NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Maksim Rudnev, Vladimir Magun, Peter Schmidt

THE STABILITY OF THE VALUE TYPOLOGY OF EUROPEANS: TESTING INVARIANCE WITH CONFIRMATORY LATENT CLASS ANALYSIS

BASIC RESEARCH PROGRAM

WORKING PAPERS

SERIES: SOCIOLOGY
WP BRP 51/SOC/2014

Maksim Rudnev[1], Vladimir Magun[2], Peter Schmidt[3]

THE STABILITY OF THE VALUE TYPOLOGY OF EUROPEANS: TESTING INVARIANCE WITH CONFIRMATORY LATENT CLASS ANALYSIS[4]

Unlike variable-centered measures, validity and stability of typologies have been rarely studied. Magun, Rudnev and Schmidt [in review] developed a value typology of the European population using data from the 4th round of the European Social Survey. The value classes showed heuristic power in the comparison of different parts of the European population, countries in particular, enabling more differentiated interpretations in a parsimonious way. The current paper tests the stability of this typology by extending the study to three time points – consecutive surveys in 2008, 2010 and 2012. Conceptually, this test coincides with measurement invariance testing. We reviewed the levels of typology measurement invariance. Then, the invariance of the value typology of Europeans was tested across three rounds of the ESS and it was found to hold configural and partial invariance. The reliability of the value classes was supported by the stability of country class probabilities across the time points as well. The correlations of the country shares of the value classes with the economic development of countries are also invariant at the three time points. The results imply that the value classification of Europeans is not ad hoc, but reflects the natural structure of European societies, and can be used in future studies.

JEL Classification: A13.
Keywords: basic human values, latent class analysis, measurement invariance, heterogeneity, European population

[1] National Research University Higher School of Economics, Laboratory for Comparative Studies of Mass Consciousness, senior fellow, Email: mrudnev@hse.ru
[2] National Research University Higher School of Economics, Laboratory for Comparative Studies of Mass Consciousness, director, and Institute of Sociology of Russian Academy of Sciences, head of department. Email: maghome@yandex.ru
[3] University of Giessen, Faculty of Social Science. Email: peter.schmidt@sowi.uni-giessen.de
[4] The research of the first two authors leading to these results has received funding from the Basic Research Program at the National Research University Higher School of Economics. We are grateful to Lisa Trierweiler for the valuable remarks and suggestions.

1. Introduction

The formation and testing of typologies has long been a neglected issue but is becoming a fast growing area of social science research [Hagenaars and Halman, 1989; Hagenaars and McCutecheon, 2002; Hancock and Samuelsen, 2008]. This method has three advantages for the analysis of values:

- First, in contrast to variable-centered methods like factor analysis, it is a holistic approach. Typologies capture the whole system of values by classifying people into classes instead of looking at the scores of distinct items, scales or latent variables;
- Second, because people are classified into types on the basis of all the item scores taken together, it is a parsimonious method;
- Finally, the differentiation between types provides a natural criterion for studying within-country value heterogeneity.

A neglected topic in research has been the validity and reliability of classifications. Nearly all typologies until now are used ad hoc and in a descriptive way and are never used again [Finch and Bronk, 2011]. For example, Lee et al. [2011] developed a typology based on a modified Schwartz instrument, but it is unknown whether this typology can be reproduced with other samples, with a second wave of a panel study, or by using other Schwartz instruments. The same problem is relevant in Klages and Gensicke's [2005] study. In other words, the validity and reliability of these classifications are questionable since they were not assessed. Typologies lacking these attributes may lead to oversimplification (in the case of artificial classification) or data-driven conclusions (in the case of natural one) that may be wrong due to inductive generalizations or due to random fluctuations in empirical data, respectively. Such typologies do not allow and do not intend to test explicit hypotheses in a confirmatory way, since their nature is predominantly exploratory. To the best of our knowledge, the development of typologies with proven validity and reliability, used more than once and by more than one author, is very rare in the social sciences. This is very unlike the variable-centered approach in which repeated assessments of measurement properties are widespread. Examples are the Big Five personality instrument and the Schwartz value measures [Schwartz, 2007; Schwartz et al., 2012]; they have been used and reproduced by hundreds of authors, and their validity and reliability are continually assessed and discussed.

Magun, Rudnev and Schmidt [in review] developed a classification of Europeans based on their values assessed in the 4th round of the European Social Survey (ESS). The purpose of this paper and its added value is to determine how robust this classification is, or, in other terms, how invariant it is. We aim to extend the validity and the robustness of the specific classification to several time points with different samples of the ESS. It is a simultaneous comparison over three time points from 2008 to 2012. The objective is to test the robustness of the initial typology across different samples of the European population assessed at different time points.

In their study, Magun, Rudnev and Schmidt classified European respondents using Schwartz's Portrait Values Questionnaire (PVQ) data gathered within the 4th round of the ESS in 2008. The data from 28 countries were pooled, weighted by their population and design weights, and classified using latent class analysis (LCA). Five value classes were found, and the first values class was labeled Growth values. Its members emphasize the importance of Openness to Change as well as Self-Transcendence values. The members of the other four classes are somehow in opposition to the Growth class and are aligned to the Social Focus – Personal Focus dimension. After determining the value classes, the authors demonstrated that every country has a share of almost every value class. The membership of the Growth values class was highly correlated with the economic development of the country, and this correlation was even higher than the correlations of the single value variables with economic development. Membership in the other four classes was higher in less economically advanced countries, and its correlations with economic development were weak and negative. To assess the robustness of the new typological approach to studying values across time points, we test the invariance of the five value classes across three time points.

Due to the existence of an exploratory study conducted by Magun, Rudnev and Schmidt for the 4th round of the ESS, it is possible to test explicit hypotheses about value typology for the 5th (2010) and 6th rounds (2012) of the ESS. Our main hypothesis is that the initial class solution, with all its properties, is robust across three time points. The first two hypotheses refer to the dimensionality and the reliability of the value class solution itself; they extrapolate the features found for the 4th round data to the 5th and 6th round data of the ESS.

H1. There are five value classes in Europe.

H2. The substantial differences between value classes are the same as have been found in the previous study, i.e. the Growth values class, the Strong and Weak Social Focus classes, and the Strong and Weak Person Focus classes.

The next two hypotheses concern the reliability of the relations between the latent classes and external variables, namely, respondent country of residence and country level of economic development. We expect that these relations discovered in the 4th round of the ESS and indicating external validity of the class structure remain the same in the 5th and 6th rounds of the ESS.

H3. The relations between the country shares of the Growth values class and the level of economic development are stable across rounds and are strongly positive. The country shares of the other value classes are negatively and weakly related to country economic development.

The period between the 4th and two subsequent rounds of the ESS was a time of economic crisis and included some time interval shortly after the crisis (2008-2012). Although values are considered to be stable, it is possible, that the crisis affected the distribution of country populations between classes. Extrapolating relations between the level of economic development and the size of the Growth values class, there is a chance that after the economic crisis took place, a share of the Growth values class decreased. Still, this is very unlikely, especially in such a short-term perspective. Attitudes, not values, are prone to change in response to changing situation, they are seen as less stable than values [Eagly and Chaiken 1993]; see, for example, the study of the effects of the economic crisis on attitudes toward immigration [Billiet, Meuleman, and de Witte, 2014]. Thus, the next hypothesis states the stability of country values.

H4. The shares of the value classes in European countries are approximately the same in the 4th, 5th and 6th rounds.

The rest of the paper is organized in three sections. In the next one, we discuss value measures, procedures of classification and levels of invariance; in the third section we test the invariance of classes across the three ESS rounds; and in the 4th section we relate an outcome invariant classification to external variables, namely, country and its economic development, in order to prove the robustness of external validity of the typology.

2. Data and Methodology

2.1. Data

The analyses are based on data from the 4th, 5th and 6th rounds of the European Social Survey (2008-2012) for 32 European countries [Jowell, Roberts, Fitzgerald, and Eva, 2007]. The data for 22 countries were available for all three ESS rounds and included in the present analyses. In addition, data from Croatia, France, Greece and Ukraine were available for the 4th and 5th ESS rounds; data from Latvia, Romania, Turkey and Lithuania for the 4th round; data from Lithuania for the 5th round; and data from Iceland and Kosovo for the 6th round. For a full list of countries, see Appendix 3. In total, data were available for 155,467 respondents. The samples of individuals were the national representative ones. The sample of countries was not random, hence, it has certain limitations in representing Europe in its entirety. The sample excludes 2,402 respondents (1.6%) who did not reply to value questions. We included in our analysis only three of the six available ESS rounds, mostly because of the technical limitations: a model that uses a numerical integration in combination with a very large sample size results in a very high computational load. From the substantial point of view, we believe that the three most recent rounds were enough to test the stability of typology.

2.2. Value measures

We employed Schwartz's approach to studying values, since it was used in the initial Magun et al. paper and because it is up to date theoretically, and an easily measureable concept. Following Schwartz, basic values are "desirable trans-situational goals, varying in importance, that serve as guiding principles in the life of a person or other social entity" [Schwartz, 1994, p. 21]. Values differ by the type of goal that they express, so values can be differentiated by an underlying goal. The central idea of Schwartz's theory is continuity of value universe and stability of relationships between values in most cultures in the world. These ideas are best represented by a value circle separated into sectors, where each sector designates a value (see Figure 1). Adjacent values in this circle share the same motivational emphases and are, therefore, compatible, while values that are further away from one another are less related or even conflicting [Schwartz, 1992]. Following the idea of continuity of values, any number of distinct values can be potentially measured depending on the instrument. Initially, Schwartz distinguished 11 basic values, but later this number changed several times. For the ESS he postulated 10 values [Schwartz, 2007].

