

Endogenous Partitions*

Klaus Desmet
SMU, NBER and CEPR

Ignacio Ortuno-Ortín
UC3M

Romain Wacziarg
UCLA and NBER

April 2023

Abstract

We develop a new method to endogenously partition society into cultural groups, without reference to predefined identity markers. To do so, we develop a measurement framework that divides people into distinct groups based on their answers to surveys of values, norms and attitudes. The method chooses divisions in order to minimize within-group antagonism, where antagonism is increasing in an individual's distance to other members of their group in answer space. We implement the method using the World Values Survey across 7 waves and 81 countries, with a particular focus on the US. We document several facts, among which: 1) A substantial reduction in within-group antagonism is obtained by partitioning society into two or three groups. 2) The reduction in within-cluster antagonism based on endogenous cultural partitions is an order of magnitude larger than that obtained based on predefined identity traits, such as gender. 3) In the US, partitions are largely driven by differences in religious and moral values - and cross-cluster differences in these values have increased over time. 4) In the US, cross-cluster differences in political ideology are the largest in our sample of countries, and have been increasing over time. We discuss the possible uses of our new measurement framework to study a range of questions in political economy.

*Desmet: Department of Economics and Cox School of Business, Southern Methodist University, 3300 Dyer, Dallas, TX 75205, kdesmet@smu.edu; Ortuno-Ortín: Department of Economics, Universidad Carlos III de Madrid, 28903 Getafe, Madrid, Spain, iortuno@eco.uc3m.es; Wacziarg: UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles CA 90095, wacziarg@ucla.edu. We thank seminar and conference participants at Bocconi University, FGV São Paulo and Berkeley-Haas, as well as Enrico Spolaore, Francesco Trebbi, Clémence Tricaud and Luigi Zingales, for useful comments.

1 Introduction

There is growing concern that disagreement on social and moral issues undermines the functioning of modern societies. In the United States, many scholars and commentators have argued that differences in values are growing, leading to social conflict, political polarization, and personal animosity. The rise of populist movements around the world has similarly created concerns over conflict between groups separated by differences in a broad range of values - including those related to religion, nationalism, redistribution, and democracy. The usual approach to societal divisions is to define them based on exogenous identity traits - such as gender, race, ethnicity or language. However, identity traits are not very predictive of cultural values, and there is a lot of cultural diversity within identity groups (Desmet, Ortuño-Ortín & Wacziarg, 2017, Desmet & Wacziarg, 2021).

In this paper, we depart from the usual approach: we consider values and norms of individuals as primitives for the construction of social divisions. Indeed, values, norms and attitudes are the building blocks of social and political preferences. Heterogeneity in values reflects the degree of social disagreement, and the evolution of values can lead to changes in the degree of observed societal antagonism and polarization. Clusters of individuals formed on the basis of differences in values also determine visible social partitions, such as interest groups and political parties. The goal of this paper is to shed light on the nature and evolution of cultural partitions.

Our approach begins directly from answers to survey questions on values, norms and attitudes. We develop a novel methodology to create natural partitions between respondents based on their answers. Partitions are obtained by assuming that agents within a group prefer to interact with others that share more similar values (homophily): they experience more antagonism when grouped with agents that have more different vectors of values from their own. Antagonism is defined as a function of differences in values between agents. Agents then choose the group where their expected antagonism is lowest. An equilibrium is a partition such that no agent would prefer to join another group. The corresponding endogenous partition of the set of respondents is characterized by a certain degree of within-group antagonism. There are many such equilibria, and we focus on the one that minimizes overall within-group antagonism. Such an equilibrium is also the one that maximizes between-group antagonism.

We implement our method using data from the seven waves of the integrated European Values Survey - World Values Survey (EVS-WVS), for a cross-section of 81 countries (focusing in particular on the United States). We consider a set of about 200 questions reflecting respondents' values. We first implement our algorithm for the case of two groups. We create the endogenous partitions and then study the demographic characteristics of the agents that form the two groups, as well as the average cultural characteristics of the two groups. We find that religious identity is an often important characteristic setting apart the groups, especially in the US. Yet in general, endogenously-formed cultural partitions feature a lot of heterogeneity in identity space: cultural groups do not clearly overlap with identity groups. We also document the qualitative differences in values between groups by summarizing a small set of responses for different groups and different countries. Not surprisingly, groups differ extensively along such cultural dimensions.

Our main finding is that the reduction in societal antagonism resulting from the creation of cultural

partitions is an order of magnitude larger than the reduction in antagonism stemming from creating identity-based groups, such as those based on gender or educational attainment. For instance, when creating two cultural partitions, the average reduction in antagonism is 41.51%. In contrast, the average reduction from creating gender partitions is 1.29% and the average reduction based on creating two groups that differ by educational attainment is 2.96%. We are, of course, not claiming that gender or education are not socially relevant dimensions of heterogeneity, but rather that the partitioning of society based on values and norms also has potential explanatory power to understand contemporary social divisions.

We document many other interesting facts. One set of findings pertains to variation across countries. In most countries, a substantial reduction in antagonism is achieved with a small number of cultural clusters. These gains tend to peter out quite rapidly when allowing for more than three cultural groups: most gains are achieved by the creation of the first two or three groups. The types of values that are most important in the formation of clusters differs substantially across countries. For example, in the US, clusters are formed mostly based on religious beliefs and attendance, and moral values. There, economic preferences do not play a big role in defining endogenous partitions. In contrast, in Nigeria, groups are formed mostly on the basis of differences in confidence in government and institutions more generally, as well as preferences over political action and satisfaction with the political system.

Another set of findings pertains to changes across time. Here we focus on the US. We document a recent increase in the degree of between-group antagonism, going hand in hand with an increase in cultural cohesion within partitions. The growing group distinctiveness is driven in large part by values reflecting religious beliefs, moral values, religious attendance, and social capital (trust, membership in religious organization, etc.). Finally, the endogenous partitions exhibit increasingly different political ideologies on a right-left scale. In fact, in the latest wave of the WVS, when considering a two-group partition, the ideological difference between these clusters is larger than in any other country in our sample.

Our paper relates to a vast literature conceptualizing and measuring social heterogeneity. One strand of this literature focuses on measuring heterogeneity based on identity markers, such as ethnicity (Alesina et al., 2003). As noted earlier, a limitation of this approach is that there is considerable heterogeneity in cultural values within ethnic groups, so identity-based partitions may not be the most relevant dimensions of heterogeneity (Desmet, Ortuño-Ortín and Wacziarg, 2017). In fact, there is considerable within-group heterogeneity in values whatever the identity metric under considerations, be it gender, race, ethnicity, education levels, income quintiles, etc. (see Desmet and Wacziarg, 2021, for the US case). This motivates a direct focus on values as a foundation for measuring cultural heterogeneity, an approach that we adopted in our own past work, that was pursued in Alesina, Tabellini and Trebbi (2017) and that we continue to follow here. Our approach to deriving endogenous social partitions is also related to the literature on endogenous country formation (Alesina and Spolaore, 1997), as well as to the literature on voting with your feet in public finance (Tiebout, 1956). Here we do not consider secession *per se* but the division of each society into cultural groups based on value-based affinity between individuals. This paper is also related to the vast literature on the changing nature of political cleavages around the world (salient recent examples include Gethin, Martinez-Toledano and

Piketty, 2021 and Bonomi, Gennaioli and Tabellini, 2021). This literature pays a lot of attention to recent changes in voting behavior, party platforms and politically salient cleavages, but devotes less attention to the nature and evolution of the individual values that underlie these changes. The latter is our main focus here. Finally, our paper is related to the literature on endogenous party formation, where party platforms reflect voter preferences over policy (Baron, 1993, Ortuño-Ortín and Roemer, 2000, Gomberg, Mahruenda and Ortuño-Ortín, 2004).

2 Conceptual Framework

In this section, we propose a conceptual micro-founded framework for the formation of values-based partitions. We also compare values-based partitions to identity-based partitions.

2.1 Partitions Based on Values

There is a set P of N individuals in society. Each individual j is characterized by a "values" vector $x_j \in R^Q$. Each individual is also characterized by a socio-demographic type or "identity" vector $f_j \in R^S$, containing information about her gender, race, education, religious denomination, political affiliation, income, and so on.

A large literature on homophily argues that individuals tend to associate disproportionately with others who are similar to them (Verbrugge, 1977; McPherson, Smith-Lovin and Cook, 2001; Jackson, 2021). In our context, similarity could be defined based on either values or identity. In what follows, we analyze the case of homophily on values. That is, individuals prefer to interact with other individuals who have similar values. More specifically, an individual j experiences a disutility from interacting with individual $k \in P$ that is increasing in the distance between the vector of values x_j and the vector of values x_k :

$$u(x_j, x_k) = u(d(x_j, x_k)) \tag{1}$$

where $d(x_j, x_k)$ is a distance metric between x_j and x_k .

Values Identification Equilibrium (VIE). When an individual chooses the group she wants to interact with, she minimizes the expected disutility from this interaction. We assume that there is a limited number of groups in society. For now, we set this number to two, and each individual associates or identifies with only one of the two groups.¹ These groups are not exogenously defined, but are endogenously formed by individuals aiming to minimize their expected disutility from interacting or identifying with others.

Consider a partition A of the population set P into two groups, group A_1 and group A_2 , with $P = A_1 \cup A_2$ and $A_1 \cap A_2 = \emptyset$. Let \mathcal{P} denote the set of all possible partitions of P into two groups, so that $A \in \mathcal{P}$. The expected disutility that an individual with values x_j experiences when identifying

¹In the empirical section, we also analyze the case of more than two groups. Another conceptually straightforward extension would be to consider individuals probabilistically identifying with more than one group.

with group A_i , and interacting with each individual of that group with equal probability, is:

$$E(A_i, x_j) = \frac{1}{|A_i|} \sum_{k \in A_i} u(x_k, x_j) \quad (2)$$

We refer to this expected disutility as the *antagonism* individual j experiences by identifying with group A_i . When deciding which group to identify with, individual j chooses group A_i over group A_{-i} if $E(A_i, x_j) \leq E(A_{-i}, x_j)$.

Definition of a Values Identification Equilibrium (VIE). A partition $A \in \mathcal{P}$ is a Values Identification Equilibrium (VIE) if for each group $A_i \in A$, $i = 1, 2$, and for each agent $j \in A_i$ we have $E(A_i, x_j) \leq E(A_{-i}, x_j)$.

Thus, a VIE is a Nash equilibrium: taking as given the identification of all other individuals, no agent wants to change her identification. One can easily prove the existence of an equilibrium, because there is a finite number of individuals.² However, there may be multiple VIEs. We denote the set of possible VIEs by \mathcal{V} .

For any VIE $A \in \mathcal{V}$, we define social antagonism as the average within-group antagonism in society:

$$E(A) = \frac{1}{N} \sum_{j \in P} E(A_{i(j)}, x_j) \quad (3)$$

where $A_{i(j)}$ denotes the group agent j belongs to. Within the set of possible VIEs, we refer to the one that minimizes average within-group antagonism as the Global VIE, and denote that VIE by A^* .

Definition of the Global VIE. A^* is the Global Values Identification Equilibrium if for each $A \in \mathcal{V}$, $E(A) \geq E(A^*)$.

A social planner wishing to choose the partition that minimizes social antagonism would choose the Global VIE. We denote the level of social antagonism in the Global VIE by $E^* = E(A^*)$.

Partitions as a way to reduce social antagonism. How much does social antagonism decline when partitioning individuals into groups? We define a society's potential antagonism as the social antagonism when everyone belongs to the same group (i.e., there is only one group P) and all individuals interact with everyone else in society with equal probability:

$$E(P) = \frac{1}{N} \sum_{j \in P} E(P, x_j) = \frac{1}{N} \sum_{j \in P} \frac{\sum_{k \in P} u(x_k, x_j)}{N} \quad (4)$$

Potential antagonism $E(P)$ is the expected distance between two randomly chosen individuals in society. It is the same as the Greenberg B index of diversity.

To measure the reduction in social antagonism when society is partitioned into two groups, we use the ratio

$$r = \frac{E(P) - E^*}{E(P)} \quad (5)$$

²Gomberg, Marhuenda and Ortuño-Ortín (2004) provide conditions for the existence of an equilibrium in the case of a continuum of agents.

That is, when individuals go from interacting with everyone else in society to interacting only with individuals of their group, social antagonism drops by a proportion r . If r is high, it is efficient to partition society into two groups: even though potential antagonism $E(P)$ in society could be high, there is a way to divide individuals into homogenous groups that achieves a large reduction in social antagonism (with two groups, one would achieve a 100% reduction if there are two cultural types of individuals).

The reduction in social antagonism r is equivalent to the well-known Φ_{ST} index of between-group differentiation:

$$\Phi_{ST} = \frac{E(P) - \sum_{i=1}^2 \frac{|A_i|}{N} E(A_i)}{E(P)} \equiv r \quad (6)$$

where $E(A_i)$ denotes social antagonism within group A_i . This result is proven in Appendix A1A.

Differences in identity and values across groups. Given a VIE, we can analyze to what extent the identity traits of individuals belonging to each group differ. For example, we can compare the share of women in A_1 and the share of women in A_2 . Similar shares would suggest that the values that most contribute to the partition in a given VIE are not strongly associated with gender. Similarly, given a VIE, we can analyze differences in cultural values across groups. For example, we can compare whether there is a large or a small difference between groups in the share of individuals who believe in God. This tells us which values contributed most to partitioning society into groups.

2.2 Partitions Based on Identity

Instead of partitions that minimize social antagonism, suppose that partitions are based purely on identity traits. Take, for instance, the case of a partition based on gender. In such a partition $A^g \in \mathcal{P}$, the first group, A_1^g , contains only men, and the second group, A_2^g , contains only women. The utility function of an individual is unchanged. That is, an individual's utility depends on the distance between his values and those of the individuals he interacts with. Gender does not enter the utility function, and social antagonism is still solely based on differences in values across individuals. In general, the gender-based partition A^g is not a VIE.

The degree of social antagonism associated with a gender-based partition is then:

$$E(A^g) = \frac{1}{N} \sum_{j \in P} E(A_{i(j)}^g, x_j) \quad (7)$$

Thus, $E(A^g)$ measures social antagonism in the case where individuals still only care about values, but where men and women are partitioned into two different groups. As before, we can measure by how much social antagonism decreases when partitioning society into women and men:

$$r^g = \frac{E(P) - E(A^g)}{E(P)} \quad (8)$$

That is, when individuals go from interacting with everyone else in society to interacting only with individuals of their gender, social antagonism drops by a proportion r^g . This is related to the intensity of the gender cleavage as measured in Desmet and Wacziarg (2021).

2.3 Comparing Identity-Based Partitions to Values-Based Partitions

How can we measure the efficiency of partitioning society by identity traits? One approach is to simply measure by how much social antagonism declines when splitting up society into, say, men and women. This is what (8) measures. Another, arguably better, approach is to measure the efficiency of identity-based partitions in terms of opportunity cost. That is, how much larger is social antagonism under an identity-based partition compared to a values-based partition? Using the example of gender, the opportunity cost of a gender-based partition is:

$$OC^g = \frac{E(A^g) - E^*}{E^*} \quad (9)$$

To illustrate how to interpret this measure in the context of our framework, suppose the opportunity cost of a gender-based partition is high. One interpretation is that individuals who care about cultural values would be unlikely to choose their gender as a main way to socially identify. Indeed, socially identifying with one’s gender would imply interacting with individuals who differ greatly on the values they hold, compared to the alternative of choosing one’s group based directly on values. Another interpretation is that instrumentalizing gender to garner political support is costly. Compared to mobilizing individuals based on their values, political coalitions based on gender would be much more heterogeneous in their underlying views. These interpretations are subject to the caveat mentioned before: in our framework identity traits do not directly enter into one’s utility function, and when interacting with others, people only care about similarity in values.

2.4 Squared Euclidean Distance and Interpretation of Values Identification

Squared Euclidean distance. So far, we have not assumed a functional form for $u(d(x_j, x_k))$. Here we assume that the disutility function (1) is given by:

$$u(x_j, x_k) = \|x_j, x_k\|^2 \quad (10)$$

where $\|x_j, x_k\|$ is the Euclidian distance between vector of values x_j and vector of values x_k .³ This is the distance metric used in our empirical application, but our framework can accommodate other distance metrics.

Using the squared Euclidean distance has several advantages. First, it allows us to use the standard k-means clustering method in order to create endogenous partitions. Second, it leads to measures of within-group and between-group heterogeneity that are closely related to those in Desmet, Ortuno-Ortín and Wacziarg (2017) and Desmet and Wacziarg (2021), as we show later. Third, adopting the squared Euclidean distance implies an additive preference structure (the distance between two vectors can be computed by adding up distances in each of the Q dimensions). Fourth, the antagonism individual j experiences by identifying with group A_i as given by (2) can be written as her distance to the mean values of the group:

$$E(A_i, x_j) = \frac{1}{|A_i|} \sum_{k \in A_i} \|x_k, x_j\|^2 = 2 \|x_j - \mu_i\|^2 \quad (11)$$

³This approach is related to that in Alesina, Tabellini and Trebbi (2017), p. 183. They compute bilateral distances between all respondents to the European Values Survey (among other surveys) using the squared Euclidian distance between vectors of individual answers. They then plot the densities of a monotonic function of these distances.

where $\mu_i = (\sum_{k \in A_i} x_k) / |A_i|$. Thus, in any VIE, an individual in group A_i is closer to the mean position of her group than to the mean position of the other group. In this case, the Global VIE A^* is the partition that minimizes total within-group variance. This is equivalent to the partition that maximizes total between-group variance, as we prove in Appendix A1B.

Alternative interpretation of values identification. Rather than an individual choosing the group that minimizes her bilateral disagreements with other individuals of the same group, under the squared Euclidean distance assumption, an individual chooses the group that minimizes her distance to the group’s mean position. As such, identification does not require an individual to know all bilateral distances to all other individuals, but only her distances to the mean positions of the groups. We can thus view the mean value μ_i as the representative culture of group i , with the cost for an individual with values x_j to identify with group i given by the distance $\|x_j - \mu_i\|^2$. This type of identification is closer to the one typically considered by economists (Akerlof and Kranton, 2000, Shayo, 2009, Bonomi, Gennaioli and Tabellini, 2019).

This interpretation of a VIE is akin to the way we might think about the formation of political platforms. In the realm of politics, μ_i could be viewed as the policy position of political party i . That position, or platform, depends on the political positions of its supporters. While there is no good general theory explaining how political parties aggregate or represent the preferences of their supporters, a reasonable assumption is that parties adopt the mean position of their supporters as their policy position, and conversely, that supporters pick the party whose platform is closest to their own values. This sort of political equilibrium is explored in Gomberg, Marhuenda & Ortuño-Ortín (2004).

3 Empirical Methodology

3.1 Implementation

We use data from all waves of the World Values Survey. Vector x_j is given by agent j ’s answers to questions on values. We first focus mostly on the set of 81 countries included in Wave 7 of the integrated WVS-EVS dataset, - paying particular attention to the United States. We find the VIE that minimizes within-group antagonism, A^* , and describe its properties. In a second step, we examine previous waves and characterize the evolution of endogenous partitions and their properties across time and countries, focusing on the US.

Selection of the sample. Consider a wave of the WVS and a given country. Suppose there are N_0 individuals and Q_0 relevant questions (in wave 7, Q_0 is 210 and N_0 ranges from 1,000 to 4,018 with a mean of 1,569).⁴ We proceed in the following steps:

1. We select the Q_0 questions that are about values and attitudes, with answers that can be ordered. The latter correspond to two types of questions: those that admit a binary answer and those

⁴Appendix A2 lists the WVS questions used in our analysis, for wave 7.

with answers ordered on a scale (we discard questions of a qualitative nature where answers do not admit a natural order).

2. We rescale each question so that the values of the possible answers are always in the interval $[0, 1]$.
3. We drop any question that is not answered by more than 30% of the individuals. Thus, we end up with a number of questions $Q \leq Q_0$.
4. We drop all the individuals who do not answer at least 70% of those Q questions. Thus we end up with a number of individuals $N \leq N_0$.

The last two filters are applied in order to obtain a sample that has the same number of questions per individual in each country, and the same number of respondents per question in each country. In practice, there is no need to eliminate too many questions or too many individuals. Indeed, the average value of Q is 193, with a minimum of 163 in the case of Egypt (the US features 198 questions, only dropping one). Similarly the average value of N is 1,558, with a range of 987 to 4,018. Thus, we eliminate on average 17 questions, and 0.7% of respondents. The resulting set of questions is very balanced across the countries in the sample, an attractive feature of using the WVS.

