brings together three out of the four variables collected in the CHES 2017 Codebook under the title "Populism, Corruption, and Internal Party Democracy Questions;" the fourth variable, *members_vs_leadership*, does not seem to belong to the C17 group. To our knowledge, the CHES authors do not justify why the four variables have been placed under the same title. The group labeled C14 collects the variables measuring the salience and clarity in a party's stance about their GAL–TAN (Green, Alternative, and Libertarian, versus Traditional, Authoritarian, and Nationalist) characteristics. Interestingly, though, the two "real" GAL–TAN variables, namely *galtan* and *galtan_new* are not in this group. The group labeled C18 has a complex structure, which remains very stable with respect to changes in the number of groups. Getting clues from this and the other major groups, we build a structural equation model in the next section.

A Latent Variable Model for the CHES 2017 Data

Exploratory factor analysis gives us an idea about how accurately different variables describe different features of the individual parties, as well as how different parties are located along several dimensions. Our conclusion is, so far, that CHES data do not identify very well cleavages among parties, and that each variable can be found in several dimensions (principal components). This conclusion is not too helpful in our quest of better defining populism or better identifying it based on parties' scores over a particular set of CHES variables. This section attempts to create a structural equation (latent-variable) model that would better describe and classify the parties in our dataset. For building this model, we rely on the grouping-by-variable result in Figure 3 and use the R package *lavaan* (Rosseel 2012).

The following hypothesized relationships define our latent variable model; The "~" symbol indicates the usual regression relationship, the symbol "= ~" indicates a measurement relationship, where the left-hand side is a latent variable, *manifested by* the latent or measured variables in the right-hand side; the symbol "~ ~" represents unexplained correlation between variables. Since all the variables in the CHES data aim at measuring some underlying, unobservable characteristics of the parties, our model cannot aim at revealing any causal relationship, but only at identifying commonalities and exposing those unobservable characteristics such as populism, nationalism, and left-right orientations.

The following hypothesized relationships describe our structural equation model:

- populism = ~ people_vs_elite + antielite_salience + corrupt_salience
- nationalism = ~ immigrate_policy + eu_asylum + ethnic_minorities + multiculturalism
- proEU = ~ position + eu_budgets
- left_right = ~ lrgen + lrecon + lrecon_clear + lrecon_salience
- ideology = ~ galtan + galtan_salience + galtan_new + galtan_new_clear + galtan_new_salience
- proEU ~ populism + nationalism + left_right
- left_right ~ ~ ideology
- nationalism ~ ~ populism

In a structural equations model, latent variables determine, rather than being determined by the observed, manifest variables. CHES variables that would measure nationalism could be the following, as defined in the CHES 2017 Codebook: *eu_asylum* (1 – party leadership strongly opposes common policy on asylum, 7 – strongly favors common policy); *immigrate_policy* (1 – strongly opposes a restrictive policy on immigration, 10 – strongly in favor of restrictive policy on immigration); *eu_dissent* (0 – party completely united on EU integration, 10 – party extremely divided); and *multicult_dissent* (0 – party is completely united on integration policy for immigrants, 10 – party extremely

divided on policy). The definitions of the first two variables show some inconsistency in establishing the scales. The scale for *eu_asylum* is low if party opposes common policy (weather such a party opposes asylum in itself or just a common policy is unclear), while the scale for *immigrate_policy* is low for a party that favors immigration. In other words, one would expect that an anti-immigration party would score low on the *eu_asylum* scale and high on the *immigrate_policy* scale.

Now, suppose our model predicts a high probability, close to one, on the *nationalism* latent variable. Is this indeed a mostly nationalist party? Hard to tell, upfront, given the inconsistency in the scales just mentioned. To tackle this issue, however, we may turn to the results of the model and try to calibrate our interpretation of the results. For instance, let us look at a party that many observers would consider as nationalist, such as France's National Front (FN; see Table 5). Our model predicts a maximum score (equal to 1) for this party, which suggests that indeed a higher predicted score on nationalism corresponds to a more nationalist party.

Table 4 gives estimated coefficients of the model. The effect of a latent variable on the first variable in its corresponding equation is scaled to 1; its effects on the remaining variables are scaled relative to the first variable. For instance, we see that a party's underlying *populism* affects the party's *antielite_salience* measure more than it affects the party's *people_vs_elite* measure.

As it turns out, the CHES data allow very little space in creating a reliable model. Our model's chi-square value, as a measure of reliability, is rather modest. However, many of the predictions of the model, as displayed in Table 5 seem reasonable. (The table shows predicted values, on a scale from 0 to 1, for only 40 out of the about 130 parties; the 40 parties have been selected for having the highest scores on the *populism* scale.) Figure 4 illustrates the Structural Equation Model, where the various boxes contain shortened latent and manifest variable names, and the direction of the arrows show the hypothesized directions of causality. The thickness and color intensity of the arrows are proportional to the strength of the respective effect. An interesting result revealed by the model is the rather weak correlation (equal to 0.2) between populism and ideology, as defined in our model.

Clustering by Parties

This section explores commonalities among parties, using a different approach: clustering by observations, rather than by variables as in a previous section. When several observations are available for one party (several expert assessments), we aggregate the observations by their median, so that each aggregate observation corresponds to one party.

