The following section introduces the data and its sources. We use the Chapel Hill Expert Survey (CHES) on European party positions in order to construct our dependent variable of *Support for Establishment/Populist Parties* (i.e. Support for Progressive or Traditionalist Populist parties) along with individual level data from the European Social Survey (ESS) (Section). The section further operationalizes our hypotheses (cultural and economic explanations for populism) and give a description of the used control variables. Following this, the statistical methodology is explained (Section).

Data & Operationalization

The CHES dataset contains information on the positions of 365 political parties in 40 European countries on european and national policy issues in the timerange 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. Since parties have changed over time, for the purposes of our analysis we only use the latest available data from 2014.

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 cleavage 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 of all these positions. The following variables have been selected in order to distinguish between progressive and traditionalist populism.

GAL-TAN

[1]The 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 abbrevations 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 affilliation 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 producess 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* [cf. @friedman2001elements 509 ff.].

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 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 repated 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.

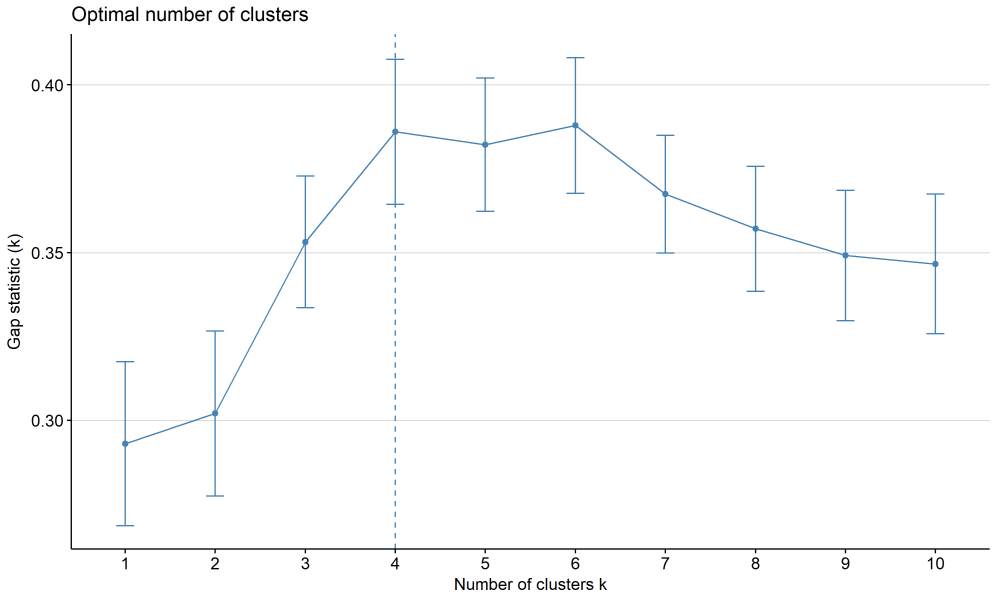
2. Centroid Update Step

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.[2] Figure 2 shows the results of the gap statistic that clearly indicate four clusters.

Figure 2: Results of Gap Statistic Method



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 3.

In order to validate the clusters, let's pick some examples and see whether the classified parties fit our theoretical expectations.

The *top-left quadrant* shows progressive populists such as:

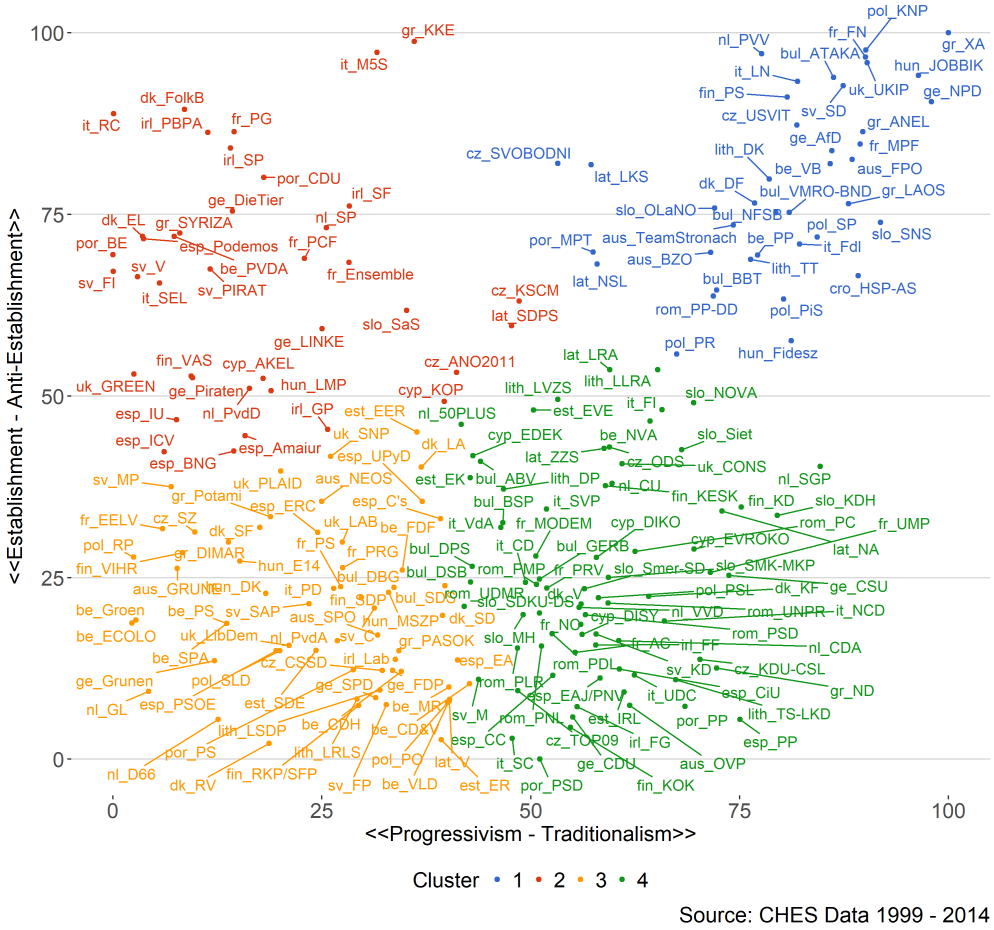
Syriza

After winning the parliamentary elections in January 2015, SYRIZA (*Coalition of the Radical Left*) attempted to carry out the difficult balancing act between extreme left-wing positions, opposition to EU imposed austerity and yet supposedly pro-European commitments. In social policy, Syriza is particularly committed to the socially disadvantaged in society with policies such as guaranteeing that unemployed, homeless and low-income people should be allowed to use the health facilities free of charge or that family reunification should be made easier for individuals with a migration background.

Podemos

[2]Elbow and average silhouette method also suggest four clusters as optimal.

Figure 3: Party Alignment of European Parties



In the general elections held on December 20, 2015, the left-wing political party Podemos that emerged from a protest movement obtained 20.68% of the votes and 69 deputies in the whole of the State. The Spanish "Indignados", the "indignants" of the Podemos movement, practice critique of globalization and capitalism often symbolized in the overarching EU bureacracy. Among other measures, they defend abortion, want to stop house evictions, suppress church privileges, promote renewable energies and are in favor of curbing nuclear energy. With regard to political parties, they propose to stop gauging, reduce subsidies and expand restrictions on connections between politicians and companies.

Red–Green Alliance (Denmark)

The Red–Green Alliance (Enhedslisten) was formed as a collaboration between the Left Socialists (VS), the Danish Communist Party (DKP) and the Socialist Labor Party (SAP) in 1989. During the last parliamentary elections in 2015, the Red–Green Alliance gained 7,8% of the popular vote. Enhedslisten does not stand in European elections, but supports Folkebevægelsen mod EU (*Popular Movement against the EU*), a heavily anti-EU political party that only competes for the European elections. The party attaches great importance to combating social inequality and poverty, as well as advocating strengthening and expanding the welfare state. Politically, the party is in favor of more space for all forms of diversity, including gender, sexuality, disability and ethnic background.

The *top-right quadrant* shows traditionalist populists such as:

AfD

The alternative for Germany (AfD) is a political party founded in 2013 in Germany. As of 2014, it gradually moved into 14 state parliaments and in the 2017 general election, the AfD received 12.6% of the vote and thus became the third strongest force and the strongest opposition party in the German Bundestag. Regarding the EU, they have been in favor of renationalization of policies that are currently situated in the EU. The AfD represents conservative-antifeminist positions in gender politics and rejects gender equality policies and relies thereby on Christian fundamentalist and nationalist ideas. According to the AfD, Islam does not belong to Germany. In particular, the party calls for a ban on minarets and the face veiling.

