Methods

The data were acquired from the ESS. Then, the relevant variables were selected and cleaned, including removal of missing values and transformation of round to corresponding year. The working data set consisted of 403,837 respondents from 39 countries across ten survey waves. After the data cleaning, the polarization metrics were calculated, and a separate PCA was performed for each round and country. The loadings and explained variance metrics derived from the first principal component were then considered for further interpretation.

The data were obtained from the website of the European Social Survey Data Portal (ESS Data Portal, 2024) using the inbuilt datafile builder wizard tool which allows to specific selection of the variables, rounds, and countries of interest and extracts the resulting data set as a CSV file. The data for our analysis comprised the following 39 European countries Albania, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Kosovo, Latvia, Lithuania, Luxembourg, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine and The United Kingdom. The data surveyed by the ESS are on a typical Likert scale. This type of scale measures attitudes, opinions, or perceptions and provides a range of options for respondents to choose from. It often ranges from strongly disagree (0) to neutral (5) to strongly agree (10), with shades in between (e.g., Jamieson, 2004). The Likert scale is often used in social science in order to quantify subjective data such as attitudes and satisfaction levels (Journal of Social Sciences, 2022). The focus on was on three variables related to opinions about migration, which ranged from 0 to 10. Thus, 0 would indicate strong resentment, 5 a neutral sentiment, and 10 a strong approval of immigration. Other values encoding questionnaire answers such as 77 = ‘Refusal’, 88 = ‘Don’t know’ and 99 = ‘No answer’ were present in the data as well. Those values were re-labeled as missing data points and subsequently removed them from the data set. The three relevant variables were imwbcnt(‘Immigrants make country worse (0) or better (10) place to live’), imueclt(‘Country's cultural life is undermined (0) or enriched (10) by immigrants’), and imgbeco(‘Immigration is bad (0) or good (10) for country's economy’).

Weights are crucial in survey data as they help to ensure that the survey results accurately reflect the population being studied. Surveys aim to gather information from a sample of the population, but it's often difficult to get a perfectly representative sample. Some groups within the population may be over-represented or under-represented in the sample due to factors like sampling design and responsiveness (some people chosen for the survey don't participate). Weights adjust the data to compensate for these imbalances, giving more influence to the responses of under-represented groups and less influence to the responses of over-represented groups (Pfeffermann, 1996; Ciol et al., 2006). The ESS provides several survey weight variables. The analysis weight (variable name anweight)corrects for differential selection probabilities within each country as specified by sample design, for nonresponse, for noncoverage, and for sampling error related to the four post-stratification variables, and takes into account differences in population size across countries. It is constructed by first deriving the design weight, then applying a post-stratification adjustment, and then a population size adjustment. Starting from Round 9, anweight is provided in the integrated data file (ESS weighting variables, 2024). For data from earlier ESS rounds, anweight was derived by multiplying *pspwght* with pweight in accordance with instructions provided by the ESS (Guide to Using Weights and Sample Design Indicators with ESS Data, 2024).

The data were analysed using the R programming language in version 4.3.2 (R Core Team, 2023). PCA was done using the inbuilt stats (R Core Team, 2023) package. Data manipulation was done using the dplyr (Wickham et al., 2023) and the glue (Hester & Bryan, 2024) packages. The 2-letter country codes were transformed to the full country names and vice versa using the countrycode (Arel-Bundock, Enevoldsen & Yetman, 2018) package. Missing data were handled using the naniar package (Tierney & Cook, 2023). Visualizations were created using the ggplot2 (Wickham, 2016) and the ggrepel (Slowikowsi, 2024) packages and composed using the Patchwork (Pedersen, 2024) package. A complementary web application was created in order to create a visual interface for data exploration using the shiny (Chang et al., 2023), shinyWidgets (Perrier et al., 2025) and shinydashboard (Chang & Borges Ribeiro, 2021) packages.

