

Measuring political protest in Western Europe: assessing cross-national equivalence

MARIO QUARANTA *

Istituto Italiano di Scienze Umane, Palazzo Strozzi, Piazza Strozzi, Firenze, Italy

This article tests the cross-national equivalence of the political protest scale, as developed by Barnes and Kaase, in 20 Western European countries using a battery of items included in the fourth wave of the European Values Study. The scale measuring the concept of political protest is widely used, but no evidence of cross-country equivalence has yet been provided in the literature. The article illustrates the concept of political protest, the relationship between concept formation, operationalization, and measurement equivalence, and the possible consequences of a lack of equivalence. It is argued that comparative research may be threatened by a lack of measurement equivalence. The spread of international surveys eases comparative designs, but at the same time enlarges the chances that we compare what is not actually comparable. The article then outlines an empirical strategy to assess the political protest scale's measurement equivalence. To assess cross-country equivalence, Mokken Scale Analysis, a nonparametric scaling method within the family of Item Response Theory models, is used. This has been shown to work better than Confirmatory Factor Analysis when dealing with dichotomous and polytomous items forming ordinal scales. The results show that the cross-country equivalence of the political protest scale depends on the type of measure the scholar wishes to build and use.

Keywords: political protest; unconventional participation; measurement equivalence; Western Europe; Mokken Scale Analysis

Introduction

Citizens often use their participatory rights to influence governments and private subjects' decisions without necessarily passing through institutional channels. So-called 'protest politics', or 'unconventional' forms of political participation, are slowly replacing more 'conventional' forms (Norris, 2002). In fact, it is argued that contemporary democracies are 'social movement societies' in which protest activities are common and participants are not easily identifiable as they are ordinary citizens (Meyer and Tarrow, 1998). But how do we measure political protest? In this article, we conceptualize political protest, build an index to measure it, and test its cross-national equivalence.

* E-mail: mario.quaranta@sumitalia.it

In the social sciences, it is well known that concepts have to ‘travel’ in order to compare phenomena (Sartori, 1970). Gerring (1999: 366) argues that ‘concept formation is a highly contextual process’, as is measurement, given that the two are intimately connected (Adcock and Collier, 2001). The aim of this article is to answer another question: is the measure of political protest equivalent across countries? It is argued that the issue of measurement equivalence remains unaddressed in political science (Jackman, 2008; Ariely and Davidov, 2012; Stegmuller, 2011),¹ in contrast with other social science disciplines, such as psychology (Van de Vijver and Tanzer, 2004). The test of measurement equivalence represents one of the most important phases of the research process because it guarantees that the analysis of a phenomenon is reliable (Jacoby, 1999). If we want to compare levels of political protest across different contexts, assessing measurement equivalence is a necessary step to provide unbiased estimates.

It is argued that the concept of political protest has one underlying dimension. A measurement model is tested to build a scale to determine individual and country scores of political protest. The ‘political protest’ scale from the European Values Study (EVS, 2011), first employed in the ‘political action’ study (Barnes and Kaase, 1979) and showing that the concept of political protest can be measured using five indicators (signing a petition, joining a boycott, attending lawful/peaceful demonstrations, joining unofficial strikes, occupying buildings or factories), is used for this purpose. This conceptualization thus follows the work of Barnes and Kaase (1979), who create a distinction between ‘conventional’ and ‘unconventional’ participation, later known as ‘protest politics’ or ‘non-violent protest behavior’, and that has been followed by, among others (Inglehart, 1990, 1997; Parry *et al.*, 1992; Norris, 2002; Benson and Rochon, 2004; Inglehart and Welzel, 2005; Dalton, 2008; Dalton *et al.*, 2010; Welzel and Deutsch, 2012).

Mokken Scale Analysis (MSA) is used to assess the cross-country equivalence of the political protest scale (Mokken, 1971; Sijtsma *et al.*, 1990, 2011; Sijtsma, 1998; Sijtsma and Molenaar, 2002; Van Schuur, 2003), a nonparametric scaling method within the family of Item Response Theory models. It has been proven to work better than Confirmatory Factor Analysis (CFA; Bollen, 1989) when dealing with dichotomous and polytomous items forming ordinal scales. To ensure the maximum degree of comparability, a set of Western European countries is used, as they are all established democracies with similar pasts and economic development levels.

The article develops as follows. First, I define the concept of political protest. Second, an outline of the problem of measurement equivalence in comparative research is offered. Third, the methodological strategy and data used are illustrated, and then the measurement equivalence analysis is presented. The last section concludes.

¹ A notable exception is the volume edited by Van Deth (1998a).

The concept and the measure of political protest

Until the late 1970s, political protest was excluded from the broader concept of political participation. In fact, intense forms of political behavior were considered irrational and disruptive (Rucht, 2007). No form of active political involvement that failed to focus on political personnel selection was considered political participation (Milbrath, 1965; Verba and Nie, 1972). The concept of political participation was based on the idea that one of its necessary components was the authoritative allocation of values (Teorell *et al.*, 2007). Despite its limits, this conceptualization had been very influential, especially in moments of great social change.

It is undoubtedly true that the concept of political participation is usually linked to the act of voting, but it should not be limited to just this. Consequently, the conceptualization of political participation was enlarged to capture innovative forms of political action. The ground-breaking ‘political action’ study (Barnes and Kaase, 1979) introduced a very important distinction in the field since ‘empirical research [had] not kept in pace the growing theoretical emphasis upon non institutionalized, non electoral political action, an emphasis that reflects the prominence of protest in the mass politics of Western democracies during the 1960s’ (Kaase and Marsh, 1979a: 27). On one hand is ‘conventional’ political participation, concerning all acts that form part of the constitutional process of interest aggregation and representation, mediated by political institutions, and which define the relationship between political authorities and citizens within the political arena. On the other is ‘unconventional’ political participation, or ‘political protest’, which is non-institutionalized direct political action, not necessarily aimed at disrupting or threatening the stability of liberal democracies. In fact, ‘direct political action generally, and political protest in particular, do not necessarily assume anti-regime protests but may form one element of an expanded repertoire of political action’ (Kaase and Marsh, 1979a: 27). Nor are they seen as having an ‘anti-system-directed orientation’ (Kaase and Marsh, 1979b: 157). As a consequence, political protest is considered ‘a means of political redress, namely [...] the use of tactics as petitions, demonstrations, boycotts, rent or tax strikes, unofficial industrial strikes, occupations of buildings, blocking of traffic, damage to property, and personal violence’ (Marsh and Kaase, 1979: 59).

All relevant studies following Barnes and Kaase’s seminal book (see Dalton, 1988, 2008; Jennings *et al.*, 1989; Parry *et al.*, 1992; Inglehart, 1997; Norris, 2002; Inglehart and Welzel, 2005; Van Deth *et al.*, 2007; Dalton *et al.*, 2010; Welzel and Deutsch, 2012) incorporated these forms in the *repertoire* of political participation, grasping changes that otherwise would not have been perceived: ‘[t]he analysis of protest politics shows that many of these forms of activity, such as petitions, demonstrations, and consumer boycotts, are fairly pervasive and have become increasingly popular during recent decades. Protest politics is on the rise as a channel of political expression and mobilization’ (Norris, 2002: 221).

