Complexity in European political spaces: Exploring dimensionality with users

Garret Binding*

April 4, 2018

Abstract

Spatial descriptions of political actors as low-dimensional ideological summaries of their positions on distinct issues are commonplace in political science. Yet questions remain regarding the structure of the political space of the mass public, both regarding the number of dimensions and their content. Drawing on data generated by a European-wide voting advice application, I estimate and describe the political spaces of 24 European countries in a Bayesian multidimensional item response theory (IRT) set-up. The results are illustrative in highlighting the heterogeneity in political spaces across Europe: first, not all countries profit equally from models with an increasing number of dimensions, even though the fielded items were explicitly selected to capture three distinct dimensions. Second, while the issue structuring a dimension is defined to be equal across countries, the relationship between the remaining issues and dimensions varies considerably across countries. Finally, the countries can be clustered according to their (dis)similarity in this regard, allowing us summarize how (dis)similar political spaces are across European countries.

*binding@ipz.uzh.ch. I would like to thank Marco Steenbergen, Lukas Stoetzer, and Thomas Willi for helpful discussions and feedback.

1 Introduction

Since the seminal contribution of Converse (2006), the notion of a low-dimensional belief system which constrains the opinion of individuals regarding many political issues has been a constant in political science. While different strands of literature have highlighted distinct types of belief systems stemming from different sources such as partisanship, values, or social status (Inglehart and Klingemann, 1976; Kuklinski and Peyton, 2007), a prominent line of argument has been that belief systems are defined by ideology (Hinich and Munger, 1992, 1996) and that we can aggregate individual belief systems to shared political spaces (Benoit and Laver, 2006; Laver, 2014).

This raises questions concerning the number of ideological dimensions necessary to adequately describe a political space, the defining content of each dimension, and whether political spaces are similar across countries. Contributions such as Kriesi et al. (2006) concerning the formation of a second dimension in Western Europe or Marks and Steenbergen (2002) concerning the (non)existence of a EU-specific dimension are examples of research focusing on the first two questions, while Rovny and Edwards (2012) or Rovny and Polk (2018) attempt to understand why political spaces differ cross-nationally in Europe.

Drawing on data collected by a large-scale voting advice application (VAA; Garzia and Marschall (2014)) fielded across Europe prior to the elections to the European parliament of 2014 (Gemenis et al., 2018), I address these questions by estimating uni- and multidimensional IRT models for 24 European countries (a unidimensional model defined by an economic dimension, a two dimensional model with an economic and a cultural dimension, and finally a three dimensional model with an addtional EU specific dimension; Bakker et al. (2012)).

Based on the comparison of the results of these models, I find that a three dimensional model including an economic, a cultural, and a EU specific dimension fits the data best and reduces the predicted error rate most in all countries.

This finding is unsurprising as the items fielded in the VAA were chosen to capture these three dimensions. However, I also find that the reduction in error rate going from lower- to higher-dimensional model is not equal across all countries. Although the items should capture three distinct dimensions, a subgroup of countries does just fine regarding their error rates when less than three dimensions are defined.

By employing an IRT model I can define the latent dimensions using a flexible a priori approach (Laver, 2014). While the relationship between a selected few of the items and the dimensions is constrained, the parameters of all other items are freely estimated. The relationship between the remaining issues and dimensions varies considerably across countries. This highlights that the political spaces across Europe differ not only in the number of dimensions necessary to adequately describe them, but also in the content of these dimensions.

Based on the item parameters linking issues to dimensions, I estimate the similarity of political spaces across European countries. This allows me to cluster the 24 countries according to whether their spaces are more or less similar to one another.

Overall, this is a largely descriptive and data-driven endeavour which remains agnostic on questions such as why the spaces of certain countries are more similar than those of others or why the same item is linked differently to the same dimension across countries. In its current version, this paper's main contribution is therefore bringing this heterogeneity center stage.

The findings can be connected to two distinct, but related strands of literature in political science. The first would connect to theories attempting to explain *why* and *how* political spaces differ across Europe. Examples of such contributions are Marks and Steenbergen (2002), Kriesi et al. (2006), Rovny and Edwards (2012), Bakker et al. (2012), or Rovny and Polk (2018).

A second more methodological perspective could focus on questions of cross-national measurement equivalence when scaling respondents. Examples of contributions in this vein are Stegmueller (2011), Davidov et al. (2011), or Davidov et al. (2014).

In the following, I discuss the IRT model, its Bayesian estimation, and data in Section 2. I present my findings regarding the number of dimensions of the political spaces, their content, and the similarity of these country-specific spaces in Section 3.

2 Model, Estimation & Data

In the following Subsections, I will describe the IRT model used to relate the items to the ideological dimensions, discuss the Bayesian estimation strategy, and briefly discuss the data.

2.1 Model

Political dimensions are latent variables, i.e. they are not directly observable. For this reason, item response theory (IRT) is ideally suited to relate the observable to the unobservable (Hambleton et al., 1991). It is established in political science as a theory and method to estimate positions of political actors on a latent scale (e.g. Martin and Quinn, 2002; Clinton et al., 2004; Treier and Hillygus, 2009). Treier and Hillygus (2009, p. 683f.) discuss the various advantages of this approach relative to other techniques such as confirmatory factor analysis or additive issue scales in more detail (see also van der Eijk and Rose, 2015, for a criticism of factor analysis).

Importantly, it outlines a parametric relationship between characteristics of an item and characteristics of a respondent regarding the probability of responding in a certain manner to an item. An item response function for a two parameter logistic model (2PL) is displayed in Figure 1. As the latent trait θ goes from negative to positive along the horizontal axis, the probability of responding with $y = 1$ increases. Respondents at $\theta = -2$ are very unlikely to choose $y = 1$, while those at $\theta = 2$ are very likely to respond with $y = 1$.

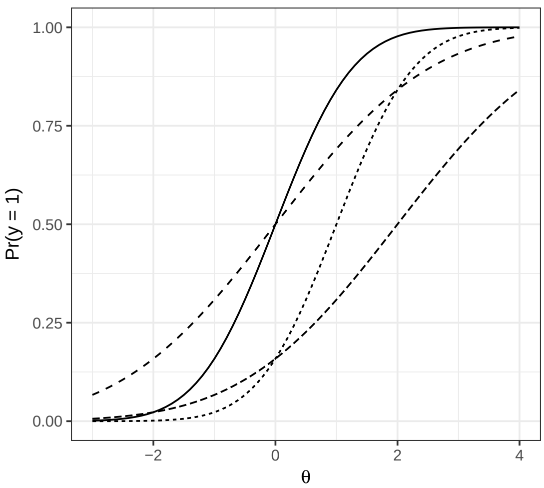


Figure 1: Binary Item Response Functions of 2PL IRT Model.

In the 2PL model, an item is defined by its difficulty and discrimination parameter. The difficulty parameter governs the point along the scale of the latent trait θ at which the probability of responding with $y = 1$ is at 0.5, i.e. the point at which a response of $y = 1$ becomes more probable than a response of $y = 0$. As the difficulty parameter changes, the response function of an item shifts along the horizontal axis. Items with high (low) absolute difficulty parameters can be seen as items where one must be more (less) extreme along the latent trait to respond with $y = 1$.

The discrimination parameter describes the slope of the function at the value of the difficulty parameter ($Pr(y = 1) = 0.5$). Steeper slopes (i.e. a higher absolute value of this parameter) indicate that the distinction between whether a respondent is on one side or another of the value of the difficulty parameter along the horizontal axis is clearer, while less steep slopes mean that the probability of responding with $y = 1$ increases more incrementally. Items with high (low) absolute discrimination parameters can be considered as more (less) polarizing with regards to the underlying latent variable.

The parameters of interest regarding the content of the latent dimensions are therefore these discrimination parameters, as they reflect the importance of issues discussed in items for the dimension. If an item is not relevant in placing a respondent along the dimension, then the discrimination parameter of this item should be indistinguishable from zero. Varying signs of these parameters across countries indicate that the relationship between an issue and the latent trait differs: while higher values of θ may imply higher probabilities of choosing right-wing response options in some countries, it may imply higher probabilities of choosing left-wing options in others (or vice versa).

As the items in this data are presented with ordinal response categories, polytomous IRT models are needed (Ostini and Nering, 2006). Drawing on Samejima (1969), a graded response model is estimated where the response of an individual $i \in 1, ..., I$ to an item $k \in 1, ..., K$ depends on an item's discrimination parameter β_k, an item's difficulty parameter α_k, and an individual's position on the latent variable θ_i (here, i's position on a political dimension), so that

$$y^*_{ik} = \alpha_k + \beta_k * \theta_i. \tag{1}$$

This model can be extended to a multidimensional set-up with d dimensions (Fox, 2010). Here, θ_i and β_k are defined as vectors in $\mathbb{R}^d$, so that

$$y^*_{ik} = \alpha_k + \boldsymbol{\beta_k \theta'_i}. \tag{2}$$

For each k, i chooses the response category $t \in 1, ..., T$ if the position y^*_{ik} is above threshold τ where the thresholds are ordered so that $\tau_1 < ... < \tau_t < ... < \tau_T$. The observed outcome (i.e. the response category chosen by the individual for an item) is given by

$$y_{ik} = \begin{cases} 1 \text{ if } y^*_{ik} < \tau_1 \\ t \text{ if } \tau_t < y^*_{ik} < \tau_{t+1} \\ T \text{ if } y^*_{ik} > \tau_T \end{cases} \tag{3}$$

Finally, the probability of i choosing t is derived using a normal ogive function as discussed in Fox (2010), so that

$$Pr[y_{ik} = 1] = \phi[\tau_t - \beta_k * \theta_i - \alpha_k] \tag{4}$$

$$Pr[y_{ik} = t] = \phi[\tau_t - \beta_k * \theta_i - \alpha_k] - \phi[\tau_{t-1} - \beta_k * \theta_i - \alpha_k] \tag{5}$$

$$Pr[y_{ik} = T] = 1 - \phi[\tau_T - \beta_k * \theta_i - \alpha_k], \tag{6}$$

where $\phi()$ is the normal CDF.

2.2 Identification, Priors & Estimation

In line with Jackman (2000, 2001, 2004) or Fox (2010), I estimate the parameters in a Bayesian set-up. There are two main reasons for this: first, it makes life easier in models with many parameters compared to maximum likelihood estimation. Second, the uncertainty in the estimates is obtained as a by-product of the Markov chain Monte Carlo (MCMC) iterations.

IRT models are not identified without additional assumptions (Rivers, 2003; Bafumi et al., 2005). Therefore, the distribution of the latent variable θ is defined to follow a standard normal distribution, so that

$$\boldsymbol{\theta_i} \sim \mathcal{N}(0, 1). \tag{7}$$

This creates two constraints per dimension, namely regarding the mean and the dispersion of the parameters, and addresses scalar invariance in the terminology of Jackman (2001) and Bafumi et al. (2005). Rotational invariance is addressed by defining the relationship between selected items and the underlying dimensions. By doing this, the values along the latent dimension obtain a more intuitive meaning, namely one in which negative (positive) values along the latent variable are associated with more left-wing (more right-wing) positions.

I define the economic dimension based on the responses to the item on the redistribution of wealth and constrain the discrimination of this item to be equal to one for the economic dimension, but zero for all other dimensions in the two and three dimensional models settings. The cultural dimension is defined by the item on abortion rights, while the EU dimension is defined by the item on whether EU membership has been a bad thing for the country or not. The discrimination parameters of these items for the other dimensions are constrained to be zero.

This means that people who disagree with economic redistribution, are against abortion rights, and disagree with the statement that EU membership has been bad for their country of reference are placed at higher values of the corresponding latent dimensions.

One benefit of this specification is that I can define the latent dimensions using a flexible a priori approach (Benoit and Laver, 2006, 2012; Laver, 2014). Prior knowledge and assumptions are employed to define probable dimensions of political spaces, while the relationships between the remaining issues and the dimensions are estimated freely.

The priors for the difficulty α_k and the discrimination β_k parameters are set to a standard normal distribution in line with previous applications (e.g. Jackman, 2004). The prior of the cutpoints τ is set so that the difference between $\tau_t - \tau_{t-1} \sim \text{lognormal}(0.5, 1)$, while the medium cutpoint is essentially fixed to 0 by a prior of $\mathcal{N}(0, 0.001)$. This was necessary because α_k and τ could otherwise move in tandem leading to multi-modal posterior distributions.

I begin with a unidimensional model characterized by the economic dimension, and then add first the cultural dimension for a two dimensional model before adding the EU dimension for a three dimensional model. These three models are estimated separately by country. While this means that questions of cross-national measurement equivalence should be less of a topic (Stegmueller, 2011), this also means that the resulting parameters cannot be directly compared across countries, but only in relation to the identifying constraints on the discrimination parameters within each country.

I estimate all models in Stan (Carpenter et al., 2016) with the No-U-Turn Sampler (NUTS) and run four chains for a total of 2000 iterations in R (R Core Team, 2017). The warmup phase is set to 500 iterations. I assess the convergence of the four chains via the $\hat{R}$ diagnostic, which is below 1.1 across all parameters, countries, and models (Gelman and Rubin, 1992).[1]

2.3 Data

The data was generated by the use of the VAA EUVox prior to the elections to the European parliament in 2014 (Gemenis et al., 2018). Essentially, VAA attempt to match voters to political actors by assessing the (dis)similarity between the opinions indicated by respondents regarding a variety of issues and those of politicians or parties. Therefore, VAA data represents a source of information on the political opinions of many. Additionally, the use of VAA is widespread with large sample sizes in most countries (Garzia and Marschall, 2014). This is an ideal condition to assess the dimensionality of the political spaces across European nations[2].