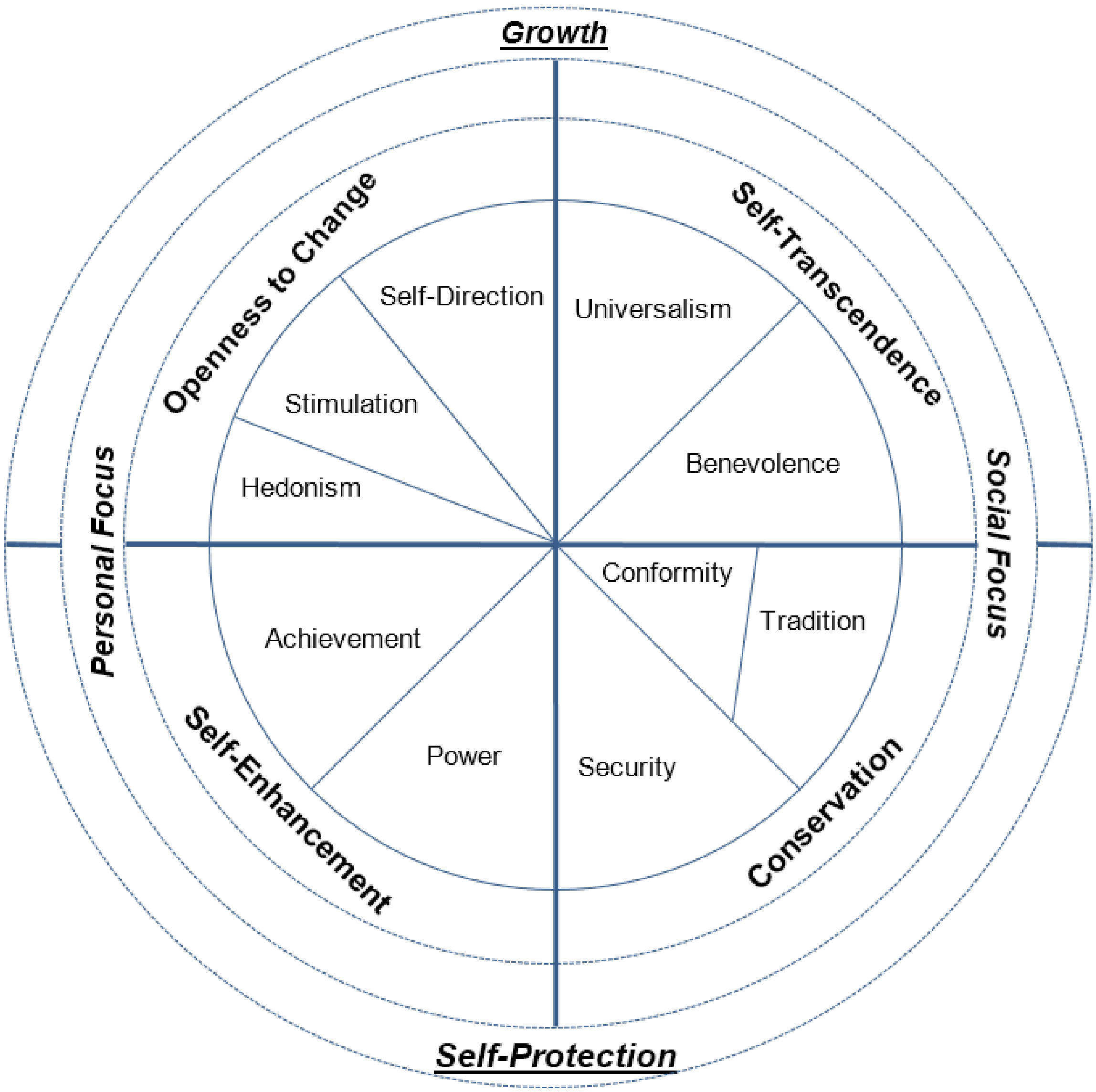

Figure 1. Schwartz value circle depicting the relations between 10 values and several value groupings [Schwartz, 1992, 2006].

Values were measured by a modified version of the Portrait Values Questionnaire (PVQ-21) developed by Schwartz [Schwartz et al., 2001; Schwartz, 2005]. Respondents were provided with 21 descriptions of people for whom different things were important, and they assessed each of the portraits using a 6-point scale ranging from "very much like me" (6 points) to "not like me at all" (1 point). The full wordings of the value portraits, as well as labels of the items used throughout the paper, are listed in Appendix 1. The PVQ-21 was designed to measure the 10 basic values which are calculated on the basis of the 21 initial items [Schwartz, 2007]. Given the dynamic relations between basic values, the same items can be used to calculate the four higher-order values and the higher-order value dimensions of Conservation – Openness and Self-Enhancement – Self-Transcendence. The scores for the two value dimensions are calculated by subtracting the individual score for Conservation from the Openness score and the score for Self-Enhancement from the Self-Transcendence score. Hence, the two value dimensions measure a preference for Openness over Conservation and for Self-Transcendence over Self-Enhancement.

2.3. Statistical procedure

To classify the respondents on the basis of their values, we used the LCA technique, first introduced by Lazarsfeld and Henry [1968]. Compared to classical clustering methods such as *k*-means, LCA is a model-based technique which takes into account measurement error, uses a probability-based approach instead of *ad hoc* criteria to estimate cluster centers, and provides a formal statistical test of the number of latent classes. LCA allows the researcher to identify a set of discrete latent classes from observed indicators [McCutcheon, 1987; Muthén and Muthén, 2010]. LCA has three types of parameters:

1. the fundamental one – the number of classes;
2. the response probabilities for each of the classes; and
3. the probabilities of classes themselves.

Response probabilities are the key parameters in LCA which define the class by representing the chances for the respondents of a given class to choose one of the responses. Probability of the class is different from response probability and refers to the size of class.

In the following analysis, LCA is based on the 21 Schwartz value items, which were treated as ordinal variables. To adjust for an individual response style influencing a person to use a certain part of the rating scale (e.g. assigning only low, high, or medium ratings to all the questions), Schwartz suggested the so-called centering procedure [Schwartz, Verkasalo, Antonovsky, and Sagiv, 1997]. Following this procedure, each value score for each individual respondent is centered by subtracting the individual average for all the 21 value items from the raw score. However, it requires the assumption that a 6-point Likert-type scale has an interval level of measurement. Recently, instead of centering, Schwartz and co-authors used a method factor that loaded on all the value items [Schwartz et al., 2012]. Adding a method factor (or random intercept, as referred to by Vermunt [2010]) to LCA allows for controlling an individual response style. Although the introduction of a method factor is more complex than centering, it does not require the assumption of the scales' continuity and it corrects for response style, keeping the initial distributions of respondent answers [Billiet and McClendon, 2000; Lubke and Muthen 2004; Van Herk, Poortinga and Verhallen, 2004]. We extended the classic LCA model by adding a method factor. All the 21 loadings of this factor were fixed at zero and the factor mean was fixed at one. The variance of the method factor for all classes and rounds was set free. Every LCA model, including LCA models with covariates, described below has this method factor.

In our LCA procedure, the data were weighted with the population weights, since we were interested in determining the all-European latent class structure and not in simply classifying the respondents in the sample. The population weight reshapes the sample of respondents to make proportions of respondents from different countries equal to the proportions of populations of these countries. The results, which were obtained using design weights only, or no weights at all, were very similar to those presented here, although conceptually it is more reasonable to use population weights, since it allows extrapolating results to most of the European populations.

The data were weighted by the design weight as well. Design weights correct for differences in probabilities of respondent selection, "thereby making the sample more representative of a 'true' sample of individuals aged 15+ in each country” [Weighting European Social Survey Data, 2013]. Therefore, this enhances the equivalence of samples across countries.

2.4. Levels of typology invariance and confirmatory latent class analysis

The general purpose of the establishing a level of measurement invariance (or equivalence) is to estimate the degree to which “the instrument measures the same concept in the same way across various subgroups of respondents” [Davidov, Cieciuch, Meulemann, Schmidt, and Billiet, 2014, p. 9]. There are several levels of measurement invariance of typologies across different groups, e.g. across different ESS rounds [Eid, Langeheine, and Diener, 2003; Kankaras, Moors, and Vermunt, 2011; Siegers, 2011].

Full invariance (structurally homogeneous model) holds when a number of classes and all the thresholds (or response probabilities) of all classes are the same across all groups. This situation is hardly empirically tenable, although highly desirable, since it fully proves the robustness of the typology.

Full invariance of specific classes is held when only some of the classes have the same response probabilities across groups. Unlike multiple group confirmatory factor analysis, it is not necessary to keep all the classes the same across groups in order to be able to compare group shares of some of the classes. That is, if a researcher has a substantial interest in only one class, and if the response probabilities for the members of this class are equal across groups, this class can be claimed robust and invariant regardless of the number and response probabilities of the other classes.