The distinction between conventional and unconventional political participation introduced a key element: the object of political action. Conventional participation aims at the political arena, public institutions, while political protest can also target other objects that are not necessarily part of the political system. As Teorell *et al.* (2007: 336) argue that in Western democracies ‘the authoritative allocations of values is *not* the sole responsible of state actors of the public sector. As a result, [...] non-governmental institutions may be targeted by citizen attempts to influence political outcomes’. Unconventional actions are extra-representational and can be undertaken to influence both the public and private sectors. Demonstrations are a clear example. Very often, they are a means of opposition to governmental decisions, such as cuts to the welfare state sector, but they are also oriented to influence public opinion or private firms. We may note in this vein the ‘anti-globalization movement’ or ‘global justice movement’ (Della Porta *et al.*, 2006; Della Porta, 2007), critical of the neoliberal economic turn, that often concluded their gatherings with large demonstrations. In fact, as Della Porta and Diani (2006: 165) argue: ‘heterogeneous and initially loosely connected groups had mobilized together, mainly against international organizations, using different strategies: from lobbying to marches, from boycotts to petitions, from strikes to netstrikes [...] demonstrators from many countries challenged the legitimacy of the decisions of some international governmental organizations and sought to hinder their plans. They did not do so through normal diplomatic channels or through elections. Rather, they sought to influence public opinion in various ways’.

Finally, the actions forming its *repertoire* can be considered hierarchically ordered (Van Deth, 1986; Kaase, 1989): ‘[t]he first threshold indicates the transition from conventional to unconventional politics. Signing petitions and participating in lawful demonstrations are unorthodox political activities but still within the bounds of accepted democratic norms. The second threshold represents the shift to direct action techniques, such as boycotts. A third level of political activities involves illegal, but nonviolent, acts. Unofficial strikes or a peaceful occupation of a building typify this step. Finally, a fourth threshold includes violent activities such as personal injury or physical damage’ (Dalton, 1988: 65).

For the purposes of this work, ‘political protest’ or ‘unconventional participation’ is considered a direct form of political participation taking place without the intermediation of institutional actors. Protest may arise from social organizations that vary in terms of their structure, membership, scope, resources, and capacity for mobilization (Della Porta and Diani, 2006). Political actions must be free and organized by civil society, not by governmental institutions looking for popular support. As a form of direct participation, protest requires an extended effort and a certain degree of conflict. Potentially, it produces high pressure on the contested actors, although it may not produce the expected outcome. Protest may also presuppose collective action (Tilly and Tarrow, 2006), although it is not necessary. Unconventional political actions are not professional activities and

must be voluntary. In order to be defined ‘political’, an unconventional action must have a target. The action must aim to influence something, either governmental institutions or actors belonging to the private sector (Teorell *et al.*, 2007).

A scale that has been used in the literature to measure the concept of political protest is that proposed by Barnes and Kaase (1979), but it has never been proven to be valid across a large number of countries. In fact, several comparative studies using this scale apply it to very different contexts that not only have different cultures, but also different historical legacies, paths of democratization, and, consequently, modes of political engagement.

These studies provide some evidence of the scale’s internal reliability but may be insufficient to ensure measurement equivalence. For instance, in ‘Democratic Phoenix’ Norris (2002: 195–196) elaborates a scale of ‘protest activism’ using the five items included in the World Value Survey, arguing that these form a distinct dimension of engagement, different from other forms of political and social involvement, such as voting.

A principal component factor analysis supports this argument. The results are clear but the analysis is run on the pooled sample without testing cross-country equivalence. Benson and Rochon (2004: 441–442) use Guttman scaling to assess the reliability of the political protest scale, but no evidence of cross-country equivalence is given. Dalton (2004: 177), analyzing the correlation between political trust and the political protest scale, says the latter is a count of five political activities. Similarly, Dalton *et al.* (2010: 61) use the same scale, providing a principal component analysis and emphasizing the fact that just one factor with an eigenvalue greater than one emerges. The following sections illustrate potential problems arising from the lack of measurement equivalence in comparative research, and delineate an empirical strategy to assess it.

The problem of cross-country measurement equivalence

In comparative studies, researchers use numerous cases to draw inferences and test theories. Among others, two elements constitute fundamental aspects of the comparative research process: concepts and measures. On one hand, concepts define the phenomena under study. Sartori (1970) warns that a potential risk for the reliability of a study, particularly relevant in comparative politics, is the problem of ‘conceptual stretching’. Comparative research is often weakened by the incorrect use of concepts that are not applicable to different contexts. On the other hand, comparative researchers build measures that should be comparable across the contexts they study. Concept formation and measurement are two strictly intertwined processes (Adcock and Collier, 2001).

The careful definition of concepts to be used in a comparative design is a necessary stage to construct empirical measures and operationalize them. Once the first step has been taken, the comparative researcher faces a great challenge: assessing measurement equivalence. As with concepts, measures should also be

valid across the different contexts to which they are applied. Cross-national research has grown dramatically over the last two decades because of international survey projects (Norris, 2009), but apparently the level of attention devoted to the assessment of the cross-national equivalence of measurement instruments has been insufficient, especially in the field of political science (King *et al.*, 1994, 2004; Jacoby, 1999; Adcock and Collier, 2001; Jackman, 2008; Ariely and Davidov, 2012). As Adcock and Collier (2001: 534) maintain: ‘this concern with context can arise when scholars are making comparisons across different world regions or distinct historical periods [...] the potential difficulty that context poses for valid measurement [...] deserves more attention in political science’. Further, the process of measure validation in comparative research is linked to theory testing (King *et al.*, 1994; Jacoby, 1999), often the final aim of scientific research, as without reliable instruments it is not possible to provide reliable results, nor draw inferences from the cases under investigation.

When we deal with several contexts, we need to be sure that the measurement instrument is able to ‘capture’ the underlying latent construct we want to measure (Jackman, 2008). The operationalization process in comparative research requires the recognition that concepts have a ‘contextual specificity’ (Adcock and Collier, 2001: 529–530). Unfortunately, this is not sufficient to construct valid measures. Dismissing the importance of measurement equivalence may have very important consequences on the comparative study’s validity. First, the conclusions drawn from a study using a non-tested measurement instrument representing a latent concept cannot be accepted as valid (Steenkamp and Baumgartner, 1998). Second, if the measurement instrument is not validated comparatively, we cannot be sure that the relationships between it and other variables are reliable.

In brief, scholars need to assess whether or not a measurement instrument works in a similar manner across countries (Steenkamp and Baumgartner, 1998; Vandenberg and Lance, 2000). The latent construct structure must be the same in all the contexts to which the measurement instrument will be applied (Byrne, 2008). This part of the research process is necessary for both descriptive and causal inference (Adcock and Collier, 2001). Thus, cross-country equivalence assessment should be a central concern for researchers testing hypotheses and theories in different contexts (Hui and Triandis, 1985). If researchers want to meaningfully compare the same concept in different contexts, they must be sure that the measure representing its underlying latent trait is comparable in all those contexts in order to avoid any potential bias.

Assessing measurement equivalence

What is measurement equivalence and how do we assess it? It implies the concepts of validity and reliability. Bollen (1989: 184) conceptualizes validity as an issue ‘concerned with whether a variable measures what it is supposed to measure’. With regard to reliability, Adcock and Collier (2001: 531) argue that ‘random

error, which occurs when repeated applications of a given measurement procedure yield inconsistent results, is conventionally labeled a problem of reliability'. Measurement equivalence can be defined as 'whether or not, under different conditions of observing and studying phenomena, measurement operations yield measures of the same attribute' (Horn and McArdle, 1992: 117), meaning that what we observe through measurement is reliable and valid.

In Classical Test Theory (CTT), measurement equivalence has three different levels (Horn and McArdle, 1992; Steenkamp and Baumgartner, 1998; Vandenberg and Lance, 2000; Byrne, 2008).² The first level of equivalence is 'configural invariance'. This mainly refers to the factor loadings structure, meaning that in all countries the latent construct shows the same configuration of factor loadings. The second level of measurement equivalence is 'metric invariance'. This type of equivalence requires that all factor loadings, measuring the strength of the relationship between items and the construct, be equal across countries. The last level of measurement equivalence is 'scalar invariance', which implies that the differences in the group means are due to the means of the underlying constructs, which ensures that the mean estimates are not biased.

