How Polarized are Citizens? Measuring Ideology from the Ground-Up

Mirko Draca,* Carlo Schwarz†

April 2, 2018

Abstract

Strong evidence has been emerging that major democracies have become more politically polarized, at least according to measures based on the ideological positions of political elites. We ask: have the general public ('citizens') followed the same pattern? Our approach is based on unsupervised machine learning models as applied to issue-position survey data. This approach firstly indicates that coherent, latent ideologies are strongly apparent in the data, with a number of major, stable types that we label as: Liberal Centrist, Conservative Centrist, Left Anarchist and Right Anarchist. Using this framework, and a resulting measure of 'citizen slant', we are then able to decompose the shift in ideological positions across the population over time. Specifically, we find evidence of a 'disappearing center' in a range of countries with citizens shifting away from centrist ideologies into anti-establishment 'anarchist' ideologies over time. This trend is especially pronounced for the US.

JEL *classification*: D72, C81.
Keywords: Polarization, Ideology, Unsupervised Learning.

*Department of Economics, University of Warwick and Centre for Economic Performance, LSE
†Department of Economics, University of Warwick and Centre for Competitive Advantage in the Global Economy (CAGE)

1 Introduction

In political terms, we seem to be living in the midst of the proverbial 'interesting times'. Across established democracies there appear to be strong trends of political populism and ideological polarization. In the US, a large body of evidence indicates that the political positions taken by elected representatives in legislatures have sharply polarized. For example, this is apparent in recent work examining partisanship in the use of political language (Jensen et al., 2012; Gentzkow et al., 2016). In particular, Gentzkow et al. (2016) isolate this increase as occurring from the mid-1990s onwards, a period when the nature of political communication changed as parties became more acutely strategic with their use of language. Further evidence of 'elite polarisation' is also found in the extensive literature (following Poole and Rosenthal, 1985) that has measured the evolving ideological positions of elected representatives using data on Congressional rollcall voting.

By comparison, the evidence about political polarisation amongst the general public (or 'citizens') is more contested than the findings that have emerged for political elites. In the US, contributions such as Fiorina and Abrams (2008) make the point that both the underlying distribution of views across issues and the level of self-identification with 'strong' political positions has been stable over time. Similar skepticism about citizen polarisation in the US is also evident in the studies of Glaeser and Ward (2006) and Ansolabehere et al. (2006) . In Europe, recent contributions by Algan et al. (2017) and Guiso et al. (2017) have documented a strong pattern of populist politics across the continent that appears to have roots in changing economic conditions. However, this populist trend is not necessarily symptomatic of ideological polarisation. For example, Algan et al. (2017) detect no significant shift in political positioning along the left-right scale in their cross-country sample and actually pick up a decline in close party identification.

In this paper we offer a new approach to measuring citizen ideology and political polarisation using unsupervised machine learning tools as applied to 'issue-position' data on individual political views. In short, the core of our approach is based on applying Latent Dirichlet Allocation (LDA) topic models (Blei et al., 2003) to individual-level survey responses across a typical range of social and economic issues. Topic models are mainly known in the social sciences for their use in the analysis of text, in particular for their capacity in identifying the latent topic structure that underpins the generation of documents across various corpora. Applications of topic modelling have thus proliferated recently with empirical studies of text data across a range of social science questions (Gentzkow et al., 2017).

In contrast, rather than analysing text we instead make discretely coded 'issue-positions' the main objects of analysis in our application. Following along from this, we then frame the latent topics as the *political ideologies* that underpin the generation of individual political beliefs amongst the general public. The unique advantage of this approach is that it is based on a probabilistic generative model of ideology, allowing individual beliefs to be explained as mixtures of latent ideologies. As such, it is a concept of ideology that is directly empirical, that is, built up from the statistical pattern of political views across the population. Our approach also allows the identified citizen ideologies - defined practically as probability distributions over issue-positions - to evolve over time such that 'within' and 'between' shifts in ideology can be clearly measured. The general approach we adopt of using topic models to analyse discrete, non-text data is closest to (and indeed, draws inspiration from) Bandiera et al. (2017)'s empirical model of behavioral manager 'types' in CEO time-use data.

We use this methodology to explore two main questions. Firstly, we are able to ask: to what extent do the general public hold beliefs that can be summarised as statistically coherent 'ideologies'? Further to this point, to what extent do the latent ideologies found in the data conform to the traditional left-right ideological line that dominates both popular discourse and classic formal models in the spirit of Downs (1957)? This assumption of systematic coherence in political views within the population has been challenged by recent critiques of the principle of retrospective voting that have explored how voters use information on political performance (Healy and Malhotra, 2013)[1], as well as puzzles about citizen political views that have emerged from research on subjects such as preferences over redistribution (e.g Ashok et al., 2015)[2].

Building further on this, the second main question we address is then: how do the empirically-based citizen ideologies we identify vary across countries and over time? In practical terms, this involves studying the factors that determine the ideological mixture of views held by citizens at the individual level and assessing the extent to which aggregate shifts can be explained in terms of changing demographics or other observables. Importantly,

[1] This literature has focused on how voters use information as part of voting decisions. For example, Achen and Bartels (2002) point out the apparent sensitivity of voters to arbitrary local events while contributions such as Wolfers (2002) and Leigh (2009) test for evidence of economic voting.

[2] For example, data on evolving political views indicates that the demand for re-distribution via taxation is not increasing despite higher economic inequality. Ashok et al. (2015) show that, in US data, this cannot be explained by a general ideological shift to the right and focus their explanations on how re-distributive preferences vary by demographic sub-group.

because our topic modelling approach allows for the mixed membership of individuals with respect to latent ideologies, this lets us parse individual ideological positions very finely. As a result, we are able to develop a measure of 'citizen slant' that captures the degree to which individuals weigh alternative ideologies within their overall beliefs, for example, the extent to which a given person is, say, 'a bit conservative and a bit liberal'. We are then able to use this measure of slant to better characterise overall patterns of political polarisation. In particular, we put forward an analysis of multidimensional polarisation over more than two ideologies following the framework of Esteban and Ray (1994) and Duclos et al. (2004).

The main empirical setting for applying our methodology is the cross-country World Values Survey (WVS) which provides a wide-ranging set of consistently asked questions from the late 1980s onwards. In answer to our first major question, a series of coherent citizen ideologies do indeed emerge from our modeling. A left-right dimension is strongly evident in the data but alongside this there is another major ideological dimension that pivots on citizen confidence in institutions. We generically label the ideological types that are characterized by low confidence in institutions as 'anarchist' but note that the broad position that this type represents is consistent with the anti-establishment or populist positions that have been the focus of recent research (Acemoglu et al., 2013; Piketty, 2018; Rodrik, 2017). The anarchist label that we use is meant to avoid pejorative interpretations of terms such as populist[3] and emphasize opposition to current institutional structures as the defining feature of this ideological type.

Our main empirical model of ideology therefore takes the shape of a 4-type model where we label the main types as Liberal Centrist, Conservative Centrist, Left Anarchist, and Right Anarchist. These types emerge as part of a clear hierarchy of empirical ideologies that becomes apparent as we allow our unsupervised learning models to identify different numbers of latent issue-position clusters. Hence, an important finding from our analysis is that anti-establishment or populist trends in opinion are potentially rooted in a formal, statistically coherent citizen ideology.

This first finding regarding the nature of the ideologies at play then feeds into our second set of findings on the variation that is evident across countries and time periods. Firstly, at the level of the latent ideologies we find that our 4 main ideological types are stable over time with limited 'within ideology' changes, as measured by the weighting of different

[3]For example, see media critiques such as 'Populism: It's the BBC's new buzzword, being used to sneer at the 'uneducated' 17 million who voted for Brexit' from the UK's *Daily Mail* (Murray (2016)).

issue-positions. The most notable finding here is an increase in the intensity of socially liberal attitudes. For example, the Conservative Centrist type shifts in their attitudes on issues such as homosexuality and the status of immigrants. The Right Anarchist type also shows signs of moving in a potentially authoritarian direction by expressing higher confidence in institutions such as the police and the armed forces, as well as more acute hostility towards immigration.

Secondly, we analyse aggregate type shares, that is, the extent to which citizens are drawing from the different ideological types when we sum across all individuals. This shows an ordering across countries that is consistent with previous studies of cross-national political attitudes, for example, northern European countries are more liberal while countries with stronger religious traditions are more conservative. In turn, this is reinforced by a sensible pattern of correlations between individual level characteristics and type shares (eg: women are more liberal and conservatism increases with age). The composition of the aggregate type shares is also stable across time for most countries. However, a notable exception is the US where the total type share for the two Anarchist types increases from around 30% in the 1989-1993 wave to 50% by the fifth WVS wave in 2005-2009. The majority of this increase is accounted for by the Right Anarchist type.

The final part of our analysis then studies individual-level 'citizen slant' and societal polarization. The citizen slant measure we calculate provides a within-person measure of ideological concentration and is constructed following a basic Gini index logic. It directly exploits the mixed membership format of our unsupervised learning framework to capture the extent to which individuals might be heterogeneous in their views (ie: 'a bit conservative and a bit liberal'). We find that the mean citizen slant across types, countries and years is relatively high at around 0.7 on a 0-1 scale. The degree of slant or within-person concentration has also increased over the time we consider. There is a slight increase in the case of Europe (of around 2-3% relative to the baseline in the initial wave) but much stronger shifts are apparent in the US. The rise in the US is also focused heavily on the Anarchist types (which increased their slant by around 15%) as well as the Centrist Conservative type (a 4.3% increase).

We then develop an overall measure of polarization that builds on the framework of Esteban and Ray (1994) and Duclos et al. (2004). This framework is novel for offering a systematic, multi-polar analysis of ideology in terms of own-group *identification* and between-group *alienation*. We find that changes in the level of polarization over time are muted, with an average 5-10% increase the our sample period. Again, the US stands out as experiencing the sharpest increase, chiefly driven by the compositional change in type shares noted above.

Related Literature. The nature of this paper's main topic (pun unintended) means that it has connections with many literatures and contributions. Some areas to highlight are the following. Firstly, there is the literature on democratic politics and populism, with recent examples that include: Acemoglu et al. (2013), Algan et al. (2017), Buisseret and Van Weelden (2017), Bursztyn et al., Dal Bo et al. (2017), Dal Bo et al. (2018), Guiso et al. (2017), and Rodrik (2017)[4]. As discussed, our work sheds light on the potential long-run ideological underpinnings of these political trends in the population.

Secondly, there is fast-growing literature that studies aspects of ideology, policy-making and political communication using tools from machine learning and natural language processing. This includes the already noted Gentzkow et al. (2016) and Jensen et al. (2012), as well as other text-based studies such as: Ash (2015), Grimmer (2009), Hansen et al. (2014) and Jelveh et al. (2015). Another branch of this overall literature (Blaydes and Grimmer (2013), Gross and Manrique-Vallier (2012)), Wang et al. (2017)) has also begun to explore the application of unsupervised learning tools to survey response data.

Finally, there is a large literature that explicitly addresses polarization and fractionalization along political, ethnic and cultural lines. This literature often focuses on measuring group structure in societies and relating this to patterns of conflict. An indicative list includes: Alesina et al. (2003), Bossert et al. (2011), Caselli and Coleman (2013), Duclos et al. (2004), Esteban and Ray (1994) and Montalvo and Reynal-Querol (2005). Our paper adds to this literature by applying this perspective to the analysis of ideologically-defined groups.

Structure. The paper is organized in the typical way. In section 2, we outline the main data used, namely the World Values Survey (WVS) as well as our approach to defining answers to survey questions as 'issue-positions'. Section 3 outlines an unsupervised learning methodology for studying this issue-position data. This includes details on how we develop a hierarchy of ideological types and select the optimal number of topics in our LDA models. Section 4 outlines the results and section 5 concludes.

[4] A range of studies that have looked at the recent determinants of voting patterns are also relevant here: Becker et al. (2017), Dippel et al. (2015), Dorn et al. (2016) and Che et al. (2016).

2 Data

World Values Survey

For our main analysis, we use data from the World Values Survey (WVS) and the European Values Study (EVS). These surveys are an output of a global research project conducted by a large network of social scientists and run via a non-profit association based in Stockholm. They have been widely deployed in social science research and some prominent studies using the data include: Alesina et al. (2013, 2001); Blanchflower and Oswald (2008); Inglehart (1997); and Norris (2016).

The WVS consists of 6 waves from 101 countries while the EVS consists of 4 Waves from 48 countries. We construct what is formally known as the Integrated Value Survey (IVS) by combining the two datasets. The resulting dataset contains the 4 EVS waves and the corresponding waves 1, 2, 4 and 5 from the WVS[5]. For the sake of simplicity we refer to this combination of the data as the 'World Values Survey (WVS)'.

The set of questions asked and countries covered differs across successive waves of the WVS. We therefore develop a sample of WVS observations based on the principle of capturing the widest range of consistently asked questions over waves and across countries. Since the first wave has limited country and question coverage[6] we construct our sample from the second wave onwards and develop a set of 17 countries in Europe and North America (Austria, Belgium, Canada, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, Malta, Netherlands, Portugal, Spain, Great Britain, United States, North Ireland) and 29 questions. The selected questions cover issues such as abortion, immigration, sexuality, the role of government, and confidence in institutions. The resulting dataset contains a total of 82,338 observations over 3 waves spanning the years from 1989 to 2010.

Construction of Features

As part of the data preparation, we unify the coding of the questions and convert them to the same scale. The intention here is to represent the answers to the survey questions as discrete 'features' for the subsequent topic modelling. Specifically, we recode the responses for each of the 29 questions into two indicator variables expressing either support or opposition to each issue, for example an indicator variable if the person believes that abortion is justifiable and

[5]There is no wave of EVS that corresponds to the 6th wave of the WVS. Therefore, our main dataset ends in 2010 so as to focus on a consistently defined set of repeated country-question cross-sections .

[6]The countries Austria and Portugal as well as 7 complete questions are not contained in the first wave.

a second indicator variable if the person opposes abortion. In cases where a person expressed neither support or opposition to an issue both binary variables are coded as zero.

Summary statistics for the 58 recoded issue positions can be found in Table 1. Importantly, the features cover a broad range of salient political issues. A number deal with what would be typically classified as 'social issues' such as abortion, prostitution and attitudes towards minority groups while three questions deal with classic economic questions relating to the role of government, private sector competition and support for the welfare state. Finally, there is a set of questions dealing with confidence in a comprehensive set of social and political institutions.

The information in Table 1 indicates a rich mix of positions across political issues. There is a current of anti-foreigner sentiment with 12.3% of respondents preferring not to have immigrants as neighbours, and this is backed up by an overwhelming 60% endorsing a priority for native workers in the allocation of scarce jobs. However, most respondents either hold liberal or neutral views on leading social issues such as abortion and prostitution. There is also a widespread lack of confidence in key institutions, with only around 35-45% expressing a favourable view of the press, parliaments, the civil service and major companies.

3 Discovering Latent Ideology using LDA

We develop an approach based on machine learning topic models that investigates the pattern of responses in the WVS data in terms of a generating structure characterized by latent political ideologies. To be clear, we will define an 'ideology' as a probability distribution of issue-position responses across questions. Since the basic methods we use are most commonly applied to the analysis of text in terms of underlying topics or subjects, we first outline how we adapt the methods to study citizen ideology. We then describe an approach for model selection, that is, discerning the number of topics or ideological types that best describe the data. Finally, we discuss how we track changes over time within our overall topic model methodology.

3.1 Discovering Citizen Ideology via Latent Dirichlet Allocation (LDA)

The basis of our approach is Latent Dirichlet Allocation (LDA) (Blei et al., 2003) which can be summarised as a Bayesian hierarchical model that defines a probabilistic structure for joint distributions of observed data and latent generating factors. It was originally developed for the unsupervised classification of text data into a user-chosen number of topics.

In the context of text-based applications LDA makes use of the fact that authors tend to use similar words when they talk about the same topic. For example, a text containing the words ‘equilibrium’ and ‘preferences’ is far more likely to be about economics than sports. LDA is therefore built on the principle of algorithmically classifying any corpus of text documents as a probabilistic mixture of underlying topics. Again, as an example, a document discussing a Pigovian tax might get classified as a mixture of a taxation and an environmental policy topic. Each LDA topic is therefore defined as a probability distribution over words. A taxation topic for example might put high weights on the words ‘tax’, ‘revenue’ and ‘IRS’. Since the LDA algorithm itself does not provide any topic labels and the standard machine learning topic labeling approaches (e.g Lau et al., 2011; Aletras et al., 2014) are not applicable in our setting, it is up to the user to interpret and judge the focus of each topic. However, some metrics for ‘topic coherence’ are available for assessing the quality of a given topic model and assisting in this process of labeling and interpretation.

At its core LDA is a clustering algorithm for discrete data. As a result, LDA can therefore be used in non-text applications, for example image classification tasks in the field of computer vision (e.g. Putthividhy et al., 2010)[7]. For our study, we apply LDA to the WVS survey responses of individuals. Instead of clustering frequently co-occurring words into a topic, LDA will combine issue positions that are frequently held together into an ‘ideological type’. Each of the respondents in the WVS will also be classified as a mixture of ideological types based on their answers to questions, for example, as 20% ‘conservative’ and 80% ‘liberal’.

Each ideological type will be described by a probability distribution over issue positions. This probability distribution describes how important the individual issue positions are for each ideological type. Our general approach is most closely analogous to Bandiera et al. (2017) who model CEO time use across discretely-defined activities.

The advantage of LDA in comparison to other clustering algorithms is that it provides a generative model of the data and thereby a quasi-microfoundation. For example, neither Principal Component Analysis (PCA) or Factor Analysis (FA) model the latent types of each individual directly. Moreover, both PCA and FA use linear transformations of the data while LDA allows for non-linear relationships. Overall, LDA is better suited for categorical

[7] See also the collection by Airoldi et al. (2014) for a diverse set applications of mixed membership modeling.

data than either PCA or FA. Another advantage of LDA is that it is a mixed membership model which describes every observation as a mixture of types rather than in terms of some attachment to a single type or category as Latent Class Analysis, k-means, or spectral clustering.

Underlying our LDA model of citizen ideology is a probabilistic model which assumes that every individual $i \in I$ can be described as a probabilistic mixture of $t \in T$ topics or 'types'. These probabilities are contained in a vector θ_i of type proportions. The latent T types are described by 'type vectors' β_t with a question response profile for each of the Q questions. The entries in the type vector give the probability of holding a particular issue-position when drawing from particular latent type. The generative process underlying the data is defined as:

1. For each individual i in the data draw ideological type proportions $\theta_i \sim Dir(\alpha)$
2. For each of the q question of individual i:
   - Draw a type assignment $z_{i,q} \sim Mult(\theta_i)$
   - Draw a response $r_{i,q}$ from $P(r_{i,q}|z_{i,q}, \beta)$

Given this generative process the probability of the observed survey responses is:

$$\prod_{i=1}^{I} P(\theta_i|\alpha) \left(\prod_{q=1}^{Q} \sum_{z_{i,q}} P(z_{i,q}|\theta_i) P(r_{i,q}|z_{i,q}, \beta) \right) \tag{1}$$

The first term describes how likely it is to observe an individuals ideological type proportions θ_i. The second term in brackets is the probability of observing the individual's i responses to the Q questions. LDA identifies ideological types by finding parameter values for α, β and θ_i such that this probability is maximized. Simply maximizing this likelihood for the relevant parameters is computationally unfeasible. LDA therefore makes use of an approximate inference algorithm. We use the inference algorithm developed by Hoffman et al. (2010, 2013) and implemented by (Pedregosa et al., 2011).

