

The following section will introduce the data and its sources. We use the Chapel Hill Expert Survey (CHES) on European party position in order to construct our dependent variable of *Support for Populist Parties* (i.e. Support for Progressive or Traditionalist Populist parties) along with individual level data from the European Social Survey (ESS) (Section 3.1). The following subsection operationalizes our hypotheses (cultural and economic explanations for populism) and subsequently a description of the used control variables is given (Section 3.2). Following this, the statistical methodology is explained (Section 3.3) and a short examination of descriptive statistics takes place (Section 3.4). Lastly, the results of estimated multinomial logistic regression models are reported and examined for their implications regarding the research hypotheses (Section 3.5).

Data & Operationalization

The CHES dataset contains information on the positions of XXX political parties in 40 European countries on European and national policy issues in the time range between 1999 and 2014. This makes the CHES data suitable for identifying the ideological party positions that can be classified as progressive and traditionalist populism within the European context.

As a first step, we selected two variables that are in line with our minimalistic definition of populism. They will be used to construct the an Establishment - Anti-Establishment Axis.

Establishment - Anti-Establishment Axis

Populism, as it is conceptualized in this study, is characterized by two main features: a disdain for the established elites that supposedly exploit the *pure* and *little* people and an opposition to the effects of globalization that brings cultures and economies closer together at the expense of the (local) working class.

Anti-Elite Sentiment

Anti-Elite Sentiment is measured with the 11-point scale (0-1) variable *antielite_salience* that indicates the salience of anti-elite rhetoric within a given party. This corresponds with Mudde and Kaltwasser's concept of populism where the "corrupt elite" is pitted against the pure people (M/K 2017: 12).

- *Salience of anti-establishment and anti-elite rhetoric*
 0. Not important at all
 1. Extremely important

Euroskepticism

Euroskepticism (*position*¹) will be used as a proxy variable for anti-globalization. Populists are consistently opposed to the European integration process, albeit for different reasons.

- *Overall orientation of the party leadership towards European integration*
 1. Strongly opposed
 2. Strongly in favor

Progressivism - Traditionalism Axis

Next, we try to identify the value cleavage between progressivism and traditionalism.

This value cleavage depicted divides *progressives*, who favor progressive social values, promote liberal lifestyles and acceptance of homosexuality, civil liberties and multiculturalism from *traditionalists* who take the opposite stance on all these positions. The following Variables have been selected in order to distinguish between progressive and traditionalist populism.

GAL-TAN

¹The Euroskepticism variable has been recoded so that higher values indicate higher opposition to European integration.

GAL-TAN is a new politics dimension invented by @hooghe2002does. The capital letters are abbreviations for a scale that is supposed to capture the new fault lines in European politics and they stand for *Green-Alternative-Libertarian* (GAL) and *Traditional-Authoritarian-Nationalist* (TAN) respectively.

- *Position of the party [...] in terms of their views on democratic freedoms and rights. “Libertarian” or “postmaterialist” parties favor expanded personal freedoms, for example, access to abortion, active euthanasia, same-sex marriage, or greater democratic participation. “Traditional” or “authoritarian” parties often reject these ideas; they value order, tradition, and stability, and believe that the government should be a firm moral authority on social and cultural issues (galtan).*
 0. Libertarian/Postmaterialist
 1. Center
 2. Traditional/Authoritarian

Social Lifestyle

The acceptance of different lifestyle is a phenomena that consistently splits traditionalists from progressives. While progressives push for the acceptance of non-traditional social lifestyles traditionalists see this push as undermining very fabric of society.

- *Position on social lifestyle (e.g. homosexuality) (sociallifestyle).*
 0. Strongly supports liberal policies
 1. Strongly opposes liberal policies

Civil Liberties

While progressives favor civil liberties and rehabilitation of criminals into society, traditionalists favor tough measures can serve as a deterrence, even at the expense of civil liberty.

- *Position on civil liberties vs. law and order (civlib_laworder).*
 0. Strongly promotes civil liberties
 1. Strongly supports tough measures to fight crime

Multiculturalism

Traditionalists usually see a looming threat from immigrants from different countries, especially when they come from non-European countries, so they favor their complete assimilation into the host country. Progressives on the other hand understand diversity as strength and favor multicultural society without assimilation.