However, CTT has been criticized for some assumptions that cannot in most cases be met. In particular, it assumes that the items measuring a latent trait are parallel, meaning that they have similar means and standard deviations. Further, it does not take into account how respondents answer to items and, therefore, does not consider their 'easiness' or 'difficulty'. Table 1 illustrates an example of hypothetical items, which are ordered to form a perfect scale along which respondents answering positively to more difficult items have also answered positively to easier items.³ According to its critics, CTT is not able to capture this feature, especially when dealing with dichotomous or ordinal items (Sijtsma and Molenaar, 2002; Van Schuur, 2003). The ordinal nature of a scale has relevant theoretical implications, as each category represents cumulative 'steps' formed by items that do not have the same probability of being chosen by respondents.

Since CTT presents these problems, I use MSA⁴ here within the framework of Item Response Theory (Mokken, 1971; Sijtsma *et al.*, 1990, 2011; Sijtsma, 1998; Sijtsma and Molenaar, 2002; Van Schuur, 2003) – a development of the Guttman scale (Guttman, 1945). MSA has three assumptions. The first concerns the unidimensionality of the latent trait. The second concerns the monotonicity of the Item Response Function, meaning that as the probability of a positive answer to an item increases, for example participation in lawful demonstrations, the latent score, for example the index of political protest, also increases. The third is that

² The standard technique is Multi Group Confirmatory Factor Analysis (Bollen, 1989).

³ Van Schuur (2003: 139–141) uses this example, with six items, to introduce Mokken Scale Analysis and its differences with principal component or factor analysis.

⁴ The aim of the article is not to present this scaling technique in detail. For introductory readings, see Sijtsma and Molenaar (2002) and Van Schuur (2003).

Table 1. Example of a perfect Guttman scale

Response pattern	Item difficulty				
	Low				High
	Item 1	Item 2	Item 3	Item 4	Item 5
1	0	0	0	0	0
2	1	0	0	0	0
3	1	1	0	0	0
4	1	1	1	0	0
5	1	1	1	1	0
6	1	1	1	1	1

Note: adapted from Van Schuur (2003).

respondents are locally independent. This means that the responses to some items by an individual depend on ability, defined by the latent trait, and not by other individual and item characteristics (Van Schuur, 2003: 145). This last assumption is often referred to as Invariant Items Ordering (Sijtsma *et al.*, 2011). It follows that when a set of items form a Mokken scale the simple sum score can be used as the latent trait score (Mokken, 1971).

MSA has several advantages over covariance-based measurement models. First, it allows the researcher to determine the probability that a respondent has answered positively to an item conditional on other items. For instance, the probability that a respondent has attended a demonstration should be higher if he or she has signed a petition. Second, MSA is a probabilistic technique and not a deterministic one. This means it takes measurement errors into account. Guttman scaling assumes that a respondent has to follow a precise pattern of answers: a respondent answering positively to a difficult item also answers positively to a less difficult item. MSA, instead, accounts for the possibility that a respondent will not follow the hypothetical hierarchy of items. Third, MSA detects the items that do not conform to a cumulative scale and drops them through an iterative pairwise process. In fact, MSA uses a hierarchical clustering procedure made up of the following steps: (1) it finds the pair of items with the highest scalability coefficient; (2) it finds the next best item in the scale and repeats step one for all the items. Fourth, it requires items forming the scale to be sufficiently homogeneous among themselves. This makes the measurement instrument more reliable. Fifth, MSA works well when applied to a small number of items, unlike CFA. It can, eventually, be used as a confirmatory test that a set of items form a unidimensional and cumulative scale across different populations.

Since the aim here is to test the presence of an ordinal scale, a Double Monotonicity Model is used according to the following strategy. First, the homogeneity of the scale is tested – does the scale measure one latent trait and can it be cumulated using two coefficients? H is the scalability coefficient for the overall scale and must be ≥ 0.30 . If the coefficient is between 0.30 and 0.40, the scale is

‘weak’, if the coefficient is between 0.40 and 0.50 it is ‘moderate’, and if the coefficient is higher than 0.50 the scale is considered ‘strong’. When the coefficient is equal to one, we have a perfect Guttman scale, meaning that all respondents follow a hierarchical pattern in answering items. H_i , which measures the scalability of the single items, must be ≥ 0.30 (Mokken, 1971; Van Schuur, 2003). Eventually, another coefficient, ρ , provides a measure of the reliability of the scale (Sijtsma and Molenaar, 1987).⁵ Second, item ordering is evaluated so as to assess whether or not respondents follow, on average, the same response pattern in the analyzed countries.⁶

Data

As mentioned, the ‘political protest’ scale can be constructed using five items in the EVS (2011). Western European countries only are selected and a cross-sectional design adopted.⁷ Previous studies suggest a most similar systems design when studying political participation, since in recently democratized countries the patterns and the extent to which several modes of participation are used are very different compared with consolidated democracies (Teorell *et al.*, 2007; Morales, 2009). Other publications (Norris, 2002; Dalton *et al.*, 2010) study different contexts. Yet it may be argued that the chances of bias in such analytical settings are very high since fully consolidated democracies and still-developing democracies present different modes of civic and political engagement. Further, the levels of development and democratization have a significant effect on levels of political protest (Dalton and Van Sickle, 2005; Dalton *et al.*, 2010). The countries included in the analysis are: Austria, Belgium, Cyprus, Denmark, Finland, France, Germany, Great Britain, Greece, Ireland, Iceland, Italy, Luxembourg, Malta, the Netherlands, Norway, Portugal, Spain, Sweden, and Switzerland.⁸

The question wording for the political participation scale items in the EVS questionnaire is as follows:⁹

Now I'd like you to look at this card. I'm going to read out some different forms of political action that people can take, and I'd like you to tell me, for each one,

⁵ This step would be sufficient to test a Monotone Homogeneity Model, but does not guarantee the presence of an ordinal scale.

⁶ The property of Invariant Item Ordering assumes that item popularity is the same in different points of the latent trait (Sijtsma *et al.*, 2011). With this step we aim to assess whether the item ordering is the same across countries.

⁷ A longitudinal design was avoided since the amount of information arising from the analysis would make the article too lengthy. However, the procedure described here can easily be applied to such a design to assess equivalence over time.

⁸ The following country abbreviations are used: AT, BE, CY, DK, FI, FR, DE, GB, GR, IE, IS, IT, LU, MT, NL, NO, PT, ES, SE, and CH.

⁹ Although other cross-national surveys are freely available, the EVS was chosen because its questionnaire includes more items about political protest. Further, the items belonging to the EVS are very similar to those used by Barnes and Kaase (1979). We use these two surveys later in the analysis.

whether you have actually done any of these things, whether you would/might do it or would not/never, under any circumstances, do it/any of them:

- Signing a petition
- Joining in boycotts
- Attending lawful/peaceful demonstrations
- Joining unofficial strikes
- Occupying buildings or factories.

The original coding scheme assigns a score of ‘three’ to those who would never carry out the political action, two to those who might, and one to those who have carried out the political action. The items have been recoded reversing the scale and assigning a zero to those who would never carry out and might carry out the actions, and one to those who have carried out the political actions.¹⁰ The index aims to measure only actual political protest, not potential protest (Marsh and Kaase, 1979: 59).

Table 2 presents the mean scores and standard deviations for each country for the forms of political action included in the analysis.¹¹ The table illustrates that means and standard deviations are not similar between the items both in the pooled sample and in the separate country samples. Further, it should be noted that ‘signing a petition’ is a form of action used frequently in Northern and Continental European countries. The highest mean score is present in Sweden, where almost 80% of the sample has signed a petition. In addition, Norway and Denmark present high levels of petitioning. In Continental Europe, France and Germany score quite high.