Polarization metrics

As in the work of DiMaggio et al. (1996), of Evans et al. (2001), and of Bramson et al. (2016), we assess opinion polarization using representative survey data as provided by the ESS. Work by Bauer (2019) attempts to give an overview of like-minded existing approaches, classifying them according to several criteria: response scale type, dimensionality, and investigated distributional characteristics. Regarding Bauer’s (2019) concept of scale types, the questions we use have eleven response options.

imbgeco, imueclt and imwbcnt all range from 0 (extremely negative) to 10 (extremely positive), with 5 being a neutral stance. However, for many of the following metrics, the fractions denoted as had to be calculated first as the relative proportions of (valid) answers for the options zero to ten regarding the survey variable in question. A proportion compares a part to the whole. It indicates what fraction of the total a particular part represents and thus ranges from 0 to 1. The weighted proportions of each level of imbgeco, imueclt and imwbcnt were calculated, taking into account the aforementioned weighting variable anweight.

According to Bauer’s (2019) concept of dimensionality our approach was unidimensional with the exception of the last metric: We measured polarization for one topic at a time as a function of the distribution of valid answers for one country and round, not including missing data, “Don’t know”, “Refusal”, and “No answer” responses. Basis for the assessment of polarization are the various principles (DiMaggio et al., 1996) and axioms (Esteban and Ray, 1994; Duclos et al., 2004) for polarization measures that previous research identified. Bramson et al. (2016) offer an overview as well as formal measures. Inspired by these publications, we used the following six distinct and unidimensional aspects of polarization with the addition of one multidimensional metric we developed ourselves.

Non-neutrality

Non-neutrality (not choosing "5" on a 0-10 scale) indicates polarization by measuring the proportion of individuals with non-impartial views. Increased non-neutrality on immigration implies fewer neutral opinions, with more people leaning towards acceptance or objection. A completely polarized society would lack any neutral stances, aligning with Esteban and Ray's (1994) axiom that polarization increases when population mass moves away from the center. Ademmer and Stöhr (2018) also used a low fraction of central responses as a polarization measure (cf. Abramowitz and Saunders, 2008; Fiorina and Abrams, 2008 in the US). Our reversed measure reflects Draca and Schwarz's (2021) "disappearing centre" effect.

Average deviation from neutrality

where

Average deviation from neutrality, similar to psychological group polarization (how far the average attitude is from the midpoint), is calculated as the absolute difference between the mean opinion and the neutral point (5). While distinct, high average deviation can paradoxically indicate consensus at an extreme, contrasting with the typical view of polarization as the opposite of consensus. The measure is normalized by for a maximum value of one. Regarding public opinion on immigration, a rising average deviation from neutrality indicates increasingly accepting or rejecting individual views. Maximum polarization in this sense occurs when everyone holds an extreme opinion ('0' or '10'). This aligns with Esteban and Ray's (1994) third axiom regarding "shifting population mass from the central mass" and relates to social psychology's group polarization concept, capturing opinion shifts "toward a more extreme point" (cf. Sunstein, 2003). However, maximal polarization here implies extremity-based consensus, contrasting with the subsequent polarization notion.

Dispersion

**Dispersion,** measured by the mean absolute deviation of an opinion distribution, serves as a basic polarization measure for bounded scales (like 0-10 in the ESS). Maximum dispersion occurs with equal halves at both extremes, while minimum dispersion reflects complete consensus. This measure aligns precisely with **Bramson et al. (2016)**.The measure's maximal value is again normalized to 1 by a factor of . Regarding public opinion on immigration, increased dispersion signifies greater individual deviation from the average attitude. Maximum polarization occurs in a society split equally between total acceptance and total objection. The dispersion principle of polarization was introduced by DiMaggio et al. (1996): "Other things being equal, the more dispersed opinion becomes, the more difficult it will be for the political system to establish and maintain centrist political consensus". Prior studies using dispersion as a polarization measure include Adams et al (2011), Bramson et al. (2016), Duffy et al. (2019), and Rapp (2016).