First, we evaluate how prone the CHES 2017 dataset is to clustering (we assess the clustering tendency of the dataset) by calculating the Hopkins statistic, which tests the (null) hypothesis that there are no adequate clusters in the data. A

value close to 0.5 indicates that the data is not suitable for clustering, while a value close to zero indicates high clusterability. We find a value of the Hopkins statistic H = 0.33, which is somewhere at the limit of clusterability.

The optimal number of clusters is determined by the *gap statistic*, which maximizes the difference (*gap*) between the sum of intra-cluster distances of the data and an amorphous artificially created dataset, one with no obvious clusters. As Figure 5 shows, the optimal number of clusters is four. Figure 6 displays the clusters and the parties in each cluster.

Table 6 shows a count of cases (parties) by category and cluster. This can give an image of how "pure" each cluster is. For instance, most of the populist parties (category 2) can be found in cluster 2.

Table 7 shows the means of the CHES variables by cluster, retaining only the variables that have the mean at the center of Cluster 2 greater than those in the other three clusters. The means are given in percentages of their scores. This table helps identifying the attributes that are common to the parties in this cluster, namely KSCM, EKRE, DLF, FN, AfD, BLAU, ANEL, XA, Fidesz, JOBBIK, FdI, LN, FvD, PVV, SGP, Korwin, Kukiz, PiS, Kotleba LSNS, Sme Rodina, SNS, SD, DUP, UKIP.

So, what can we say about these parties? Table 7 indicates that the *ethnic_minorities* variable has the highest score, suggesting that these parties tend to oppose more rights for ethnic minorities; they advocate stronger restrictions on immigration and stronger policies of assimilation of immigrants as opposed to multiculturalism; ideologically, they lean toward extreme right; they also outscore the other clusters in the variables that are supposed to measure populism as per the CHES Codebook, namely the variables *people_vs_elite* and *members_vs_leadership*. A higher value of the *people_vs_elite* variable reflects the stance that "the people," not politicians should make the important decisions; a higher value of the *members_vs_leadership* variable reflects the stance that the leadership of a party should have control over policies, not the members. A similar message is conveyed by the high value of the *galtan* variable: these parties lean toward the TAN (Traditional, Authoritarian, and Nationalist) side on the GAL–TAN scale.

Table 8 gives the means of the variables at the center of Cluster 4; the parties in this cluster are EA/EH Bildu, IU, Podemos, PSOE, SDE, EELV, Insoumis, Grunen, DIMAR, Potami, SYRIZA, DK, E14, MLP, PM, FI, MDP, PD, SI, D66, GL, Nowo, Razem, TR, BE/O Bloco, PAN, FI, MP, V, Greens, LibDem. While this cluster is less well defined than cluster 2, the values displayed in Table 8 still identify some commonalities. As compared to the other clusters, these parties support a common EU policy concerning political asylum; while their stance on supporting the EU project is less than spectacular, it is still greater than the one of the other clusters. These parties also tend to support EU authority over the member states budgetary policies.

Cluster Validation

A good clustering algorithm should identify clusters that are compact (low distance between within-cluster cases) and well separated (high distance between clusters), if such clusters exist in the data. Various validation statistics, such as the Dunn index, are used by some clustering algorithms to determine the optimal number of clusters. Thus, further investigation of such validation statistics is not necessary. An exception, though, is the silhouette index, which takes values between -1 and 1; a silhouette index closer to 1 indicates a more compact cluster, a desirable feature of clustering. Figure 7 shows the average silhouette statistics by group; it indicates that clusters 2 is the most compact, while cluster 4 is more heterogeneous.

A Recursive Partitioning Model

The purpose of this section is to determine the variables that best predict the populist character of a political party or, from a different perspective, to identify the attributes that make a party belong to a certain category. We use the CHES 2017 data and an additional, original classification of EU parties by expert opinion. We classify the parties into three categories: 0 = "not populist," 1 = "populist-left," and 2 = "populist-right."

A *recursive partitioning* model (Breiman et al. 1983), creates a decision tree by repeatedly splitting variables into ranges. Figure 8 displays a simpler version, while Figure 9 displays a more complex version of the model. The shaded boxes at the bottom of the diagram in Figure 8 display the following information: the number at the top is the most likely of the three categories, 0, 1, or 2; the three numbers in the middle row give the probabilities for each of the categories and, finally, the number at the bottom of the box gives the percentage of observations that have the combination of characteristics leading to the respective box.

This combination is shown as the path from the top of the diagram to the respective box at the bottom. Let's take, for example, box 43: it represents 1% of the parties in the dataset; these parties have been predicted as left-populist (coded as 1); the probability that these parties are indeed in category 1 is 0.94, the probability that they are not populist is 0.06, and the probability that they are right-populist is about 0. The parties in box 43 have the following attributes, starting from the top of the tree: they favor common asylum policies in various degrees (*eu_asylum* $\geq$ 2), they somehow oppose European integration or are neutral (*position* < 3.9), they take a clear stand towards left-oriented economic policies, and they consider immigration to be an important topic.