In Southern European countries, this mode of action is less popular compared with other contexts – the lowest scores are in Malta, Portugal, Cyprus, and Greece. As far as ‘joining in boycotts’ is concerned, the same pattern is present in the selected countries. Iceland and Finland have the highest scores, while in Southern Europe this form of action is less present. ‘Attending a lawful demonstration’ is, instead, a more common form of political action – highest in France, Italy, and the Nordic countries. It is lower in Portugal, Great Britain, and Ireland. With regard to ‘joining unofficial strikes’, the highest percentages of respondents taking part in such actions are observed in France and Denmark. Conversely, in Cyprus, Portugal, Germany, and Sweden, we find the lowest levels of participation in this form. The last form of political action, ‘occupying buildings or factories’, is most frequent in France and Italy, while in Finland and Iceland it is marginal among the five forms of action.

¹⁰ As is known, non-response is considered a normal problem in survey research. The solution to this problem is multiple imputation (Rubin, 1987; King *et al.*, 2001). A “chained equations” imputation approach (Raghunathan *et al.*, 2001) and socio-demographic variables are used to perform the imputation (gender, age and education).

¹¹ Items are ordered as in the EVS questionnaire.

Table 2. Item means and standard deviations for each country and for the pooled sample

Country	Petition		Boycotts		Demonstrations		Strikes		Occupying	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
AT	0.48	0.50	0.09	0.28	0.16	0.36	0.04	0.20	0.02	0.15
BE	0.61	0.49	0.11	0.31	0.30	0.46	0.07	0.26	0.04	0.19
CH	0.62	0.48	0.13	0.34	0.25	0.43	0.04	0.19	0.02	0.15
CY	0.19	0.39	0.03	0.16	0.23	0.42	0.00	0.06	0.00	0.06
DE	0.55	0.50	0.08	0.27	0.29	0.45	0.02	0.15	0.02	0.13
DK	0.66	0.47	0.18	0.38	0.34	0.47	0.19	0.39	0.02	0.14
ES	0.36	0.48	0.06	0.24	0.36	0.48	0.08	0.27	0.02	0.15
FI	0.50	0.50	0.23	0.42	0.13	0.34	0.04	0.20	0.01	0.07
FR	0.68	0.47	0.16	0.36	0.45	0.50	0.11	0.32	0.09	0.28
GB	0.65	0.48	0.13	0.34	0.14	0.35	0.07	0.25	0.02	0.13
GR	0.19	0.39	0.06	0.24	0.22	0.42	0.04	0.20	0.06	0.24
IE	0.52	0.50	0.09	0.29	0.15	0.35	0.04	0.19	0.02	0.14
IS	0.54	0.50	0.30	0.46	0.27	0.44	0.04	0.19	0.00	0.07
IT	0.49	0.50	0.12	0.33	0.37	0.48	0.08	0.27	0.09	0.29
LU	0.59	0.49	0.12	0.33	0.35	0.48	0.07	0.26	0.02	0.14
MT	0.32	0.47	0.08	0.28	0.16	0.36	0.04	0.20	0.01	0.11
NL	0.50	0.50	0.11	0.31	0.21	0.41	0.03	0.18	0.02	0.15
NO	0.73	0.44	0.18	0.38	0.30	0.46	0.04	0.19	0.01	0.12
PT	0.21	0.41	0.04	0.21	0.11	0.32	0.02	0.13	0.01	0.08
SE	0.78	0.41	0.22	0.41	0.21	0.40	0.02	0.13	0.01	0.09
Pooled	0.51	0.50	0.12	0.32	0.25	0.43	0.06	0.23	0.03	0.16

AT = Austria; BE = Belgium; CH = Switzerland; CY = Cyprus; DE = Germany; DK = Denmark; ES = Spain; FI = Finland; FR = France; GB = Great Britain; GR = Greece; IE = Ireland; IS = Iceland; IT = Italy; LU = Luxembourg; MT = Malta, NL = the Netherlands; NO = Norway; PT = Portugal; SE = Sweden.

Analysis

This section presents the results of the MSA.¹² Table 3 shows the H coefficients for each country and the pooled sample. The coefficient for the pooled sample is about 0.56, which tells that if we do not take countries' heterogeneity into account the items can be summed up in a scale.

The same analysis can be carried out for the coefficients estimated on the separate country samples. In almost all cases, the coefficient is higher than 0.5, which represents the threshold for a strong scale. The coefficients are highest in Malta, Ireland, and Germany and lowest in Denmark.

Looking at the standard errors and Z statistics,¹³ it is noteworthy that the coefficients are highly statistically significant. Further, ρ coefficients confirm that the scale

¹² The analysis was carried out using the library 'Mokken' in R (Van der Ark, 2007).

¹³ It has a normal standard distribution for a large N .

Table 3. *H* coefficients with standard errors, *Z* statistics, and ρ for each country and for the pooled sample

	<i>H</i>	SE	<i>Z</i>	ρ
AT	0.61	0.03	28.5	0.68
BE	0.52	0.03	25.3	0.64
CH	0.61	0.03	23.4	0.65
CY	0.56	0.04	19.0	0.67
DE	0.65	0.02	29.0	0.65
DK	0.42	0.02	23.7	0.58
ES	0.57	0.02	32.7	0.71
FI	0.55	0.03	20.9	0.64
FR	0.58	0.02	33.1	0.72
GB	0.59	0.03	28.2	0.69
GR	0.50	0.02	36.6	0.70
IE	0.66	0.04	23.7	0.70
IS	0.58	0.03	21.4	0.69
IT	0.56	0.02	35.3	0.72
LU	0.56	0.02	26.9	0.64
MT	0.66	0.03	34.5	0.73
NL	0.58	0.03	27.4	0.66
NO	0.55	0.03	18.0	0.59
PT	0.53	0.04	25.8	0.67
SE	0.56	0.04	19.0	0.65
Pooled	0.57	0.01	121.1	0.66

AT = Austria; BE = Belgium; CH = Switzerland; CY = Cyprus;
 DE = Germany; DK = Denmark; ES = Spain; FI = Finland; FR = France;
 GB = Great Britain; GR = Greece; IE = Ireland; IS = Iceland; IT = Italy;
 LU = Luxembourg; MT = Malta, NL = the Netherlands; NO = Norway;
 PT = Portugal; SE = Sweden.

is a reliable one. If the cross-national equivalence analysis of the political protest scale were to be stopped here, it would be accepted, as the *H* coefficients are high in all the countries. However, the *H* coefficient only measures the homogeneity of the scale and the distance from a perfect Guttman model. Therefore, in order to further assess the cross-national equivalence of the scale, it is important to take into account the *H* coefficients for the items (Table 4).

In almost all cases, the *H* coefficients of the items are ≥ 0.30 and highly statistically significant. Only one case presents *H* coefficients below the suggested threshold: Portugal for the items 'strikes' and 'occupying'. Therefore, the Mokken scale should not be accepted for this case. According to the analysis of the scale's *H* coefficients and the items' *H* coefficients, if we are to meaningfully compare the sum scores of the political protest scale we should eliminate Portugal. The monotonicity assumption is not violated in the separate country samples. Having confirmed the presence of a unidimensional construct measuring the underlying concept of political protest and its cross-national equivalence for a

Table 4. *H* coefficients and *Z* statistics of the five items for each country and for the pooled sample