After removing rogue entries from the data, the entries of more than 400'000 users remain in the data (Andreadis (2014)). Of the remaining entries, I sampled 750 users at random per country. While it is possible to run these models with more users per country, the computational cost and duration of estimating a total of 72 models (24 countries with three models) including 750 individual, 60 item-specific, and four cutpoint parameters for each run is not negligible.

Users were asked to respond to a total of thirty items and could indicate responses on an ordinal scale with response options from 'completely agree' (= 1) to 'completely disagree' (= 5) with a 'neither agree, nor disagree' option in the middle. According to Wheatley (2016), these items were selected based on the Chapel Hill Expert Survey with the goal of capturing three distinct dimension, namely an economic dimension, a cultural dimension, and a EU dimension (Bakker et al., 2015). Additional items were included to capture

[1] So as to accelerate convergence, I set starting values for $\boldsymbol{\theta_i}$ and β_k based on the responses to the identifying items for $\boldsymbol{\theta_i}$ and based on the expected directions of the discrimination parameters for β_k.

[2] For a more detailed description of data cleaning, items, and sample demographics, see Section A in the Appendix.

country-specific topics. The items should therefore not be considered as a random sample from a large pool of items concerning many issues, but as tailor-made to a conception of a three dimensional political space across Europe with added leeway for country specific topics.

It is important to highlight that this sample is not a random sample from a larger population (Pianzola, 2014). Users self-select into the use of VAA and this selection bias is reflected in significant differences in the distributions regarding age, education, and sex between the reduced sample of 750 users and Eurostat data (Eurostat, 2017, see Table A.4 in the Appendix).

This may lead to an overestimation of the discrimination parameters of the items, because this sample corresponds to highly interested and highly motivated respondents for whom the linkages between issues and an underlying dimension may well be stronger than for others (Jessee, 2016). But it should not lead to biased results regarding the sign of these parameters, as it is to be assumed that VAA users don't significantly differ from the rest of the population in the relationship between their value on the latent trait and their probabilities for different items and response options.

This is supported by the fact that of the 21 items fielded cross-nationally, (nearly) none of the discrimination parameters is surprising in its direction regarding the dimension the item was originally selected for across Europe (i.e. an economic item's discrimination parameter regarding the economic dimension). In the few exceptions where the discrimination parameter for an item on the dimension it was selected for does not correspond to the assumed relationship, I argue that this difference is not due to the sample, but due to issues relating differently to the same dimensions cross-nationally[3].

3 Results

In the following Subsections, I will give a detailed overview of my findings regarding the dimensionality of political spaces across Europe. To address the question concerning the number of dimensions necessary to describe a political space, I focus on a comparison of the expected log pointwise predictive density (ELPD; Gelman et al. (2014); Piironen

[3]Put differently, the monotonicity assumption concerning the direction of the relationship between the value of a latent trait and the probability of a response option should be similar for VAA users and a broader population (Fox, 2010, p.35).

and Vehtari (2017); Vehtari et al. (2017)) and the error rates of the one, two, and three dimensional models estimated by country. Given the selection of items meant to capture the three distinct dimensions of economy, culture, and EU (Wheatley, 2016; Gemenis et al., 2018), it is not surprising that a three dimensional model fits the data best in all countries as measured by the ELPD. However, the improvement in error rates with increasing dimensionality is not equal for all countries. A subgroup of countries does just fine regarding their error rates when less than three dimensions are defined.

Turning to the content of these three dimensions across Europe, I focus on the resulting item discrimination parameters in the three dimensional models. Because the values of these parameters cannot be directly compared across countries, I recode them to either positive, neutral, or negative depending on the resulting 95% credible intervals of the parameters. This amounts to assuming configural invariance across countries in the relationships between between items and dimensions (Steenkamp and Baumgartner, 1998; Stegmueller, 2011).

First, I calculate the variance ratio of the discrimination parameters per item and dimension. Items with a low variance ratio indicate that the relationship between the issue specified in the item and a dimension is very similar across Europe, while items with a high ratio often differ across countries. Then, I plot the share of discrimination parameters which are positive, neutral, or negative across all countries and dimensions. For most items, the discrimination parameters for the dimensions they were selected for don't differ much across Europe, while the parameters of the same items for other dimensions do. However, items gauging the opinion of respondents concerning IMF loans and economic redistribution in the European Union do differ in their discrimination parameters for the dimension they were originally selected for. This highlights that the same response to an item can place you at opposite ends of a latent scale depending on the political space you're moving in or be relevant for an underlying dimension in some spaces, but irrelevant in others.

Based on the relationship between items and dimensions, the proximity of political spaces across Europe can be estimated which allows us to identify countries with very similar spaces - such as Greece and Cyprus - and countries with very dissimilar spaces - such as Ireland and Hungary. In the construction of this proximity measure, the discrimination parameters of all items for all dimensions are considered. I believe that this is important so as to consider the full complexity of the relationships between issues and dimensions, as e.g. a cultural issue may also be of importance for an economic or EU dimension[4].

3.1 How Many Dimensions Are Enough?

Figure 2 displays the comparison of the ELPD of different model specifications across countries (Gelman et al., 2014; Piironen and Vehtari, 2017; Vehtari et al., 2017). The ELPD can be understood as a summary statistic that indicates the fit of the model to the data and includes a penalty for over-fitting the model to the data. This measure can be estimated using different methods (LOO and WAIC; Vehtari et al. (2017))[5].

Models with a better fit to the data should have higher ELPD values than those with a worse fit. According to the discussion in Vehtari et al. (2017), Figure 2 displays the difference in ELPD including estimated standard errors of the difference (the bars display $\pm 1.96 s.e.$). Positive values indicate that the second model has a higher ELPD than the first model. Across all countries, a inverted V shape emerges. This indicates that the ELPD increases most when going from a one to a three dimensional model, while it increases least when going from a two to a three dimensional model (with the exception of Slovakia). Based on this measure, a three dimensional model is best suited to fit the data in all countries, even if the improvement in fit going from a two to a three dimensional model is not equally large across Europe (see e.g. Greece where the inclusion of a separate EU dimension only marginally increases the difference in ELPD).

An alternative measure to assess the number of dimensions is to compare the observed data to simulated data based on the posterior distributions of the parameters (Gelman et al., 1996). Figure 3 displays the error rates across all responses including 95% credible intervals based on 1000 simulations from the posterior distribution. So as to have a meaningful yardstick to compare the error rates of the models to, I estimate a baseline error rate of a zero dimensional model in which the probability of choosing a response option is governed by the marginal distribution of response options across all items (as suggested in Marble and Tyler, 2017). While an error rate of one implies that the model gets every

[4]Results in which only the item parameters on the dimensions these items were selected for are considered are reported in Section D in the Appendix.

[5]I decided against the use of the deviance information criterion (DIC; Spiegelhalter et al. (2002)) as a measure of model fit based on the discussion in Spiegelhalter et al. (2014) and Gelman et al. (2014).

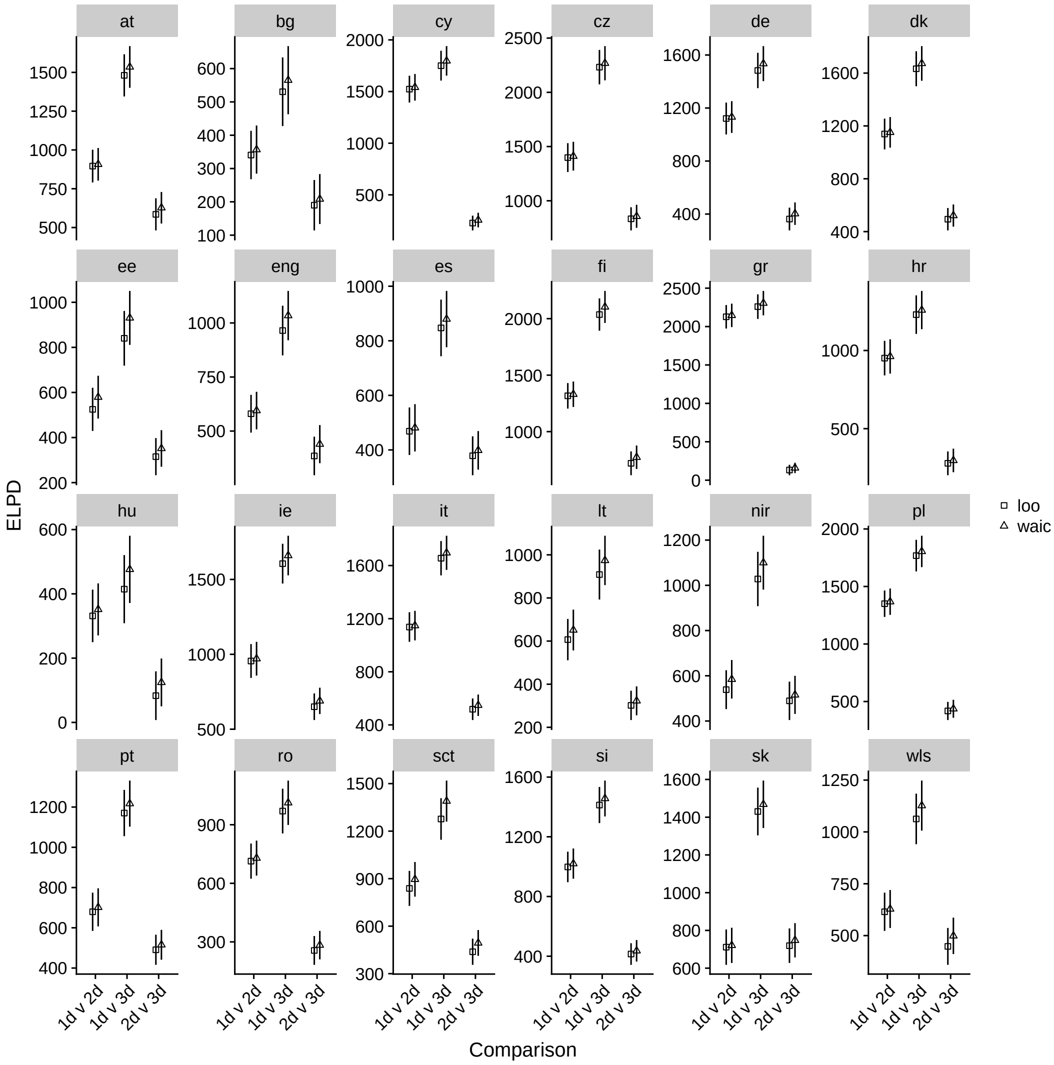


Figure 2: Comparison of ELPD across Models.

response wrong, an error rate of zero would mean that the model predicts every response correctly [6].

In a zero dimensional model, the country-wise error rate in prediction is somewhere between 70 and 80 percent, which reflects that we get a share of roughly $\frac{1}{\text{number of response options}}$ right given five response options. As the models become more complex and include more dimensions, the error rate decreases across all countries. In all countries, the lowest error rate results when using a three dimensional model.

While the ranking of error rates is equal across countries, the difference in error rates with increasingly complex models isn't. For example, the error rate drops by around 15%

[6] In a further iteration of this paper, it would be sensible to do this based on the item-specific marginal distribution, but then the cut-points τ of the IRT model would also need to be estimated separately by item.

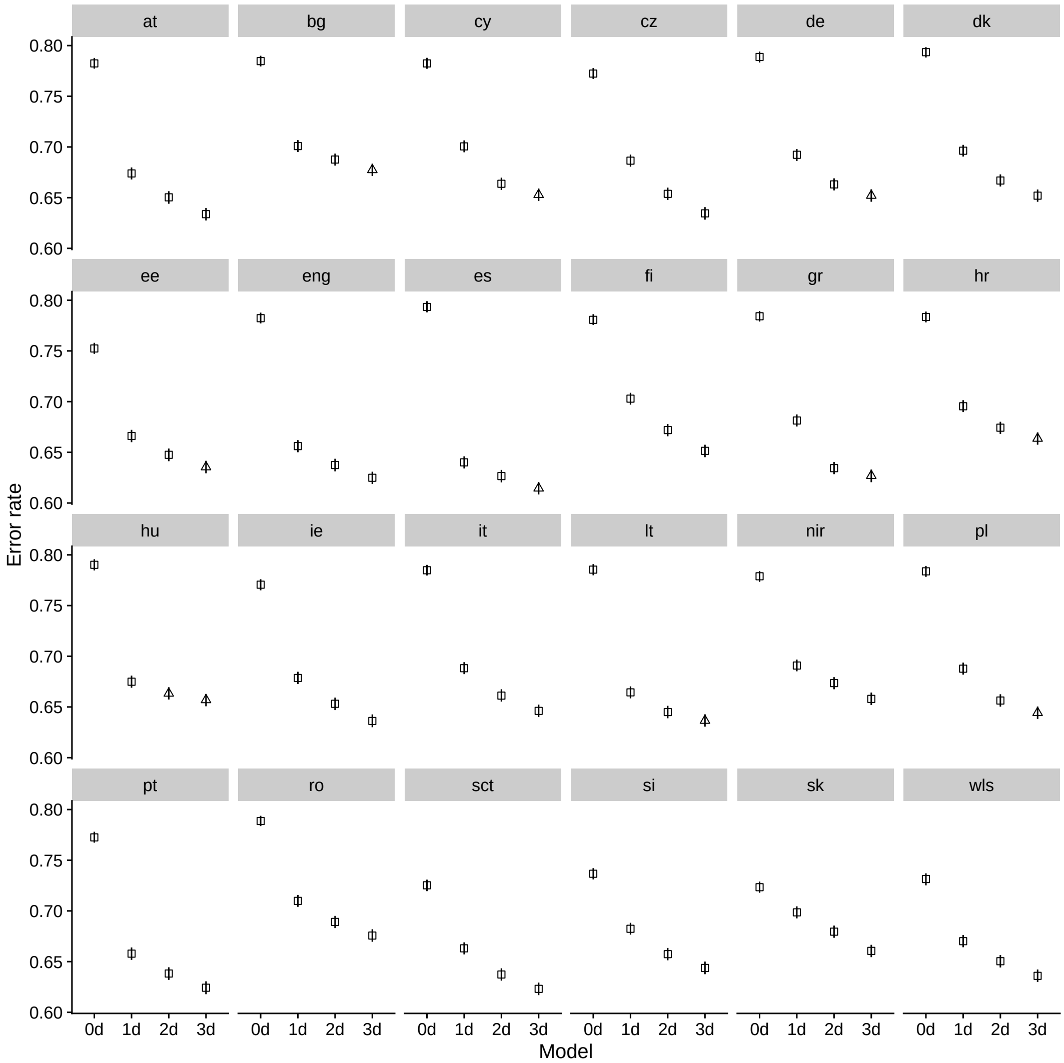


Figure 3: Comparison of Error Rates across Models.