An interesting insight produced by the recursive partitioning algorithm is that the right-populist parties identified in boxes 23, 51, 55, and 109 have less defined attributes, with probabilities between 0.59 and 0.67; on the contrary, those identified in box 7 are well defined (probability = 0.9). Moreover, box 7 is at the end of a very simple path: *eu_asylum* < 2 and *lrgen* > 7.3. In other words, they are to the far right on the ideology scale and oppose a common asylum policy.

An unexpected outcome of this model is how important the variable *eu_asylum* is in predicting the stance of a party between being populist or not.

We can use now the recursive partitioning model (Therneau and Atkinson 2019) to predict a party's belonging to one of our three categories: not populist (0), left-populist (1), and right-populist (2). Table 9 is a contingency table, showing the actual categories as determined by expert assessment as row names and those predicted by the model as column names. The degree of accuracy of this prediction can be assessed using the Rand index; the closer the Rand index is to 1, the better the prediction. Our calculated Rand index is equal to 0.76, showing a reasonably good fit.

Discussion

We use CHES 2017 data to identify common features of parties deemed to be "populist" in political science literature.

The first part of the analysis examines the missing-value structure of the data and finds that there is a substantial amount of missing values in the CHES 2017 dataset. We find that, for this dataset, imputation is the least invasive method to deal with missing data. We checked our results on several versions of imputed data and the results were very similar. The second motive of our investigation sets to characterize the profiles of various European parties by two methods. First, we try to identify patterns in the data that would classify European parties according to underlying, latent categories; a "latent variable" model allows us to predict the proportion of *populism*, *nationalist*, and other ideological features for each party in the dataset; clustering analysis finds at least one cluster that comprises most of the known right-wing, nationalist, and populist parties. Second, we create a new variable in the dataset named *populism* by expert opinion, a variable we then use as a dependent variable in a recursive partitioning model. The outcome of the model shows a high degree of accuracy in predicting three classes (non-populist parties, left-populist, and right-populist). Perhaps the reason for this high level of accuracy could be an over-fitting of the model. The recursive partitioning model, however, even being over-fit, is useful because it helps us identify the variables that are instrumental in classifying the parties.

Overall, our analysis is descriptive in nature given the nature of the CHES 2017 data, but it sheds light on the similarities and differences between European parties. It reveals, surprisingly sometimes, less understood features of some parties; it provokes reflection on the various shades of gray surrounding some political parties that are, otherwise, in the very spotlight of today's European politics.

References

Breiman L, Friedman JH, Olshen RA, Stone CJ (1983) Classification and regression trees. In *CHES 2017*.

Chapel Hill Expert Survey – CHES (2017) *Codebook*. Available at: https://tinyurl.com/y8 cdryjg.

Honaker J, King G, Blackwell M (2011) Amelia II: A program for missing data. *Journal of Statistical Software* 45(7): 1–47.

Hooghe L, Bakker R, Brigevich A, De Vries C, Edwards E, Marks G, et al. (2010) Reliability and validity of the 2002 and 2006 chapel hill expert surveys on party positioning. *European Journal of Political Research* 49(5): 687–703.

Jamshidian M, Jalal S (2010) Tests of homoscedasticity, normality, and missing completely at random for incomplete multivariate data. *Psychometrika* 75(4): 649–674.

Mohan K, Pearl J (2014) On the testability of models with missing data. In S Kaski, J Corander (eds.), *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics* (volume 33), 643–650. Reykjavik, Iceland: PMLR.

Polk J, Rovny J, Bakker R, Edwards E, Hooghe L, Jolly S, et al. (2017) Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 chapel hill expert survey data. *Research & Politics* 4(1): 1–9.

Revelle W (1979) Hierarchical cluster analysis and the internal structure of tests. *Multivariate Behavioral Research* 14(1): 57–74.

Revelle W (2018) *Psych: Procedures for psychological, psychometric, and personality research*. Evanston, Illinois: Northwestern University.

Rosseel Y (2012) Lavaan: An R package for structural equation modeling. *Journal of Statistical Software* 48(2): 1–36.

Therneau T, Atkinson B (2019) *Rpart: Recursive partitioning and regression trees*. Available at: https://CRAN.R-project.org/package=rpart.

Tierney N, Cook D, McBain M, Fay C (2018) *Naniar: data structures, summaries, and visualisations for missing data*. Available at: https://CRAN.R-project.org/package= naniar

Table 1. *Descriptive Statistics for Observed Variables*

Variable	Mean	SD	Min	Max	Missing
multicult_dissent	2.53	2.25	0	10	28.41
ethnic_minorities	4.90	3.18	0	10	23.75
immigrate_dissent	2.74	2.50	0	10	23.25
eu_budgets	3.51	1.96	1	7	21.96
multiculturalism	5.86	3.09	0	10	20.71
multicult_salience	6.13	2.74	0	10	19.79
eu_asylum	4.23	2.18	1	8	18.50
eu_dissent	2.47	2.33	0	10	17.90
corrupt_salience	5.09	3.28	0	10	16.14
people_vs_elite	5.11	2.99	0	10	16.14
galtan_new_clear	7.17	2.52	0	10	15.64
members_vs_leadership	6.76	2.54	0	10	15.50
galtan_new_salience	6.69	2.47	0	10	14.85
immigra_salience	6.45	2.60	0	10	13.51
immigrate_policy	5.66	3.14	0	10	12.82
galtan_salience	6.49	2.49	0	10	12.04
antielite_salience	4.92	3.34	0	10	11.53
galtan_clear	7.31	2.38	0	10	11.02
galtan_new	5.34	3.09	0	10	10.61
lrecon_clear	6.20	2.71	0	10	10.52
lrecon	4.94	2.69	0	10	9.64
position	4.74	2.07	1	7	9.55
galtan	5.08	3.17	0	10	9.36
lrgen	5.37	2.69	0	10	7.20
eu_salience	6.29	2.50	0	10	6.83
lrecon_salience	6.24	2.61	0	10	4.43