Moderate divergence

with

and

and

and

As Bramson et al. (2016) note, the preceding three aspects don't fully encompass polarization. They propose measures considering group divergence, internal consensus, and size parity, assuming group existence. While US research often uses self-identified partisans as exogenous groups (suitable for the two-party system), this doesn't readily apply to Europe. Bramson et al. (2016) suggest endogenous group formation based on distribution. Lorenz (2017) identified five endogenous groups in ESS opinion distributions: extreme left, moderate left, neutrals, moderate right, and extreme right. We operationalize analogous groups per item: 0 (full acceptors), 1-4 (moderate acceptors), 5 (neutrals), 6-9 (moderate opponents), and 10 (full opponents). Separating scale mid- and endpoints acknowledges their distinct treatment by respondents, with midpoints as neutral and endpoints representing "the most extreme instances" (Tourangeau, 2018). To define the group-based polarization aspects described by Bramson et al. (2016) we use two endogenous groups: The moderate accepting group and the moderate opposing group , as well as the corresponding mean attitudes and

Moderate divergence is then assessed by the absolute difference of group means of the moderate accepting group and the moderate opposing group, as described in Bramson et al. (2016). The factor normalizes the measure. Regarding public opinion on immigration, increased moderate divergence signifies a greater gap between the average views of moderately accepting and moderately opposing individuals. Maximum polarization in this sense occurs when these positions are furthest apart. Prior studies using this measure include DiMaggio et al. (1996) and Fiorina and Abrams (2008). It reflects DiMaggio et al.'s (1996) bimodality principle: "the greater the extent to which opinions move toward separate modes (and the more separate those modes become), the more likely it is that social conflict will ensue".

Moderate group consensus

where

and

**Moderate group consensus** is measured by the mean absolute deviation (MAD) within the two moderate groups. Unlike with the dispersion metric, higher group consensus corresponds to lower MAD within these groups. For public opinion on immigration, increasing moderate group consensus means greater agreement among members of each moderate group. Maximum polarization in this sense occurs when each group perfectly agrees on a single opinion. This aspect was introduced by **Bramson et al. (2016)** and relates to the identification aspect in **Duclos et al.'s (2004)** identification-alienation framework, indicating the coherence of moderate stances.

Moderate size parity

}

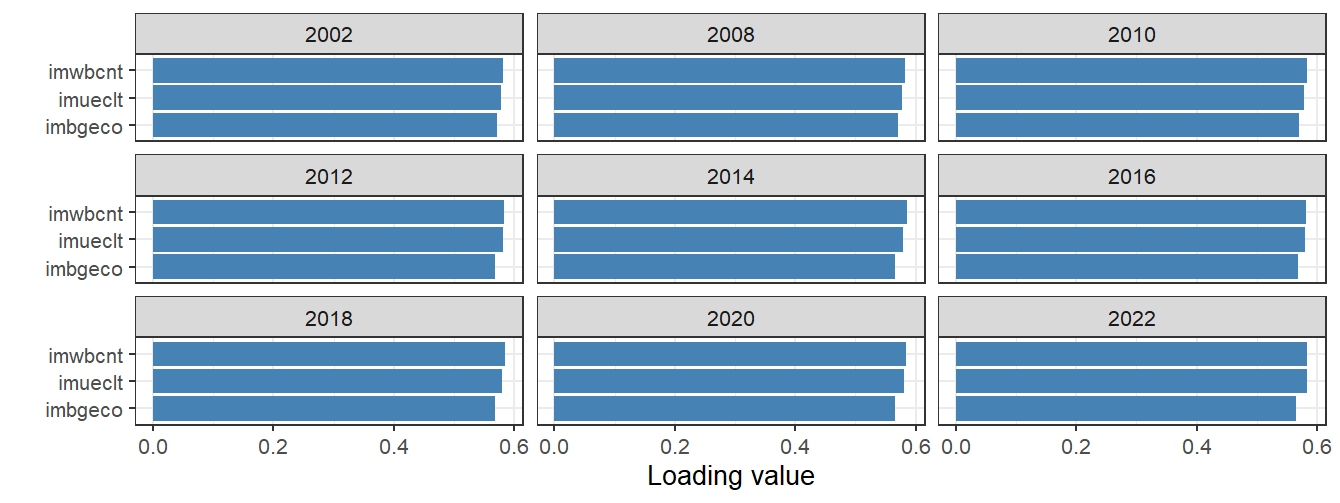
**Moderate size parity** is the ratio of the smaller to the larger moderate group's size. A parity of 1 signifies equally sized groups, indicating maximum polarization in terms of parity. This is a simplified version of **Bramson et al.'s (2016)** measure. For public opinion on immigration, increasing moderate size parity means the number of moderately accepting and opposing individuals becomes more balanced. Maximum polarization occurs when both groups are equal in size. This aspect, introduced by **Bramson et al. (2016)**, conceptually relates to **DiMaggio et al.'s (1996)** bimodality principle and the alienation effect in **Duclos et al.'s (2004)** identification-alienation framework, where equal group size enhances alienation.