In our application the assumptions of the independence of responses do not strictly hold. If a question has been answered the same question cannot be answered again by the same person. The inference of LDA is nonetheless still valid, since the bias in $P(r_{i,q}|z_{i,q})$ is identical for all types. Therefore, $z_{i,q}$ still represents the correct probability of a person belonging to one ideological type. Only the interpretation of the β vector changes. We discuss this point in detail in Appendix C.

3.2 Determining the Optimal Number of Types

LDA makes it possible to estimate any number of ideological types. Therefore, the question of model selection is crucial for understanding which level of topic model best describes the data. To find the optimal number of ideological types we modify the methods for understanding topic cohesion which have been developed in recent years (e.g. Chang et al., 2009; David Newman et al., 2010; Aletras and Stevenson, 2013; Lau et al., 2014). Following standard k-fold cross-validation principles, we randomly split the data from the largest wave in our sample (wave 5) into 10 folds (each 10% of the data). Nine folds are then grouped into a training sample and the remaining becomes the test sample. Afterwards, we fit LDA models with different numbers of topics to the training sample and formally evaluate them according to statistical criteria relative to the test sample. In each run of LDA a different test sample is chosen. The resulting ideological types for the different models can be found in Table 2.

We then chose the optimal number of ideological types based on the cohesion of the generated types. A type is more coherent if the issue positions with the largest weight for that type also frequently appear together in the held-out survey responses of WVS participants. The intuition behind this is that more coherent ideological types should put more weight on issue positions that people frequently hold together, e.g. the co-occurrence of the views that abortion and suicide are not justifiable. This approach is preferable to evaluating the likelihood or perplexity of the model in the hold out data, since the hold-out likelihood is not necessary a good predictor for human judgment of topic cohesion (see for example Chang et al., 2009).

As a measure of co-occurrence of issue positions, we use Normalized Pointwise Mutual Information (NPMI). NPMI is defined as:

$$NPMI_{i,j} = \frac{PMI_{i,j}}{-\ln(p(i,j))} = \frac{\ln\left(\frac{p(i,j)}{p(i)\cdot p(j)}\right)}{-\ln(p(i,j))} \tag{2}$$

Pointwise Mutual Information (PMI) is simply defined as the log-ratio of the joint and marginal probabilities. Hence, PMI measures how probable it is that two features i and j appear together in comparison to how often we would expect them to appear together if the features were independent from each other. NPMI additionally normalizes PMI between

$[-1, 1]$. If two features always appear together their NPMI will be 1. In the case where two features never appear together their NPMI will be -1.[8]

The average NPMI for all pairwise combinations of the B most important issue positions of an ideological type t is then given by:

$$\overline{NPMI_t} = \frac{\sum_i^B \sum_{j \neq i}(NPMI_{i,j})}{B \cdot (B-1)} \tag{3}$$

Similarly, the overall cohesion for a model with M ideological types can therefore be calculated from the hold-out sample as:

$$Cohesion_m = \frac{\sum_t^M \overline{NPMI_t}}{M} \tag{4}$$

Follow the findings of Lau and Baldwin (2016) we average our measure of cohesion over different number of features $B \in (5, 10, 15, 20)$. The type cohesion scores can be found in Figure 2. As we discuss later, based on these scores we choose the 4 type LDA specification as our benchmark model, since it seems to best describe the pattern of responses across citizens.

3.2.1 Dynamic Type Models - Ideological Change Over Time.

The three waves of the WVS that we use stretch over 20 years. For our analysis we therefore want to allow for the ideological types to change over time. We do this by fitting LDA models separately to the 3 waves in our sample and only linking the ideological types together afterwards based on the similarity of their issue positions. Our approach is more generic than a dynamic topic model (Blei and Lafferty, 2006) or continuous topic model (Wang et al., 2008) since we neither impose any assumptions on the dynamics of the ideological types nor on the shares of the types over time. The general structure of our approach is most closely related to the topic chains suggested in Kim and Oh (2011) and has the advantage of allowing for completely different ideological patterns to emerge in each wave. But, as we will see, the topic structure in our WVS data displays a remarkable level of stability over time.

4 Results

We report our results across four linked sub-sections. In the first sub-section, we show the results of our LDA models in terms of different variants of type model - from 2-types to

[8] More details on the topic cohesion literature and an example for the calculation of NPMI can be found in Appendix F.

5-types. The results here indicate a coherent hierarchy of types across the models such that types can be seen to 'split off' into related families as we move to higher order models. The second sub-section then applies the NPMI model selection criteria outlined above to the different orders of type models with the conclusion that the 4-type model is the most preferred.

We then use the 4-type model as our main vehicle of analysis in the third sub-section, focusing on within-type and between-type differences over time. To guide the reader, this boils down to a close study of the β type vectors in the LDA model, that is, the probability distribution of issue-positions per estimated type. In the final subsection we focus on how the distribution of type proportions - essentially the θ_i values outputted by the LDA model - play out over countries and time. In turn, this leads to our analysis of within-person slant and country-level polarization.

4.1 Hierarchy of Ideological Types.

In Table 2 we summarize the results of various orders of LDA model, reporting the 'top ten' features for each type. These top ten features represent the issue-positions with the highest probability values in the β type vectors and are effectively the defining features of each ideological type. We present the results as separate panels in the table per order of type model.

Panel (a) shows the results for the basic 2-type model in the first column. These two types are distinguished by stances on social issues - for example, a liberal attitude towards minority groups (eg: reporting 'no problems' with neighbours who are homosexuals or immigrants) by one type and conservative positions on social issues such as abortion and prostitution by the other type. We therefore label these types in panel (a) generically as 'Left' and 'Right'. Across the 58 features the β topic vectors for these types have a correlation of 0.39, indicating that they have some common components.

The second column of panel (a) then reports the top features for the 3-type model. Two 'Left' and 'Right' types distinguished mainly by their positions on social issues such as sexuality, race and abortion are still apparent. However, the most striking result from this model is the nature of the third type. Rather than being a simple mixture of the basic Left-Right types of the earlier model the third type draws on a qualitatively different set of issue-positions for its top features. Specifically, the third type draws heavily on features that represent low confidence in major institutions such as parliament, the civil service and major companies. We provide a more detailed discussion of the rationale for our type labels in the next sub-section but here we flag type 3 as an 'Anarchist' type to reflect this type's opposition to the current workings of major social institutions. In contrast, the main left and right types in the 3-type model report confidence in institutions across the majority of features in this category. We therefore label these types as 'Liberal Centrist' and 'Conservative Centrist' to reflect their contrasting positions on social issues but common pattern of support for established political institutions.

The top features for the 4-type model are reported in the third column of panel (a), Table 2. The type structure continues to evolve here. Most notably, two anarchist types now become apparent, again distinguished by contrasting views on social issues but similar positions in terms of confidence (or lack thereof) in institutions. These are labeled 'Left Anarchist' and 'Right Anarchist' to reflect this. Intuitively, the top ten features reported in panel (c) suggest a splitting of the Anarchist type from the 3-type model has occurred.

We can validate this by examining the cross-model correlations in the weights on issue-positions in the β type vectors. These correlations are useful for indicating how close the individual types in the 4-type model are to those in the lower order 3-type order. We report these in Figure 1. In line with the intuitive 'eyeballing' of the top features, the Left Anarchist and Right Anarchist types are most strongly correlated with the Anarchist type from the 3-type model, with correlation measures of 0.92 and 0.64 respectively. This splitting of the Anarchist type is reinforced by the continuity in the Liberal Centrist and Conservative Centrist types as we go from the 3-type to 4-type model. These two types can be tracked across the different hierarchies of type model, with correlations of 0.98 (Liberal Centrist) and 0.87 (Conservative Centrist) across the models.

The top features for a further 5-type model are reported in panel (b) of Table 2. Greater subtlety in the types becomes evident here. The set of Liberal Centrist, Conservative Centrist and Left Anarchist types remain intact relative to the 4-type model but there appears to a splitting of the Right Anarchist type. Two variants of the Right Anarchist emerge. One variant still expresses a lack of confidence in institutions but appears to be liberal on social issues and is notably pro-competition, one of the only instances when an economy-focused feature makes it into the top ten features for a type across our models. We label this type as 'Market Liberal'[9].

[9]Our nickname for this anarchistic, pro-market and socially liberal type is 'George Mason University (GMU) Blogger' in reference to the well-known *Marginal Revolution* blog.

The other variant of the Right Anarchist is not socially liberal, with a string of conservative positions on minorities and social issues amongst its top ten features. The correlations indicate that this type is strongly correlated (0.644) with the original Right Anarchist from the 4-type model but negatively correlated with the Liberal Centrist (-0.157) and Left Anarchist types (-0.306) types. In line with, our label for this type is 'Right Anarchist (Hard Social Conservative)'.

Further results on potential 6-type and 7-type models are reported in the Appendix Table 7. These indicate a further splitting of the right-wing types (6-type model) and the emergence of a nihilistic 'Super Anarchist' type (7-type model). We defer detailed discussion of these models for the appendix but note here that (as might be expected) intuitive labels for types in the higher order models are less obvious. So far, the results presented above indicate that both the low-order (2 or 3 type) and higher-order (4 plus types) models offer plausible sets of types and, considered together, can be interpreted in terms of coherent hierarchy. We therefore turn to the question of formal model selection using the NPMI framework outlined above.

4.2 Model Selection and Type Labeling.

4.2.1 Automatic Model Selection

Our NPMI framework for assessing model cohesion is based on comparing predictions of feature co-occurrence in hold-out data. Simply put, the approach asks: to what extent do the (say) top 10 features suggested by our type models occur together in data held out from the original estimation of the given model? Figure 2 reports the results of this exercise. The x-axis denotes the order of model we are estimating (going from a 1-type model up to a 10-type model) while the y-axis denotes the resulting cohesion score.

The cohesion scores show a inverse u-shape pattern. At first, the cohesion score increases with the number of topics. After the number of topics increases beyond 7 the cohesions scores begin to fall. Overall, the most cohesive models (M) appear to be either the 4 or 7 type model. We decided to use the 4 type model, since it delivers nearly identical fit to the 7 type model but with a smaller number of model parameters. Our analysis from this point therefore employs the 4-type model composed of the Liberal Centrist, Conservative Centrist, Left Anarchist and Right Anarchist types.

4.2.2 Alternative Models

In our appendix we discuss two further modeling issues that relate to the quality of the information provided via the LDA framework. Firstly, in Appendix D we look at the sensitivity of our baseline 4-type model to the removal and addition of features. The basic model is very robust to the removal of features with types from iterative 'leave one out' models showing a high correlation with the types in the original model. The relative ordering of β weights is also preserved when we substantially widen the feature set (ie: add lots more questions). Both of these exercises provide re-assurance that our overall baseline feature set is comprehensive enough to identify stable types in the data.

The second modeling issue that we examine (in Appendix E) is a comparison of our LDA approach with other unsupervised learning methods. Specifically, we apply principal components analysis (PCA), factor analysis (FA) and k-means clustering to the same discretized feature data as our LDA models. As we discuss in Appendix E, these alternative approaches are distinct from LDA in that they are linear methods and capture mixed membership relationships in a less intuitive way. For example, a method such as PCA will pick out linear combinations of features with the highest degrees of variance in the data and therefore may not parse more complex data generating processes.

This is borne out in the types derived from these models which are reported in Appendix Tables 13, 14 and 15. The PCA models tend to identify conservative and anti-establishment types as part of the main model components, with no clear centrist or socially liberal types emerging. The identified types also display some counter-intuitive groupings of features. The FA and k-means models produce similar results. Further to this, no plausible hierarchy or 'family' of types emerges from these alternative models. Again, this provides re-assurance that our LDA models - which are, after all, specifically intended for the analysis of discrete multinomial data - identify stable and interpretable types that are difficult to pin down using other methods.

4.2.3 Type Labeling

The labeling of our LDA-derived types is a point for discussion. An important advantage of our approach is that it is based on ideologies that emerge from the 'bottom-up' collection of views amongst the general public. The topics that we identify are empirical ideologies and may not necessarily have a tight mapping to traditional taxonomies of ideology[10].

Our labeling is therefore based on capturing the main empirical differences in issue-positions between types. Arguably, the main issue here is the labeling of types 3 and 4, which we have dubbed Left and Right Anarchist respectively. These two types are strongly distinguished by issue-positions that hinge on (low) confidence in institutions[11]. We use the term 'anarchist' to denote a pattern of opposition to current structures of political authority and hierarchical organization. That is, our use of the term is meant to be distinct from historical uses of the label, as per early socialist or syndicalist thinkers such as, *inter alia*, Proudhon, Bakunin or Kropotkin.

Other plausible labels for these types are 'Populist' or 'anti-Establishment'. In particular, the fact that recent studies of populism (such as Algan et al. (2017)) have directly leveraged data on institutional trust provides some foundation for such a branding. However, we choose Anarchist as our label for this type because (i) it is a more generic and potentially neutral term for the concept of an opposition to existing institutional structures[12], and (ii) the types that we identify are clearly apparent from the early 1990s, thereby pre-dating the latest wave of populist politics. That is, our empirical results indicate that there may be some clear historical roots to the current populist trend, extending at least as far back as the late 1980s.

4.3 Changing Ideologies?

Given the baseline 4-type model established above our next exercise examines within and between-type shifts across the different waves of the WVS. Recall that our approach here is to estimate the 4-type topic model separately for each wave and compare the β type vectors over time.

The first point to note is that our main types are stable and repeat themselves across waves. This is evident in Table 3a which reports the correlations between the separately estimated types across waves. It is straightforward to pin down similar types across waves since the correlations are high with, for example, the nominated Liberal Centrist type showing

[10]These taxonomies, covered in texts such as Vincent (2009) and Geoghegan (2003), are centered on 'classical' ideologies (eg: liberalism, conservatism, socialism) that are often explicitly articulated as bodies of thought by key writers (Locke, Burke, Stuart Mill, Marx), as well as modern 'post-materialist' ideologies (eg: environmentalism, feminism).

[11]See Appendix Table 9 for a breakdown of the largest differences in β issue-position weights across pairs of types. This shows the points of separation between the Centrist and Anarchist types.

[12]As mentioned the term 'populist' can be considered pejorative - see the blunt critique of UK's *Daily Mail* (Murray (2016)). The term 'anti-establishment' is subject to similar concerns, with competing claims of who the elite or establishment are.

a correlation of 0.97 between waves 2 and 3 or the the Right Anarchist type reporting a correlation of of 0.94 betweens waves 2 and 5.

These high correlations also imply that there are fairly limited 'within-type' shifts over time, as measured by the probability weights in the β type vectors. Since we are using the same issue-position features across waves we can directly report the shifts in probabilities per feature. To facilitate the interpretation we have re-scaled the β vectors as described in Appendix C[13].

These probabilities can be interpreted as statistics for the 'likelihood of expression' for a given issue-position conditional on drawing on a latent type. For example, a (re-scaled) β weight of 0.72 for 'Confidence: Labor Unions' within the Liberal Centrist type indicates that an individual drawing on this type to form their issue-position will express confidence in this institution 72% of the time.

The ten largest shifts in probabilities for the Wave 2-5 difference are shown in Figure 3 for each type. The baseline numbers are also reported in Appendix Table 8 along with the changes. The most noticeable trend is an increase in socially liberal attitudes across types with, for example, the Conservative Centrist increasing their weights on issue-positions such as 'No problem neighbours: Homosexual' and 'No problem neighbours: People with AIDS'. Also notable is the the Right Anarchist type, which shows higher confidence in the police and armed forces over time, along with more intense hostility towards immigration. Some of these changes are nominally large, with 15-20% increases in liberal attitudes on gay rights for the Conservative Centrist and 50-60% increases in confidence for the armed forces and police for the Right Anarchist. However, the overall changes in the β weights have not been pervasive enough to drastically shift the between-wave correlations evident in Table 3a[14].

The between type differences can also be summarized using correlations across the β type vectors within the WVS waves and we show these in Table 3b. The increase in the intensity of socially liberal issue-positions is now most clearly seen via the increasing closeness between the Conservative Centrist type and the the two left-wing types. Between waves 2 and 5 the negative correlation with the Left Anarchist type moderates (going from -0.599 to -0.341) while the correlation with the Liberal Centrist type strengthens (from 0.446 to

[13]Since the rescaling of the β vectors is a non-linear transformation, the changes between the re-scaled β vectors and unscaled β vectors can differ.

[14]In the case of Right Anarchist attitudes towards the the police and armed forces it should be noted that this shift brings this type closer to the mean β for these issue-positions displayed by the two Centrist types.

0.547). Hence, at the type level defined by the β-vectors, we can say that there has been some convergence in the overall ideologies driven by attitudes on social issues.

4.4 Analysis of Type Shares

4.4.1 Determinants of Type Shares and Country Differences.

We start out the analysis of the θ_i individual type shares by studying the micro-level determinants. In Table 4 we run some simple regressions of the type shares on individual characteristics. Since the dependent variable is a continuous share the regression results have the interpretation of telling us how the intensity of ideological positions changes with different covariates.

We run four regressions corresponding to each type. These indicate some intuitively plausible correlations - women are more liberal and centrist, with a magnitude of 0.7 - 1.9% points, and the unemployed have higher shares in the two anarchist ideologies. Furthermore, there are clear shifts in the distribution of type shares over time. After conditioning on covariates it is evident that the Liberal Centrist share increases by around 5.7% points after Wave 2 with the Conservative Centrist share falling by a similar amount. Following a similar pattern, the Left Anarchist share rises in Waves 4 and 5 while the Right Anarchist share falls.

Hence, across the sample of countries the net result is a growth in the share of the two left-wing types (ie: Liberal Centrist and Left Anarchist). However, there are also significant country-level factors evident from the individual level analysis. The country fixed effects in Table 4 account for 50-75% of the explained variation and we plot the country-level means by type in Figure 4. This again shows some intuitively plausible relationships - northern European countries (eg: Denmark, Finland, Netherlands) are more liberal while countries with strong religious traditions (Malta, Ireland, Portugal) are more conservative.

To summarize the changes across countries over time we implement some splits along different ideological dimensions. Firstly, in Figure 5a we examine the left-right distinction and pool the type shares for the left-wing Liberal Centrist and Left Anarchist types. The plot of changes in these pooled type shares between Waves 2 (1989-1993) and 5 (2005-2009) shows that most countries have moved left ideologically, with a mean shift of 9% points. In Figure 5b we then plot the changes for the pooled Left and Right Anarchist types. This provides an indicator of the overall strength of anti-establishment ideological sentiment across countries. The results show a large increase in the Anarchist type shares for the US (around 18% points), with significant increases also evident for the Anglo-Celtic domains (Great Britain, Northern Ireland) and the Netherlands. In net terms however, the anti-establishment trend is more muted across countries, with a mean increase of only 1.5%.

In Figure 6, we further probe the sharp increase in the Anarchist ideologies for the US. The clearest development is the growth in the US Right Anarchist share, which increases from a 24.2% share in Wave 2 (1989-1993) to 35.5% in Wave 5 (2005-2009). Note here though that this increase actually took place as a single step change between Waves 4 (1999-2004) and 5 (2005-2009). By comparison, the rise of the US Left Anarchist share from 7.4% to 14.5% was more gradual across the waves.

Overall, these country-level findings are generally consistent with other international studies of shifts in political attitudes (Inglehart (1997); Inglehart et al. (2010)). Taken together with the within-type analysis, the basic message on ideological change that follows from our methodology so far is one of a stable structure of ideologies over 20 years and some increase in social liberalism. This increase in social liberalism has occurred both on the intensive margin (ie: the weights on liberal issue-positions in the β vectors) as well as the extensive margin (the growing individual-level type shares for the Liberal Centrist and Left Anarchist types). The other major development in the data so far is the strong tilt towards anti-establishment Anarchist ideologies in some countries - particularly the US.