Missing answers. Let $x_j = \{x_{j1}, x_{j2}, \dots, x_{jQ}\}$ be the vector of all the answers of individual j , $j = 1, \dots, N$. Even after steps 3 and 4, there may still be a few missing answers in x_j . Let TS be set of individuals without missing answers, and $X_{TS} = \{x_j : x_j \in TS\}$. We use a machine learning algorithm to impute values to all missing responses.⁵ This algorithm is based on the training sample X_{TS} , and does not any use demographic / identity information on respondents. In this way, we obtain a data matrix $X = \{x_1, x_2, \dots, x_N\}$ with no missing values.⁶ In practice, in most countries the TS set contains a very high proportion of individuals, another nice property of the WVS (which does not hold to the same extent in the GSS, a US survey of values and attitudes).

Principal components analysis. The next step is to reduce the dimensionality of the answer space by using principal component analysis (PCA). The advantages are fourfold. First, using PCA avoids the unnecessary duplication of questions that captures similar values and are likely to be answered similarly by different respondents (for example, there are separate questions for belief in Heaven and in Hell - with very highly correlated answers). Second, by construction PCA produces dimensions that are orthogonal to each other, allowing an interpretation of the resulting measures of antagonism as minimizing within-cluster variance in answers (as captured by principal component positions - see

⁵We use the Mathematica (version 13.0.1) command "SynthesizeMissingValues" to replace missing values. The training sample used was formed by the answers given by the set individuals with no missing answers (TS). We set the level of performance to "Quality" to maximize the synthesis quality. For each country and wave, Mathematica chooses the best machine learning algorithm from among "Multinomial", "Kernel Density Estimation", "Decision Tree", and "Gaussian Mixture". In wave 7, the average number of individuals in the training sample was 582 (a minimum of 47 individuals in the case of New Zealand and a maximum of 2,723 for Canada).

⁶In rare cases, the algorithm can replace a missing value with a value that lies outside the $[0, 1]$ interval. In such cases, we assign a value of 0 or 1, depending on which is closest.

Appendix A1B). Third, with fewer dimensions, finding a VIE is computationally less costly. Fourth, to the extent there is measurement error in the way individuals answer WVS questions, the use of principal components helps mitigate the problem.⁷

For each country and wave, we compute the principal components of matrix $X = \{x_1, x_2, \dots, x_N\}$. For any individual j we write the vector of her position on the different PC dimensions as $p_j = \{p_{j1}, p_{j2}, \dots, p_{jQ}\}$.⁸ We can use either the answers themselves or any number of principal components to find VIEs. In practice, we do the latter, and consider alternatively the first, first two, first three and first 75 principal components to create the endogenous partitions.⁹ In the first case, the distance between individual j and individual k is given by:

$$u(p_{j1}, p_{k1}) = \|p_{j1}, p_{k1}\|^2 = (p_{j1} - p_{k1})^2 \quad (12)$$

In the second case the corresponding distance is:

$$u(\{p_{j1}, p_{j2}\}, \{p_{k1}, p_{k2}\}) = (p_{j1} - p_{k1})^2 + (p_{j2} - p_{k2})^2 \quad (13)$$

and so on for 3 and 75 dimensions.

Finding the VIEs and A^* . The next step consists of finding the best VIE equilibria for each country, A^* . We use the k-means clustering algorithm to find VIEs.¹⁰ The distance metric is the squared Euclidian distance, implying that in a VIE each individual is at a smaller distance from the mean answers of their own group than from the mean of any other group. We use the sampling weights of the WVS to ensure our underlying sample is nationally representative. This applies whenever there is aggregation between individuals, at all stages of the implementation of our clustering method.¹¹ We double-checked that the clusters obtained indeed constitute a VIE and that no individual prefers to be in a different cluster. We allow successively for 2, 3, 4 and 5 clusters.

Since we have 81 countries, we consider four different numbers of clusters, and we consider four different numbers of principal components, we need to find the best VIE (A^*) 1,296 times. For each country, number of clusters and number of principal components, the algorithm is run 1,000 times, each run starting from a different random point. Thus, for each country, we get several partitions, but in practice many of them are very similar, i.e. there are minor differences in the sets of individuals they include. We select the partition with the lowest antagonism level A , and we claim that this is

⁷We have verified that at least for wave 7 and the United States, the results obtained *without* first reducing the question dimensionality using PCA are very similar to those obtained using PCA, in the sense that we obtain very similar clusters in both cases.

⁸In all the cases, the number of questions is less than the number of individuals, so matrix $P = \{p_1, \dots, p_N\}$ has Q columns.

⁹The case of 75 PCs is basically equivalent to considering all PCs. We use 75 PCs instead of all of them because it is computationally less taxing.

¹⁰This is implemented in *Mathematica*.

¹¹In practice, only 52 of the 81 countries in our sample have WVS sample weights, For three of them (including the US) and for wave 7, we checked that the use of weights did not result in clusters that are different from those obtained without weights - nor does it modify our findings relating to the total level of within-group antagonism obtained from these clusters.

partition A^* . While we cannot be absolutely certain that this is the global minimum, it is our best candidate for a global minimum.

Computing descriptive statistics. The final step is to compute statistics describing the endogenous partitions. For each country, we compute E^* , PA and r , among other statistics. We also examine the demographic characteristics of each cluster - we consider gender, age, income deciles, years of education, political ideology, whether the respondent belongs to a religious denomination, and self-reported social class. We also characterize the cultural make-up of each cluster by examining the mean of a selection of cultural values for each partition.

3.2 Validation

We conducted two validation exercises to assess whether our methodology is likely to produce sensible results.

The first exercise begins with data from three countries: the United States, China and Zimbabwe (we chose these countries because they belong to different continents). We pooled the respondents from wave 7 of the WVS for these three countries. We then ran principal components analysis on this joint sample, finding that the first PC explains 13% of the variance in answers, while the first 10 PCs explain 40%. Plotting individuals along the first two PCs makes the three countries appear distinctly (Figure 1, Panel A). We next run our algorithm on the pooled data (Figure 1, Panel B), with three clusters. Our goal is to see if our algorithm recovers the three countries that underlie the pooled data. Indeed, we recover very accurately the three countries: 96.5% of the individuals from China are classified in cluster 1 (in blue in the figure), 98.2% of the individuals from Zimbabwe are classified as belonging to cluster 3 (in green in the figure) and 79.5% of the individuals from the US belong to cluster 2 (in orange in the figure). Most of the remaining US individuals (18.2%) are assigned to cluster 3 (the "Zimbabwe cluster").¹² Overall, our algorithm does a remarkable job at recovering the three underlying countries. Moreover, using common methods from cluster analysis to assess the optimal number of partitions in the pooled data, we find that the optimal number of partitions tends to be equal to either three or four (a fourth cluster typically divides the US sample into two further groups).¹³

The second exercise exploits WVS data from reunified Germany, from wave 3 (where the data were collected in 1997) and wave 7 (where the data were collected in 2018). Here we attempt to separately detect respondents from East and West Germany. Our hypothesis is that a decades-long separation under very different political regimes would have increased the degree of cultural distinctiveness between the two areas, and that a subsequent period of 21 years under reunification would have blunted such a distinctiveness, either because of cultural change or because of migration across the two regions. Given the expected cultural similarity between East and West Germans, we do not expect to

¹²This 18.2% of the Americans sample of respondents consists of 246 Whites, 58 Blacks, and 74 Hispanics. As a percentage of the total of each group in the US they are: 14% of White respondents, 27.6% of Black respondents, and 16% of Hispanic respondents.

¹³The optimality criteria we checked for include Silhouette, Calinski-Harabasz, Davies-Bouldin, Dunn, R2-Elbow, and Standard Deviation-Elbow. All yield either 3 or 4 as the optimal number of clusters.

be able to separate the two groups as neatly as in the previous validation exercise. Our goal here is instead to assess, in a comparison across time, whether East Germans are more easily detectable by our algorithm in 1997 compared to 2018. We find that this is indeed the case. Figure 2, Panel A plots all respondents from Germany in the dimensions of the first two PCs for wave 3, and Panel B does the same for wave 7. Red dots represent residents of the former GDR/DDR, and blue dots are residents of the former FRG/BRD.¹⁴ We find that, in both waves, East and West Germans differ mostly along the dimensions of the second PC, which contains many questions on politics and institutions. It is also apparent from Figure 2 that East Germans are much more bunched together in wave 3 than in wave 7. This is confirmed using our endogenous partitions methodology with two clusters: we find that, in wave 3 of the WVS, 65.2% of the individuals in cluster 2 live in East Germany, while only 27.9% of the individuals in cluster 1 live in East Germany. Individuals residing in the former East Germany therefore make up most of cluster 2 (similarly cluster 1 is predominantly composed of West Germans). This distinctiveness is much reduced in wave 7 of the WVS, where 20.1% of cluster 2 individuals hail from the East, while 9.5% of cluster 1 individuals hail from the East.¹⁵

4 Endogenous Partitions: Empirical Findings

4.1 Partitions and Antagonism

4.1.1 Partitions into Two Clusters

Our focus in this subsection is on data from wave 7 of the WVS. Figure 3 presents plots of the endogenous partitions of individuals into two clusters, based on the first two principal components of answers, for a selection of seven countries among our sample of 81.¹⁶ Each dot in the plot represents an individual’s position along the two principal components, the colors indicate the cluster to which they belong, and the solid line represents the hyperplane separating the two clusters. Solid black dots represent the mean of each cluster along the two dimensions. In some countries, like the USA, the clusters are formed mostly along a single dimension (the hyperplane is almost a vertical line). When looking at the underlying questions that have large weights in this component for the US, they tend to be questions related to religious values and morality.¹⁷ For several other countries, however, both dimensions matter to separate individuals into clusters (the hyperplane is tilted) - see for instance the cases of Germany, Nigeria and Ethiopia. Naturally, since the first principal component, by definition,

¹⁴For wave 7, all respondents from Berlin as assigned to the West, as we lack the ability to assess whether they reside on either side of the former wall. For wave 3, we can correctly assign respondents to either East or West Berlin. In that wave, there are 35 respondents residing in West Berlin and 88 respondents residing in East Berlin.

¹⁵In wave 7 only 15.5% of the respondents were from the East, while in Wave 3 they were 45%. In waves 3 and 7, 76.7% and 73.3% of the respondents from East Germany, respectively, belong to cluster 2. Qualitatively similar results are found when allowing for more than two clusters. Then, there tends to be one or two predominantly East German clusters in wave 3, but no longer in wave 7.

¹⁶Appendix A3 lists the WVS questions that receive the highest weights in the first two principal components (for wave 7), for a selection of 7 illustrative countries. These are the questions that play the most important role in setting apart the endogenous partitions discussed in this subsection.

¹⁷Specifically, the 5 questions with the largest weight in the first principal component for the US are: Believe in: hell; Believe in: heaven; How important is God in your life; Important child qualities: religious faith; Believe in: God..

accounts for the greatest share of the variance in answers, groups tend to be formed mostly along this component (exceptions exist, like Peru, where the second component is the main determinant of the clustering).

Table 1 presents the data and summary statistics for three central objects of interest: the reduction in antagonism achieved by creating two clusters (r), average antagonism per capita and the size of the largest cluster. We consider partitions based on successively larger numbers of principal components (1, 2, and 75). When moving from one to two clusters, and partitioning based on the first principal component of answers to wave 7 of the WVS, we find that social antagonism as defined in Section 2.1, is reduced by 68.49% on average. The reduction that is achieved with more principal components becomes progressively smaller, because more dimensions of heterogeneity make it more difficult to efficiently group individuals: two individuals can share an affinity along one dimension but not the other, and if they end up in the same cluster, they will differ along this second dimension, leading to greater within-group antagonism.

When using the first two principal components to form clusters, the average reduction in antagonism from going to two clusters is 41.51% - with some variation across countries. For reference, if we instead were to cluster on the sole basis of gender, the average reduction in antagonism that would be achieved would be a mere 1.29% (see Section 2.2 for details on the method used to obtain this number, r^g). If we divided the sample into individuals above and below the median education level, antagonism reduction would be 2.96% on average. In other words, creating clusters based on values achieves a reduction in antagonism that is an order of magnitude larger than partitions based on exogenous identity and demographic characteristics, like gender and education.

Looking at individual country examples can also be informative. For instance, dividing South Korean respondents into two clusters based on the first two principal components reduces within-group antagonism by 51.39%. If South Korea were instead divided on the basis of gender, the reduction in antagonism would be 3.24% (the highest number among our sample of countries), and the education-based reduction would be 2.96%. For the US, a values-based partition would reduce within-cluster antagonism by 41.10% (the gender-based reduction would be 1.07%, and the education-based reduction would be 4.89%). There is also substantial variation in the amount of antagonism per capita that remains after dividing society into two clusters: this quantity ranges from 2.24 units of antagonism in Egypt to 9.83 in Thailand (the US scores quite highly, at 6.56). Another interesting observation is that group sizes are relatively balanced, as the average size of the largest cluster is around 62%. The US stands out in this respect, as the largest group contains 67.5% of the respondents. These observations tend to hold in relative terms across countries when varying the number of principal components underlying the endogenous partitions.

4.1.2 Partitions into Three or More Clusters

How much of a reduction in social antagonism is achieved by allowing for more clusters? In this subsection, we explore what happens to social antagonism when we allow for 3, 4 and 5 clusters. It is obvious that within-group antagonism will decline as we allow for more clusters (by definition antagonism is zero if each individual is its own cluster).

Table 2 shows the results when the answer space is collapsed to the first two principal components of WVS answers, and we allow for three, four and five clusters.¹⁸ Figure 4 represents these results graphically for a selection of seven countries, based on the first two principal components (for the US, we also display clusters based on three principal components). In the US, the cluster on the right of Figure 3 now gets divided into two, mostly along lines of the second principal component (capturing questions on trust, political participation, interest in politics, and social capital including church attendance). It is notable that neither of the first principal components that define cultural partitions have anything to do with economic redistribution: both of the defining dimensions of cultural partitioning in the US concern questions related to religious values and morality, trust, social capital, election fairness and religious attendance. This may explain the salience of many of these issues in the US political debate in recent years.

Figure 5 displays the partitions obtained when we allow for four clusters, for a selection of seven countries. Results are quite similar to those obtained with three clusters (some of the latter simply get further divided). One notable subdivision is apparent for the US when allowing for four clusters: the two large horizontal clusters on the right side of Figure 4 (corresponding to religious individuals, with high or low levels of trust and social capital) get subdivided into three clusters in Figure 5, while the cluster of largely secular respondents on the left side of the figure remains largely unaffected. The ideological distance between the mean of latter cluster and the mean of the cluster with the average position located farthest to the right, is equal to 3.04 (on a left-right scale from 1 to 10). To put this in context, the sample mean ideological distance between the most distant clusters for all other countries is 1.51, so the US stands out starkly in terms of ideological distance between the two farthest clusters.¹⁹

We saw in the previous subsection that the reduction in within-cluster antagonism from going to two clusters is 41.51%. Figure 6 shows the reduction in antagonism that is achieved by moving to successively more clusters (when clusters are formed using the first two principal components of WVS answers in wave 7). An additional 20% or so of antagonism is eliminated by going to three clusters, and gains diminish to about 9% and 6%, respectively, when allowing for four and five clusters. In other words, on average across countries, by allowing for five clusters, we can reduce antagonism by 76.3% compared to a society that is not partitioned. The US is very close to this average. South Korea, Mexico and Thailand are among countries with larger reductions from going from zero to five clusters. But in some cases, like South Korea, the biggest decline is obtained by allowing for only two clusters, indicating that a division of society into two groups already achieves a very large reduction

¹⁸Results based on different numbers of principal components are also available and do not differ materially - except of course that the level of societal antagonism left over after partitioning is larger the greater the number of principal components under consideration.

¹⁹There is an interesting distinct group in the Ethiopian sample, on the left side of the corresponding figure (see Figure 5, Panel B). This group consists of 118 individuals, 79 of whom are from the Oromo ethnic group, and 68 of whom are Muslims (this is 57% of the group, whereas Muslims represent 30% of the Ethiopian population). 81 of them live in the same region, the group is quite gender- and age-balanced, but it is very distinct in terms of political ideology - the group is very right-winged. 56 of the individuals reported a 10 on the 1-10 left-right scale. This example shows that our clustering method is able to detect a very culturally distinct group. In fact, this group is already set apart when allowing for only three clusters (Figure 4 Panel B).

in within-group antagonism. Thailand gets the lowest reduction in antagonism from going from two to five clusters. Countries like Bangladesh and Ethiopia achieve high reductions going from two to three clusters. In sum, there is substantial cross-country variation in antagonism reduction when allowing for successively larger numbers of partitions.

4.2 Identity and Values Differences between Clusters

In this subsection, we discuss the demographic and cultural characteristics of the partitions obtained from Wave 7 of the WVS. We focus on two cases: those corresponding to two and three partitions. We also focus on divisions obtained from a consideration of two or three principal components.

4.2.1 Identity Differences across Partitions

Consider first the clustering of society into two cultural groups. Table 3 Panel A displays the mean difference in various identity traits across the two cultural partitions, for the 81 countries in our sample. The largest average differences between clusters are found for gender, where the two clusters exhibit a 7.69 percentage point average difference in the proportion of men, and especially religion, with a 20.74 percentage point difference in the proportion of respondents who do not belong to any religious denomination. Panel B shows the underlying data for the seven baseline countries, to get a better notion of heterogeneity. Figure 7 displays the distribution of differences in six identity traits across 81 countries in the case of two clusters. Country labels indicate the positions of seven illustrative countries. In general, the mode of the distribution occurs at levels of identity trait differences between groups that is relatively low. Gender and religious affiliation differences across groups are particularly pronounced for Korea and the US. In South Korea, there is a 22.37 percentage point difference in the proportion of males between the two groups. The US exhibits a 28.31 percentage point difference between groups in the proportion of those who do not belong to any religious denomination. Another notable fact is the large ideological difference between groups in the US: on a 1-10 left-right scale of ideology, the difference between the two groups in the US is 2.14 points - higher than in any other country in the sample (the sample mean is 0.65).

Despite these patterns, there is substantial variation in identity traits within cultural groups: identity-based groupings differ significantly from values-based groupings. There are particularly weak between-group differences in income and education. This is not what we would expect if the most relevant dimension of social heterogeneity had to do with preferences for redistribution and taxation.

4.2.2 Values Differences across Partitions

Figure 8 contains information on the distribution of differences in values between clusters (again based on a two-cluster partition). For tractability, we selected a set of 18 questions from wave 7 of the WVS (about 10% of the total), which are roughly representative of the span of issues covered by the survey. Table 4 Panels A and B shows summary statistics for 9 of these 18 questions, and the data for seven countries. Not surprisingly, the difference in values between clusters tends to be larger than the difference in identity traits: after all, clusters are formed on the basis of differences in cultural values. What is interesting here is the heterogeneity across values and across countries.

Some values do not seem to be very divisive and others are. For instance, views on whether the family is important do not differ much between clusters in most countries (on a scale from 0 to 1, the mean absolute difference between clusters is about 0.024 with a standard deviation of 0.024, across 81 countries). A low value is also typical for a question on whether success is due to hard work or luck (the mean difference is 0.065). In contrast, questions on religion and associated values (homosexuality, abortion) tend to display higher mean differences between clusters.

Turning to heterogeneity across countries, we also uncover interesting patterns (Table 4 Panel B and Figure 8). In the US, questions on religion, homosexuality, abortion as well as government versus individual responsibility, and immigration policy, are the most divisive. Questions on the importance of religion, belief in Heaven and religious attendance are particularly divisive in Korea, but the clusters are very similar on all other questions. In Germany, questions on religion and associated moral issues seem paramount. In Nigeria, the question on confidence in government displays big between-cluster differences, while in Ethiopia questions on homosexuality and abortion display big between-cluster differences. China displays relatively small differences on many dimensions, except when it comes to questions on generalized trust, abortion and homosexuality.

4.3 Changes over Time: The Case of the US

In this subsection, we explore time variation in antagonism across successive waves of the WVS in the US. We focus on the case of two clusters, obtained using principal components that were themselves computed separately wave-by-wave.

4.3.1 Evolution of Antagonism over Time

Figure 9 displays r , the reduction in antagonism that results from creating two clusters. No matter how many principal components are used to create the clusters, we see interesting patterns. r tended to increase between the first two waves of the WVS, and then remained elevated. There is also a spike in r in the latest wave of the survey (i.e. between 2011 and 2017). How can we interpret this finding? A higher value of r indicates that a higher within-group reduction in antagonism is achieved. This means that the share of between-group cultural antagonism in total societal antagonism is higher: when captured by endogenously formed cultural partitions, between-group divisions in the US have become relatively more pronounced since 2011. Recent work has documented the increasing political divide in the US, as differences in values between Democrats and Republicans have grown in the last decades (Desmet and Wacziarg, 2021) and affective polarization between individuals of different parties has increased too (Boxell, Gentzkow and Shapiro, 2022). A possible interpretation of these trends is that the underlying cultural divisions in US society have recently become more pronounced when looking at the fundamental building blocks of political preferences (cultural values).