Non-significant differences in error rates with increasing dimensionality marked by a △.

in Spain when going from a zero to a one dimensional model, but it only drops by around 5% in Slovakia. If we are interested in predicting the responses of users and wish to use their position on a latent dimension to do so, it becomes apparent that some countries have more to gain from more complex models than others, possibly indicating that additional dimensions only marginally contribute to explaining responses in these countries.

All countries benefit from a first dimension defined by economic conflict as the error rate drops significantly compared to the error rate in the zero dimensional model. Afterwards, the effects of adding dimensions are more mixed. An additional cultural dimension again reduces the error rate in all countries except for Hungary where the error rate's credible intervals of the one and two dimensional model overlap (marked by a $\triangle$). The group of countries without a significant reduction in error rate grows when adding a third dimension regarding the European Union to the two dimensional model. In Bulgaria, Cyprus, Germany, Estonia, Spain, Greece, Croatia, Hungary, Latvia, and Poland, an additional EU dimension does not significantly improve the fit of the model to the data based on this measure.

Overall, the comparison of both ELPD and error rates across countries and model specifications has highlighted that a three dimensional model including an economic, a cultural, and a EU dimension is best fit to the data. This is unsurprising based on the selection criteria for the items in the first place. However, the comparison of the error rates of the models based on predictions stemming from the posterior distributions of all parameters shows that not all countries profit equally from increasingly complex model specifications. In these countries, the share of explained variation in responses by the inclusion of an additional dimension is small.

3.2 Content of Dimensions

I wish to turn to the varying content of dimensions across countries. I focus on the recoded discrimination parameters of items for the three dimensions. These tell us whether the issue captured by an item is relevant for the placement of an actor along a dimension or not (is the parameter neutral or not) and whether higher (lower) values along the dimension are associated with higher (lower) probabilities of certain response options (is the parameter positive or negative).

Figure 4 displays the variance ratio of discrimination parameters faceted by the dimension an item was originally selected for. Different symbols indicate the discrimination parameters of these items for different dimensions. The ratio is calculated as $1-$ the share of countries with the modal parameter across all countries. Therefore, values at zero indicate that there is no variation in this parameter across countries, while values at one indicate that there is a lot of variation across countries.

The identifying items which are constrained to be equal across countries do not differ cross-nationally and therefore have a ratio of zero (abortion_right, redistribute_wealth, and eu_bad). Figure 4 shows that there is least variation across countries regarding the discrimination parameters of the items for the dimension they were originally selected for. In other words, the parameters of cultural items are similar for the cultural dimension across Europe, as are those of economic items for the economic dimension and the EU items' parameters for the EU dimension. There is much more variation in the parameters for the dimension the items were not originally selected for[7].

There are two notable exceptions to this, namely the economic item regarding IMF loans and the EU item regarding economic redistribution within the European Union. Both have a high ratio for the dimensions they were selected for (roughly 0.25 for the imf_loan item for the economic dimension and roughly 0.35 for the redistribute_eu item for the EU dimension). Based on these results, it is unrealistic to assume that these items measure the same latent dimension across countries in the same way. At the same time, these items exhibit a low ratio for the parameter for other dimensions. The discrimination parameter of the IMF item is equal across Europe regarding the EU dimension, while

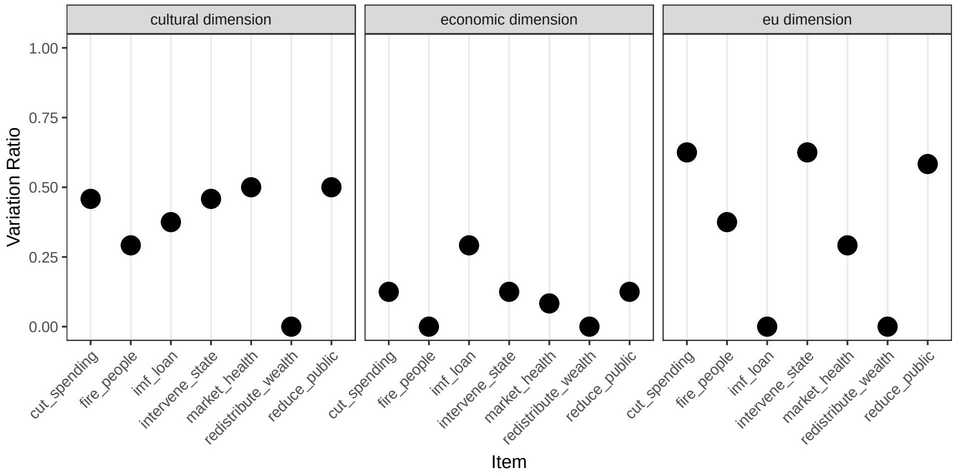
[7] A more detailed discussion of the variation in these 63 parameters across countries (21 common items and three dimensions) can be found in Section C of the Appendix.

the parameter of the item concerning redistribution in the EU is equal for the economic dimension.

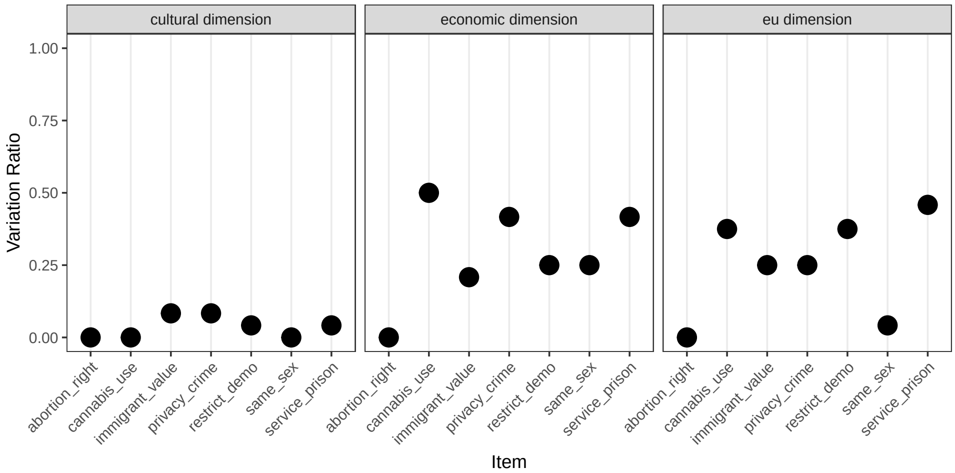
The variation across item parameters, countries, and dimensions is shown in somewhat more detail in Figure 5. The share of countries with a positive, neutral, or negative item parameter is shown separately across items and dimensions. Again, focusing on the identifying items can aid in interpretation. For example, the economic item of redistribute_wealth has a neutral discrimination parameter for the cultural and the EU dimension across all countries, while the parameter is positive for the economic dimension.

Continuing with the IMF loan item and the redistribution in the EU item, we can analyze the variation in their discrimination parameters across dimensions based on the share of countries with a positive, neutral, or negative parameter. While the IMF item was originally assigned to the economic dimension, five of the 24 countries (0.21) have a neutral discrimination parameter for this dimension, while two have a positive parameter contrary to the assumed relationship between the item and the dimension (Hungary and Poland). Only roughly two thirds of the countries display the expected relationship between the redistribution in the EU item and the underlying EU dimension, while this item is of no relevance for this dimension in six countries and has the opposite relationship in three of the countries (Bulgaria, Greece and Hungary).

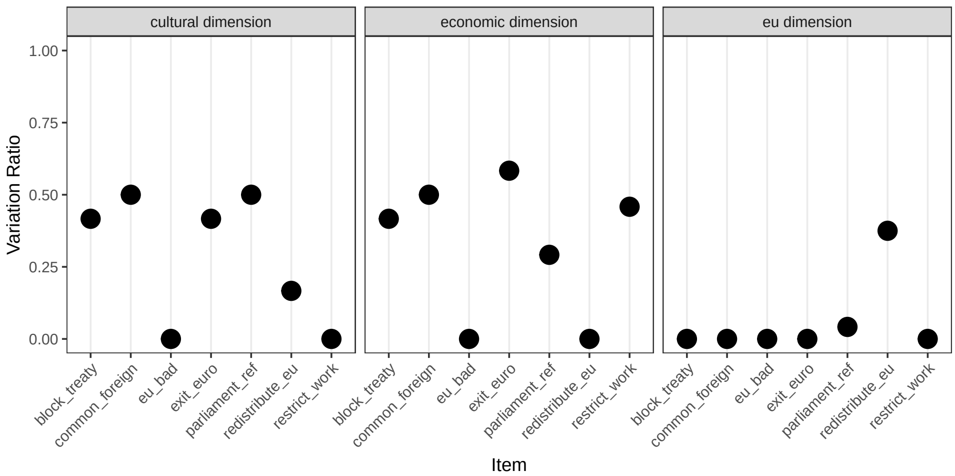
Overall, the analysis of the discrimination parameters underlines that there is a lot of variation concerning the relationship between issues and dimensions across countries. The variation is smallest for the discrimination parameters for the dimensions which items were originally selected for (with some exceptions), but increases when concentrating on the parameters for the dimensions the items were not selected for. Even though the underlying dimensions are defined equally across countries, the estimated relationships between issues and dimensions are not. Choosing the same response for the same item may place a respondent at a different end of a latent dimension across countries or be of no help in placing a respondent along a dimension in others.



(a) **Economic** items.

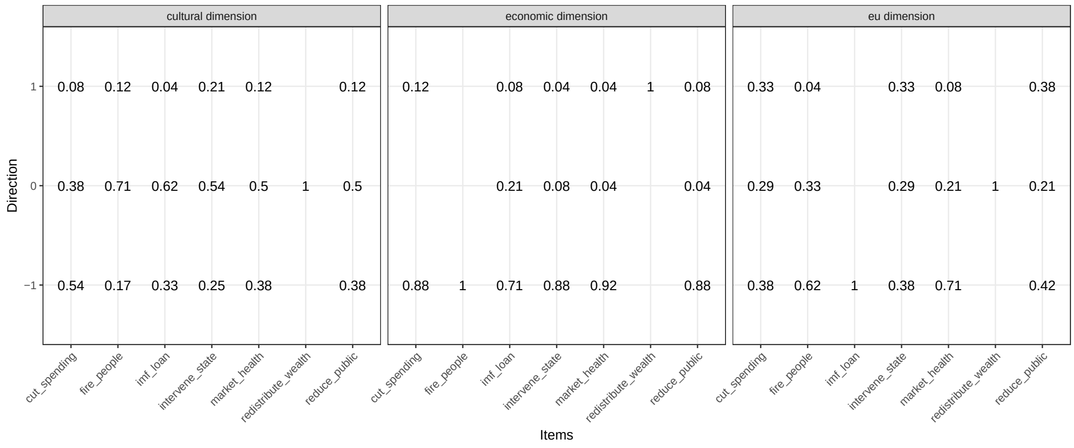


(b) **Cultural** items.

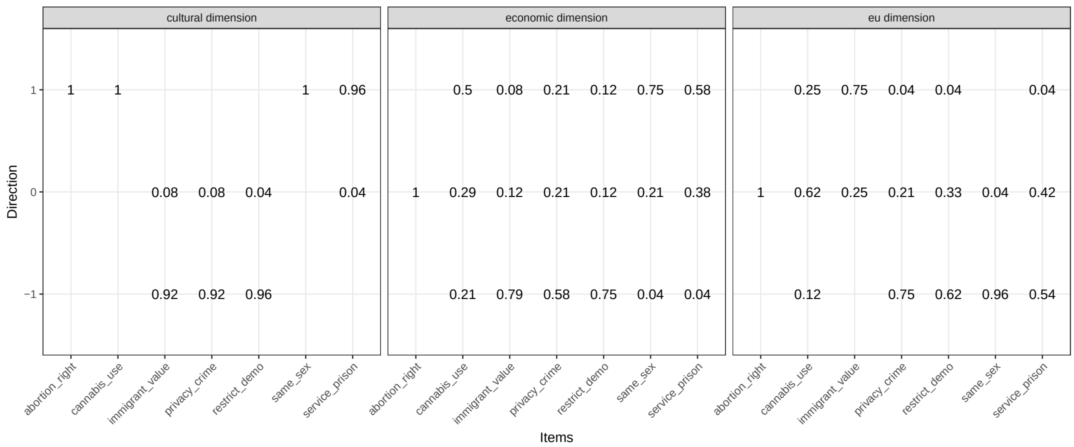


(c) **EU** items.