Partial invariance is another way to deal with the data in case the full invariance was not confirmed. It is similar to partial factor invariance (either metric or scalar). In this case, a researcher may allow some response probabilities to be different across groups [Eid et al., 2003]. Steenkamp and Baumgartner [1998] suggested that, for the factor models, two items’ loadings or two intercepts that are equal across groups are enough to keep the latent factor unbiased at the metric and scalar invariance level, respectively. However, it is not clear how many response probabilities should be held equal and how many may be allowed to vary across groups in order to keep the class membership unbiased. Further statistical experiments are needed to determine this.

Partial invariance of specific classes is an even lower level of invariance that is held when only some classes have some response probabilities that are equal across groups.

It is possible to test other intermediate levels of invariance between the full invariance and no invariance. Sometimes equality constraints in testing measurement invariance are referred to as too strict and unrealistic, since they require exact equality between parameters across groups [Davidov et al., 2014]. It is possible to get *approximate invariance* for each of the levels that does not require the strict equality of probabilities across groups, instead it allows for a small difference between probabilities across groups. The range of response probability differences across classes should be set based on former studies or substantial theorizing. This approximate invariance has been initiated in the context of Bayesian approaches [Muthén and Asparouhov, 2013] in which a researcher should set the prior probabilities' variance of differences between parameters across classes.

Configural invariance (or *construct equivalence,* or the *heterogeneous model* as referred to by McCutcheon, 1987) means a similarity of a general configuration of class response probabilities across groups. It can be assessed in two ways: with independently estimated models in each group or with a single model that allows differences between groups, i.e. the multiple group LCA model or an LCA with a group as a predictor covariate without restrictions. Configural invariance implies satisfying two requirements: there should be the same number of latent classes in each group and similar patterns of class response probabilities in all groups. The literature does not discuss statistical criteria for the similarity of patterns; we suggest using a correlation of class profiles (i.e. the whole set of response probabilities of a class) between groups and a comparison of response probabilities between-classes ranks across groups.

Configural invariance of specific classes. Sometimes it is not even necessary to obtain the same number of classes to proceed with invariance testing; such a situation is possible when it is important, from a substantial point of view, to test invariance of some classes only [Kankaras et al., 2011]. In this case, a different number of classes are allowed in different groups, and it has already been shown that full invariance does not hold.

No invariance occurs when the classes obtained in different groups with the same items have notably different response probabilities, which leads to a different meaning of classes across groups.

A traditional procedure of invariance testing is described by McCutcheon [1987], who suggested simultaneous or confirmatory multiple group LCA. Confirmatory LCA (CLCA) is a relatively rarely used method that mimics the logic of a confirmatory factor analysis. Until now there have been few studies using CLCA beyond a couple methodological applications [cf. Eid et al., 2003]. To test the invariance, a multiple group LCA, which builds several LCA models in all the groups simultaneously, is computed. It allows for setting different kinds of constraints, mainly equality of response probabilities of the corresponding classes across groups [Kankaras, 2011, Siegers, 2011]. However, we found the multiple group approach to be computationally too demanding so, in this paper, we turned to a group-as-covariate approach. Instead of treating the group variable as an indicator of a group in a multiple group LCA, we added a dummy variable for each group except the reference one as a predictor of value items given the value class and (sizes of) latent classes themselves in a single-group LCA. The chosen model is more parsimonious since it estimates the unified item response probability for all groups together and the effect of group, whereas the multiple group LCA estimates response probabilities for each group separately. A drawback of the group-as-covariate approach is that all the groups are compared to the reference one and are not compared to each other. This problem is easy to resolve if we have a small number of groups by changing the reference group and repeating the computations: in this case, the model fit stays exactly the same and the parameters reflecting necessary differences are estimated. However, this strategy could be tedious when there are many groups to consider.

The strategy of analysis includes the comparison of the fit statistics for models with different sets of constraints. A model corresponding to a configural level of invariance does not constrain the effects of group and thus gives a general overview of the degree of higher levels of invariance: non-significant effects of a group variable indicate invariance of an item's class response probabilities between reference group and the other groups, significant effects indicate non-invariant items. Testing of the higher levels of invariance involves constraining some or all the effects of group to be zero (this tests the hypothesis that the group has no effect on some or all of the response probabilities). A model selection problem is found in the fact that the fit statistics are not standardized, so judgments about which model is the most appropriate can only be made based on the relative values of the fit indices and the likelihood. Specifically, the comparison is done using the likelihood ratio test with a scale factor correction implemented for likelihoods and obtained with the maximum likelihood robust (MLR) estimator. However, some authors have pointed out that the likelihood ratio test has a low power in large samples, thus high sample size can make the test significant [Kelloway, 1995]. This is why the LRT test must be used cautiously with large sample sizes.

We started with the estimation of the number of classes using exploratory LCA models in three groups (i.e. ESS rounds) independently, and if it was confirmed to be invariant across rounds, we compared the response probabilities by correlating class profiles and ranks of the items between classes. Then we proceeded with the confirmatory approach, assessing configural (or heterogeneous or unrestricted) invariance with a single group LCA model including the variable “ESS round” as a predictor. This was used as a baseline model and provided hints when choosing a set of constraints. Next, the fully invariant (homogeneous) model was estimated, and if it was significantly worse than the configural (heterogeneous) model, we had to introduce a subset of constraints freeing the parameters that appeared non-invariant in the configural model estimates, and fixing the ones that were invariant to be equal across groups.

The models were computed using an analysis of the mixture type in the Mplus software version 7.11 [Muthén and Muthén, 2010] and maximum likelihood robust estimation, which is robust to non-normality and non-independence when estimating standard errors and chi-square statistics. By default, Mplus uses full information maximum likelihood for treatment of missing values.

When assessing classification invariance within the multiple group LCA framework, both Kankaras [2011] and Siegers [2011] were interested in finding a class solution that would be comparable across countries. Our case was different. First, we were not interested in cross-country comparability, since the typology we were looking for was pan-European. Second, we were interested in testing a certain class solution across time points. The grouping variable was the ESS round, which was the time when the data were gathered. This is why we emphasized the comparisons of the prototypical solution based on the data from ESS round 4 [Magun et al., in review] with the latter rounds' solutions and looked for the extent to which this original solution held in the data of the 5th and 6th ESS rounds.

3. Results

3.1. Testing the number of value classes across the three ESS rounds

In order to identify an optimal number of classes, 10 similar models were computed differing only in a number of classes, i.e. from 1 to 10. This was repeated for each ESS round separately. The fit statistics are listed in table 1. This part of the study was conducted in an exploratory way; however, its purpose was confirmatory, testing the hypothesis of whether there are the same number of classes in each of the three ESS rounds data. (Alternative hypotheses include an indeterminate number of solutions with the number of classes other than 5, so it was not possible to perform this test in a fully confirmatory way).

The usual way of identifying the number of classes is by choosing a model with the lowest Bayesian information criterion (BIC) or Akaike information criterion (AIC), where the smaller values of these indices point to the better fit of the model. In the present analysis, each step which adds one more class to the model leads to smaller BIC and AIC. At the same time, the reduction in the BIC and the AIC becomes increasingly smaller with every step, which makes it hard to determine whether the decrease of BIC and AIC values is substantially important or not. For these reasons, we applied the Vuong-Lo-Mendell-Rubin (VLMR) likelihood ratio test, a measure that provides a formal testing of the difference in model fit [Lo, Mendell, and Rubin, 2001]. The VLMR test identifies whether the fit of a model with *k* classes is significantly higher than the fit of a model with *k-1* classes. If the former is not higher, it is not necessary to add an extra class and, following the parsimony rule, we can conclude that *k-1* is the optimal number of classes for a given LCA model. Therefore, significant values of the VLMR test show that the *k*-1 class solution is no better than the *k* class solution, thus, the *k*-1 model is the one to choose.

The significance of the VLMR test presented in Table 1 demonstrates a very similar pattern in each of the three ESS rounds: it is significant until the number of classes is 6. When the number of classes is 6, the VLMR becomes insignificant at the 0.05 level, indicating that the 6-class solution does not have a better fit than the 5-class solution. Fewer than 5 classes is not a choice as well, since the models with 4 classes or less have significantly poorer model fit. Therefore, the 5-class solution is optimal for all three ESS rounds. The entropy measure demonstrates a degree of certainty of classification, and this value becomes lower in solutions with more than 5 classes, indicating the appropriateness of the 5-class solution as well.

Taken altogether, we can conclude that the 5-class solution is the best solution for each of the three ESS rounds. This finding was confirmed with the tests that are independent and exploratory in nature.