4.3.2 Changes in Identity Differences between Clusters

Figure 10 displays the differences in identity traits between the two clusters in the US over time, using endogenous partitions obtained with the first two principal components of answers to successive waves. We do so for 9 identity markers, adding the difference in the share of Democrats and the differences

in the share of Whites to the list of seven identity cleavages discussed in Section 4.2.1. We uncover a notable pattern: political and ideological differences between clusters have increased sharply over time. For example the ideological difference, on a scale of 1 to 10, has increased steadily from 0.68 points in 1981 to 2.13 points in 2017, with the biggest increase occurring between waves 6 (2011) and 7 (2017). The share of democrats in one of the 1995 clusters was 4.71 percentage points higher than in the other, and this difference had grown to 28.75 percentage points by 2017. Thus, the endogenous partitions have become much more politically patterned: partitions derived endogenously from cultural values are much more predictive of political and ideological positions today than they were in the past.

These statements do not hold when it comes to the remaining identity cleavages that we consider. For instance, there no systematic tendency for the clusters to differ more in terms of gender, age, schooling and income. The increase in racial differences across clusters is quantitatively modest (the difference in the share of Whites across clusters only increased by about 2 percentage points). One cleavage that displays a reduction in between-cluster difference is age, with a 6 year decline in the absolute age difference between clusters from 1981 to 2017. Income differences between clusters also display, if anything, a downward trend, consistent with both a lower salience of economic position as a determinant of cultural partitions, and with the facts reported in Gethin, Martínez-Toledano and Piketty (2021) concerning the gradual reversal of the income gradient between left and right in the US - which is still in process. Of course, it is important to keep in mind that our values-based partitions are *not* the same as political partitions. However, to the extent that cultural partitions are more reflective of political divisions today than in the past, it is not unreasonable to argue that the decline in income differences between cultural clusters is related to the partial reversal of the income gradient in political affiliation.

4.3.3 Changes in Cultural Differences between Clusters

Figure 11 displays the difference in cultural values between clusters over time in the US, for a selection of 12 WVS questions. We focus again on partitions obtained using the first two principal components in each wave of the WVS. For all of these values, we see increasing differences between clusters. These changes are particularly pronounced for questions on religion and associated values: whether religion is important, whether homosexuality or abortion are justifiable, etc. For instance, when it comes to belief in Heaven, the difference between the two clusters went from 0.25 in 1981 to 0.64 in 2017 - a large increase. Similar magnitudes are obtained for other questions on religion and moral values. We also see pronounced increases in difference relating to confidence in government and belief in a democratic political system, especially between 2011 and 2017. These increasing differences also exist when it comes to economic values, such as questions on the role of hard work versus luck in explaining life success, the responsibility of the individual versus the government in providing for people (for these questions, the increase is pronounced mostly between waves 6 and 7).

However, the latter variables do not feature prominently among the first two principal components that serve to determine the two cultural clusters. Indeed, Appendix A4 displays the five most weighted questions for the first three principal components, for the US, across the seven waves for which we have US data. We find considerable stability across time in the types of questions that matter most to determine variation in the first three principal components - especially for the last five waves and

especially for the first two principal components.²⁰ The first dimension mostly focuses on religious values and associated values (views on homosexuality and abortion). The second mostly focuses on values related to trust, social capital and membership in religious organizations. In other words, we do not see major changes in the set of questions that play the biggest roles in setting apart the clusters. Despite this, we do see that, in the US, the clusters have grown wider apart over time in terms of cultural values.

5 Conclusion

In this paper, we have documented a number of notable facts, among which:

1. A substantial reduction in within-group antagonism can be achieved by creating a small number of cultural clusters, typically two or three.
2. Dividing individuals into endogenous clusters based on their affinities in terms of cultural values achieves a reduction in within-group cultural antagonism that is an order of magnitude larger than that achieved using conventional markers of identity, such as gender.
3. In the US, clusters are formed mostly based on religious beliefs, religious attendance, and moral values. Economic preferences do not play a significant role in defining endogenous partitions. This is not necessarily the case in other countries, where a heterogeneous range of questions helps define clusters.
4. In the US, cultural antagonism between groups has increased, and conversely, cultural groups have become more homogeneous over time.
5. Ideological, religious and political differences between clusters have increased sharply in the US across WVS waves. In particular, the ideological difference between clusters is now at the highest level of all countries in our sample.

This paper is meant more as a proof of concept than an attempt to analyze the causes and consequences of cultural antagonism. There are several more analytical uses to which our measurement framework could apply.

One set of questions relates to the fragmentation and polarization of the political system: if dividing society into two groups results in a bigger reduction in within-group cultural antagonism, society can be said to become more culturally polarized. This in turn could explain the rise of political polarization, to the extent that political parties are responsive to the distribution of voter beliefs. Correspondingly, if a division of society into two cultural groups becomes less effective at reducing within-group antagonism, this could be a sign of growing political fragmentation (requiring more than two parties to effectively represent voter beliefs).

²⁰Wave 1 stands out as different, with a strong weight given to important attributes of jobs in the first two principal components. This is perhaps due to the specific labor market conditions that prevailed at the time of the 1981 wave of the WVS, namely an incipient sustained recession.

Another set of questions concerns the degree of alignment of political platforms with voter values and preferences. By characterizing the distribution of values between and within cultural groups, our framework can allow a comparison with the values and preferences embedded in political platforms, to analyze whether the latter adequately capture the values of various groups - and whether there is any scope for a political realignment. For instance, the potential for a third party to emerge in the US can be assessed, and the nature of that party's political platform can be surmised by studying the values of groups of voters that such a party would be likely to attract (by homophily with its platform). The platforms of existing parties can also be reexamined in light of the distribution of values among the respondents to a representative survey. Trends in group-level values over time can help assess the likely fortunes of existing political parties whose platforms may lag behind the evolution of voters' cultural packages - and help anticipate transformations in the political landscape.

Our framework could also be applied to study issues related to identity politics. If platforms and parties are increasingly created to appeal to certain identity groups, as many commentators have observed, the costs of such identity politics in terms of within-group cultural heterogeneity can be assessed. Suppose that groups are formed exclusively on the basis of identity (race, gender) as opposed to cultural values. Is the resulting level of within-group antagonism now higher? To the extent that groups formed on the basis of values differ greatly from groups formed on the basis of identity, our approach can help assess whether identity-based approaches to politics are likely to fail or to succeed.

Finally, a set of issues relates to the geographic distribution of the groups obtained based on endogenous partitions. If groups are strongly geographically patterned (e.g. most of group 1 respondents are in region 1), then the scope for political secession based on heterogeneity in values becomes greater (Alesina and Spolaore, 1997). Our method would allow us to identify significant differences in values between residents of "secessionist" regions and the rest, in countries with strong regional secessionist movements, such as Spain, Canada and the United Kingdom.

References

- Akerlof, G. and R. Kranton (2000), "Economics and Identity", *The Quarterly Journal of Economics*, 115(3): 715-753.
- Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat and R. Wacziarg (2003) "Fractionalization", *Journal of Economic Growth*, (8), 2: 155-194.
- Alesina, A. and E. Spolaore (1997) "On the Number and Size of Nations", *The Quarterly Journal of Economics*, 112(4): 1027-1056.
- Alesina, A., G. Tabellini and F. Trebbi (2017), "Is Europe an Optimal Political Area?", *Brookings Papers on Economic Activity*, Spring: 169-213.
- Baron, D. (1993) "Government Formation and Endogenous Parties", *The American Political Science Review*, 87(1): 34-47.
- Bell, A., P. J. Richerson and R. McElreath (2009), "Culture Rather than Genes Provides Greater Scope for the Evolution of Large-Scale Human Prosociality," *PNAS*, 106(2): 17671-17674.
- Bonomi, G., N. Gennaioli and G. Tabellini (2019), "Identity, Beliefs, and Political Conflict", *The Quarterly Journal of Economics*, 136(4): 2371-2411.
- Boxell, L., M. Gentzkow and J. M. Shapiro (2022), "Cross-Country Trends in Affective Polarization", *The Review of Economics and Statistics*, forthcoming.
- Desmet, K., I. Ortuño-Ortín and R. Wacziarg (2017), "Culture, Ethnicity and Diversity", *American Economic Review*, 107(9): 2479-2513.
- Desmet, K. and R. Wacziarg (2021), "The Cultural Divide", *The Economic Journal*, (131)637: 2058-2088.
- DiMaggio, P., J. Evans and B. Bryson (1996), "Have Americans' Social Attitudes Become More Polarized", *American Journal of Sociology*, 102(3): 690-755.
- Egan, P. J. (2020), "Identity as Dependent Variable: How Americans Shift Their Identities to Align with Their Politics", *American Journal of Political Science*, 64(3): 699-716.
- Enke, B. (2020), "Moral Values and Voting", *Journal of Political Economy*, 128(10): 3679-3729.
- Gethin, A., C. Martínez-Toledano, T. Piketty (2021), "Brahmin Left Versus Merchant Right: Changing Political Cleavages in 21 Western Democracies, 1948–2020", *The Quarterly Journal of Economics*, 137(1): 1-48.
- Gomberg, A. M., F. Marhuenda and I. Ortuño-Ortín (2004), "A Model of Endogenous Political Party Platforms", *Economic Theory*, 24: 373-394.
- Greenberg, J. H. (1956). "The Measurement of Linguistic Diversity", *Language*, 32(1): 109-115.
- Inglehart, R. (1997). *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*. Princeton University Press.
- Inglehart, R. and W. E. Baker (2000), "Modernization, Cultural Change and the Persistence of Traditional Values", *American Sociological Review*, 65: 19-51.
- Jackson, M. (2022), "Inequality's Economic and Social Roots: The Role of Social Networks and Homophily". *Working Paper*, Stanford University.
- Mason, L. (2016), "A Cross-Cutting Calm: How Social Sorting Drives Affective Polarization", *Public Opinion Quarterly*, 80(S1): 351-377.

- Mason, L. (2018), "Ideologues without Issues: The Polarizing Consequences of Ideological Identities", *Public Opinion Quarterly*, 82(S1): 866-887.
- McPherson, M., Smith-Lovin, L. and J. C. Cook (2001), "Birds of a Feather: Homophily in Social Networks", *Annual Review of Sociology*, 27(1): 415-444.
- Norris, P. and Inglehart, R. (2019), *Cultural Backlash: Trump, Brexit, and Authoritarian Populism*. Cambridge: Cambridge University Press.
- Olivier, J., Thoenig, M. and Verdier, T. (2008), "Globalization and the Dynamics of Cultural Identity", *Journal of International Economics*, 76(2): 356-370.
- Ortuño-Ortín, I. and J. E. Roemer (2000), "Endogenous Party Formation and the Effect of Income Distribution on Policy", *Working Paper*.
- Shayo, M. (2009), "A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution", *American Political Science Review*, 103(2): 147-174.
- Tiebout, C. M. (1956), "A Pure Theory of Local Expenditures", *Journal of Political Economy*, 64(5): 416-424.
- Verbrugge, L. (1977), "The Structure of Adult Friendship Choices", *Social Forces*, 56(2): 576-597.

Appendix

Appendix A1. Link between Antagonism and Existing Measures of Heterogeneity

A1A. Antagonism, Cultural Fractionalization and Cultural Differentiation.

Suppose the cultural values of individuals are in a Q -dimensional space. Thus, an individual is characterized by the vector $x_j = \{x_j^1, x_j^2, \dots, x_j^Q\}$. Total antagonism in a given cluster A_i is given by the total pairwise distances of the N_i individuals in that cluster, i.e.:

$$E(A_i) = \frac{1}{N_i} \sum_{j=1}^{N_i} E(A_i, x_j) = \frac{1}{N_i} \sum_{j=1}^{N_i} \frac{\sum_{k=1}^{N_i} d(x_j, x_k)}{N_i} \quad (14)$$

The average (per capita) antagonism in cluster A_i is:

$$E(A_i) = \frac{1}{N_i^2} \sum_{j=1}^{N_i} \sum_{k=1}^{N_i} d(x_j, x_k) \quad (15)$$

Consider first the one-dimensional case. We can define the CF index of cultural fractionalization that incorporates a consideration of distances (Desmet, Ortuño-Ortín and Wacziarg, 2017, Appendix, page 7):

$$CF_D^{A_i} = \sum_{j=1}^{N_i} \sum_{k=1}^{N_i} w_j w_k d_{jk} \quad (16)$$

Here CF denotes cultural fractionalization, superscript A_i denotes the specific cluster under consideration, subscript D denotes that we are accounting for distances in answers. $CF_D^{A_i}$ is Greenberg's B index: the expected distance between the answers given by two randomly picked individuals. In our case $w_j = w_k = 1/N_i$, so:

$$E(A_i) = CF_D^{A_i} \quad (17)$$

In the Q -dimensional case, $CF_D^{A_i}$ is the average CF over the Q dimensions (see Desmet, Ortuño-Ortín and Wacziarg, 2017). In the case of squared Euclidean distance, we have $d(x_j, x_k) = \|x_j - x_k\|^2$, so:

$$E(A_i) = \frac{1}{N_i} \sum_{j=1}^{N_i} E(A_i, x_j) = \frac{1}{N_i^2} \sum_{j=1}^{N_i} \sum_{k=1}^{N_i} d(x_j, x_k) = \frac{1}{N_i^2} \sum_{j=1}^{N_i} \sum_{k=1}^{N_i} \sum_{q=1}^Q (x_j^q - x_k^q)^2 \quad (18)$$

The index $CF_D^{A_i}$ is now:

$$CF_D^{A_i} = \sum_{q=1}^Q \frac{1}{Q} \sum_{j=1}^{N_i} \sum_{k=1}^{N_i} \frac{1}{N_i^2} (x_j^q - x_k^q)^2 = \frac{1}{Q} E(A_i) \quad (19)$$

Next, consider G clusters, A_1, A_2, \dots, A_G , with number of individuals in cluster g given by N_g . We write the mean position in cluster g as μ_g . The set of all individuals is $P = A_1 \cup A_2 \dots \cup A_G$. In this case, within cluster CF is:

$$CF_D^W = \sum_{g=1}^G \frac{N_g}{N} CF_D^{A_g} = \sum_{g=1}^G \frac{N_g}{N} \frac{E(A_g)}{Q} \quad (20)$$

Now we can relate antagonism to the Φ_{ST} as defined on page 7 of the appendix of Desmet, Ortuno-Ortín and Wacziarg (2017) - the so-called "index of genetic differentiation". We have:

$$\Phi_{ST} = \frac{CF_D^P - CF_D^W}{CF_D^P} = \frac{\frac{E(P)}{Q} - \sum_{g=1}^G \frac{N_g}{N} \frac{E(A_g)}{Q}}{\frac{E(P)}{Q}} = \frac{E(P) - \sum_{g=1}^G \frac{N_g}{N} E(A_g)}{E(P)} \equiv r \quad (21)$$

It is immediate to see that, given clusters G , the clusters $A_1^*, A_2^* \dots A_G^*$ that minimize total within-group antagonism are the ones that maximize the index Φ_{ST} (i.e. between-group antagonism) - which also happens to be the indicator r .

A1B. Antagonism, Mean Position and Between-Cluster Variance in the Squared Euclidian Case.

We next show the well-known fact that in the case of Squared Euclidean distances, we have:

$$E(A_i) = \frac{1}{N_i^2} \sum_{j=1}^{N_i} \sum_{k=1}^{N_i} \|x_j - x_k\|^2 = \frac{2}{N_i} \sum_{j=1}^{N_i} \|x_j - \mu_i\|^2 \quad (22)$$

where μ_i denotes the vector of mean answers within group A_i :

$$\mu_i = \frac{\sum_{j=1}^{N_i} x_j}{N_i} \quad (23)$$

Proof:

$$\begin{aligned} E(A_i) &= \frac{1}{N_i^2} \sum_{k=1}^{N_i} \sum_{j=1}^{N_i} \|x_i - x_j\|^2 = \frac{1}{N_i^2} \sum_{k=1}^{N_i} \sum_{j=1}^{N_i} \sum_{q=1}^Q (x_k^q - x_j^q)^2 = \frac{1}{N_i^2} \sum_{k=1}^{N_i} \sum_{j=1}^{N_i} \sum_{q=1}^Q (x_k^q - \mu_i^q - x_j^q + \mu_i^q)^2 \\ &= \frac{1}{N_i^2} \left(\sum_{k=1}^{N_i} \sum_{j=1}^{N_i} \left(\left(\sum_{q=1}^Q (x_k^q - \mu_i^q)^2 \right) + \left(\sum_{q=1}^Q (x_j^q - \mu_i^q)^2 \right) \right) - 2 \sum_{q=1}^Q \left(\sum_{k=1}^{N_i} \sum_{j=1}^{N_i} (x_k^q - \mu_i^q) (x_j^q - \mu_i^q) \right) \right) \\ &= \frac{2}{N_i} \sum_{k=1}^{N_i} \sum_{q=1}^Q (x_k^q - \mu_i^q)^2 - 0 = \frac{2}{N_i} \sum_{k=1}^{N_i} \|x_k - \mu_i\|^2 \end{aligned} \quad (24)$$

We can also relate the antagonism in A_i with the variance of the positions $\{x_1, \dots, x_{N_i}\}$. In the one-dimensional case the variance of A_i is:

$$Var(A_i) = \frac{1}{N_i} \sum_{j=1}^{N_i} (x_j - \mu_i)^2 \quad (25)$$

Thus, $E(A_i) = 2Var(A_i)$. With G clusters, A_1, A_2, \dots, A_G we have:

$$\sum_{g=1}^G E(A_g) = 2 \sum_{g=1}^G Var(A_g) \quad (26)$$

Thus, total within cluster antagonism is closely related to weighted total within cluster variance under the assumption that the distance metric is squared Euclidian. Another way to state this result is that the clusters $A_1^*, A_2^*, \dots, A_G^*$ that minimize total antagonism in society are the one that maximize the index Φ_{ST} . Since total variance is constant, minimizing within-cluster antagonism is equivalent to maximizing in between clusters variance.

In the multidimensional case ($Q > 1$), the interpretation is not exactly the same. We have:

$$E(A_i) = \frac{2}{N_i} \sum_{j=1}^{N_i} \sum_{q=1}^Q (x_j^q - \mu_i^q)^2 = \frac{2}{N_i} \sum_{q=1}^Q \sum_{j=1}^{N_i} (x_j^q - \mu_i^q)^2 = 2 \sum_{q=1}^Q Var(x_i^q) \quad (27)$$

where $Var(x_i^q)$ stands for (sample) variance in dimension q . Obviously, $\sum_{q=1}^Q Var(x_i^q)$ is not the "variance" of the random vector x_i , since it ignores covariances. However, if we apply our method using principal components of the questions, the Q dimensions are by construction uncorrelated, and $Cov(x^q, x^{q'}) = 0$. In this case, results obtained in the one-dimensional case apply for $Q > 1$ as well, and minimizing within-clusters antagonism is equivalent to maximizing between-clusters variance.