Table 2. *Missing Values by Country*

country_name	nMissing	nValsByCountry	prMiss
it	1538	4680	32.9
por	492	2496	19.7
ge	723	3744	19.3
nl	888	5070	17.5
hu	760	4576	16.6
esp	563	3640	15.5
uk	471	3276	14.4
sk	716	5200	13.8
pl	701	5148	13.6
swe	588	4680	12.6
fr	384	5200	7.4
gr	192	3042	6.3
est	115	2340	4.9
cz	142	3276	4.3

Table 3. *Top Twenty Parties by Missing Values*

Party	Party ID	Missing (%)
VdA(UV)	847	87.9
LKR	350	76.4
PIRAT	1611	72.0
SVP	827	71.3
BLAU	351	61.8
Siet	2816	55.8
MPT	1209	51.9
PDR	1251	48.1
CD	843	42.8
50PLUS	1020	40.5
SF	1150	39.8
CC	517	39.6
TR	2613	38.5
MLP	2313	36.9
AP	852	36.7
Denk	1050	36.2
SI	850	35.6
EA/EH Bildu	507	35.2
FI	1612	33.3
SMK-MKP	2804	33.3

Table 4. *Structural Equation Model Coefficients*

term	estimate	std.error	statistic	p.value
populism = ~ people_vs_elite	1.000	0.000	NA	NA
populism = ~ antielite_salience	1.432	0.132	10.873	0
populism = ~ corrupt_salience	0.734	0.109	6.729	0
nationalism = ~ immigrate_policy	1.000	0.000	NA	NA
nationalism = ~ eu_asylum	-0.854	0.055	-15.433	0
nationalism = ~ ethnic_minorities	0.943	0.043	21.793	0
nationalism = ~ multiculturalism	1.016	0.028	35.943	0
proEU = ~ position	1.000	0.000	NA	NA
proEU = ~ eu_budgets	0.952	0.047	20.277	0
left_right = ~ lrgen	1.000	0.000	NA	NA
left_right = ~ lrecon	0.705	0.057	12.273	0
left_right = ~ lrecon_clear	-0.248	0.070	-3.520	0
left_right = ~ lrecon_salience	-0.258	0.070	-3.665	0
ideology = ~ galtan	1.000	0.000	NA	NA

Table 5. *Predicted Latent Variables in the Structural Equation Model*

party	populism	nationalism	proEU	left_right	ideology
PiS	0.939	0.937	0.159	0.991	0.906
UKIP	0.938	0.942	0.030	0.896	0.937
DLF	0.935	0.943	0.031	0.915	0.924
AfD	0.934	0.941	0.066	0.880	0.908
Kotleba LSNS	0.917	0.949	0.031	0.996	0.950
ANO	0.911	0.758	0.275	0.114	0.590
XA	0.911	0.950	0.031	0.999	0.951
FN	0.909	0.950	0.032	0.999	0.947
Insoumis	0.901	0.164	0.078	0.053	0.163
PVV	0.897	0.948	0.032	0.996	0.944
Fidesz	0.892	0.935	0.096	0.995	0.948
Korwin	0.891	0.940	0.032	0.985	0.916
SD	0.888	0.941	0.047	0.739	0.928
FvD	0.881	0.942	0.033	0.965	0.906
M5S	0.880	0.630	0.174	0.290	0.467
KKE	0.866	0.261	0.036	0.004	0.353
Podemos	0.861	0.094	0.440	0.216	0.069
EKRE	0.855	0.947	0.073	0.959	0.950
Kukiz	0.846	0.904	0.143	0.875	0.899
SYRIZA	0.845	0.096	0.636	0.110	0.073
Sme Rodina	0.841	0.918	0.141	0.450	0.901
BLAU	0.807	0.873	0.160	0.896	0.825
EVE	0.798	0.593	0.657	0.665	0.465
ANEL	0.797	0.876	0.276	0.845	0.900
Greens	0.784	0.089	0.845	0.117	0.051
OLaNO-NOVA	0.782	0.827	0.330	0.506	0.739
Razem	0.780	0.081	0.796	0.051	0.068
PDR	0.780	0.451	0.432	0.339	0.515
SF	0.774	0.307	0.675	0.063	0.355
JOBBIK	0.767	0.933	0.139	0.948	0.907
Lab	0.765	0.320	0.221	0.100	0.316
Pirates	0.764	0.270	0.498	0.339	0.128
EA/EH Bildu	0.759	0.098	0.428	0.045	0.091
LN	0.754	0.929	0.074	0.968	0.905
LMP	0.721	0.259	0.692	0.426	0.133
EK	0.708	0.424	0.673	0.316	0.583
Linke	0.707	0.212	0.427	0.055	0.212
PIRAT	0.699	0.252	0.333	0.563	0.208
PCF	0.694	0.218	0.189	0.014	0.214
LREM	0.677	0.351	0.895	0.285	0.255

