ROYAL SOCIETY OPEN SCIENCE

royalsocietypublishing.org/journal/rsos

Research

Cite this article: Young DJ, Ackland J, Kapounek A, Madsen JK, Greening LJ, de-Wit L. 2026 A new measure of issue polarization using *k*-means clustering: US trends 1988–2024 and predictors of polarization across the world. *R. Soc. Open Sci.* **13**: 251428.
https://doi.org/10.1098/rsos.251428

Received: 29 July 2025
Accepted: 22 October 2025

Subject Category:
Psychology and cognitive neuroscience

Subject Areas:
Psychology

Keywords:
polarization, issue, *k*-means, clustering, ethnic fractionalization, political

Author for correspondence:
David Jack Young
e-mail: dy286@cam.ac.uk

Supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.8190166.

THE ROYAL SOCIETY PUBLISHING

A new measure of issue polarization using *k*-means clustering: US trends 1988–2024 and predictors of polarization across the world

David Jack Young[1], James Ackland[1], Andreas Kapounek[1], Jens Koed Madsen[2], Lara Jane Greening[1] and Lee de-Wit[1]

[1]Department of Psychology, University of Cambridge, Cambridge, UK
[2]Department of Psychological and Behavioural Science, London School of Economics and Political Science, London, UK

DJY, 0000-0002-2707-8284; JA, 0000-0002-4722-6301; AK, 0009-0003-0445-4421; JKM, 0000-0003-2405-8496; LJG, 0000-0001-6952-7116; Ld-W, 0000-0003-3048-2875

Political issue polarization worries scholars and the public alike. To understand what drives political issue polarization, longitudinal analyses and cross-national comparative research are necessary, but difficult to implement using current measures. We propose a new technique for measuring political issue polarization which is well suited to longitudinal and comparative analyses, using a machine learning algorithm called *k*-means clustering, which identifies coherent groups of politically-like-minded citizens from the bottom up. We analyse the between-cluster separation, within-cluster cohesion and size parity of the clusters to quantify a society's political issue polarization. Using American National Election Studies data, we find that polarization increased in the USA from 1988 to 2024, driven by a period of rising separation between clusters from 2008 to 2020. Using World and European Values Survey data, we find that across the world, mass issue polarization is driven primarily by disagreement over cultural issues, but manifests differently depending on a society's level of Human Development Index (HDI), with lower-HDI countries seeing culturally conservative clusters account for a majority of citizens, and higher-HDI countries having more culturally liberal and equally sized clusters. Different societal-level predictors, including ethnic fractionalization, wealth inequality and HDI, are associated with different aspects of polarization.

1. Introduction

Political polarization is a multi-faceted topic which interests a large multi-disciplinary body of scholars, spanning political science [1–5], psychology [6–9], philosophy [10–12], economics [13,14], sociology [15–18], geography [19], environmental science [20], computer science [21] and beyond. It is also a topic of immense popular concern, with the World Economic Forum [22] naming 'societal polarization' the third-greatest short-term risk facing humanity in 2024, and books about polarization becoming best-sellers (e.g. [23]). Much research has focused on whether polarization is rising in the USA [17,24], or has looked at the potential drivers of polarization, considering political [25], economic [26,27], technological [13,28–30] and institutional factors [31], as well as rational [32] and irrational aspects of human psychology [4,10,33].

Polarization has a straightforward dictionary definition: 'division into two sharply distinct opposites' (Merriam-Webster [34]);[1] it can refer to the *fact* of something being in a polarized state, or the act of it *becoming* so. A major worry is that if democratic societies become divided into camps which are too different in terms of what they want from government, democratic institutions may erode, with parties emboldened by their supporters to rewrite rules and norms in order keep opponents from power [37].

However, the accumulation of scholarly knowledge about polarization is complicated by enormous heterogeneity regarding what the term is used to refer to. A prominent division is made between *affective polarization*, which refers to how much the major political groups within a society *dislike* each other [2], and *issue polarization*, which concerns the level of *disagreement* over political issues, either among the public [3,38] or among elites [25]. Even within discussions focused on one of these sub-types of polarization, confusion can occur due to the numerous different approaches that have been taken to operationalizing and measuring levels of polarization.

In recent years, promising steps have been taken to develop measures which allow us to perform meaningful comparative (i.e. comparing different societies) and longitudinal (i.e. tracking trends in the same society over time) analyses of affective polarization [1,14,39–41] and issue polarization among elites [42–44]. But while there is a wide range of existing measures of issue polarization among the public ('mass issue polarization'), each of these has distinctive limitations involving validity, intuitiveness and whether they are well suited to affording comparisons between societies and over time. These limitations restrict our ability to meaningfully track mass issue polarization trends and identify driving factors.

A prominent index of mass issue polarization has recently been proposed by the Edelman Trust Barometer [45], where a country's level of polarization is based on whether citizens say they think their country is divided, and whether they think those divisions can be overcome. But these perceptions do not provide an *objective* assessment of the extent of political division within a country, particularly due to the phenomenon of 'false polarization', the finding that people consistently over-estimate the degree of issue separation between members of different political groups [46]. Therefore perceptions of division may not be informative about genuine mass issue polarization.

Many measures of mass issue polarization utilize data on citizens' positions towards political issues, and therefore offer a more objective indication of division, but nevertheless possess other limitations. For example, Fiorina [24,38] measures polarization by calculating the proportion of respondents who either say 'don't know' or give a middle-of-the-scale response when asked to give their position on political issues. They average the proportion of such responses across multiple issues and suggest that polarization is indicated by a low proportion of these responses. Yet, this arguably conflates the issue of polarization with extremity—if everyone in a society had the *same* position on every issue, or very similar positions, they would very clearly *not* be polarized, yet they would not necessarily provide a high proportion of middle-of-the-scale or 'don't know' responses, as everyone may have a relatively strong, confidently held opinion.

Another approach is to measure the *variance* in the positions taken across a society towards a set of issues [27,47,48]. The limitation of this approach is that it does not take into account the multi-dimensional structure of people's issue positions. Imagine one society where everyone's political issue positions are effectively random—the average variance is likely to be quite high, because for each issue, the full range of issue positions is taken. But consider another society which is divided into two

[1] See also similar definitions given in different dictionaries: 'the act of dividing something, especially something that contains different people or opinions, into two completely opposing groups' [35], 'The accentuation of a difference between two things or groups; division into two sharply contrasting groups or sets of beliefs or opinions; an instance of this' [36].

groups, with everyone in one group slightly disagreeing with a set of issue position statements, and everyone in the other group slightly agreeing—the average variance may be much lower, because only two, not-too-distant positions are taken towards each issue. Yet, the society is clearly more polarized now because there are two groups who consistently disagree with each other across a range of issues, rather than the unstructured free-for-all of cross-cutting agreements and disagreements of the first scenario.

Another approach, which is to calculate the *bimodality* of positions across individual issues [3], has a similar limitation, because even when we find bimodal distributions for individual issues, we cannot be sure whether this is because of society-wide polarization between groups who *consistently* disagree with each other across the full range of issues, or if citizens are grouped together differently across issues.

Inter-item correlations across issues can be an insightful metric [16,17], with high correlations taken to indicate polarization. This data point helps us to understand how opinions are distributed within societies, and is of clear relevance to mass issue polarization. However, it is somewhat removed from the view of polarization as division between two groups, as it makes no reference to groups. Furthermore, theoretically, increased correlations could occur in some cases when more citizens adopt the same singular ideology if opinions are initially quite randomly distributed, or if three or more ideological groups form, which might not cause polarization in the form of sharp division between two groups. There is therefore a need for measures which are concerned with division between two groups specifically.

Some such measures have been deployed in polarization research, including measures based on the division between the two major political groups in a society on policy issues (e.g. [3,49–52]). This research has typically been conducted in the USA, where identifying these groups appears straightforward: take self-identified Democrats and Republicans, and compare their issue positions. Researchers have quantified the degree of polarization in terms of the average distance between them across issues [49–52], and others have looked at the degree of overlap [3].

But using self-identified groups like this has two limitations. For one, there may be a confound of 'sorting' when this approach is used to track longitudinal trends [17,53]. Suppose that, over time, the distribution of political beliefs in a society does not change, but people get better at figuring out which party's positions lie closest to their own, and so become more likely to identify with the party nearest to their preferences. If we then measured the average distance between supporters of the two parties over time, it would increase (and the degree of overlap decrease), just because there are fewer people who identify with the 'wrong' party as time goes on, meaning polarization appears to increase even though the underlying distribution of issue positions has not changed.

The second limitation is that outside the USA it is significantly harder to determine which two political groups to apply the measure to. This is because the USA is a global outlier in having only two large parties, who are roughly equal in size. Most other democracies have more than two parties, which differ in size, and may form superordinate groups like blocs and coalitions, and some democracies have only one large party with numerous smaller parties (e.g. South Africa and Japan), whereas non-democracies may have only one party (e.g. China) or none (e.g. Saudi Arabia). In these cases, it is difficult, even impossible, to identify the society's two major political groups. This makes it difficult to deploy this measure in a comparable way across different countries.

Despite these challenges, we suggest in this article that this broad way of measuring of polarization—calculating the distance between the two major political groups in a country over political issues—can be adapted in a way that overcomes these limitations, to provide a clear and intuitive measure of polarization well suited to longitudinal and comparative analyses. This approach is attractive because it clearly aligns with the standard definition of polarization by measuring how much a society's political groups have become split into 'two sharply distinct opposites' in terms of policy preferences. With a topic as inter-disciplinary and of interest to the general as polarization, it may prove useful to utilize measures which align with common understandings of the term.

We suggest that rather than operationalizing these groups as political parties, blocs or movements which participants self-identify as supporting, progress can be made by adapting the measure to instead operationalize them as statistical 'clusters' of people within a country's political issue position space, which can be identified from the bottom-up using clustering algorithms. This avoids the need to use self-identification data, averting the problem with longitudinal analyses caused by the confound of sorting, as well as avoiding the need to identify two major political groups prior to analysis, which avoids the problem with comparative analyses caused by the difficulty of doing this in most countries.

Specifically, we use the *k*-means clustering algorithm to split samples of political issue position data into *two* clusters of citizens. We then perform some comparisons of the two clusters to estimate the level of polarization within the country from which the sample was taken. The *k*-means clustering algorithm is designed to partition points within continuous multi-dimensional feature spaces into groups such that the points in the same group are maximally similar, and points in different groups maximally different, in terms of their Euclidean distance from one another [54], and is there well suited to identifying coherent sub-groups of similar individuals within multi-dimensional spaces. While clustering algorithms are not often used in social sciences, they are, for example, increasingly used in computational psychiatry to group individuals with similar psychiatric and phenotypic profiles [55,56] to aid diagnosis and treatment.

It is worth noting though that *k*-means assigns *all* points to one cluster or another—no individuals can be 'left out'. This creates an important qualitative difference between our approach and using people's self-identified party affiliations to define a society's major political groups, as even in the USA, many people do not self-identify as supporting the Democrats or Republicans, and so would usually be excluded from analyses of polarization levels. Yet, it is advantageous to capture variation in political opinion data from these people, rather than throwing it out—a large, apolitical 'centre', for instance, does represent a depolarizing force for society at large.

Once we have found the clusters in a given sample of political opinion data using *k*-means, we calculate the average distance between the issue positions of the people in each cluster, a measure we call 'Separation'. The more issues the clusters take different positions towards, and the greater the distance between those positions, the higher this measure will score.

However, this is not the only measure we calculate to inform our analysis of polarization. We suggest that two other features of the clusters can exacerbate or diminish the level of mass issue polarization, giving us three facets of polarization we consider.

The second facet of polarization is how 'spread out' each cluster is across the issue position space—in our view, the greater the spread of views within each cluster, the less a society can be said to be polarized, because there is a greater level of internal disagreement within each political group. We quantify the degree of spread within clusters using a measure we call 'Dispersion', measured as the average distance between each individual and the centre of their cluster, across issues.

The third facet of polarization is how equally sized the clusters are—in our view, if one cluster is much larger than the other, then the society, rather than being polarized, has for the most part coalesced around one set of opinions, which is held by the majority, but with a minority group who are in dissent. We quantify how equally sized the clusters are using Shannon entropy [57][2] and term the resulting measure 'Equality-of-Size'.