Appendix A2. List of Questions Used in the Empirical Analysis (Wave 7)

Category	Variable name (Common Dictionary)	Variable name (WVS 7)	Common Dictionary: Variable label
Environment	B008	Q111	Protecting environment vs. Economic growth
Family	D057	Q32	Being a housewife just as fulfilling
Family	D026_03	Q37	Duty towards society to have children
Family	D081	Q36	Homosexual couples are as good parents as other couples
Family	D001_B	Q58	How much do you trust your family (4-point scale)
Family	D026_05	Q38	It is child's duty to take care of ill parent
Family	D078	Q31	Men make better business executives than women do
Family	D059	Q29	Men make better political leaders than women do
Family	D054	Q27	One of main goals in life has been to make my parents proud
Family	D061	Q28	Pre-school child suffers with working mother
Family	D066_B	Q35	Problem if women have more income than husband (B)
Family	D060	Q30	University is more important for a boy than for a girl
National Identity	G052	Q121	Evaluate the impact of immigrants on the development of [your country]
National Identity	G257	Q257	How close do you feel: to country
National Identity	G256	Q256	How close do you feel: to your county, region, district
National Identity	G062	Q258	How close you feel: Continent; e.g. Europe, Asia etc.
National Identity	G063	Q259	How close you feel: World
National Identity	G255	Q255	How close you feel: Your [village, town or city]
National Identity	G006	Q254	How proud of nationality
National Identity	G007_36_B	Q63	Trust: People of another nationality (B)
National Identity	G007_35_B	Q62	Trust: People of another religion (B)
National Identity	G007_33_B	Q60	Trust: People you know personally (B)
National Identity	G007_34_B	Q61	Trust: People you meet for the first time (B)
National Identity	G007_18_B	Q59	Trust: Your neighborhood (B)
Perceptions of life	A106	Q105	Active/Inactive membership of any other organization
Perceptions of life	A100	Q96	Active/Inactive membership of art, music, educational
Perceptions of life	A105	Q101	Active/Inactive membership of charitable/humanitarian organization

Perceptions of life	A098	Q94	Active/Inactive membership of church or religious organization
Perceptions of life	A103	Q99	Active/Inactive membership of environmental organization
Perceptions of life	A101	Q97	Active/Inactive membership of labour unions
Perceptions of life	A102	Q98	Active/Inactive membership of political party
Perceptions of life	A104	Q100	Active/Inactive membership of professional organization
Perceptions of life	A099	Q95	Active/Inactive membership of sport or recreation
Perceptions of life	A106B	Q102	Active/Inactive membership: Consumer organization
Perceptions of life	A106C	Q103	Active/Inactive membership: Self-help group, mutual aid group
Perceptions of life	A173	Q48	How much freedom of choice and control
Perceptions of life	A062	Q200	How often discusses political matters with friends
Perceptions of life	A039	Q14	Important child qualities: determination perseverance
Perceptions of life	A032	Q10	Important child qualities: feeling of responsibility
Perceptions of life	A027	Q7	Important child qualities: Good manners
Perceptions of life	A030	Q9	Important child qualities: Hard work
Perceptions of life	A034	Q11	Important child qualities: imagination
Perceptions of life	A029	Q8	Important child qualities: independence
Perceptions of life	A042	Q17	Important child qualities: obedience
Perceptions of life	A040	Q15	Important child qualities: religious faith
Perceptions of life	A038	Q13	Important child qualities: thrift saving money and things
Perceptions of life	A035	Q12	Important child qualities: tolerance and respect for other people
Perceptions of life	A041	Q16	Important child qualities: unselfishness
Perceptions of life	A001	Q1	Important in life: Family
Perceptions of life	A002	Q2	Important in life: Friends
Perceptions of life	A003	Q3	Important in life: Leisure time
Perceptions of life	A004	Q4	Important in life: Politics
Perceptions of life	A006	Q6	Important in life: Religion
Perceptions of life	A005	Q5	Important in life: Work
Perceptions of life	A080_02	Q103R	Member: Belong to self-help group, mutual aid group
Perceptions of life	A071B	Q99R	Member: Belong to conservation, the environment, ecology
Perceptions of life	A078	Q102R	Member: Belong to consumer groups
Perceptions of life	A066	Q96R	Member: Belong to education, arts, music or cultural activities
Perceptions of life	A080_01	Q101R	Member: Belong to humanitarian or charitable organization

Perceptions of life	A067	Q97R	Member: Belong to labor unions
Perceptions of life	A079	Q105R	Member: Belong to other groups
Perceptions of life	A068	Q98R	Member: Belong to political parties
Perceptions of life	A072	Q100R	Member: Belong to professional associations
Perceptions of life	A065	Q94R	Member: Belong to religious organization
Perceptions of life	A074	Q95R	Member: Belong to sports or recreation
Perceptions of life	A075	Q104R	Member: Belong to women's group
Perceptions of life	A165	Q57	Most people can be trusted
Perceptions of life	A124_08	Q18	Neighbors: Drug addicts
Perceptions of life	A124_03	Q24	Neighbors: Heavy drinkers
Perceptions of life	A124_09	Q22	Neighbors: Homosexuals
Perceptions of life	A124_06	Q21	Neighbors: Immigrants/foreign workers
Perceptions of life	A124_02	Q19	Neighbors: People of a different race
Perceptions of life	A124_12	Q23	Neighbors: People of a different religion
Perceptions of life	A124_07	Q20	Neighbors: People who have AIDS
Perceptions of life	A124_43	Q26	Neighbors: People who speak a different language
Perceptions of life	A124_42	Q25	Neighbors: Unmarried couples living together
Politics and Society	E001	Q152	Aims of country: first choice
Politics and Society	E003	Q154	Aims of respondent: first choice
Politics and Society	E034	Q42	Basic kinds of attitudes concerning society
Politics and Society	E218	Q159	Because of science and technology, there will be more opportunities for the next generation
Politics and Society	E039	Q109	Competition good or harmful
Politics and Society	E069_02	Q65	Confidence: Armed Forces
Politics and Society	E069_41	Q78	Confidence: Banks
Politics and Society	E069_40	Q81	Confidence: Charitable or humanitarian organizations
Politics and Society	E069_01	Q64	Confidence: Churches
Politics and Society	E069_43	Q82_CIS	Confidence: CIS
Politics and Society	E069_59	Q82_GULFCOOP	Confidence: Cooperation Council for the Arab states of Gulf (GCC)
Politics and Society	E069_45	Q84	Confidence: International Monetary Found (IMF)
Politics and Society	E069_17	Q70	Confidence: Justice System/Courts
Politics and Society	E069_05	Q68	Confidence: Labor Unions
Politics and Society	E069_13	Q77	Confidence: Major Companies

Politics and Society	E069_18A	Q82	Confidence: Major regional organization (combined from country-specific)
Politics and Society	E069_19	Q86	Confidence: NATO
Politics and Society	E069_07	Q73	Confidence: Parliament
Politics and Society	E069_10	Q67	Confidence: Television
Politics and Society	E069_29	Q82_APEC	Confidence: The APEC
Politics and Society	E069_21	Q82_ARABLEAGUE	Confidence: The Arab League
Politics and Society	E069_22	Q82_ASEAN	Confidence: The Association of South East Asian Nations -ASEAN
Politics and Society	E069_08	Q74	Confidence: The Civil Services
Politics and Society	E069_14	Q79	Confidence: The Environmental Protection Movement
Politics and Society	E069_18	Q82	Confidence: The European Union
Politics and Society	E069_30	Q82_TLC	Confidence: The Free Commerce Treaty (Tratado de libre comercio)
Politics and Society	E069_11	Q71	Confidence: The Government
Politics and Society	E069_26	Q82_MERCOSUR	Confidence: The Mercosur
Politics and Society	E069_24	Q82_NAFTA	Confidence: The NAFTA
Politics and Society	E069_56	Q82_OEA	Confidence: The Organization of American States (OAE)
Politics and Society	E069_55	Q82_ISLCOOP	Confidence: The Organization of the Islamic World
Politics and Society	E069_06	Q69	Confidence: The Police
Politics and Society	E069_12	Q72	Confidence: The Political Parties
Politics and Society	E069_04	Q66	Confidence: The Press
Politics and Society	E069_27	Q82_SAARC	Confidence: The SAARC
Politics and Society	E069_20	Q83	Confidence: The United Nations
Politics and Society	E069_15	Q80	Confidence: The Women's Movement
Politics and Society	E069_54	Q75	Confidence: Universities
Politics and Society	E229	Q246	Democracy: Civil rights protect people's liberty against oppression.
Politics and Society	E224	Q241	Democracy: Governments tax the rich and subsidize the poor.
Politics and Society	E226	Q243	Democracy: People choose their leaders in free elections.
Politics and Society	E233B	Q248	Democracy: People obey their rulers
Politics and Society	E227	Q244	Democracy: People receive state aid for unemployment.
Politics and Society	E225	Q242	Democracy: Religious authorities interpret the laws.
Politics and Society	E228	Q245	Democracy: The army takes over when government is incompetent.
Politics and Society	E233A	Q247	Democracy: The state makes people's incomes equal
Politics and Society	E233	Q249	Democracy: Women have the same rights as men.

Politics and Society	E236	Q251	Democracticness in own country
Politics and Society	E018	Q45	Future changes: Greater respect for authority
Politics and Society	E015	Q43	Future changes: Less importance placed on work
Politics and Society	E016	Q44	Future changes: More emphasis on technology
Politics and Society	E037	Q108	Government responsibility
Politics and Society	E040	Q110	Hard work brings success
Politics and Society	E265_06	Q229	How often in country's elections: Election officials are fair
Politics and Society	E265_05	Q228	How often in country's elections: Journalists provide fair coverage of elections
Politics and Society	E265_02	Q225	How often in country's elections: Opposition candidates are prevented from running
Politics and Society	E265_07	Q230	How often in country's elections: Rich people buy elections
Politics and Society	E265_03	Q226	How often in country's elections: TV news favors the governing party
Politics and Society	E265_04	Q227	How often in country's elections: Voters are bribed
Politics and Society	E265_09	Q232	How often in country's elections: Voters are offered a genuine choice in the elections
Politics and Society	E265_08	Q231	How often in country's elections: Voters are threatened with violence at the polls
Politics and Society	E265_01	Q224	How often in country's elections: Votes are counted fairly
Politics and Society	E235	Q250	Importance of democracy
Politics and Society	E035	Q106	Income equality
Politics and Society	E023	Q199	Interest in politics
Politics and Society	E290	Q194	Justifiable: Political violence
Politics and Society	E005	Q156	Most important: first choice
Politics and Society	E027	Q211	Political action: attending lawful/peaceful demonstrations
Politics and Society	E026	Q210	Political action: joining in boycotts
Politics and Society	E028	Q212	Political action: joining unofficial strikes
Politics and Society	E025	Q209	Political action: Signing a petition
Politics and Society	E117	Q238	Political system: Having a democratic political system
Politics and Society	E114	Q235	Political system: Having a strong leader
Politics and Society	E115	Q236	Political system: Having experts make decisions
Politics and Society	E116	Q237	Political system: Having the army rule
Politics and Society	E036	Q107	Private vs state ownership of business
Politics and Society	E124	Q253	Respect for individual human rights nowadays
Politics and Society	E111_01	Q252	Satisfaction with the political system
Politics and Society	E217	Q158	Science and technology are making our lives healthier, easier, and more comfortable

Politics and Society	E266	Q234	Some people think that having honest elections makes a lot of difference in their lives
Politics and Society	E234	Q163	The world is better off, or worse off, because of science and technology
Politics and Society	E263	Q221	Vote in elections: local level
Politics and Society	E264	Q222	Vote in elections: National level
Politics and Society	E220	Q160	We depend too much on science and not enough on faith
Politics and Society	E012	Q151	Willingness to fight for country
Religion and Morale	F050	Q165	Believe in: God
Religion and Morale	F054	Q168	Believe in: heaven
Religion and Morale	F053	Q167	Believe in: hell
Religion and Morale	F051	Q166	Believe in: life after death
Religion and Morale	F063	Q164	How important is God in your life
Religion and Morale	F028	Q171	How often do you attend religious services
Religion and Morale	F028B	Q172	How often do you pray
Religion and Morale	F120	Q184	Justifiable: Abortion
Religion and Morale	F115	Q178	Justifiable: Avoiding a fare on public transport
Religion and Morale	F116	Q180	Justifiable: Cheating on taxes
Religion and Morale	F114A	Q177	Justifiable: Claiming government benefits to which you are not entitled
Religion and Morale	F144_02	Q195	Justifiable: Death penalty
Religion and Morale	F121	Q185	Justifiable: Divorce
Religion and Morale	F122	Q188	Justifiable: Euthanasia
Religion and Morale	F199	Q189	Justifiable: For a man to beat his wife
Religion and Morale	F132	Q193	Justifiable: Having casual sex
Religion and Morale	F118	Q182	Justifiable: Homosexuality
Religion and Morale	F114C	Q190	Justifiable: Parents beating children
Religion and Morale	F119	Q183	Justifiable: Prostitution
Religion and Morale	F135A	Q186	Justifiable: Sex before marriage
Religion and Morale	F117	Q181	Justifiable: Someone accepting a bribe
Religion and Morale	F114B	Q179	Justifiable: Stealing property
Religion and Morale	F123	Q187	Justifiable: Suicide
Religion and Morale	F114D	Q191	Justifiable: Violence against other people
Religion and Morale	F200	Q174	Meaning of religion: To follow religious norms and ceremonies vs To do good to other people
Religion and Morale	F201	Q175	Meaning of religion: To make sense of life after death vs To make sense of life in this world

Religion and Morale	F034	Q173	Religious person
Religion and Morale	F203	Q170	The only acceptable religion is my religion
Religion and Morale	F202	Q169	Whenever science and religion conflict, religion is always right
Science	I002	Q162	It is not important for me to know about science in my daily life
Science	I001	Q161	One of the bad effects of science is that it breaks down people's ideas of right and wrong
Security	H011	Q198	Government has the right: Collect information about anyone living in [COUNTRY] without their knowledge
Security	H009	Q196	Government has the right: Keep people under video surveillance in public areas
Security	H010	Q197	Government has the right: Monitor all e-mails and any other information exchanged on the Internet
Security	H003_03	Q141	Things done for reasons of security: Carried a knife, gun or other weapon
Security	H003_01	Q139	Things done for reasons of security: Didn't carry much money
Security	H003_02	Q140	Things done for reasons of security: Preferred not to go out at night
Security	H006_05	Q148	Worries: A civil war
Security	H006_04	Q147	Worries: A terrorist attack
Security	H006_03	Q146	Worries: A war involving my country
Security	H006_01	Q142	Worries: Losing my job or not finding a job
Security	H006_02	Q143	Worries: Not being able to give one's children a good education
Work	C002	Q34	Jobs scarce: Employers should give priority to (nation) people than immigrants (3 categories)
Work	C002_01	C002_01	Jobs scarce: Employers should give priority to (nation) people than immigrants (5-point scale)
Work	C001	Q33	Jobs scarce: Men should have more right to a job than women (3 categories)
Work	C001_01	C001_01	Jobs scarce: Men should have more right to a job than women (5-point scale)
Work	C038	Q39	People who don't work turn lazy
Work	C039	Q40	Work is a duty towards society
Work	C041	Q41	Work should come first even if it means less spare time

Appendix A3. Questions With the Highest Weights in the First Two Principal Components (WVS wave 7)

The following list displays, for each of 7 illustrative countries, the 5 questions receiving the largest weights in each of the first two principal components (in decreasing order of weight).

BRAZIL:

PC 1: Member: Belong to religious organization; Active/Inactive membership of church or religious organization; Believe in: heaven; Important child qualities: obedience; Important child qualities: religious faith.

PC 2: Confidence: The Environmental Protection Movement; Confidence: The Women's Movement; Important in life: Politics; Confidence: Justice System/Courts; Confidence: Charitable or humanitarian organizations.

CHINA:

PC 1: Neighbors: Unmarried couples living together; Neighbors: People of a different religion; Neighbors: Immigrants/foreign workers; Neighbors: Homosexuals; Neighbors: People who speak a different language.

PC 2: Believe in: God; Most people can be trusted; Believe in: heaven; Believe in: hell; Confidence: The Civil Services.

ETHIOPIA:

PC 1: Important child qualities: religious faith; Important child qualities: obedience; Political system: Having a strong leader; Confidence: Parliament; Confidence: The Government.

PC 2: Confidence: The Government; Confidence: Parliament; Confidence: The Police; Confidence: Justice System/Courts; Member: Belong to self-help group, mutual aid group.

GERMANY:

PC 1: Believe in: God; Believe in: heaven; How important is God in your life; Believe in: life after death; Believe in: hell.

PC 2: Member: Belong to religious organization; Believe in: God; Active/Inactive membership of church or religious organization; Believe in: life after death; How important is God in your life.

NIGERIA:

PC 1: Political action: attending lawful/peaceful demonstrations; Political action: joining unofficial strikes; Satisfaction with the political system; Neighbors: Unmarried couples living together; Important child qualities: religious faith.

PC 2: Confidence: The Political Parties; Interest in politics; Confidence: The Government; Confidence: The Civil Services; Confidence: Justice System/Courts.

SOUTH KOREA:

PC 1: How often do you attend religious services; Religious person; How often do you pray; Meaning of religion: To follow religious norms and ceremonies vs To do good to other people; Meaning of religion: To make sense of life after death vs To make sense of life in this world.

PC 2: Political action: Signing a petition; Democracy: The army takes over when government is incompetent; Political action: attending lawful/peaceful demonstrations; Important child qualities: thrift saving money and things; Democracy: Religious authorities interpret the laws.

UNITED STATES:

PC 1: Believe in: hell; Believe in: heaven; How important is God in your life; Important child qualities: religious faith; Believe in: God.

PC 2: Most people can be trusted; How often do you attend religious services; Political action: attending lawful/peaceful demonstrations; Interest in politics; How often do you pray.

Appendix A4. Questions With the Highest Weights in the First Three Principal Components, United States, across Seven WVS Waves

The following list displays, for the United States, the 5 questions receiving the largest weights in each of the first three principal components (in decreasing order of weight), across successive waves of the WVS:

WAVE 1

DIMENSION 1: How often do you attend religious services; Important for successful marriage: Religious beliefs; Abortion when woman not married; Abortion if not wanting more children; Member: Belong to none.

DIMENSION 2: Voluntary work: Unpaid work none; Political action: joining in boycotts; Political action: Signing a petition; Member: Belong to none; Important child qualities: Good manners.

DIMENSION 3: Important in a job: good hours; Important in a job: generous holidays; Important in a job: not too much pressure; Most people can be trusted; Voluntary work: Unpaid work none.

WAVE 2

DIMENSION 1: Believe in: resurrection of the dead; Important child qualities: religious faith; Member: Belong to religious organization; Get comfort and strength from religion; Life is meaningful because God exists.

DIMENSION 2: Churches speak out on: third world problems; Churches speak out on: ecology and environmental issues; Important in a job: an opportunity to use initiative; Churches speak out on: disarmament; Churches speak out on: racial discrimination.

DIMENSION 3: Important in a job: a responsible job; Important in a job: a respected job; Important in a job: good chances for promotion; Important in a job: meeting people; Important in a job: a job that meets one's abilities.

WAVE 3

DIMENSION 1: Believe in: hell; Important child qualities: religious faith; Believe in: devil; Believe in: heaven; Get comfort and strength from religion.

DIMENSION 2: Important in a job: an opportunity to use initiative; Active/Inactive membership of charitable/humanitarian organization; Active/Inactive membership of professional organization; Active/Inactive membership of church or religious organization; Active/Inactive membership of art, music, educational.

DIMENSION 3: Important in a job: good hours; Important in a job: a responsible job; Important in a job: generous holidays; Important in a job: not too much pressure; Important in a job: a job that meets one's abilities.

WAVE 4

DIMENSION 1: Spend time with people at your church, mosque or synagogue; How often do you attend religious services; Important in life: Religion; Pray to God outside of religious services (i); Justifiable: Homosexuality.

DIMENSION 2: Most people can be trusted; Political action: joining in boycotts; Political action: attending lawful/peaceful demonstrations; Spend time with people at sport, culture, communal organization; How often discusses political matters with friends.

DIMENSION 3: Important in a job: generous holidays; Important in a job: a responsible job; Important in a job: not too much pressure; Important in a job: a respected job; Important in a job: a job that meets one's abilities.

WAVE5

DIMENSION 1: How often do you attend religious services; Woman as a single parent; Justifiable: Homosexuality; Better if more people with strong religious beliefs in public office; Important in life: Religion.
DIMENSION 2: Political action recently done: signing a petition; Most people can be trusted; Political action: joining in boycotts; Interest in politics; Political action: attending lawful/peaceful demonstrations.
DIMENSION 3: Churches give answers: moral problems; Churches give answers: the problems of family life; Churches give answers: the social problems; Churches give answers: people's spiritual needs; Woman as a single parent.

WAVE 6

DIMENSION 1: Important child qualities: religious faith; Believe in: hell; Active/Inactive membership of church or religious organization; How important is God in your life; We depend too much on science and not enough on faith.
DIMENSION 2: Vote in elections: National level; Vote in elections: local level; Most people can be trusted; Interest in politics; Political action: Signing a petition.
DIMENSION 3: Government responsibility; Democracy: Governments tax the rich and subsidize the poor.; Democracy: The state makes people's incomes equal; Things done for reasons of security: Didn't carry much money; Important child qualities: tolerance and respect for other people.

WAVE7

DIMENSION 1: Believe in: hell; Believe in: heaven; How important is God in your life; Important child qualities: religious faith; Believe in: God.
DIMENSION 2: Most people can be trusted; How often do you attend religious services; Political action: attending lawful/peaceful demonstrations; Interest in politics; How often do you pray.
DIMENSION 3: Political system: Having a strong leader; Worries: Not being able to give one's children a good education; Worries: Losing my job or not finding a job; Worries: A civil war; Political system: Having the army rule...