4.4.2 'Citizen Slant' - Within-Person Concentration

Our analysis so far has focused on changes at the level of the ideological types as well as the total shares in the types across the sample. However, for the analysis of the polarization the loadings of individuals on the four types is of key importance. In particular, the 'mixed membership' structure of our approach means that two countries with the same overall type distribution can have completely different individual type compositions.

For example, imagine a country which has an overall 50% share in Type 1 and 50% share in a second Type 2. This country can either consist of completely identical individuals with 50% shares in the two types or it could consist of half the population holding a 100% share in Type 1 and another half with a 100% share in Type 2. These two possible type compositions obviously have very different implications for the understanding of societal polarization. A country with two separate sets of 'pure' homogeneous types is plausibly more vulnerable to political conflict than a country where there is more ideological heterogeneity at the individual-level.

We therefore develop a measure of how strongly an individual is loading on one of the four ideological types by constructing a Gini-style measure of within-person concentration or 'slant'. We define the within-person concentration G_i of individual i as:

$$G_i = \frac{\sum_{t=1}^{4} \sum_{s \neq t} |\theta_i^t - \theta_i^s|}{2(n-1) \sum_{t=1}^{4} \theta_i^t} \tag{5}$$

where θ_i^t and θ_i^s are the types shares of individual i. Intuitively, this approach is aggregating the absolute pairwise difference in ideological shares that exist at the individual level. The measure of within-person concentration G_i is monotonically increasing the more an individual loads on one of the ideological types. If a person has a 100% share in one type then G_i will be equal to 1, while $G_i = 0$ implies shares of 25% in all types.

In this way this our measure is able to capture how segregated type shares are on a within-individual basis. Furthermore, it allows us to analyze which groups exhibit a particularly high tendency towards within-person concentration. We plot the (Wave 5) country means of the G_i measure in Figure 7 along with the changes between Wave 2 and 5. This shows that G_i is relatively high across the sample with a mean of around 0.75. However, between-country differences are not dramatic. There is only an 8% gap between the the most and least concentrated countries and the ordering does not suggest that any particular ideological types are driving concentration. That is, amongst the most concentrated or 'slanted' countries we see cases of both relatively Conservative and Liberal countries defined in terms of the mean type shares seen earlier. The major, positive country level shifts in slant over the waves occurred in Denmark, Finland and the US (Figure 6a) but the changes were muted for most countries.

To study the importance of individual characteristics on within-person concentration (as well as the development of G_i over time) we estimate the following regression equation:

$$\ln(G_{icw}) = X_{icw}\delta + \tau_w + \mu_c + \epsilon_{icw} \tag{6}$$

where X_{icw} is a vector of covariates, τ_w are wave dummies, μ_c are the country dummies and $\ln(G_{icw})$ is the natural logarithm of G_i. The log-specification makes it easier to analyze the magnitude of the coefficients. The results are reported in Table 5 with controls for the type shares and with the Liberal Centrist set as the baseline type. The purpose of the type share controls is to allow us to study whether G_i concentration is increasing with shares of certain types. The results indicate that the Left Anarchist is the least concentrated type followed by the Right Anarchist. In turn, this means that the individuals with larger shares in either of the two anarchist ideologies are more likely to mix different ideological types together than the centrist types.

After controlling for the available individual characteristics we find an 2.7% increase in G in wave 4 and an 1.6% increase in wave 5[15]. The results for the analysis of the US are similar overall except that the increases of G concentration in Waves 4 and 5 are far larger, standing at around 4% (Column 3).

To further probe the increases in G over time we estimate eq. (6) separately for individuals conditional on their main type and also broken down according to the US and non-US samples.

The results in Table 6 show that the increases in G within the US are predominantly driven by the two Anarchist types, both of which exhibit increases in concentration above 14%. To clarify, note that the message from the earlier table was that the Anarchist types are less concentrated in the cross-section (hence the positive coefficients on these type variables in the associated regressions). In contrast, the regressions in Table 6 track how concentration developed over time on a per type basis. The simple story then is that, where they are held, Anarchist views are becoming more concentrated and we are seeing 'purer' cases of Anarchist views being held across the population.

In effect, this evidence implies that the Anarchist types have become an even more dominant ideology for people who had already shown a lack of trust in social and political institutions. While in earlier waves this section of the population might still have shown large type shares in Centrist ideologies this potentially moderating centrist influence became less potent in more recent years. The findings for the US also contrast fairly strongly with the results for the non-US sample where the increase in concentration for the Anarchist types is more muted (at around 2-4%) and in any case is matched by increases for the Liberal Centrist type as well (see Table 6, panel (b), first column).

4.5 Societal Polarization

While the above measure of within-person concentration describes the strength of the individual loadings on the four ideological types, it does not necessarily speak to the strength of the division between citizens inside a society. A society in which there are sub-groups of

[15] We suppress the reporting of the individual attribute coefficients in eq. (6) to avoid clutter. The basic result for these covariates is that only gender (female) and unemployment contribute significantly to within-person concentration but with small magnitudes. They enter with positive and negative signs respectively.

individuals that heavily load on one ideological type may not necessarily be dramatically polarized. The extent of polarization would hinge on how big these 'purist' sub-groups are relative to the full set of ideological sub-groups across the population. As an example, the country-level type share plots in Figure 4 indicate that some countries are characterised by widely represented types with aggregate type shares around the 50% mark, such as Liberal Centrists in Denmark and Conservative Centrists in Malta. At face value, these countries could be plausibly classified as 'unipolar' and less vulnerable to group conflict no matter how concentrated the different types are in terms of citizen slant.

We therefore study polarization by adapting the measures proposed by Esteban and Ray (1994) and Duclos et al. (2004) to our setting with 4 ideological types. These measures have the novel feature of accommodating two forces that define polarization as a general concept. Firstly, there is *identification* which occurs amongst individuals with a common characteristic and is an increasing function of the total number of common individuals (that is, group size). Secondly, there is *alienation* which accounts for the social detachment that individuals feel towards others who do not share some common characteristic. Again, the strength of the alienation effect will depend on (relative) group size as well as the 'distance' between groups in the key characteristic of concern.

Using the example of income inequality, Esteban and Ray (1994) prove that any measure of polarization P that accurately accounts for own-group identification as well as alienation in relation to an out-group and fulfills 3 'reasonable' axioms has to be of the form[16]:

$$P(\pi, y) = K \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_i^{1+\alpha} \pi_j |y_i - y_j| \tag{7}$$

where π are the number of people in the groups, y is amount of income of each group, K is a normalizing constant and α is the polarization sensitivity, which measure how strongly the polarization measure reacts to the group sizes.

This general polarization measure P was constructed for a one dimensional variable y, for example, income. Polarization in our case obviously has to be measured over all 4 ideological types. Practically, we divide people into meaningful ideological groups based on their dominant type share. That is, we allocate individuals to one of our 4 groups based on their highest type share at the i-level. We then add up the θ_i type shares amongst the defined group members to get the mean type share, defined as $\tilde{\theta}_t$. This differs from the mean

[16]The Axioms put forward in Esteban and Ray (1994) are explained in more detail in Appendix G.

type shares $\bar{\theta}_t$ we have presented earlier on the basis that we are only taking the mean over individuals with the same dominant type rather over the whole population.

Given this modification the polarization measure is defined as:

$$P(\pi, \theta) = K \sum_{t=1}^{4} \sum_{j=1}^{4} \pi_t^{1+\alpha} \pi_j (|\tilde{\theta}_{t1} - \tilde{\theta}_{j_1}| + |\tilde{\theta}_{t2} - \tilde{\theta}_{j_2}| + |\tilde{\theta}_{t3} - \tilde{\theta}_{j_3}| + |\tilde{\theta}_{t4} - \tilde{\theta}_{j_4}|) \tag{8}$$

where π_t and π_j are the number of people who have the dominant type share t and j. The means of the type shares in the each of the 4 dominant type groups are $\tilde{\theta}_t$ for own type and $\tilde{\theta}_j$ for a generic other type. The second subscript on $\tilde{\theta}_t$ and $\tilde{\theta}_j$ represents the dominant type group we are conditioning on when calculating absolute distance between groups.

As an example, consider setting type t as the Liberal Centrist and j is the Conservative Centrist. We index the Liberal Centrist as the type 1 in the second conditioning subscript. The calculation $|\tilde{\theta}_{t1} - \tilde{\theta}_{j_1}|$ then represents the (absolute) difference between the mean Liberal Centrist type share for dominant Liberal Centrist individuals and the mean Conservative Centrist type share for dominant Conservative Centrist individuals. This can be interpreted as a measure of how close different ideological groups are despite their contrasting dominant type shares. That is, a Liberal Centrist and a Conservative Centrist are more likely to 'get along' if they have high minority type shares in each other's ideology.

The other components of $P(\pi, \theta)$ are the polarization sensitivity parameter α, which we fix at $\alpha = 0.5$, and the constant $K = (\sum_{t=1}^{4} \pi_t)^{-(2+\alpha)}$ that serves to normalize the polarization measure by population size. We provide more detail and show how P varies with different values of α in Appendix G.

Intuitively, the polarization measure will be largest for the case where there are two major type share groups of identical size who exhibit completely different type shares. An example of this would be a bipolar Liberal Centrist and Right Anarchist society where each type group had very small minority shares in the other type. This provides a natural link back to our earlier measure of citizen slant. Since an increase in citizen slant implies an increase in the means for $\tilde{\theta}_t$ and $\tilde{\theta}_j$, absolute differences in type shares between the groups increase and polarization rises due to stronger alienation effects.

It is also useful to note how polarization also depends on the relative sizes of the groups within an population, as measured by π_t and π_j. For example, given the same between group differences in types, a country in which 2 groups each make up 50% of the population will be more polarized than a country with 4 groups each making up 25% of the population.

We calculate the polarization measure separately for each country and wave in our sample. The ranking of the countries based on their polarization in each wave is shown in Figure 8. The ranking of countries according to Wave 5 polarization levels is distinct from the earlier ranking for citizen slant. Denmark, which has the lowest level of polarization, provides an instructive example of how the P polarization measure combines information. The inputs into the result for Denmark are its high Liberal Centrist type share (approximately 0.6 - see Figure 4) and high level of within-person concentration or slant (which intensifies over time - see Figure 7). Hence the low Danish P measure reflects a case of ideological consensus where there is a major plural type (Liberal Centrist) that is strongly held by individuals (as manifested in high slant).

The US, which sits at the top of the polarization ranking in Wave 5, provides a sharp comparison that again illustrates the mechanics of the P measure. It has a relatively even spread of type shares, with shares of around 30% for the Liberal Centrist, Conservative Centrist and Right Anarchist types. Hence, the group size effect picked up by the π_t and π_j terms is stronger in the US compared to unipolar cases such as Denmark. Overall, the increase in polarization in the US is mainly driven by the rise in slant over time (Figure 4), which contributes by intensifying the alienation effect. The changes in dominant type composition in the US only have a minor influence on the polarization measure.

However, it should be noted that, across countries, the changes in polarization over time are not dramatic, with most of the shifts occurring in the 5-10% range relative to baseline values in Wave 2. A key point to note is that the defining feature of some of the polarization episodes seen in the data is the qualitative content of developments. The US is the banner example here since the increase in polarization was driven by an increase in the presence of Anarchist types. Hence the US experience with polarization has the extra qualitative dimension of also reflecting the trend of a 'disappearing center' which is likely to have additional consequences for social cohesion over and above the increase in P that we measure statistically.

5 Conclusion

In this paper, we have proposed a new way to identify the latent ideologies of individuals from survey data. Our approach does not presuppose any ideological structure for individuals. Nonetheless, we are able to identify interpretable, consistent and stable ideological types in the data. The findings generally align with the left-right framing frequently used in the social sciences but we also identify anti-establishment 'anarchist' ideologies that are characterized by their distrust in important societal institutions.

The framework we develop allows us to put forward rich new measures of citizen polarization based on multiple ideologies. The measurements we take from the data indicate that changes in polarization across countries are muted, falling in the 5-10% range. A comparison with earlier findings on elite polarization is useful here. Even in the US - which experienced the largest increase in citizen polarization - the shifts that we pick up are not as dramatic as those found in noted studies of elite polarization such as Gentzkow et al. (2016) and Poole and Rosenthal (1985). That said, the central qualitative feature of the US experience - a 'disappearing center' and a rise of anarchist, anti-establishment types plausibly reflects a deeper threat to social cohesion. For example, polarization driven by the growth of anarchist types is arguably a greater threat to political stability than a statistically equal polarization underpinned by strong growth in the Liberal Centrist and Conservative Centrist types.

Overall, our findings shed light on recent trends in polarization in the advanced democracies. The framework we provide is rich enough to account for multiple developments that have previously been hard to reconcile with each other. In particular, these developments include the growing liberalism of advanced democracies on social issues which we can track on the basis of of an intensive margin (ie: the within-type shifts in β weights that take place across all types) as well as the extensive margin (the increase in θ type shares for liberal types). A second development that we are able to track is the growth of both the left and right wing Anarchist types who represent a natural constituency for populist or anti-establishment movements. Our framework is able to capture the shared roots of these types in their low confidence in institutions as well as measure the hardening of these attitudes via our index of citizen slant. Hence, the enriched ideological spectrum we uncover formally takes into account the anti-establishment leanings of citizens and has the potential to better explain divides across political issues and the 'interesting times' we currently find ourselves in.

References

D. Acemoglu, G. Egorov, and K. Sonin. A political theory of populism. *The Quarterly Journal of Economics*, 128(2):771–805, 2013.

C. H. Achen and L. M. Bartels. Ignorance and bliss in democratic politics: Party competition with uninformed voters. *Prepared for presentation at the Annual Meeting of the Midwest Political Science Association, Chicago*, 2002.

E. M. Airoldi, D. Blei, E. A. Erosheva, and S. E. Fienberg. *Handbook of mixed membership models and their applications*. CRC Press, 2014.

A. Alesina, E. L. Glaeser, and B. Sacerdote. Why doesn't the united states have a european-style welfare state? *Brookings Papers on Economic Activity*, 2001(2):187–277, 2001. ISSN 1533-4465.

A. Alesina, A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg. Fractionalization. *Journal of Economic growth*, 8(2):155–194, 2003.

A. Alesina, P. Giuliano, and N. Nunn. On the origins of gender roles: Women and the plough. *Quarterly Journal of Economics*, 128(2):469–530, 2013. ISSN 0033-5533.

N. Aletras and M. Stevenson. Evaluating topic coherence using distributional semantics. *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013)*, 2013.

N. Aletras, T. Baldwin, J. H. Lau, and M. Stevenson. Representing topics labels for exploring digital libraries. *Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries,IEEE Press*, pages 239–248, 2014.

Y. Algan, S. Guriev, E. Papaioannou, and E. Passari. The european trust crisis and the rise of populism. *CEPR Discussion Papers*, 2017.

S. Ansolabehere, J. Rodden, and J. M. Snyder. Purple america. *Journal of Economic Perspectives*, 20(2):97–118, 2006. ISSN 0895-3309.

E. Ash. The political economy of tax laws in the us states. Technical report, Working Paper, Columbia University, 2015.

V. Ashok, I. Kuziemko, and E. Washington. Support for redistribution in an age of rising inequality: New stylized facts and some tentative explanations. *National Bureau of Economic Research*, 2015.

O. Bandiera, S. Hansen, A. Prat, and R. Sadun. Ceo behavior and firm performance. *National Bureau of Economic Research*, 2017.

S. O. Becker, T. Fetzer, and D. Novy. Who voted for brexit? a comprehensive district-level analysis. *Economic Policy*, 32(92):601–650, 2017.

D. G. Blanchflower and A. J. Oswald. Is well-being u-shaped over the life cycle? *Social science & medicine (1982)*, 66(8):1733–1749, 2008. ISSN 0277-9536.

L. Blaydes and J. Grimmer. Political cultures: Exploring the long-run determinants of values transmission. In *Annual Meeting of the International Political Economy Society, Claremont Graduate School, Claremont, CA*, 2013.

D. M. Blei and J. D. Lafferty. Dynamic topic models. *Proceedings of the 23rd international conference on Machine learning*, pages 113–120, 2006.

D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.

W. Bossert, C. D'Ambrosio, and E. La Ferrara. A generalized index of fractionalization. *Economica*, 78(312):723–750, 2011.

P. Buisseret and R. Van Weelden. Crashing the party? elites, outsiders, and elections. In *Wallis conference*, 2017.

L. Bursztyn, G. Egorov, and S. Fiorin. From extreme to mainstream: How social norms unravel.

F. Caselli and W. J. Coleman. On the theory of ethnic conflict. *Journal of the European Economic Association*, 11(suppl. 1):161–192, 2013.

J. Chang, J. Boyd-Graber, D. M. Blei, C. Wang, and S. M. Gerrish. Reading tea leaves: How humans interpret topic models. *Advances in neural information processing systems*, pages 288–296, 2009.

Y. Che, Y. Lu, J. R. Pierce, P. K. Schott, and Z. Tao. Does trade liberalization with china influence us elections? Technical report, National Bureau of Economic Research, 2016.

E. Dal Bo, F. Finan, O. Folke, J. Rickne, and T. Persson. Economic losers and political winners: The rise of the radical right in sweden. Technical report, mimeo Uppsala University, 2017.

E. Dal Bo, P. Dal Bo, and E. Eyster. The demand for bad policy when voters underappreciate equilibrium effects. *The Review of Economic Studies*, 85(2):964–998, 2018.

David Newman, Jey Han Lau, Karl Grieser, and Timothy Baldwin. Automatic evaluation of topic coherence. *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL,*, pages 100–108, 2010.

C. Ding and X. He. K-means clustering via principal component analysis. In *Proceedings of the twenty-first international conference on Machine learning*, page 29. ACM, 2004.

C. Dippel, R. Gold, and S. Heblich. Globalization and its (dis-) content: Trade shocks and voting behavior. Technical report, National Bureau of Economic Research, 2015.

D. Dorn, G. Hanson, K. Majlesi, et al. Importing political polarization? the electoral consequences of rising trade exposure. Technical report, National Bureau of Economic Research, 2016.

A. Downs. An economic theory of political action in a democracy. *Journal of Political Economy*, 65(2):135–150, 1957. ISSN 00223808.

J.-Y. Duclos, J. Esteban, and D. Ray. Polarization: Concepts, measurement, estimation. *Econometrica*, 72(6):1737–1772, 2004. ISSN 00129682.

J.-M. Esteban and D. Ray. On the measurement of polarization. *Econometrica*, pages 819–851, 1994. ISSN 00129682.

M. P. Fiorina and S. J. Abrams. Political polarization in the american public. *Annual Review of Political Science*, 11:563–588, 2008. ISSN 1094-2939.

M. Gentzkow, J. M. Shapiro, and M. Taddy. Measuring polarization in high-dimensional data: Method and application to congressional speech. *National Bureau of Economic Research*, 2016.

M. Gentzkow, B. T. Kelly, and M. Taddy. Text as data. *The Journal of Economic Literature (forthcoming)*, 2017.

V. Geoghegan. *Political ideologies: An introduction.* Routledge, 2003.

E. L. Glaeser and B. A. Ward. Myths and realities of american political geography. *Journal of Economic Perspectives*, 20(2):119–144, 2006. ISSN 0895-3309.

J. Grimmer. A bayesian hierarchical topic model for political texts: Measuring expressed agendas in senate press releases. *Political Analysis*, 18(1):1–35, 2009.

J. H. Gross and D. Manrique-Vallier. A mixed-membership approach to the assessment of political ideology from survey responses. In *Individual Presentation, Society for Political Methodology, 29th Annual Summer Meeting, Chapel Hill, NC.* Citeseer, 2012.

L. Guiso, H. Herrera, M. Morelli, and T. Sonno. Demand and supply of populism. *IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University*, 2017.