- *Position on integration of immigrants and asylum seekers (multiculturalism vs. assimilation) (multiculturalism).*
 0. Strongly favors multiculturalism
 1. Strongly favors assimilation

Left-Right Scale

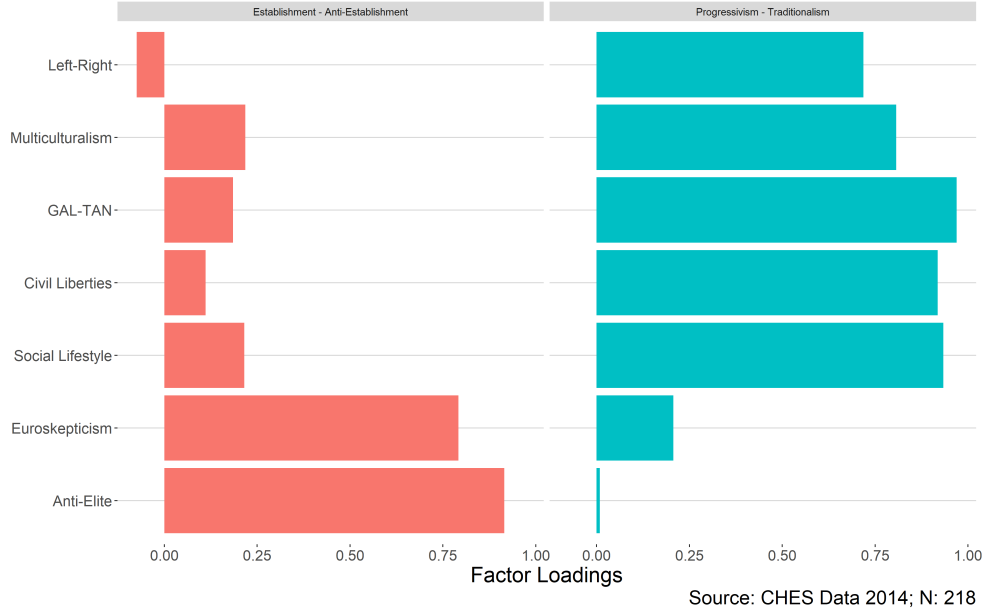
Lastly, a general left-right scale is added to this dimension. While our definition of the Progressive-Traditionalist Axis is mostly based on value differences, it's not just that. Party affiliation with a set of ideas matters as well and therefore we also include a measure of ideology through this scale.

- *Position of the party [...] in terms of its overall ideological stance (lrgen).*
 0. Extreme left
 1. Center
 2. Extreme right

Having selected the variables, a maximum likelihood factor analysis with varimax rotation is conducted in order to estimate whether our proposed dimensions are being measured by the relevant variables.

Based on the Kaiser-Criterion, two distinct dimensions are extracted explaining a total variance of 0.78%. The extracted scales are then summed into two scales *Establishment vs. Anti-Establishment* and *Progressive vs. Traditionalism*, each standardized from 0 to 100 points to facilitate easy interpretation.

Figure 1: Factor Analysis of CHES Data



As a next step, we want to extract our traditionalist and progressive populist parties. This will be done with the help of *k-means clustering*. K-Means clustering is a very popular form of unsupervised machine learning that helps with classification problems. The algorithm produces a k number of clusters (classification groups), where k is specified by the researcher. K-Means clustering estimates a centroid (i.e. a center) for each group that has the highest *intra-class similarity* within a given cluster (i.e. smallest distance from the centroid) and the lowest *inter-class similarity* with other specified cluster (i.e. maximized distance from other cluster centroids). The resulting clusters have minimal *within cluster variation* and a maximum of *between cluster variation*.

The classical algorithm for k-means clustering is the Hartigan-Wong algorithm [-@hartigan1979algorithm], where the the total within-cluster variation is defined as the sum of squared (Euclidean) distances between data points and the corresponding centroid:

$$W(C_k) = \sum_{x_i \in C_k} (x_i - \mu_k)^2$$

Where x_i is a data point belonging to the cluster C_k and μ_k is the mean of values that are classified as cluster C_k (centroid).

Each data point x_i is classified as a specific cluster so that the sum of squares euclidian distance of the observation to their assigned cluster centroid μ_k is minimized.

$$Within - SS = \sum_{k=1}^k W(C_k) = \sum_{k=1}^k \sum_{x_i \in C_k} (x_i - \mu_k)^2$$

Finally, the total within-cluster sum of square (Within-SS) measures the appropriateness of the clustering based on how much it can be minimized.