Country	Petition		Boycotts		Demonstrations		Strikes		Occupying	
	<i>H</i>	<i>Z</i>	<i>H</i>	<i>Z</i>	<i>H</i>	<i>Z</i>	<i>H</i>	<i>Z</i>	<i>H</i>	<i>Z</i>
AT	0.74	17.6	0.58	21.0	0.63	22.5	0.52	16.8	0.47	12.9
BE	0.62	15.6	0.47	17.2	0.60	20.7	0.41	14.4	0.42	11.8
CH	0.66	13.3	0.57	17.4	0.64	19.6	0.59	14.1	0.51	10.6
CY	0.58	17.3	0.56	9.5	0.59	17.3	0.37	4.3	0.30	3.5
DE	0.67	19.8	0.64	19.6	0.66	23.4	0.56	14.6	0.67	16.1
DK	0.53	14.4	0.38	16.8	0.49	20.3	0.32	14.0	0.43	7.2
ES	0.52	22.9	0.59	20.3	0.57	24.9	0.59	20.7	0.63	15.0
FI	0.70	16.6	0.55	17.4	0.50	15.2	0.31	6.6	0.57	5.6
FR	0.75	21.3	0.48	19.7	0.70	25.6	0.47	18.9	0.50	18.5
GB	0.84	15.3	0.56	23.1	0.58	23.2	0.42	14.6	0.63	14.0
GR	0.53	25.8	0.46	22.0	0.62	28.6	0.49	21.1	0.36	17.3
IE	0.84	14.9	0.60	17.5	0.66	18.7	0.59	14.3	0.51	10.2
IS	0.64	15.5	0.54	17.2	0.56	17.4	0.56	7.6	0.74	5.1
IT	0.61	22.4	0.53	22.6	0.64	26.6	0.51	20.3	0.47	20.0
LU	0.62	18.2	0.52	18.1	0.63	22.8	0.46	14.6	0.40	8.1
MT	0.69	22.4	0.72	27.9	0.65	26.2	0.56	18.6	0.60	13.4
NL	0.73	19.4	0.57	20.0	0.61	21.9	0.36	10.7	0.51	13.4
NO	0.72	12.0	0.50	13.6	0.58	16.2	0.31	5.6	0.44	6.2
PT	0.63	21.7	0.51	17.7	0.60	23.3	0.23	6.5	0.12	2.3
SE	0.68	9.2	0.57	18.1	0.52	16.5	0.41	6.1	0.60	7.4
Pooled	0.66	80.9	0.53	84.7	0.60	96.4	0.46	64.3	0.48	53.2

AT = Austria; BE = Belgium; CH = Switzerland; CY = Cyprus; DE = Germany; DK = Denmark; ES = Spain; FI = Finland; FR = France; GB = Great Britain; GR = Greece; IE = Ireland; IS = Iceland; IT = Italy; LU = Luxembourg; MT = Malta, NL = the Netherlands; NO = Norway; PT = Portugal; SE = Sweden.

specific number of cases, an additive scaling procedure is used to construct the scale. The individual scores on each item are summed to obtain the overall political protest index.

The index ranges from zero to five, where zero represents an individual who has never engaged nor thought of engaging in any of the five forms of unconventional political participation and five represents an individual who has engaged in all of the forms. This scale may be used similarly across countries if its ordering is not taken into account.¹⁴ Figure 1 shows the mean point estimates with 95% confidence intervals for the comparable countries and the pooled sample.¹⁵

¹⁴ An example of cross-national equivalence of a Mokken scale can be found in Van der Meer *et al.* (2009).

¹⁵ A simulation procedure is used to estimate the uncertainties based on Bayesian inference. See Kerman and Gelman (2007).

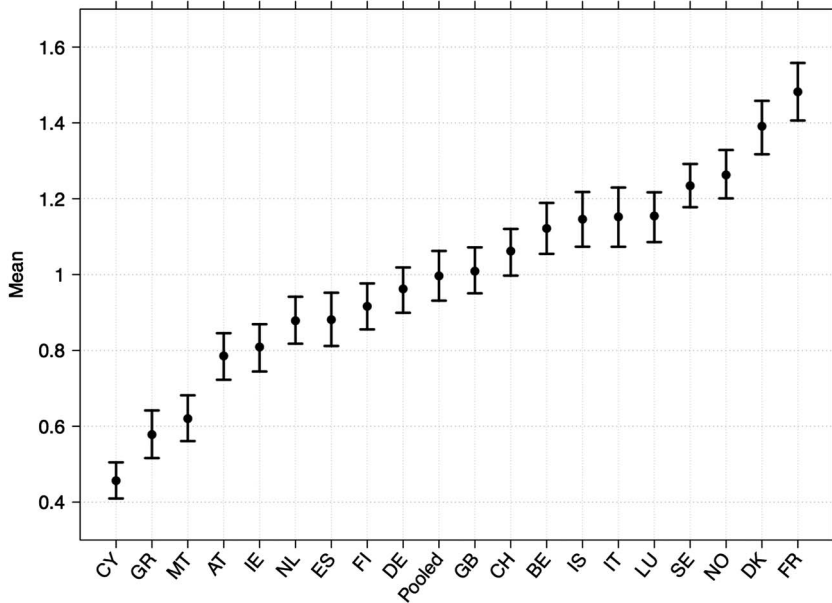


Figure 1 'Political protest' index means with 95% confidence intervals for each country and for the pooled sample

Within the set of countries where the items form a homogeneous scale, there is a substantial amount of variation. The index mean for the pooled sample is about one and countries can be classified in three distinctive groups. There is a low unconventional participation group of countries composed by Cyprus, Greece, Malta, Austria, Ireland, and the Netherlands. Then, the group of average countries with mean scores that match the pooled sample average: Spain, Finland, Germany, Great Britain, Switzerland, Belgium, Iceland, Luxembourg, and Italy. Eventually, four countries, Norway, Sweden Denmark, and France, show a mean score above the average.

Although the scale is comparable across these countries, it is not yet clear whether, for example, respondents scoring three in each country have engaged in the same actions. This can be assessed by looking at the Invariant Item Ordering violations in the single samples. Significant violations were found for Spain and Sweden. This means that in these two countries we do not have an ordinal scale when using five items.

The second step of the analysis tests whether the item ordering is the same in the different samples so that we can say, for instance, that a respondent scoring two in France has engaged in the same actions as a respondent scoring two in Italy. We can now tell how many citizens have engaged in zero, one, two, three, four, and five political actions.¹⁶ To further explore the political protest scale's

¹⁶ It would possible to use a linear model if the latent trait were considered as continuous, or a Poisson model, if considered as a set of counts.

Table 5. Item ordering for each country and for the pooled sample

Country	Ordering				
AT	P	D	B	S	O
BE	P	D	B	S	O
CH	P	D	B	S	O
CY	D	P	B	O	S
DE	P	D	B	S	O
DK	P	D	S	B	O
ES	P	D	S	B	O
FI	P	B	D	S	O
FR	P	D	B	S	O
GB	P	D	B	S	O
GR	D	P	O	B	S
IE	P	D	B	S	O
IS	P	B	D	S	O
IT	P	D	B	O	S
LU	P	D	B	S	O
MT	P	D	B	S	O
NL	P	D	B	S	O
NO	P	D	B	S	O
PT	P	D	B	S	O
SE	P	B	D	S	O
Pooled	P	D	B	S	O

AT = Austria; BE = Belgium; CH = Switzerland; CY = Cyprus;
 DE = Germany; DK = Denmark; ES = Spain; FI = Finland; FR = France;
 GB = Great Britain; GR = Greece; IE = Ireland; IS = Iceland; IT = Italy;
 LU = Luxembourg; MT = Malta, NL = the Netherlands; NO = Norway;
 PT = Portugal; SE = Sweden; P = signing a petition; D = attending lawful
 demonstrations; B = joining in boycotts; S = joining unofficial strikes;
 O = occupying buildings or factories.

cross-country equivalence, the analysis should check how the items are ordered in the country samples in order to assess whether the points of the scale hold the same meaning in different contexts. Table 5 includes the item ordering, which can make an optimized Mokken scale. In most cases, the item ordering is the same as in the pooled sample. Five patterns describing different item orderings emerge from the data. However, to use the scale as an ordinal one and compare it across countries, only those countries showing the same item ordering should be used. The countries that may be compared using a five-item scale with six points, with an ordering equal to (1) petition, (2) demonstrations, (3) boycotts, (4) strikes, and (5) occupying, are: Austria, Belgium, France, Germany, Great Britain, Ireland, Malta, the Netherlands, Norway, and Switzerland. Finland and Iceland can also be compared, but follow a different item ordering. Denmark, Greece, and Italy cannot be compared with the other cases since they all present unique item orderings. For Spain, Sweden, and Portugal no ordinal scale may be built, as the