Explained variance of the first principal component

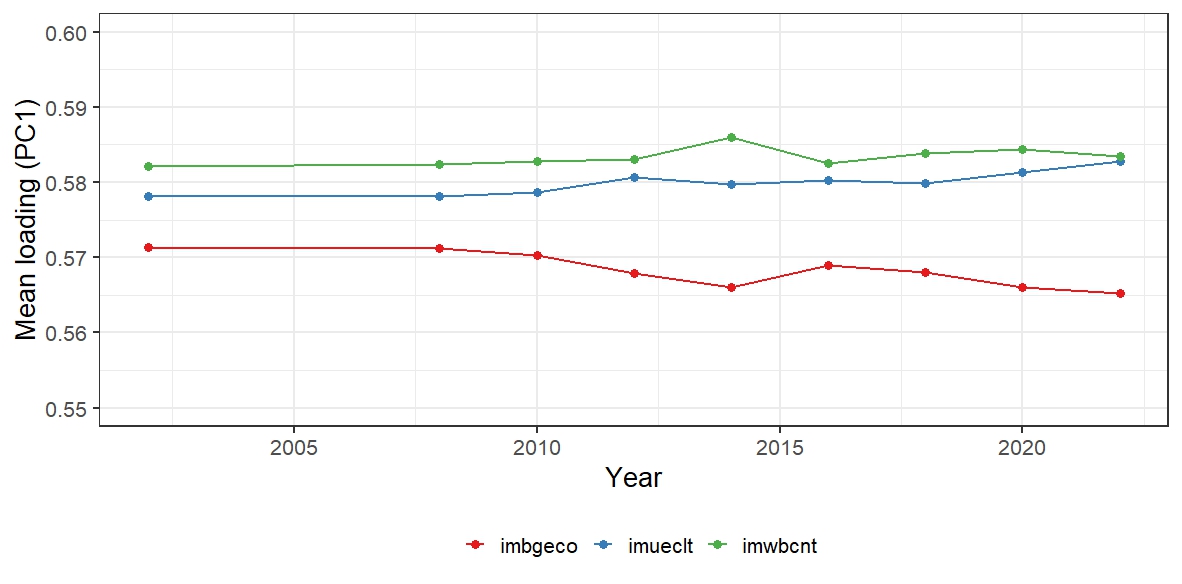
While the previous six metrics are unidimensional in Bauer’s (2019) classification, this seventh metrics follows a multi-dimensional approach by considering imbgeco, imueclt and imwbcnt simultaneously using PCA.

PCA transforms a dataset with potentially correlated variables into a new set of uncorrelated variables: The so-called principal components. These principal components are ordered so that the first few retain most of the variation present in all of the original variables, allowing to simplify complex datasets while preserving crucial information. As the first step of performing PCA, the covariance matrix is calculated to understand the relationships between the considered variables imbgeco, imueclt and imwbcnt. That is, by which amount and direction the variables vary by themselves and together. Then, the correlation matrix is derived from the covariance matrix by standardization. The weights were incorporated by creating the correlation matrix using the analysis weight variable (anweight), resulting in the weighted correlation matrix based on which the following steps of the PCA were done. Next, the eigenvectors and eigenvalues of the covariance matrix are computed using methods from linear algebra. The eigenvectors become the directions of the principal components, and the eigenvalues indicate the amount of variance explained by each component. By selecting only the top few principal components - those with the highest explained variance -, it is possible to reduce the dimensionality of the original data (e.g., Abdi & Williams, 2010; Gewers et al., 2018). In essence, PCA is a way to find the most important patterns in a complex dataset by finding the directions of greatest variance, and then using those directions to represent the data in a simpler way. To use an analogy, one could think of a shadow shining a light on a cloud of dots representing data points. The shadow on the wall is a simplified version of the cloud. PCA is like finding the best angle to shine the light, so the shadow captures as much of the original shape as possible. Now, imagine stretching the cloud along its longest stretch and squeezing it along its shortest stretch. PCA is like finding the right stretches and squeezes to simplify the cloud.