S. Hansen, M. McMahon, and A. Prat. Transparency and deliberation within the fomc: a computational linguistics approach. *The Quarterly Journal of Economics*, 2014.

A. Healy and N. Malhotra. Retrospective voting reconsidered. *Annual Review of Political Science*, 16:285–306, 2013. ISSN 1094-2939.

M. Hoffman, F. R. Bach, and D. M. Blei. Online learning for latent dirichlet allocation. In *advances in neural information processing systems*, pages 856–864, 2010.

M. D. Hoffman, D. M. Blei, C. Wang, and J. Paisley. Stochastic variational inference. *the Journal of machine Learning research*, 14(1):1303–1347, 2013. ISSN 1532-4435.

R. Inglehart. *Modernization and postmodernization: Cultural, economic, and political change in 43 societies.* Princeton University Press, 1997.

R. F. Inglehart, M. Basanez, and A. Moreno. *Human Values and Beliefs: A Cross-Cultural Sourcebook.* University of Michigan Press, Ann Arbor, 2010. ISBN 978-0-472-10833-6.

Z. Jelveh, B. Kogut, and S. Naidu. Political language in economics. *Working Paper*, 2015.

J. Jensen, S. Naidu, E. Kaplan, L. Wilse-Samson, D. Gergen, M. Zuckerman, and A. Spirling. Political polarization and the dynamics of political language: Evidence from 130 years of partisan speech [with comments and discussion]. *Brookings Papers on Economic Activity,*, Fall:1–81, 2012.

D. Kim and A. Oh. Topic chains for understanding a news corpus. *International Conference on Intelligent Text Processing and Computational Linguistics, Springer*, pages 163–176, 2011.

J. H. Lau and T. Baldwin. The sensitivity of topic coherence evaluation to topic cardinality. *Proceedings of NAACL-HLT*, pages 483–487, 2016.

J. H. Lau, K. Grieser, D. Newman, and T. Baldwin. Automatic labelling of topic models. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1, Association for Computational Linguistics*, pages 1536–1545, 2011.

J. H. Lau, D. Newman, and T. Baldwin. Machine reading tea leaves: Automatically evaluating topic coherence and topic model quality. *Proceedings ofthe 14th Conference ofthe European Chapter ofthe Association for Computational Linguistic*, pages 530–539, 2014.

A. Leigh. Does the world economy swing national elections? *Oxford Bulletin of Economics and Statistics*, 71(2):163–181, 2009. ISSN 1468-0084.

J. G. Montalvo and M. Reynal-Querol. Ethnic polarization, potential conflict, and civil wars. *American economic review*, 95(3):796–816, 2005.

D. Murray. Populism: It's the bbc's new buzzword, being used to sneer at the 'uneducated' 17 million who voted for brexit. *Daily Mail*, 16:December, 2016.

P. Norris. *Democratic deficit: Critical citizens revisited.* Cambridge University Press, Cambridge [etc.], 2016. ISBN 978-0521127448.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, and V. Dubourg. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12(Oct):2825–2830, 2011.

T. Piketty. Brahmin left versus merchant right: Rising inequality and the changing structure of political conflict. *WID Working Paper Series*, 2018.

K. T. Poole and H. Rosenthal. A spatial model for legislative roll call analysis. *American Journal of Political Science*, pages 357–384, 1985. ISSN 1540-5907.

D. Putthividhy, H. T. Attias, and S. S. Nagarajan. Topic regression multi-modal latent dirichlet allocation for image annotation. *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 3408–3415, 2010.

D. Rodrik. Populism and the economics of globalization. Technical report, National Bureau of Economic Research, 2017.

A. Vincent. *Modern political ideologies*. John Wiley & Sons, 2009.

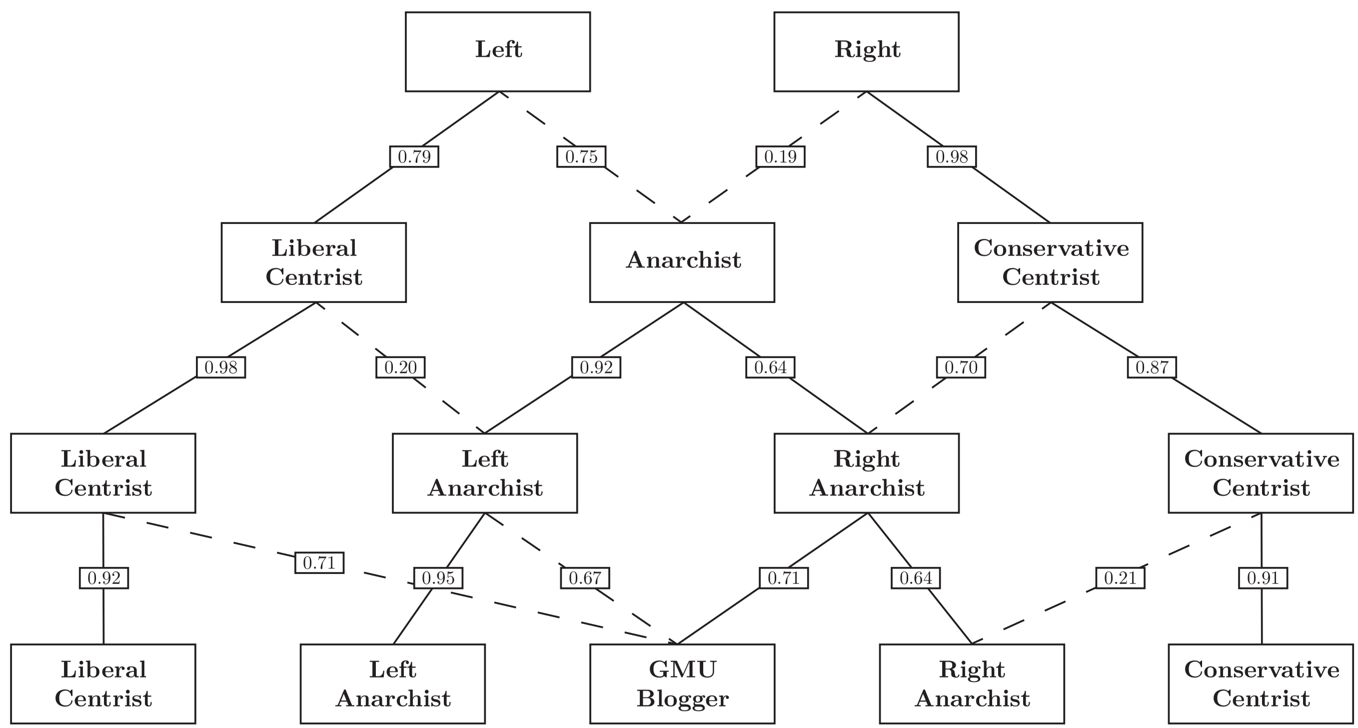
C. Wang, D. M. Blei, and D. Heckerman. Continuous time dynamic topic models. *Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence*, pages 579–586, 2008.

Y. S. Wang, R. L. Matsueda, E. A. Erosheva, et al. A variational em method for mixed membership models with multivariate rank data: An analysis of public policy preferences. *The Annals of Applied Statistics*, 11(3):1452–1480, 2017.

J. Wolfers. Are voters rational? evidence from gubernatorial elections. *Work. pap., The Wharton School*, 2002.

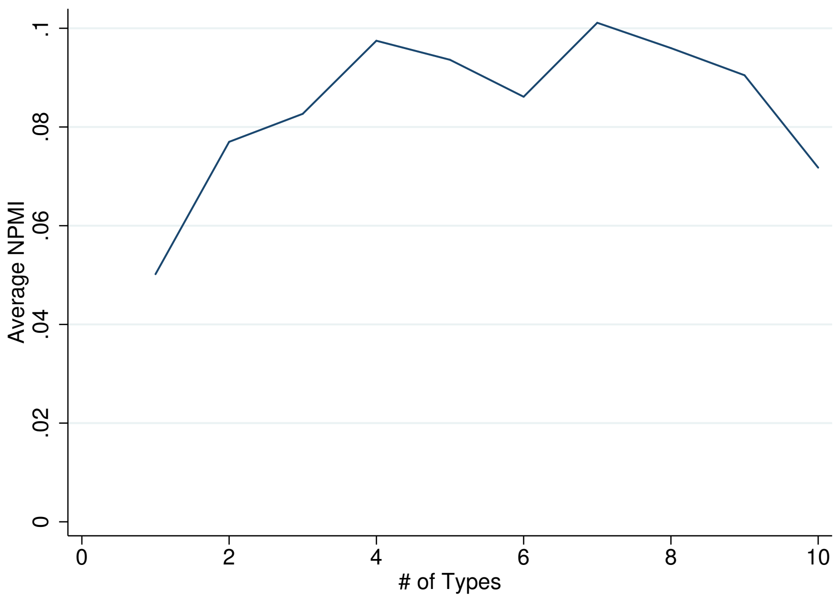
6 Figures

Figure 1: Hierarchy of Types as created by LDA



Notes: This graphic shows the hierarchy of types as created by Latent Dirichlet Allocation (LDA) as per the different orders of topic model reported in Table 2.The values reported amongst the lines connecting the boxes record the correlation in the β issue-position probability vectors across types.

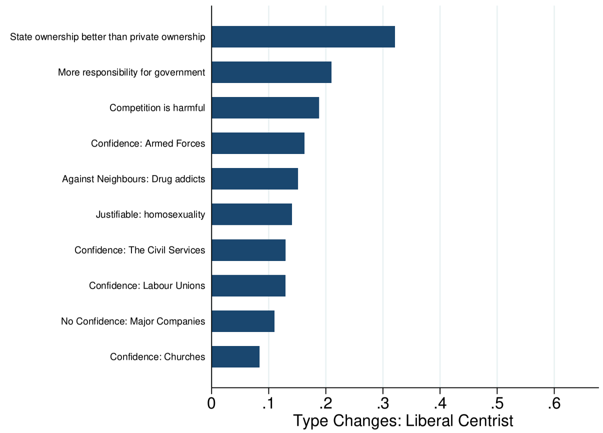
Figure 2: Average Cohesion of Ideological Types for Different LDA Models



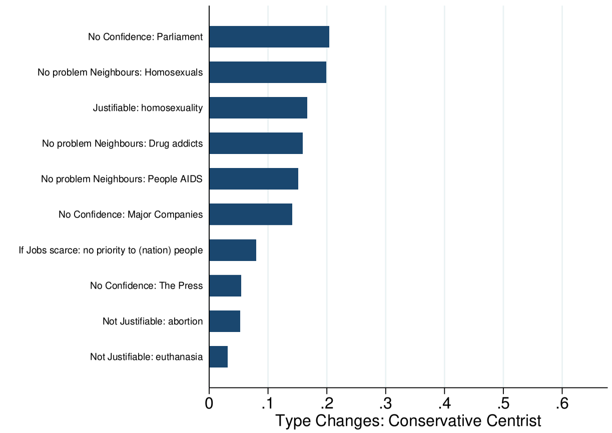
Notes: This figure show the topic cohesion scores calculated for models with $M \in \{1-10\}$ types for the 5th wave. The topic cohesion is calculated for different number of features $B \in (5, 10, 15, 20)$. Afterwards the average over the different values of B is taken.

Figure 3: Within-Type Changes in Issue-Position Weights (Wave2 to 5)

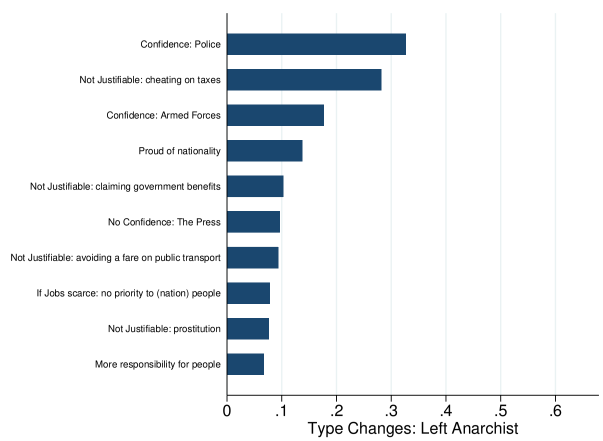
(a) Centrist Liberal



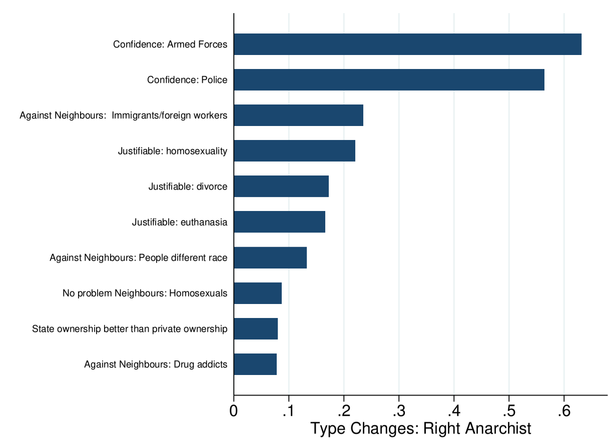
(b) Centrist Conservative



(c) Left Anarchist



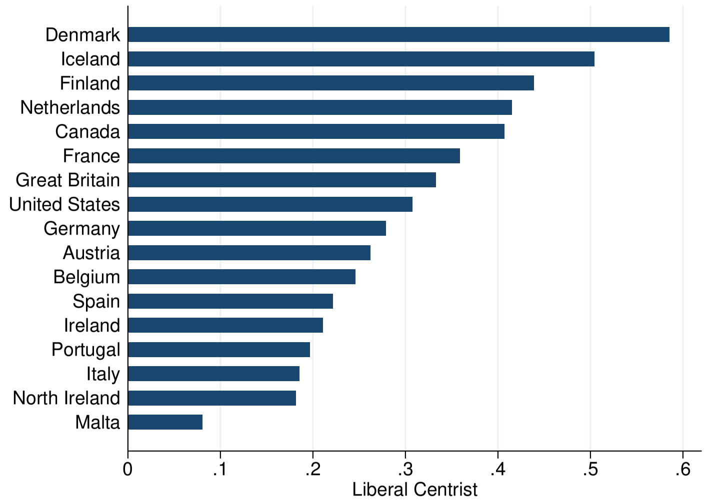
(d) Right Anarchist



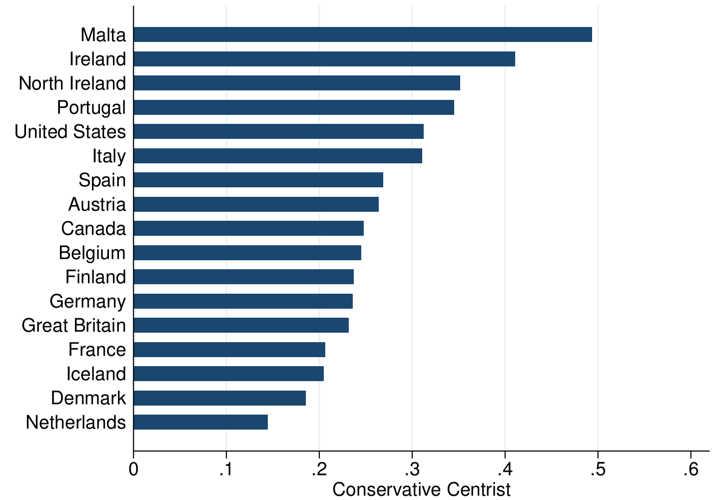
Notes: This figure reports the largest changes in the β issue position weights per ideological type from Waves 2 (1989-1993) to Wave 5 (2005-2009). We report the top ten changes per type amongst the 58 possible features. The scale is set to 0-0.60 to facilitate direct comparisons across types.

Figure 4: Country-Level Type Shares

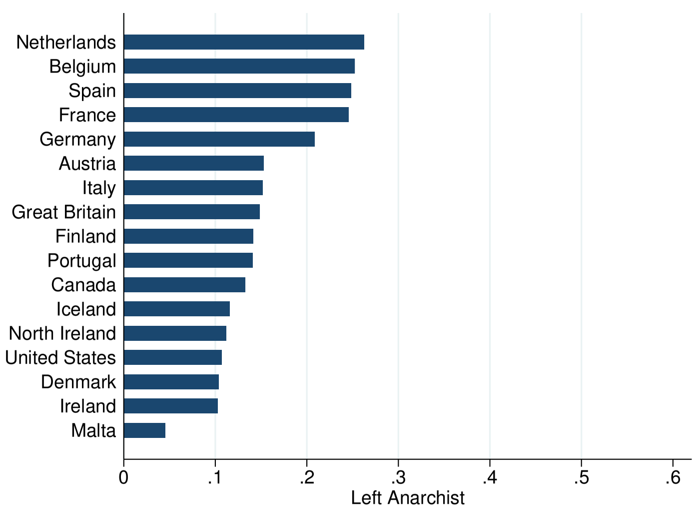
(a) Liberal Centrist



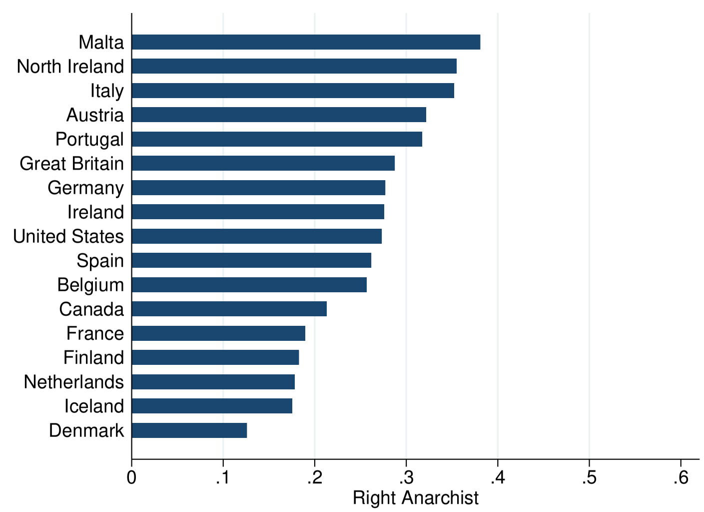
(b) Conservative Centrist



(c) Left Anarchist



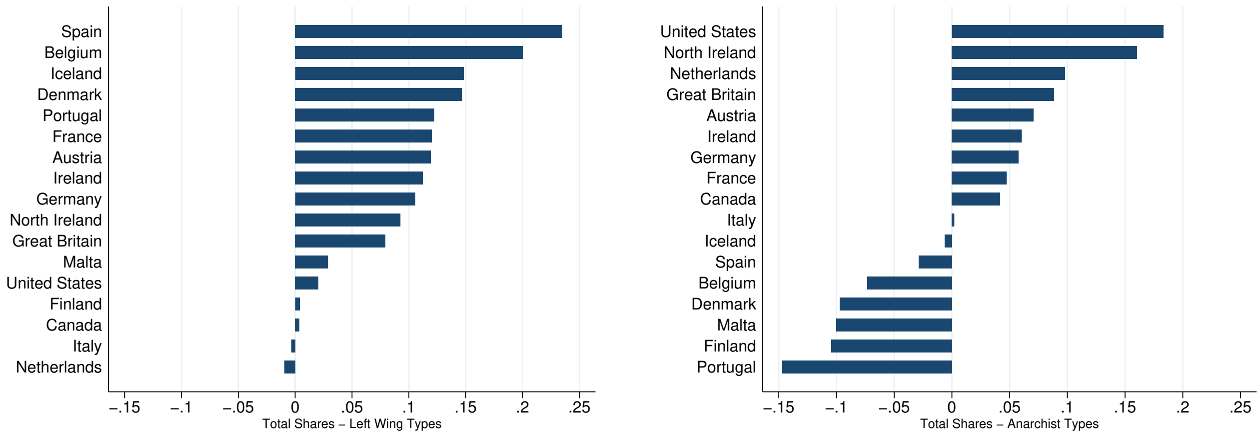
(d) Right Anarchist



Notes: This figure shows the mean country-level θ type shares aggregated over individuals. Country means calculated using WVS sample weights.