Table 1. Fit statistics for exploratory LCA models obtained separately from the 4th, 5th and 6th ESS rounds data. Each row represents an independent model

Number of classes	Number of parameters	Log-likelihood	AIC	BIC	Entropy	Significance of likelihood ratio VLMR test (*p* values)
ESS Round 4 (2008)						
1	106	-1699838	3399888	3400834	-	-
2	213	-1634999	3270424	3272325	0.81	0.00
3	320	-1609434	3219508	3222364	0.81	0.00
4	427	-1589266	3179386	3183197	0.81	0.00
5	534	-1580538	3162145	3165213	0.81	0.00
6	641	-1573548	3148377	3154098	0.80	0.56
7	748	-1567665	3136826	3143501	0.80	0.58
8	855	-1563055	3127820	3135451	0.79	0.37
9-10	Models did not converge					
ESS Round 5 (2010)						
1	106	-1690709	3381631	3382577	-	-
2	213	-1630935	3262296	3264197	0.80	0
3	320	-1605615	3211870	3214726	0.80	0
4	427	-1588096	3177046	3180857	0.80	0

5	534	-1580586	3162240	3167005	0.80	0
6	641	-1573726	3148734	3154455	0.79	0.74
7	748	-1475234	2951965	2956208	0.79	0.10
8-10	Models did not converge					
ESS Round 6 (2012)						
1	106	-1391768	2783747	2784673	-	-
2	213	-1345113	2690652	2692513	0.78	0
3	320	-1326422	2653485	2656280	0.79	0
4	427	-1310896	2622646	2626376	0.79	0
5	534	-1303815	2608697	2613362	0.79	0.01
6	641	-1298253	2597787	2603387	0.79	0.41
7	748	-1293584	2588665	2595199	0.78	0.49
8	855	-1290278	2582265	2589734	0.79	0.76
9-10	Models did not converge					

Note: *AIC – Akaike information criterion, BIC – Sample adjusted Bayesian information criterion, Entropy – a measure of uncertainty of classification.*

3.2. Testing the content of the value classes across the three ESS rounds

Configural (heterogeneous) models. As we found the same number of classes present in all three ESS rounds, we now turn to examining the similarity of their content. First, we assess the response probabilities from three independent exploratory models and then repeat the analysis using a single confirmatory model that uses the ESS round as a covariate.

Class profiles, i.e. the whole set of response probabilities, were compared for the similar classes across the three ESS rounds. The correlations are very high, ranging from 0.976 for the Strong Personal Focus class in rounds 5 and 6 to 0.997 for the Strong Social Focus class in rounds 4 and 5. Hence, the value profiles of the classes are very alike for the three ESS rounds. Figure 2 demonstrates cross-round similarity between the classes as described by average scores on the two higher-order value dimensions. The averages for all the classes are rather similar although there are fluctuations between rounds. The Weak Personal Focus class is the most stable one, the Strong Personal Focus and Growth classes show a little fluctuation, and the two Social Focus classes demonstrate larger fluctuations between rounds.

For reasons of simplicity, we will not describe the differences between the specific response probabilities in detail here. Specifically, the six categories of 21 items for five classes compared between three rounds would result in about 2,000 comparisons. Instead of this, we considered two responses to each item only, namely, the responses "very much like me" and "like me", summing up the probabilities of these responses, and compared them between rounds. In addition, a difference in the rank of class by the item importance was computed. It reflects the logic of interpretation of the classes[5]. Comparisons of response probabilities for the corresponding classes between rounds as well as the difference in ranks are listed in Appendix 2. Although there are some significant differences in absolute values of response probabilities between rounds, there are few differences in class ranks exceeding 1 for the corresponding classes between ESS rounds 4 and 6. There are no differences at all in class ranks between rounds 4 and 5. So, the between-round differences indicate only minor changes in the value profiles of each class and do not affect the general interpretation of each class in its relations to the other classes.

In general, we can see that the number of classes is the same, value profiles of the classes are very alike, and the ranks of class response probabilities are very similar across ESS rounds. These facts are enough for the conclusion to be reached that at least configural invariance of value classes is supported. However, there is another, more parsimonious way to test the configural invariance which is necessary for testing the higher levels of invariance.

[5] For example, there is a class with the highest importance of items belonging to Openness to Change domain, and all of them are expected to have the 1st rank among the other classes.

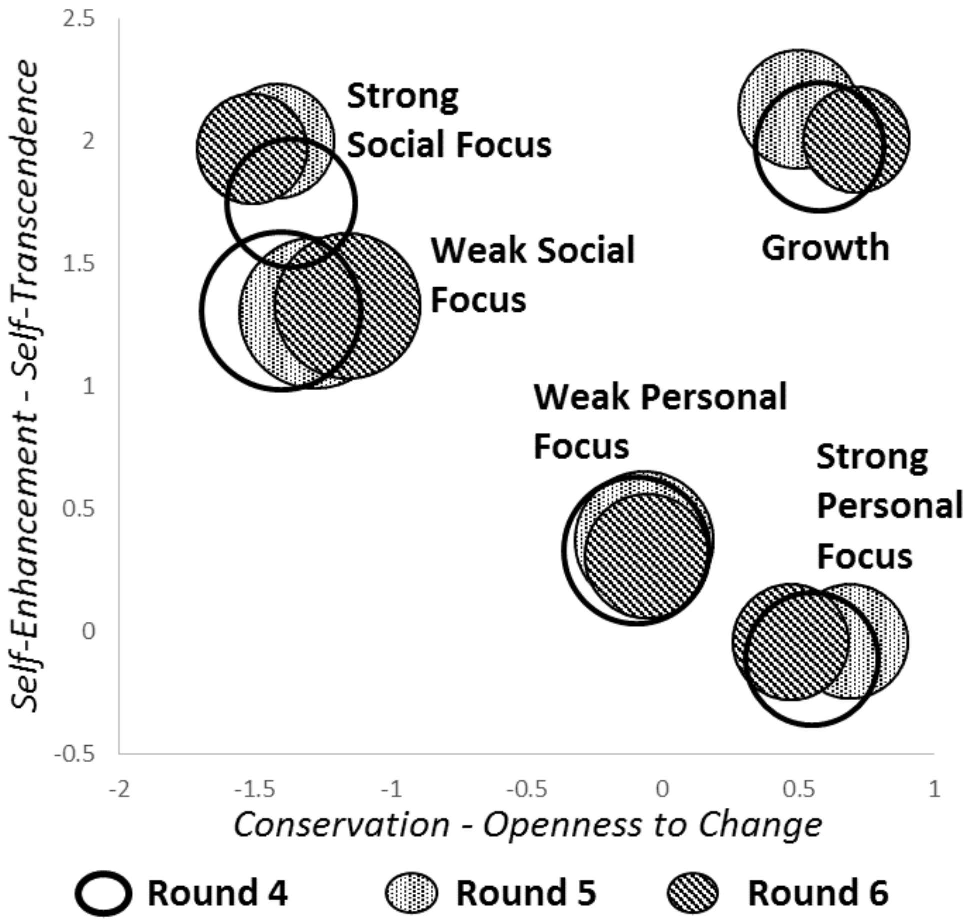

Figure 2. Value classes in the space of the Schwartz higher-order value dimensions. The location is determined by a mean score on both dimensions; the size of the bubbles corresponds to the proportion of the class size in the population.

This is a single group LCA model including ESS round variable as a covariate, using round 4 as a reference group for covariate and dummies for rounds 5 and 6. Since it is a configural model, none of the round effects are constrained. To estimate the difference between the 5th and 6th ESS rounds, the model was recalculated with the 5th round as a reference group. This model (model M1) generally reproduces the three 5-class models described above. The fit statistics are listed in Table 2 (the fit statistics are of minor interest at this point since none of them are standardized). The parameter estimates are presented in Table 3. The effects of the ESS round repeat, in many respects, the differences found in the three independent models (see Appendix 2), demonstrating the same difference of response probabilities across rounds in a more efficient way. The magnitude of the effects indirectly refers to differences between rounds in response probabilities for the corresponding classes: the value of 0.5 corresponds approximately to difference of response probabilities between rounds, which is not higher than 12p.p. (it corresponds to a lower difference when it is applied to the comparison of very low and very high response probabilities, e.g. 0.5 effect converts to a difference of 2p.p. for probabilities about 5%). In Table 3, a negative effect in the "ESS-5 vs. ESS-4" column implies that the distribution of the class response probabilities for the given value item decreased in importance in ESS-5 as compared to ESS-4. A positive effect means that the response probabilities of the current value item increased in importance.

Almost a quarter of regression coefficients are significantly different from zero at the p<.001 level[6], and most of them are indicative of the cross-round non-invariance of the Growth values class (8% of all coefficients), the Weak Social Focus, and the Weak Personal Focus classes (7% and 5%, respectively). The least invariant items are "follow rules", "modesty" and "success". There are no differences between rounds in terms of the degree of invariance since all the rounds have the same number of invariant and non-invariant items.

The magnitude of effects is relatively low – out of 315 there are only 5 effects that are higher than 0.5 in absolute value and 7 effects in the range of 0.4-0.5. All the other effects, i.e. 96%, are less than 0.4, which at a maximum point corresponds to 10p.p. difference in response probabilities. For example, the significant coefficient of -.55, which demonstrates differences between the 4th and 5th round in response probabilities for the item "own decisions" given membership in the Growth values class (see Table 2), translates into a 9% difference in terms of probabilities to respond with "very much like me" or "like me" (see Appendix 2).

Taken altogether we can conclude that configural invariance is fully supported since the cross-round correlations of the class profiles are very high, and the between-class ranks based on response probabilities are very similar in the different ESS rounds as well. In addition to these *relative* measures of similarity in response probabilities, the

[6] The large confidence interval or 99.9% was chosen for two reasons: first, it corresponds to a very large sample size involved in computing standard errors; and second, the magnitude of the significant coefficients at the p<.001 level is not lower than .2, which translates into maximum of 5 p.p. difference in response probabilities between rounds, which traditionally could be considered negligible.

group-as-covariate approach provided us with the coefficients demonstrating the *absolute* differences between class response probabilities across rounds. These coefficients also indicate the high similarity of profiles. Since the results reveal the high level of invariance between ESS rounds and despite the fact that some of the round effects are significant (they may not have a significant impact on the overall model fit), it is reasonable to test the fully invariant model.