As mentioned before, each principal component (PC) has a certain amount of explained variance. We focussed on the first PC and assumed that it would reflect a general migration attitude, which we could confirm by looking at its variable loadings (Fig. xx)



**Fig 2.**x Mean PCA loadings across all countries and years. The data were centered and scaled. PC1 seemed to be a general migration attitude with larger values indicating a pro-immigration attitude and smaller values indicating an anti-immigration attitude.



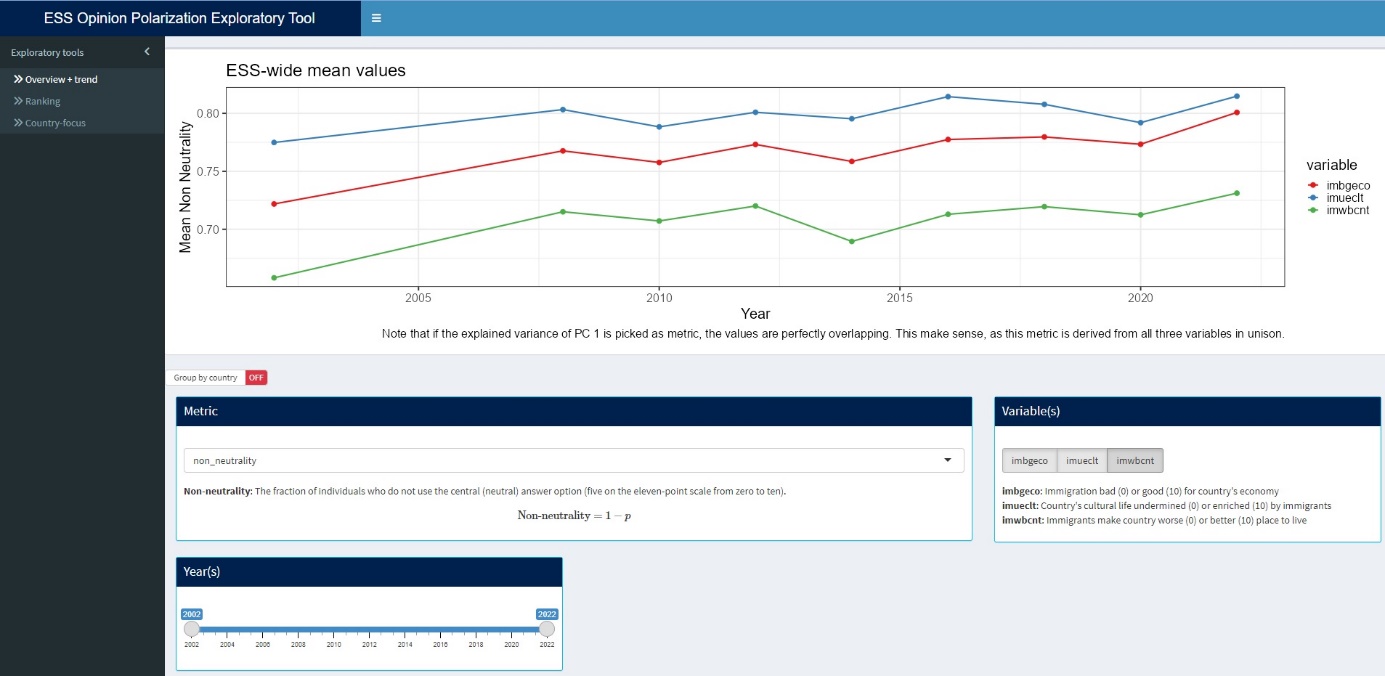
**Fig 2.x** Mean PC1 Loadings of imbgeco, imueclt & imwbcnt across all countries over time. The corresponding loading values of the three variables have been consistently between 0.55 and 0.60 over all rounds.