Front National

UKIP

The *bottom-left and bottom-right quadrants* show progressive and traditionalist establishment parties:

German SPD and CDU, Labour Party and Conservatives UK, SOME FRENCHIES

Given that the distinction between kinds of establishment parties is not of greater interest to us, we will merge progressive and traditionalsit establishment parties into a single establishment party group.

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

Dependent Variable: Support for Establishment/Populist Parties

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 this timerange, because in this way we capture the years after the European financial crisis (2008-09). Two variables will be used to measure our dependent variable *Support for Establishment/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 alghorithm. 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.

After merging of the datasets is completed, we are initially left with 130155 respondents from 22 European countries. However, given that there are many missing values for respondents who gave no indication on whether they support a political party (32.99%), we are ultimately left with 87238 cases for the purposes of our descriptive analysis and 68403 for the multinomial logistic regression analysis (where all missing values were deleted listwise).[3]

Independent Variables: Cultural and Economic Explanations

Next, the hypotheses will be operationalized with corresponding ESS variables. *Economic deprivation* will be captured with two variables: *Economic Insecurity* (4-point scale) and the dummy variables *Unemployment*, and *Welfare*. The *cultural value hypothesis* will be measured with an index for *Anti-Immigration Sentiment* and four *Schwartz Human Value* dimensions, all ranging from 0 to 10: *Openness*, *Self-Transendence*, *Self-Enhancement* and *Conservation*, where the former two are associated with liberal values and the latter two are associated with conservative values [cf. @schwartz1994there; @schwartz2005human]. The models further

[3]Note that is certainly a high amount of missing values. Imputation could be considered in this case, however, this would be beyond the scope of this seminar paper.

include common socio-demographic control variables, for example, *Age*, *Education* and *Sex* but also includes a *Left-Right Scale*, *Religiosity*, *Government Satisfaction*, *Ethnic Minority Status*, *Trust in Global Governance* and the dummy variable *Rural vs. Urban*. Lastly, regional dummies (East, West, North and Southern Europe) and time dummies for each year will be included in the model as controls (2010, 2012, 2014 and 2016).

A more detailed description of the used variables can be requested from the authors.

Statistical Methodology

Multinomial logistic regression expands upon binary logistic regression, so that it is possible to predict three or more outcomes of a variable. Each category of the variable of interest is compared to the *reference category*, which is specified by the researcher, with the consequence that estimated parameters (logits and/or odds ratios) are interpreted in reference to that category [cf. @hosmer2013applied 269].

For this example, we assume that the multinomial model is run over a categorical response category Y, coded as 0, 1, or 2, where 0 is used as reference categaory. Similar to the logistic regression model, where the the logit is understood as $Y = 1$ versus $Y = 0$, a multinomial logistic regression compares the outcomes $Y = 2$ and $Y = 1$ to $Y = 0$. Therefore, two distinct logit functions are estimated:

$$g_1(x) = ln\left[\frac{Pr(Y=1|x)}{Pr(Y=0|x)}\right]$$

$$g_2(x) = ln\left[\frac{Pr(Y=2|x)}{Pr(Y=0|x)}\right]$$

This model is almost analogous to a logistic regression model, except that the multinomial logit equations contrast each of response category Y with reference category $Y = 0$. When Y has a binary outcomethe multinomial logit model reduces to the usual logistic regression model.

The probabilities for each Y-Category is calculated as follows:

$$Pr(Y = 0|x) = \frac{1}{1 + e^{g_1(x)} + e^{g_2(x)}}$$

$$Pr(Y = 1|x) = \frac{e^{g_1(x)}}{1 + e^{g_1(x)} + e^{g_2(x)}}$$

$$Pr(Y = 2|x) = \frac{e^{g_2(x)}}{1 + e^{g_1(x)} + e^{g_2(x)}}$$

Theoretically a multilevel model would have been needed to estimate the model properly, but given that there are some countries with almost none or no populist supporters in our data, this would lead to problems. In order to still account for the hierarchical order of our data, we decided to use regional variables of Europe, based on the four classifications: Eastern, Western, Southern and Northern Europe.