Full invariance. This is the same as the model just described with constraints imposed on the effects of the variable, ESS round, on class response probabilities. These effects are set to zero, i.e. the response probabilities for corresponding classes are kept the same across rounds. The full invariance model (M2 in Table 3) is the most restrictive one and constrains all the class probabilities across ESS rounds. Expectedly, the fit statistics for the constrained model are much worse than for the unconstrained models and the likelihood ratio test (LRT) is significant, indicating that the unconstrained model significantly better describes the data than the fully constrained one. Since the configural model has definitely demonstrated a similarity of value class structure across rounds, it is reasonable to turn to the model with fewer equality constraints across rounds and test it against the unrestricted one.

Partial invariance. The partial invariance model is an intermediate one between the configural and fully invariant model. We fixed the effects of ESS round to zero for the most invariant items detected in the configural model and kept free the effects for the least invariant ones, i.e. for those items which were significantly different from zero in the configural model (see Table 3). The LRT between the partial invariance and the configural invariance model is significant. It formally rejects the hypothesis about the partial invariance of value classes between ESS rounds. However, as we noted above, the LRT is sensitive to a large sample size, making any change in the model significant. In the present study, the sample size is huge (well over 150,000), therefore, the results of the LRT could be biased and the other model fit statistics would need to be examined. BIC and AIC increased only slightly: BIC increased by 0.01p.p. as compared to the configural model (it was a 0.03p.p. increase for the full invariance model); AIC increased by 0.02p.p. (it was a 0.33p.p. increase for the full invariance model). The results for BIC and

AIC demonstrate that the differences in model fit between the configural and partially invariant models are very small.

Based on the comparison of the fit statistics of the three models considered here (M1, M2 and M3), we can stop at this point and select the partial invariance model as the final one. As these analyses have demonstrated, the general meaning of five value classes expressed in class response probabilities is very similar between the three ESS rounds. This is evidence of a stability of value classes across samples, which implies that the typology developed in our earlier work is feasible. Although the measurement of classes is not fully invariant across rounds and the degree to which shares of classes can be directly compared across rounds is open to discussion, the meaning of the classes is stable.

Table 2. Fit statistics for LCA with ESS round as a predictor (ESS round 4 is the reference group)

Model	Npar	AIC	BIC	-2LL	LL Ratio Test significance
M1. Configural invariance model. Class response probabilities can differ (effects of ESS round number on response probabilities are estimated freely)	752	8753581	8758665	8752077	[baseline]
M2. Full invariance model. Response probabilities for all the classes are constrained to be equal across three rounds (effects of ESS round variable on response probabilities are fixed equal to zero)	542	8782091	8763596	8757124	0.000
M3. Partial invariance model. Some response probabilities are constrained to be equal across ESS rounds (the ones that are not significantly different from zero in Table 3)	593	8755408	8759417	8754222	0.000

Notes: *AIC – Akaike information criterion, BIC – Sample adjusted Bayesian information criterion, Entropy – a measure of uncertainty of classification.*

Table 3. The effects of ESS round on the class response probabilities for five value classes

Value class	Growth			Strong Social Focus			Weak Social Focus			Weak Personal Focus			Strong Personal Focus		
Item\ESS round	ESS-5 vs ESS-4	ESS-6 vs ESS-4	ESS-6 vs ESS-5	ESS-5 vs ESS-4	ESS-6 vs ESS-4	ESS-6 vs ESS-5	ESS-5 vs ESS-4	ESS-6 vs ESS-4	ESS-6 vs ESS-5	ESS-5 vs ESS-4	ESS-6 vs ESS-4	ESS-6 vs ESS-5	ESS-5 vs ESS-4	ESS-6 vs ESS-4	ESS-6 vs ESS-5
Openness to Change															
Creative	-0.04	**-0.27***	-0.23	0.17	0.09	-0.08	-0.1	**-0.3***	**-0.2***	0.21	0.02	-0.19	0.03	0.09	0.06
Own decisions	-0.04	**-0.55***	**-0.51***	0.12	-0.23	**-0.35***	-0.13	**-0.46***	**-0.33***	**0.27***	-0.09	**-0.37***	-0.1	-0.06	0.04
New things	-0.03	**-0.36***	**-0.34***	0.06	0.2	0.15	-0.14	**-0.26***	-0.12	0.17	0.02	-0.15	-0.004	0.09	0.1
Adventures	-0.04	**-0.43***	**-0.39***	0.22	0.14	-0.08	**-0.27***	**-0.44***	-0.17	0.17	0.05	-0.12	0.01	0.01	-0.001
Good time	-0.06	0.001	0.06	0.06	0.24	0.18	-0.14	-0.22	-0.09	**0.26***	0.05	-0.22	0.02	0.01	-0.01
Fun	-0.05	**-0.3***	-0.24	-0.04	-0.05	-0.02	**-0.23***	**-0.36***	-0.13	0.04	-0.17	-0.21	0.09	0.06	-0.03
Conservation															
Secure surroundings	-0.16	-0.25	-0.1	0.11	-0.06	-0.18	-0.08	-0.07	0.01	0.15	-0.06	-0.21	-0.22	-0.2	0.03
Security by government	**-0.25***	**-0.57***	**-0.32***	0.13	-0.01	-0.14	-0.18	-0.05	0.13	0.10	-0.01	-0.11	**-0.3***	-0.17	0.14
Rules	-0.1	0.01	0.11	**0.42***	0.1	**-0.32***	**0.28***	0.20	-0.08	**0.46***	0.06	**-0.4***	0.24	-0.01	-0.24
Behave properly	-0.14	0.04	0.18	0.19	-0.01	-0.2	0.01	-0.24	**-0.24***	**0.28***	0.01	-0.27	-0.01	-0.3	-0.29
Tradition	-0.01	-0.13	-0.12	0.08	-0.03	-0.1	0.08	0.15	0.07	0.21	0.03	-0.19	0.06	-0.19	-0.25
Modesty	**-0.26***	-0.23	0.03	0.04	-0.19	-0.23	0.1	**-0.39***	**-0.49***	**0.26***	-0.07	**-0.33***	0.01	**-0.35***	**-0.36***
Self-Transcendence															
Help people around	-0.16	**-0.50***	**-0.35***	-0.06	-0.24	-0.17	-0.08	**-0.36***	**-0.28***	0.20	-0.04	-0.24	-0.18	**-0.35***	-0.17
Friends	-0.13	**-0.33***	-0.2	-0.19	-0.15	0.03	**-0.22***	**-0.39***	-0.17	0.08	-0.13	-0.2	-0.26	**-0.42***	-0.16
Understanding	0.02	**-0.33***	**-0.35***	0.10	0.04	-0.06	0.03	-0.21	**-0.24***	**0.37***	0.06	**-0.31***	0.11	-0.05	-0.17
Nature	0.001	**-0.32***	**-0.33***	0.09	-0.21	**-0.3***	-0.09	**-0.26***	-0.17	0.21	0.07	-0.14	-0.04	-0.1	-0.06
Equality	**-0.35***	**-0.39***	-0.04	-0.07	0.05	0.13	-0.11	**-0.28***	-0.16	0.12	-0.19	**-0.31***	**-0.31***	-0.25	0.06
Self-Enhancement															
Abilities	0.16	-0.23	**-0.39***	0.18	0.12	-0.06	-0.03	**-0.22***	-0.19	**0.30***	-0.02	**-0.31***	-0.13	-0.08	0.04
Success	0.22	**-0.39***	**-0.61***	**0.26***	0.05	-0.21	-0.02	**-0.23***	-0.21	**0.35***	0.01	**-0.34***	-0.08	-0.11	-0.04
Wealth	0.05	-0.07	-0.12	**0.26***	**0.4***	0.14	-0.17	0.06	0.23	**0.3***	0.04	-0.27	0.03	0.18	0.16
Respect	0.05	-0.15	-0.19	**0.28***	0.21	-0.07	-0.04	0.01	0.04	0.26	-0.03	**-0.29***	-0.004	-0.28	-0.28

Note: 1) The group noted after “vs.” is a reference group; 2) * - effect is significant at *p*<.001 level.

3.3. The invariant value class solution

In the following two sections we describe in more detail the LCA model M3 that is partially invariant across ESS rounds and that was accepted as the most appropriate. The significant effects of the ESS round provide an indication of the cross-round *differences* allowed in that model between similar classes (Table 3). And now we describe the value profiles of these classes averaged across rounds.

The resulting classes are described in the LCA output in terms of thresholds, which are logged ratios of the respondent's probability to give a certain answer compared to the probability of choosing the last option in the set of responses. Thresholds were converted into probabilities for a respondent to provide a certain answer given this respondent's class membership.

The sums of probabilities of the respondents' answers "very much like me" and "like me" for each of the 21 value items conditioned by the class membership are listed in Figure 3. For instance, given that the respondent is a member of class 1 (Growth), there is an 85% probability that this respondent would claim his or her similarity to a person who believes that people should be treated equally and should have equal opportunities.