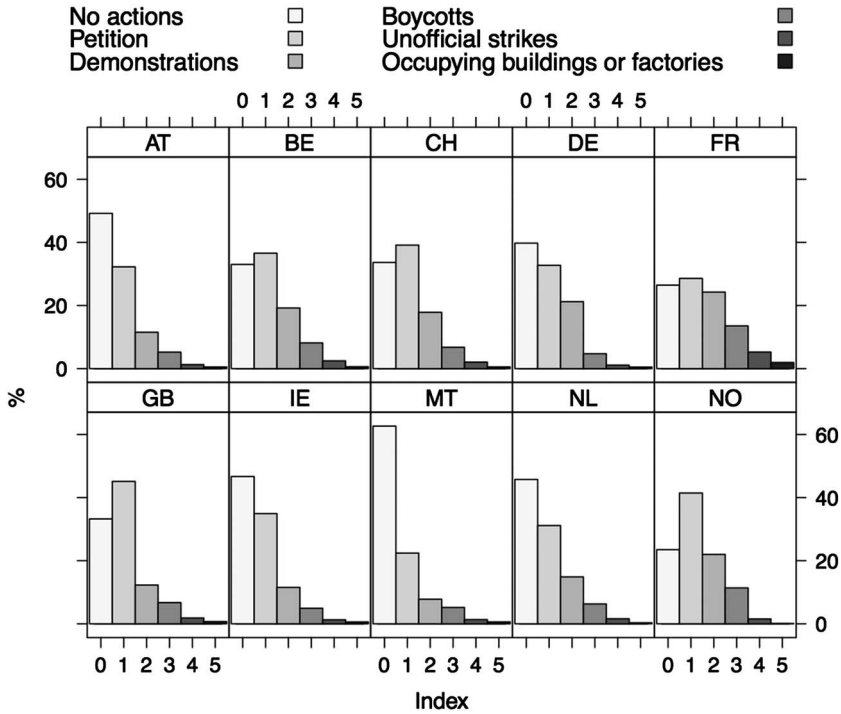


Figure 2 'Political protest' index distribution in the comparable countries

first two have different item orderings compared with the other countries, and the latter has an inadequate *H*.

The levels of political protest in the pooled sample and in the country samples can now be addressed. Disregarding the countries' heterogeneity, we can say that those who never engaged in political protest nor thought of doing so represent about 40% of the sample. Those who signed a petition represent about 34%, those who attended lawful demonstrations 16%, the percentage of respondents who boycotted products is seven, while those who joined an unofficial strike and occupied a factory or a building account for, respectively, two and almost 1% (see Figure 2). The highest percentage of inactive respondents is in Malta (62%), while the lowest is in Norway (23%). As far as petitions are concerned, the country in which these are the most used is Great Britain, while respondents use them the least in Malta. The highest numbers of demonstrators, boycotters, strikers, and occupiers are found in France, and the lowest in, respectively, Malta, Ireland, Germany, and Norway.

If a comparative researcher is interested in comparing more cases, he or she may enlarge the scope by discarding an item and checking whether the new set of items forms a 'more' comparable scale, a sub-scale of political protest. Table 6 demonstrates that removing the item 'boycotts' would lead to more cases in which the items form a Mokken scale. All the *H* coefficients for the scale are higher than

Table 6. *H* coefficients, standard errors, *Z* statistics, items' *H* coefficients for the political protest scale constructed with four items for each country and for the pooled sample

Country	<i>H</i>	SE	<i>Z</i>	Items <i>H</i>				Ordering			
				P	D	S	O				
AT	0.64	0.04	19.7	0.67	0.68	0.59	0.51	P	D	S	O
BE	0.56	0.03	18.7	0.58	0.60	0.53	0.48	P	D	S	O
CH	0.65	0.04	15.9	0.63	0.70	0.63	0.57	P	D	S	O
CY	0.56	0.04	16.4	0.56	0.59	0.52	0.32	D	P	O	S
DE	0.66	0.03	21.5	0.64	0.68	0.60	0.67	P	D	S	O
DK	0.46	0.03	16.9	0.47	0.49	0.44	0.43	P	D	S	O
ES	0.56	0.03	26.1	0.48	0.54	0.72	0.71	P	D	S	O
FI	0.54	0.05	11.6	0.68	0.58	0.28	0.57	P	D	S	O
FR	0.66	0.02	26.5	0.70	0.69	0.60	0.62	P	D	S	O
GB	0.62	0.04	17.1	0.79	0.62	0.46	0.65	P	D	S	O
GR	0.53	0.03	29.4	0.51	0.60	0.59	0.41	D	P	O	S
IE	0.72	0.05	16.3	0.81	0.75	0.65	0.54	P	D	S	O
IS	0.65	0.05	12.7	0.66	0.65	0.58	0.77	P	D	S	O
IT	0.58	0.03	27.2	0.56	0.61	0.59	0.57	P	D	O	S
LU	0.59	0.03	20.0	0.57	0.64	0.61	0.43	P	D	S	O
MT	0.61	0.04	21.9	0.63	0.62	0.58	0.58	P	D	S	O
NL	0.60	0.03	19.1	0.68	0.64	0.42	0.55	P	D	S	O
NO	0.62	0.05	11.8	0.69	0.69	0.43	0.45	P	D	S	O
PT	0.54	0.04	19.1	0.61	0.61	0.27	0.10	P	D	S	O
SE	0.53	0.07	7.6	0.52	0.58	0.44	0.50	P	D	S	O
Pooled	0.59	0.01	87.3	0.61	0.63	0.56	0.53	P	D	S	O

AT = Austria; BE = Belgium; CH = Switzerland; CY = Cyprus; DE = Germany; DK = Denmark; ES = Spain; FI = Finland; FR = France; GB = Great Britain; GR = Greece; IE = Ireland; IS = Iceland; IT = Italy; LU = Luxembourg; MT = Malta, NL = the Netherlands; NO = Norway; PT = Portugal; SE = Sweden; P = signing a petition; D = attending lawful demonstrations; S = joining unofficial strikes; O = occupying buildings or factories.

0.4, and in most cases they are higher than 0.5. Further, they are all highly significant. The *H* coefficients of the items are lower than 0.30, as before, for Portugal for the items 'strike' and 'occupying'. However, the item ordering is now similar across more countries. In fact, using five actions the scale is comparable for 10 countries, while excluding the item 'boycotts' makes the scale comparable for up to 16 countries, thus allowing a broader scope of analysis.

Welzel and Deutsch (2012: 469) argue that a scale of political protest can be constructed using three items, getting rid of 'joining unofficial strikes' and 'occupying buildings or factories', for several reasons: 'First, these activities are closer to violence, so including them blurs the focus on non-violent protest. Second, these activities stick out from the others as being by far the least popular ones. They are used in every sample by such minor proportions of the respondents (consistently below 5%) that

responses are fully within the margin of sampling error'. In line with this argument, the scale was reanalyzed using three items.¹⁷ The *H* coefficient for the whole scale is high (above 0.50) in all the countries, showing that this sub-scale is very strong. *Z* statistics strengthen the reliability of the scale since all coefficients are highly statistically significant. The *H* coefficients of the items also tell that the three-item scale is acceptable since all values are >0.30. The most popular item ordering is: (1) 'petition', (2) 'demonstrations', and (3) 'boycotts'. This response pattern is the same in 15 countries, while it varies for the other five. This analysis shows that reducing the number of items allows the scale to be applied in all 20 Western European countries, if item ordering is not taken into account. This means that a cumulative scale can be constructed, ranging from zero to three, where zero represents an individual who has not engaged in any political action and three represents a respondent who has engaged in all actions. To analyze political protest in a comparative perspective while being certain that each point of the scale has the same meaning in all countries, Cyprus, Finland, Greece, Iceland, and Sweden should be excluded.