Our suggestion that separation between groups needs to be considered within the context of the Dispersion within the groups and Equality-of-Size is consistent with several prior approaches. All three are noted as prior measures of inter-group polarization by Bramson *et al*. [58], with their 'Group Divergence' sense of polarization analogous to between-cluster Separation, 'Group Consensus' analogous to within-cluster Dispersion, and 'Size Parity' analogous to Equality-of-Size. Furthermore, two sets of authors have previously argued that we should consider these three aspects of polarization, or close analogues of them, simultaneously when measuring polarization. Esteban & Ray [59, p.824] proposed that societal polarization was indicated by 'heterogeneity across groups', i.e. high between-cluster Separation, 'homogeneity within each group', i.e. low within-cluster Dispersion, and 'a small number of significantly sized groups', which implies that smaller dissenting groups contribute less towards society being polarized, a logic captured by our Equality-of-Size measure. Bauer [60], in a review of measures of polarization, posits a 'general principle': 'Polarisation is at a maximum when individuals cluster in two cells [and] when the two cells represent the endpoints on [the measured scale(s)] [...]. Polarisation is at a mini[m]um when all individuals cluster in a single cell' (p. 3), which points towards groups forming identifiable clusters, i.e. having low within-cluster Dispersion, as well as groups being far apart, i.e. having high between-cluster Separation, and people not all being in the same group, i.e. low size inequality between clusters (high Equality-of-Size). Overall, we suggest that mass issue polarization is greatest when a society can be split into two opinion clusters which are equally sized and internally cohesive, but far apart from each other. While previous reviews and theory articles justify this approach [58–60], we go beyond the prior literature by suggesting a method for identifying the groups required to deploy the measures in practice (*k*-means), a practical issue which

[2]Where entropy, H, is given by the formula $H = -\sum p(x)\log p(x)$.

had not previously been solved, and presenting a novel empirical analysis of polarization data using all three measures.

The purpose of this article then is to introduce our approach to measuring mass issue polarization using these three facets—Separation, Dispersion and Equality-of-Size—and to demonstrate how it can be used to shed light on important questions. In both studies, we also perform some additional analyses to better understand how polarization manifests in the studied samples.

The first question, addressed in Study 1, is whether mass issue polarization has increased in recent decades in the USA. Some analysts suggest it has [17,49,52] others suggest it has not [3,24,38,61]. Given the heterogeneity of approaches taken, and noted limitations of previous measures, it is difficult to resolve this debate. We apply our methodology to American National Election Studies (ANES) data from 1988 to 2024.

In Study 2, we address the question of which societal, institutional and technological factors are associated with higher levels of polarization. To address this question properly, comparative analysis across countries which meaningfully differ in terms of these factors is necessary, and since our approach does not require countries to only have two parties to submit to analysis, it provides a tractable and intuitive way of measuring mass issue polarization cross-nationally. We study the association between our three measures of polarization and 14 candidate factors across 114 samples of World Values Survey (WVS) and European Values Survey (EVS) data taken from 57 countries between 1999 and 2018, and gain some broader contextual insights from a larger sample of 247 samples from 105 countries, from 1999 to 2022.

2. General method

In both studies, we analyse samples of multi-dimensional issue position data from questionnaires where a sample of participants have been asked to state their positions on numerous political issues.

Our first data pre-processing step is to impute missing or non-applicable values. To find the clusters, we then standardize the variables using *z*-scoring, and reduce the resulting dataset to three dimensions using principal components analysis (PCA). Standardization is a recommended precursor step to PCA [62], and we use PCA for two reasons. First, dimensionality reduction improves the power of clustering analyses [63], and second it reduces bias introduced by item selection by mapping measured items onto latent ideological dimensions. We use PCA as a standard, simple, and well-known method for dimensionality reduction. The choice of three dimensions was taken because in both studies, our issues are grouped into three thematic domains (we also conducted our analyses after reducing the data to the first two principal components, and without performing any data reduction before clustering, finding highly correlated polarization scores, all within-study $r \geq 0.995$; in both studies that follow, the choice of dimensionality reduction has no impact on statistical significance except for one case noted in the text—see footnote 5). We then use k-means to split the sample into two clusters. We use the Hartigan–Wong k-means algorithm, which performs better than alternatives in finding global rather than local minima [64]. We specify $k = 2$, with 1000 random starts, and a maximum of 100 of iterations.

To calculate Separation and Dispersion we return to the post-imputed dataset and normalize the variables, such that responses at the theoretical end-points of the scale score 0 or 1, with all other responses evenly spaced between, and use the resulting dataset without any further standardization or dimensionality reduction. This is so that the scores will be on a comparable scale across countries and timepoints. To calculate Separation, we first find the mean position of each cluster on each issue, we then find the absolute difference between these means for each issue, then we average these absolute differences. This measure therefore ranges from 0 to 1, and is higher the further apart are the mean positions of the clusters. Therefore, higher Separation scores indicate higher polarization.

To calculate Dispersion, we first find each person's residual difference between their position and the mean position of their cluster for each individual issue, we then find the mean absolute residual across issues for each individual, and finally average these mean absolute residuals across individuals. Dispersion therefore ranges from 0 to 1, and is higher when there is more disagreement within clusters. As such, higher Dispersion scores indicate *lower* polarization.

To calculate Equality-of-Size we only need look at the proportion of the sample assigned to each cluster. We randomly assign one cluster to be 'Cluster 1' and the other to be 'Cluster 2', then apply the formula Equality-of-Size = $P_1 * \log(1/P_1) + P_2 * \log(1/P_2)$ (using log base = 2), which is a restatement of the formula for Shannon entropy, where P_1 and P_2 are the proportions of the sample in Cluster 1 and

Cluster 2, respectively. Equality-of-Size therefore also ranges from 0 to 1, and is higher the closer the clusters come to splitting the population 50:50. As such, higher Equality-of-Size scores indicate higher polarization. Further information about our methodology can be found in electronic supplementary material, appendix A, and study-specific details are available in the sections for each study.

All analyses are performed in R [65]. Our scripts and electronic supplementary material can be found via: https://osf.io/kzd23/?view_only=423559a7c7274c269817978beaa05f18. The electronic supplementary material and references together contain citations for each dataset used.

3. Study 1

In Study 1, we measure Separation, Dispersion and Equality-of-Size over time in the USA from 1988 to 2024. We also construct two alternative measures of inter-group mass issue polarization to afford a robust analysis of whether US issue polarization has increased, and perform some further analyses to explore how polarization manifests at each time point.

3.1. Method

We use 10 waves of data from the ANES. The first nine waves, from 1988 to 2020, come from the ANES Times Series Cumulative Data File [66], comprising 30 044 observations (minimum $n = 1212$ per wave, in 2004); the tenth wave, from 2024, comes from the Time Series Study Preliminary Release [67], using the $n = 4964$ who completed both the pre-election and post-election questionnaires (as we require data from both).

To define our political issue space, we initially long-listed variables that had face validity as measures of belief about policy-relevant issues. We also enforced a metrical constraint that items must be measured by a numeric rating or Likert-style scale with at least four levels. Finally, we removed from the shortlist any item that had over 60% *NA* responses (missing responses, and any non-standard response, like 'don't know' or 'haven't thought about it much', were coded *NA*) in any of the waves.[3] This left us with 14 items spread across three domains, shown in table 1.

We use the waves from Presidential election years, with the exception that because four of the African-American inequality items were missing from 1996 (VCF9039, VCF9040, VCF9041, VCF9042), we used 1994 instead.

3.1.1. Alternative measures

We also calculate two alternative measures of mass issue polarization, using the same 14 issues, based on respondents' self-reported party affiliation and ideology. We use the post-imputation data normalized onto a 0–1 scale as before, but also with some variables rotated so that higher scores always indicate a right-wing position; the variables were not standardized and no PCA was conducted. For each year, we then calculated the following measures:

Separation between Democrats and Republicans. We found the mean absolute distance between the mean position of self-identified Democrats and Republicans on each issue. Party identities were taken from participants' response to the question 'Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?' (VCF0301); those who identified as Independent but who, when asked the follow-up 'Do you think of yourself as closer to the Republican or Democratic party?', chose one of the parties, were classified as identifying with that party.

Separation between liberals and conservatives. We found the mean absolute distance between the mean position of self-identified liberals and conservatives on each issue. Ideological identities were taken from participants' response to the question 'We hear a lot of talk these days about liberals and conservatives. Here is a 7-point scale on which the political views that people might hold are arranged from extremely liberal to extremely conservative. Where would you place yourself on this scale, or haven't you thought much about this?' (VCF0803), participants who selected any response containing 'liberal' were classified as liberal, and any response containing 'conservative' as conservative.

[3]In 2000 and 2004, four of the questions were only given to a randomly selected half of the sample, so this constraint ensured those questions could be retained. Outside these eight cases, the percentage of *NA* responses ranges from 0.8% to 20.8% with a median of 10.8%. *NA* responses were treated as missing and imputed, as discussed in §2.

Table 1. ANES policy items used.

item	statement
social domain	
abortion (VCF0838) (R)	'by law, abortion should never be permitted' (1). 'The law should permit abortion only in case of rape, incest, or when the woman's life is in danger' (2). 'The law should permit abortion for reasons other than rape, incest, or danger to the woman's life, but only after the need for the abortion has been clearly established' (3). 'By law, a woman should always be able to obtain an abortion as a matter of personal choice' (4)
moral flexibility (VCF0852)	'the world is always changing and we should adjust our view of moral behaviour to those changes' (5-point Likert)
traditionalism (VCF0853) (R)	'this country would have many fewer problems if there were more emphasis on traditional family ties' (5-point Likert)
economic domain	
health insurance (VCF0806)	'there should be a government insurance plan' versus 'medical expenses should be paid by individuals' (7-point scale)
guaranteed job and income (VCF0809)	govt should 'see to it that every person has a job and a good standard of living' versus 'let each person get ahead on his/their own' (7-point scale)
services versus spending (VCF0839) (R)	govt should 'provide fewer services, even in areas such as health and education, in order to reduce spending' versus 'provide many more services even if it means an increase in spending' (7-point scale)
equal opportunities (VCF9013)	'our society should do whatever is necessary to make sure that everyone has an equal opportunity to succeed' (5-point Likert)
equal chances (VCF9016) (R)	'it is not really that big a problem if some people have more of a chance in life than others' (5-point Likert)
equality (VCF9017) (R)	'this country would be better off if we worried less about how equal people are' (5-point Likert)
African-American inequality[a]	
government help (VCF0830)	govt should 'make every effort to improve the social and economic position of [African Americans]' versus 'not make any special effort to help [African Americans] because they should help themselves' (7-point scale)
social mobility (VCF9039)	'generations of slavery and discrimination have created conditions that make it difficult for [African Americans] to work their way out of the lower class' (5-point Likert)
no special favours (VCF9040) (R)	'Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. [African Americans] should do the same without any special favors' (5-point Likert)
lack of effort (VCF9041) (R)	'it's really a matter of some people not trying hard enough; if [African Americans] would only try harder they could be just as well off as whites' (5-point Likert)
less than deserved (VCF9042)	'over the past few years [African Americans] have gotten less than they deserve' (5-point Likert)

(R) indicates items that were reverse-coded so that high scores indicate a position typically considered to be conservative.
[a]Where the term 'African Americans' is written in square brackets, the original ANES wording uses a term (the word 'black' as a pluralized noun) which can be considered offensive [68] and which we have therefore chosen not to reproduce.

3.2. Results

3.2.1. Primary analysis

Our primary analysis concerns how the levels of each facet of polarization have changed from 1988 to 2024. Separation, Dispersion and Equality-of-Size scores for each year are shown in figure 1.

As figure 1 illustrates, Separation increased between 1988 and 2024 by 0.14 scale points, from 0.22 to 0.36 (a 64% increase in percentage terms), while Dispersion and Equality-of-Size effectively remained constant. Separation did not steadily increase, but rather saw one period of sustained increase from

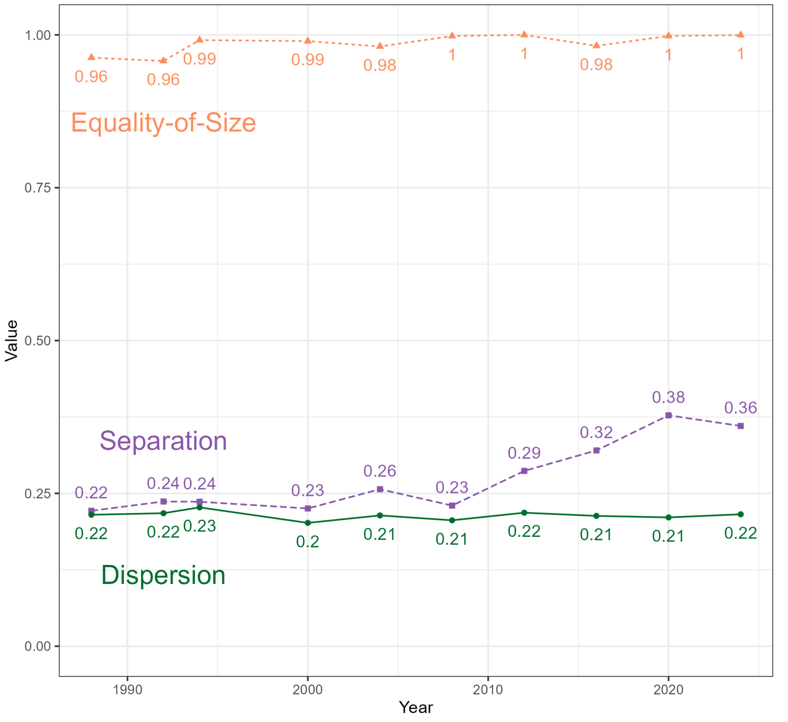

Figure 1. Mean cluster-based Equality-of-Size, Separation and Dispersion over time, Study 1 (ANES).