Table 6. *Number of Parties by Category and Cluster*

	1	2	3	4
0	22	2	35	25
1	6	1	2	5
2	4	21	5	1

Table 7. *Variable Means by Cluster, in Percentage, Highest of Cluster 2*

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster 4
ethnic_minorities	33.6	93.0	56.6	16.1
multiculturalism	34.2	92.7	62.1	11.8
galtan_new	35.5	92.6	61.9	11.4
immigrate_policy	32.0	92.6	63.5	12.1
galtan	35.2	91.9	62.7	11.7
immigra_salience	17.9	91.7	48.2	45.0
multicult_salience	21.0	89.5	48.0	45.0
lrgen	24.3	89.4	66.5	18.5
galtan_new_salience	19.3	89.0	34.6	67.0
antielite_salience	59.2	87.8	22.8	45.1
galtan_new_clear	15.8	87.2	36.1	72.3
members_vs_leadership	38.0	84.4	59.0	20.6
galtan_salience	15.1	81.6	43.7	70.3
people_vs_elite	66.5	76.6	17.6	61.6
corrupt_salience	54.1	64.4	31.7	60.6

Table 8. *Variable Means by Cluster, in Percentage, Highest of Cluster 4*

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster 4
eu_asylum	53.3	6.4	55.7	82.5
galtan_clear	16.3	77.4	38.9	78.1
position	48.0	7.9	68.2	69.7
eu_budgets	52.1	9.5	65.4	67.0
lrecon_clear	41.4	23.8	60.2	66.5
eu_salience	36.2	59.2	47.5	60.1

Table 9. *Contingency Table Comparing Predicted vs. True Categories*

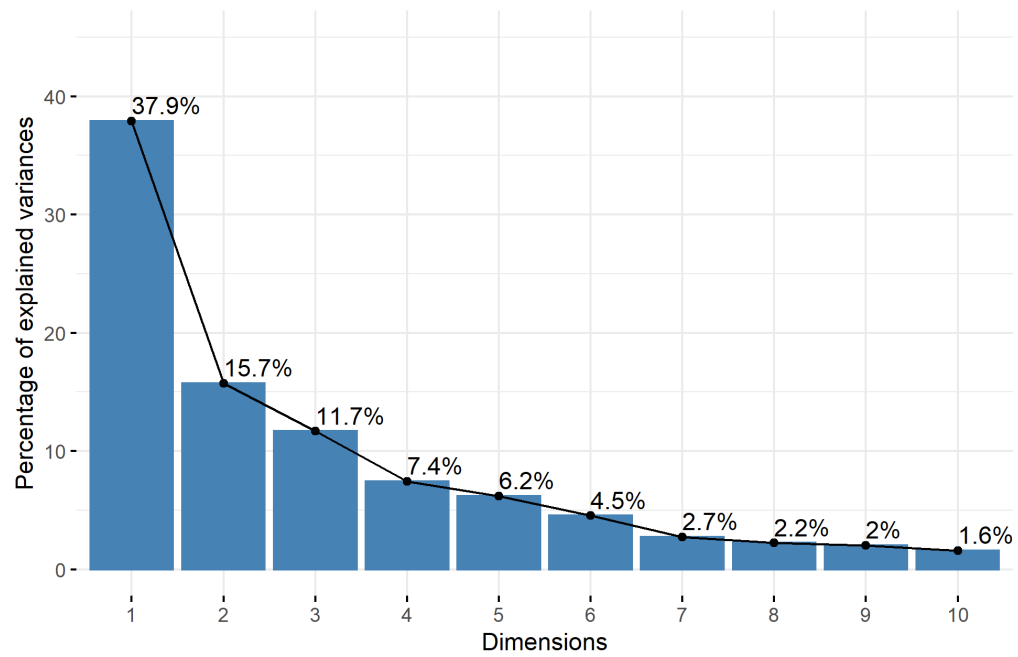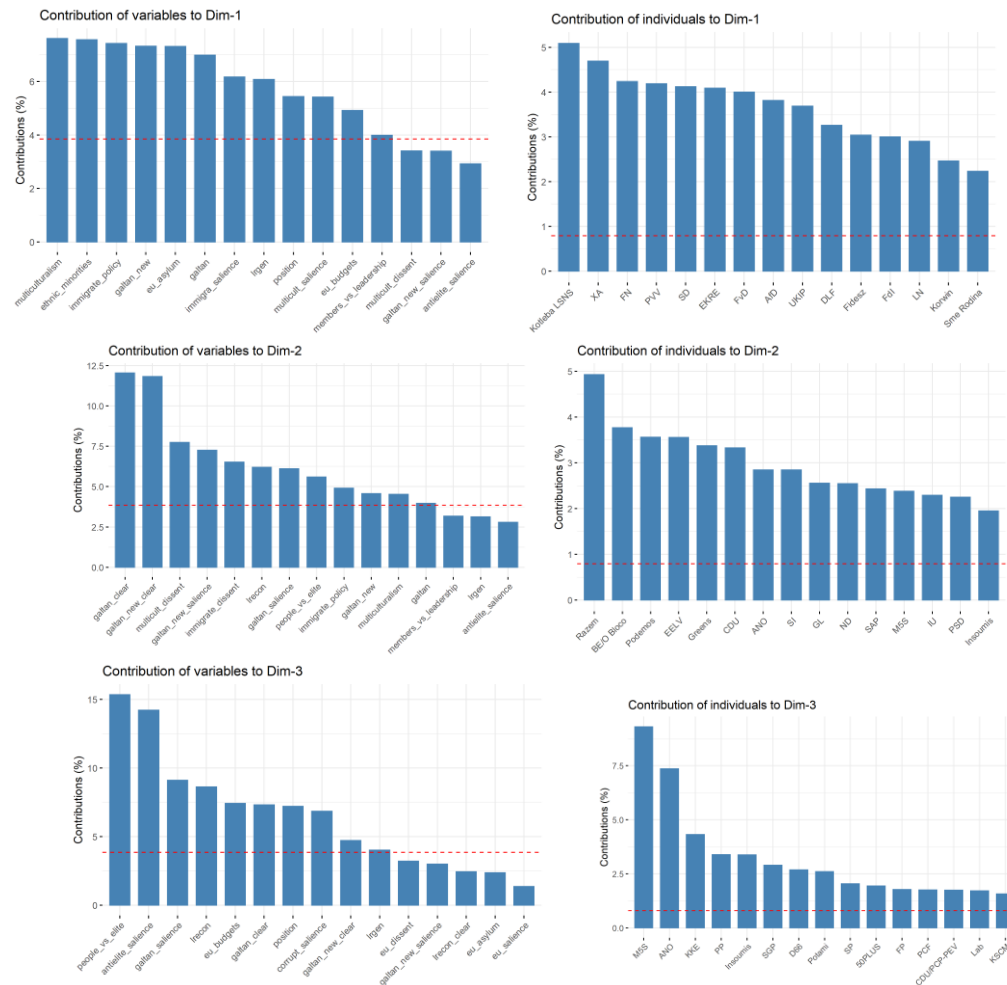|	0	1	2
0	82	1	1
1	5	9	0
2	3	0	28