On the country level, the loadings of PC1 for the three considered migration variables ranged similarly stable between 0.54 and 0.60, with some smaller and larger differences between the countries (see appendix). The consistency of the loading values of PC1 ensured that there was no between-year variation of the meaning of PC1. The seventh metric was thus the explained variance of the first PC for a given year and country. This, in turn, made it also possible to use the slope of the explained variance of PC1 later on as an indicator for an ongoing trend (or the lack thereof) of issue alignment or heterogenization regarding the three migration-opinion variables.

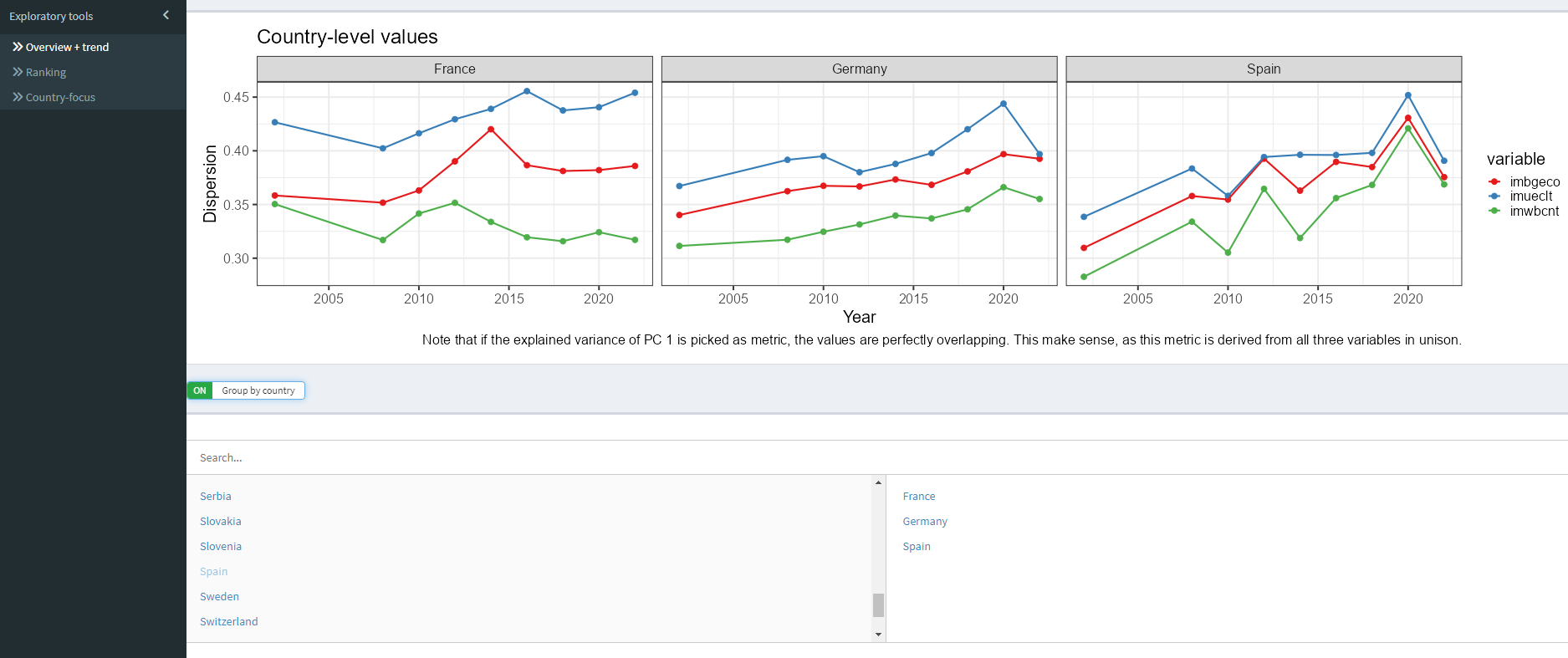
We were now able to identify trends in the form of a slope coefficient by assessing the explained variance of PC1 over time using linear regression. Linear regression attempts to find a linear relationship between a dependent variable (the variable one wants to predict or model) and one or more independent variables (the variables use to make the prediction). In simple terms, it tries to fit a straight line to a set of data points. The "best fit" line is determined by minimizing the differences between the predicted values (from the line) and the actual values of the dependent variable. A common method for doing this is the "least squares" method (e.g., Su, Yan & Tsai, 2012; Kumari & Yadev, 2018). For only one independent variable, the line follows the form where m is the slope coefficient describing the average change in y for a 1-unit change in x. For the research question at hand, linear regression was used to model the explained variance of the first principal component over time. This, in turn, made it possible to use the corresponding slope coefficient as an indicator for an ongoing trend (or the lack thereof) of issue alignment or heterogenization regarding the three migration-opinion variables.