2008 to 2020, with a slight fall from 2020 to 2024. Therefore, Americans have come to form more distinct clusters in terms of their policy views over the past four decades, but these clusters have not become less cohesive, nor has one come to predominate in the population. This means that, overall, polarization has increased, though effectively only from 2008 to 2020. For a bootstrapped analysis of these trends, which reinforces the conclusions drawn here, see electronic supplementary material, appendix B.

3.2.2. Secondary analyses

We conducted a series of secondary analyses to better understand how polarization manifests in the USA, how that has changed over time, and how our measures of polarization relate to others. To begin, figure 2 shows the separation between self-identified liberals and conservatives, and between Democrats and Republicans, with a similar period of consistently increasing separation from 2008 to 2020.

Figure 3 shows the mean positions of the clusters across all issues in each year, after all responses are linearly transformed to range from 0 to 1, where higher scores indicate a more right-wing position. The points show cluster-level means for each year. The regression lines are fitted not to the mean for each year, but to the individual-level responses found within each cluster in each year, leading to confidence intervals which are too small to see.

In each year, one cluster is more to the right, and one more to the left. The regression lines show that the left-wing cluster moves over time to be situated in more of a left-wing position ($b = -0.0046$, $se = 0.0001$, $p < 0.001$), whereas the right-wing cluster moves to be situated more to the right, though the amount of movement is negligible ($b = 0.0004$, $se = 0.0001$, $p < 0.001$). This pattern suggests mean opinion in the US has become more left-wing, while simultaneously the clusters have spread further apart.

We then sought to determine whether any particular issues were driving the increased separation between clusters. To that end, figure 4 depicts the mean positions of each cluster for each issue. The issues are ordered from left-to-right and top-to-bottom in descending order of how much further the positions of the clusters have become separated over time. This was determined by performing a regression analysis of participants' responses, separately for each issue, which included as predictors an interaction between cluster and year; we then inspected the coefficient of the year:cluster interaction term. This term measures the extent to which the progression of time has amplified the distance between clusters. Notably, this term is statistically significant for every issue (all $p < 0.001$), showing

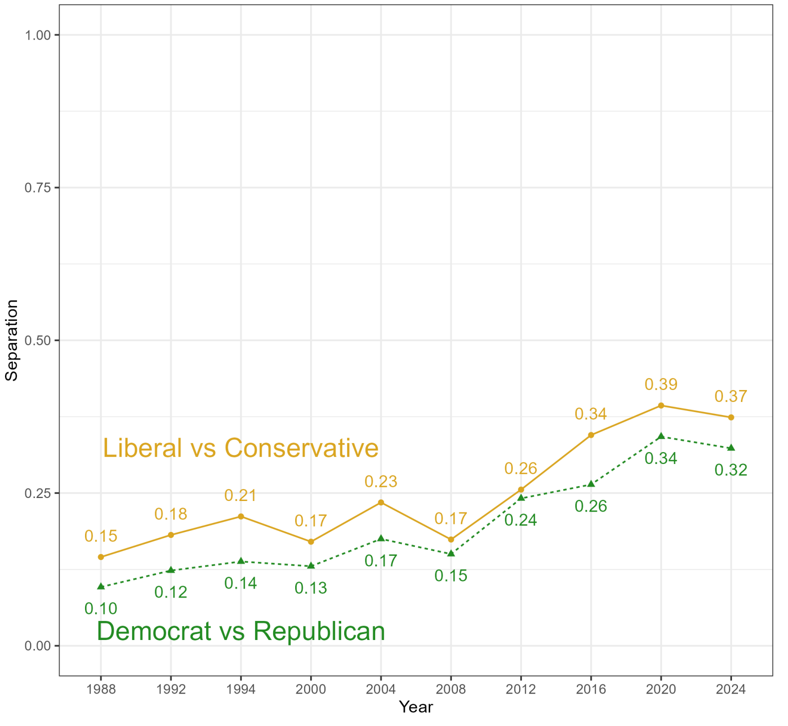

Figure 2. Separation between liberals and conservatives and between Democrats and Republicans over time, Study 1 (ANES).

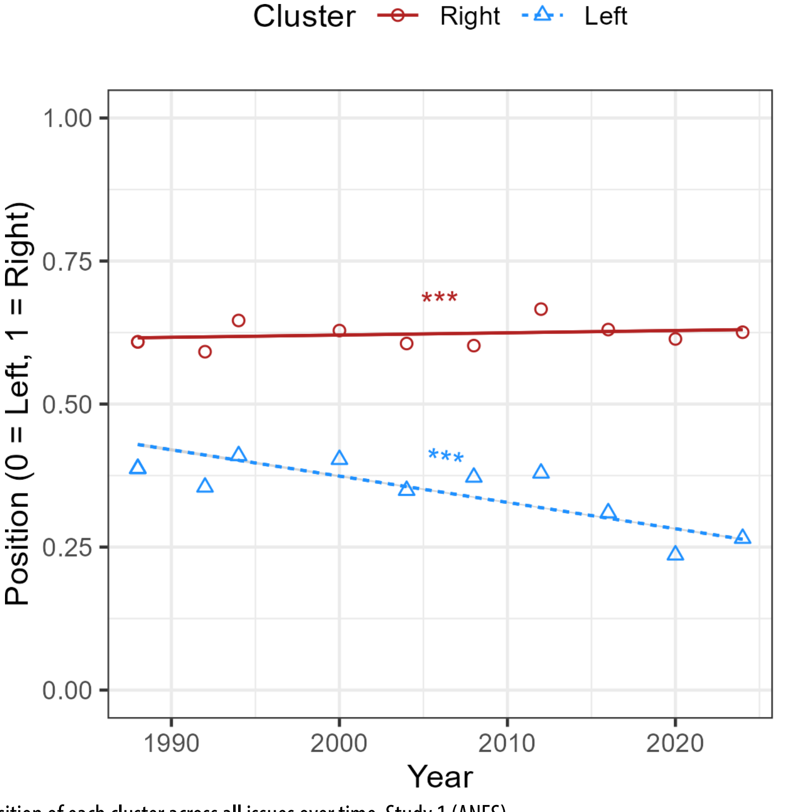

Figure 3. Mean position of each cluster across all issues over time, Study 1 (ANES).

that the clusters have separated over time on every issue. Seven issues show 'true' polarization in the sense of both clusters moving in *opposing* directions over time, whereas for the other seven, both clusters have moved in the same direction, but at different rates, causing them to spread apart—for most of these issues, both clusters have moved to the left, with the left-leaning cluster moving further, but for one issue, 'Equal Opportunities', both clusters have moved right (though both clusters still have a very 'left-wing' position on this issue).

How well do the clusters we have identified match people's self-reported partisan and ideological identities? Since we know that those in the left-wing cluster 'should' be more likely to identify as liberals and Democrats, due to their left-wing views, and those in the right-wing cluster as conservatives and Republicans, analysing the proportion of each cluster who identifies with the expected group gives an insight into levels of sorting in the USA over time. Figure 5 shows that, over time, sorting has increased, as the proportions in each cluster who identify with the expected party and with the expected ideological group have increased.

3.3. Discussion

Our analysis finds that mass issue polarization in the USA has increased—our measure of between-cluster Separation has gone up, driven by a period of sustained increase from 2008 to 2020, while our measures of within-cluster Dispersion and Equality-of-Size have remained effectively constant. Therefore, it seems the US population has become more differentiated into two opinion clusters over the past four decades, without those clusters becoming less cohesive or one becoming more dominant in terms of size.

We find concordant evidence that the separation between Democrats and Republicans, and between self-identified liberals and conservatives has increased too, strengthening the evidence for increased mass issue polarization. While previous work using such self-labels was inconclusive on the question of polarization because of the possibility of a confound through increasing sorting, our measure can disambiguate these processes, demonstrating in fact that both sorting *and* latent attitude separation have occurred. Our analysis also provides a replication and extension of the finding that Americans with left-leaning views are less likely to describe themselves as 'liberal' than as a Democrat [69], although this trend has diminished over time.

We find that the increased polarization is driven by increased separation between clusters on *all* issues studied. Some items have changed over time from being ones where both clusters effectively have the same position to ones where there is substantial disagreement—in particular, attitudes to abortion, but also the 'traditionalism' item. What this effectively means is that whereas in 1988, people with generally right-leaning opinions did not necessarily have a more restrictive opinion on abortion than people with generally left-leaning opinions, nowadays right-leaning views tend to be 'packaged' with restrictive views on abortion. This trend echoes DellaPosta's [17] argument that the USA has witnessed 'pluralistic collapse', with people who disagree on some issues becoming less likely to have any points of agreement on other issues. Our results extend DellaPosta's findings by showing that this trend emerges using an alternative cluster-based methodology.

The time period we study is characterized by a period of effective plateau in between-cluster Separation (1988–2008) followed by a period of consistent increase (2008–2020). This is theoretically interesting, as it implies that whichever factors have driven increased mass issue polarization in the USA, their impact is not consistent over time. Rather, it appears that polarizing forces began increasing from the late 2000s and sustained into the 2010s. To complicate things further, the latest datapoint, for 2024, suggests the period of increasing polarization may have ended (though it will require several more datapoints before we can be confident asserting this), but polarization levels still remain high relative to earlier years. Any explanation of the USA's increase in mass issue polarization over the past four decades would therefore need to explain why polarization increased *at certain times*, and did not at others, in a way that fits this timeline.

Yet, even if we successfully determined which factors have driven the increase in US mass issue polarization, these findings may not explain patterns of polarization in other countries. This is because several systemic factors which might moderate the effect of polarizing factors are constant in the USA—for example, the presence of only two major political parties. In Study 2, we therefore perform a cross-national analysis of the predictors of Separation, Dispersion and Equality-of-Size, as well as trying to map how polarization manifests in different countries. We also explore whether the trend of increasing polarization we found in the USA is seen more broadly across the world, by measuring longitudinal trends for Separation, Dispersion and Equality-of-Size.

4. Study 2

Study 2 takes data from the Integrated Values Survey (IVS), which combines data from the WVS [70] and EVS [71]. We are able to perform the clustering and polarization calculations using data

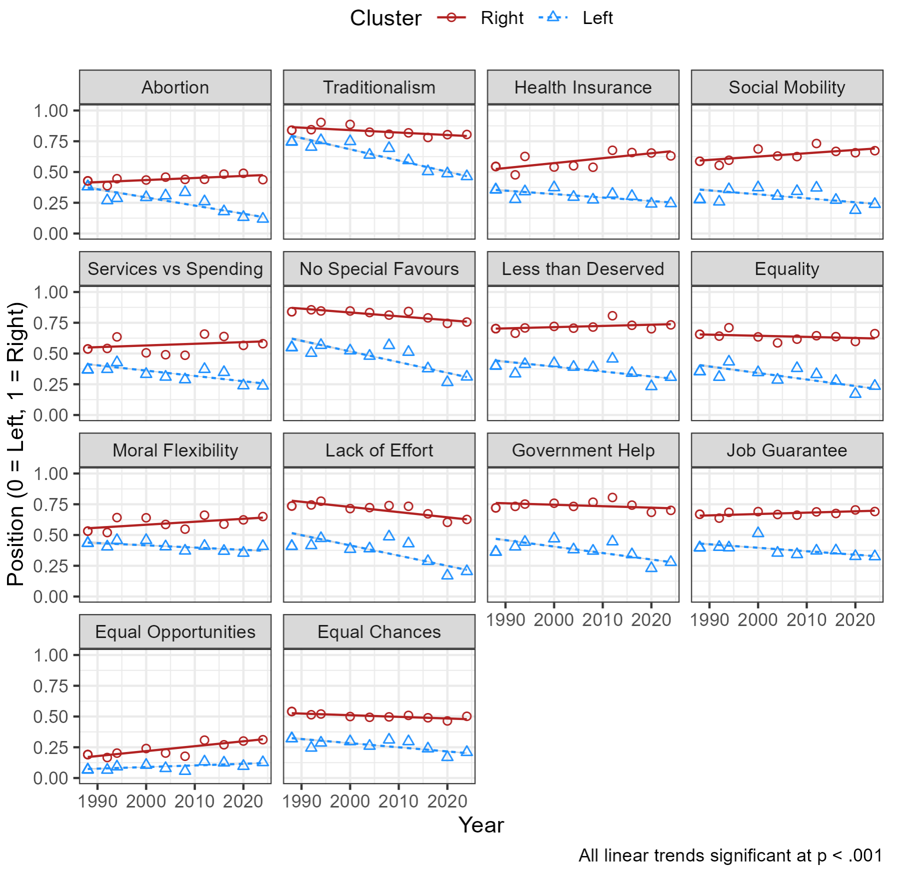

Figure 4. The mean position of each cluster on each issue over time, Study 1 (ANES).

from 364 142 unique individuals across 247 samples from 105 countries, from 1999 to 2022, with each country surveyed between one and five times within the timeframe. For our analysis of the predictors of polarization, however, we use a subset of the data which only includes the countries and years for which data are available for the full set of the 14 predictors of polarization we investigate (detailed below), which comprises 173 513 unique individuals spread across 114 samples from 57 countries surveyed from 1999 to 2018, to ensure comparability across analyses.