Since there are too many differences between classes in terms of items, and these differences are very consistent within value categories, we demonstrate differences between classes in terms of value categories. For example, class 1, as compared to the other value classes, has the lowest probabilities for five of the six value items measuring Conservation. One exception (modesty item), like the other exceptions, does not change the general interpretation. Overall, exceptions are found for about 6% of comparisons.

The members of class 1 (15-16% of the population of the three ESS rounds) are characterized by two minima; they display the weakest commitment to both Conservation and Self-Enhancement values. They also indicate relatively strong commitment (i.e. the second highest probabilities) to values which belong to the categories of Openness and Self-Transcendence. In Schwartz's terms, the members of this class prefer Growth values over Self-Protection. In short, this class may be labeled "Growth".

The members of class 2 (17-19% of the population of the three ESS rounds) are characterized by two maxima; they indicate the strongest commitment to both Conservation and Self-Transcendence values and moderate commitment to the values of both Openness and Self-Enhancement. Commitment to Conservation and Self-Transcendence values emphasize the strong social focus of the people who share them, and these values are clearly preferred by these class members over the personally focused Openness and Self-Enhancement values. The concise label for this class would be "Strong Social Focus".

Class 3 (the largest one, consisting of 26-29% of the population of the three ESS rounds) is similar to class 2 in its value profile although with a slightly lower importance of all value categories for its members. Its members indicate the weakest commitment to Openness and relatively strong commitment (i.e. the second highest probabilities) to Conservation values. Moreover, they have contrasting levels of commitment to Self-Transcendence and Self-Enhancement values. Their commitment to Self-Transcendence values is relatively strong (i.e. the second to highest probabilities), and their commitment to Self-Enhancement is relatively weak (i.e. the second to lowest probabilities). As this class is quite similar to the Strong Social Focus class while differing only with a degree of preference for socially focused values, we designate it the "Weak Social Focus" class.

The members of class 4 (22-24% of the population of the three ESS rounds) are characterized by their lower (i.e. second or third lowest probabilities) commitment to Openness along with their lower (i.e. second to the lowest scores) commitment to Conservation values. They also indicate the weakest commitment to Self-Transcendence and a relatively strong (i.e. second highest probabilities) commitment to Self-Enhancement values. This class is labeled "Weak Personal Focus".

The members of class 5 (15-18% of the population of the three ESS rounds) are characterized by two maxima; they indicate the strongest commitment to both Openness and Self-Enhancement values, and they are moderately committed to Conservation and display only relatively low commitment (i.e., second to lowest level) to Self-Transcendence values. In contrast to the strong socially focused value class, this class can be labeled "Strong Personal Focus".

Differences in the class sizes between rounds reported above were calculated using the effects of the variable round on class and are not larger than 3p.p. These differences are mostly due to differences in the samples of countries, which varied between ESS rounds. The other reason for these differences is non-invariance of some of the items.

The specific class response probabilities in the partial invariance model (M3) are very similar to the ones that have been found and described using the data from the 4th ESS round only [Magun et al., in review]. The findings from the present analysis are similar enough to the ones found in the aforementioned study to allow us to keep exactly the same interpretations of the classes. Moreover, the main conclusion of the current study is confirmed: the value typology of Europeans is robust and stable across ESS rounds. Taking together the support of partial invariance and the rejection of full invariance, it may be further hypothesized that the lack of full invariance originates from the minor problems of value measurement and not from the instability of the value typology itself, as a construct.

Figure 3. The estimated probabilities of the respondents' answers "very much like me" and "like me" conditioned by the class membership (LCA model M3 partially invariant across ESS rounds; the class value profiles are averaged across three ESS rounds).

3.4. Testing the stability of country effects on value classes across ESS rounds

To test the hypotheses H3 and H4 concerning external variables, namely, country and country's economic development and the cross-round stability of relations of value class probability, we focus on the differences of country effects on respondent class membership between rounds, using both the ESS round and country dummies as well as their interactions as predictors of the LCA class membership.

We employed a 3-step approach initially proposed by Vermunt [2010] and described in Asparouhov and Muthen [2013]. This technique allowed us to account for the uncertainty of class membership and to avoid disturbing the classification procedure by adding covariates into an LCA model.

1. The first step is described above and involves the partially invariant LCA model for three ESS rounds.
2. In the second step, the estimated class membership is assigned with uncertainty rates. These are based on the average probability of class members to be a member of this class and the sizes of the classes (for details, see Asparouhov and Muthen, 2013).
3. In the third step, the covariate (the interaction between country of residence and ESS round, referring to the changeability of country effects between rounds) was included in the model as a predictor, so that the LCA model is fixed, uncertainty of class membership is accounted for, and the predictors of interest are added to the model. In this step, a multinomial logistic regression is run. To achieve convergence of the regression model, some of its parameters were fixed to -15, because -15 on the logarithmic scale translates into a value which is very close to zero.

The regression coefficients of the resulting model are the interactions between the ESS round and country, reflecting the differences of country class membership probabilities between rounds[7]. These regression coefficients are hard to interpret, since each of them have three reference groups: ESS round 4, one of the countries, and one of

[7] The probability of membership is very close to descriptive shares of classes in each country. However, the membership is predicted in LCA with some degree of uncertainty which is not accounted for when working with membership instead of membership probability. That is why we use the term "membership probability".

the classes. For the interpretation, country and round coefficients were converted to probabilities of class membership given residence in a certain country and round of ESS. These calculated probabilities for each of the three ESS rounds are listed in Appendix 3. They indicate that, in each round, all of the 32 countries are internally diverse in their value class composition, and most of the countries have a non-zero probability of having members of all five value classes in its population. Residents of Kosovo, Romania, Russia, Slovakia, Turkey and Ukraine have a close to zero probability of being a member of the Growth class, but their residents have non-zero chances to be represented in the four other classes. Just like in the ESS round 4, the most contrasting differences between countries are found in the probabilities of membership in the Growth class. Its membership in most Nordic (varying from 24-30% for the different rounds) and Western European (31-35%) countries is remarkably higher than in Mediterranean (10-15%) and Post-Communist countries (6-7%), and most of these differences are statistically significant at $p<0.01$. All the other classes are better represented in the Mediterranean and Post-Communist than in the Nordic and Western European countries (although the country differences of membership probability are not very salient for these classes). For example, average membership probability for the Weak Social Focus class for Post-Communist and Mediterranean countries for all three rounds is 29-33% and 27-30%, respectively; and it is lower in the Nordic and Western European countries, with 22-24% and 24-30% probability of membership, respectively. Statistically significant negative correlations between country membership probability for the Growth class and each of the four other classes confirm the gap between the Growth class and all the others. (All the correlations between country membership probabilities for the other four classes are insignificant.)

The membership probabilities for the individual countries are rather robust between rounds. The average difference of country probabilities for corresponding classes between ESS rounds is 4%. In most cases, the effects of round on country probability are small and translate into variations of not more than 10%. The difference of class country probabilities exceeds 10% only in 8 out of 160 comparisons between rounds. The correlations between country membership probabilities in different rounds are stable, ranging from 0.70 to 0.98 with an average of 0.87; all of them are statistically significant (p<0.001). Out of 22 countries that participated in all three ESS rounds, Growth class probabilities slightly decreased in 9 countries, increased in two countries and the other countries demonstrated an inconsistent tendency (increase in 5th but decrease in 6th ESS rounds, or decrease in 5th and increase in 6th ESS round). Coming back to the alternative hypothesis, considering influence of the economic crisis on country value class probabilities, we can conclude that a consistent tendency of decreasing Growth values class has not been found.

The correlations of calculated country probabilities of the value classes with gross national income (GNI) per capita indicating country economic development are remarkably stable across all three rounds (Table 4). They indicate that in more economically advanced countries, probability of membership in the Growth class is much higher (the magnitudes of these coefficients are very high, 0.80-0.90) and that in economically less advanced countries, the probabilities of membership are higher in all the classes except the Growth one (the coefficients of the growth class are not so high by their magnitude and not so consistent across rounds).

Table 4. Correlations between calculated country value class probabilities and country GNI per capita (in the years 2008, 2010 and 2012 assessed in ESS rounds 4, 5 and 6, respectively)

Class label	Round 4	Round 5	Round 6
Growth	0.91*	0.86*	0.81*
Strong Social Focus	-0.48*	-0.43*	-0.51*
Weak Social Focus	-0.39*	-0.49*	-0.46*
Weak Personal Focus	-0.38*	-0.35	-0.27
Strong Personal Focus	-0.46*	-0.48*	-0.27
N	28	26	23

** Significantly different from zero at p< 0.05 or stricter level. Norway is excluded as an outlier.*

4. Discussion and Conclusions

The analyses presented in this paper have demonstrated a high level of robustness for the value classification of Europeans, initially described by Magun, Rudnev and Schmidt [in review], across three ESS rounds.

Three completely independent LCA models based on the data from the 4^{th}, 5^{th} and 6^{th} ESS rounds detected the same number of value classes initially reported for one ESS round (i.e. ESS round 4). It has been demonstrated that the value classes found in each of the three ESS rounds under consideration have very high cross-round correlations in their corresponding value class profiles, very similar between-class item ranks, and similar response probabilities that differ by no more than 10%, with a few exceptions. The substantial similarity of classes between the ESS rounds is clearly demonstrated by very similar average scores of the corresponding classes on two higher-order value dimensions. Based on the results reported here, the exact same substantial interpretation of the typologies found previously for one ESS round can be extended to encompass all three ESS rounds. Moreover, an exploratory analysis supported configural or construct invariance, pointing out the feasibility of the proposed value typology.