Since the International Social Survey Programme (ISSP, 2007) and the European Social Survey (ESS, 2011) contain the same three items, the scale is validated using these surveys.¹⁸ The MSA on the ISSP data (Citizenship survey) shows that the scale is homogeneous in all the countries ($H \geq 0.30$). The *H* coefficients of the items are all above the suggested threshold. Conversely, the MSA on the ESS data (fourth round) suggests that Denmark, Germany, and Finland should all be discarded on the basis of low *H* coefficients for items. Figure 3 presents the mean scores of the cumulative scales, built using three items and the EVS, ISSP, and ESS data, for each country where the data are available and *H* coefficients are above the threshold of 0.30.¹⁹

Surprisingly, the scores are not similar among the three data sets used. It seems the ISSP data overestimate the mean scores, while the ESS data underestimate them. The EVS data always present scores that fall in between the other two data sets. The item ordering for both the ISSP and ESS data is different from that found using the EVS data. The item ordering for the two data sets is most frequently: (1) 'petition', (2) 'boycotts', and (3) 'demonstrations'.²⁰ It may be possible that the differences in the means and item orderings are due to different question wordings in the three questionnaires and not to systematic bias. In fact, the EVS question allows respondents to express the intention to engage in forms of political protest; the ISSP question has four possible answers;²¹ and the ESS questions only ask whether or not the

¹⁷ The idea that 'unofficial strikes' and 'occupying buildings or factories' are actions close to violence is not shared by the author, although this reduced scale is worth testing.

¹⁸ The two data sets are multiply imputed. See footnote 8. Results are shown in the Appendix.

¹⁹ In the ISSP, Belgium is Flanders and Germany is West Germany. This figure does not take item ordering into account, as it would not be correct to calculate mean scores for ordinal scales.

²⁰ The results are reported in the 'Appendix' section.

²¹ The question is worded as follows: 'Here are some different forms of political and social actions that people can take. Please indicate, for each one, whether you have done any of these things in the past; whether you have done it in the more distant past; whether you have not done it but might do it; or have

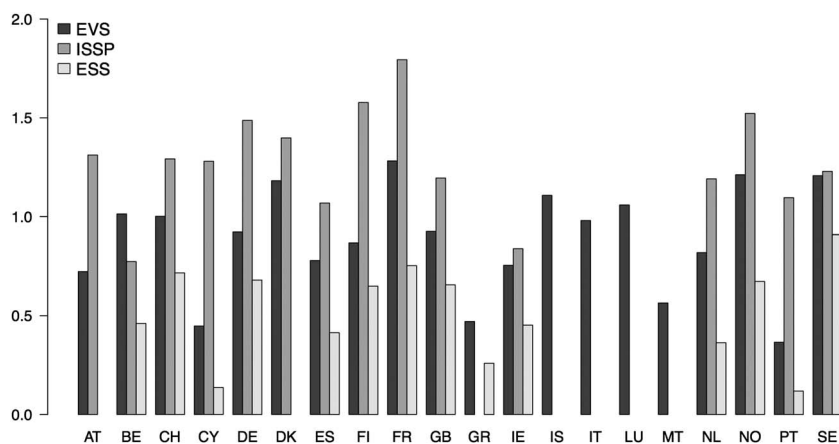


Figure 3 Mean scores of the political protest index using three items and the EVS, ISSP, and ESS data for the available countries. EVS = European Values Study; ISSP = International Social Survey Programme; ESS = European Social Survey

respondent has engaged in the actions in the last 12 months.²² Besides, it should be underlined that the three- and five-item scales built using the EVS data have similar mean scores, as they are highly correlated.²³ This is due to the hierarchical property of the items. In fact, those who answered positively to ‘joining unofficial strikes’ and ‘occupying buildings or factories’ also answered the other three items positively.

Conclusion

This article analyzed the concept of political protest and provided an empirical strategy, first, to measure its underlying latent trait and, second, to assess its cross-national equivalence in a set of 20 Western European countries. As argued, despite the tradition of concept formation and analysis within the field of political science (Sartori *et al.*, 1975; Sartori, 1984), as also demonstrated by recent publications (Goertz, 2006; Collier and Gerring, 2008; Mair, 2008), the importance of testing measurement instruments in comparative research is less prominent in the literature (Van Deth, 1998b; Jackman, 2008; Ariely and Davidov, 2012; Stegmüller, 2011), even though it is a fundamental stage of empirical research ensuring that analyses carried out are unbiased (Jacoby, 1999).

not it and would never, under any circumstances, do it: a) Signed a petition; b) Boycotted, or deliberately bought, certain products for political, ethical or environmental reasons; c) Took part in a demonstration’. The first two categories are collapsed into one and the last two into another.

²² Question wording is: ‘There are different ways of trying to improve things in [country] or help prevent things from going wrong. During the last 12 months, have you done any of the following? Have you: a) signed a petition? b) boycotted certain products c) taken part in a lawful public demonstration? Yes; No.’

²³ Kendall’s rank correlation (τ) between the two scales is above 0.90 in each country.

The article thus assessed the cross-national measurement equivalence of the political protest scale. First, according to the existing literature, the concept of political protest was outlined, distinguishing it from conventional political participation. Barnes and Kaase (1979) created the distinction between conventional and unconventional political participation, arguing that the latter addresses not only political institutions but also private subjects using more intense forms of political action. Accordingly, political participation can be seen as a ladder made up of several rungs of intensity, ranging from legal conventional participation, such as voting or campaigning, to unconventional participation activities, such as demonstrations, boycotts, or occupations of buildings (Dalton, 1988, 2008). Unconventional political participation includes direct political protest activities that are not mediated by institutions (Della Porta and Diani, 2006; Tilly and Tarrow, 2006), with different degrees of *'legality'*, that is, their conformity to legal norms relevant for a given type of behavior, and their *legitimacy*, that is, the extent to which a given population at a given point in time approves of or disapproves them' (Kaase and Marsh, 1979a: 45). Since political protest can be conceptualized as an ordinal continuum, MSA was used to assess its unidimensionality, the homogeneity of the items capturing the underlying latent trait and its cross-national equivalence. Techniques belonging to CTT were avoided, and the article focused on a model of ordinal unidimensional measurement more suited to dichotomous items (Van Schuur, 2003).

It was shown that when using the items included in the EVS (2011) the number of comparable Western European countries varies depending on what we intend to measure. If the goal is simply to verify that the scale has a sufficient degree of homogeneity, the scale may be applied to 19 out of 20 countries. If, instead, each step of the scale must have the same meaning, only 10 cases out of 20 may be used.

It was also shown that using the four- or three-item scales maximizes the comparability of the concept of political protest, since they make the ordinal measure fully comparable in 16 countries, albeit while reducing its intension, as fewer items are employed in the scale. This demonstrates that measurement equivalence should be an important concern for comparative researchers because analyses depend on how we choose measures.

The article aimed to contribute to the literature by providing a systematic study about the cross-national equivalence of the protest scale, which was absent in the literature, although this measure is widely used. It was argued that the assessment of cross-national measurement equivalence is a very relevant component of comparative analysis. The growing possibilities that international surveys provide to include larger numbers of countries (Norris, 2009) require that the comparative researcher tests measurement instruments, in particular given the popularity of statistical methods, such as hierarchical models, suited for quantitative comparative research (Steenbergen and Jones, 2002; Gelman and Hill, 2006). Therefore, assessing the equivalence of measurement instruments ensures that we use measures with the same construct, that is, the concept is measurable with the same set of items, across populations. It follows that the measurement instrument

thus has the same meaning in different contexts. Further, and most importantly, it increases the probability that the inferences drawn are correct. If we do not test measurement instruments, we cannot be sure whether the relations between dependent and independent variables are the product of real phenomena or just chance.