To create the issue position space used for the clustering, we use 14 items (selection criteria below) spanning three domains: Social, Economic and Democratic Attitudes (it is only a coincidence that both Study 1 and Study 2 feature 14 items spanning three domains). Variable selection in Study 2 followed similar criteria to those in Study 1, with an initial longlist of variables selected from the IVS codebook according to face validity as political issue beliefs. Care was taken to include issues that might be relevant to politics in non-Western contexts, and we again required items to be measured by a numerical rating or Likert-style scale of four levels or more. We chose from this longlist the 14 variables available in every WVS and EVS wave from 1999 and retained data from all samples which contained at least some data on each variable—this provides a good balance between coverage across different kinds of items and over time, while ensuring consistent geographical coverage throughout the time period (before 1999, some of the items are available only in the WVS). Our items were split across three domains, as table 2 shows. A table of scores for each sample is available in electronic supplementary material, appendix C.

4.1. Initial results

Since the regression analyses require that we use only around half the samples for which we can calculate polarization scores (see above), we first report some results which can utilize the whole dataset to give an impression of how polarization manifests across the world. These results are also useful to bear in mind when interpreting the regression results that follow.

Firstly, Separation scores (M = 0.18, SD = 0.04) and Dispersion scores (M = 0.20, SD = 0.02) are lower and less variable than Equality-of-Size scores (M = 0.88, SD = 0.17). Secondly, there is a trade-off

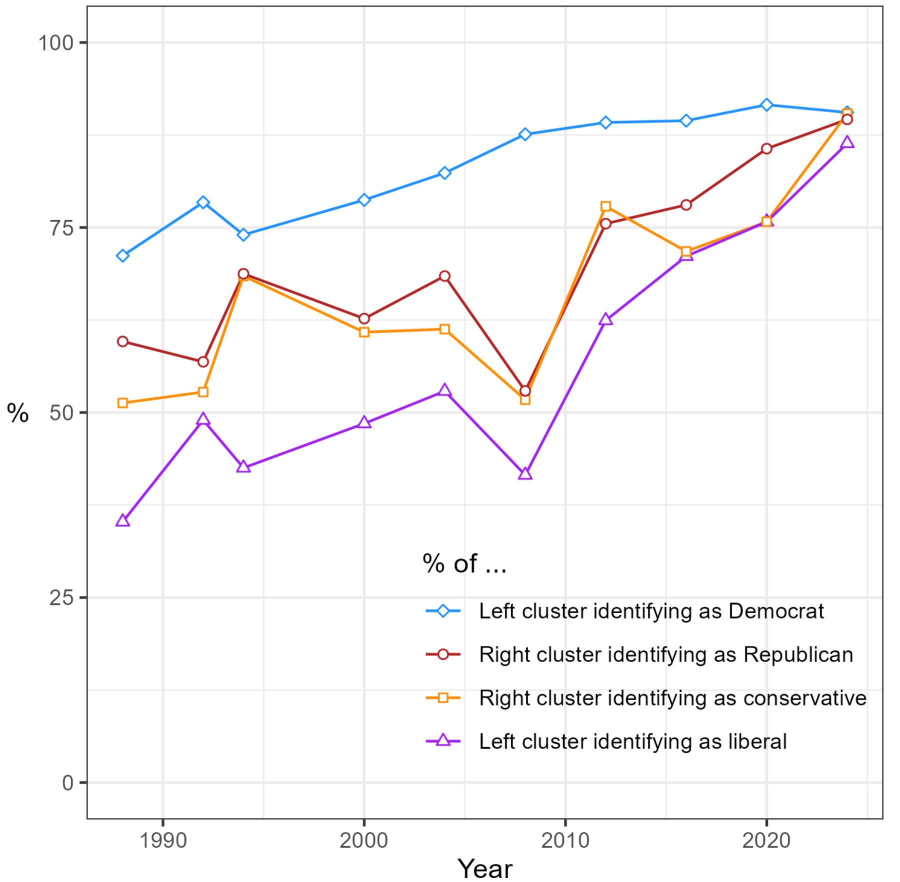

Figure 5. Party and ideological identities of members of each cluster over time, Study 1 (ANES).

between the different facets of polarization: Separation is correlated negatively with Equality-of-Size, $r(245) = -0.31$ [−0.19, −0.42], $p < 0.001$, and positively with Dispersion, $r(245) = 0.31$ [0.20, 0.42], $p < 0.001$, while Equality-of-Size and Dispersion are positively correlated with each other, $r(245) = 0.16$ [0.04, 0.28], $p = 0.011$. This suggests that in countries with well-separated clusters, the clusters tend to simultaneously be less cohesive and more imbalanced in terms of size.

Is polarization increasing over time throughout the world, like we saw in the USA in Study 1? To analyse this, we first reduced our dataset down to include only countries for whom we collect data at least twice (giving us 218 samples from 76 countries). We then ran, for each measure of polarization, a multi-level regression using Year, the year the sample was taken, as a fixed-effects independent variable, and a random intercept for Country. The random intercept lets us avoid potential confounding due to countries who differ in their level of polarization being sampled at different times. We found statistically significant, positive coefficients for Separation ($b = 0.0008$, $se = 0.0003$, $p = 0.003$) and Dispersion ($b = 0.0005$, $se = 0.0001$, $p < 0.001$), but a null effect for Equality-of-Size ($b = -0.0004$, $se = 0.0008$, $p = 0.564$). Giving that Separation increases polarization but Dispersion decreases it, this means there is no clear evidence that polarization is increasing on a global scale. All coefficients are also very small given the 0–1 scale of the measures, but given these indicate the average *annual* change, larger differences might be expected to occur if these trends persist over longer time periods.

Thirdly, it is primarily cultural issues which drive divisions between the clusters. First taking each sample and issue separately, we found Welch's *d*, a measure of effect size which is a variant of Cohen's *d* that does not assume equal variances between groups, of the difference in positions taken by the people in each cluster, as well as the 95% confidence interval. We then averaged Welch's *d* and upper and lower 95% confidence interval bounds across samples for each issue. This gives an impression of which issues cause the most division between clusters. The results are shown in figure 6, clearly showing that it is disagreement over cultural issues, like the acceptability of abortion and homosexuality, that drive the most division between political opinion clusters across the world.

Table 2. Items used in Study 2 (IVS).

item	question
social domain	
bribery justifiable (F117)	'someone accepting a bribe in the course of their duties' can 'always be justified' versus 'never be justified' (10-point scale)
homosexuality justifiable (F118)	'homosexuality' can 'always be justified' versus 'never be justified' (10-point scale)
prostitution justifiable (F119)	'prostitution' can 'always be justified' versus 'never be justified' (10-point scale)
abortion justifiable (F120)	'abortion' can 'always be justified' versus 'never be justified' (10-point scale)
divorce justifiable (F121)	'divorce' can 'always be justified' versus 'never be justified' (10-point scale)
euthanasia justifiable (F122)	'euthanasia (terminating the life of the incurably sick)' can 'always be justified' versus 'never be justified' (10-point scale)
economic domain	
pro-inequality (E035)	'incomes should be made more equal' versus 'there should be greater incentives for individual effort' (10-point scale)
pro-nationalization (E036)	'private ownership of business and industry should be increased' versus 'government ownership of business and industry should be increased' (10-point scale)
pro-paternalism (E037)	'individuals should take more responsibility for providing for themselves' versus 'the state should take more responsibility to ensure that everyone is provided for' (10-point scale)
anti-competition (E039)	'competition is good' versus 'competition is harmful' (10-point scale)
democratic domain	
strong leader bad (E114)	'having a strong leader who does not have to bother with parliament and elections' is a 'very good' versus 'very bad way of governing this country' (4-point scale)
expert rule bad (E115)	'having experts, not government, make decisions according to what they think is best for the country' is a 'very good' versus 'very bad way of governing this country' (4-point scale)
army rule bad (E116)	'having the army rule the country' is a 'very good' versus 'very bad way of governing this country' (4-point scale)
democracy bad (E117)	'having a democratic political system' is a 'very good' versus 'very bad way of governing this country' (4-point scale)

4.2. Predictor analysis

4.2.1. Method

We tested 14 variables which have a theoretical or empirical basis as possible predictors of polarization, classified into six groups, most of which are taken from external data sources (see electronic supplementary material, appendix D, for detailed information on sources). Polarization could stem from *within-country socioeconomic differences* (e.g. [27]), so we include Wealth Inequality (using the Gini coefficient), Ethnic Fractionalization and Urban-Rural Fractionalization. *Cultural values* could affect tolerance of dissenting views and therefore polarization, so we include Secular Cultural Values and Emancipative Cultural Values, which are found within the IVS dataset. Welzel [72] explains that Secular values pertain to a society's rejection of 'quasi-divine sources of authority over people, including the authority of religion, the nation, the state, and conformity norms' (p. 58), whereas Emancipative values pertain to 'how strongly people claim authority over their lives for themselves' (p. 59), which involves embracing an egalitarian worldview and placing 'an emphasis on freedom of choice' (p. 67) in how people live their lives. Similarly, *pluralistic political institutions* could affect polarization by permitting open discussion of competing viewpoints, so we include level of Democracy, Civil Liberties and Freedom of the Press, all coded so that higher scores means the society is more pluralistic. *Elite cues* may drive polarization [73], and so polarized countries may be so just because they have fewer parties, so we analyse the effect of the Effective Number of Electoral Parties. Further, *Economic turmoil* may provoke polarization [26], so we included Unemployment, Inflation and Growth; to counteract extreme positive skew, we perform a signed log transformation of the Inflation scores. We also include two miscellaneous predictors: Human Development Index (HDI), as this correlates

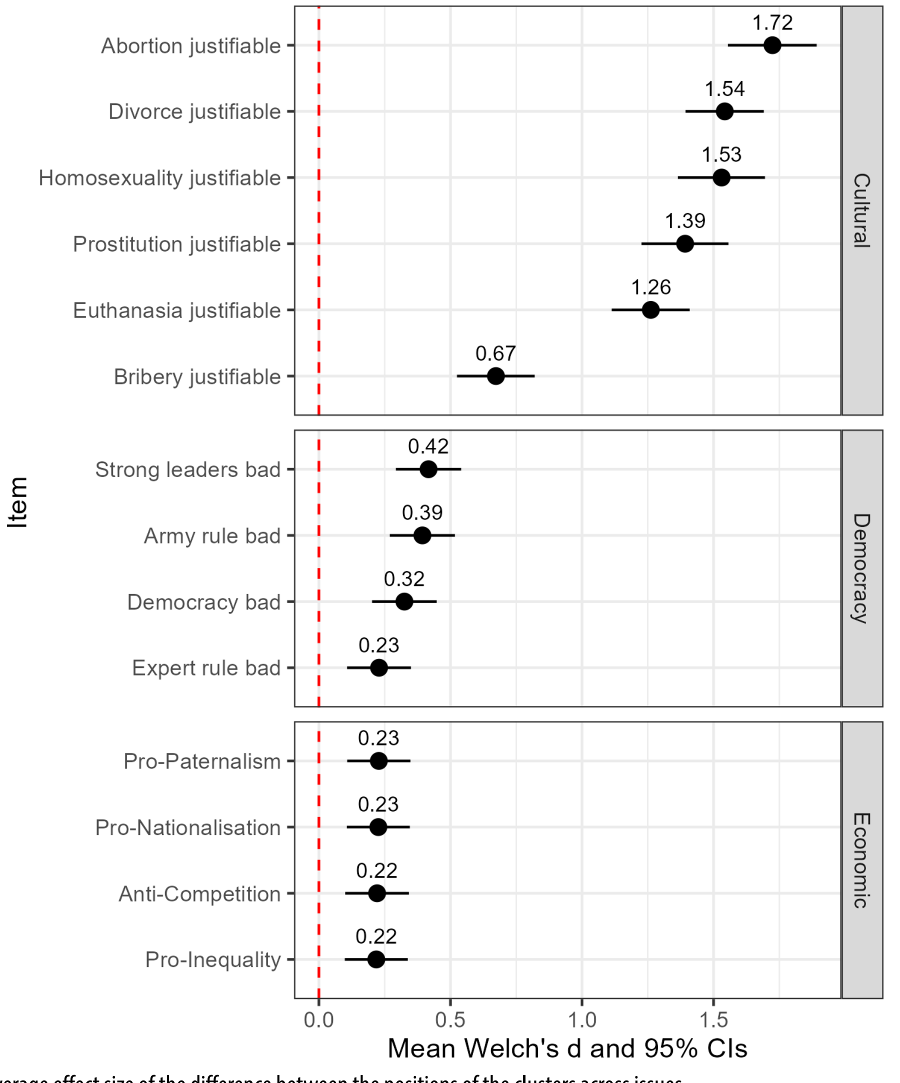

Figure 6. Average effect size of the difference between the positions of the clusters across issues.

with liberal cultural values and pluralistic institutions but is conceptually different, and Internet Users per Capita, as previous authors have contested whether Internet usage is related to various kinds of polarization [13,74]. For context, figure 7 shows the correlation matrix for the predictors, demonstrating that there are strong correlations between several pairs of variables.