Figure 5: Changes of Types over Time

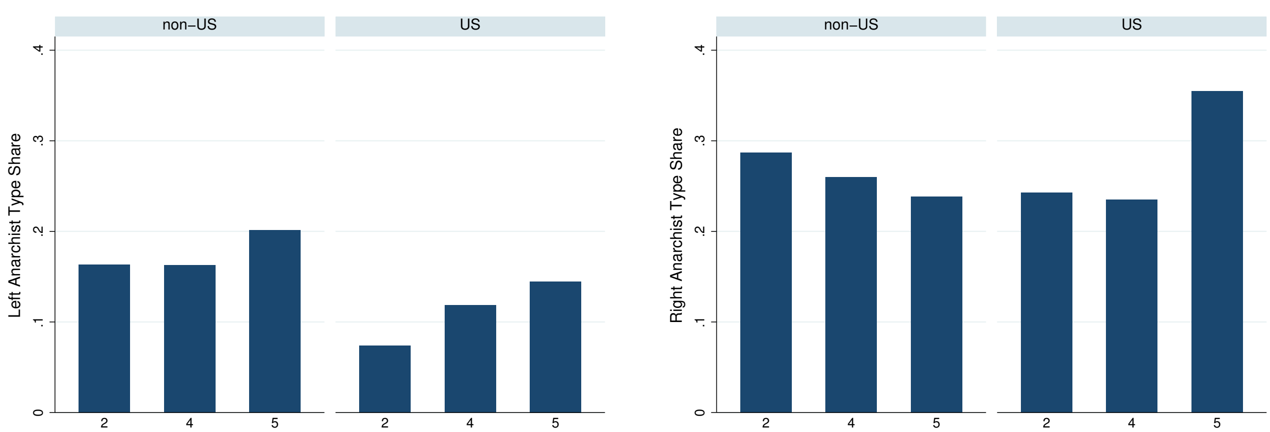
(a) Change in Left-Wing Types **(b) Change in Anarchist Share**



Notes: This figure shows the change in θ type shares by country between Waves 2 (1989-1993) and 5 (2005-2009) in the WVS. In 4(a) we pool the type shares for the Liberal Centrist and Left Anarchist types. In figure 4(b) we show the pooled change in the Left Anarchist and Right Anarchist types.

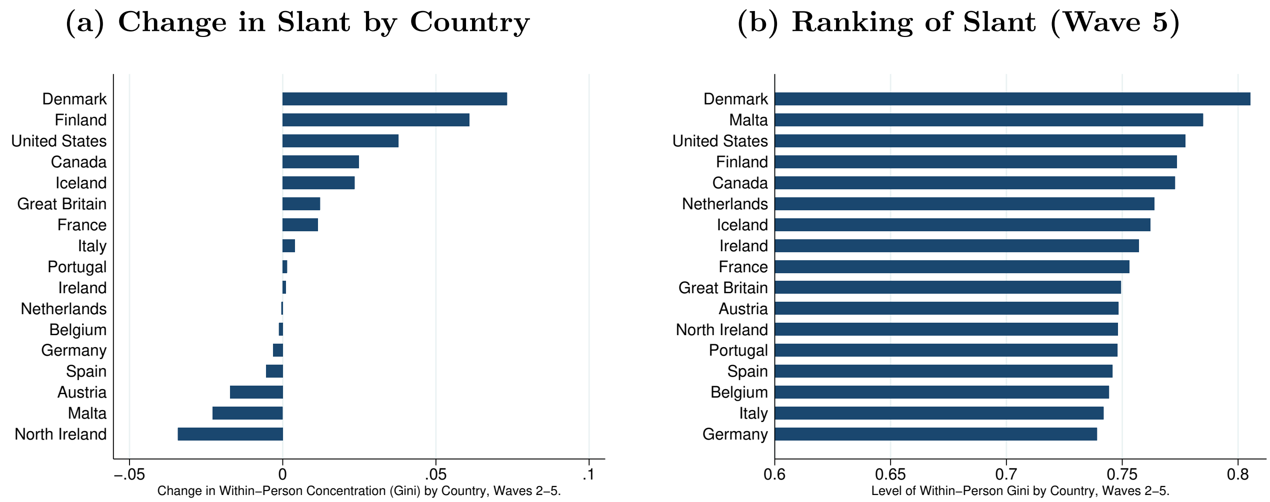
Figure 6: Type Shares - US vs non-US

(a) Left Anarchist Types **(b) Right Anarchist Types**



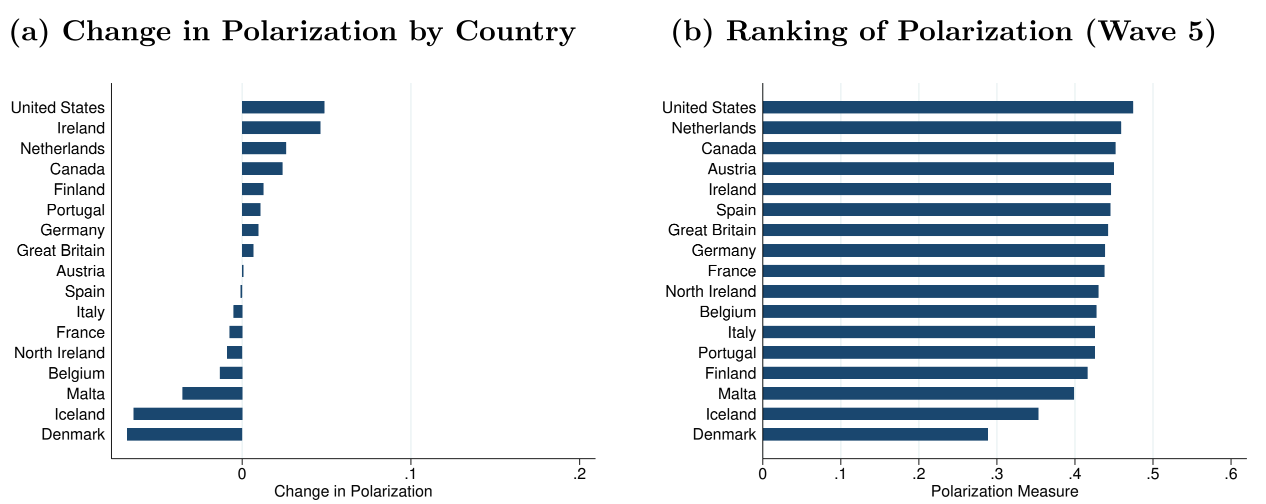
Notes: This figure compares the levels of θ type shares across waves for the Left Anarchist and Right Anarchist types. We pool all 16 non-US countries (effectively all Western European countries apart from Iceland and Canada) and contrast them to the US. The pooling for the non-US sample is based on WVS sample weights. The timing of the waves is Wave 2 (1989-1993), Wave 4 (1999-2004) and Wave 5 (2005-2009).

Figure 7: Citizen Slant by Country



Notes: Panel (a) shows the shows the change in our G_i Gini within-person ideological concentration measure ('citizen slant') from Waves 2 (1989-1993) to 5 (2005-2009) by country. Panel (b) shows the level of the within-person Gini measure by country in Wave 5. Country means calculated using WVS sample weights.

Figure 8: Polarization by Country



Notes: Panel (a) shows the change in country-level polarization measures from Waves 2 (1989-1993) to 5 (2005-2009) calculated following Esteban and Ray (1994). Panel (b) shows the level of the country-level polarization measure in Wave 5.

7 Tables

Table 1: Summary Statistics, WVS Questions

(a) (Part 1/2)

Code	Question	Scale	Share For	Share Against
	On this list are various groups of people. Could you please mention any that you would not like to have as neighbors?	{0,1}		
A124_02	People of a different race		0.097	0.903
A124_06	Immigrants / foreign workers		0.123	0.877
A124_07	People who have AIDS		0.208	0.782
A124_08	Drug addicts		0.638	0.362
A124_09	Homosexuals		0.217	0.781
C002	Do you agree, disagree or neither agree nor disagree with the following statements?: “When jobs are scarce, employers should give priority to people of this country over immigrants.”	{1–3}	0.600	0.305
E036	Rate your view on a 1 to 10 scale between the positions: “Private ownership of business and industry should be increased” vs. “Government ownership of business and industry should be increased”	{1–10}	0.745	0.255
E037	Rate your view on a 1 to 10 scale between the positions: “Government should take more responsibility to ensure that everyone is provided for” vs. “People should take more responsibility to provide for themselves”	{1–10}	0.376	0.469
E039	Rate your view on a 1 to 10 scale between the positions: “Competition is good. It stimulates people to work hard and develop new ideas” vs. “Competition is harmful. It brings out the worst in people”	{1–10}	0.785	0.215

(b) (Part 2/2)

Code	Question	Scale	Share For	Share Against
	Could you tell me how much confidence you have in these organizations:	{1–4}		
E069_01	Church		0.519	0.481
E069_02	Armed Forces		0.567	0.433
E069_04	The Press		0.356	0.644
E069_05	Labour Unions		0.385	0.615
E069_06	The Police		0.704	0.296
E069_07	Parliament		0.413	0.587
E069_08	The Civil Services		0.451	0.549
E069_13	Major Companies		0.432	0.568
E069_17	Justice System / Courts		0.533	0.466
	Please tell me for each of the following actions whether you think it can always be justified, never be justified, or something in between:	{1–10}		
F114	Claiming government benefits		0.076	0.869
F115	Avoiding a fare on public transport		0.086	0.842
F116	Cheating on taxes		0.106	0.828
F117	Someone accepting a bribe		0.035	0.931
F118	Homosexuality		0.407	0.432
F119	Prostitution		0.196	0.663
F120	Abortion		0.348	0.458
F121	Divorce		0.496	0.280
F122	Euthanasia		0.418	0.430
F123	Suicide		0.149	0.730
G006	How proud are you of your nationality?	{1–4}	0.885	0.115

Notes: This table reports summary statistics for the recoded questions from the WVS. The third column reports the original coding of the question in the WVS. Questions with a binary or 1–4 coding are recoded into two indicator variables expressing either support or opposition to each issue. Questions with 1–3 or 1–10 allow for a neutral coding if the answer is coded as 3 or 5 in which case both indicator variables are coded as zero. The fourth (fifth) column contains the share of people are coded as a positive (negative) response to the question.

Table 2: Hierarchy of Types (Top Ten Features)

(a) 2-4 Type Model

2 Type Model	3 Type Model	4 Type Model
Left	**Liberal Centrist**	**Liberal Centrist**
No problem Neighbours: Homosexuals	Confidence: Police	Confidence: Police
No problem Neighbours: People different race	Confidence: Justice System/Courts	No problem Neighbours: Homosexuals
No problem Neighbours: People AIDS	No problem Neighbours: Homosexuals	No problem Neighbours: People AIDS
No problem Neighbours: Immigrants/foreign workers	No problem Neighbours: People different race	No problem Neighbours: People different race
Justifiable: divorce	No problem Neighbours: People AIDS	No problem Neighbours: Immigrants/foreign workers
Not Justifiable: someone accepting a bribe	No problem Neighbours: Immigrants/foreign workers	Not Justifiable: someone accepting a bribe
Justifiable: euthanasia	Proud of nationality	Proud of nationality
Justifiable: homosexuality	Not Justifiable: someone accepting a bribe	Not Justifiable: claiming government benefits
Not Justifiable: claiming government benefits	Confidence: The Civil Services	Justifiable: divorce
Proud of nationality	Not Justifiable: cheating on taxes	Not Justifiable: cheating on taxes
Right	**Conservative Centrist**	**Conservative Centrist**
Not Justifiable: someone accepting a bribe	Not Justifiable: homosexuality	Confidence: Police
Not Justifiable: suicide	Not Justifiable: abortion	Confidence: Churches
Proud of nationality	Not Justifiable: suicide	Not Justifiable: suicide
Not Justifiable: prostitution	Not Justifiable: prostitution	Proud of nationality
Not Justifiable: avoiding a fare on public transport	Proud of nationality	Confidence: Armed Forces
Not Justifiable: claiming government benefits	Not Justifiable: someone accepting a bribe	Not Justifiable: prostitution
Not Justifiable: cheating on taxes	Not Justifiable: avoiding a fare on public transport	Not Justifiable: abortion
Not Justifiable: abortion	Not Justifiable: claiming government benefits	Not Justifiable: someone accepting a bribe
Not Justifiable: homosexuality	Not Justifiable: cheating on taxes	Confidence: Justice System/Courts
No problem Neighbours: People different race	Not Justifiable: euthanasia	Confidence: The Civil Services
	Anarchist	**Left Anarchist**
	No Confidence: Civil Services	Justifiable: divorce
	No Confidence: Parliament	No Confidence: Churches
	No Confidence: Churches	Justifiable: euthanasia
	No Confidence: Major Companies	No Confidence: Armed Forces
	No Confidence: Justice System/Courts	No Confidence: Parliament
	No Confidence: The Press	No Confidence: Civil Services
	No problem Neighbours: Homosexuals	No problem Neighbours: Homosexuals
	No problem Neighbours: People different race	Justifiable: homosexuality
	No problem Neighbours: People AIDS	Justifiable: abortion
	No Confidence: Labour Unions	No problem Neighbours: People different race
		Right Anarchist
		No Confidence: Parliament
		No Confidence: Civil Services
		No Confidence: Labour Unions
		No Confidence: The Press
		No Confidence: Justice System/Courts
		Not Justifiable: someone accepting a bribe
		Not Justifiable: avoiding a fare on public transport
		Not Justifiable: claiming government benefits
		Not Justifiable: suicide
		Not Justifiable: cheating on taxes

(b) 4-5 Type Model

4 Type Model	5 Type Model
Liberal Centrist	**Liberal Centrist**
Confidence: Police	Confidence: The Civil Services
No problem Neighbours: Homosexuals	Confidence: Justice System/Courts
No problem Neighbours: People AIDS	Confidence: Police
No problem Neighbours: People different race	Confidence: Parliament
No problem Neighbours: Immigrants/foreign workers	Proud of nationality
Not Justifiable: someone accepting a bribe	Confidence: Armed Forces
Proud of nationality	Confidence: Labour Unions
Not Justifiable: claiming government benefits	No problem Neighbours: People different race
Justifiable: divorce	Not Justifiable: someone accepting a bribe
Not Justifiable: cheating on taxes	No problem Neighbours: Immigrants/foreign workers
Conservative Centrist	**Conservative Centrist**
Confidence: Police	Not Justifiable: abortion
Confidence: Churches	Not Justifiable: prostitution
Not Justifiable: suicide	Not Justifiable: suicide
Proud of nationality	Not Justifiable: euthanasia
Confidence: Armed Forces	No problem Neighbours: People different race
Not Justifiable: prostitution	No problem Neighbours: Immigrants/foreign workers
Not Justifiable: abortion	Not Justifiable: someone accepting a bribe
Not Justifiable: someone accepting a bribe	Not Justifiable: avoiding a fare on public transport
Confidence: Justice System/Courts	Not Justifiable: cheating on taxes
Confidence: The Civil Services	Proud of nationality
Left Anarchist	**Left Anarchist**
Justifiable: divorce	No Confidence: Armed Forces
No Confidence: Churches	No Confidence: Churches
Justifiable: euthanasia	No Confidence: Police
No Confidence: Armed Forces	No Confidence: Major Companies
No Confidence: Parliament	No Confidence: Justice System/Courts
No Confidence: Civil Services	No Confidence: Parliament
No problem Neighbours: Homosexuals	No Confidence: Civil Services
Justifiable: homosexuality	No problem Neighbours: Homosexuals
Justifiable: abortion	No problem Neighbours: People different race
No problem Neighbours: People different race	No problem Neighbours: People AIDS
Right Anarchist	**Right Anarchist**
No Confidence: Parliament	Against Neighbours: People AIDS
No Confidence: Civil Services	Against Neighbours: Homosexuals
No Confidence: Labour Unions	Against Neighbours: Immigrants/foreign workers
No Confidence: The Press	If Jobs scarce: priority to (nation) people
No Confidence: Justice System/Courts	Against Neighbours: Drug addicts
Not Justifiable: someone accepting a bribe	Not Justifiable: homosexuality
Not Justifiable: avoiding a fare on public transport	Not Justifiable: suicide
Not Justifiable: claiming government benefits	Against Neighbours: People different race
Not Justifiable: suicide	Not Justifiable: prostitution
Not Justifiable: cheating on taxes	Proud of nationality
	Market Liberal
	No Confidence: Parliament
	No problem Neighbours: People different race
	Not Justifiable: claiming government benefits
	No problem Neighbours: Homosexuals
	Competition is good
	No Confidence: The Press
	Not Justifiable: someone accepting a bribe
	Proud of nationality
	No problem Neighbours: People AIDS
	Not Justifiable: cheating on taxes

Notes: This table reports the 10 most important features for a n-type LDA model, where $n \in \{2, 3, 4, 5\}$.

Table 3: Type Correlations

(a) Between-Wave Type Correlations

	Centrist Liberal Wave 2	Centrist Conservative Wave 2	Left Anarchist Wave 2	Right Anarchist Wave 2
Wave 4	0.973	0.985	0.963	0.981
Wave 5	0.935	0.963	0.943	0.939

Notes: This table shows the correlation of the β issue-position probability weights across types estimated in separate waves. That is, we identify 4 types in the initial Wave 2 (1989-1993) and correlate their β weights with the most similar types estimated separately on Waves 4 (1989-1993) and 5 (2004-2009).

(b) Within-Wave Type Correlations

	Centrist Liberal	Centrist Conservative	Left Anarchist	Right Anarchist
Wave 2				
Centrist Liberal	1.000			
Centrist Conservative	0.446	1.000		
Left Anarchist	0.215	-0.599	1.000	
Right Anarchist	0.231	0.256	0.161	1.000
Wave 4				
	Centrist Liberal	Centrist Conservative	Left Anarchist	Right Anarchist
Centrist Liberal	1.000			
Centrist Conservative	0.475	1.000		
Left Anarchist	0.336	-0.502	1.000	
Right Anarchist	0.291	0.345	0.178	1.000
Wave 5				
	Centrist Liberal	Centrist Conservative	Left Anarchist	Right Anarchist
Centrist Liberal	1.000			
Centrist Conservative	0.547	1.000		
Left Anarchist	0.282	-0.341	1.000	
Right Anarchist	0.201	0.319	0.317	1.000

Notes: This table shows the correlation of the β issue-position probability weights amongst types in the same wave. That is, we estimate our 4 types using data on a single wave and then correlate the β weights across pairs of types in the same wave.

Table 4: Correlates of Individual-level Type Shares

VARIABLES	(1) Liberal Centrist	(2) Conservative Centrist	(3) Left Anarchist	(4) Right Anarchist
Female	0.019***	0.007***	-0.015***	-0.011***
	(0.002)	(0.002)	(0.002)	(0.002)
Age	-0.003***	0.004***	-0.003***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Unemployed	-0.072***	-0.002	0.049***	0.025***
	(0.005)	(0.005)	(0.004)	(0.005)
Wave 4	0.070***	-0.063***	0.020***	-0.027***
	(0.003)	(0.003)	(0.002)	(0.003)
Wave 5	0.057***	-0.077***	0.059***	-0.039***
	(0.003)	(0.003)	(0.002)	(0.003)
Observations	81,141	81,141	81,141	81,141
R-squared	0.141	0.106	0.113	0.056
Country FE	YES	YES	YES	YES

Notes: Each column reports the regression results for individual level regression. The dependent variable are the type shares for one of the 4 types created by LDA. Robust standard errors are used. Significance levels: *** p<0.01, ** p<0.05, and * p<0.1. The data come from the World Value Survey and the European Value Survey.