Table 1 - Main Statistics of Interest for the Two-Cluster Case, using different numbers of principal components (81 Countries)

Country	ISO2	Source	First Principal Component			First Two Principal Components			First 75 Principal Components		
			(PA- E*)/PA	Antagonism PC	% Size largest	(PA- E*)/PA	Antagonism PC	% Size largest	(PA- E*)/PA	Antagonism PC	% Size largest
Andorra	AD	WVS7	70.21	2.94	50.40	42.26	4.88	50.30	7.24	28.67	50.20
Albania	AL	EVS5	62.75	1.54	56.08	33.90	2.86	54.07	4.30	25.02	66.68
Armenia	AM	EVS5	64.66	1.61	58.94	40.69	2.80	53.20	4.20	28.89	51.41
Argentina	AR	WVS7	62.33	2.20	59.48	31.38	4.16	86.50	4.86	31.12	51.88
Austria	AT	EVS5	64.79	3.26	59.02	37.23	5.68	55.86	6.93	30.86	57.24
Australia	AU	WVS7	69.85	3.60	55.73	45.08	5.69	59.33	8.34	31.37	56.62
Azerbaijan	AZ	EVS5	65.40	1.90	56.59	38.02	3.49	54.26	4.85	28.49	53.66
Bos & Herzeg	BA	EVS5	64.86	2.40	51.80	36.38	4.30	52.78	4.91	32.63	52.74
Bangladesh	BD	WVS7	71.46	1.95	67.39	37.44	3.74	66.97	5.20	27.86	66.97
Bulgaria	BG	EVS5	70.08	2.05	50.21	37.57	3.83	51.05	4.99	29.64	50.66
Bolivia	BO	WVS7	65.28	1.89	65.84	37.82	3.26	65.59	4.21	29.98	65.01
Brazil	BR	WVS7	64.10	2.18	59.51	37.39	3.75	59.45	4.56	31.13	61.65
Belarus	BY	EVS5	63.88	2.84	59.77	42.07	4.32	57.99	6.06	30.21	58.68
Canada	CA	WVS7	67.92	3.76	58.21	42.11	6.16	55.42	8.39	31.44	54.84
Switzerland	CH	EVS5	66.70	2.95	59.51	40.23	4.90	57.04	7.00	28.28	56.90
Chile	CL	WVS7	85.72	3.39	85.87	46.64	6.24	85.70	8.91	33.93	85.70
China	CN	WVS7	66.77	1.42	54.39	39.43	2.42	54.04	4.37	22.41	57.80
Colombia	CO	WVS7	85.64	4.63	68.62	57.82	6.87	68.62	11.14	36.00	68.82
Cyprus	CY	WVS7	69.13	2.56	77.39	39.74	4.56	77.41	6.27	29.44	77.03
Czechia	CZ	EVS5	70.25	2.26	71.90	38.30	4.15	71.90	5.25	31.23	73.21
Germany	DE	WVS7	65.79	2.72	55.82	37.56	4.94	57.86	6.83	27.39	56.67
Denmark	DK	EVS5	64.11	2.18	54.38	35.81	3.90	54.84	5.52	26.08	51.18
Ecuador	EC	WVS7	67.77	2.58	75.71	42.37	4.18	76.79	5.64	31.71	76.04
Estonia	EE	EVS5	65.76	2.50	54.82	38.70	4.26	58.07	5.96	27.93	59.60
Egypt	EG	WVS7	64.37	1.22	59.66	35.84	2.24	50.42	3.83	21.48	53.87
Spain	ES	EVS5	69.92	4.07	50.23	50.04	5.70	50.11	8.88	32.37	50.46
Ethiopia	ET	WVS7	64.34	2.81	85.92	34.88	5.23	83.47	5.61	34.32	95.01

Country	ISO2	Source	First Principal Component			First Two Principal Components			First 75 Principal Components		
			(PA- E*)/PA	Antagonism PC	% Size largest	(PA- E*)/PA	Antagonism PC	% Size largest	(PA- E*)/PA	Antagonism PC	% Size largest
Finland	FI	EVS5	65.36	3.28	54.21	37.81	5.66	53.66	7.56	28.67	55.61
France	FR	EVS5	68.19	3.54	60.20	39.67	6.08	60.65	7.75	31.37	60.49
United Kingdom	GB	EVS5	65.55	3.06	52.96	39.82	5.45	59.05	7.38	29.74	58.90
Georgia	GE	EVS5	68.07	2.00	59.11	38.90	3.71	60.24	5.23	28.88	60.77
Greece	GR	WVS7	66.17	3.19	51.72	46.47	4.55	51.80	7.64	28.12	51.65
Guatemala	GT	WVS7	76.21	5.02	73.15	54.72	7.08	73.65	11.07	35.56	73.91
Hong Kong	HK	WVS7	72.36	3.50	82.80	49.32	5.31	85.42	9.32	28.49	85.93
Croatia	HR	EVS5	67.34	3.33	62.74	45.74	4.92	62.45	7.49	30.38	62.75
Hungary	HU	EVS5	68.14	3.18	50.58	42.06	5.25	51.82	7.06	31.47	53.16
Indonesia	ID	WVS7	71.56	2.44	69.07	43.54	4.06	68.79	6.12	29.27	68.59
Iraq	IQ	WVS7	65.35	2.99	58.88	37.53	5.23	59.55	5.86	33.83	61.47
Iran	IR	WVS7	66.29	3.34	68.42	47.11	4.71	67.82	7.04	31.68	66.36
Iceland	IS	EVS5	68.46	3.01	60.94	44.43	4.63	61.27	7.80	26.67	60.27
Italy	IT	EVS5	66.73	3.07	60.13	40.43	5.11	61.69	7.19	28.91	61.18
Jordan	JO	WVS7	62.06	1.48	55.17	33.85	2.71	54.08	3.33	27.92	56.00
Japan	JP	WVS7	65.89	1.76	56.44	35.89	3.26	56.44	4.83	24.69	56.97
Kyrgyzstan	KG	WVS7	61.78	2.55	54.41	38.66	4.07	54.47	5.33	30.08	54.54
South Korea	KR	WVS7	79.98	2.79	69.72	51.39	4.34	69.72	8.84	25.86	69.48
Kazakhstan	KZ	WVS7	66.56	2.39	51.79	36.95	4.32	52.60	5.20	31.11	53.25
Lebanon	LB	WVS7	69.39	2.53	62.17	40.00	4.68	58.08	6.24	30.57	58.25
Lithuania	LT	EVS5	64.90	2.11	51.75	35.65	3.94	54.26	5.27	27.10	53.23
Montenegro	ME	EVS5	79.71	3.31	53.64	50.29	5.26	52.72	8.59	31.63	55.69
N. Macedonia	MK	EVS5	63.85	3.53	55.25	40.55	5.61	55.07	7.03	33.55	51.87
Myanmar	MM	WVS7	66.03	1.94	52.17	35.85	3.61	56.58	4.34	30.06	54.67
Macao	MO	WVS7	67.25	2.84	52.20	43.74	4.89	55.34	7.32	29.62	55.93
Mexico	MX	WVS7	79.57	3.44	83.52	50.84	5.45	84.02	8.34	33.66	84.87
Malaysia	MY	WVS7	77.42	4.09	73.04	46.57	6.83	72.89	9.68	33.23	73.57
Nigeria	NG	WVS7	63.13	2.99	53.89	37.02	5.33	63.53	6.05	33.05	62.22
Nicaragua	NI	WVS7	75.11	4.23	71.25	50.64	6.31	74.67	8.95	36.06	74.67

Country	ISO2	Source	First Principal Component			First Two Principal Components			First 75 Principal Components		
			(PA-E*)/PA	Antagonism PC	% Size largest	(PA-E*)/PA	Antagonism PC	% Size largest	(PA-E*)/PA	Antagonism PC	% Size largest
Netherlands	NL	EVS5	71.03	3.18	65.20	43.76	5.37	65.60	8.44	28.09	65.05
Norway	NO	EVS5	67.57	2.66	55.27	41.42	4.40	59.06	6.90	26.98	59.15
New Zealand	NZ	WVS7	71.27	3.74	64.99	44.17	6.11	65.86	7.96	34.23	65.66
Peru	PE	WVS7	63.87	1.28	54.36	32.13	2.60	59.47	3.42	26.32	56.11
Philippines	PH	WVS7	71.72	2.47	75.11	42.58	4.16	74.30	6.08	30.07	75.15
Pakistan	PK	WVS7	65.29	3.93	75.06	40.91	6.58	72.29	7.82	34.71	73.70
Poland	PL	EVS5	67.62	3.26	61.44	44.27	4.98	61.63	7.89	28.14	60.14
Puerto Rico	PR	WVS7	64.10	2.64	67.53	35.71	4.76	68.69	4.86	35.78	68.07
Portugal	PT	EVS5	65.71	2.53	61.85	40.79	4.16	63.06	6.08	28.78	63.21
Romania	RO	WVS7	63.29	2.48	50.35	37.38	4.20	50.73	5.08	31.50	52.72
Serbia	RS	WVS7	69.43	4.09	82.57	47.68	6.14	79.48	8.95	35.10	80.28
Russian Fed.	RU	WVS7	65.36	2.28	52.99	36.62	4.08	53.15	5.09	29.59	53.22
Sweden	SE	EVS5	66.22	2.46	59.62	38.50	4.26	61.95	6.19	27.10	64.40
Singapore	SG	WVS7	68.75	2.63	56.79	41.61	4.37	56.04	6.81	26.86	55.49
Slovenia	SI	EVS5	70.35	2.99	53.50	44.67	4.75	52.80	7.35	29.08	52.76
Slovakia	SK	EVS5	72.80	3.41	53.09	45.37	5.48	53.70	8.15	31.34	56.14
Thailand	TH	WVS7	84.74	7.97	66.29	69.01	9.83	66.62	18.14	37.59	66.61
Tajikistan	TJ	WVS7	80.86	4.84	78.08	54.78	7.16	78.17	11.52	34.50	77.83
Tunisia	TN	WVS7	60.29	2.98	77.03	35.56	5.29	74.27	6.50	30.62	79.78
Türkiye	TR	WVS7	71.77	3.17	57.06	45.35	5.04	57.27	8.11	28.34	57.60
Taiwan	TW	WVS7	65.35	2.59	54.38	36.62	4.72	51.08	6.15	28.56	51.48
Ukraine	UA	WVS7	65.66	2.35	86.26	34.92	4.52	55.37	4.96	32.77	88.81
USA	US	WVS7	69.10	3.89	69.24	41.10	6.56	67.50	8.32	32.95	67.81
Viet Nam	VN	WVS7	69.49	1.98	51.08	39.92	3.58	63.67	5.61	26.40	66.92
Zimbabwe	ZW	WVS7	63.56	2.92	54.24	35.14	5.42	77.70	5.81	33.34	54.40
Average			68.49	2.90	61.83	41.51	4.81	62.36	6.79	30.25	62.58
Standard Dev.			5.42	0.98	9.92	6.24	1.22	9.81	2.18	3.12	10.29
Min			60.29	1.22	50.21	31.38	2.24	50.11	3.33	21.48	50.20
Max			85.72	7.97	86.26	69.01	9.83	86.50	18.14	37.59	95.01

Table 2 - Reduction in Antagonism ($r=(PA-E^*)/PA$) from Going to Different Numbers of Clusters, using the first two principal components baseline

Country	ISO2	Two Clusters	Three Clusters	Four Clusters	Five Clusters	2-to5 spread
Andorra	AD	42.26	62.48	70.22	76.66	34.40
Albania	AL	33.91	53.38	63.29	70.12	36.21
Armenia	AM	40.69	60.06	73.30	78.27	37.58
Argentina	AR	34.86	61.19	71.90	77.93	43.07
Austria	AT	37.23	59.19	68.68	73.66	36.43
Australia	AU	45.08	63.75	71.47	76.66	31.58
Azerbaijan	AZ	38.02	59.70	70.24	76.32	38.30
Bosnia & Herzegovina	BA	36.38	59.64	69.15	75.52	39.14
Bangladesh	BD	37.44	59.82	70.74	75.61	38.17
Bulgaria	BG	37.57	60.02	67.85	73.77	36.20
Bolivia	BO	37.82	58.36	67.06	72.91	35.09
Brazil	BR	37.39	56.10	66.18	72.41	35.02
Belarus	BY	42.07	57.80	67.94	73.76	31.69
Canada	CA	42.11	63.78	71.88	77.66	35.55
Switzerland	CH	40.23	59.83	69.32	74.81	34.58
Chile	CL	46.64	75.39	82.07	85.49	38.85
China	CN	39.43	55.57	65.53	71.53	32.10
Colombia	CO	57.82	74.02	78.92	83.36	25.53
Cyprus	CY	39.74	64.35	71.57	76.82	37.08
Czechia	CZ	38.30	61.36	68.84	74.10	35.80
Germany	DE	36.96	60.53	70.08	74.76	37.79
Denmark	DK	35.81	57.19	66.45	73.01	37.20
Ecuador	EC	42.37	63.27	70.47	76.17	33.80
Estonia	EE	38.70	61.87	69.20	75.15	36.44
Egypt	EG	35.84	57.23	66.47	73.74	37.90
Spain	ES	50.04	64.98	72.25	76.84	26.80
Ethiopia	ET	34.88	62.10	70.16	76.93	42.05

Country	ISO2	Two Clusters	Three Clusters	Four Clusters	Five Clusters	2-to5 spread
Finland	FI	37.81	59.59	68.06	74.00	36.19
France	FR	39.67	63.43	70.64	77.22	37.55
United Kingdom	GB	39.82	63.84	71.02	76.62	36.80
Georgia	GE	38.90	57.56	69.75	74.81	35.91
Greece	GR	46.47	61.33	69.99	75.58	29.11
Guatemala	GT	54.72	73.71	80.74	84.86	30.14
Hong Kong	HK	49.32	74.45	79.88	83.68	34.36
Croatia	HR	45.74	59.12	68.44	74.92	29.18
Hungary	HU	42.06	61.98	70.64	75.91	33.85
Indonesia	ID	43.54	63.98	71.73	77.06	33.51
Iraq	IQ	37.54	60.15	68.43	74.43	36.90
Iran	IR	47.11	63.45	71.93	77.39	30.29
Iceland	IS	44.43	61.04	71.66	76.08	31.64
Italy	IT	40.43	57.72	67.49	72.78	32.35
Jordan	JO	33.85	53.94	64.04	70.94	37.09
Japan	JP	35.93	56.43	66.56	71.93	36.00
Kyrgyzstan	KG	38.66	58.19	67.20	73.02	34.35
South Korea	KR	51.39	68.01	74.78	79.29	27.90
Kazakhstan	KZ	36.95	58.23	66.77	73.34	36.39
Lebanon	LB	40.00	60.05	72.46	78.09	38.10
Lithuania	LT	35.65	59.05	66.58	72.38	36.73
Montenegro	ME	50.29	70.78	77.87	82.03	31.75
North Macedonia	MK	40.55	56.97	67.07	73.41	32.86
Myanmar	MM	35.85	56.53	66.17	72.26	36.40
Macao	MO	43.74	68.15	76.38	81.38	37.64
Mexico	MX	50.84	70.47	75.78	80.31	29.47
Malaysia	MY	46.57	69.91	75.16	80.05	33.48
Nigeria	NG	37.02	59.29	69.17	74.61	37.59
Nicaragua	NI	50.64	70.28	76.74	81.83	31.19
Netherlands	NL	43.76	65.41	72.21	77.34	33.58

Country	ISO2	Two Clusters	Three Clusters	Four Clusters	Five Clusters	2-to5 spread
Norway	NO	41.42	61.42	71.73	76.60	35.18
New Zealand	NZ	44.17	63.68	71.51	76.50	32.32
Peru	PE	33.81	56.40	65.54	70.97	37.15
Philippines	PH	42.58	64.43	71.63	77.60	35.02
Pakistan	PK	40.91	62.53	72.12	77.73	36.83
Poland	PL	44.27	58.47	68.83	75.04	30.77
Puerto Rico	PR	35.71	59.41	68.66	74.49	38.78
Portugal	PT	40.79	60.42	68.96	75.13	34.35
Romania	RO	38.39	57.55	68.03	73.54	35.16
Serbia	RS	39.70	56.69	66.18	73.27	33.57
Russian Federation	RU	45.73	59.38	68.34	73.39	27.66
Sweden	SE	38.50	59.21	69.12	74.99	36.48
Singapore	SG	41.61	58.84	69.10	74.83	33.22
Slovenia	SI	44.67	61.48	70.44	75.47	30.80
Slovakia	SK	45.37	61.48	71.70	76.94	31.58
Thailand	TH	69.01	76.86	82.21	85.49	16.47
Tajikistan	TJ	54.78	73.18	78.76	83.33	28.55
Tunisia	TN	35.57	62.64	73.89	78.93	43.36
Türkiye	TR	45.35	59.90	69.80	75.46	30.11
Taiwan	TW	36.62	64.08	71.40	77.05	40.43
Ukraine	UA	42.80	63.55	71.93	77.16	34.36
USA	US	41.10	61.80	70.44	75.21	34.11
Viet Nam	VN	40.48	68.65	78.40	82.40	41.92
Zimbabwe	ZW	35.57	64.36	71.65	77.46	41.90
Average		41.70	62.13	70.83	76.30	34.60
Standard Dev.		6.05	5.10	4.02	3.41	4.23
Min		33.81	53.38	63.29	70.12	16.47
Max		69.01	76.86	82.21	85.49	43.36

Table 3 - Differences in Identity Traits between Two Clusters (based on first 2 PCs)

Panel A - Summary Statistics (81 countries, WVS Wave 7 and EVS Wave 5)

	Men (Percent)	Age (Years)	Income (1-10)	Education (Years)	Ideology (1-10)	No Religion (Percent)	Social Class (% in lowest two classes)
Average	7.693	4.985	0.376	0.488	0.649	20.740	4.870
Standard Deviation	5.769	3.788	0.295	0.500	0.567	21.814	6.266
Minimum	0.370	0.144	0.000	0.000	0.000	0.000	0.000
Maximum	22.539	15.806	1.451	2.383	2.136	77.201	26.296

Panel B - Six Selected Country Examples (WVS Wave 7)

	Men (Percent)	Age (Years)	Income (1-10)	Education (Years)	Ideology (1-10)	No Religion (Percent)	Social Class (% in lowest two classes)
Brazil	6.561	8.388	0.589	0.153	0.455	26.452	1.009
China	5.648	7.354	0.405	0.251	-	5.454	0.478
Germany	7.668	3.470	0.214	0.131	0.645	46.563	6.220
Ethiopia	8.411	1.657	0.005	0.667	0.328	0.990	3.424
South Korea	22.373	6.398	0.064	0.135	0.598	74.700	1.844
Nigeria	8.032	1.034	0.108	0.178	0.104	0.271	2.824
USA	14.609	2.027	0.111	1.137	2.136	28.314	9.977

Table 4 - Differences in Cultural Values between Two Clusters (based on first 2 PCs)

Panel A - Summary Statistics (up to 81 countries, WVS Wave 7 and EVS Wave 5)

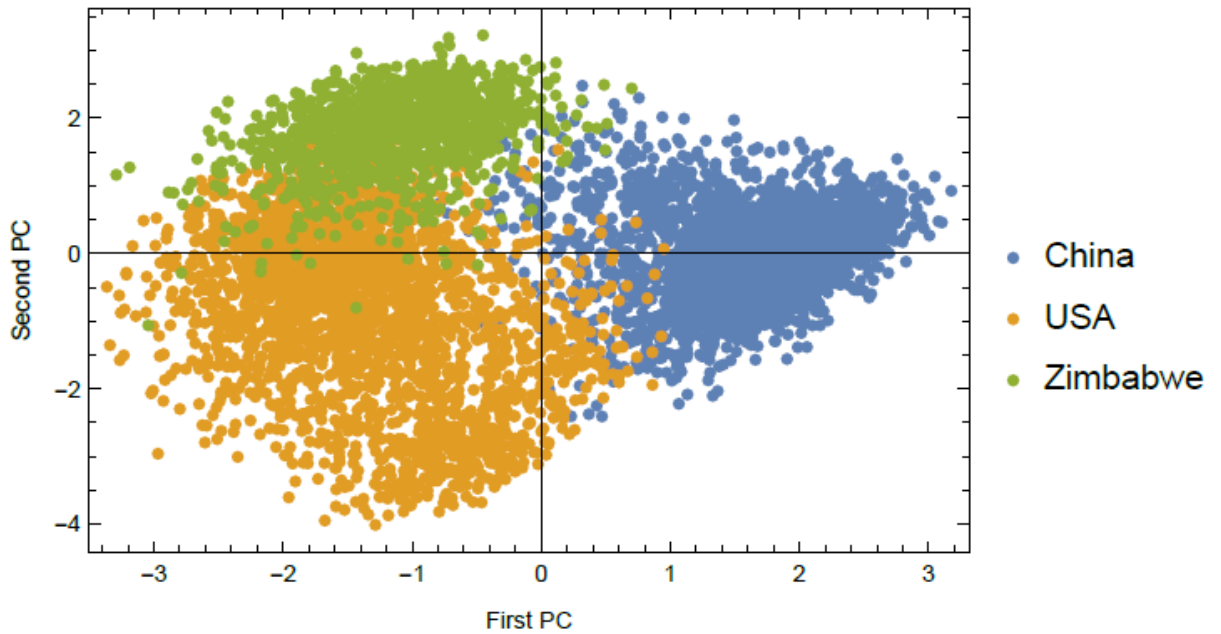
	Important in life: Religion A006	How often do you attend religious services F028	Most people can be trusted A165	Justifiable: Homo-sexuality F118	Justifiable: Abortion F120	Hard work brings success E040	Confidence in government E069_11	Willingness to fight for country E012	Priority to nationals over immigrants for jobs C002	Election officials are fair E265_06
Obs.	81	81	81	78	81	51	80	80	81	78
Average	0.204	0.205	0.088	0.173	0.169	0.065	0.117	0.097	0.085	0.082
St Dev	0.168	0.147	0.078	0.129	0.121	0.049	0.100	0.084	0.075	0.075
Min	0.000	0.000	0.000	0.000	0.003	0.000	0.001	0.001	0.001	0.000
Max	0.552	0.623	0.346	0.495	0.393	0.230	0.485	0.452	0.342	0.386