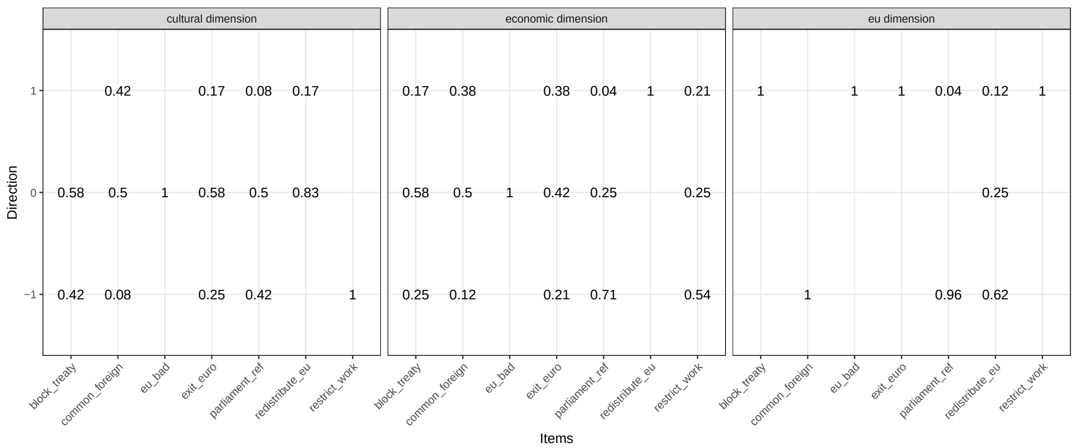
Figure 4: Variance ratio of discrimination parameters across countries and dimensions.



(a) **Economic** items.



(b) **Cultural** items.



(c) **EU** items.

Figure 5: Distribution of discrimination parameters across countries and dimensions.

3.3 Proximity of Spaces

After having discussed the question of how many dimensions are sensible to describe the political spaces of Europe and the content of these dimensions, I want to finish with an analysis of the proximity of political spaces across Europe.

The distance between political spaces based on the sign of all 63 item parameters is estimated using the Euclidean metric. Therefore, the distance of each item parameter is equally important in calculating the distance between countries. The resulting values are normalized so that the maximal distance is one and the minimal distance is zero (when comparing country A with country A or if all item parameters in country B were equal to those in country A).

The distances between countries are displayed in Figure 6. Both the ranking and the value of the distances per country with other countries is of interest. For example, this Figure shows not only that Cyprus and Greece have very similar political spaces as defined by the estimated discrimination parameters, but also that the proximity between these countries is much larger than between Cyprus and Greece and all other countries. It also showcases that the most similar countries for certain countries are still pretty dissimilar. For example, the closest country to Hungary is Latvia with a distance of roughly 0.6 units, while Ireland, Scotland, Wales, and England all lie within 0.2 to 0.4 units apart from one another.

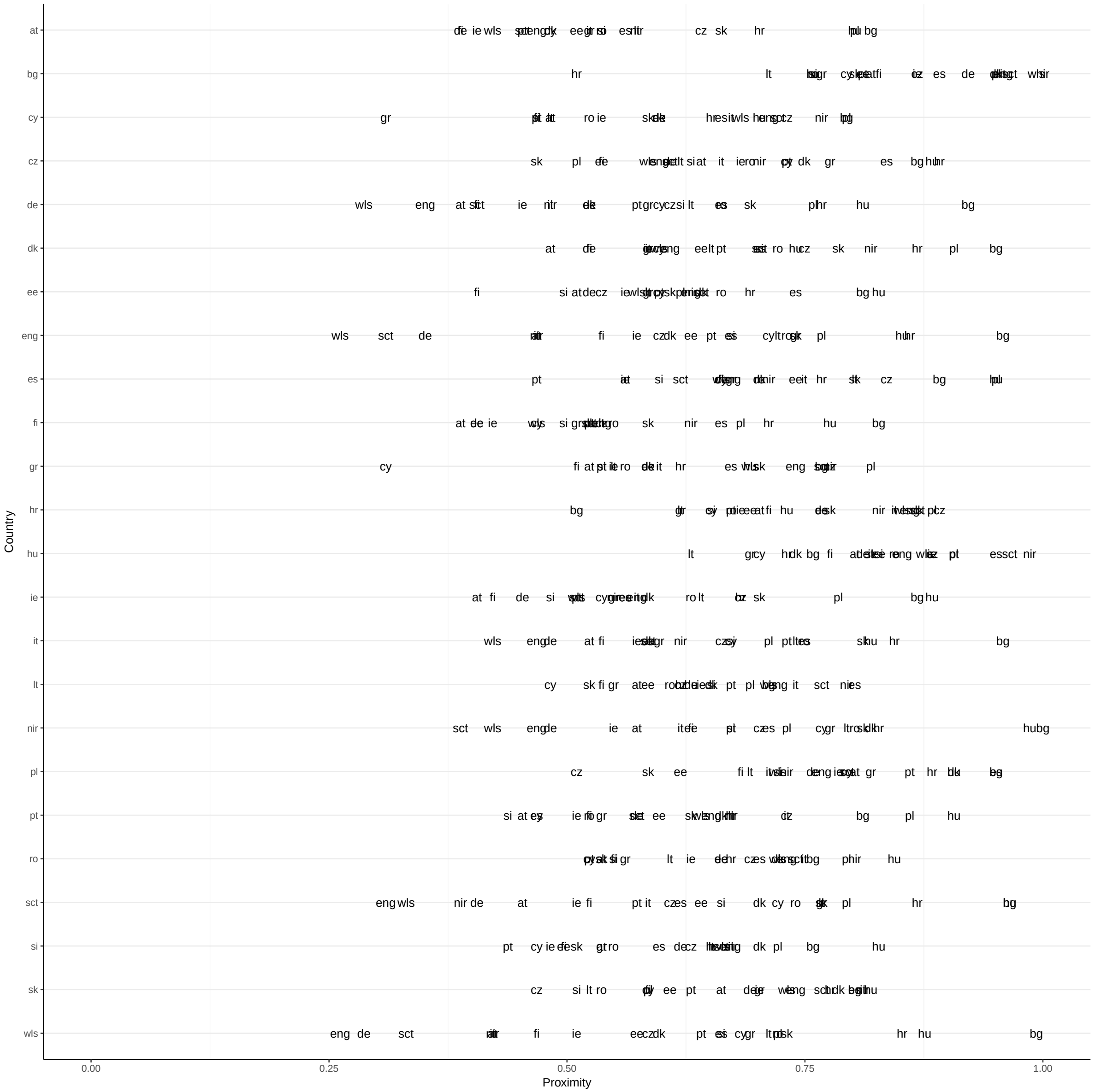


Figure 6: Proximity of political spaces across Europe.

The Euclidean distances between countries can be displayed in a dendogram in Figure 7 with the countries clustered into two, three, four, or five groups depending on the sequence of hierarchical splits. This gives an impression as to which countries have more (or less) similar political spaces across Europe. At a first stage, Hungary, Bulgaria, Latvia, and Croatia are split away from the other 24 countries, before Poland, the Czech Republic, Slovakia, Northern Ireland, Scotland, Wales, and England join an own distinct cluster. Then, this cluster is split into two, separating the countries of the United Kingdom from the Eastern European countries. Finally, the four countries which split off at the beginning are separated into two groups: Bulgaria and Croatia, and Hungary and Latvia. This clustering could be continued up to a 24 clusters (where each country is in a distinct cluster), but for the time being it seems sensible to leave it at this number of groups. While I do not attempt to explain the variation in political spaces at this stage, certain splits align well with previous findings regarding Eastern versus Western European political spaces (Rovny and Edwards, 2012).

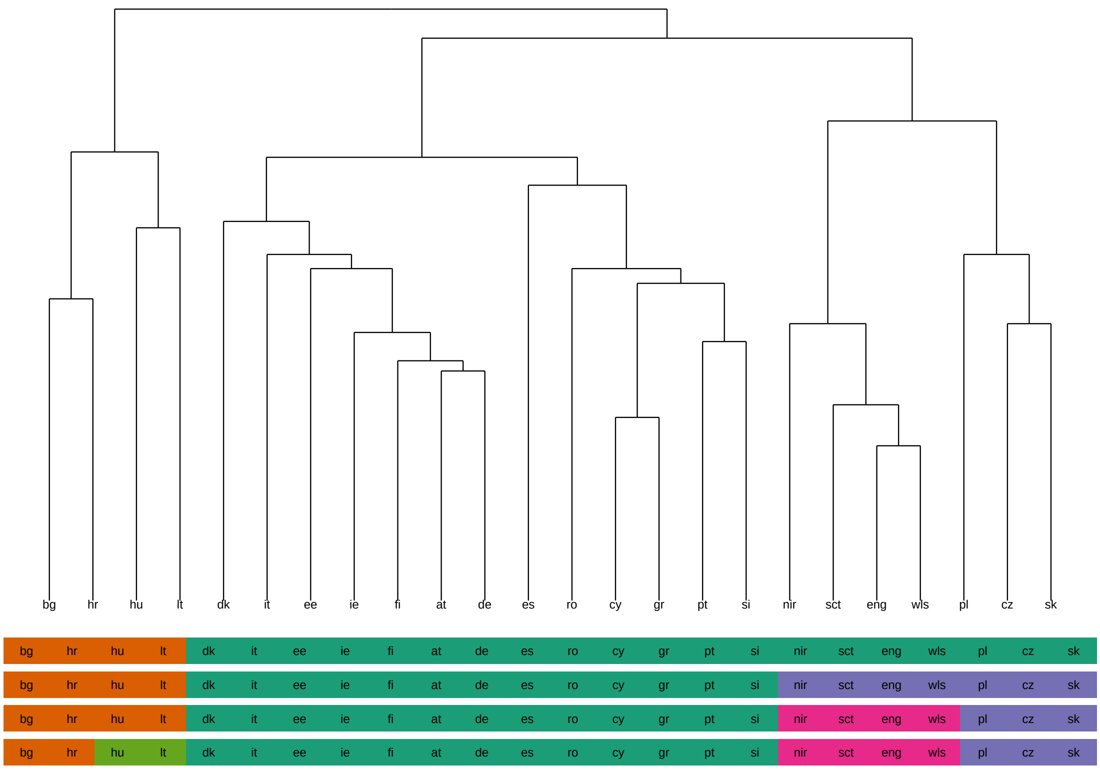


Figure 7: Clustering of political spaces across Europe based on proximity.

4 Conclusion

I focus on three questions in this paper: what number of dimensions are necessary to describe the political spaces of European countries? Is the defining content of the dimensions similar across Europe? Finally, which countries are most similar to others regarding the composition of their political space?

A three dimensional model including an economic, a cultural, and a EU specific dimension fits the data best and reduces the predicted error rate most in all countries. However, the improvement in error rates with increasing dimensionality is not equal for all countries. A subgroup of countries have no significant decrease in error rates when more than one or two dimensions are defined. These countries are Hungary when going from a one to a two dimensional model and Hungary, Bulgaria, Cyprus, Germany, Estonia, Spain, Greece, Croatia, Latvia, and Poland when going from a two to a three dimensional model.

While the relationship between a selected few of the items and the dimensions is constrained to be equal across Europe, the parameters of all other items are freely estimated. The relationship between the remaining issues and dimensions varies considerably across countries. This highlights that the political spaces across Europe differ not only in the number of dimensions necessary to adequately describe them, but also in the content of these dimensions. Choosing the same response for the same item may place a respondent at a different end of a latent dimension across countries, while it the response may be of no aid in placing a respondent along a dimension in others.

Finally, I estimate the similarity of political spaces across European countries based on the item parameters linking issues to dimensions. This allows me to cluster the 24 countries according to whether their spaces are more or less similar to one another. The political spaces of Hungary, Bulgaria, Latvia, and Croatia are most dissimilar to the spaces of all other countries, followed by a group consisting of Poland, the Czech Republic, Slovakia, Northern Ireland, Scotland, Wales, and England.

In its current version, this paper's main contribution is bringing the heterogeneity in political spaces across Europe center stage. In a next iteration, the framing of the paper should become clearer than it currently is.

To do

- Framing + Focus.
- Robustness:
  - Re-estimate model without the country-specific items, because the inclusion of some items can influence the parameters of others.
  - Re-estimate model with item-specific τ, because they currently only reflect the marginal distribution of responses across all items.
  - Check composition of data for political composition by country, because this may bias the discrimination parameters.
  - Check data cleaning. Do we care whether foreigners filled out the survey?

References

Andreadis, I. (2014), Data quality and data cleaning, *in* 'Matching Voters with Parties and Candidates Voting Advice Applications in a Comparative Perspective.'.

Bafumi, J., Gelman, A., Park, D. K. and Kaplan, N. (2005), 'Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation', *Political Analysis* **13**(2), 171–187.

Bakker, R., de Vries, C., Edwards, E., Hooghe, L., Jolly, S., Marks, G., Polk, J., Rovny, J., Steenbergen, M. and Vachudova, M. A. (2015), 'Measuring party positions in Europe The Chapel Hill expert survey trend file, 1999–2010', *Party Politics* **21**(1), 143–152.