External validity of the typology was supported by relating value class membership with country of residence and country level of economic development. The evidence regarding the validity of the typologies based on the data from the 5^{th} and 6^{th} ESS rounds appeared the same as has been detected for the 4^{th} ESS round data. The correlations between country class probabilities, with the level of economic development (measured by GNI per capita), are relatively stable across three consecutive ESS rounds. Probability of membership in the Growth class is much higher in more economically advanced countries, and this is true for all three ESS rounds. The country probabilities of membership in the other four classes are higher for the less developed countries.

In most countries, the effects of round on country class membership are small and translate into variations of not more than 10%. It is noteworthy, that in spite of the economic crisis, a consistent tendency in value change was not found. Values are claimed to change only in really harsh conditions, such as psychological trauma, war, migration, etc., and probably this was not the case with the mentioned crisis. The other reason might be that values require a longer time to change, and probably there will be a postponed effect of the crisis. However at this point, such an effect on value classes was not detected.

As we mentioned above, the sample of countries analyzed here is not representative of Europe in its entirety and, strictly speaking, is limited to a certain set of countries that participated in each ESS round. However, the latent class solution is surprisingly stable given that the set of countries differs from round to round. It also supports the robustness of the value class solution and confirms the stability of the relationship between the value classes and different kinds of countries.

Confirmatory analysis was performed using single group LCA models that use ESS round as a covariate. These results supported configural invariance as well. Some deviations from strict full invariance were demonstrated as well. BIC and AIC were used to detect whether partial invariance holds, and inspection of these values confirmed that this was indeed the case. At the same time, however, strict statistical criteria for accepting or rejecting the hypotheses about a certain level of invariance in the context of latent classes have not yet been developed. Due to a lack of empirical studies dealing with latent class invariance, we cannot conclude if there is enough similarity between response probabilities to be entirely certain that the class membership is unbiased across the three ESS rounds.

At this point, it is useful to differentiate the concepts of measurement invariance and construct invariance. Construct invariance is referred to as configural measurement invariance; it is a feature of typology validity, implying that the classification is feasible and reflects the social reality as being different from an accidental product of the data analysis. The measurement invariance assesses the precision with which the constructs are measured in different groups. Thus, the construct invariance is a matter of theoretical constructs validity, whereas measurement invariance refers to a degree of between-group validity of measurement. Kuha and Moustaki [2013] share the view that "even when all the true measurement probabilities are such that each latent class would be given the same qualitative interpretation in every group, the class probabilities estimated under equivalence can still be substantially biased" (p. 21). Therefore, although a degree of measurement invariance of the value classes is still an open issue, the substantial similarity of the value classes across ESS rounds is notable and provides evidence of construct invariance, that is, the existence of the stable and feasible 5-class structure of the European population, which also means that this structure is not simply an accidental product of the 4th round ESS data.

Although the results of this study do not allow strict conclusions to be made about the specific sources of non-invariance of some items, we suggest that given the feasibility of classification, full invariance of value classes has not emerged and this, first of all, is due to the measurement issues and differences in samples across the three ESS rounds.

Overall, we can conclude that there are five value types in the European population: a Growth class, emphasizing Openness to Change and Self-Transcendence values; in opposition to this first value type, classes along the Social-Personal focus dimension, consisting of two Social Focus classes, emphasizing a combination of Conservation and Self-Transcendence values (at the expense of Openness to Change and Self-Enhancement values); and conversely, the two Personal Focus classes emphasizing the combination of Openness to Change and Self-Enhancement (at the expense of Conservation and Self-Transcendence values).

These results clearly imply that the value classification of Europeans is not an ad hoc study and, although its measurement varies in different ESS rounds, it reflects a stable and feasible value-based structure of the European population that can be used in future studies.

References

Asparouhov, T. and Muthen, B. 2013. “Auxiliary variables in mixture modeling: 3-step approaches using Mplus”. Web note 15: Version 7.

Billiet, J. B. and McClendon, J. M., 2000. “Modeling acquiescence in measurement models for two balanced sets of items. Structural Equation Modeling”, vol. 7, pp. 608-628.

Billiet, J., Meuleman, B., & De Witte, H. 2014. “The relationship between ethnic threat and economic insecurity in times of economic crisis: Analysis of European Social Survey data”. Migration Studies, vol. 2, pp. 135-161.

Davidov, E., J. Cieciuch, B. Meuleman, P. Schmidt and J. Billiet. 2014. “Measurement equivalence in cross-national research”. Annual Review of Sociology, vol. 40.

Eagly, A. H., and Chaiken, S. 1993. The psychology of attitudes. Harcourt Brace Jovanovich College Publishers.

Eid, M., Langeheine, R.and Diener, E. 2003. "Comparing typological structures across cultures by multigroup latent class analysis. A primer". Journal of Cross-Cultural Psychology, vol. 34, no. 2, pp. 195-210.

Finch, W. H. and Bronk, K. C. 2011. "Conducting confirmatory latent class analysis using MPLUS". Structural Equation Modeling, vol. 18, pp. 132-151.

Hagenaars J. A. and Halman, L. C. 1989. "Searching for ideal types: The potentials of latent class analysis". European Sociological Review, vol. 5, pp. 81-96.

Hagenaars, J. A. and McCutcheon, A. L. (Eds.) 2002. Applied latent class analysis. Cambridge: Cambridge University Press.

Hancock, G. R. and Samuelsen, K. M. (Eds.) 2008. Advances in latent variable mixture models. IAP.

Jowell, R., Roberts, C., Fitzgerald, R., and Eva, G. (Eds.) 2007. Measuring attitudes cross-nationally: Lessons from the European Social Survey. London: Sage.

Kankaras, M., Moors, G. and Vermunt, J. 2011. "Testing for measurement invariance with latent class analysis", in E. Davidov, J. Billet and P. Schmidt (Eds.), Cross-cultural analysis: Methods and applications. London: Taylor and Francis, pp. 359-384.

Kelloway, E. K. 1995. "Structural equation modelling in perspective". Journal of Organizational Behavior, vol. 16, no. 3, pp. 215-224.

Klages, H. and Gensicke, T. 2005. „Wertewandel und die Big Five Dimensionen [Changing values and the Big Five dimensions]", in S. Schumann (Ed.), Persönlichkeit. Eine vergessene Grösse der empirischen Sozialforschung, Wiesbaden, Germany: VS Verlag, pp. 279-299.

Kuha, J. and Moustaki, I. 2013. "Non-equivalence of measurement in latent variable modelling of multigroup data: A sensitivity analysis", viewed September 27, 2013, <http://ssrn.com/abstract=2332071>

Lazarsfeld, P. F. and Henry, N. W. 1968. Latent structure analysis. Boston: Houghton Mifflin.

Lee, J., Soutar, G. N., Daly, T. M. and Louviere, J. J. 2011. “Schwartz values clusters in the United States and China”. Journal of Cross-Cultural Psychology, vol. 42, pp. 234-252.

Lo, Y., Mendell, N. R. and Rubin, D. B. 2001. „Testing the number of components in a normal mixture”. Biometrika, vol. 88, pp. 767-778.

Lubke, G. H. and Muthén, B. O. 2004. “Applying multigroup confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons”. Structural Equation Modeling, vol. 11, pp. 514-534.

Magun V., Rudnev M. and Schmidt P. In review. Within- and between-country value diversity in Europe: A person-centered approach.

McCutcheon, A. L. 1987. Latent class analysis. Sage University Paper Series on Quantitative Applications in the Social Sciences. Series no. 07-064. Newbury Park, CA: Sage.

Muthén, B. and Asparouhov, T. 2013. “BSEM measurement invariance analysis.” Webnote number 17. <http://statmodel.com/examples/webnotes/webnote17.pdf>

Muthén, L. K. and Muthén, B. O. 2012. Mplus user’s guide. Seventh edition. Los Angeles, CA: Muthén & Muthén.

Schwartz, S. H. 1992. “Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries”. Advances in Experimental Social Psychology, vol. 25, pp. 1-65.

Schwartz, S. H. 1994. Are there universal aspects in the content and structure of values? Journal of Social Issues, vol. 50, pp. 19-45.

Schwartz, S. H. 2005. “Basic human values: Their content and structure across countries”, in A. Tamayo, and J. B. Porto (Eds.), Valores e Comportamento nas Organizac ¸Boes [Values and behavior in organizations] Petrópolis, Brazil: Vozes, pp. 21–55.

Schwartz, S. H. 2007. “Value orientations: Measurement, antecedents and consequences across nations”, in R. Jowell, C. Roberts, R. Fitzgerald and G. Eva (Eds.), Measuring attitudes cross-nationally: Lessons from the European Social Survey. London: Sage Publications, pp. 169-204.

Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lцnnqvist, J.-E., Demirutku, K., Dirilen-Gumus, O. and Konty, M. 2012. „Refining the theory of basic individual values”. Journal of Personality and Social Psychology. Advance online publication. DOI: 10.1037/a0029393

Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M. and Owens, V. 2001. “Extending the cross-cultural validity of the theory of basic human values with a different method of measurement”. Journal of Cross-Cultural Psychology, vol. 32, pp. 519-542.