Panel B - Selected Examples (WVS Wave 7)

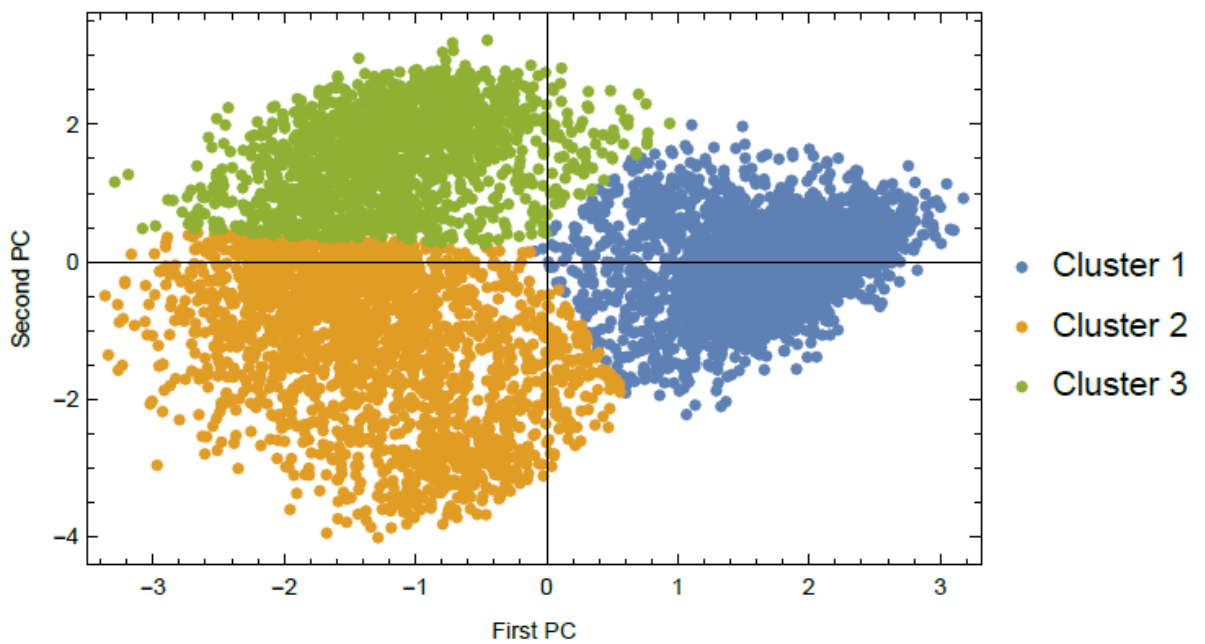
	Important in life: Religion A006	How often do you attend religious services F028	Most people can be trusted A165	Justifiable: Homo-sexuality F118	Justifiable: Abortion F120	Hard work brings success E040	Confidence in government E069_11	Willingness to fight for country E012	Priority to nationals over immigrants for jobs C002	Election officials are fair E265_06
Brazil	0.224	0.320	0.030	0.207	0.206	0.075	0.079	0.122	0.159	0.106
China	0.057	0.019	0.226	0.165	0.138	0.022	0.052	0.010	0.086	-
Germany	0.456	0.362	0.018	0.236	0.316	0.055	0.102	0.083	0.013	0.018
Ethiopia	0.039	0.020	0.050	0.286	0.274	0.129	0.092	0.027	0.070	0.092
South Korea	0.325	0.623	0.078	0.029	0.062	0.033	0.069	0.045	0.005	0.050
Nigeria	0.021	0.061	0.083	0.063	0.063	0.036	0.123	0.024	0.001	0.005
USA	0.520	0.371	0.181	0.436	0.391	0.111	0.237	0.179	0.117	0.090

Figure 1 – Validation: Recovering Countries from Pooled Data from China, the US and Zimbabwe

Panel A – Individuals from the Three Countries along the First and Second PCs



Panel B – Individuals from the Three Endogenous Partitions along the First and Second PC

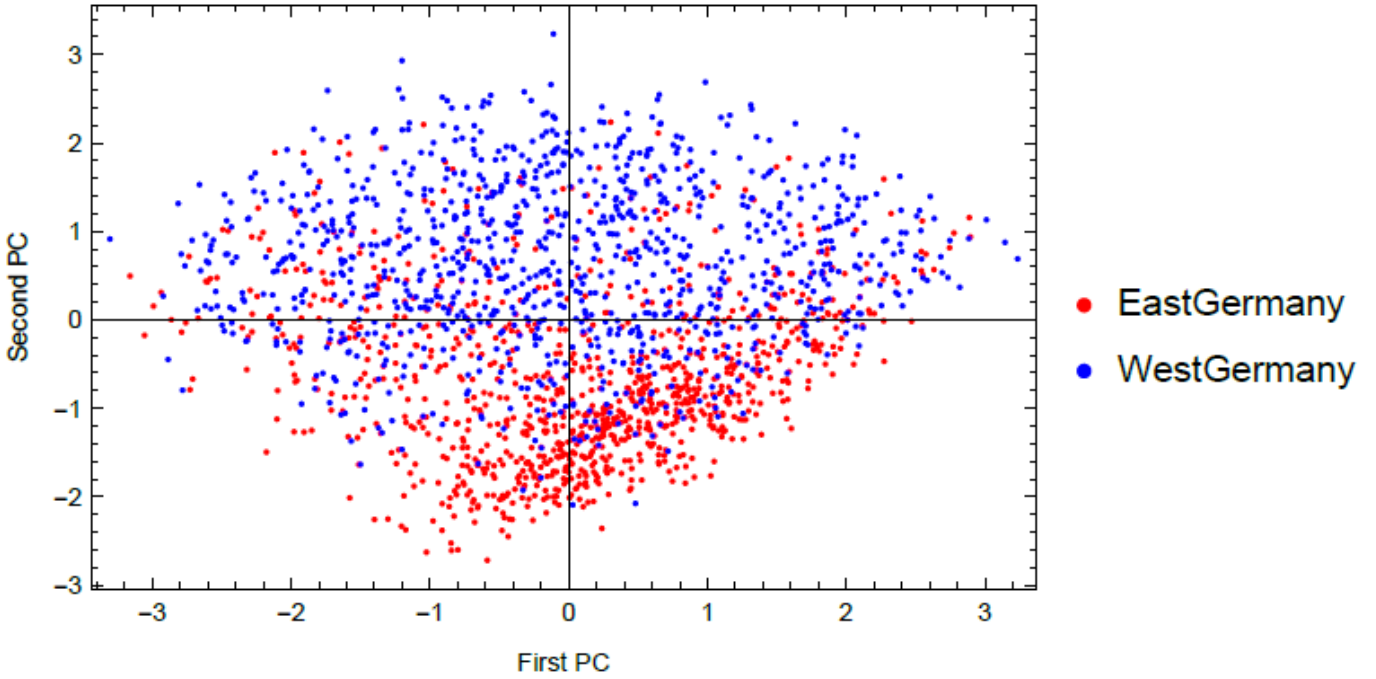


Percentages of respondents from each country that belong to each cluster:

□	China	USA	Zimbabwe
Cluster 1	96.5	2.2	0
Cluster 2	1	79.5	1.7
Cluster 3	2.4	18.2	98.2

Figure 2 – Validation: Detecting East and West Germany

Panel A – East and West Germans along the First and Second PCs, Wave 3



Panel B – East and West Germans along the First and Second PCs, Wave 7

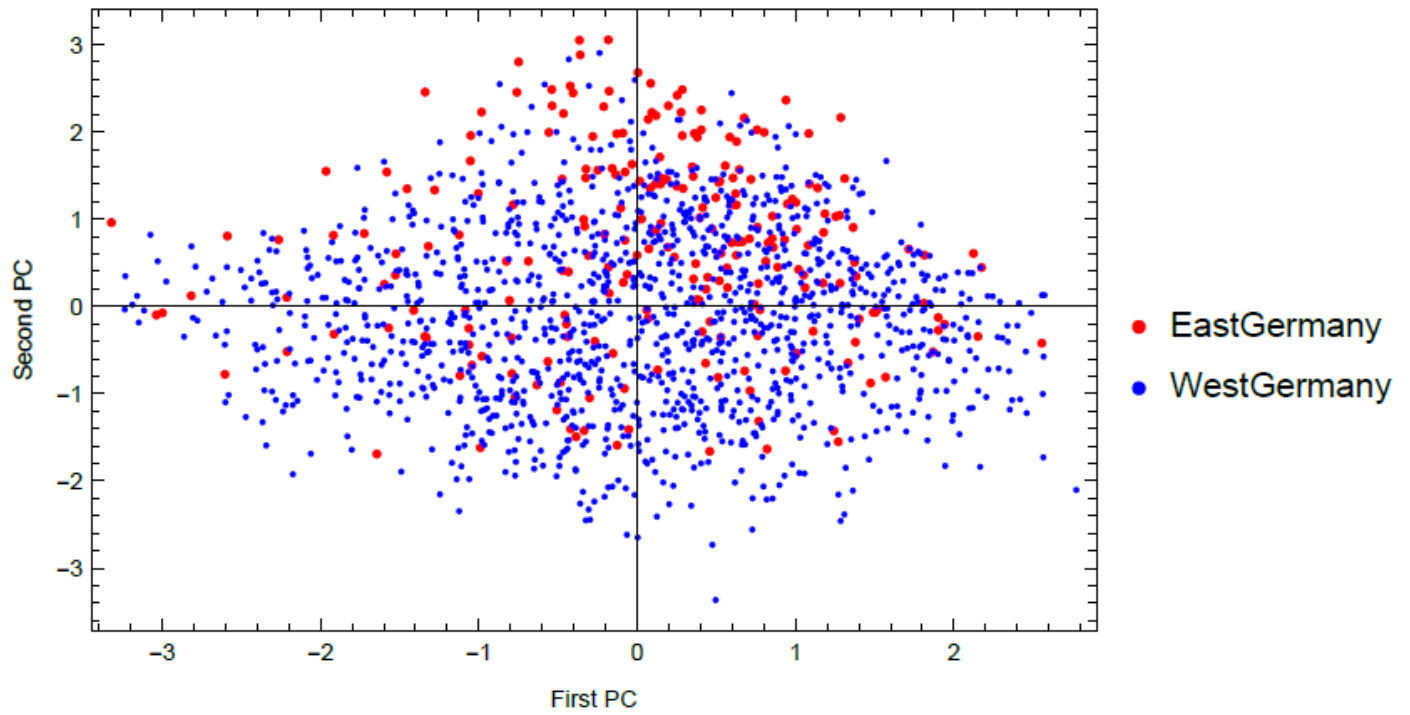
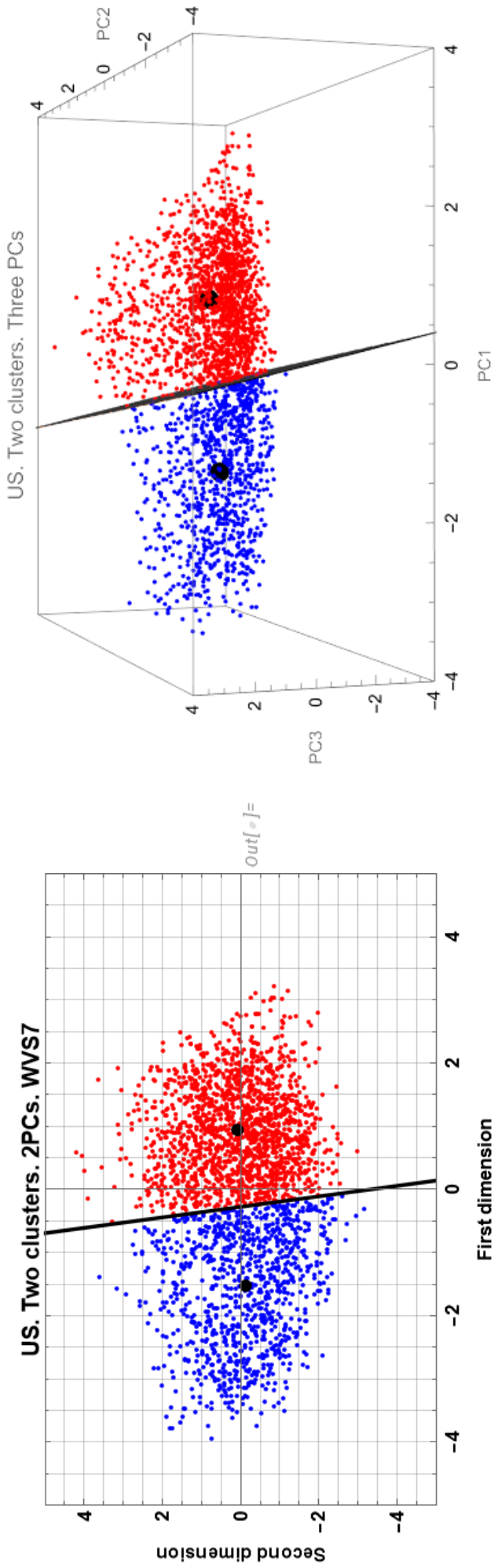


Figure 3 – Endogenous Partitions for a Selection of 7 Countries: The Case of 2 Clusters

Panel A – USA, 2 and 3 Principal Components



Panel B – Selection of Other Countries (Wave 7)

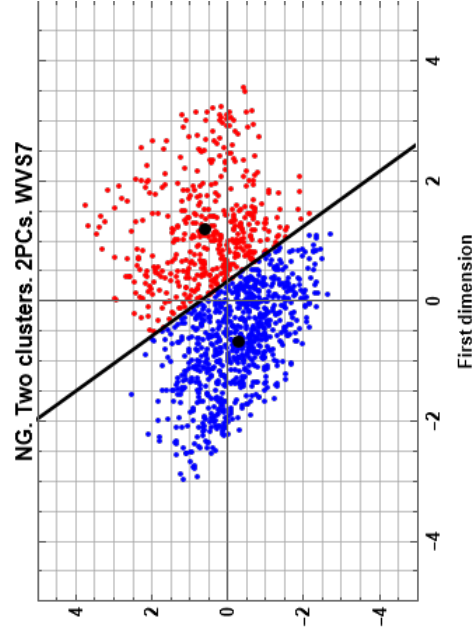
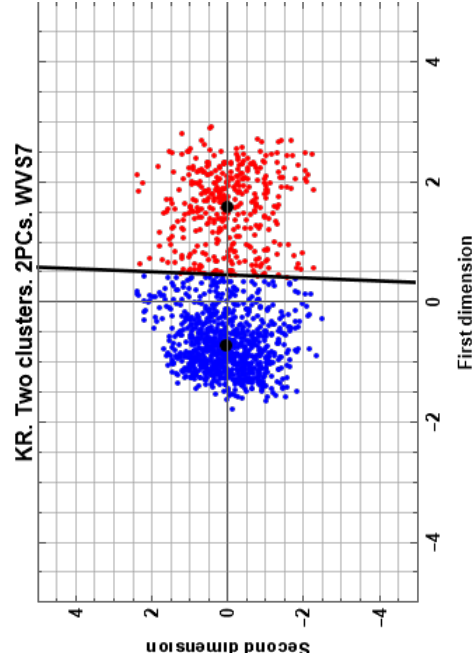
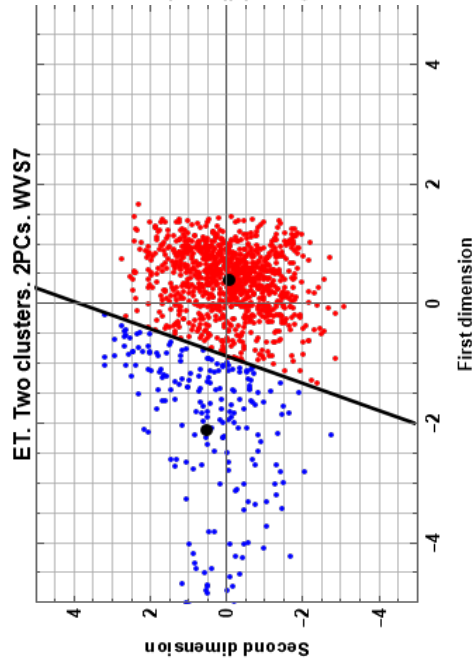
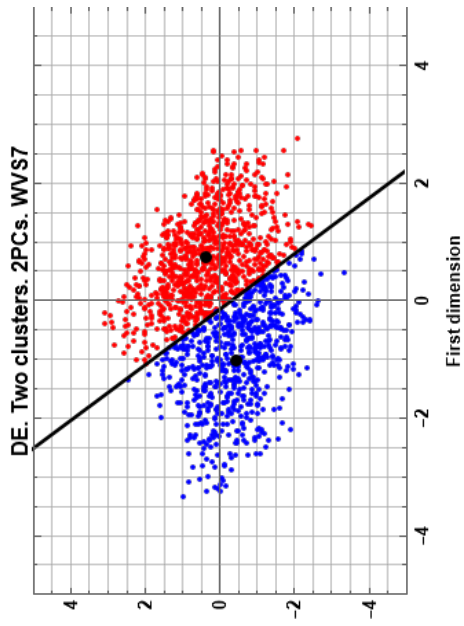
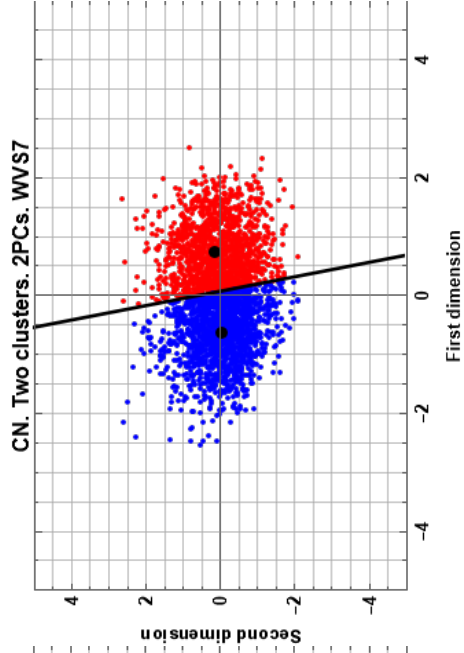
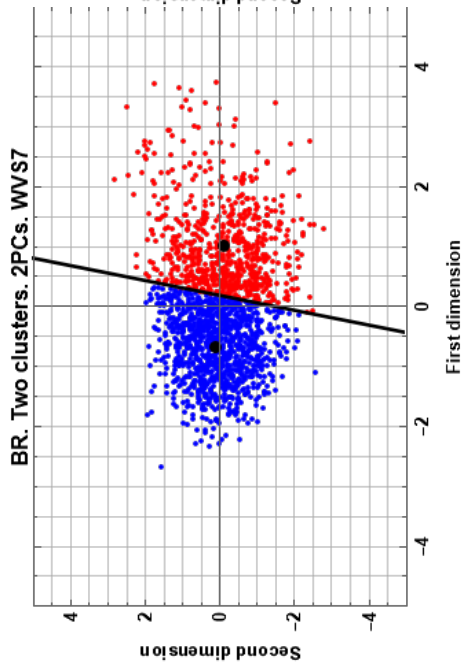
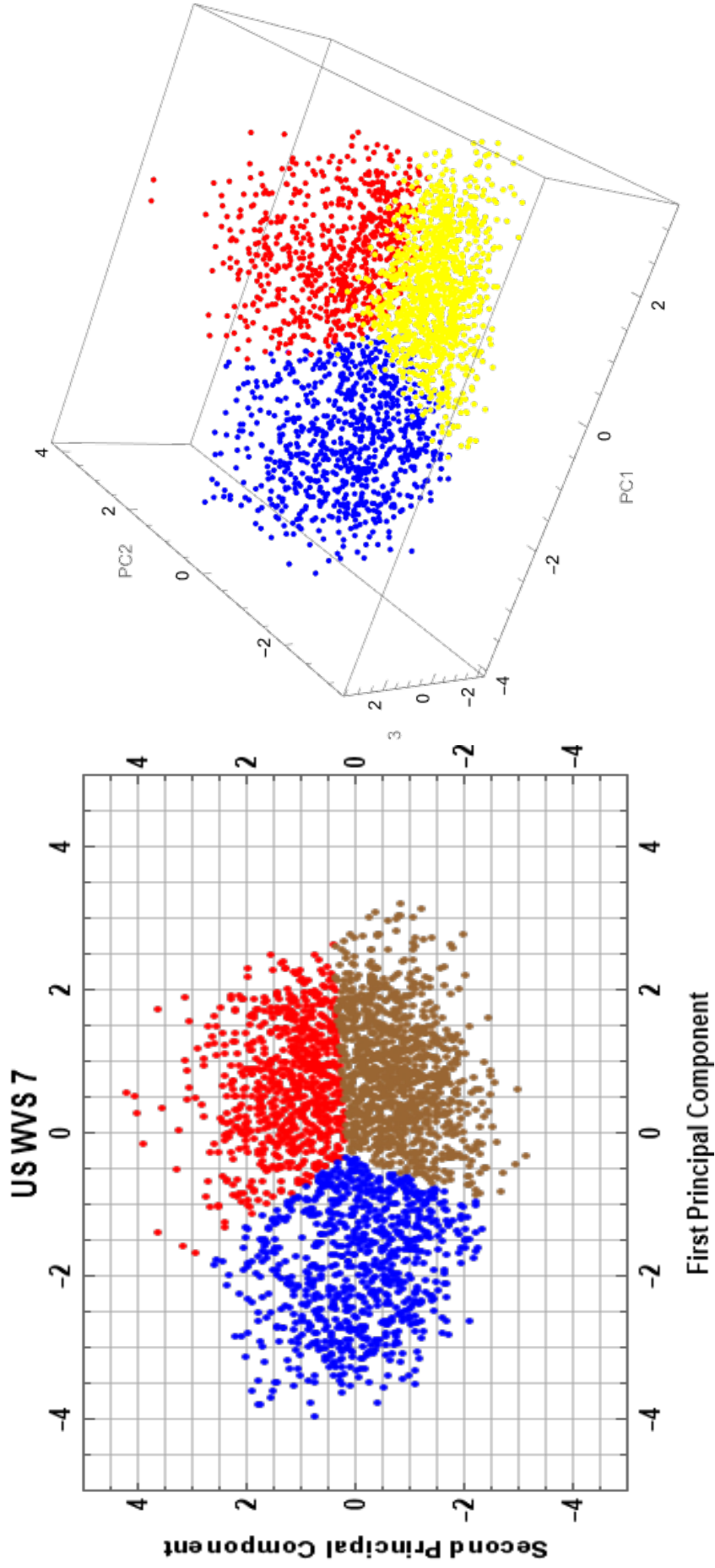