Figure 1. *Scree Plot for Factor Analysis*

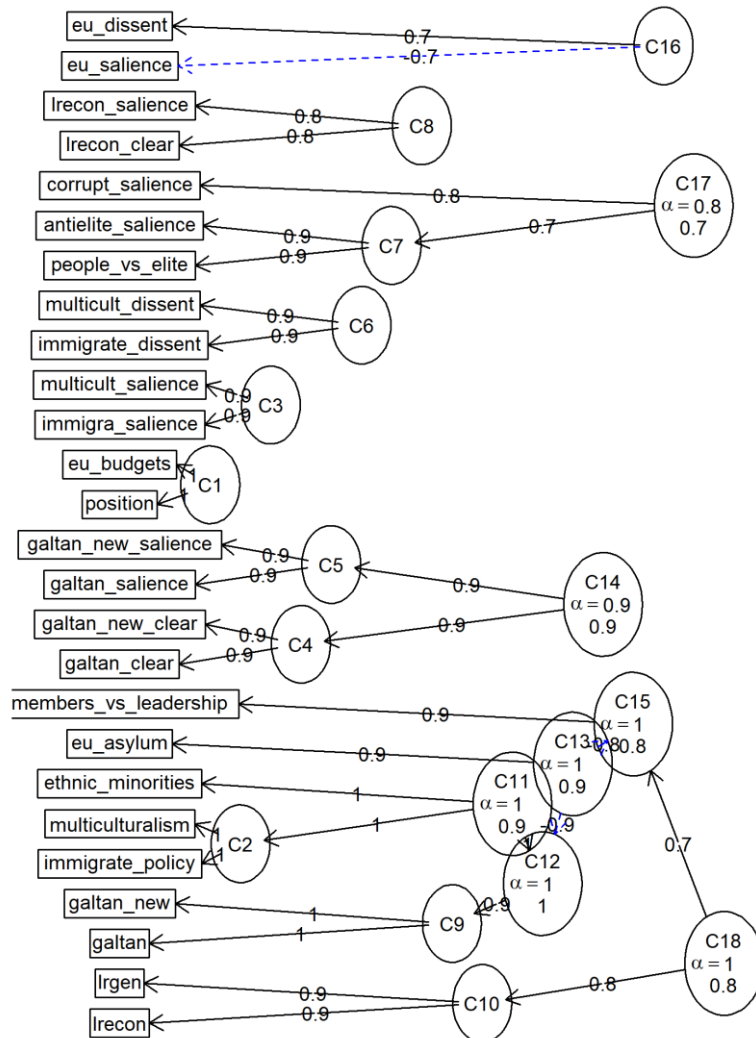**Figure 2.** *Contributions of Variables and Parties to Factors*