Schwartz, S. H., Verkasalo, M., Antonovsky, A., and Sagiv, L. 1997. „Value priorities and social desirability: Much substance, some style”. British Journal of Social Psychology, vol. 36, pp. 3-18.

Siegers P. 2011. “A multiple group latent class analysis of religious orientations in Europe”, in E. Davidov, J. Billet, and P. Schmidt (Eds.), Cross-cultural analysis: Methods and applications, London: Taylor and Francis, pp. 359-384

Steenkamp, J.-B. E.M. and Baumgartner, H. 1998. “Assessing measurement invariance in cross-national consumer research”. Journal of Consumer Research, vol. 25, pp. 78-90.

Van Herk, H., Poortinga, Y. H. and Verhallen, T. M. M. 2004. “Response styles in rating scales: Evidence of method bias in data from six EU countries”. Journal of Cross-Cultural Psychology, vol. 35, pp. 346-360.

Vermunt, J. K. 2010. “Latent class modeling with covariates: Two improved three-step approaches”. Political Analysis, vol. 18, pp. 450-469.

Weighting European Social Survey Data. Viewed 11 March 2014 <http://www.europeansocialsurvey.org/docs/methodology/ESS_weighting_data.pdf>

Appendix 1. ESS Portrait Values Questionnaire item wording, their labels and the values they are supposed to measure

Item label	Item label	Item wording
OPENNESS TO CHANGE		
Self-Direction	(creative)	Thinking up new ideas and being creative is important to him. He likes to do things in his own original way.
	(own decisions)	It is important to him to make his own decisions about what he does. He likes to be free and not depend on others.
Stimulation	(new things)	He likes surprises and is always looking for new things to do. He thinks it is important to do lots of different things in life.
	(adventures)	He looks for adventures and likes to take risks. He wants to have an exciting life.
Hedonism	(good time)	Having a good time is important to him. He likes to “spoil” himself.
	(fun)	He seeks every chance he can to have fun. It is important to him to do things that give him pleasure.
CONSERVATION		
Security	(secure surroundings)	It is important to him to live in secure surroundings. He avoids anything that might endanger his safety.
	(safety government)	It is important to him that the government ensures his safety against all threats. He wants the state to be strong so it can defend its citizens.
Conformity	(follow rules)	He believes that people should do what they're told. He thinks people should follow rules at all times, even when no one is watching.
	(behave properly)	It is important to him always to behave properly. He wants to avoid doing anything people would say is wrong.
Tradition	(traditions)	Tradition is important to him. He tries to follow the customs handed down by his religion or his family.
	(modest)	It is important to him to be humble and modest. He tries not to draw attention to himself.
SELF-TRANSCENDENCE		

Benevolence	(help people)	It's very important to him to help the people around him. He wants to care for their well-being.
	(friends)	It is important to him to be loyal to his friends. He wants to devote himself to people close to him.
Universalism	(understand)	It is important to him to listen to people who are different from him. Even when he disagrees with them, he still wants to understand them.
	(nature)	He strongly believes that people should care for nature. Looking after the environment is important to him.
	(equally)	He thinks it is important that every person in the world should be treated equally. He believes everyone should have equal opportunities in life.
SELF-ENHANCEMENT		
Achievement	(abilities)	It's important to him to show his abilities. He wants people to admire what he does.
	(success)	Being very successful is important to him. He hopes people will recognize his achievements.
Power	(rich)	It is important to him to be rich. He wants to have a lot of money and expensive things.
	(respect)	It is important to him to get respect from others. He wants people to do what he says.

Appendix 2. Differences in classifications of European population into five value classes by three independent LCA models for three ESS rounds

Item	Growth s		Strong Social Focus		Weak Social Focus		Weak Personal Focus		Strong Personal Focus	
	ESS-5	ESS-6	ESS-5	ESS-6	ESS-5	ESS-6	ESS-5	ESS-6	ESS-5	ESS-6
Openness to Change										
Creative		7% (+0.5)	-7%	-5%		6%				(-0.5)
Own decisions		9% (+0.5)				11%			5%	(-0.5)
New things		11% (+1)	-6%	-9% (-1)		5%		(+1)		(-1)
Adventures		6%								
Good time				-8% (-0.5)		5%		(+0.5)		
Fun		9%				6%				
Conservation										
Secure surroundings										6%
Security by government	5%	12% (+0.5)						(-0.5)	6%	
Rules			-9%		-5%	-6%	-9%		-9%	
Behave properly										8%
Tradition		5%				-6%	-5%			
Modesty	7%	5%				6%				6%
Self-Transcendence										
Help people around		9%				6%			6%	9%
Friends		5% (+0.5)			5%	5%			7%	8% (-0.5)
Understanding		8%				5%	-7%			
Nature		8%						-5%		
Equality		6% (+0.5)							8%	6% (-0.5)
Self-Enhancement										
Abilities			-5%	-5%		5%			5%	
Success		7%	-9%	-6%						
Wealth			-5%	-7%						
Respect			-8%	-6%			-5%			5%

Note. The differences are described in the table through deviations of the response probabilities for answers "very much like me" and "like me" in the 5th and 6th ESS rounds from the 4th ESS round; the cross-round deviations of the class rank of the item importance are noted in parentheses. Empty cells refer to the lack of differences given a 95% confidence interval.

Appendix 3. Probabilities of value class membership given the country of respondent's residence for the three ESS rounds

Class	Growth			Strong Social Focus			Weak Social Focus			Weak Person Focus			Strong Person Focus		
ESS round	**4th**	**5th**	**6th**	**4th**	**5th**	**6th**	**4th**	**5th**	**6th**	**4th**	**5th**	**6th**	**4th**	**5th**	**6th**
Belgium	26%	26%	23%	11%	9%	8%	26%	28%	30%	26%	24%	28%	11%	12%	11%
Bulgaria	5%	3%	4%	28%	24%	24%	29%	33%	35%	19%	22%	19%	19%	19%	19%
Croatia	9%	6%		23%	23%		30%	37%		23%	19%		16%	15%	
Cyprus	12%	9%	11%	16%	25%	29%	35%	36%	29%	14%	12%	10%	24%	18%	22%
Czech Rep.	7%	6%	5%	14%	11%	11%	26%	27%	28%	29%	32%	33%	24%	24%	22%
Denmark	35%	37%	32%	11%	13%	11%	19%	21%	20%	16%	12%	17%	18%	17%	20%
Estonia	17%	18%	16%	16%	10%	13%	28%	36%	36%	21%	24%	22%	18%	12%	13%
Finland	29%	29%	33%	15%	12%	14%	24%	25%	27%	21%	22%	16%	11%	12%	10%
France	38%	41%		28%	25%		14%	14%		12%	12%		8%	9%	
Germany	32%	30%	31%	12%	14%	12%	28%	31%	31%	14%	13%	12%	14%	13%	14%
Greece	8%	10%		17%	21%		26%	31%		24%	18%		25%	19%	
Hungary	10%	12%	8%	21%	21%	15%	24%	25%	24%	22%	19%	29%	24%	24%	25%
Iceland			46%			8%			21%			10%			16%
Ireland	21%	15%	15%	18%	17%	14%	30%	21%	33%	17%	28%	24%	15%	18%	15%
Israel	6%	9%	9%	23%	21%	18%	18%	19%	22%	28%	28%	20%	25%	23%	31%
Kosovo			1%			31%			30%			15%			23%
Latvia	6%			13%			26%			21%			34%		
Lithuania		5%			17%			29%			26%			22%	
Netherlands	26%	28%	24%	7%	8%	8%	22%	20%	22%	32%	32%	32%	13%	12%	13%
Norway	22%	30%	26%	9%	7%	8%	29%	33%	33%	27%	17%	18%	14%	13%	15%
Poland	9%	6%	6%	13%	14%	16%	42%	44%	46%	24%	22%	19%	13%	15%	13%
Portugal	10%	10%	7%	10%	8%	7%	26%	24%	29%	43%	50%	46%	10%	8%	11%
Romania	2%			16%			21%			41%			20%		
Russia	3%	3%	2%	21%	17%	20%	30%	29%	24%	25%	29%	28%	21%	23%	26%
Slovakia	2%	3%	2%	19%	17%	16%	39%	40%	39%	29%	25%	27%	11%	16%	16%
Slovenia	16%	12%	12%	15%	15%	19%	27%	28%	38%	28%	31%	17%	15%	14%	14%
Spain	21%	28%	25%	22%	17%	26%	35%	33%	34%	15%	13%	8%	8%	8%	7%
Sweden	38%	46%	38%	8%	10%	13%	17%	18%	20%	25%	14%	16%	13%	12%	13%
Switzerland	37%	34%	32%	13%	13%	11%	21%	26%	26%	14%	14%	14%	15%	13%	17%
Turkey	3%			16%			24%			42%			16%		
Ukraine	2%	3%		31%	26%		27%	25%		17%	22%		23%	24%	
UK	27%	23%	22%	13%	14%	14%	28%	32%	36%	17%	16%	15%	14%	15%	14%

Note. An empty cell indicates that the country did not participate in the corresponding ESS round.

Correspondence regarding this paper should be addressed to:

Maksim Rudnev
National Research University Higher School of Economics
20 Myasnitskaya St.
101000 Moscow
Russia
E-mail: mrudnev@hse.ru