Bakker, R., Jolly, S. and Polk, J. (2012), 'Complexity in the European party space: Exploring dimensionality with experts', *European Union Politics* **13**(2), 219–245.

Benoit, K. and Laver, M. (2006), *Party Policy in Modern Democracies*, Routledge.

Benoit, K. and Laver, M. (2012), 'The dimensionality of political space: Epistemological and methodological considerations', *European Union Politics* **13**(2), 194–218.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P. and Riddell, A. (2016), 'Stan: A Probabilistic Programming Language', *JSS Journal of Statistical Software* . in press.

Clinton, J., Jackman, S. and Rivers, D. (2004), 'The Statistical Analysis of Roll Call Data', *American Political Science Review* **98**(2), 355–370.

Converse, P. E. (2006), 'The nature of belief systems in mass publics (1964)', *Critical Review* **18**(1-3), 1–74.

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P. and Billiet, J. (2014), 'Measurement Equivalence in Cross-National Research', *Annual Review of Sociology* **40**(1), 55–75.

Davidov, E., Schmidt, P. and Billiet, J. (2011), *Cross-Cultural Analysis: Methods and Applications*, Routledge.

Eurostat (2017), 'Population by sex, age and educational attainment level (1 000)', http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsa_pgaed&lang=en.

Fox, J.-P. (2010), *Bayesian Item Response Modeling: Theory and Applications*, Springer Science & Business Media.

Garzia, D. and Marschall, S., eds (2014), *Matching Voters with Parties and Candidates Voting Advice Applications in a Comparative Perspective*, ECPR Press.

Gelman, A., Hwang, J. and Vehtari, A. (2014), 'Understanding predictive information criteria for Bayesian models', *Statistics and Computing* **24**(6), 997–1016.

Gelman, A., Meng, X.-L. and Stern, H. (1996), 'POSTERIOR PREDICTIVE ASSESSMENT OF MODEL FITNESS VIA REALIZED DISCREPANCIES', *Statistica Sinica* **6**(4), 733–760.

Gelman, A. and Rubin, D. B. (1992), 'Inference from Iterative Simulation Using Multiple Sequences', *Statistical Science* **7**(4), 457–472.

Gemenis, K., Mendez, F. and Wheatley, J. (2018), 'Political parties' policy positions for the 2014 elections to the European Parliament', https://datorium.gesis.org/xmlui/handle/10.7802/1594.

Hambleton, R. K., Swaminathan, H. and Rogers, H. J. (1991), *Fundamentals of Item Response Theory*, SAGE.

Hinich, M. J. and Munger, M. C. (1992), 'A Spatial Theory of Ideology', *Journal of Theoretical Politics* **4**(1), 5–30.

Hinich, M. J. and Munger, M. C. (1996), *Ideology and the Theory of Political Choice*, The University Of Michigan Press, Ann Arbor.

Inglehart, R. and Klingemann, H.-D. (1976), 'Party identification, ideological preference and the left-right dimension among Western mass publics', *Party identification and beyond: representations of voting and party competition* pp. 243–273.

Jackman, S. (2000), 'Estimation and Inference Are Missing Data Problems: Unifying Social Science Statistics via Bayesian Simulation', *Political Analysis* **8**(4), 307–332.

Jackman, S. (2001), 'Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference, and Model Checking', *Political Analysis* **9**(3), 227–241.

Jackman, S. (2004), 'Bayesian Analysis for Political Research', *Annual Review of Political Science* **7**(1), 483–505.

Jessee, S. (2016), '(How) Can We Estimate the Ideology of Citizens and Political Elites on the Same Scale?', *American Journal of Political Science* pp. n/a–n/a.

Kriesi, H., Grande, E., Lachat, R., Dolezal, M., Bornschier, S. and Frey, T. (2006), 'Globalization and the transformation of the national political space: Six European countries compared', *European Journal of Political Research* **45**(6), 921–956.

Kuklinski, J. H. and Peyton, B. (2007), Belief Systems and Political Decision Making, *in* R. J. Dalton and H.-D. Klingemann, eds, 'The Oxford Handbook of Political Behavior', Oxford University Press, Oxford.

Laver, M. (2014), 'Measuring Policy Positions in Political Space', *Annual Review of Political Science* **17**(1), 207–223.

Marble, W. and Tyler, M. (2017), How Much Should We Trust Ideal Point Estimates from Surveys?

Marks, G. and Steenbergen, M. (2002), 'Understanding Political Contestation in the European Union', *Comparative Political Studies* **35**(8), 879–892.

Martin, A. D. and Quinn, K. M. (2002), 'Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999', *Political Analysis* **10**(2), 134–153.

Ostini, R. and Nering, M. L. (2006), *Polytomous Item Response Theory Models*, SAGE.

Pianzola, J. (2014), 'Selection biases in voting advice application research', *Electoral Studies* **36**, 272–280.

Piironen, J. and Vehtari, A. (2017), 'Comparison of Bayesian predictive methods for model selection', *Statistics and Computing* **27**(3), 711–735.

R Core Team (2017), *R: A Language and Environment for Statistical Computing*, Vienna, Austria.

Rivers, D. (2003), Identification Of Multidimensional Spatial Voting Models.

Rovny, J. and Edwards, E. E. (2012), 'Struggle over Dimensionality: Party Competition in Western and Eastern Europe', *East European Politics and Societies* **26**(1), 56–74.

Rovny, J. and Polk, J. (2018), 'New wine in old bottles: Explaining the dimensional structure of European party systems', *Party Politics* p. 1354068817752518.

Samejima, F. (1969), 'Estimation of Latent Ability Using a Response Pattern of Graded Scores', *Psychometrika Monographs* (34).

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van Der Linde, A. (2002), 'Bayesian measures of model complexity and fit', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**(4), 583–639.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2014), 'The deviance information criterion: 12 years on', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **76**(3), 485–493.

Steenkamp, J.-B. E. M. and Baumgartner, H. (1998), 'Assessing Measurement Invariance in Cross-National Consumer Research', *Journal of Consumer Research* **25**(1), 78–107.

Stegmueller, D. (2011), 'Apples and Oranges? The Problem of Equivalence in Comparative Research', *Political Analysis* **19**(4), 471–487.

Treier, S. and Hillygus, D. S. (2009), 'The Nature of Political Ideology in the Contemporary Electorate', *Public Opinion Quarterly* **73**(4), 679–703.

van der Eijk, C. and Rose, J. (2015), 'Risky Business: Factor Analysis of Survey Data – Assessing the Probability of Incorrect Dimensionalisation', *PLOS ONE* **10**(3), e0118900.

Vehtari, A., Gelman, A. and Gabry, J. (2017), 'Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC', *Statistics and Computing* **27**(5), 1413–1432.

Wheatley, J. (2016), Reconceptualising Dimensions of Political Competition in Europe in an Age of Globalisation, Prague.

A Data Description

The data was generated by the use of the voting advice application (VAA) EUvox prior to the elections to the European parliament in 2014 (Gemenis et al., 2018). Essentially, VAA attempt to match voters to political actors by assessing the (dis)similarity between the opinions indicated by respondents regarding a variety of issues and those of politicians or parties. Therefore, VAA data represents a source of information on the political opinions of many. Additionally, the use of VAA is widespread with large sample sizes in most countries (Garzia and Marschall, 2014).

This assessment is highlighted by Table A.1 which displays the number of users by country included in the full dataset. In total, the VAA was used more than nine hundred thousand times across Europe as can be seen in the third column of the Table. However, this initial number includes rogue entries, i.e. duplicates or other records which are similarly undesirable in analysis. I followed the suggestions of Andreadis (2014) and especially those of Wheatley (2016) to clean the data of rogue entries.

I removed entries where:

1. the time taken to complete all issue statements was less than 120 seconds;
2. the VAA was completed by smartphone;
3. the response time to any one issue statement was less than one second;
4. the response time to three or more statements was less than two seconds;
5. the response to any one item was missing;
6. the user answered ten successive issue statements in the same way;
7. the user indicated a birth year prior to 1920;
8. the same combination of IP address, gender, and date of birth of the user appeared multiple times across entries.

The third column in Table A.1 reflects the number of entries per country after data cleaning. Around 51% of all entries are cleaned out of the data by country, so that the data is reduced to around 400 thousand entries. The smallest share of rogue entries based on this cleaning procedure is observed in Romania with around 39%, while the largest share of rogue entries is in Denmark with roughly 61%.

Table A.1: Number of Users

Abbreviation	Country	Nr. of Users	Nr. of Users (cleaned)	Sample Size
at	Austria	10'669	4'732	750
bg	Bulgaria	7'255	3'566	750
cy	Cyprus	5'198	2'605	750
cz	Czech Republic	28'679	15'028	750
de	Germany	9'727	4'808	750
dk	Denmark	127'342	48'841	750
ee	Estonia	18'387	9'490	750
eng	England	103'291	44'584	750
es	Spain	281'559	110'244	750
fi	Finland	8'299	4'075	750
gr	Greece	63'923	34'465	750
hr	Croatia	7'823	3'569	750
hu	Hungary	6'732	2'924	750
ie	Ireland	9'811	4'973	750
it	Italy	36'943	18'778	750
lt	Latvia	9'127	4'433	750
nir	Northern Ireland	3'614	1'486	750
pl	Poland	74'022	37'677	750
pt	Portugal	55'263	31'459	750
ro	Romania	9'585	5'799	750
sct	Scotland	9'921	4'357	750
si	Slovenia	3'890	1'760	750
sk	Slovakia	7'301	4'115	750
wls	Wales	5'804	2'546	750
total		904'165	406'314	18'000

Contrary to Wheatley (2016), I did not exclude non-native respondents from the data, but I did exclude respondents who failed to indicate a response on any one of the issue statements. While I believe that the inclusion of non-natives should not be consequential for the analysis, the exclusion of those not responding to any one of the items could further accentuate the self-selection bias in VAA (Pianzola, 2014). This could lead to an overestimation of the discrimination parameters of the items, as these users may well be the most motivated and most interested of all.

However, the results of a χ^2-test comparing the distribution of the socio-demographic and political variables concerning age, sex, education, and self-reported political interest displayed in Table A.3 indicate that the reduced sample only differs significantly from the full, but cleaned dataset for a selected few countries and variables (regarding political interest in Austria, education in Greece, age and education in Italy, sex in Latvia, and interest in Portugal).

Nonetheless, the selection bias inherent to VAA data based on the fact that users self-select into the use of these applications is visible in Tables A.2 and A.3. Table A.2 gives an overview of the socio-demographics of the pooled cleaned data. Across Europe, the users who remain in the data after cleaning are predominantely young, well educated, interested in politics, and male (Garzia and Marschall, 2014).

The results of a χ^2-test comparing the distribution of the socio-demographic variables in the reduced sample (the 750 users after data cleaning) to those reported by Eurostat (2017) reported in Table A.4 underlines that the analyzed sample differs significantly from the population of the country on age, sex, and education.

Table A.2: Demographics of Full Cleaned Sample

Age	
y15-19	0.05
y20-24	0.17
y25-29	0.15
y30-34	0.13
y35-39	0.11
y40-44	0.09
y45-49	0.08
y50-54	0.07
y55-59	0.06
y60-64	0.04
y65-69	0.05
ISCED11	
ed0-2	0.26
ed3_4	0.14
ed5-8	0.61
Interest	
0 - very interested	0.39
1 - somewhat interested	0.43
2 - little interest	0.14
3 - no interest	0.02
4 - prefer not to say	0.02
Sex	
f	0.30
m	0.70

In total, 30 items were included in the survey (Wheatley, 2016; Gemenis et al., 2018). Of these, nine were included to capture country-specific topics, while the remaining 21 items were fielded in all countries with a (more or less) consistent wording. Table A.5 includes both the abbreviated names of the items fielded across Europe as well their description. These 21 items were selected to capture three distinct political dimensions, namely an economic dimension, a cultural dimension, and a dimension regarding the European Union according to the Chapel Hill Expert Survey (Bakker et al., 2015) and findings regarding the dimensionality of political spaces of parties in Europe (Bakker et al., 2012).

The dimension which each item is attributed to is also described in Table A.5. Users had five response options, ranging from 'completely agree' (= 1) to 'completely disagree' (= 5) including a 'neither agree, nor disagree' option. Because the items are not worded in the same direction, higher (lower) responses do not always indicate more right-wing (more left-wing) positions on the underlying latent dimension[8].

[8]The country-specific items are included in estimation, but not considered in subsequent analysis. Therefore, I excluded them from Table A.5.