Web app for data exploration

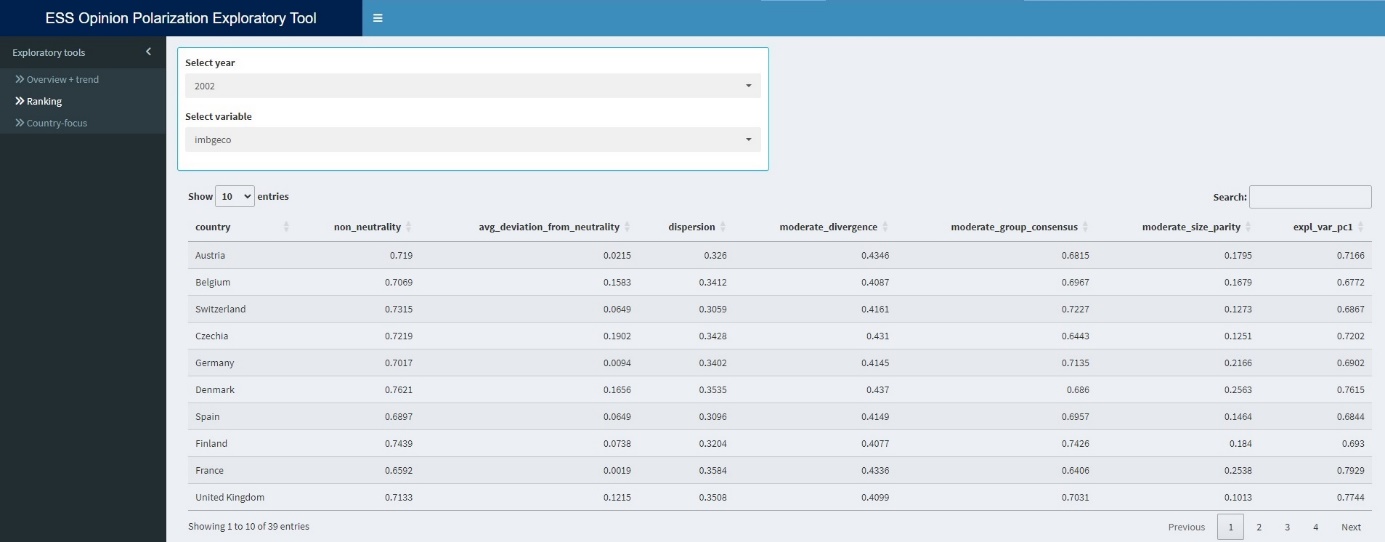
Investigating polarization dynamics across seven distinct metrics observed over ten time points spanning two decades (2002-2022) for three key variables presented a substantial analytical challenge, particularly when conducted at both continental and individual country levels. Consequently, a dedicated tool with the ability to facilitate effective data exploration, filtering, and visualization was built prior to the substantive analysis. With the data obtained from the ESS, this web application allowed to assess the temporal development of each metric and variable on both the continental and country levels. Additionally, it made it possible to rank the data using any of the metrics and variables for a given year. Lastly, it allowed to focus on a given country-year-variable combination, yielding insights into all metrics in comparison to the European average along with a view into the distribution of opinions (0 to 10 on the Likert scale) for the same time frame (Figs. 2.xx to 2.xx).