Figure 4 – Endogenous Partitions for a Selection of 7 Countries: The Case of 3 Clusters

Panel A – USA, 2 and 3 Principal Components

US. Two clusters. Three PCs



Panel B – Selection of Other Countries (2 PCs)

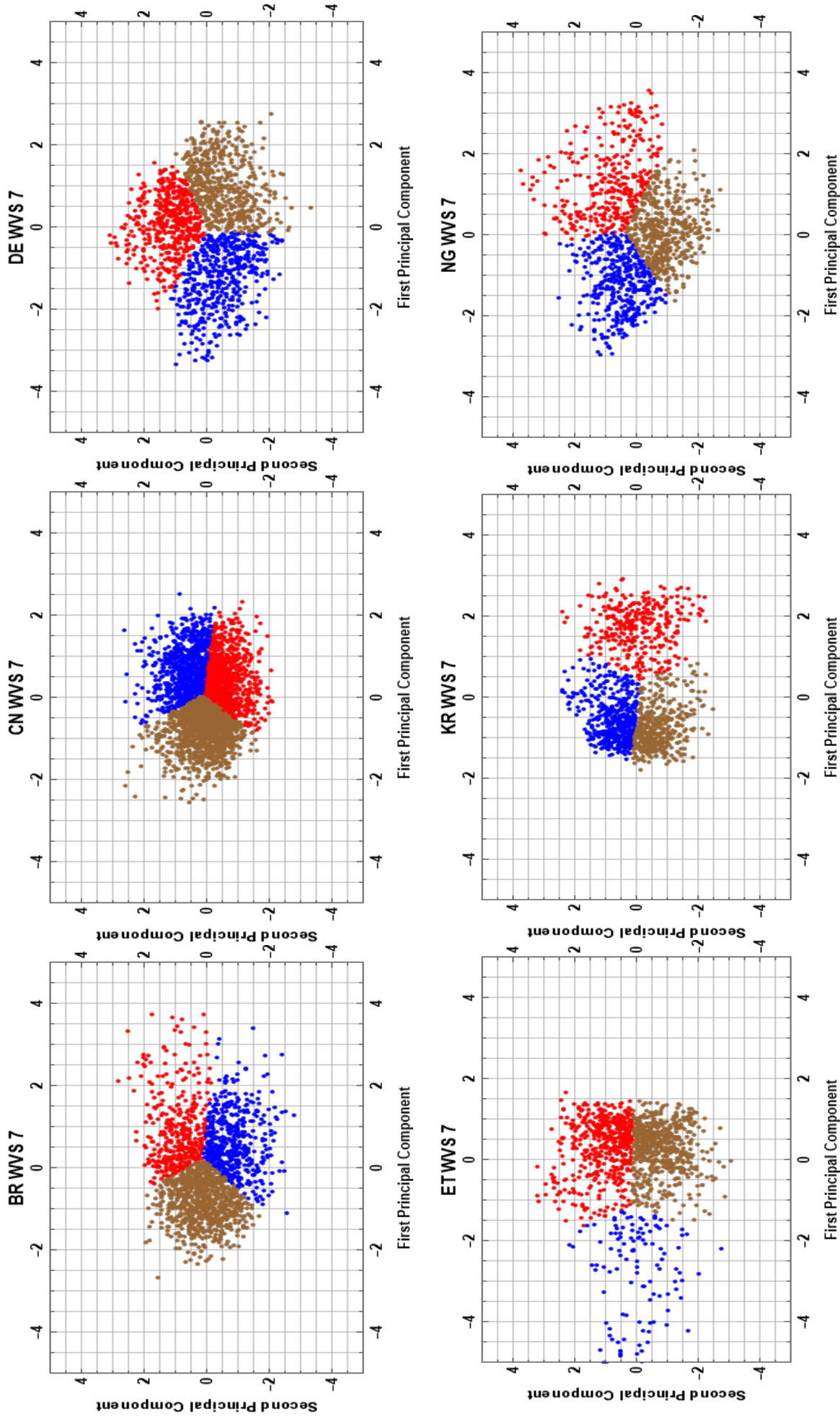
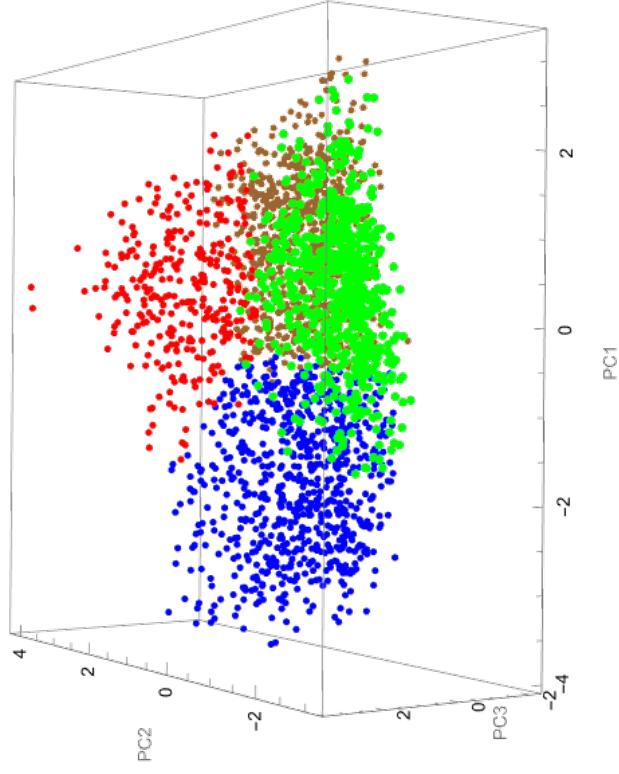
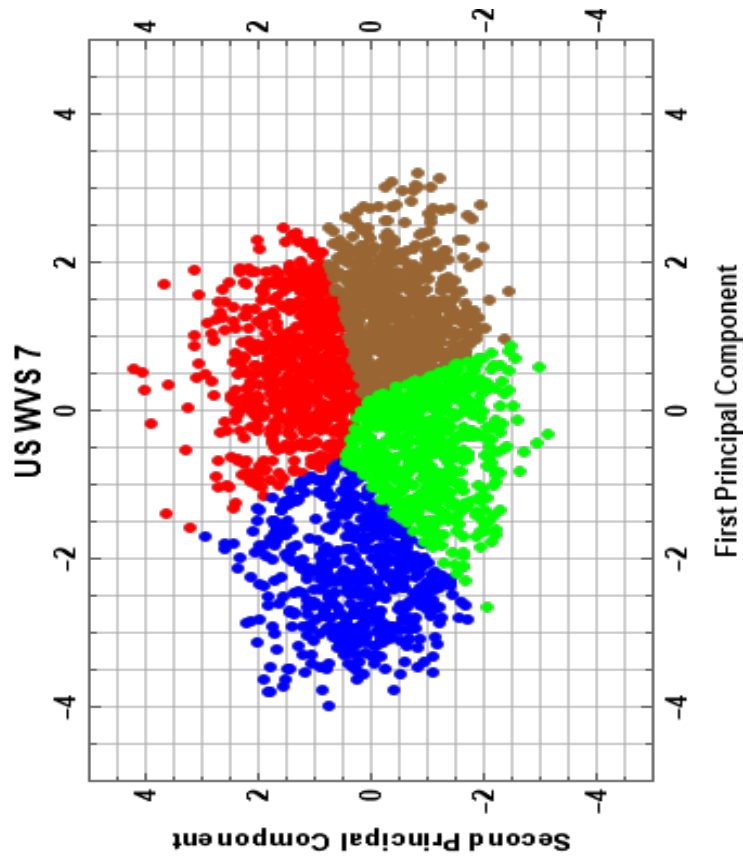


Figure 5 – Endogenous Partitions for a Selection of 7 Countries: The Case of 4 Clusters
Panel A – USA - Two and Three Principal Components

US: Four clusters. Three PCs



Panel B – Selection of Other Countries (2 PCs)

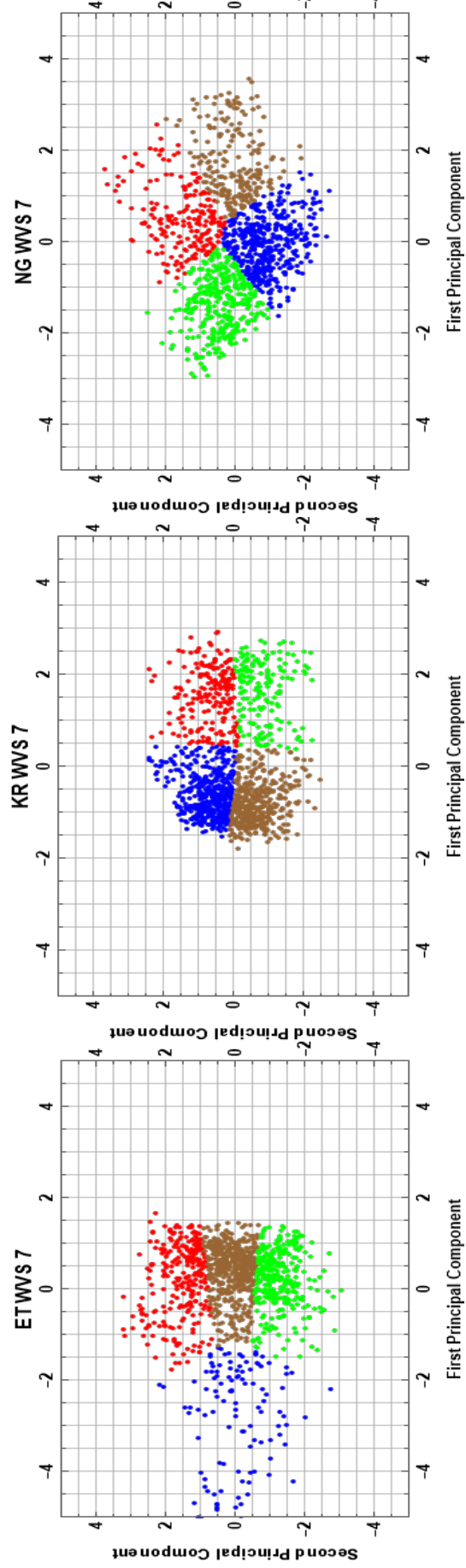
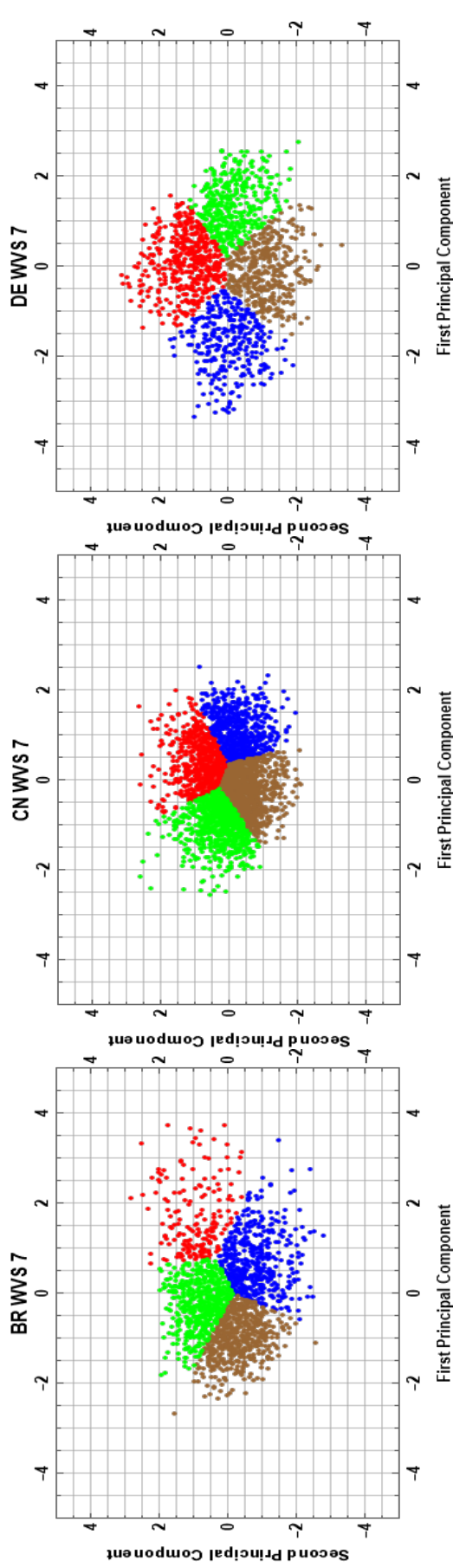
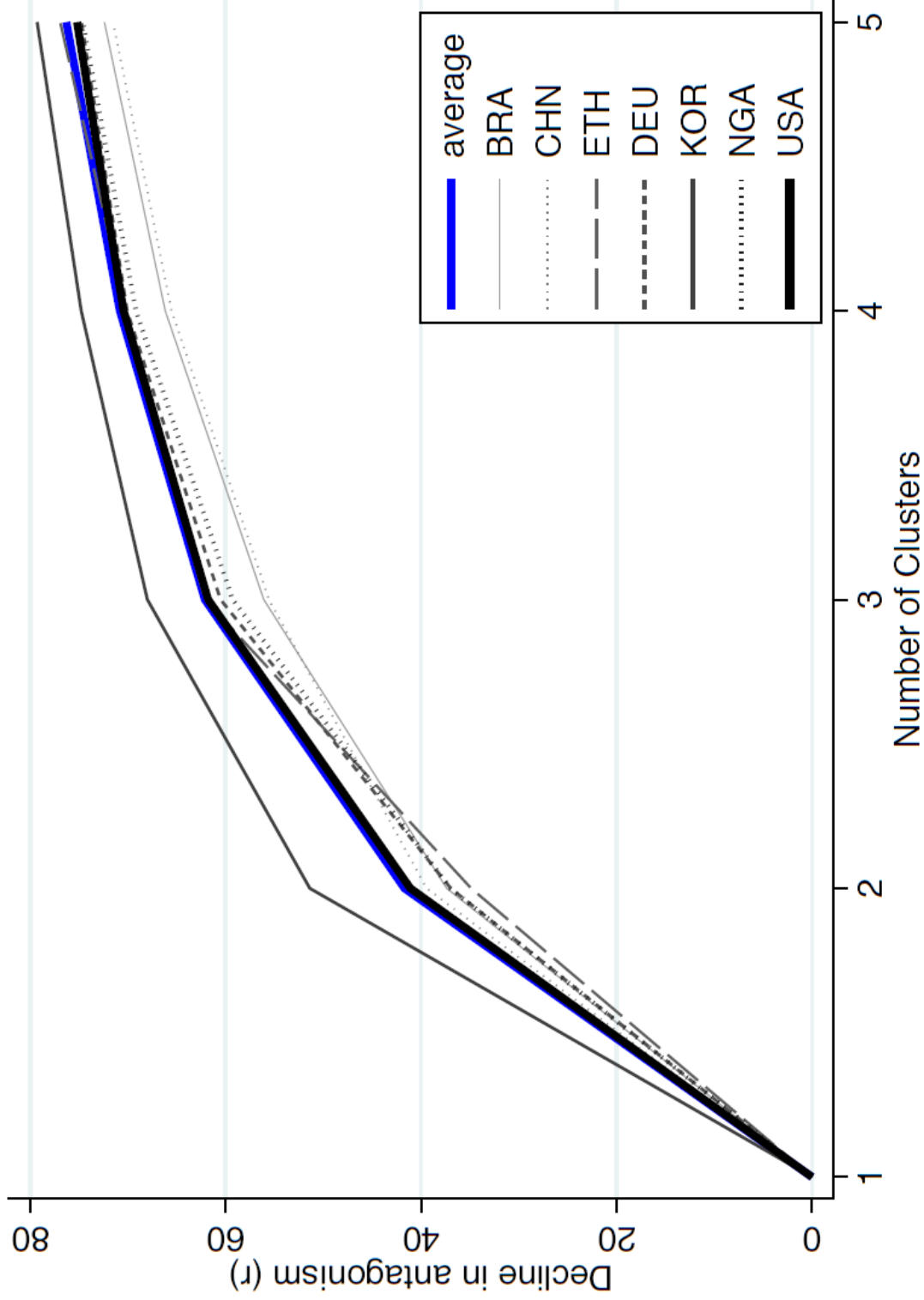
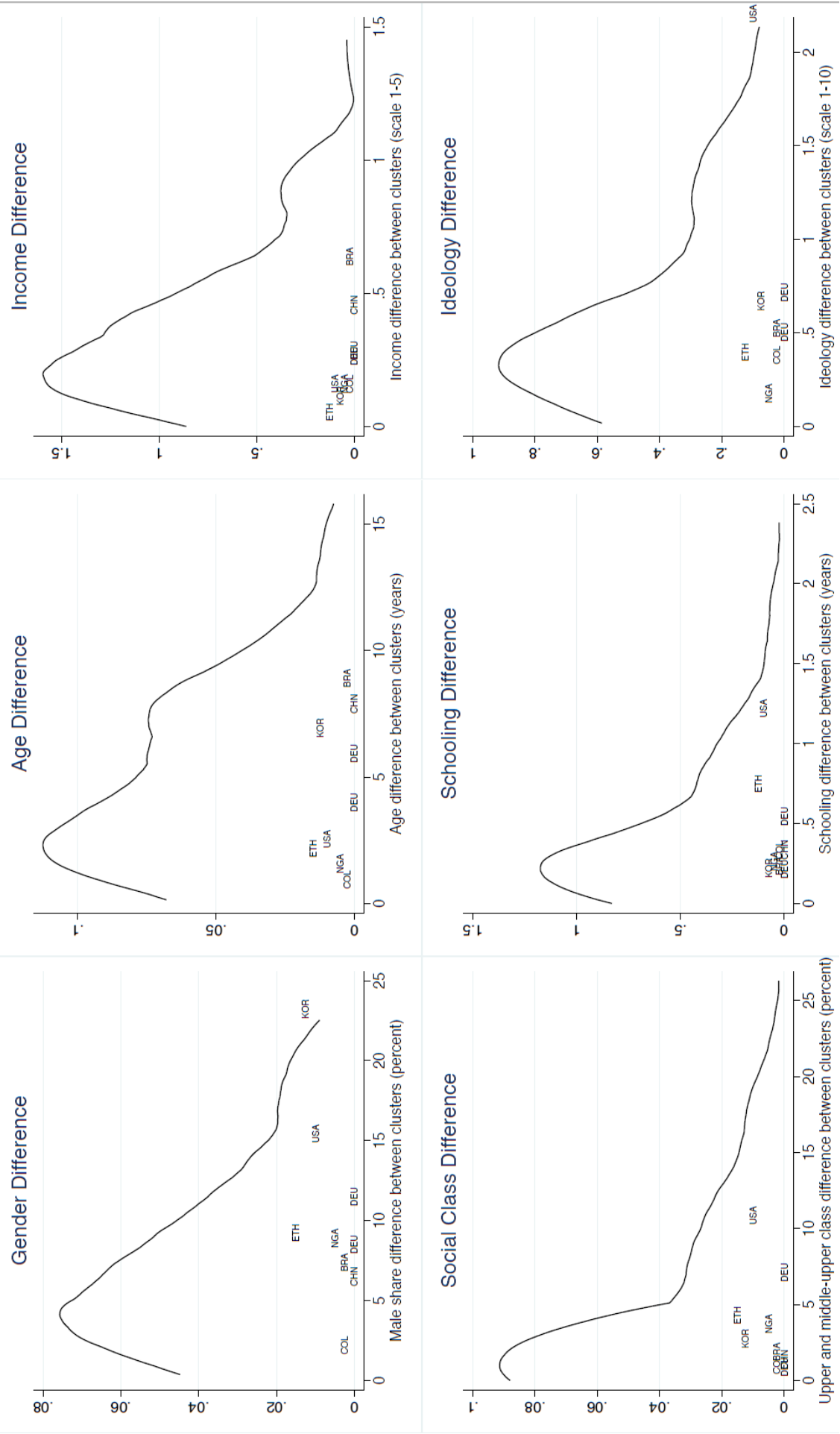


Figure 6 – Reduction in Within-Group Antagonism ($r=(PA-E^*)/PA$), for various numbers of clusters and for 7 illustrative countries



- Note:
- Clusters are based on 2 principal components
 - Average refers to the average reduction in within-cluster antagonism over the 81 countries in sample.