Table A.3: Comparison of Sample Demographics to Full Cleaned Sample

Country	Variable	χ^2	DF	P
at	age	16.22	10	0.09
at	sex	0.06	1	0.80
at	edu	0.71	2	0.70
at	interest	35.52	4	**0.00**
bg	age	14.19	10	0.17
bg	sex	0.01	1	0.90
bg	edu	0.53	2	0.77
bg	interest	3.09	4	0.54
cy	age	5.32	10	0.87
cy	sex	0.00	1	1.00
cy	edu	1.20	2	0.55
cy	interest	1.96	4	0.74
cz	age	11.77	10	0.30
cz	sex	0.45	1	0.50
cz	edu	0.84	2	0.66
cz	interest	4.78	4	0.31
de	age	6.00	10	0.81
de	sex	1.80	1	0.18
de	edu	0.14	2	0.93
de	interest	6.01	4	0.20
dk	age	6.71	10	0.75
dk	sex	0.11	1	0.74
dk	edu	4.31	2	0.12
dk	interest	8.67	4	0.07
ee	age	1.75	10	1.00
ee	sex	0.03	1	0.86
ee	edu	1.69	2	0.43
ee	interest	1.59	4	0.81
eng	age	11.14	10	0.35
eng	sex	1.36	1	0.24
eng	edu	3.39	2	0.18
eng	interest	0.69	4	0.95
es	age	11.08	10	0.35
es	sex	0.69	1	0.41
es	edu	0.27	2	0.88

es	interest	1.61	4	0.81
fi	age	9.56	10	0.48
fi	sex	0.25	1	0.62
fi	edu	2.13	2	0.34
fi	interest	1.40	4	0.84
gr	age	9.30	10	0.50
gr	sex	0.36	1	0.55
gr	edu	14.39	2	**0.00**
gr	interest	3.84	4	0.43
hr	age	14.59	10	0.15
hr	sex	0.41	1	0.52
hr	edu	1.19	2	0.55
hr	interest	2.66	4	0.62
hu	age	9.48	10	0.49
hu	sex	0.13	1	0.72
hu	edu	0.67	2	0.71
hu	interest	2.92	5	0.71
ie	age	4.41	10	0.93
ie	sex	0.01	1	0.91
ie	edu	1.57	2	0.46
ie	interest	1.09	4	0.90
it	age	20.17	10	**0.03**
it	sex	0.80	1	0.37
it	edu	7.19	2	**0.03**
it	interest	0.70	4	0.95
lt	age	6.12	10	0.81
lt	sex	3.87	1	**0.05**
lt	edu	0.19	2	0.91
lt	interest	2.31	4	0.68
nir	age	3.60	10	0.96
nir	sex	0.84	1	0.36
nir	edu	2.01	2	0.37
nir	interest	2.40	4	0.66
pl	age	5.40	10	0.86
pl	sex	0.00	1	0.99
pl	edu	2.44	2	0.29
pl	interest	2.00	4	0.74
pt	age	17.39	10	0.07

pt	sex	0.01	1	0.94
pt	edu	1.80	2	0.41
pt	interest	10.25	4	**0.04**
ro	age	4.31	10	0.93
ro	sex	0.37	1	0.54
ro	edu	1.88	2	0.39
ro	interest	2.42	4	0.66
sct	age	8.86	10	0.55
sct	sex	0.15	1	0.70
sct	edu	5.42	2	0.07
sct	interest	2.16	4	0.71
si	age	4.22	10	0.94
si	sex	1.63	1	0.20
si	edu	1.78	2	0.41
si	interest	1.43	4	0.84
sk	age	8.85	10	0.55
sk	sex	0.55	1	0.46
sk	edu	0.93	2	0.63
sk	interest	8.55	4	0.07
wls	age	10.53	10	0.40
wls	sex	0.00	1	0.98
wls	edu	3.57	2	0.17
wls	interest	3.33	4	0.50

P values below 0.05 marked in bold.

Table A.4: Comparison of Sample Demographics to Eurostat

Country	Variable	χ^2	DF	P
at	age	103.14	10	**0.00**
at	sex	131.68	1	**0.00**
at	edu	550.63	2	**0.00**
bg	age	306.01	10	**0.00**
bg	sex	85.04	1	**0.00**
bg	edu	865.61	2	**0.00**
cy	age	160.89	10	**0.00**
cy	sex	270.02	1	**0.00**
cy	edu	739.96	2	**0.00**
cz	age	73.50	10	**0.00**
cz	sex	242.42	1	**0.00**
cz	edu	1'457.36	2	**0.00**
de	age	338.56	10	**0.00**
de	sex	265.28	1	**0.00**
de	edu	392.41	2	**0.00**
dk	age	57.10	10	**0.00**
dk	sex	51.09	1	**0.00**
dk	edu	12.88	2	**0.00**
ee	age	109.26	10	**0.00**
ee	sex	57.98	1	**0.00**
ee	edu	439.57	2	**0.00**
eng	age	64.76	10	**0.00**
eng	sex	122.08	1	**0.00**
eng	edu	445.22	2	**0.00**
es	age	277.35	10	**0.00**
es	sex	115.61	1	**0.00**
es	edu	327.06	2	**0.00**
fi	age	106.29	10	**0.00**
fi	sex	150.32	1	**0.00**
fi	edu	142.46	2	**0.00**
gr	age	160.42	10	**0.00**
gr	sex	241.62	1	**0.00**
gr	edu	978.24	2	**0.00**
hr	age	448.26	10	**0.00**
hr	sex	133.97	1	**0.00**

hr	edu	1'176.35	2	**0.00**
hu	age	297.54	10	**0.00**
hu	sex	126.76	1	**0.00**
hu	edu	270.15	2	**0.00**
ie	age	149.29	10	**0.00**
ie	sex	160.92	1	**0.00**
ie	edu	868.00	2	**0.00**
it	age	555.68	10	**0.00**
it	sex	145.78	1	**0.00**
it	edu	735.53	2	**0.00**
lt	age	760.20	10	**0.00**
lt	sex	8.27	1	**0.00**
lt	edu	852.11	2	**0.00**
nir	age	154.76	10	**0.00**
nir	sex	147.49	1	**0.00**
nir	edu	625.36	2	**0.00**
pl	age	504.05	10	**0.00**
pl	sex	191.25	1	**0.00**
pl	edu	1'048.40	2	**0.00**
pt	age	202.25	10	**0.00**
pt	sex	75.23	1	**0.00**
pt	edu	1'777.94	2	**0.00**
ro	age	125.32	10	**0.00**
ro	sex	120.53	1	**0.00**
ro	edu	231.42	2	**0.00**
sct	age	81.20	10	**0.00**
sct	sex	99.07	1	**0.00**
sct	edu	401.89	2	**0.00**
si	age	285.34	10	**0.00**
si	sex	112.71	1	**0.00**
si	edu	932.10	2	**0.00**
sk	age	219.56	10	**0.00**
sk	sex	164.53	1	**0.00**
sk	edu	1'734.97	2	**0.00**
wls	age	64.69	10	**0.00**
wls	sex	82.39	1	**0.00**
wls	edu	393.33	2	**0.00**

P values below 0.05 marked in bold.

Table A.5: Item Description

Item	Description	Dimension
abortion_right	**Women should be free to decide on matters of abortion.**	culture
block_treaty	A single member state should be able to block a treaty change even if all the other member states agree to it.	EU
cannabis_use	The recreational use of cannabis should be legal.	culture
common_foreign	There should be a common EU foreign policy even if this limits the capacity of Country X to act independently.	EU
cut_spending	Cutting government spending is a good way to solve the economic crisis.	economy
eu_bad	**Overall EU membership has been a bad thing for the Country X.**	EU
exit_euro	Country X should exit the Euro (Eurozone countries)/ never adopt the Euro (non-Eurozone countries)/ introduce the Euro (SE).	EU
fire_people	It should be easy for companies to fire people./ (SE) Employers should be able to make more exceptions from LAS 1 when downsizing.	economy
imf_loan	External loans from institutions such as the IMF are a good solution to crisis situations.	economy
immigrant_value	Immigrants must adapt to the values and culture of Country X.	culture
intervene_state	The state should intervene as little as possible in the economy.	economy
market_health	Free market competition makes the health care system function better.	economy
parliament_ref	EU treaties should be decided by [name of national parliament] rather than by citizens in a referendum.	EU
privacy_crime	Restrictions on citizen privacy are acceptable in order to combat crime.	culture
redistribute_eu	The EU should redistribute resources from richer to poorer EU regions.	EU
redistribute_wealth	**Wealth should be redistributed from the richest people to the poorest.**	economy
reduce_public	The number of public sector employees should be reduced.	economy
restrict_demo	To maintain public order governments should be able to restrict demonstrations.	culture
restrict_work	The right of EU citizens to work in Country X should be restricted.	EU
same_sex	Same sex couples should enjoy the same rights as heterosexual couples (FR)/ to adopt children (SENL)/ to marry (remaining countries).	culture
service_prison	Less serious crimes should be punished with community service not imprisonment.	culture

Identifying items marked in bold.

B Error Rate by Interest & Education

In Figures B.1 and B.2, the error rates by political interest and education of respondents by model and country are shown. Respondents with higher interest or education have lower error rates, indicating that the estimated models are better suited to predict their responses than those of other respondents. The responses of these people are more constrained than those of others (Jessee, 2016).

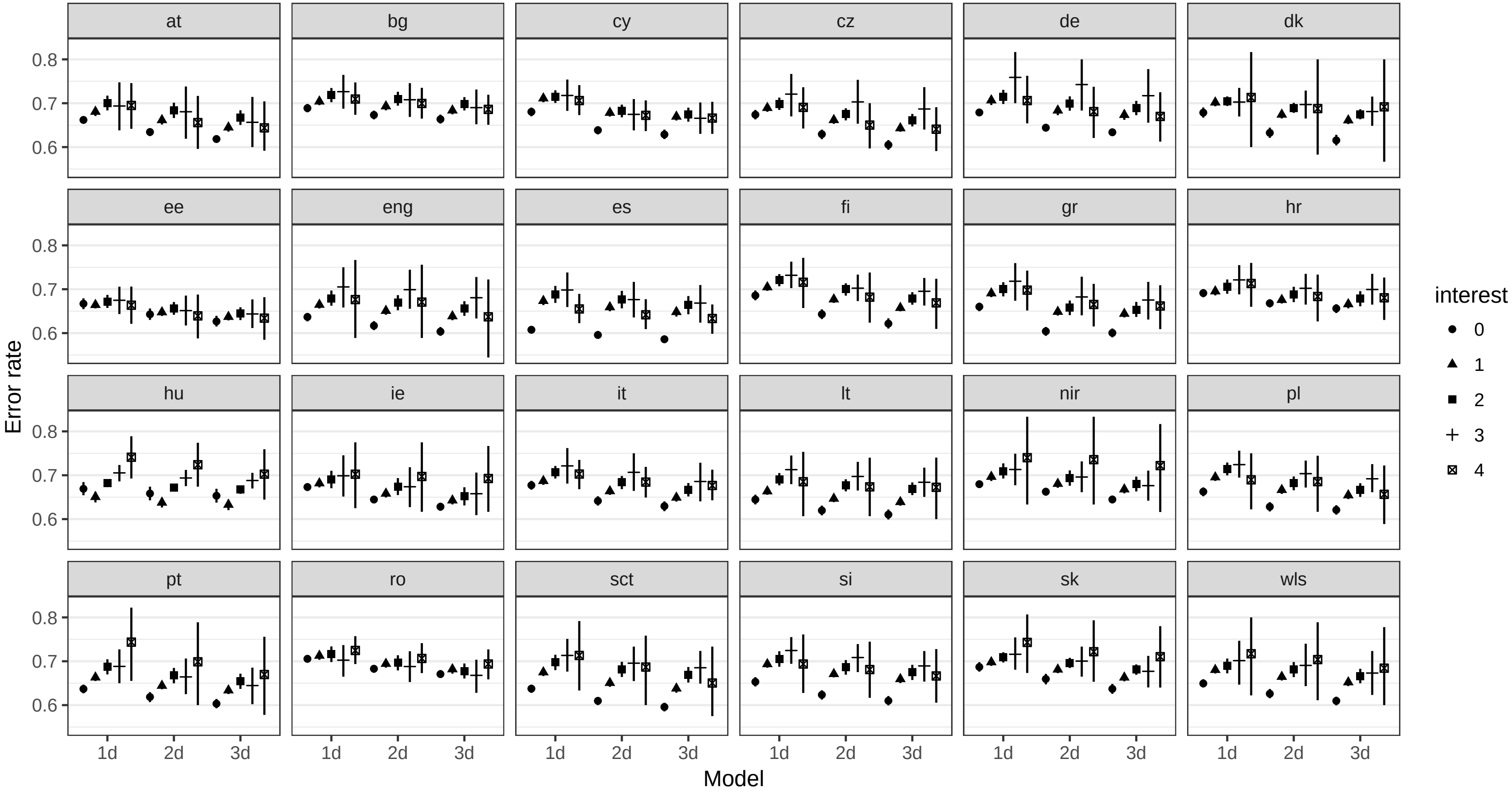


Figure B.1: Comparison of Error Rates across Models by Political Interest.