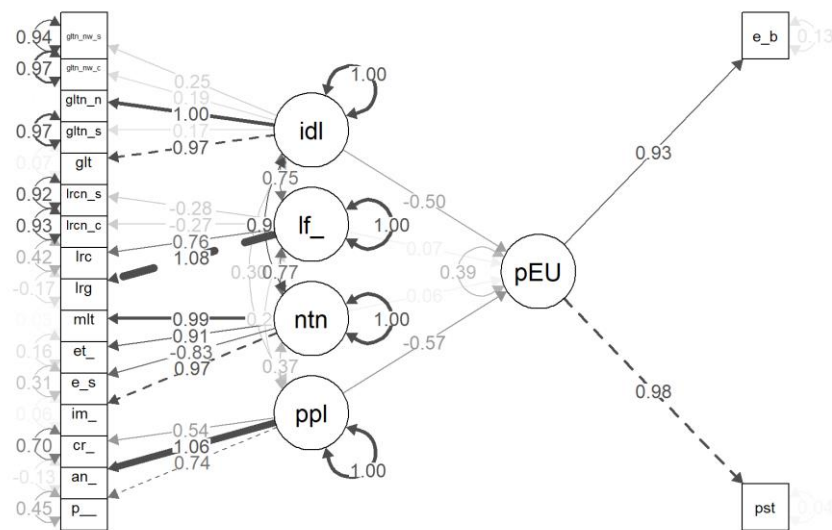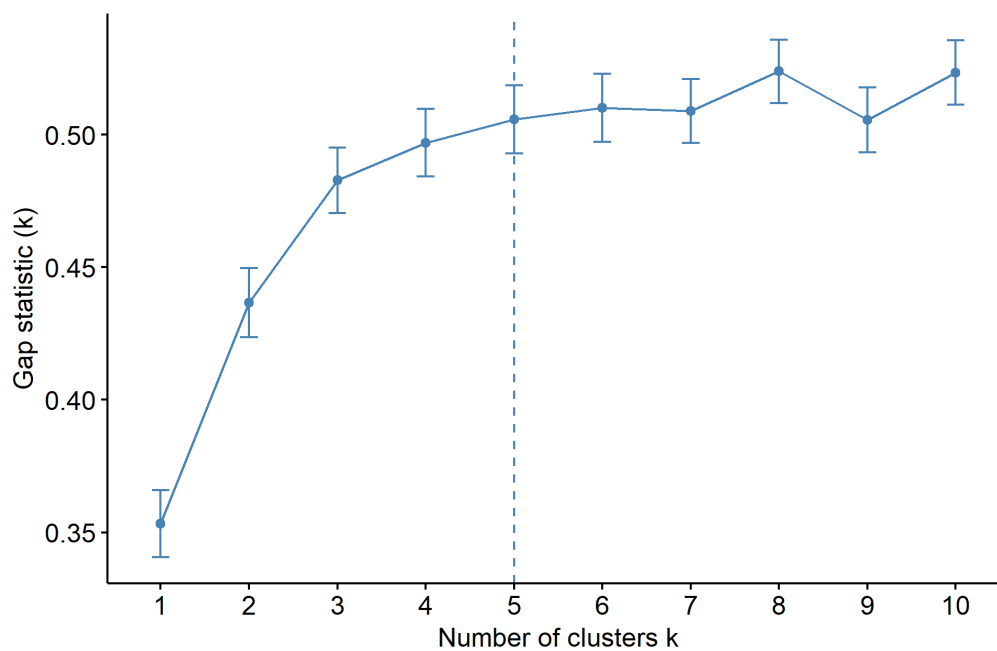**Figure 3.** *A Cluster Analysis of Variables*