Data are not available for every predictor for each country and year, so analysis is restricted to those samples for which data on all predictors are available—see electronic supplementary material, appendix C, for a list of which samples and which are not included in the regression analyses. To increase our coverage, for all the variables except Unemployment, Inflation and Growth, we take the nearest non-missing datapoint from the external data source within 5 years of the IVS sample year; for these three economic variables, we average across all non-missing data from within 5 years before the IVS sample year.

We conducted one univariate OLS regression for each combination of predictor and measure of polarization (Separation, Dispersion and Equality-of-Size), with each predictor standardized using *z*-scoring. With 14 regressions per dependent variable, this, problematically, creates 14 chances to find a significant predictor for each measure of polarization, therefore we have corrected the significance threshold for each predictor using the Holm–Bonferroni method. As we detected heteroskedasticity in most of the regression models, significance tests of the coefficients were performed with robust standard errors. We also planned to iteratively remove any predictor values with Cook's distances above 1 before performing the regressions, but found none.

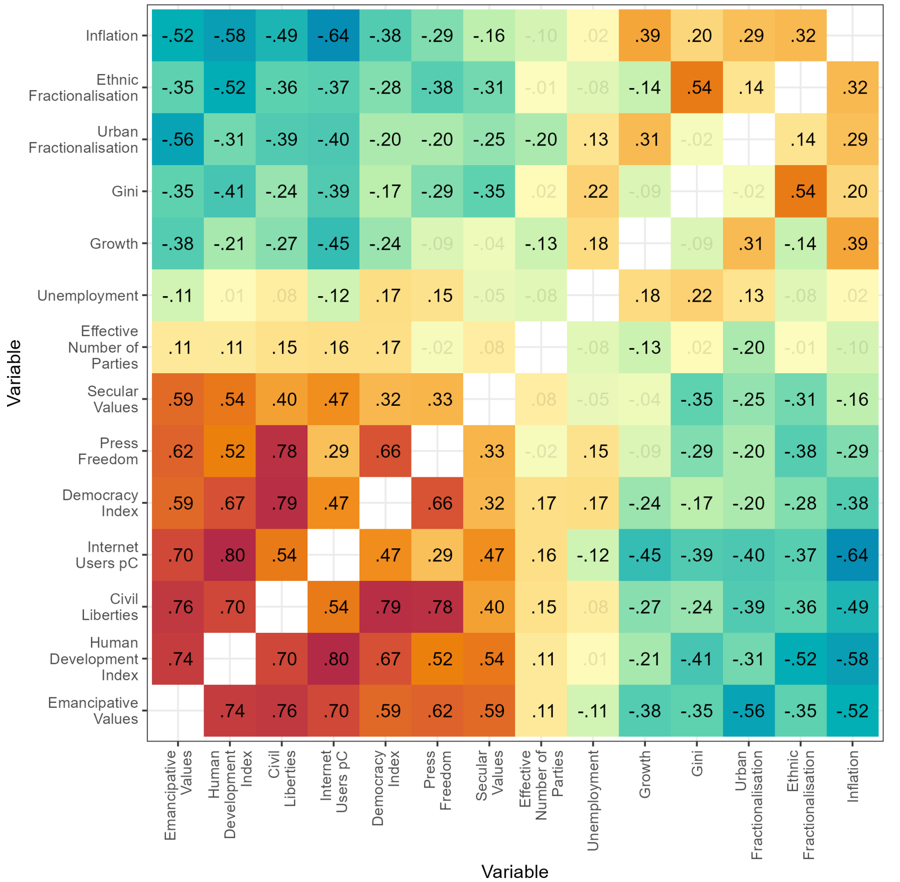

Figure 7. Correlation matrix for predictors used in regression analyses, Study 2 (EVS and WVS). Lightened correlation coefficients are non-significant at $p < 0.05$.

4.2.2. Results

Tables 3–5 show the results for each measure of polarization, with predictors significant at the Bonferroni–Holm-corrected level in bold.

Notably, no predictor is simultaneously associated with higher Separation, higher Equality-of-Size and lower Dispersion. Therefore, no single factor is associated with higher polarization across all three facets. Only one predictor is predictive of all facets in any direction—Ethnic Fractionalization, which is associated with higher Separation, but also lower Equality-of-Size and higher Dispersion. Therefore countries with higher Ethnic Fractionalization seem to have more-separated, but less cohesive and equally sized clusters. See electronic supplementary material, appendix E, for details of a bootstrapped analysis which reinforces these results.

4.2.3. Follow-up analyses

We conducted some follow-up analyses to further explore and verify the robustness of the associations highlighted by our regressions.

Firstly, we looked more closely at the relationship between Ethnic Fractionalization and between-cluster Separation. The basic trend is visualized in figure 8. We first explored whether the effect remained after controlling for HDI, given that Ethnic Fractionalization is negatively correlated with HDI (see figure 7), and HDI approaches significance. Therefore, we ran a multivariate OLS regression which included both variables as additive predictors, again using robust standard errors and standardizing the predictors using *z*-scores. Ethnic Fractionalization remained a significant predictor, $b = 0.015$, $se = 0.004$, $p = 0.0001$, $R^2 = 0.210$.

Further, since Ethnic Fractionalization is higher in certain regions of the world (i.e. Africa and Asia), we also tested whether accounting for a country's region might explain away the effect. Therefore, we ran a multi-level regression which included region as a random intercept, and group-mean-centred Ethnic Fractionalization (i.e. the deviation of a country's Ethnic Fractionalization from its region's

Table 3. Predictors of between-cluster Separation, Study 2 (EVS and WVS).

predictor	β (*se*)	R^2	*p*-value
ethnic fractionalization	**0.016 (0.003**)	**0.210**	**< 0.001 (*a* = 0.00357)**
human development index	−0.009 (0.004)	0.063	0.043 (*a* = 0.00385)
inflation	0.006 (0.004)	0.029	0.107 (*a* = 0.00417)
gini	0.004 (0.003)	0.013	0.244 (*a* = 0.00455)
growth	−0.003 (0.003)	0.010	0.255 (*a* = 0.005)
unemployment	−0.003 (0.004)	0.010	0.409 (*a* = 0.00556)
internet users per capita	−0.003 (0.004)	0.006	0.456 (*a* = 0.00625)
effective number of parties	−0.001 (0.003)	0.002	0.628 (*a* = 0.00714)
urban fractionalization	0.001 (0.003)	0.001	0.708 (*a* = 0.00833)
press freedom	−0.001 (0.004)	0.002	0.713 (*a* = 0.010)
democracy index	−0.002 (0.004)	0.002	0.715 (*a* = 0.0125)
emancipative values	0.001 (0.003)	0.001	0.731 (*a* = 0.01667)
secular values	−0.001 (0.004)	0.000	0.861 (*a* = 0.025)
civil liberties	0.000 (0.004)	0.000	0.938 (*a* = 0.050)

a indicates the Bonferroni–Holm-corrected alpha level.

Table 4. Predictors of within-cluster Dispersion, Study 2 (EVS and WVS).

predictor	β (*se*)	R^2	*p*-value
gini	**0.011 (0.001**)	**0.268**	**< 0.001 (*a* = 0.00357**)
ethnic fractionalization	**0.008 (0.002**)	**0.165**	**< 0.001 (*a* = 0.00385**)
growth	−0.004 (0.002)	0.033	0.028 (*a* = 0.00417)
effective number of parties	0.002 (0.002)	0.014	0.216 (*a* = 0.00455)
press freedom	0.002 (0.003)	0.014	0.363 (*a* = 0.005)
democracy index	0.002 (0.002)	0.009	0.388 (*a* = 0.00556)
emancipative values	−0.001 (0.002)	0.003	0.533 (*a* = 0.00625)
inflation	0.001 (0.002)	0.003	0.559 (*a* = 0.00714)
human development index	−0.001 (0.002)	0.004	0.574 (*a* = 0.00833)
civil liberties	0.001 (0.002)	0.002	0.616 (*a* = 0.010)
internet users per capita	−0.001 (0.002)	0.002	0.626 (*a* = 0.0125)
unemployment	−0.001 (0.002)	0.002	0.683 (*a* = 0.01667)
urban fractionalization	0.000 (0.002)	0.000	0.931 (*a* = 0.025)
secular values	0.000 (0.002)	0.000	0.949 (*a* = 0.050)

a indicates the Bonferroni–Holm-corrected alpha level.

mean) as a fixed-effects predictor. We used three different systems for categorizing countries from the World Bank's Development Indicators, as implemented in the *countrycode* R package [75]. The effect of Ethnic Fractionalization was statistically significant when using all three systems: five continents, b = 0.079 (se = 0.015), $p < 0.001$; seven regions, b = 0.0059 (se = 0.016), $p < 0.001$; and 23 regions, b = 0.061 (se = 0.018), p = 0.001.

We were also interested in better understanding the nature of the association between Equality-of-Size and its predictors. Since our initial regression analysed identified 10 significant predictors (see table 5), but many of these correlate strongly with each other (see figure 7), we conducted a follow-up OLS regression which included all the significant predictors as standardized, additive predictors, using robust standard errors. Only two variables maintained statistical significance and the direction of their effect: HDI (b = 0.135, se = 0.024, $p < 0.001$) and Secular values (b = 0.077, se = 0.014, $p < 0.001$). Internet

Table 5. Predictors of cluster size Equality-of-Size, Study 2 (EVS and WVS).

predictor	β (*se*)	R^2	*p*-value
human development index	**0.107 (0.016)**	**0.415**	**< 0.001 ($a = 0.00357$)**
secular values	**0.103 (0.017)**	**0.388**	**< 0.001 ($a = 0.00385$)**
emancipative values	**0.084 (0.016)**	**0.258**	**< 0.001 ($a = 0.00417$)**
civil liberties	**0.081 (0.018)**	**0.239**	**< 0.001 ($a = 0.00455$)**
internet users per capita	**0.070 (0.016)**	**0.180**	**< 0.001 ($a = 0.005$)**
urban fractionalization	**−0.046 (0.013)**	**0.078**	**< 0.001 ($a = 0.00556$)**
inflation	**−0.056 (0.016)**	**0.115**	**< 0.001 ($a = 0.00625$)**
press freedom	**0.055 (0.016)**	**0.112**	**0.001 ($a = 0.00714$)**
democracy index	**0.061 (0.020)**	**0.136**	**0.002 ($a = 0.00833$)**
ethnic fractionalization	**−0.053 (0.019)**	**0.101**	**0.007 ($a = 0.010$)**
unemployment	0.019 (0.015)	0.013	0.214 ($a = 0.0125$)
effective number of parties	0.014 (0.015)	0.007	0.337 ($a = 0.01667$)
gini	−0.003 (0.011)	0.000	0.782 ($a = 0.025$)
growth	−0.004 (0.013)	0.000	0.789 ($a = 0.050$)

a indicates the Bonferroni–Holm-corrected alpha level.

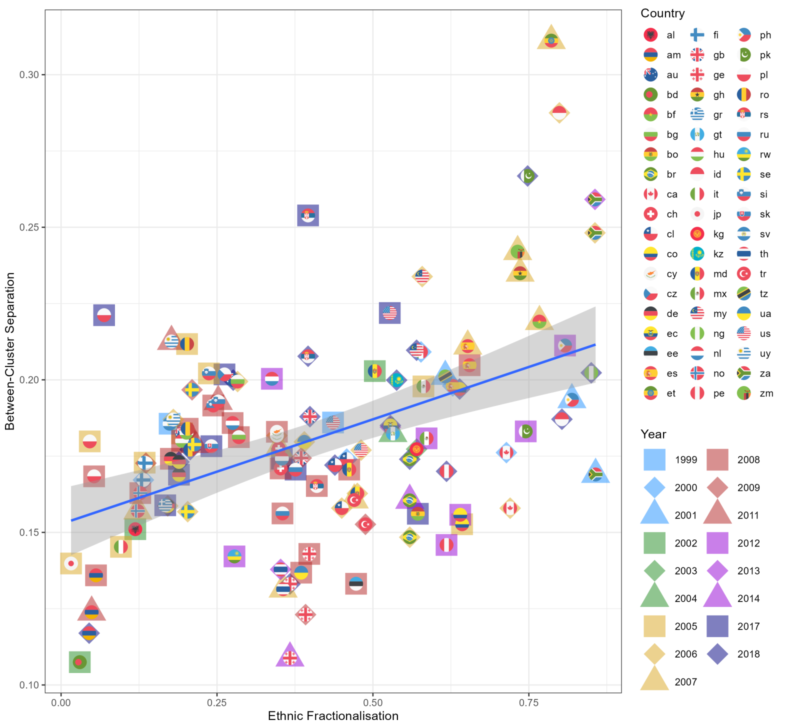

Figure 8. The relationship between Ethnic Fractionalization and between-cluster Separation, Study 2 (EVS and WVS).