Table 5: Correlates of 'Citizen Slant' (Gini Concentration)

VARIABLES	(1) log Gini	(2) log Gini	(3) log Gini US	(4) log Gini US
Cons. Centrist	-0.003		-0.029***	
	(0.002)		(0.010)	
L. Anarchist	-0.047***		-0.090***	
	(0.003)		(0.017)	
R. Anarchist	-0.039***		-0.070***	
	(0.002)		(0.010)	
Wave 4	0.027***	0.028***	0.042***	0.040***
	(0.002)	(0.002)	(0.009)	(0.009)
Wave 5	0.013***	0.012***	0.062***	0.048***
	(0.002)	(0.002)	(0.010)	(0.010)
Baseline Gini	0.746	0.746	0.736	0.736
Observations	81,141	81,141	4,197	4,197
R-squared	0.017	0.009	0.031	0.012
Country FE	YES	YES	YES	YES

Notes: Each column reports the regression results for individual level regression. The dependent variable is the Gini Coefficient of the individual type shares as a measure of polarization. Column (1) and (2) use all data and column (3) and (4) restrict the sample to the USA. Robust standard errors are used. Significance levels: *** p<0.01, ** p<0.05, and * p<0.1. The data come from the World Value Survey and the European Value Survey.

Table 6: 'Citizen Slant' - US vs non-US Comparison

Panel A: United States				
	(1)	(2)	(3)	(4)
VARIABLES	Cent. Liberal	Cent. Conservative	Left Anarchist	Right Anarchist
Wave 4	0.033**	0.020	0.088**	0.093***
	(0.013)	(0.015)	(0.041)	(0.019)
Wave 5	-0.012	0.043**	0.151***	0.142***
	(0.019)	(0.020)	(0.039)	(0.016)
Baseline Gini	.769	.752	.662	.672
Observations	1,398	1,383	283	1,133
R-squared	0.017	0.016	0.086	0.089

Panel B: Non United States				
	(1)	(2)	(3)	(4)
VARIABLES	Cent. Liberal	Cent. Conservative	Left Anarchist	Right Anarchist
Wave 4	0.051***	0.005	0.023***	0.040***
	(0.004)	(0.004)	(0.006)	(0.004)
Wave 5	0.038***	-0.007*	0.023***	0.015***
	(0.004)	(0.004)	(0.006)	(0.004)
Baseline Gini	.748	.770	.722	.733
Observations	24,595	20,725	11,399	20,225
R-squared	0.015	0.013	0.012	0.008
Country FE	YES	YES	YES	YES

Panel C: All Countries				
	(1)	(2)	(3)	(4)
VARIABLES	Cent. Liberal	Cent. Conservative	Left Anarchist	Right Anarchist
Wave 4	0.049***	0.007*	0.024***	0.043***
	(0.004)	(0.004)	(0.006)	(0.004)
Wave 5	0.033***	-0.004	0.025***	0.022***
	(0.004)	(0.004)	(0.006)	(0.004)
Baseline Gini	.750	.768	.721	.730
Observations	25,993	22,108	11,682	21,358
R-squared	0.014	0.013	0.012	0.009
Country FE	YES	YES	YES	YES

Notes: Each column reports the regression results for individual level regression. The dependent variable is the Gini Coefficient of the individual type shares as a measure of polarization. Column (1) use all US data and column (2), (3) and (4) restrict the sample to the individuals based on their dominant type. Robust standard errors are used. Significance levels: *** p<0.01, ** p<0.05, and * p<0.1. The data come from the World Value Survey and the European Value Survey.

A Additional Type Hierarchy Information

Table 7: Extended Hierarchy of Types (Top Ten Features)

6 Type Model	7 Type Model
LIBERAL CENTRIST	**LIBERAL CENTRIST**
Confidence: The Civil Services	Confidence: The Civil Services
Confidence: Parliament	Confidence: Parliament
Confidence: Police	Confidence: Police
Confidence: Justice System/Courts	Confidence: Justice System/Courts
Proud of nationality	Proud of nationality
No problem Neighbours: People different race	No problem Neighbours: People different race
No problem Neighbours: Homosexuals	No problem Neighbours: Homosexuals
No problem Neighbours: Immigrants/foreign workers	No problem Neighbours: Immigrants/foreign workers
Not Justifiable: someone accepting a bribe	Not Justifiable: someone accepting a bribe
No problem Neighbours: People AIDS	Justifiable: divorce
'SOFT' SOCIAL CONSERVATIVE	**'SOFT' SOCIAL CONSERVATIVE**
Not Justifiable: abortion	Confidence: Police
Confidence: Police	Not Justifiable: abortion
Confidence: Churches	Confidence: Churches
Not Justifiable: prostitution	Not Justifiable: euthanasia
Confidence: Armed Forces	Confidence: Armed Forces
Not Justifiable: suicide	Not Justifiable: suicide
Not Justifiable: someone accepting a bribe	Not Justifiable: prostitution
Not Justifiable: cheating on taxes	No problem Neighbours: People different race
Not Justifiable: euthanasia	Not Justifiable: someone accepting a bribe
Not Justifiable: avoiding a fare on public transport	Not Justifiable: cheating on taxes
LEFT ANARCHIST	**LEFT ANARCHIST**
No Confidence: Armed Forces	No Confidence: Armed Forces
No Confidence: Churches	No Confidence: Churches
Justifiable: divorce	Justifiable: divorce
Justifiable: homosexuality	No Confidence: Major Companies
No problem Neighbours: Homosexuals	Justifiable: homosexuality
Justifiable: euthanasia	No Confidence: Parliament
No Confidence: Major Companies	No problem Neighbours: Homosexuals
No problem Neighbours: People AIDS	No problem Neighbours: People AIDS
No problem Neighbours: People different race	Justifiable: euthanasia
No problem Neighbours: Immigrants/foreign workers	No problem Neighbours: People different race
RIGHT ANARCHIST 'HARD' SOCIAL CONSERVATIVE	**'HARD' SOCIAL CONSERVATIVE**
Against Neighbours: People AIDS	Against Neighbours: People AIDS
Against Neighbours:Immigrants/foreign workers	Against Neighbours:Homosexuals
Against Neighbours:Homosexuals	Against Neighbours:Immigrants/foreign workers
Against Neighbours: People different race	Against Neighbours: Drug addicts
If Jobs scarce: priority to (nation) people	If Jobs scarce: priority to (nation) people
Against Neighbours: Drug addicts	Against Neighbours: People different race
Not Justifiable: homosexuality	Not Justifiable: homosexuality
Confidence: Armed Forces	Proud of nationality
Proud of nationality	Confidence: Armed Forces
Justifiable: accepting a bribe	Not Justifiable: someone accepting a bribe
RIGHT ANARCHIST (GMU BLOGGER)	**RIGHT ANARCHIST (GMU BLOGGER)**
No Confidence: Parliament	No Confidence: Parliament
Not Justifiable: claiming government benefits	Confidence: Police
Competition is good	No problem Neighbours: People different race
Not Justifiable: someone accepting a bribe	No problem Neighbours: Homosexuals
No Confidence: The Press	No Confidence: The Press
Proud of nationality	Competition is good
Confidence: Police	Proud of nationality
No Confidence: Civil Services	Confidence: Armed Forces
Confidence: Armed Forces	No problem Neighbours: People AIDS
Not Justifiable: cheating on taxes	Not Justifiable: someone accepting a bribe
EXTRA RIGHT WING TYPE	**EXTRA RIGHT WING TYPE**
No Confidence: Armed Forces	No Confidence: Parliament
No Confidence: Parliament	No Confidence: Civil Services
No Confidence: Justice System/Courts	No Confidence: Justice System/Courts
No Confidence: Civil Services	No Confidence: Labour Unions
No Confidence: The Press	No Confidence: Armed Forces
No Confidence: Labour Unions	Not Justifiable: suicide
Not Justifiable: suicide	No Confidence: Major Companies
No Confidence: Major Companies	No Confidence: The Press
Not Justifiable: someone accepting a bribe	No problem Neighbours: People different race
No Confidence: Police	Not Justifiable: prostitution
	SUPER ANARCHIST 'Rage Against the Machine'
	Justifiable: avoiding a fare on public transport
	Justifiable: cheating on taxes
	Justifiable: claiming government benefits
	Justifiable: accepting a bribe
	Justifiable: euthanasia
	If Jobs scarce: priority to (nation) people
	Justifiable: prostitution
	Proud of nationality
	Justifiable: divorce
	No problem Neighbours: People different race

Notes: This table reports the 10 most important features for a n-type LDA model, where $n \in \{6, 7\}$.

Table 8: Issues of Increasing Importance between Wave 2 and Wave 5

Question	Baseline	Change
Liberal Centrist		
State ownership better than private ownership	0.012	0.321
More responsibility for government	0.217	0.209
Competition is harmful	0.000	0.188
Confidence: Armed Forces	0.697	0.162
Against Neighbours: Drug addicts	0.673	0.151
Justifiable: homosexuality	0.710	0.141
Confidence: The Civil Services	0.688	0.130
Confidence: Labour Unions	0.591	0.129
No Confidence: Major Companies	0.495	0.110
Confidence: Churches	0.561	0.084
Conservative Centrist		
No Confidence: Parliament	0.036	0.204
No problem Neighbours: Homosexuals	0.597	0.199
Justifiable: homosexuality	0.000	0.167
No problem Neighbours: Drug addicts	0.431	0.159
No problem Neighbours: People AIDS	0.653	0.151
No Confidence: Major Companies	0.269	0.141
If Jobs scarce: no priority to (nation) people	0.318	0.079
No Confidence: The Press	0.509	0.054
Not Justifiable: abortion	0.834	0.052
Not Justifiable: euthanasia	0.837	0.031
Left Anarchist		
Confidence: Police	0.041	0.327
Not Justifiable: cheating on taxes	0.409	0.282
Confidence: Armed Forces	0.002	0.177
Proud of nationality	0.586	0.137
Not Justifiable: claiming government benefits	0.633	0.102
No Confidence: The Press	0.721	0.096
Not Justifiable: avoiding a fare on public transport	0.549	0.093
If Jobs scarce: no priority to (nation) people	0.591	0.078
Not Justifiable: prostitution	0.094	0.076
More responsibility for people	0.260	0.067
Right Anarchist		
Confidence: Armed Forces	0.028	0.632
Confidence: Police	0.000	0.564
Against Neighbours: Immigrants/foreign workers	0.351	0.235
Justifiable: homosexuality	0.000	0.220
Justifiable: divorce	0.136	0.172
Justifiable: euthanasia	0.140	0.166
Against Neighbours: People different race	0.301	0.133
No problem Neighbours: Homosexuals	0.645	0.087
State ownership better than private ownership	0.297	0.080
Against Neighbours: Drug addicts	0.788	0.078

Notes: This table reports the 10 feature of each type which show the biggest increase in weight from wave 2 to wave 5. Column 2 reports the baseline value in wave 2 and column 3 reports the change from wave 2 to wave 5.

Table 9: Issue Position Differences between Types (4 Type Model)

	Difference Lib. Centrist	Difference Cons. Centrist	Difference Left Anarchist	Difference Right Anarchist
L. Centrist		Justifiable: divorce Justifiable: euthanasia Justifiable: homosexuality Justifiable: abortion No Confidence: Churches Justifiable: prostitution No problem Neighbours: Homosexuals If Jobs scarce: no priority to natives Against Neighbours: Drug addicts Justifiable: suicide	Confidence: Justice System/Courts Confidence: Police Confidence: Armed Forces Confidence: The Civil Services Confidence: Parliament Confidence: Major Companies Against Neighbours: Drug addicts Confidence: Churches Not Justifiable avoiding fare transport Competition is good	Confidence: Justice System/Courts Justifiable: divorce Confidence: The Civil Services Justifiable: homosexuality Justifiable: abortion Confidence: Police Justifiable: euthanasia Confidence: Parliament Confidence: Labour Unions Confidence: Major Companies
C. Centrist	Not Justifiable: abortion Not Justifiable: euthanasia Not Justifiable: homosexuality Not Justifiable: divorce Confidence: Churches Not Justifiable: prostitution Not Justifiable: suicide If Jobs scarce: priority to natives No problem Neighbours: Drug addicts Against Neighbours: Homosexuals		Confidence: Churches Not Justifiable: abortion Not Justifiable: euthanasia Confidence: The Civil Services Confidence: Justice System/Courts Not Justifiable: prostitution Confidence: Armed Forces Confidence: Police Confidence: Parliament Not Justifiable: homosexuality	Confidence: Justice System/Courts Confidence: The Civil Services Confidence: Major Companies Confidence: Parliament Confidence: Police Confidence: Labour Unions Confidence: Churches Confidence: Press Confidence: Armed Forces Not Justifiable: euthanasia
L. Anarchist	No Confidence: Justice System/Courts No Confidence: Armed Forces No Confidence: Civil Services No Confidence: Parliament No Confidence: Police No Confidence: Churches No problem Neighbours: Drug addicts No Confidence: Major Companies Justifiable: avoiding fare on pub. transport Not proud of nationality	No Confidence: Churches No Confidence: Civil Services Justifiable: divorce No Confidence: Armed Forces No Confidence: Parliament No Confidence: Justice System/Courts Justifiable: euthanasia Justifiable: homosexuality Justifiable: abortion No Confidence: Major Companies		Justifiable: divorce Justifiable: homosexuality Justifiable: abortion Justifiable: euthanasia No problem Neighbours: Drug addicts Justifiable: prostitution No Confidence: Churches No Confidence: Armed Forces Justifiable: suicide Justifiable: avoiding fare transport
R. Anarchist	No Confidence: Justice System/Courts No Confidence: Civil Services No Confidence: Parliament Not Justifiable: abortion No Confidence: Labour Unions Not Justifiable: homosexuality No Confidence: The Press No Confidence: Major Companies If Jobs scarce: priority to natives Not Justifiable: euthanasia	No Confidence: Civil Services No Confidence: Parliament No Confidence: Justice System/Courts No Confidence: Major Companies No Confidence: Labour Unions No Confidence: The Press No Confidence: Churches No Confidence: Police No Confidence: Armed Forces Against Neighbours: Drug addicts	Not Justifiable: prostitution Not Justifiable: abortion Not Justifiable: suicide Against Neighbours: Drug addicts Not Justifiable: homosexuality If Jobs scarce: priority to natives Not Justifiable: euthanasia Not Justifiable: avoiding fare transport Against Neighbours: People AIDS Against Neighbours: Immigrants	

Notes: This table reports the 10 features for which there exists the largest differences between the 4 ideological types created by LDA. The model is fitted to the 5th wave of the sample. The type labels are chosen by the author.

B Additional Details on the Selection of Question from the WVS

This section describes in more detail, the selection process that lead to the 29 questions that are used in the paper. There are 6 waves of the World Value Survey (WVS) and 4 waves of the European Value Survey (EVS). The 4 Waves of the EVS correspond to the 1st, 2nd, 4th and 5th wave of the WVS. When constructing the Integrated Value Survey by combining the WVS and EVS we excluded the 1st wave since, it contained a smaller set of countries and questions. The Integrated Value Survey (WVS) in total contains 971 different items grouped in 13 different categories (number of questions in brackets): Environment (25), Family (64), National Identity (105), Perceptions of life (210), Politics and Society (267), Religion and Morale (122), Science (2), Security (22), Socio-demographics (38), Special Indexes (3), Structure of the file (25), Sylatech module (42) and Work (46). The categories socio-demographics, special indexes, structure of the study, Sylatech module and work do not contain any question concerning the values of people.

We limited the set of questions to those question which were consistently asked in the 2nd, 4th and 5th wave of WVS. This already reduced the set of possible questions down to 92. From these 92 questions we chose our 29 based on which questions seems most important for the evaluation of a persons ideological type.[17] The excluded question are listed in Table 10

In Appendix D we show that the selection of these 29 questions is not crucial for our findings and that the ideological types are very similar if we use all 92 questions. We further show that also removing any of the 29 questions from our data has no bearing on our results.

A further point is that LDA does not allow for missing responses in the data. If we simply excluded all observations with any missing responses and restricted ourselves to observations with complete sets of answers, we would need to drop sizable fractions of the WVS data. We instead impute a small set of missing responses with the sample mean of the non-missing data in the same wave. This treatment of missing data allows LDA to use the information from this observation. Moreover, the imputation has only a minimal effect on the LDA classification, since the sample mean does not influence the classification of each individual. Imputation with the mean is also preferable to an alternative approach

[17] The question concerning "Confidence: Justice System/Courts" did not get asked in the 4th wave of the WVS. Since we have responses for this question from the EVS we still decided to include it in our baseline question set. The question is not included in the set of 92 questions.

Table 10: List of Excluded Questions

Code	Questions	Code	Questions
A001	Important in life: Family	D057	Being a housewife just as fulfilling
A002	Important in life: Friends	E001	Aims of country: first choice
A003	Important in life: Leisure time	E002	Aims of country: second choice
A004	Important in life: Politics	E003	Aims of respondent: first choice
A005	Important in life: Work	E004	Aims of respondent: second choice
A006	Important in life: Religion	E005	Most important: first choice
A008	Feeling of happiness	E006	Most important: second choice
A009	State of health (subjective)	E012	Willingness to fight for country
A029	Important child qualities: independence	E015	Future changes: Less importance placed on work
A030	Important child qualities: hard work	E016	Future changes: More emphasis on technology
A032	Important child qualities: feeling of responsibility	E018	Future changes: Greater respect for authority
A034	Important child qualities: imagination	E019	Future changes: More emphasis on family life
A035	Important child qualities: tolerance and respect for other people	E022	Opinion about scientific advances
A038	Important child qualities: thrift saving money and things	E023	Interest in politics
A039	Important child qualities: determination perseverance	E025	Political action: signing a petition
A040	Important child qualities: religious faith	E026	Political action: joining in boycotts
A041	Important child qualities: unselfishness	E027	Political action: attending lawful/peaceful demonstrations
A042	Important child qualities: obedience	E033	Self positioning in political scale
A124_03	Neighbours: Heavy drinkers	E035	Income equality
A124_05	Neighbours: Muslims	E069_11	Confidence: The Government
A165	Most people can be trusted	E069_12	Confidence: The Political Parties
A170	Satisfaction with your life	E069_18	Confidence: The European Union
A173	How much freedom of choice and control	F001	Thinking about meaning and purpose of life
B001	Would give part of my income for the environment	F025	Religious denomination
B002	Increase in taxes if used to prevent environmental pollution	F028	How often do you attend religious services
B003	Government should reduce environmental pollution	F034	Religious person
C001	Jobs scarce: Men should have more right to a job than women	F035	Churches give answers: moral problems
C006	Satisfaction with financial situation of household	F036	Churches give answers: the problems of family life
C059	Fairness: One secretary is paid more	F037	Churches give answers: peoples spiritual needs
D018	Child needs a home with father and mother	F038	Churches give answers: the social problems
D022	Marriage is an out-dated institution	F063	How important is God in your life
D023	Woman as a single parent	F065	Moments of prayer, meditation...

Notes: This table contains the questions that were excluded from baseline LDA model.

where we would simply replace all missing responses with 0s, because the 0s would bias the classification.

C Appendix: Interpretation of the β Vectors

One difference between our application and the standard use of LDA is that in our case features can only appear once for each observation, i.e. people can only answer each questions once, while words can appear more than once in an document. As already discussed in the main part of the paper this does not influence on the validity of LDA, since LDA exploits how often features appear relative to each other.

However, this difference influences the interpretation of the β vectors. The β vectors capture the probability that a response is drawn in each of the 58 draws (questions) asked to an individual, e.g. how likely it is that an individual will answer that he is opposed to abortion in each of the 58 draws. Therefore, the β vectors do not take into account that once an person has answered a questions the same person cannot answer the same question again.