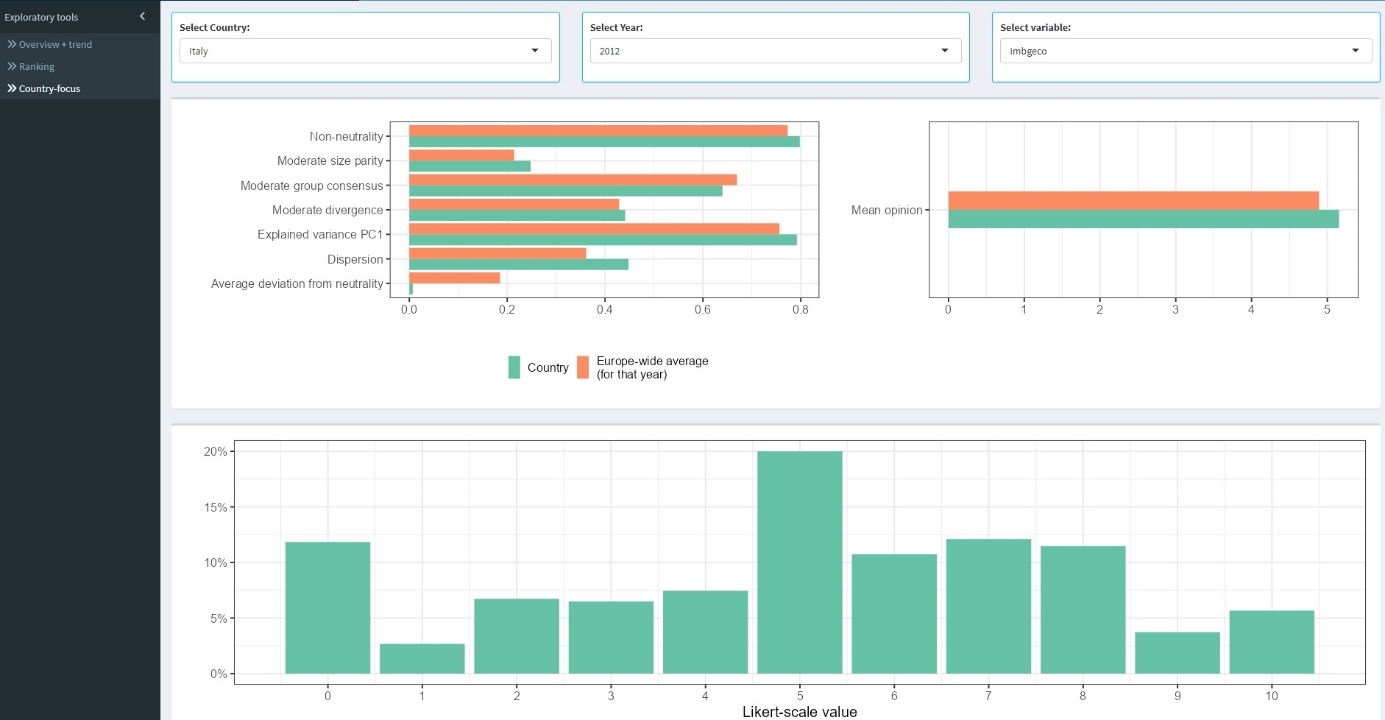


**Fig 2.x** The “Overview + trend” page of the exploratory web app, showing the development of the non-neutrality metric for all the opinion variables between 2002 and 2022 using the Europe-wide average.



**Fig 2.x** The “Overview + trend” page of the exploratory web app, showing the development of the dispersion metric for all the opinion variables between 2002 and 2022 separately for France, Germany and Spain.

**Fig 2.x** The “Ranking” page of the exploratory web app, showing all metrics for imbgeco in 2022. This table could now be sorted based on any of the metric, resulting in a corresponding country-ranking.



**Fig 2.x** The “Country-focus” page of the exploratory web app, showing the development of all metrics for imbgeco in 2012 for Italy in comparison with the Europe-wide averages. The bottom plot shows the distribution of the Likert-Scale values for the same time frame.