Users per Capita maintained statistical significance but flipped the direction of its effect ($b = -0.073$, se = 0.021, $p = 0.001$).[4] All other effects were non-significant ($p > 0.093$).[5]

[4]The effect of Internet Users per Capita flips to negative as soon as HDI is controlled for, with its effect becoming $b = -0.044$ (se = 0.020), $p = 0.027$ when HDI is used as the only additional predictor.

[5]Inflation has a statistically significant negative effect when we do not use PCA to reduce the dimensionality of the issue spaces before clustering, $b = -0.029$, $se = 0.015$, $p = 0.049$. We do not consider this pattern of results to be a reliable indicator of an effect. All other effects show the same pattern of significance across different approaches to reduction as noted in §2.

To better understand the effect of HDI on Equality-of-Size in particular, we present a visualization in figure 9. The figure depicts the 'Cultural Liberalism' of each cluster, which is calculated by averaging their members' mean positions on the justifiability of abortion, divorce, homosexuality, prostitution and euthanasia, where higher scores indicate a greater perception of justifiability. The difference between clusters on these scores explains 48.9% of the variance in between-cluster Separation scores, which is unsurprising given the strong contribution of cultural issues to between-cluster division shown in figure 6. We plot the 'Cultural Liberalism' of the majority cluster (containing >50% of the sample) and the minority cluster (containing <50%) against the HDI of the nation to which they belong, and also visualize the proportion of the sample within each cluster.

As figure 9 shows, countries with lower HDI scores typically have large culturally conservative majority clusters and small culturally liberal minority clusters, whereas countries with higher HDI scores have more equally sized and more liberal clusters, and in the most highly developed countries (i.e. above 0.850), it is normally the socially liberal cluster that is in the majority.

4.3. Discussion

Our analysis of how mass issue polarization manifests in different countries across the world, and its predictors, returns several novel insights. Firstly, we find that in countries with lower Human Development, the political landscape often takes the form of disagreement between culturally conservative majorities and culturally liberal minorities, but in countries with higher Human Development, we tend to see more equally sized and culturally liberal clusters, with culturally liberal clusters usually in the majority in the most highly developed countries. Secondly, we find that higher between-cluster Separation tends to be traded off with higher within-cluster Dispersion and lower Equality-of-Size, which means that in countries where clusters are highly separated, the level of polarization is often counterbalanced by internal disagreements within the clusters and one cluster tending to dominate, rather than society being equally split between clusters. Thirdly, and unsurprisingly given the preceding insight, no single factor is uniformly predictive of higher polarization as we conceptualize it, with no predictor simultaneously associated with higher Separation, lower Dispersion and higher Equality-of-Size.

Instead we find that different predictors are associated with each of our three measures. We find a clear association between increased Ethnic Fractionalization and increased between-cluster Separation. Ethnic Fractionalization measures the probability of two randomly selected individuals from a particular country at a particular time belonging to different ethnic groups [76]. Since politically salient norms around things like fairness can differ significantly across socioeconomic groups within a country [77], Ethnic Fractionalization could lead to lower coordination around shared norms, creating more-separated opinion clusters. We note, however, that Ethnic Fractionalization is only one of numerous ways to measure a country's ethnic diversity [78]. Ethnic Fractionalization has the advantage of capturing many potential within-country divides, but replicating this analysis with other operationalizations of ethnic diversity would be useful. This finding reinforces the argument that when conducting comparative research we should explore not just differences between countries but within countries too [79].

The association we find between Secular values and higher HDI fits well with the observation that richer countries tend to be more liberal and democratic [80–83], though there is disagreement about the causal direction here [82]. Our findings suggest that liberalization does not occur uniformly, but may happen because liberal minorities gradually increase in size in concert with greater development (our analysis cannot speak to the causal direction). Numerous previous studies of social change suggest this emergence of liberal attitudes within a society can provoke backlash from those who maintain conservative positions [84,85].

With Equality-of-Size, we find that HDI and Secular values are predictive of more-equally sized clusters. Both effects reflect the pattern visible in figure 9, whereby countries with the lowest HDI tend to have culturally conservative majority clusters that are much larger than their culturally liberal minority counterparts. Both clusters become more liberal and more equally sized as HDI increases. The fact that the clusters simultaneously become more liberal and equally sized with increasing HDI may explain why Secular values are predictive of increased Equality-of-Size, as Secular values and Cultural Liberalism measure similar latent traits—indeed we find that a country's mean level of Cultural Liberalism correlates strongly with the IVS estimate of their Secularism, $r(112) = 0.65$ [0.54,

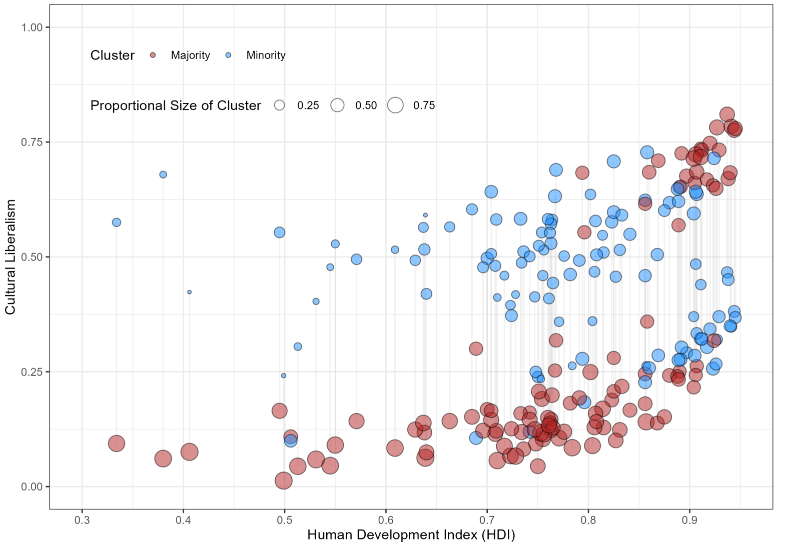

Figure 9. The proportional size and cultural liberalism of clusters by HDI. The pair of clusters for one sample are joined by a grey line.

0.75], $p < 0.001$. Thus countries with more equally sized, liberal clusters will also be more liberal and Secular overall.

Our analysis of the predictors of Equality-of-Size also found that Internet Users per Capita changed from having a positive effect in its univariate regression to having a negative effect when HDI was controlled for. This flip makes us cautious to interpret the effect and concerned about the potential for it to be a statistical artefact. Since Internet Users per Capita is very strongly correlated with HDI ($r = 0.80$; see figure 7), its positive univariate effect may have been indexing for HDI. The negative multivariate effect could therefore indicate a genuine effect whereby greater Internet connectivity makes clusters less equally sized than would be expected based on their HDI. But it could also be an artefact, especially given that when two correlated predictors are included in a regression, estimates of their coefficients can be inaccurate [86]. Overall, our results do not show any evidence of the Internet increasing mass political issue polarization, but further study which is more focused on this question is necessary before definitive conclusions can be drawn.

Finally, we found that higher levels of wealth inequality were associated with higher levels of within-cluster Dispersion (wealth inequality was the most robust predictor of Dispersion—see additional analyses in electronic supplementary material, appendix F). One reason for this could be that higher levels of wealth inequality drive greater disagreement within a society over economic and governmental issues, perhaps between those who do and do not benefit from societal systems which produce and sustain wealth inequality. Indeed, Gu & Wang [27], also using WVS data, find that higher levels of wealth inequality (using the Gini coefficient) are associated with higher population-level variance in attitudes towards economic issues, but not cultural issues. Since our clusters are primarily divided along the cultural axis, this could cause people who agree on cultural issues, and therefore find themselves in the same cluster, to disagree with each other more on economic and governmental issues, increasing within-cluster Dispersion. In fact, follow-up analyses (see electronic supplementary material, appendix F) suggest that the issues for which we see greater within-cluster dispersion at higher levels of wealth inequality are primarily the economic and governmental issues. While Gu & Wang [27] regard the increased variance in economic issues at higher levels of wealth inequality to show that wealth inequality is associated with greater polarization, our analysis suggests that increased variance in economic issue positions actually *depolarizes* societies by introducing greater disagreement within opinion clusters.

5. General discussion

The novel methodology we have introduced in this article for measuring mass issue polarization using *k*-means clustering returns several important and novel insights. In Study 1, analysing polarization levels in the USA from 1988 to 2024, we found clear evidence that polarization has increased, mostly due to a period of continually rising polarization from 2008 to 2020. We see that between-cluster

Separation has increased, while within-cluster Dispersion and Equality-of-Size have remained virtually unchanged, meaning that America's left-leaning cluster and right-leaning cluster have drifted further apart while both remaining internally cohesive and equal in size. Notably, the position of the left-leaning cluster has shifted further to the left since 1988 than the right-leaning cluster has shifted to the right, consistent with US opinion in general moving to the left while becoming more polarized. Increasing between-cluster Separation was driven by all 14 issues in our dataset, but issues which have become substantially more polarized include abortion, the emphasis that should be placed on 'traditional family ties', and government funding of healthcare. We also see clear evidence for 'sorting', with the alignment between a person's cluster and their party identity and ideological identity increasing over time.

Study 2 provides insights into what kinds of factors might explain how polarization could increase in a society over time, like we saw in Study 1. But rather than focus on the specific case of the USA, we perform a cross-national analysis so as to attain more robust and generalizable results. Performing a cross-national study also allowed us to test whether the trend of increasing polarization we saw in the USA is present on a global scale. Therefore in Study 2, we explored the patterns and predictors of polarization across the world. We find that cultural issues are the primary drivers of division between political opinion clusters. Moreover, the HDI affects the pattern of polarization, with lower HDI countries typically seeing disagreement between culturally conservative majorities and culturally liberal minorities, while higher HDI countries have more liberal clusters which are more equal in size, and the countries with the highest HDI scores usually see culturally liberal clusters in the majority. We find that higher levels of between-cluster Separation tend to be associated with higher levels of within-cluster Dispersion and lower Equality-of-Size, suggesting that there are trade-offs between the different facets of polarization. Accordingly, we find that no single predictor is simultaneously associated with higher Separation, lower Dispersion, and higher Equality-of-Size. We find that Ethnic Fractionalization is predictive of higher between-cluster Separation, wealth inequality (measured by the Gini coefficient) is predictive of higher within-cluster Dispersion, and both the HDI and Secular values are predictive of greater Equality-of-Size.

There are a few notable differences between the results of Study 1 and Study 2. First, the pattern of polarization we see in the USA is not replicated globally—the US clusters are evenly sized, but in other countries, particularly those lower in HDI, unequally sized clusters are more common. Similarly, whereas on a global scale division between clusters is centred on cultural issues (Study 2), division between the US clusters is more broad-based—cultural issues remain important, with abortion being the most divisive issue, but economic issues are also important, with topics like health insurance and social mobility among the most divisive. Lastly, Study 2 did not find that polarization was increasing over time globally, as while between-cluster Separation was increasing, within-cluster Dispersion was too, but in contrast, Study 1 found US polarization was increasing over time, as Separation increased while Equality-of-Size and Dispersion remained constant.

One of the clearest contributions of this analysis is to provide evidence that the USA has indeed become more polarized over salient political issues over the last few decades. Since our measure of between-cluster Separation avoids the potential confound that 'sorting' can cause when people's party self-identifications are used (and indeed, we find clear evidence of increased sorting), this provides convincing evidence that US mass issue polarization has increased. We note that our analysis only considers ANES data, meaning that if there is some systematic sampling error which means ANES data are not representative of US opinion as a whole, this pattern of increased polarization might not be generalizable to the US population. But given that the ANES utilizes random-probability multi-modal sampling [87], it seems unlikely a large systematic error could befall multiple waves of data.

The temporal specificity of the trend, with polarization increasing from 2008 to 2020 but not at other times, adds important constraints on any explanatory theory of *why* US mass issue polarization has increased. In particular, this seems to endanger explanatory theories of polarization which point towards any supposedly universal feature of human cognition, like a tendency towards 'tribalism' in the way we reason about political issues (i.e. adopting certain beliefs simply to signal to a particular political group that we are 'one of them'; see [10]), as a polarizing universal feature of cognition would be most likely to create steady, uniform increases in polarization. We expect that successful explanations of polarization will have to emphasize interactions between psychological factors and time-varying 'environmental' features, possibly including factors like material conditions, communication networks and elite rhetoric. Also of particular value is the evidence from Study 2 that polarization manifests differently in countries with different levels of Human Development. This suggests that developing explanatory theories of mass issue polarization with global applicability will be difficult, as they will need to explain the emergence of different patterns in different places. Another valuable finding—that it is *cultural* issues, rather than governmental or economic—which primarily drive division between clusters, illuminates further what explanatory theories of polarization will need to explain.