As a results, the β still capture which groups are more likely to exhibit an ideological position, but the values do not have a natural interpretation within our setting. If necessary one can scale up the β probabilities to give them a more natural interpretation within our setting. To do this one has to calculate the probability that a feature shows up in any of the 58 draws of the LDA taking into account that a question can only be answered once. Given this intuition $P_{f,t}$, the overall probability that a feature f appears if the chosen type is t, can be expressed as $P_{f,t} = \sum_{d=1}^{58}(1-\beta_{f,t})^{d-1}\beta_{f,t}$, where d is the number of the draw (question) and $\beta_{f,t}$ is the value of the β vector for feature f and type t. In this expression $(1-\beta_{f,t})^{d-1}$ is the probability that the response has not been given in any previous draw and $\beta_{f,t}$ is the probability that the response will be given in the current draw.

As an example to illustrate this calculation consider the question of "Confidence in the Police". In the 5th wave, the liberal centrist has a value of $\beta_{14,1} = 0.0403$ and the value for the left anarchist is $\beta_{14,3} = 0.0078$. This difference in the β values translate into the following overall differences in probability. While a liberal centrist will express confidence in the police with a probability of 90.8%, the probability that a left anarchist will express similar views is only 35.9%.

This scaling up does not take into account some features are mutually exclusive. Hence, the scaled up probability of the features "Confidence in the Police" and "No Confidence in the Police" will not necessarily add up to 1.

D Sensitivity to Removal and Addition of Features

In this section we analyze how sensitive our baseline 4 type model is to the removal and addition of features. The exercises we run here can be interpreted as a leverage or influence analysis on the statistical definition of our ideological clusters. We are unaware of formal model robustness statistics of this nature in the literature on LDA. Hence, while we think that the exercises below are promising in terms of the robustness of the basic clusters that they reveal, they should be considered indicative.

D.1. 'Leave One Out' Clusters.

As a first exercise, we re-estimate the 4-topic model removing 1 of the 29 questions (2 of the 58 features) at a time. Afterwards, we compare the original model to the new 'leave one out' model based on the similarity of the β vectors, as measured by their correlation. Table 11 reports the results of this exercise.

Overall, independent of the particular removed question, we find high correlations between the different β vectors. This is strongest for the Liberal Centrist type which has an average correlation of 0.969 between the original and leave one out models across all dropped questions. This indicates that the types generate by LDA are very closely comparable across the different models. The highest degrees of sensitivity relate to the confidence in institutions questions (where the β correlations are between 0.70-0.80 for three of the types). Another point of sensitivity is questions relating to foreigners / immigration in the case of the Right Anarchist. Given the centrality of the confidence and immigration questions to the character of different types these sensitivities are within expectations. This leads to the next issue of how the types might change when we add more information.

D.2. Widening the Feature Set.

In the next exercise we investigate how the structure of our clusters changes when we include additional features in the topic model. As described in Appendix B, there are a total of 92 questions that are available across all 3 waves of the WVS used in this paper. As an additional robustness check, we therefore include all these 92 questions in our topic model and create a extended type hierarchy. We then correlate the weights on the β positions between the original and extended models where they overlap.

Table 11: Sensitivity Removal of Features - 'Leave One Out' Exercise.

Question Code	Removed Question	Type 1 Lib. Centrist	Type 2 Cons. Centrist	Type 3 Left Anarchist	Type 4 Right Anarchist
A124_02	Against Neighbours: People different race	0.970	0.898	0.737	0.475
A124_06	Against Neighbours: Immigrants/foreign workers	0.994	0.999	0.944	0.478
A124_07	Against Neighbours: People AIDS	0.971	0.896	0.751	0.991
A124_08	Against Neighbours: Drug addicts	0.957	0.631	0.868	0.806
A124_09	Against Neighbours: Homosexuals	0.971	0.861	0.802	0.407
	Average Neighbours:	0.973	0.857	0.820	0.631
C002	If Jobs scarce: priority to (nation) people	0.980	0.919	0.962	0.976
E036	Private better than state ownership	0.983	0.925	0.951	0.968
E037	More responsibility for government	0.993	0.969	0.946	0.973
E039	Competition is good	0.972	0.867	0.933	0.943
	Average Economics:	0.982	0.920	0.948	0.965
E069_01	Confidence: Churches	0.979	0.879	0.954	0.965
E069_02	Confidence: Armed Forces	0.963	0.846	0.151	0.812
E069_04	Confidence: Press	0.965	0.398	0.835	0.740
E069_05	Confidence: Labour Unions	0.965	0.808	0.802	-0.221
E069_06	Confidence: Police	0.944	0.711	0.880	0.849
E069_07	Confidence: Parliament	0.982	0.781	0.875	0.823
E069_08	Confidence: The Civil Services	0.974	0.761	0.869	0.813
E069_13	Confidence: Major Companies	0.975	0.752	0.868	0.816
E069_17	Confidence: Justice System/Courts	0.958	0.771	0.894	0.868
	Average Confidence	0.967	0.745	0.792	0.718
F114	Justifiable: claiming government benefits	0.961	0.703	0.876	0.809
F115	Justifiable: avoiding a fare on public transport	0.968	0.696	0.853	0.748
F116	Justifiable: cheating on taxes	0.964	0.691	0.859	0.759
F117	Justifiable: accepting a bribe	0.965	0.710	0.860	0.787
	Average Fairness Values	0.964	0.700	0.862	0.776
F118	Justifiable: homosexuality	0.958	0.699	0.908	0.911
F119	Justifiable: prostitution	0.975	0.645	0.857	0.672
F120	Justifiable: abortion	0.944	0.674	0.910	0.837
F121	Justifiable: divorce	0.968	0.594	0.781	0.595
F122	Justifiable: euthanasia	0.958	0.682	0.807	0.599
F123	Justifiable: suicide	0.986	0.857	0.855	0.802
	Average Social Values	0.965	0.692	0.853	0.736
G006	Proud of nationality	0.955	0.903	0.676	0.432
	Average All:	0.969	0.777	0.837	0.739

Notes: This table reports the correlation of the β vectors of our baseline model with topic models in which 1 of the 29 features of the baseline model got removed.

Practically, this exercise allows us to ask whether the relative ordering of the β issue-position weights changes as we add more features to the model. Note that this is more of an 'add them all in' rather than an iterative 'add one in' exercise. We adopt this approach both for the sake of brevity as well to see how our original 4-type model is affected by a large, lateral addition of information. The concern would be that the addition of many extra features would fundamentally change the structure of the clusters and shift the ordering of the initial set of features.

Table 12 reports the correlations between the β-vectors from the baseline type hierarchy and those from the extended-feature type hierarchy. Obviously, the correlation coefficients can only be calculated on the basis of the 29 original questions used in the baseline hierarchy. The correlations are very high across all the hierarchy models. The only exception is the Type 3' (nominated Left Anarchist) in the 4-type model. As we show, Type 3' has comparable correlations with the original model Liberal Centrist, Left Anarchist and Right Anarchist types. Interestingly, the Left Anarchist type re-emerges as cleanly defined when we move to the 5-type model.

Overall, we find these results to be encouraging. The same basic type structure is intact even when adding in a large amount of information. This is compatible with the idea that the extra questions/features fit in as new responses that tap into a stable set of underlying latent types.

We stress though that both of the exercises we present here are indicative with limited formal precedents in the LDA literature. One interesting pattern here is that the Centrist types are less sensitive to changes in features relative to the Anarchist types. This fits with the intuition that the Centrist types are well-established and better defined with the Anarchist types still being more fluid. The tendency of the Anarchist types to split as we consider higher-order models (eg: 5, 6, and 7-type models) is also consistent with assessment.

Table 12: Sensitivity to Additional Features

2 Type Model					
	Type 1'	Type 2'			
Left	0.9821				
Right		0.9816			
3 Type Model					
	Type 1'	Type 2'	Type 3'		
Lib. Centrist	0.8222				
Cons. Centrist		0.8005			
Anarchist			0.8302		
4 Type Model					
	Type 1'	Type 2'	Type 3'	Type 4'	
Lib. Centrist	0.8748				
Cons. Centrist		0.9693			
Left Anarchist	*0.687*		0.6622		
Right Anarchist			*0.7835*	0.9496	
5 Type Model					
	Type 1'	Type 2'	Type 3'	Type 4'	Type 5'
Lib. Centrist	0.825				
'Hard' Conservative		0.911			
'Soft' Conservative			0.8914		
GMU Blogger				0.8768	
Left Anarchist					0.9168

Notes: This table reports the correlation of the β vectors of the type hierarchy from the main paper and the type hierarchy of a topic model including all 92 consistent questions from the WVS. The prime' notation indicates the types estimated using the 92 feature topic model. We report the highest cross-model correlations for the overlapping β weights, except for the 4-type Left Anarchist case where (in the interests of exposition) we report the three highest correlations.

E Comparison of LDA to PCA, Factor Analysis and k-means

This section provides a comparison between Latent Dirichlet Allocation (LDA) and the other alternative machine learning dimensionality reduction techniques, specifically Principal Component Analysis (PCA), Factor Analysis (FA) and k-means clustering. At their core all of these techniques aim to reduce high dimensional data to a set of more easily interpretable topics, components, factors or clusters. Differences arises in the way these lower dimensional representation of the data are constructed.

As we have outlined in detail in the main part of the paper LDA relies on a generative model that makes assumptions about the data generating process and allow for an direct interpretation of the latent objects as topics. Furthermore, the LDA model was specifically designed for the analysis of sparse multinomial data.

PCA on the other hand relies on a truncated singular value decomposition to derive components that explain the maximum possible amount of variance in the data while keeping all components orthogonal to each other. Truncated singular value decomposition is based on decomposing the original $O \times F$ data matrix D of rank R with O observation and F features into three matrices such that $D = U\Sigma W^T$, where U is a $O \times R$ orthogonal matrix, W^T is a $R \times F$ orthogonal matrix, and Σ is a $R \times R$ diagonal matrix. Afterwards, PCA truncates the resulting matrices by removing the rows and columns associated with the smallest eigenvalues in the matrix Σ. This truncation process reduces the dimensions of the matrices to a user-chosen number of components C, such that U becomes U_C of dimension $O \times C$, Σ becomes Σ_C of dimension $C \times C$, and W^T becomes W_C^T of dimension $C \times F$.

Each of the resulting components are orthogonal to each other and represent a linear combination of the original data weighted by eigenvectors. This highlights two important limitations of PCA for our application. Neither is it obvious that the ideological types (components) we want to find in the data should be orthogonal to each other nor are they necessarily a linear combination of the data. As a result the ideological type hierarchy created by PCA (see Table 13) is less coherent than the types created by LDA.

Similar problems arise when using FA. FA represents the original data as a linear combination of factors such that $D = C + \beta \cdot F + \epsilon$, where D is the original data matrix, C is a vector of constants F is the factor matrix, β are the factor loadings and ϵ a vector of Gaussian noise. The advantage of FA in comparison to PCA is that it accounts for random measurement error through the ϵ vector and hence allows for heteroscedastic noise. Nevertheless, FA is still uses a linear model to decompose the data. Due to the linear model the ideological type generate by FA (see Table 14) are less coherent than the LDA results.

Last, k-means is a clustering algorithm that minimizes the distance of the original data to a user chosen number of centroids. As any other clustering algorithm, k-means assigns each observation to a unique cluster. This seems counterintuitive in our case, since people do not necessarily subscribe to a single political ideology. For example, people might be liberal when it comes to social issues but conservative with regard to economic questions. While LDA captures this its mixture of ideological types, k-means cannot account for this.[18] Moreover, as discussed by Ding and He (2004) k-means clustering represents a discrete cluster solution to the components derived by PCA. As such k-means suffers from similar short comings as PCA and the derived ideological types (see Table 15) also are as less coherent in comparison to LDA.

[18]PCA and FA are also allow for 'mixed membership' through different component and factor loadings.

Table 13: **Hierarchy of Types (Top Ten Features) as created by PCA**

2 Type Model	3 Type Model	4 Type Model
Type 1	**Type 1**	**Type 1**
No Confidence: Churches	No Confidence: Churches	No Confidence: Churches
No Confidence: Civil Services	No Confidence: Civil Services	No Confidence: Civil Services
No Confidence: Parliament	No Confidence: Parliament	No Confidence: Parliament
No Confidence: Armed Forces	No Confidence: Armed Forces	No Confidence: Armed Forces
No Confidence: Justice System/Courts	No Confidence: Justice System/Courts	No Confidence: Justice System/Courts
No Confidence: Police	No Confidence: Police	No Confidence: Police
No Confidence: Major Companies	No Confidence: Major Companies	No Confidence: Major Companies
Justifiable: euthanasia	Justifiable: euthanasia	Justifiable: euthanasia
Justifiable: abortion	Justifiable: abortion	Justifiable: abortion
Justifiable: divorce	. Justifiable: divorce	. Justifiable: divorce
Type 2	**Type 2**	**Type 2**
Not Justifiable: abortion	Not Justifiable: abortion	Not Justifiable: abortion
Not Justifiable: homosexuality	Not Justifiable: homosexuality	Not Justifiable: homosexuality
Not Justifiable: euthanasia	Not Justifiable: euthanasia	Not Justifiable: euthanasia
No Confidence: Justice System/Courts	No Confidence: Justice System/Courts	No Confidence: Justice System/Courts
No Confidence: Parliament	No Confidence: Parliament	No Confidence: Parliament
No Confidence: Civil Services	No Confidence: Civil Services	No Confidence: Civil Services
Not Justifiable: divorce	Not Justifiable: divorce	Not Justifiable: divorce
No Confidence: Labour Unions	No Confidence: Labour Unions	No Confidence: Labour Unions
Not Justifiable: prostitution	Not Justifiable: prostitution	Not Justifiable: prostitution
No Confidence: The Press	No Confidence: The Press	No Confidence: The Press
	Type 3	**Type 3**
	Against Neighbours: Drug addicts	Against Neighbours: Drug addicts
	More responsibility for people	More responsibility for people
	If Jobs scarce: priority to (nation) people	If Jobs scarce: priority to (nation) people
	Against Neighbours: People AIDS	Against Neighbours: People AIDS
	Private better than state ownership	Private better than state ownership
	Competition is good	Competition is good
	Justifiable: euthanasia	Justifiable: euthanasia
	Against Neighbours: Immigrants/foreign workers	Against Neighbours: Immigrants/foreign workers
	Confidence: Armed Forces	Confidence: Armed Forces
	Against Neighbours: Homosexuals	Against Neighbours: Homosexuals
		Type 4
		More responsibility for government
		Confidence: Labour Unions
		State ownership better than private ownership
		Against Neighbours: People AIDS
		Competition is harmful
		Confidence: Press
		Against Neighbours: Homosexuals
		If Jobs scarce: priority to (nation) people
		Against Neighbours: Immigrants/foreign workers
		Against Neighbours: People different race

Notes: This table reports the 10 most important features for a n-type Principle Component Analysis model, where $n \in \{2, 3, 4\}$.

Table 14: Hierarchy of Types (Top Ten Features) as created by FA

2 Type Model	3 Type Model	4 Type Model
Type 1	**Type 1**	**Type 1**
No Confidence: Churches	No Confidence: Civil Services	No Confidence: Civil Services
No Confidence: Civil Services	No Confidence: Churches	No Confidence: Churches
No Confidence: Parliament	No Confidence: Parliament	No Confidence: Parliament
No Confidence: Armed Forces	No Confidence: Justice System/Courts	No Confidence: Justice System/Courts
No Confidence: Justice System/Courts	No Confidence: Armed Forces	Justifiable: abortion
Justifiable: abortion	No Confidence: Police	No Confidence: Police
No Confidence: Police	Justifiable: abortion	No Confidence: Armed Forces
Justifiable: divorce	Justifiable: euthanasia	Justifiable: euthanasia
Justifiable: euthanasia	Justifiable: divorce	Justifiable: divorce
Justifiable: homosexuality	No Confidence: Major Companies	Justifiable: homosexuality
Type 2	**Type 2**	**Type 2**
Not Justifiable: abortion	Not Justifiable: abortion	Not Justifiable: abortion
No Confidence: Parliament	No Confidence: Parliament	No Confidence: Parliament
No Confidence: Civil Services	No Confidence: Civil Services	No Confidence: Civil Services
Not Justifiable: homosexuality	Not Justifiable: homosexuality	Not Justifiable: homosexuality
No Confidence: Justice System/Courts	Not Justifiable: euthanasia	No Confidence: Justice System/Courts
Not Justifiable: euthanasia	No Confidence: Justice System/Courts	Not Justifiable: euthanasia
Not Justifiable: divorce	Not Justifiable: divorce	Not Justifiable: divorce
No Confidence: Labour Unions	Not Justifiable: prostitution	No Confidence: Labour Unions
Not Justifiable: prostitution	No Confidence: Labour Unions	Not Justifiable: prostitution
No Confidence: Police	Not Justifiable: suicide	No Confidence: The Press
	Type 3	**Type 3**
	Against Neighbours: Immigrants/foreign workers	Against Neighbours: Immigrants/foreign workers
	Against Neighbours: People different race	Against Neighbours: People different race
	Against Neighbours: People AIDS	Against Neighbours: People AIDS
	Against Neighbours: Homosexuals	Against Neighbours: Homosexuals
	If Jobs scarce: priority to (nation) people	If Jobs scarce: priority to (nation) people
	Not Justifiable: homosexuality	Not Justifiable: homosexuality
	Against Neighbours: Drug addicts	Against Neighbours: Drug addicts
	Not Justifiable: abortion	Not Justifiable: abortion
	Not Justifiable: divorce	Not Justifiable: divorce
	No Confidence: Justice System/Courts	No Confidence: Justice System/Courts
		Type 4
		Not Justifiable: cheating on taxes
		Not Justifiable: claiming government benefits
		Not Justifiable: avoiding a fare on public transport
		Not Justifiable: someone accepting a bribe
		Justifiable: homosexuality
		No problem Neighbours: Homosexuals
		No problem Neighbours: People AIDS
		Justifiable: divorce
		No Confidence: Major Companies
		Competition is good

Notes: This table reports the 10 most important features for a n-type Factor Analysis model, where $n \in \{2, 3, 4\}$.