Figure 7 – Identity Differences between Clusters: Distribution across Countries, case of 2 Clusters and 2 PCs (81 countries)



Note: Country labels are for illustrative purposes and refer to their values along the x-axis

Figure 8 – Values Differences between Clusters: Distribution across Countries, case of 2 Clusters and 2 PCs (81 countries)

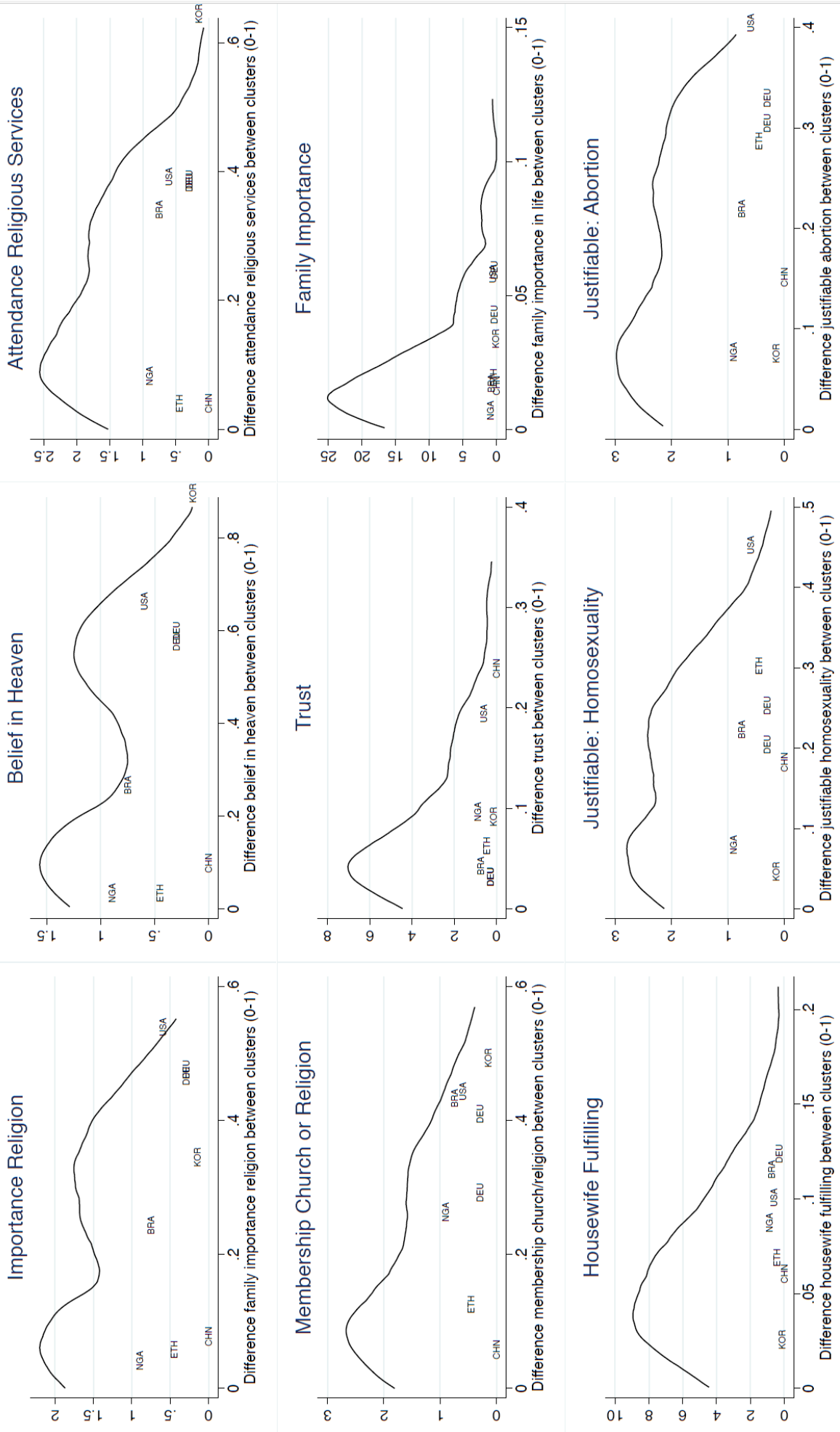


Figure 8 (contd.)

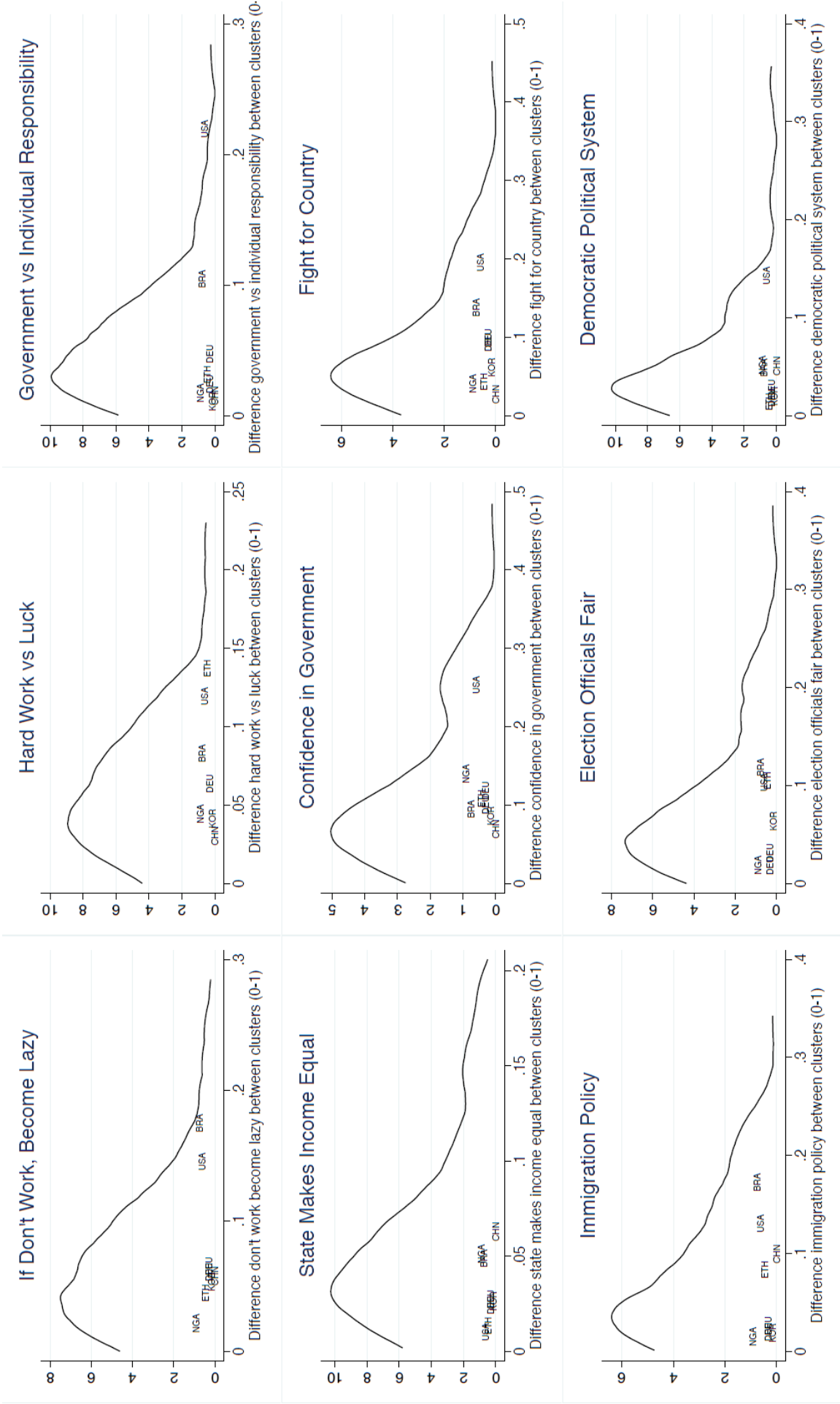


Figure 9 – Reduction in Antagonism between Clusters in the US: Evolution over Time (two-clusters, various PCs)

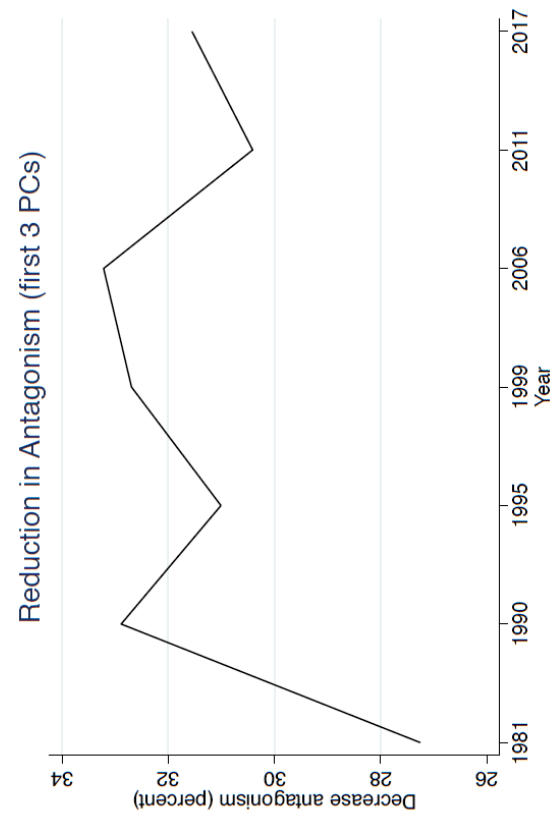
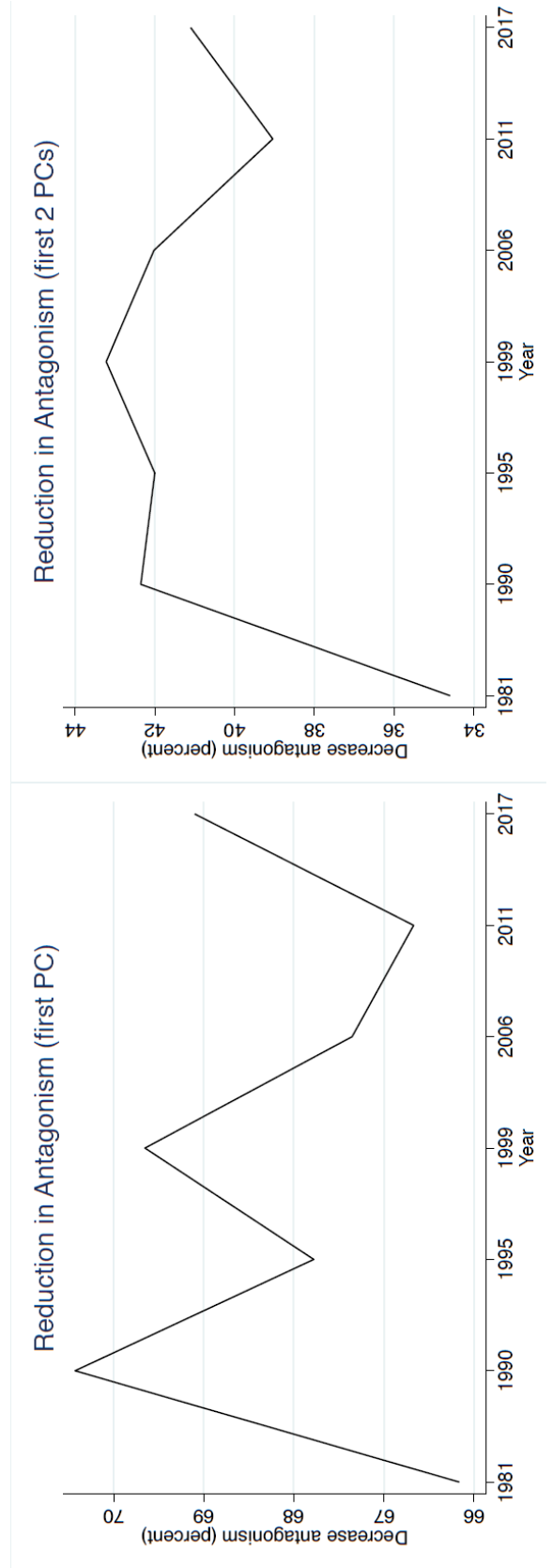


Figure 10 – Identity Differences between Clusters in the US: Evolution over Time, case of 2 Clusters and 2 PCs

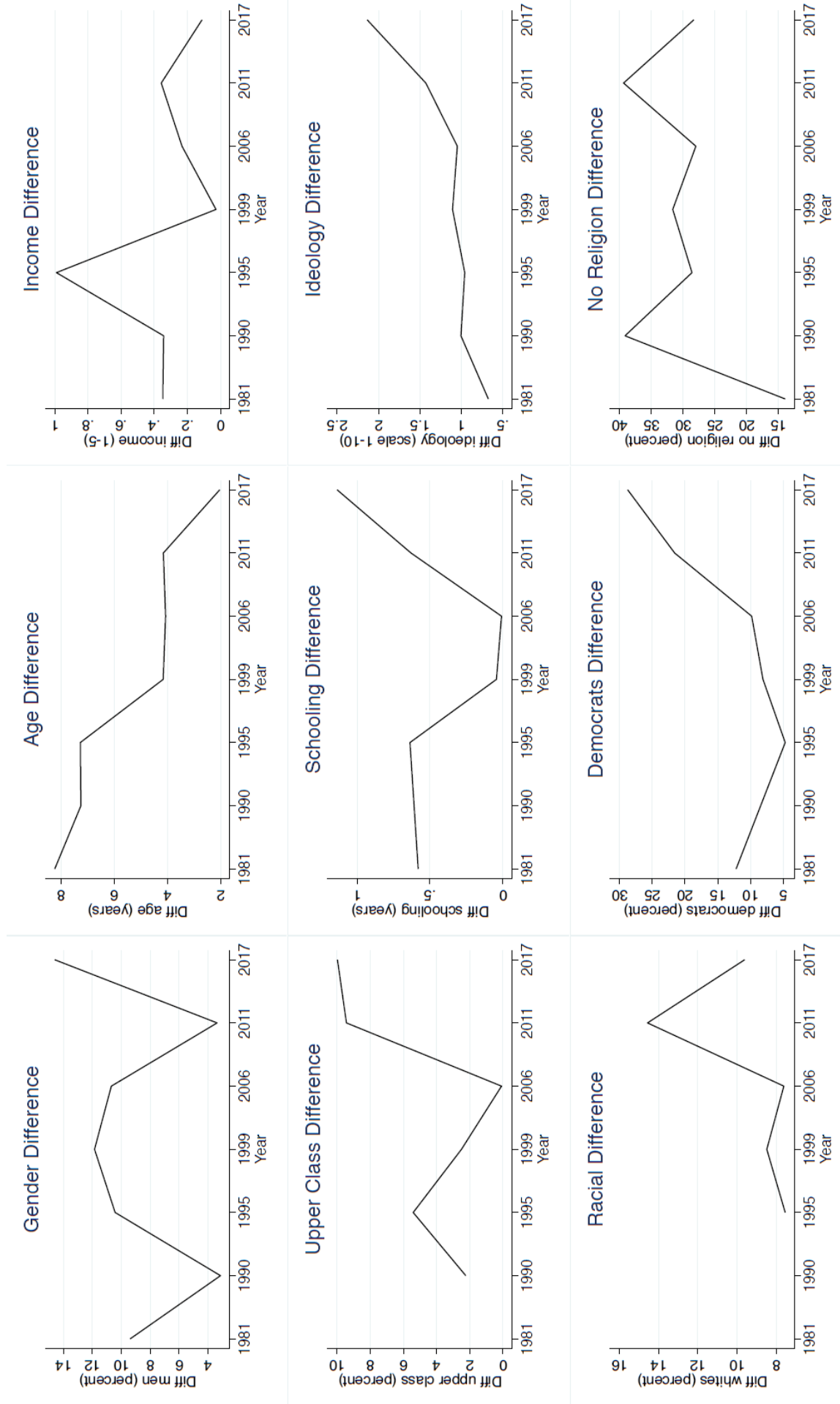


Figure 11 – Values Differences between Clusters in the US: Evolution over Time, case of 2 Clusters and 2 PCs

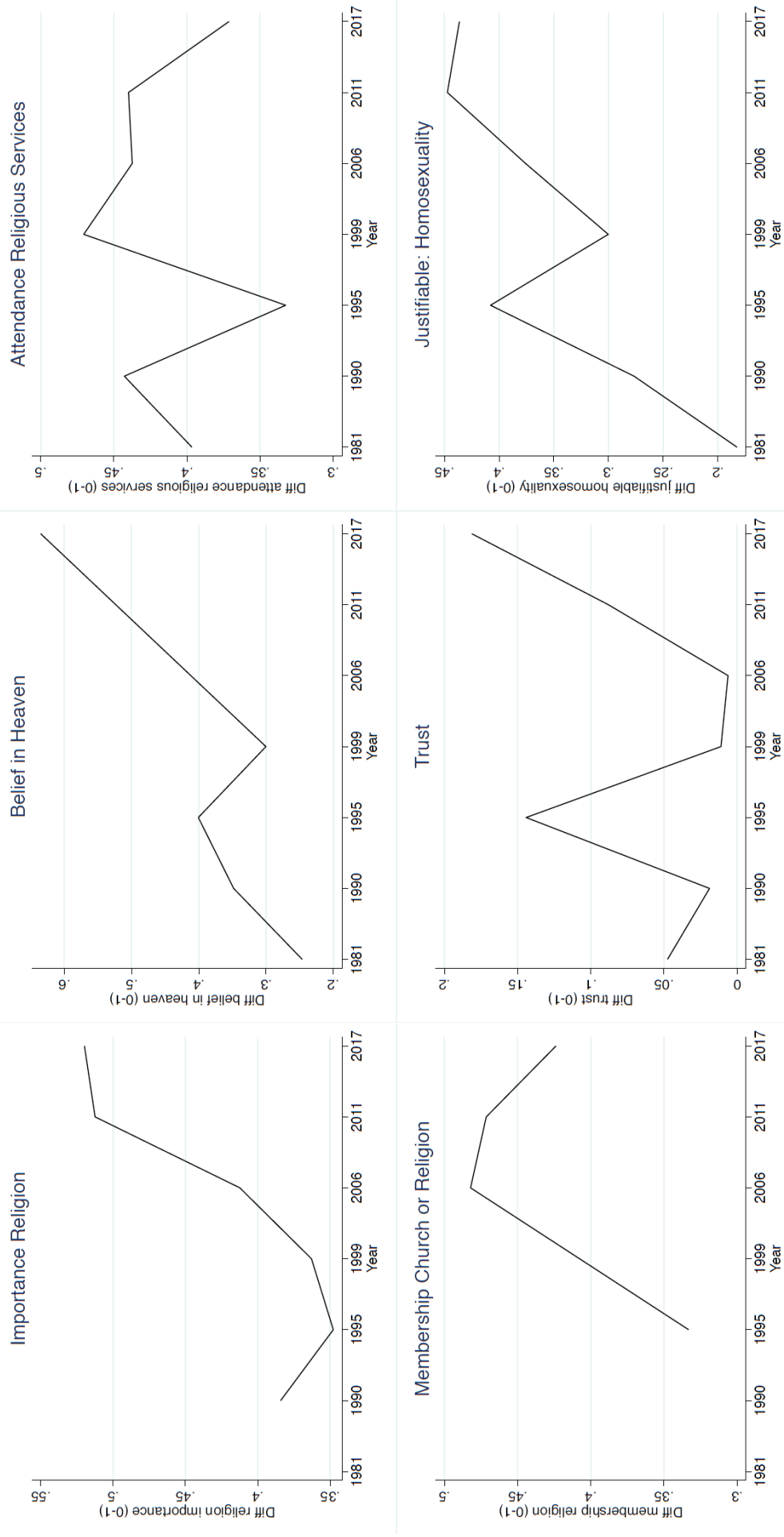


Figure 11 (contd.)