We can highlight numerous limitations of our analyses. Firstly, regarding the regression analyses used in Study 2, there are significant limitations that need to be borne in mind when interpreting the results. For one, our dataset does not sample countries at a sufficient number of timepoints to perform lagged analyses which might allow us to test causal hypotheses; our analyses conflate between-country and within-country variation, and while it is tempting to reach for causal explanations, ultimately we can only find correlational evidence in this study. The associations we do identify could ultimately be explained away by unmeasured back-door variables, or even reverse causation. There is also imprecision in the contemporaneity of our independent and dependent variables, as when a datapoint for a particular independent variable was not available for the specific year that the country sampled, we allowed data from within 5 years of that year to be used. We furthermore note that our results are specific to the countries and timepoints studied, and we cannot be sure they generalize to other countries or timepoints—though we would also point out that practically speaking, our analysis perhaps utilizes the maximum number of issues, countries and timepoints that is currently feasible for cross-national analyses of mass issue polarization. This reinforces the need to continue efforts to collect cross-national political opinion data, and indeed, to increase the frequency and geographical coverage of these efforts, and to ensure a wide range of politically salient issues are surveyed.

There is also a need to be careful in interpreting the clusters we find. Although we refer to the clusters as left-leaning or right-leaning (Study 1), and culturally liberal or culturally conservative (Study 2), it is important to remember that these clusters are not directly analogous to what we typically think of as, for example, left-leaning groups in society, because k-means clustering forces *everyone* in the sample into one group or another. More usually, we would think of left-leaning or right-leaning groups as being just two of *many* groups in society, with many people, perhaps even most, not belonging to either one. Therefore, thinking in terms of clusters requires a qualitatively different way of thinking about political groups. More generally, efforts to characterize our Study 2 clusters in greater detail would be valuable in future research—particularly, how they relate to other indicators of political positioning like party ID or ideological self-placements.

Some might suggest that polarization can manifest between more than two clusters, and that, for example, if a society is divided into three groups, like contemporary France, where the National Assembly is split almost equally between parliamentary blocs from the left, centre and right, this also represents a kind of polarization. Exploring approaches with $k > 2$ clusters is an avenue for future research, but we would contend that division between three or more groups represents something qualitatively different from polarization—fractionalization, perhaps—and note that some studies have noted this difference, for example, by referring to 'tripolarization' when studying affective polarization between three parties [88].

It is perhaps also a limitation that rather than having one number to measure the level of polarization, we have to consider three. Because in the USA Dispersion and Equality-of-Size are effectively flat during the studied time period, this concern is not so pertinent for Study 1. But for Study 2 it might be helpful to have one combined 'aggregate' measure. However, calculating such a measure presents several critical problems. For one, if we simply averaged a sample's Dispersion, Separation and Equality-of-Size, we would implicitly be assuming that, for example, a 0.10 increase in Dispersion decreases the overall level of polarization by exactly the same amount that a 0.10 increase in Separation increases it. But there is no real justification for making such an assumption. Furthermore, as we have no objective yardstick for determining how polarized any sample 'really' is, there also seems no principled way of figuring out how the different facets should be weighted, i.e. determining how much of an increase in Separation would cancel out a 0.10 increase in Dispersion. The fact that each facet of polarization has distinct correlates perhaps suggests that it is more appropriate to consider polarization as a multi-faceted phenomenon anyway, and that while there might be a pull to define the essence of polarization, it may be best to avoid this instinct [89]. Rather, it may be more appropriate to develop more complicated explanatory models that embrace distinct manifestations of polarization.

A final limitation regarding our method concerns generalizing to out-of-dataset policy issues. The effects we observe are effectively tied to the issues we are able to analyse, and there is no guarantee that they would equally apply to other issues. However, given the quite high number of issues we are able to use (14 in both datasets), the fact that each study's issues span multiple different domains of political opinion, and that our clustering is based on principal components of the issue spaces which therefore aim to capture underlying regularities in political opinions, there is a good chance that variance in many salient out-of-dataset issues is already being captured, since they will correlate with measured attitudes. For example, immigration is a major driver of political disagreement in the UK, but it is not an issue included in Study 2; yet, immigration attitudes likely share variance with the cultural items we do include, because both are correlated with right-wing authoritarianism [90]. However, it remains the case that polarization in some of our countries may revolve around issues which have not been studied here.

We think clustering approaches provide an intuitive way of capturing issue polarization, but recognize our approach is only a starting point. *K*-means is an ideal starting point, as it is a relatively simple and well-known clustering method. Fuzzy clustering algorithms, which probabilistically assign participants to clusters rather than 'crisply' categorizing them in one or the other, could be explored as an alternative in future research, though there would be a need to justify *a priori* a method for setting the 'fuzziness' parameter, which controls how crisply the boundaries of the clusters are delineated, and our measures would need to be adapted to integrate the probabilistic cluster assignments. The HBDSCAN clustering algorithm, which allows individuals who do not fit well in any cluster to remain unassigned, could also be worth exploring. One limitation of *k*-means is that there is no standard approach for integrating case weights, which will introduce sampling error if the characteristics of our samples are not representative of the national populations from which they are drawn. Although the ANES, EVS and WVS use quota-sampling to construct their samples, and therefore the samples are likely to be adequately representative, it would be useful for future research to explore clustering methods which can integrate case weights to improve upon this. There is also scope for alternative approaches regarding the assumptions made about our issue space—we treat the distance between the end-points of the response scale for each question as equivalent, but some may think that for certain questions, the end-points represent a more extreme range of opinions than others. Similarly, we treat Likert scale responses as continuous, with equally spaced intervals, and we do not consider the items to have differing 'difficulties', i.e. for some to be 'harder' for the average respondent to agree with than others because the mid-point represents a more extreme opinion. All of these assumptions are open to challenge, and the ideal analysis would probably involve items which had been carefully developed to have psychometric properties that avoid these limitations. But working with the data that are available, we think these assumptions are reasonable, parsimonious and common to many other prominent analyses.

To conclude, we have developed a new technique for measuring mass issue polarization using *k*-means clustering, which allows us to measure three different facets of polarization. Among many other findings, our results provide convincing evidence that US mass issue polarization increased in the period 1988–2024, and that different facets of polarization are associated with different socioeconomic predictors. We also illustrate how *k*-means can be used to understand the varied ways in which mass issue polarization manifests across the world. Ultimately, we highlight that the study of mass issue polarization needs to distinctly model and define the many meanings of the term 'polarization'.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. We analyse data from several publicly accessible third-party datasets. We therefore cannot share any data ourselves, but we include links and citations to all the datasets we use, and share all our code in an OSF project: https://osf.io/kzd23/?view_only=423559a7c7274c269817978beaa05f18.

Electronic supplementary material is available online [91].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. D.J.Y.: conceptualization, data curation, formal analysis, investigation, methodology, project administration, supervision, validation, visualization, writing—original draft, writing—review and editing; J.A.: conceptualization, data curation, formal analysis, investigation, methodology, validation, writing—review and editing; A.K.: conceptualization, formal analysis, investigation, methodology, validation, writing—review and editing; J.K.M.: conceptualization, funding acquisition, supervision, writing—review and editing; L.J.G.: formal analysis, investigation, methodology, visualization, writing—review and editing; L.d.-W.: conceptualization, funding acquisition, project administration, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. The completion of this work was assisted by a grant from the Templeton World Charity Foundation awarded to J.K.M. and L.d.-W. (TWCF number: 31453, https://doi.org/10.54224/31453). A.K. acknowledges support from a Leverhulme Trust International Professorship grant to the University of Cambridge. D.J.Y., J.A. and L.J.G. acknowledge support from ESRC DTP Studentships.

Acknowledgements. We thank Royce Carroll, Sander van der Linden, Mark Brandt, Adam Harris and Max Falkenberg for helpful comments on earlier versions of this work. We also thank three anonymous reviewers for their helpful and constructive comments.

References

1. Druckman JN, Levendusky MS. 2019 What do we measure when we measure affective polarization? *Public Opin. Q.* **83**, 114–122. (doi:10.1093/poq/nfz003)
2. Iyengar S, Lelkes Y, Levendusky M, Malhotra N, Westwood SJ. 2019 The origins and consequences of affective polarization in the United States. *Annu. Rev. Polit. Sci.* **22**, 129–146. (doi:10.1146/annurev-polisci-051117-073034)
3. Lelkes Y. 2016 Mass polarization: manifestations and measurements. *Public Opin. Q.* **80**, 392–410. (doi:10.1093/poq/nfw005)
4. Mason L. 2018 Ideologues without issues: the polarizing consequences of ideological identities. *Public Opin. Q.* **82**, 866–887. (doi:10.1093/poq/nfy005)
5. Orr LV, Fowler A, Huber GA. 2023 Is affective polarization driven by identity, loyalty, or substance? *Am. J. Pol. Sci.* **67**, 948–962. (doi:10.1111/ajps.12796)
6. Brown GDA, Lewandowsky S, Huang Z. 2022 Social sampling and expressed attitudes: authenticity preference and social extremeness aversion lead to social norm effects and polarization. *Psychol. Rev.* **129**, 18–48. (doi:10.1037/rev0000342)
7. Jost JT, Baldassarri DS, Druckman JN. 2022 Cognitive–motivational mechanisms of political polarization in social-communicative contexts. *Nat. Rev. Psychol.* **1**, 560–576. (doi:10.1038/s44159-022-00093-5)
8. Rollwage M, Zmigrod L, de-Wit L, Dolan RJ, Fleming SM. 2019 What underlies political polarization? A manifesto for computational political psychology. *Trends Cogn. Sci.* **23**, 820–822. (doi:10.1016/j.tics.2019.07.006)
9. Van Bavel JJ, Rathje S, Harris E, Robertson C, Sternisko A. 2021 How social media shapes polarization. *Trends Cogn. Sci.* **25**, 913–916. (doi:10.1016/j.tics.2021.07.013)
10. Funkhouser E. 2020 A tribal mind: beliefs that signal group identity or commitment. *Mind Lang.* **37**, 444–464. (doi:10.1111/mila.12326)
11. Levy N. 2021 *Bad beliefs: why they happen to good people*, 1st edn. Oxford, UK: Oxford University Press. (doi:10.1093/oso/9780192895325.001.0001)
12. Talisse RB. 2019 *Overdoing democracy*. Oxford, UK: Oxford University Press.
13. Boxell L, Gentzkow M, Shapiro JM. 2017 Greater internet use is not associated with faster growth in political polarization among US demographic groups. *Proc. Natl Acad. Sci. USA* **114**, 10612–10617. (doi:10.1073/pnas.1706588114)
14. Boxell L, Gentzkow M, Shapiro JM. 2022 Cross-country trends in affective polarization. *Rev. Econ. Stat.* **106**, 1–60. (doi:10.1162/rest_a_01160)
15. Baldassarri D, Bearman P. 2007 Dynamics of political polarization. *Am. Sociol. Rev.* **72**, 784–811. (doi:10.1177/000312240707200507)
16. Baldassarri D, Gelman A. 2008 Partisans without constraint: political polarization and trends in American public opinion. *Am. J. Sociol.* **114**, 408–446. (doi:10.1086/590649)
17. DellaPosta D. 2020 Pluralistic collapse: the 'oil spill' model of mass opinion polarization. *Am. Sociol. Rev.* **85**, 507–536. (doi:10.1177/0003122420922989)
18. Keijzer M, Mäs M, Flache A. 2024 Polarization on social media: micro-level evidence and macro-level implications. *J. Artif. Soc. Soc. Simul.* **27**, 7. (doi:10.18564/jasss.5298)
19. Luca D, Terrero-Davila J, Stein J, Lee N. 2023 Progressive cities: urban–rural polarisation of social values and economic development around the world. *Urban Stud.* **60**, 2329–2350. (doi:10.1177/00420980221148388)
20. Falkenberg M *et al*. 2022 Growing polarization around climate change on social media. *Nat. Clim. Chang.* **12**, 1114–1121. (doi:10.1038/s41558-022-01527-x)
21. Axelrod R, Daymude JJ, Forrest S. 2021 Preventing extreme polarization of political attitudes. *Proc. Natl Acad. Sci. USA* **118**, e2102139118. (doi:10.1073/pnas.2102139118)
22. World Economic Forum. 2024 *The Global Risks Report 2024 (19th Edition): insight report*. World Economic Forum. See https://www.weforum.org/publications/ global-risks-report-2024/.
23. Klein E. 2020 *Why we're polarized*. Chicago, IL: University of Chicago Press.
24. Fiorina MP. 2016 Has the American public polarized?. In *Unstable majorities: polarization, party sorting, and political stalemate*, pp. 33–54. Stanford, CA: Hoover Institution Press.
25. Layman GC, Carsey TM, Horowitz JM. 2006 Party polarization in American politics: characteristics, causes, and consequences. *Annu. Rev. Polit. Sci.* **9**, 83–110. (doi:10.1146/annurev.polisci.9.070204.105138)
26. Funke M, Schularick M, Trebesch C. 2016 Going to extremes: politics after financial crises, 1870–2014. *Eur. Econ. Rev.* **88**, 227–260. (doi:10.1016/j.euroecorev.2016.03.006)
27. Gu Y, Wang Z. 2022 Income inequality and global political polarization: the economic origin of political polarization in the world. *J. Chin. Polit. Sci.* **27**, 375–398. (doi:10.1007/s11366-021-09772-1)
28. Kubin E, von Sikorski C. 2021 The role of (social) media in political polarization: a systematic review. *Ann. Int. Commun. Assoc.* **45**, 1–19. (doi:10.1080/23808985.2021.1976070)