Table 15: Hierarchy of Types (Top Ten Features) as created by k-means

2 Type Model	3 Type Model	4 Type Model
Type 1	**Type 1**	**Type 1**
Not Justifiable: someone accepting a bribe	No problem Neighbours: People different race	No problem Neighbours: People different race
No problem Neighbours: People different race	No Confidence: Parliament	No problem Neighbours: Homosexuals
No Confidence: Parliament	Not Justifiable: someone accepting a bribe	Confidence: Police
No problem Neighbours: Homosexuals	No problem Neighbours: Homosexuals	No problem Neighbours: Immigrants/foreign workers
No problem Neighbours: Immigrants/foreign workers	No problem Neighbours: People AIDS	Not Justifiable: someone accepting a bribe
Not Justifiable: claiming government benefits	No problem Neighbours: Immigrants/foreign workers	No problem Neighbours: People AIDS
No problem Neighbours: People AIDS	No Confidence: Civil Services	Proud of nationality
Proud of nationality	No Confidence: Major Companies	Not Justifiable: claiming government benefits
No Confidence: Civil Services	Not Justifiable: claiming government benefits	Not Justifiable: cheating on taxes
Not Justifiable: cheating on taxes	No Confidence: The Press	Confidence: Justice System/Courts
Type 2	**Type 2**	**Type 2**
Proud of nationality	No problem Neighbours: People different race	Proud of nationality
Not Justifiable: someone accepting a bribe	Confidence: Police	Not Justifiable: someone accepting a bribe
Confidence: Police	No problem Neighbours: Homosexuals	Confidence: Police
No problem Neighbours: People different race	No problem Neighbours: Immigrants/foreign workers	Not Justifiable: cheating on taxes
Not Justifiable: cheating on taxes	Proud of nationality	Not Justifiable: claiming government benefits
Not Justifiable: claiming government benefits	Not Justifiable: someone accepting a bribe	Not Justifiable: avoiding a fare on public transport
No problem Neighbours: Immigrants/foreign workers	No problem Neighbours: People AIDS	Not Justifiable: suicide
Not Justifiable: avoiding a fare on public transport	Not Justifiable: claiming government benefits	No problem Neighbours: People different race
Confidence: Armed Forces	Not Justifiable: cheating on taxes	Confidence: Armed Forces
No problem Neighbours: Homosexuals	Confidence: Justice System/Courts	Not Justifiable: prostitution
	Type 3	**Type 3**
	Not Justifiable: someone accepting a bribe	Not Justifiable: someone accepting a bribe
	Proud of nationality	Proud of nationality
	Not Justifiable: cheating on taxes	Not Justifiable: claiming government benefits
	Not Justifiable: suicide	Not Justifiable: cheating on taxes
	Not Justifiable: claiming government benefits	Not Justifiable: avoiding a fare on public transport
	Not Justifiable: avoiding a fare on public transport	No Confidence: Parliament
	Not Justifiable: prostitution	Not Justifiable: suicide
	No problem Neighbours: People different race	No problem Neighbours: People different race
	Confidence: Police	Not Justifiable: prostitution
	Not Justifiable: abortion	No Confidence: The Press
		Type 4
		No problem Neighbours: Homosexuals
		No problem Neighbours: People different race
		No Confidence: Parliament
		No problem Neighbours: People AIDS
		No problem Neighbours: Immigrants/foreign workers
		Not Justifiable: someone accepting a bribe
		No Confidence: Civil Services
		No Confidence: Churches
		No Confidence: Major Companies
		Not Justifiable: claiming government benefits

Notes: This table reports the 10 most important features for a n-type k-means model, where $n \in \{2, 3, 4\}$.

F Additional Details on Topic Cohesion

F.1. Automatic Evaluation of Topic Model Cohesion

The main theme of the literature on the cohesion of topic models is that humans judge topics to be more consistent on the basis of word co-occurrence (Chang et al., 2009; David Newman et al., 2010; Lau et al., 2014; Lau and Baldwin, 2016). Consider for example, a topic containing words like 'labor', 'wage' and 'firm', which often appear together in a text, will be judged as highly coherent by humans. An alternative topic that contains words like 'inflation', 'agriculture' and 'hospital' appears incoherent, since these words are not used together as frequently.

Given this approach, it is possible to automatically calculate measures of topic cohesion that are highly correlated with human judgment. These measures are usually based on the most frequently occurring words in each topic. One standard approach is to calculate how often words appear together using the Wikipedia corpus (David Newman et al., 2010). The title and sub-sections of the Wikipedia article are used as 'tags' for discrete, human-curated topics. The more frequently that words within an LDA-derived topic appear together in a Wikipedia article (or within a sub-section of an article) then the more coherent the automatically defined topic is judged to be.

In our specific case of using survey response data, there is no equivalent, human curated outside corpus available to guide analysis. We therefore take the approach of using hold-out samples from within our data to calculate the cohesion scores. Our method thereby exploits the same intuition normally used in the literature on topic model cohesion. The key here is the β issue-position weights can be used as predictions of feature co-occurrence in the hold-out data. A political ideology is judged to be more coherent, if people frequently hold issue-position's together. We use Normalized Pointwise Mutual Information (NPMI) as our score of topic cohesion since NPMI has been shown to outperform other information metrics such as PMI or Pairwise Log Conditional Probability (LCP) and shows similar performance to pairwise distributional similarity (Aletras and Stevenson, 2013; Lau et al., 2014).

F.2. Making Sense of the NPMI Values

The calculation of the NPMI is based on the independent and joint probabilities of given features i and j. The probability $p(i)$ for example could capture the share of the population that believes abortion is not justifiable, while $p(j)$ captures the probability that a person has confidence in the church. The joint probability $p(i,j)$ then captures how many people believe that abortion is unjustifiable and have confidence in the church at the same time. The larger the joint $p(i,j)$ is in relation to $p(i)$ and $p(j)$, the higher is NPMI score of the two features.

Re-capping the basic equation from the main paper NPMI is defined as:

$$NPMI_{i,j} = \frac{PMI_{i,j}}{-\ln(p(i,j))} = \frac{\ln\left(\frac{p(i,j)}{p(i)\cdot p(j)}\right)}{-\ln(p(i,j))} \tag{9}$$

As an illustration, Table 16 shows two examples of NPMI scores for different values of $p(i)$ and $p(j)$, as well as different joint probabilities $p(i,j)$. In the first example both features appear with a probability of 0.2. In the situation where all people who are against abortion also have confidence in the church the joint probability of the features is 0.2 and the NPMI value will be 1. If the two features were independent of each other one would expect them to appear together in the data with a frequency of $(0.2 \cdot 0.2) = 0.04$. In this situation the calculated NPMI will be 0. If the joint probability is larger than the probability in the case of independence then NPMI will be positive, with the converse applying. The final two rows of Example 1 in Table 16 illustrate this relationship.

A technical point to note here is that the exact value of the NPMI depends on the individual as well as the joint probabilities. This is illustrated via the second example reported in Table 16. Note that in both Example 1 and Example 2 the third row cases are characterized by joint probability that is 50% larger than in the case of independence. The PMI is identical across the two different 'third row' cases but the NPMI is different. Two pairs of feature will only have the same NPMI if $log_{p(i,j)}(p(i), p(j)) = log_{p(x,y)}(p(x), p(y))$. In other words, the NPMI is identical, if you have to raise the joint probability to same power to get the product of the individual probabilities.

Table 16: Example Calculation NPMI

Case	$p(i)$	$p(j)$	$p(i,j)$	PMI	$NPMI$
Example 1					
Perfect Co-Occurrence	0.2	0.2	0.2	1.609	1
Independence	0.2	0.2	0.04	0	0
$p(i,j) > Independence$	0.2	0.2	0.06	0.405	0.244
$p(i,j) < Independence$	0.2	0.2	0.02	-0.693	-0.177
Example 2					
Perfect Co-Occurrence	0.6	0.6	0.6	0.511	1
Independence	0.6	0.6	0.36	0	0
$p(i,j) > Independence$	0.6	0.6	0.54	0.405	0.658
$p(i,j) < Independence$	0.6	0.6	0.18	-0.693	-0.404

G Additional Details on the Polarization Measure

The Esteban and Ray (1994) measure of polarization is based on three axioms. These three axioms aim to capture sensible assumptions about how own-group identification and out-group alienation contribute to an overall index of polarization.

Figure 9 illustrates the three axioms of Esteban and Ray (1994) graphically. The first axiom states that polarization increases if two small masses b and c that are close to each other are joined at their midpoint (see panel (a) of Figure 9). The intuition behind this axiom is that the joining of the masses increases the own-group identification of the now joined smaller masses, while the average distance and out-group alienation with respect to other major societal group a stays unchanged.

The second axiom states that polarization increases if a small mass of people b moves closer to the side of the spectrum where fewer people are concentrated (see panel (b) of Figure 9). Put simply, this change increases polarization because while the mass b has moved closer to group c it has also moved further away from another group a. Since mass a is larger than mass c, the overall alienation effect increases.

The third axiom states that polarization increases if mass is shifted equally from a central mass b to two lateral masses a and c that are each equally far away from the central mass (see panel (c) of Figure 9). This axiom captures the effect of the disappearing center. If mass shifts equally from the center to the fringes of the spectrum the own-group identification at the fringes increases while the overall out-group alienation increases as well.

Esteban and Ray (1994) prove that any measure of polarization that fulfills these three axioms must be of the form:

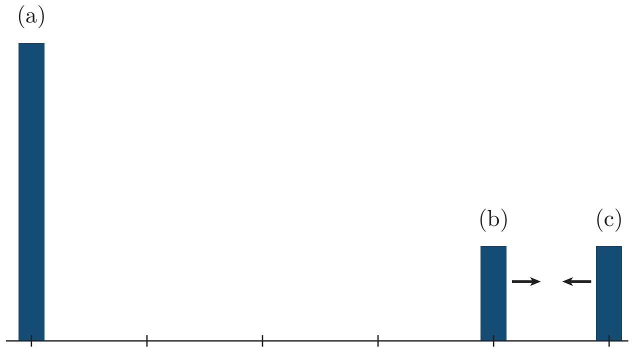
$$P(\pi, y) = K \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_i^{1+\alpha} \pi_j |y_i - y_j| \qquad (10)$$

The axioms hold for values of $\alpha \in [0, 1.6]$. For $\alpha = 0$ the polarization measure will be identical to the Gini coefficient. The sensitivity parameter α also influences the maximal possible value of the polarization measure. Esteban and Ray (1994) suggest a potential fourth axiom that would make it possible to narrow the possible interval of $\alpha \in [1, 1.6]$.

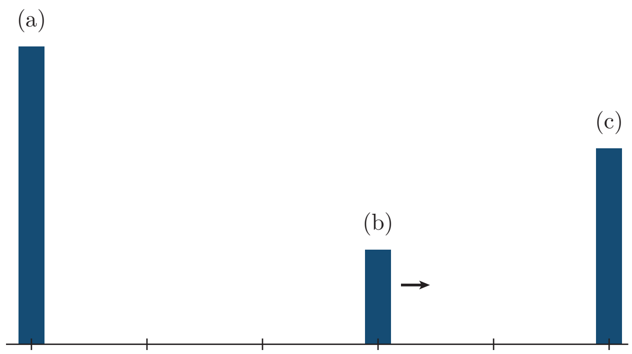
This fourth axiom is illustrated in Figure 10. The axiom states that moving mass from a small mass a to a larger mass c will increase polarization. Hence, the axiom makes an assumption on the importance of small groups within a society. On the one hand moving mass from a to c reduced the distance between the groups and therefore lowered polarization.

Figure 9: Axioms of Esteban & Ray 1994

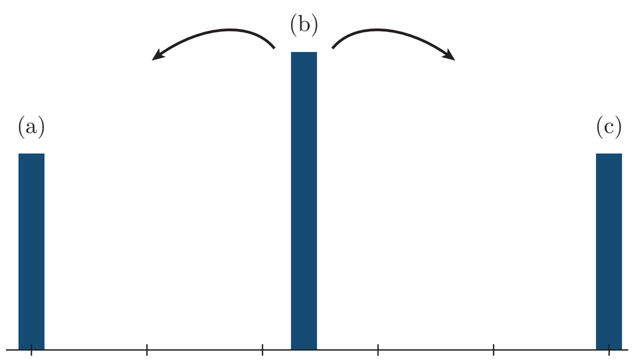
(a) Axiom 1



(b) Axiom 2



(c) Axiom 3

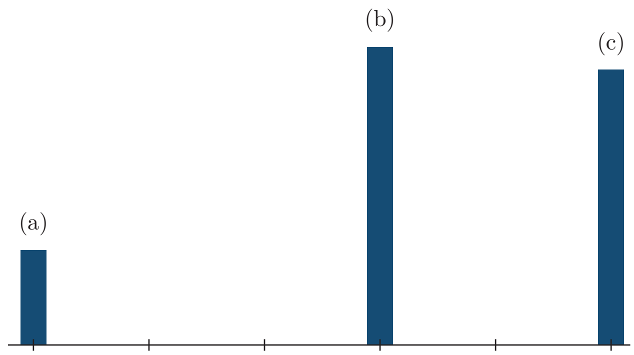


Notes: This figure illustrates the 3 main axioms use in Esteban and Ray (1994) to derive the polarization measure.

On the other hand the mass a is small in comparison to b and c and hence the effect of group a for overall polarization might be negligible, while increasing the mass of c can increase societal tension.

The polarization sensitivity parameter α here captures the relative sizes of a and c for which polarization will increase. The larger is α the smaller is the importance of a for overall polarization. It is a priori not clear whether this axiom is sensible in our context. Hence, we do not restrict the range of polarization sensitivity to $\alpha > 1$.

Figure 10: Additional Axiom of Esteban & Ray 1994



Notes: This figure illustrates the 4th axiom suggested in Esteban and Ray (1994). This axioms is not necessary to derive the form of the polarization measure but it allows for restrictions to the possible range of α.

G.1. Extending the Esteban and Ray (1994) Measure to Higher Dimension

The Esteban and Ray (1994) measure was originally constructed for a one-dimensional indicators (e.g. the income distribution). Our measure extends the measure to the four dimensions of our ideological type space. We assume that an individual identifies with groups based on his or her dominant type share, since in our model the four ideological types are the most natural line for group delineations.

Theoretically, it would also be possible to define groups based on discrete intervals of the type share distribution, such that a type would be defined by a specific interval in the four-dimensional ideological type space (e.g. [0,0.1] Liberal Centrist, [0.2,0.3] Conservative Centrist, [0.4,0.5] Left and Right Anarchist). This would obviously lead to a far greater number of ideological groups. The problem with this approach is that it is not obvious to decide on an interval length such that we can plausibly assume sufficient degrees of separation between these groups.

If the groups are defined by the dominant type share of each individual it is intuitive that the alienation between these groups will be based on differences in type shares. The only alteration to the original measure then is the fact that in our case the groups can differ along four dimensions rather than a single variable y. We hence define the overall out-group alienation as the sum of the type share differences between different groups.

G.2. Robustness Ray Measure

So far we have not addressed the question of the choice of α. As explained above any $\alpha \in [0, 1.6]$ leads to a measure of polarization that fulfills the axioms of Esteban and Ray (1994). As a robustness exercise we calculate the Esteban-Ray measure for several values of α. Table 17 reports the ranking our countries by their polarization over the three waves conditional on the choice of α. It is important to note that the values of the polarization measure are not comparable across different α, since dependent on α the maximal possible polarization level varies.

Our main finding for the rising level of polarization in the US holds for all except the largest values of α. As long as $\alpha < 1$ the US emerges as the most polarized country in our sample. The results for $\alpha = 1.6$ differ, since for high values of α the importance of small groups in society is diminished. Hence, in this case the polarization P measure for the US - where we observe four comparably sized ideological groups - is lower than for other values of α. In contrast, measured polarization is higher in countries with one large ideological group, e.g. the Conservative Centrist in Malta or Liberal Centrist in Denmark.

Overall, the results seem to point towards the fact that values of $\alpha < 1$ lead to a more balanced polarization ranking across countries. The fact that for $\alpha = 1.6$ countries such as Denmark, Iceland, Finland and Canada - all of which are usually considered harmonious societies - end up at top of the ranking seems counterintuitive. Based on these findings we set $\alpha = 0.5$ as the baseline value for polarization sensitivity in our main P measure.

Table 17: Ray Measure for different α

Panel A: Wave 2

$\alpha = 0$		$\alpha = 0.5$		$\alpha = 1$		$\alpha = 1.6$	
Country	Ray Measure	Country	Ray Measure	Country	Ray Measure	Country	Ray Measure
France	0.871	Austria	0.449	Malta	0.291	Malta	0.184
Spain	0.871	Spain	0.446	North Ireland	0.260	North Ireland	0.146
Belgium	0.867	France	0.445	Portugal	0.253	Portugal	0.146
Austria	0.838	Belgium	0.441	Austria	0.250	Ireland	0.139
Germany	0.835	North Ireland	0.439	Netherlands	0.246	Netherlands	0.138
Italy	0.834	Great Britain	0.436	United States	0.244	United States	0.128
Netherlands	0.830	Malta	0.434	Ireland	0.240	Austria	0.128
Great Britain	0.825	Netherlands	0.433	Canada	0.239	Iceland	0.127
Canada	0.806	Italy	0.431	Iceland	0.238	Denmark	0.127
North Ireland	0.795	Germany	0.429	Great Britain	0.236	Canada	0.125
Finland	0.778	Canada	0.428	Spain	0.234	Great Britain	0.116
Iceland	0.776	United States	0.426	France	0.233	Italy	0.112
United States	0.770	Iceland	0.418	Italy	0.230	Finland	0.112
Portugal	0.729	Portugal	0.415	Belgium	0.228	Spain	0.111
Ireland	0.719	Finland	0.404	Germany	0.226	France	0.110
Malta	0.690	Ireland	0.400	Finland	0.219	Germany	0.108
Denmark	0.652	Denmark	0.357	Denmark	0.214	Belgium	0.106

Panel B: Wave 4

$\alpha = 0$		$\alpha = 0.5$		$\alpha = 1$		$\alpha = 1.6$	
Country	Ray Measure	Country	Ray Measure	Country	Ray Measure	Country	Ray Measure
Spain	0.951	Spain	0.476	Malta	0.309	Malta	0.201
Austria	0.911	Austria	0.469	North Ireland	0.256	Iceland	0.161
France	0.896	Great Britain	0.466	United States	0.252	Denmark	0.158
Great Britain	0.894	United States	0.459	Ireland	0.251	Netherlands	0.142
Germany	0.889	France	0.456	Great Britain	0.250	Canada	0.135
Belgium	0.881	Germany	0.454	Canada	0.249	North Ireland	0.135
Italy	0.873	Italy	0.453	Austria	0.247	Ireland	0.133
United States	0.867	Malta	0.453	Finland	0.243	Finland	0.133
North Ireland	0.835	North Ireland	0.451	Italy	0.243	United States	0.126
Canada	0.826	Belgium	0.442	Iceland	0.243	Portugal	0.125
Ireland	0.813	Ireland	0.441	Netherlands	0.241	Great Britain	0.122
Finland	0.801	Canada	0.439	Spain	0.239	Italy	0.120
Portugal	0.799	Finland	0.427	France	0.238	Austria	0.117
Netherlands	0.765	Portugal	0.422	Portugal	0.236	France	0.113
Malta	0.699	Netherlands	0.409	Germany	0.236	Germany	0.109
Iceland	0.669	Iceland	0.377	Belgium	0.222	Spain	0.104
Denmark	0.578	Denmark	0.326	Denmark	0.220	Belgium	0.098

Panel C: Wave 5

$\alpha = 0$		$\alpha = 0.5$		$\alpha = 1$		$\alpha = 1.6$	
Country	Ray Measure	Country	Ray Measure	Country	Ray Measure	Country	Ray Measure
United States	0.911	United States	0.475	Malta	0.267	Malta	0.173
Netherlands	0.895	Netherlands	0.459	United States	0.255	Denmark	0.155
Austria	0.895	Canada	0.452	Canada	0.247	Iceland	0.142
Spain	0.880	Austria	0.450	Finland	0.244	Finland	0.141
Germany	0.873	Ireland	0.446	Netherlands	0.244	Canada	0.125
Canada	0.863	Spain	0.445	Ireland	0.243	North Ireland	0.125
Great Britain	0.862	Great Britain	0.442	North Ireland	0.240	United States	0.125
France	0.856	Germany	0.439	Portugal	0.236	Ireland	0.122
Ireland	0.853	France	0.438	Great Britain	0.234	Portugal	0.121
Belgium	0.836	North Ireland	0.430	France	0.231	Netherlands	0.119
Italy	0.822	Belgium	0.428	Spain	0.229	France	0.112
North Ireland	0.811	Italy	0.425	Italy	0.228	Great Britain	0.112
Portugal	0.800	Portugal	0.425	Austria	0.228	Italy	0.111
Finland	0.778	Finland	0.416	Belgium	0.225	Belgium	0.107
Iceland	0.660	Malta	0.399	Germany	0.221	Spain	0.104
Malta	0.643	Iceland	0.353	Iceland	0.219	Austria	0.102
Denmark	0.509	Denmark	0.289	Denmark	0.203	Germany	0.098

Notes: This table reports the polarization measure for different α. For more details see text.