Figure 4. *Graph Showing the Estimated Model*

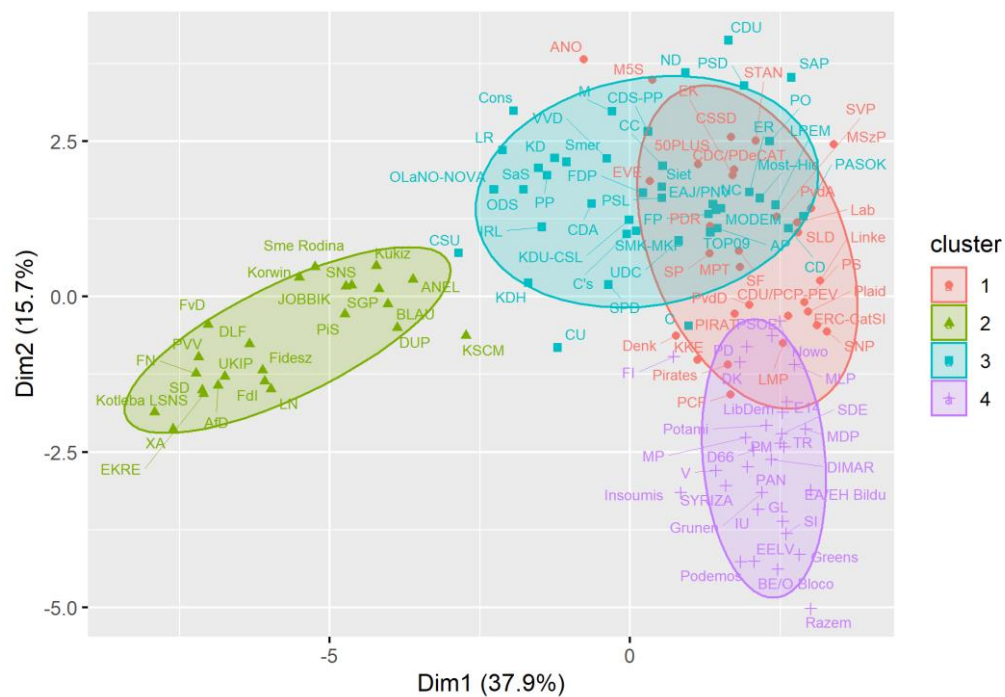**Figure 5.** *Optimal Number of Clusters Using the Gap Statistic Method*
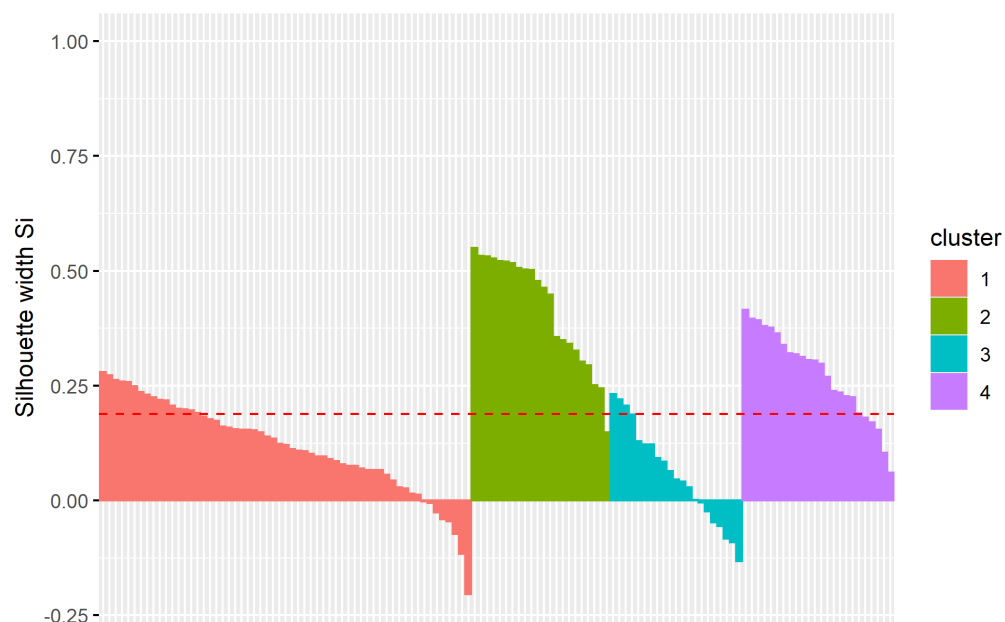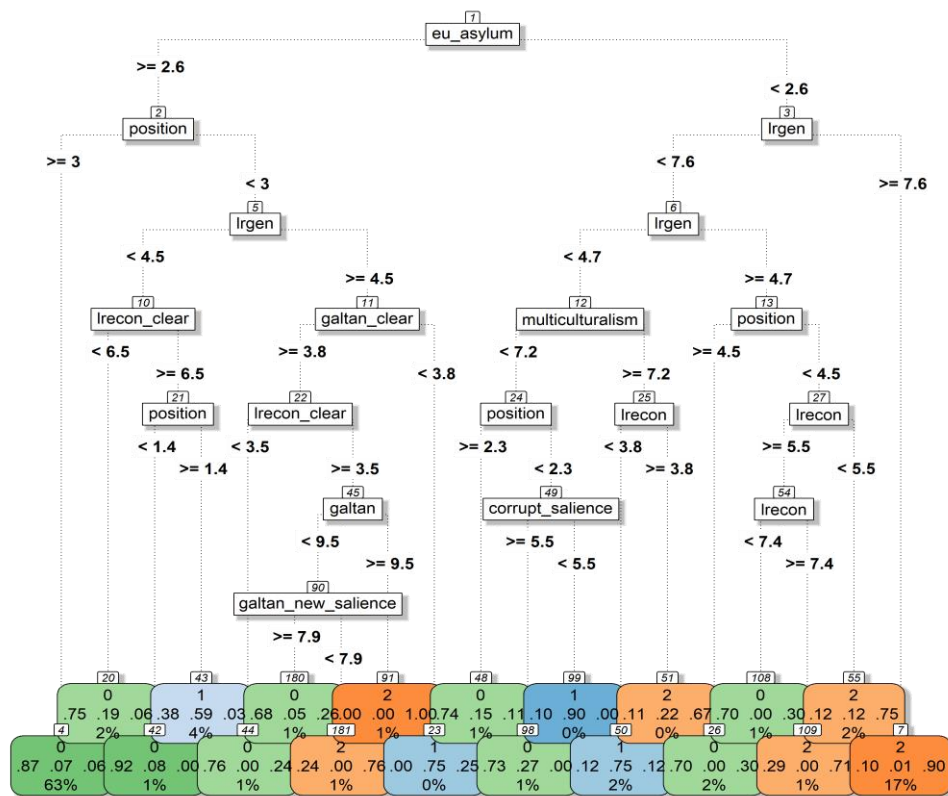
Figure 6. *Clusters of European Parties*

Figure 7. *Silhouette Diagram with Four Clusters*

Figure 8. *A Simplified Recursive Partitioning Diagram*

Figure 9. *Amore Complex Recursive Partitioning Diagram*