Now the algorithm can come into use. As a first step, the algorithm randomly selects k points from the given data that will be used as centroids. Next, two steps will be repeated iteratively until convergence is achieved:

1. Cluster Assignment Step

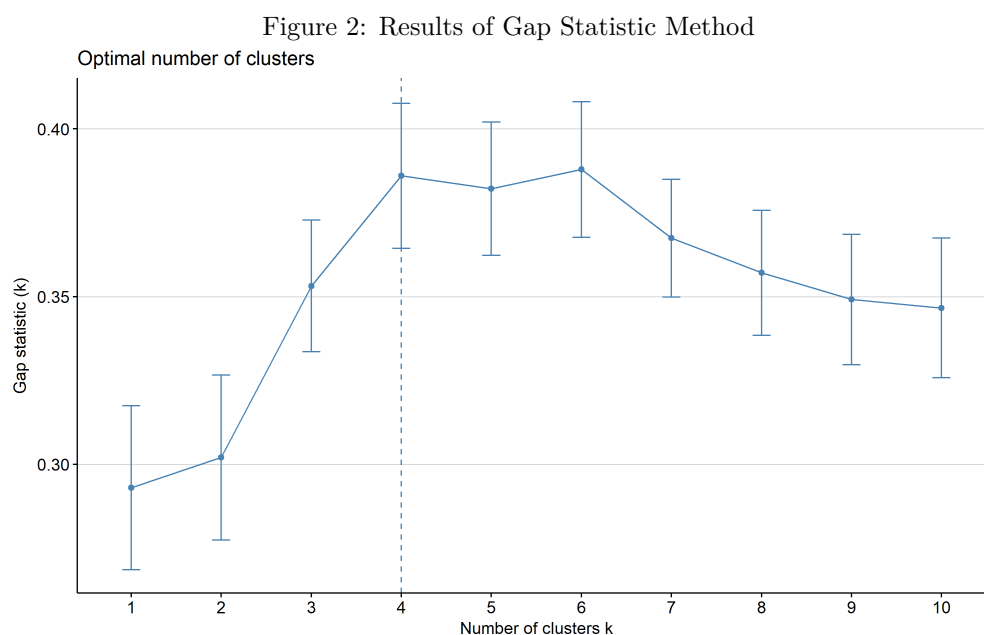
Using Euclidean distance, the distances to the centroid are calculated and the data points are classified to be part of a cluster.

2. Centroid Update Step

In this step, a new centroid is calculated based on the estimated clusters. These centroids serve as new starting point and all data points are reassigned.

The algorithm converges when the clusters do not change in the next iteration (the last two iteration produce the same clusters with the same data points within them).

Finally, the two scales *Establishment vs. Anti-Establishment* and *Progressive vs. Traditionalism* are handed over to the K-means clustering algorithm. Based on the Gap Statistic method [cf. tibshirani2001estimating], four clusters are suggested as the optimal number of clusters.[^Elbow and average silhouette method also suggest four clusters as optimal.] Figure shows the results of the gap statistic that clearly indicate four clusters.



The four clusters that are estimated with the help of the k-means algorithm can be named as traditionalist and progressive populist parties as well as their two establishment counterparts (establishment progressives and traditionalists). Together with the clustering method, the two dimensions can be used to visualize the ideological position of each European party and its classification, which is illustrated in Figure .

TODO:

Clustering Methodology

Top Left shows progressive populists such as:

DESCRIBE THEM

Top Right shows traditionalist populists such as:

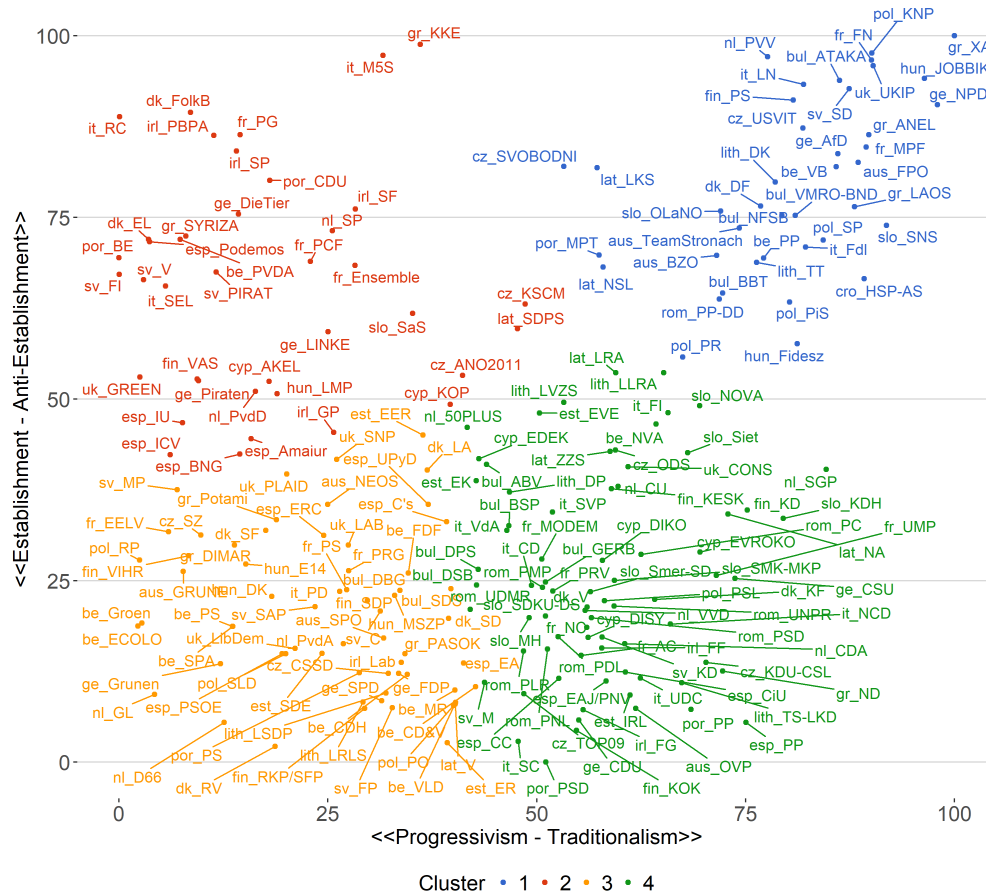
DESCRIBE THEM

Bottom left and bottom right shows progressive and traditionalist establishment parties:

DESCRIBE THEM

Our measure of populism correlates well with different similar classification methods.

Figure 3: Party Alignment of European Parties
Classification of European Parties
 Based on K-Nearest Neighbour Clustering



Source: CHES Data 1999 - 2014

(TODO:) CORRELATION WITH OTHER MEASURES (TODO:) ONLY INCLUDE PARTIES THAT ARE USED LATER ON

A full list of used parties as well as their respective affiliations can be found in the appendix.

(TODO: TABLE that shows the individuals scores)

Dependent Variable: Support for Populist Parties

TODO:

The European Social Survey data is such and such.

- from when to when
- which countries
- how many parties
- actual operationalization

After the successful classification, we combine the clusters from the CHES data with the *European Social Survey* (ESS) Round 5 – 8. We decided to use only these dates, as we expect the years after the European

financial crisis (2008-09) to be more homogenous in terms of populism. Two variables will be used to measure our dependent variable *Support for Populist Parties*:

1. *What party did you vote for in the last national election?*
2. *Which party is closest to your views?*

A respondent that either voted for or indicated that they feel closest to a specific party, will be classified as either supporting a progressive or traditionalist populist or an establishment party, based on the clusters generated by the K-Means algorithm. If it is the case that a person voted for a party but felt close to a different party, we decided to classify said person as a supporter of the party that it felt most close to (thus ranking their vote as less indicative of their support). This is based on the assumption that many voters have an incentive to vote strategically and they might end up voting for an establishment party even though they actually support a populist party (TODO: CITATION).

After the merging is completed we are left with XXX respondents from 24 European countries.

TODO: Descriptives?

Independent Variables: Cultural and Economic Explanations

TODO: Description of INDEPENDENT VARIABLES?

Measuring this or that (just check Inglehart on how they have done it)

The models include standard social and demographic controls, including sex, age, education, and ethnicity. Economic inequality was monitored through experience of unemployment, households dependent upon social benefits (excluding pensions) for their main source of income, urbanization, and subjective feelings of income insecurity. Principal Component Factor analysis with Varimax Rotation was used to determine the dimensionality of a range of cultural items which were included in all waves of the ESS and which were expected to be particularly salient for the division between Populist and Cosmopolitan Liberal values. Five values scales were produced through this process, including scales on attitudes towards immigration, trust in global governance, trust in national governance, authoritarian values, and left-right ideological self-placement. The value scales were each standardized to 100-points, for ease of comparison. All models were checked by tolerance tests to be free of problems of multicollinearity. The inclusion of items consistently asked across all rounds of the ESS maximized the size of the pooled sample of populist voters, and therefore strengthened confidence in the reliability of the results, although unfortunately it also restricted the full range of items which ideally could be included, for example concerning gender equality. The descriptive means and standard deviations of all the items are presented in Appendix C.

Statistical Methodology

Here comes a description of the multinomial model (might need to change the name of the title)

Descriptives

TODO:

- Populist Parties by Region
- Populist Parties by Country (map)

Descriptive Statistics and Exploratory Evaluation of the Hypotheses

Multinomial Logistic Regression

Here comes the Analysis Part