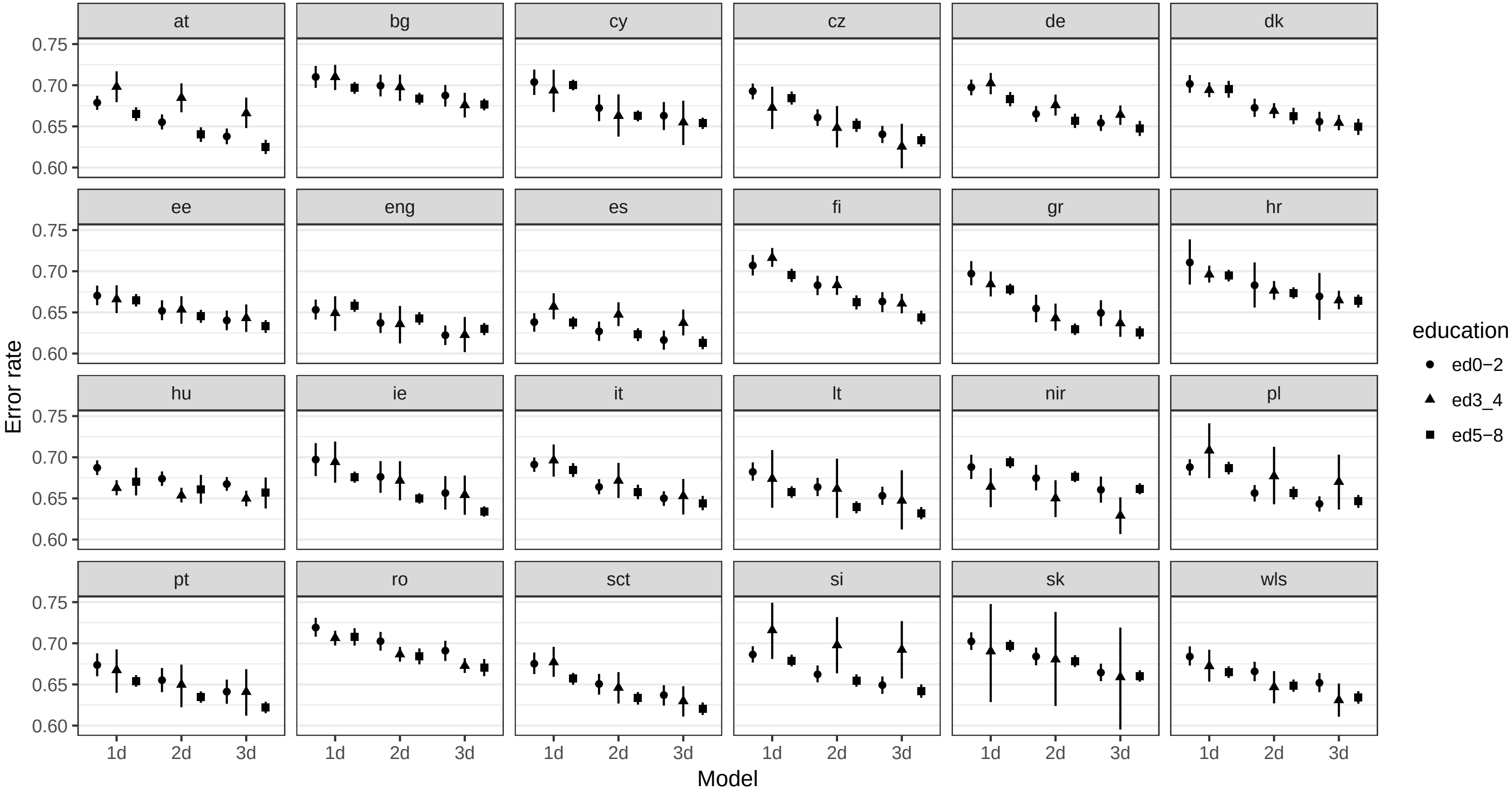


Figure B.2: Comparison of Error Rates across Models by Education.

C Item Parameters

Figures C.1, C.2, and C.3 display the sign of the item discrimination parameters across countries when estimating a three dimensional model with an economic, a cultural, and a EU dimension. Figure C.1 plots all economic items, Figure C.2 all cultural items, and Figure C.3 all EU items.

The discrimination parameters of the identifying items (redistribute_wealth, abortion_right, and eu_bad) do not vary across countries because they are constrained to be equal. More variation is apparent in the parameters of the other items. In most cases, there is less variation in the parameter for the dimension the item was selected for then for the other dimensions. For example, the cannabis_use item in Figure C.2 has a positive discrimination parameter in all countries for the cultural dimension, but a lot more variation in its parameter for the remaining two dimensions.

In other words, responding with 'completely disagree' to an item for the liberalization of cannabis use will consistently place you to the right of a cultural dimension, while it may have either no influence or place you to the left along an economic or EU dimension.

These Figures lead back to the question whether the findings are biased because of self-selection into VAA (Pianzola, 2014). Of all economic items, the only countries in which the discrimination parameters for the designated dimension have a different sign than expected are Bulgaria, Croatia, Hungary, and Poland. Concerning the cultural items, all discrimination parameters are as they are to be expected. Finally, only Spain, Bulgaria, Greece, and Hungary differ in this manner for some of the EU items. Additionally, there are some countries for each item type in which some items have a discrimination parameter which is not significantly different from zero for the dimension they were intended for. Overall, I would argue that this means that the composition of the non-random sample does not lead to biased estimated parameters, but that cross-national differences exist in the relationships between issues and dimensions.

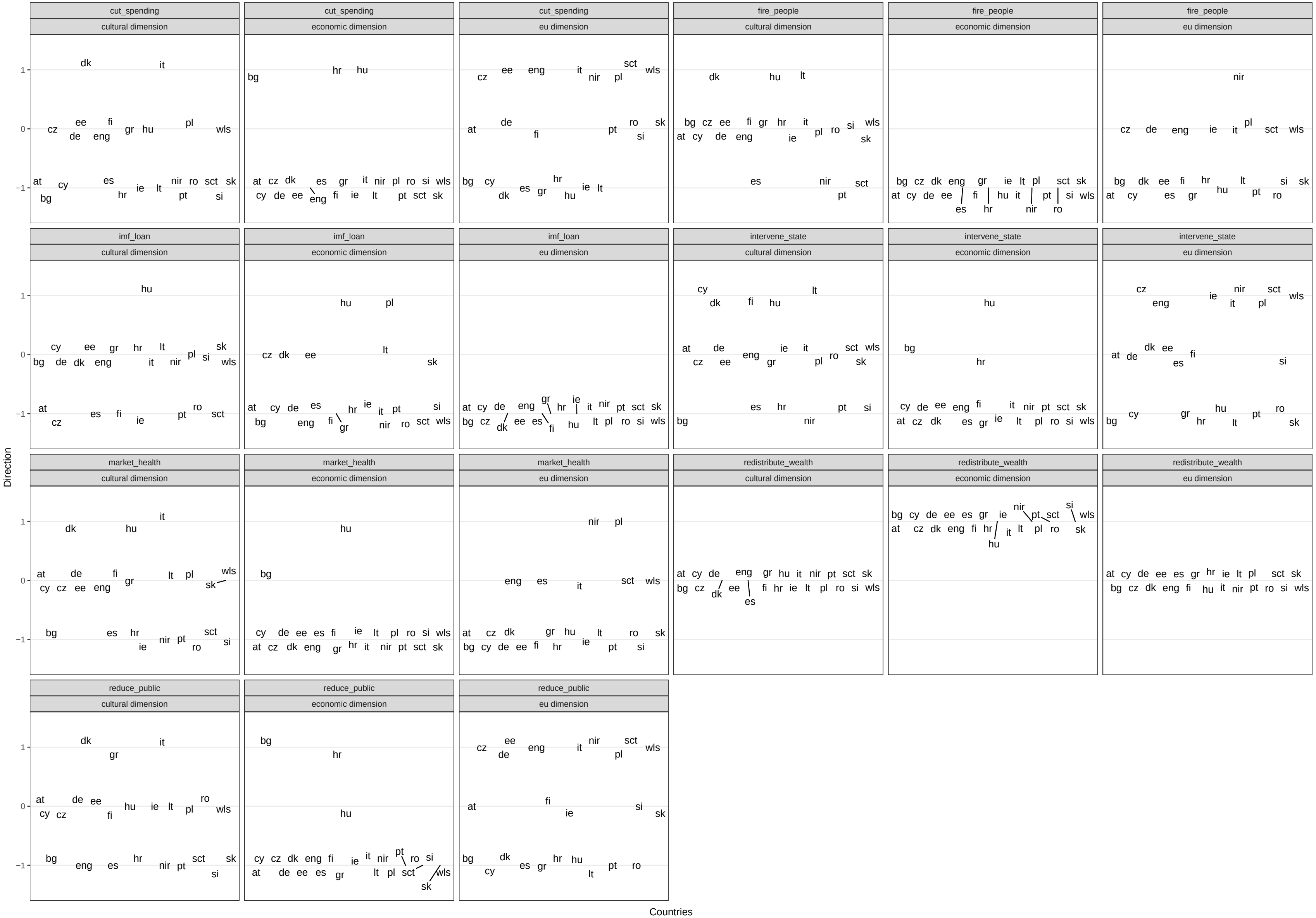


Figure C.1: Discrimination parameters across countries and dimensions for **economic** items.

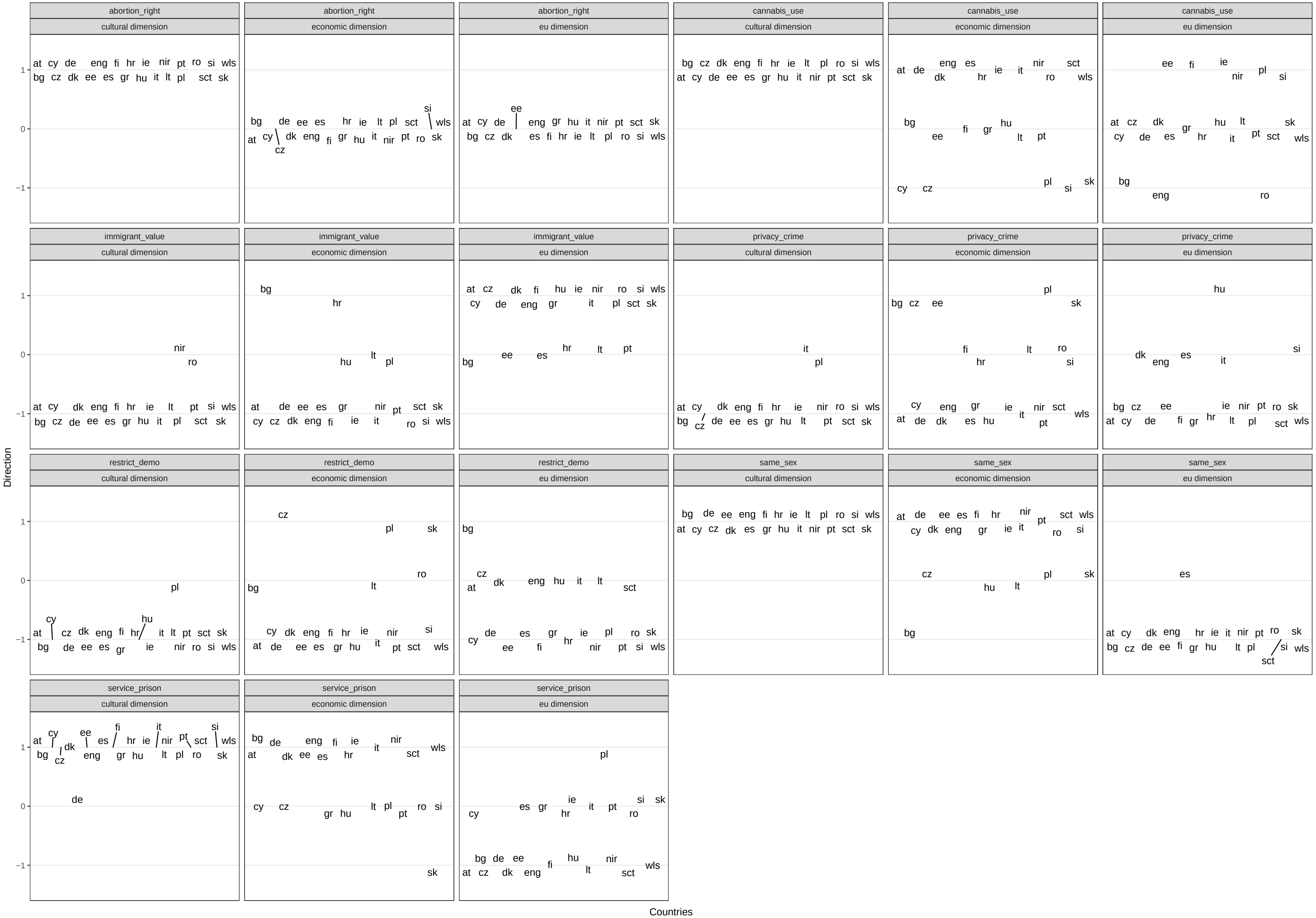


Figure C.2: Discrimination parameters across countries and dimensions for **cultural** items.