29. Milli S, Carroll M, Wang Y, Pandey S, Zhao S, Dragan AD. 2025 Engagement, user satisfaction, and the amplification of divisive content on social media. *PNAS Nexus* **4**, pgaf062. (doi:10.1093/pnasnexus/pgaf062)
30. Zhao W, Walasek L, Brown GDA. 2023 The evolution of polarization in online conversation: Twitter users' opinions about the COVID-19 pandemic become more politicized over time. *Hum. Behav. Emerg. Technol.* **2023**, 1–14. (doi:10.1155/2023/9094933)
31. Bernaerts K, Blanckaert B, Caluwaerts D. 2023 Institutional design and polarization. Do consensus democracies fare better in fighting polarization than majoritarian democracies? *Democratization* **30**, 153–172. (doi:10.1080/13510347.2022.2117300)
32. Young DJ, Madsen JK, de-Wit LH. 2025 Belief polarization can be caused by disagreements over source independence: computational modelling, experimental evidence, and applicability to real-world politics. *Cognition* **259**, 106126. (doi:10.1016/j.cognition.2025.106126)
33. Graham J, Haidt J, Nosek BA. 2009 Liberals and conservatives rely on different sets of moral foundations. *J. Personal. Soc. Psychol.* **96**, 1029–1046. (doi:10.1037/a0015141)
34. Merriam-Webster. 2025 Polarization. In *Merriam-Webster.com dictionary*. See https://www.merriam-webster.com/dictionary/polarization.
35. Cambridge Dictionary. 2025 Polarization. In *Cambridge dictionary (online)*. See https://dictionary.cambridge.org/dictionary/english/polarization.
36. Oxford English Dictionary. 2025 Polarization, II. figurative. In *Oxford English dictionary (online)*. (doi:10.1093/OED/3132747865)
37. Levitsky S, Ziblatt D. 2018 *How democracies die*. New York, NY: Penguin Random House.
38. Fiorina MP, Abrams SJ. 2008 Political polarization in the American public. *Annu. Rev. Polit. Sci.* **11**, 563–588. (doi:10.1146/annurev.polisci.11.053106.153836)
39. Reiljan A. 2020 'Fear and loathing across party lines' (also) in Europe: affective polarisation in European party systems. *Eur. J. Polit. Res.* **59**, 376–396. (doi:10.1111/1475-6765.12351)
40. Tichelbaecker T, Gidron N, Horne W, Adams J. 2023 What do we measure when we measure affective polarization across countries? *Public Opin. Q.* **87**, 803–815. (doi:10.1093/poq/nfad033)
41. Wagner M. 2021 Affective polarization in multiparty systems. *Elect. Stud.* **69**, 102199. (doi:10.1016/j.electstud.2020.102199)
42. Dalton RJ. 2021 Modeling ideological polarization in democratic party systems. *Elect. Stud.* **72**, 102346. (doi:10.1016/j.electstud.2021.102346)
43. Emanuele V, Marino B. 2024 Party system ideological polarization in Western Europe: data, trends, drivers, and links with other key party system properties (1945–2021). *Polit. Res. Exch.* **6**, 2399095. (doi:10.1080/2474736X.2024.2399095)
44. Koedam J, Binding G, Steenbergen MR. 2025 Multidimensional party polarization in Europe: cross-cutting divides and effective dimensionality. *Br. J. Polit. Sci.* **55**, e24. (doi:10.1017/S0007123424000474)
45. Edelman. 2023 *Edelman Trust Barometer: global report*. See https://www.edelman.com/trust/2023/trust-barometer.
46. Fernbach PM, Van Boven L. 2022 False polarization: cognitive mechanisms and potential solutions. *Curr. Opin. Psychol.* **43**, 1–6. (doi:10.1016/j.copsyc.2021.06.005)
47. Lindqvist E, Östling R. 2010 Political polarization and the size of government. *Am. Polit. Sci. Rev.* **104**, 543–565. (doi:10.1017/s0003055410000262)
48. Perrett S. 2021 A divided kingdom? Variation in polarization, sorting, and dimensional alignment among the British public, 1986–2018. *Br. J. Sociol.* **72**, 992–1014. (doi:10.1111/1468-4446.12873)
49. Abramowitz AI, Saunders KL. 2008 Is polarization a myth? *J. Polit.* **70**, 542–555. (doi:10.1017/s0022381608080493)
50. Bafumi J, Shapiro RY. 2009 A new partisan voter. *J. Polit.* **71**, 1–24. (doi:10.1017/S0022381608090014)
51. Dimock M, Doherty C, Kiley J, Oates R. 2014 *Political polarization in the American public: how increasing ideological uniformity and partisan antipathy affect politics, compromise, and everyday life*. Washington, DC: Pew Research Center. See https://www.pewresearch.org/wp-content/uploads/sites/4/2014/06/6-12-2014-Political-Polarization-Release.pdf.
52. Webster SW, Abramowitz AI. 2017 The ideological foundations of affective polarization in the U.S. electorate. *Am. Polit. Res.* **45**, 621–647. (doi:10.1177/1532673x17703132)
53. Mason L. 2015 'I disrespectfully agree': the differential effects of partisan sorting on social and issue polarization. *Am. J. Pol. Sci.* **59**, 128–145. (doi:10.1111/ajps.12089)
54. Morissette L, Chartier S. 2013 The k-means clustering technique: general considerations and implementation in mathematica. *Tutorials Quant. Methods Psychol.* **9**, 15–24. (doi:10.20982/tqmp.09.1.p015)
55. Astle DE, Bathelt J, Holmes J, CALM Team. 2019 Remapping the cognitive and neural profiles of children who struggle at school. *Dev. Sci.* **22**, e12747. (doi:10.1111/desc.12747)
56. Bathelt J, Holmes J, Astle DE, Holmes J, Gathercole S, Astle D, Manly T, Kievit R. 2018 Data-driven subtyping of executive function-related behavioral problems in children. *J. Am. Acad. Child Adolesc. Psychiatry* **57**, 252–262. (doi:10.1016/j.jaac.2018.01.014)
57. Shannon CE. 1948 A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 623–656. (doi:10.1002/j.1538-7305.1948.tb00917.x)
58. Bramson A, Grim P, Singer DJ, Berger WJ, Sack G, Fisher S, Flocken C, Holman B. 2017 Understanding polarization: meanings, measures, and model evaluation. *Philos. Sci.* **84**, 115–159. (doi:10.1086/688938)
59. Esteban JM, Ray D. 1994 On the measurement of polarization. *Econometrica* **62**, 819–851. (doi:10.2307/2951734)
60. Bauer PC. 2019 Conceptualizing and measuring polarization: a review. *SocArXiv*. (doi:10.31235/osf.io/e5vp8)
61. Benkler Y, Faris R, Roberts H. 2018 Polarization in American politics. In *Network propaganda* (eds Y Benkler, R Faris, H Roberts), pp. 295–310. New York, NY: Oxford University Press. (doi:10.1093/oso/9780190923624.003.0010)
62. Greenacre M, Groenen PJF, Hastie T, D'Enza AI, Markos A, Tuzhilina E. 2022 Principal component analysis. *Nat. Rev. Methods Prim.* **2**, 100. (doi:10.1038/s43586-022-00184-w)

63. Dalmaijer ES, Nord CL, Astle DE. 2022 Statistical power for cluster analysis. *BMC Bioinform.* **23**, 205. (doi:10.1186/s12859-022-04675-1)
64. Slonim N, Aharoni E, Crammer K. 2013 Hartigan's K-means versus Lloyd's K-means—is it time for a change? In *Proc. 23rd Int. Joint Conf. on Artificial Intelligence*, pp. 1677–1684. Cambridge, MA: AAAI Press. See https://www.ijcai.org/Proceedings/13/Papers/249.pdf.
65. R Core Team. 2025 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See https://www.R-project.org/.
66. American National Election Studies. 2022 ANES time series cumulative data file (version September 16, 2022 version) [Dataset and documentation]. See https://electionstudies.org/.
67. American National Election Studies. 2025 ANES 2024 time series study preliminary release: combined pre-election and post-election data [Dataset]. See https://electionstudies.org/.
68. The National Archives Catalog. 2024 Reparative description preferred terms: black person. See https://www.archives.gov/research/catalog/lcdrg/appendix/black-person (accessed 20 May 2024).
69. Schiffer AJ. 2000 I'm not that liberal: explaining conservative democratic identification. *Polit. Behav.* **22**, 293–310. (doi:10.1023/A:1010626029987)
70. Haerpfer C (ed). 2022 World Values Survey trend file (1981-2022) cross-national data-set. Data file version 2.0.0. Madrid, Spain/Vienna, Austria: JD Systems Institute & WVSA Secretariat. ()
71. EVS. 2022 EVS trend file (1981-2017) (data file ZA7503; version 3.0.0). GESIS Data Archive. See https://doi.org/doi:10.4232/1.14021.
72. Welzel C. 2013 *Freedom rising*. Cambridge, UK: Cambridge University Press. (doi:10.1017/CBO9781139540919)
73. Druckman JN, Peterson E, Slothuus R. 2013 How elite partisan polarization affects public opinion formation. *Am. Polit. Sci. Rev.* **107**, 57–79. (doi:10.1017/s0003055412000500)
74. Sunstein CR. 2018 *#Republic: divided democracy in the age of social media*. Princeton, NJ: Princeton University Press.
75. Arel-Bundock V, Enevoldsen N, Yetman C. 2018 countrycode: an R package to convert country names and country codes. *J. Open Source Softw.* **3**, 848. (doi:10.21105/joss.00848)
76. Drazanova L. 2020 Introducing the Historical Index of Ethnic Fractionalization (HIEF) dataset: accounting for longitudinal changes in ethnic diversity. *J. Open Humanit. Data* **6**, 6. (doi:10.5334/johd.16)
77. Rochat P, Dias MDG, Guo L, Broesch T, Passos-Ferreira C, Winning A, Berg B. 2009 Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *J. Cross Cult. Psychol.* **40**, 416–442. (doi:10.1177/0022022109332844)
78. Van Staveren I, Pervaiz Z. 2017 Is it ethnic fractionalization or social exclusion, which affects social cohesion? *Soc. Indic. Res.* **130**, 711–731. (doi:10.1007/s11205-015-1205-1)
79. Ghai S. 2021 It's time to reimagine sample diversity and retire the WEIRD dichotomy. *Nat. Hum. Behav.* **5**, 971–972. (doi:10.1038/s41562-021-01175-9)
80. Inglehart R, Welzel C. 2005 *Modernization, cultural change, and democracy*. Cambridge, UK: Cambridge University Press. (doi:10.1017/CBO9780511790881)
81. Lipset SM. 1959 Some social requisites of democracy: economic development and political legitimacy. *Am. Polit. Sci. Rev.* **53**, 69–105. (doi:10.2307/1951731)
82. Ruck DJ, Bentley RA, Lawson DJ. 2018 Religious change preceded economic change in the 20th century. *Sci. Adv.* **4**, eaar8680. (doi:10.1126/sciadv.aar8680)
83. Treisman D. 2020 Economic development and democracy: predispositions and triggers. *Annu. Rev. Polit. Sci.* **23**, 241–257. (doi:10.1146/annurev-polisci-050718-043546)
84. Baker R. 2019 *The structure of moral revolutions*. Cambridge, MA: MIT Press.
85. Norris P, Inglehart R. 2019 *Cultural backlash: Trump, Brexit, and authoritarian populism*. Cambridge, UK: Cambridge University Press. (doi:10.1017/9781108595841)
86. Gregorich M, Strohmaier S, Dunkler D, Heinze G. 2021 Regression with highly correlated predictors: variable omission is not the solution. *Int. J. Environ. Res. Public Health* **18**, 4259. (doi:10.3390/ijerph18084259)
87. DeBell M, Ambsary M, Brader T, Brock S, Good C, Kamens J, Maisel N, Pinto S. 2022 Methodology report for the ANES 2020 Time Series Study. See https://electionstudies.org/wp-content/uploads/2022/08/anes_timeseries_2020_methodology_report.pdf.
88. Reiljan A, Ryan A. 2021 Ideological tripolarization, partisan tribalism and institutional trust: the foundations of affective polarization in the Swedish multiparty system. *Scand. Polit. Stud.* **44**, 195–219. (doi:10.1111/1467-9477.12194)
89. Brick C, Hood B, Ekroll V, de-Wit L. 2021 Illusory essences: a bias holding back theorizing in psychological science. *Perspect. Psychol. Sci.* **17**, 491–506. (doi:10.1177/1745691621991838)
90. Peresman A, Carroll R, Bäck H. 2021 Authoritarianism and immigration attitudes in the UK. *Polit. Stud.* **71**, 1–18. (doi:10.1177/00323217211032438)
91. Young DJ, Ackland J, Kapounek A, Madsen JK, Greening LJ, de-Wit L. 2025 Supplementary material from: a new measure of issue polarisation using k-means clustering: US trends 1988-2024 and predictors of polarisation across the world. Figshare (doi:10.6084/m9.figshare.c.8190166)