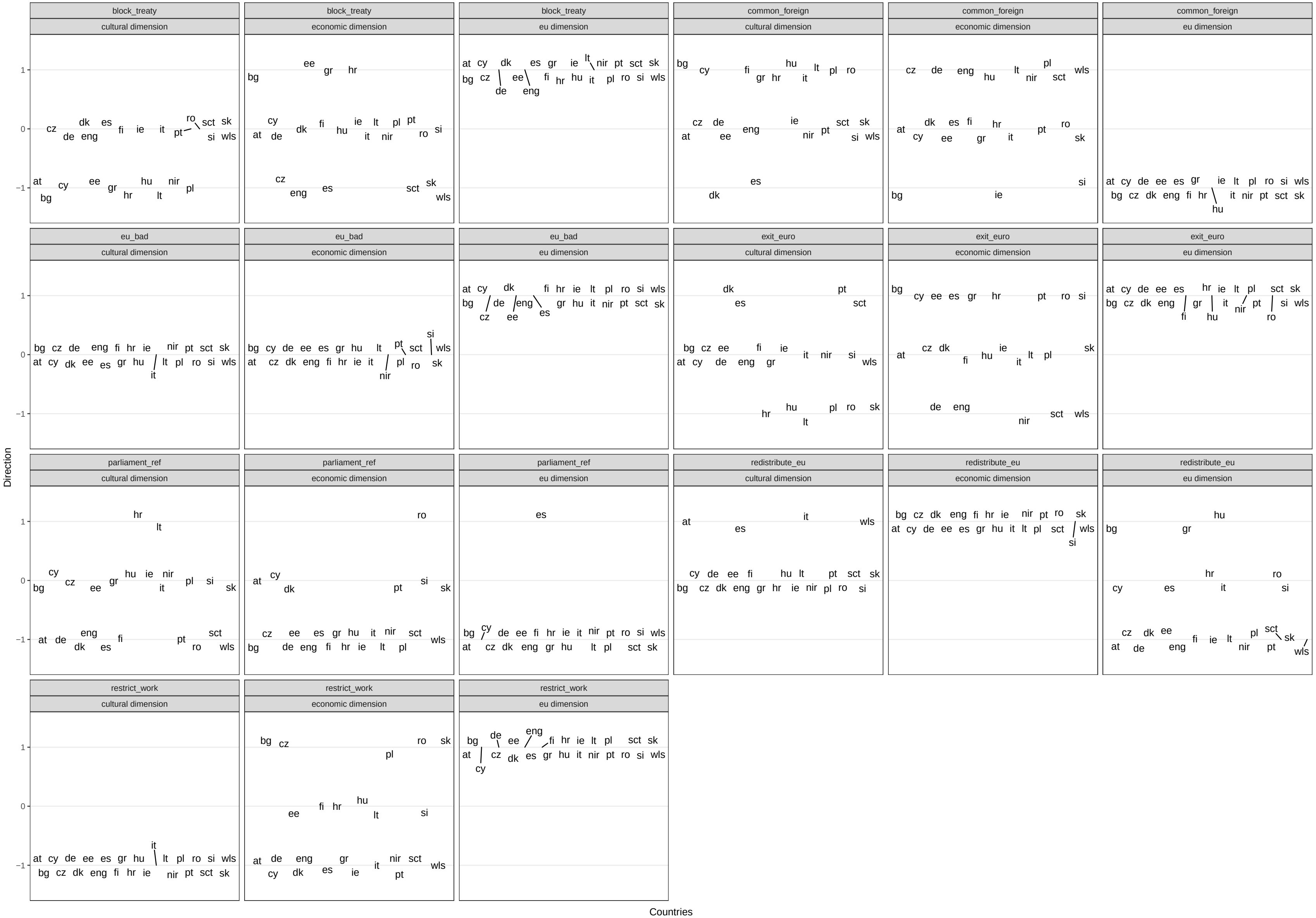


Figure C.3: Discrimination parameters across countries and dimensions for **EU** items.

D Proximity with Reduced Results

This is the same analysis as in Subsections 3.3 and 3.2, but only considering the discrimination parameters of items for the dimensions they were originally selected for (i.e. the discrimination parameter of a cultural item for the cultural dimension, not for the economic or EU dimension).

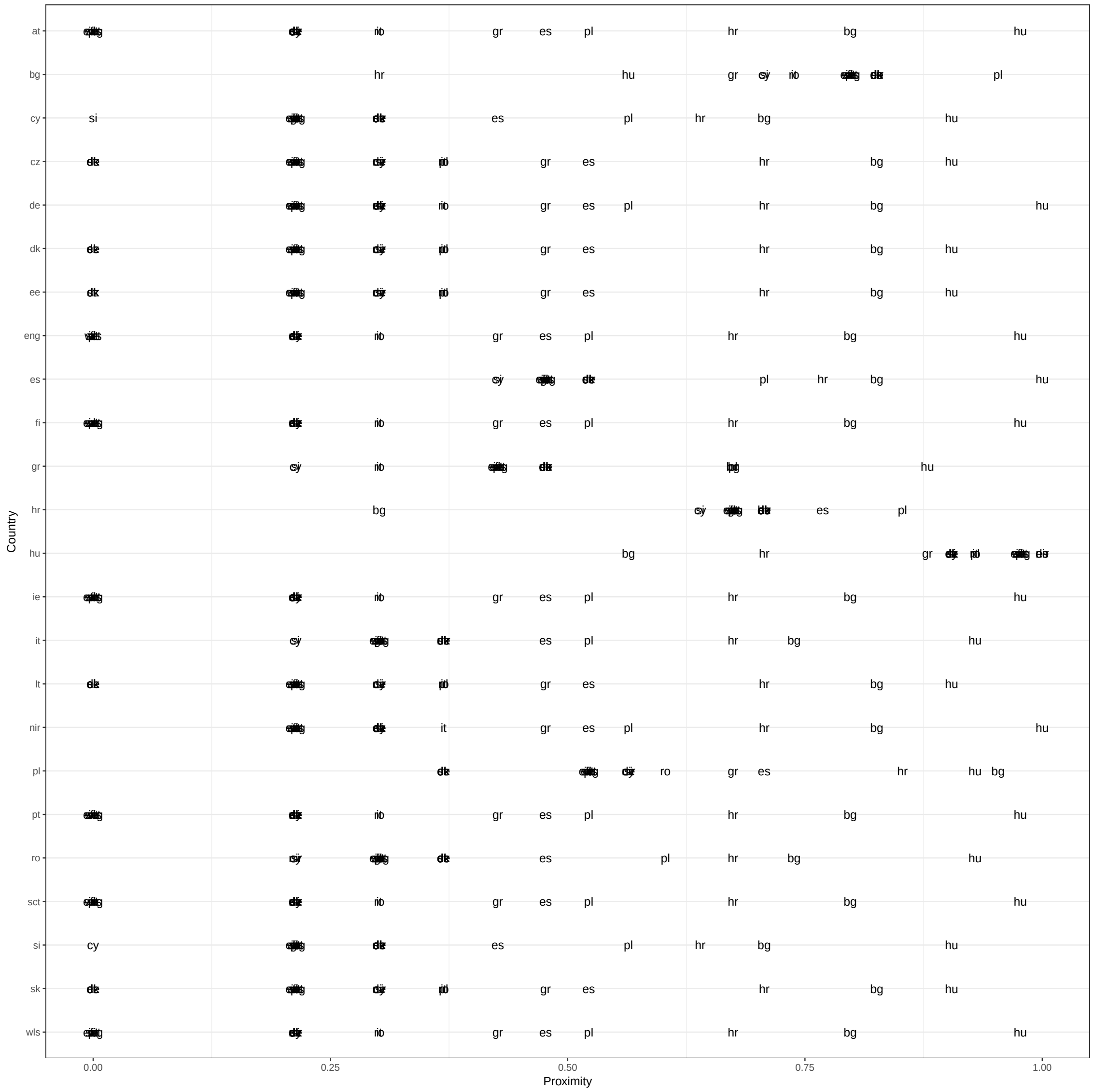


Figure D.1: Proximity of political spaces across Europe.

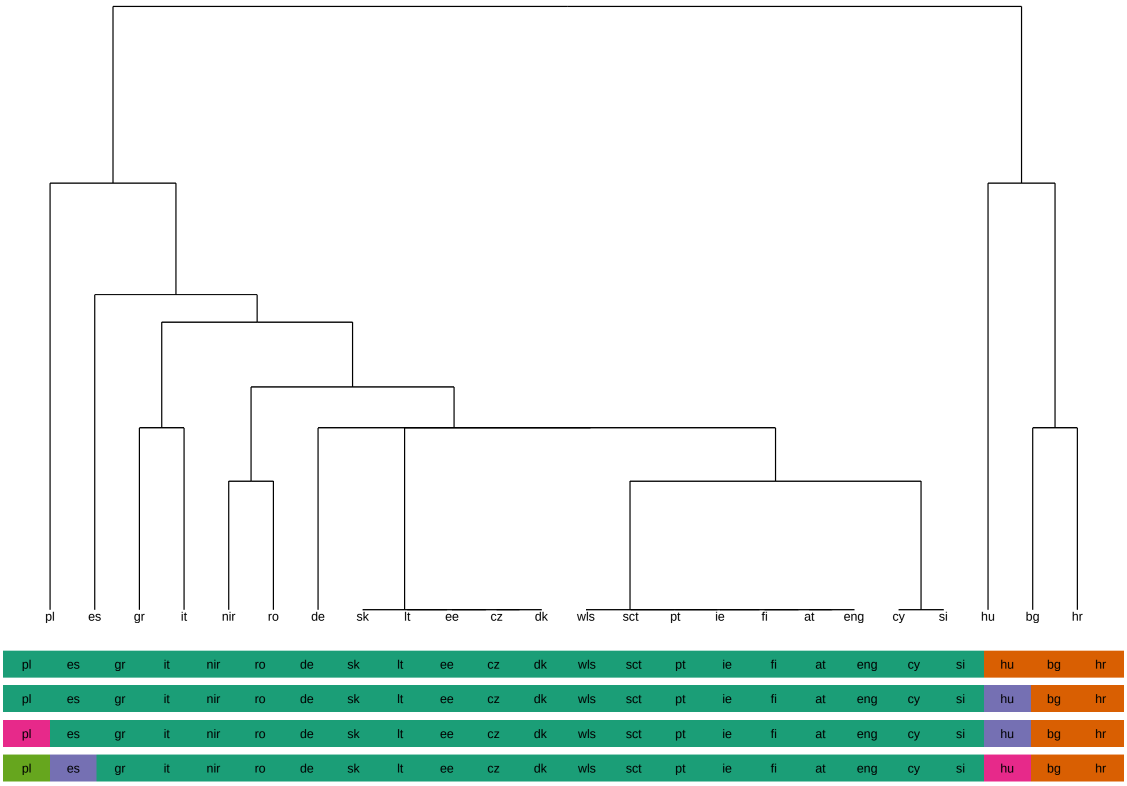
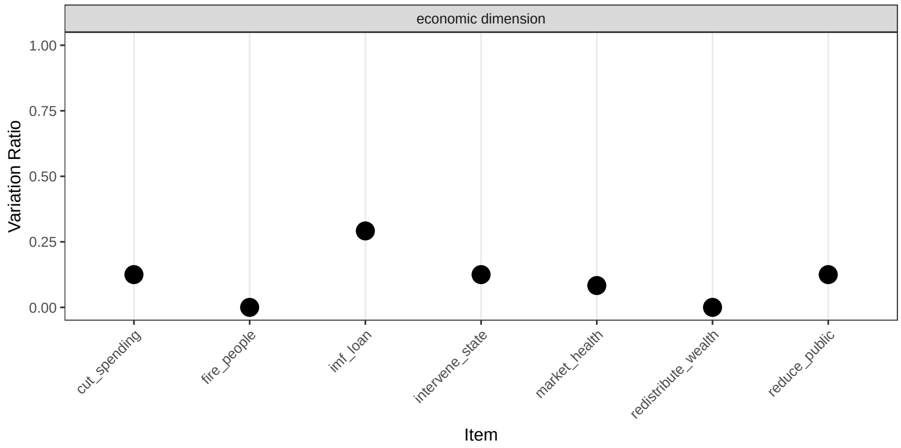
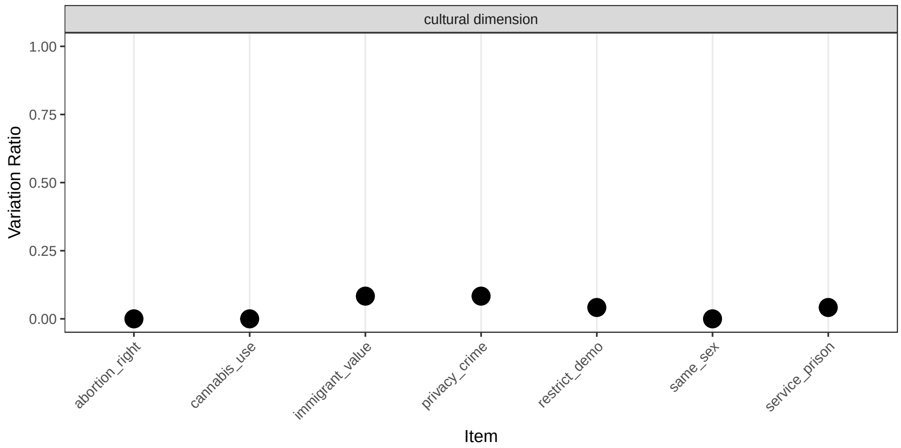


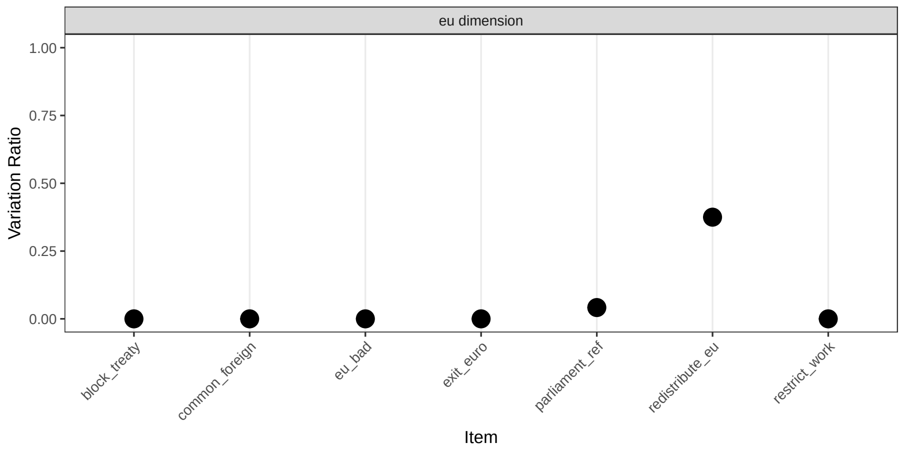
Figure D.2: Clustering of political spaces across Europe based on proximity.



(a) **Economic** items.



(b) **Cultural** items.



(c) **EU** items.

Figure D.3: Variance ratio of discrimination parameters across countries and dimensions.