Acknowledgments

I would like to thank the Editors and the anonymous reviewers for providing valuable feedback and comments to previous versions of the article.

References

- Adcock, R. and D. Collier (2001), 'Measurement validity: a shared standard for qualitative and quantitative research', *American Political Science Review* 95(3): 529–546.
- Ariely, G. and E. Davidov (2012). 'Assessment of measurement equivalence with cross-national and longitudinal surveys in political science', *European Political Science* 11(3): 363–377.
- Barnes, S.H. and M. Kaase (eds) (1979), *Political Action: Mass Participation in Five Western Democracies*, Beverly Hills: Sage.
- Benson, M. and T.R. Rochon (2004), 'Interpersonal trust and the magnitude of protest: a micro and macro level approach', *Comparative Political Studies* 37(4): 435–457.
- Bollen, K.A. (1989), *Structural Equations with Latent Variables*, New York: John Wiley.
- Byrne, B.M. (2008), 'Testing for multigroup equivalence of a measuring instrument: a walk through the process', *Psicothema* 20(4): 872–882.
- Collier, D. and J. Gerring (eds) (2008), *Concepts and Method in Social Science. The Tradition of Giovanni Sartori*, London: Routledge.
- Dalton, R.J. (1988), *Citizen Politics in Western Democracies. Public Opinion and Political Parties in the United States, Great Britain, West Germany, and France*, Chatham: Chatham House.
- (2004), *Democratic Challenges, Democratic Choices: The Erosion of Political Support in Advanced Industrial Societies*, Oxford: Oxford University Press.
- (2008), *Citizen Politics: Public Opinion and Political Parties in Advanced Industrial Democracies*, Chatham: Chatham House.
- Dalton, R.J. and A. Van Sickle 2005. 'The roots of political protest: a contextual analysis of protest behavior', Presented at the Annual Meeting of the International Studies Association, March 2005, Honolulu, HI.
- Dalton, R.J., A. Van Sickle and S. Weldon (2010), 'The individual-institutional nexus of protest behaviour', *British Journal of Political Science* 40(1): 51–73.
- Della Porta, D. (ed.) (2007), *The Global Justice Movement. Cross-National and Transnational Perspectives*, Boulder-London: Paradigm Publishers.
- Della Porta, D., M. Andretta, L. Mosca and H. Reiter (2006), *Globalization from Below. Transnational Activists and Protest Networks*, Minneapolis: University of Minnesota Press.
- Della Porta, D. and M. Diani (2006), *Social Movements: An Introduction*, London: Wiley-Blackwell.
- European Social Survey (ESS) (2011), *European Social Survey Cumulative File, ESS 1–4*, Bergen: Norwegian Social Science Data Services.
- European Values Study (EVS) (2011), *European Values Study 1981–2008, Longitudinal Data File*, GESIS Data Archive, Cologne, Germany, ZA4804 Data File Version 1.0.0 (2011-04-30). doi:10.4232/1.4804.
- Gelman, A. and J. Hill (2006), *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge: Cambridge University Press.
- Gerring, J. (1999), 'What makes a concept good? A criterial framework for understanding concept formation in the social sciences', *Polity* 31(3): 357–393.
- Goertz, G. (2006), *Social Science Concepts: A User's Guide*, Princeton: Princeton University Press.
- Guttman, L. (1945), 'A basis for analyzing test-retest reliability', *Psychometrika* 10(4): 255–282.

- Horn, J.L. and J.J. McArdle (1992), 'A practical and theoretical guide to measurement invariance in aging research', *Experimental Aging Research* 18(3): 117–144.
- Hui, C.H. and H.C. Triandis (1985), 'Measurement in cross-cultural psychology: a review and comparison of strategies', *Journal of Cross Cultural Psychology* 16(2): 131–152.
- Inglehart, R. (1990), *Culture Shift in Advanced Industrial Society*, Princeton: Princeton University Press.
- (1997), *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*, Princeton: Princeton University Press.
- Inglehart, R. and C. Welzel (2005), *Modernization, Cultural Change and Democracy*, New York: Cambridge University Press.
- International Social Survey Programme (ISSP) (2007), ZA3950: *International Social Survey Programme 2004: Citizenship (ISSP 2004)*, Universitaet zu Koeln: Zentralarchiv fuer Empirische Sozialforschung.
- Jackman, S. (2008), 'Measurement', in J. Box-Steffensmeier, H.E. Brady and D. Collier (eds), *The Oxford Handbook of Political Methodology*, Oxford: Oxford University Press, pp. 119–151.
- Jacoby, W.G. (1999), 'Levels of measurement and political research: an optimistic view', *American Journal of Political Science* 43(1): 271–301.
- Jennings, K.M., J.W. Van Deth, S.H. Barnes, D. Fuchs, F.J. Heunks, R. Inglehart, M. Kaase, H.D. Klingemann and J.J.A. Thomassen (eds) (1989), *Continuities in Political Action: A Longitudinal Study of Political Orientations in Three Western Democracies*, Berlin and New York: De Gruyter.
- Kaase, M. (1989), 'Appendix C. The cumulateness and dimensionality of the participation scales', in K.M. Jennings, J.W. Van Deth, S.H. Barnes, D. Fuchs, F.J. Heunks, R. Inglehart, M. Kaase, H.-D. Klingemann and J.J.A. Thomassen (eds), *Continuities in Political Action: A Longitudinal Study of Political Orientations in Three Western Democracies*, Berlin and New York: De Gruyter, pp. 393–396.
- Kaase, M. and A. Marsh (1979a), 'Political action. A theoretical perspective', in S.H. Barnes and M. Kaase (eds), *Political Action: Mass Participation in Five Western Democracies*, Beverly Hills: Sage, pp. 27–56.
- (1979b), 'Political action repertory. Changes over time and a new typology', in S.H. Barnes and M. Kaase (eds), *Political Action: Mass Participation in Five Western Democracies*, Beverly Hills: Sage, pp. 137–166.
- Kerman, J. and A. Gelman (2007), 'Manipulating and summarizing posterior simulation using random variable objects', *Statistics and Computing* 17(3): 235–244.
- King, G., R.O. Keohane and S. Verba (1994), *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton: Princeton University Press.
- King, G., C.J.L. Murray, J.A. Salomon and A. Tandon (2004), 'Enhancing the validity and cross-cultural comparability of measurement in survey research', *American Political Science Review* 98(1): 191–207.
- King, G., J. Honacker, J. Anne and S. Kenneth (2001), 'Analyzing incomplete political science data: an alternative algorithm for multiple imputation', *American Political Science Review* 95(1): 49–69.
- Mair, P. (2008), 'Concepts and concepts formation', in D. Della Porta and M. Keating (eds), *Approaches and Methodologies in the Social Sciences*, Cambridge: Cambridge University Press, pp. 177–197.
- Marsh, A. and M. Kaase (1979), 'Measuring political action', in S.H. Barnes and M. Kaase (eds), *Political Action: Mass Participation in Five Western Democracies*, Beverly Hills: Sage, pp. 57–97.
- Meyer, D.S. and S. Tarrow (1998), 'A movement society: contentious politics for a new century', in D.S. Meyer and S. Tarrow (eds), *The Social Movement Society: Contentious Politics for a New Century*, Lanham: Rowman & Littlefield, pp. 1–28.
- Milbrath, L.W. (1965), *Political Participation: How and Why Do People Get Involved in Politics*, Chicago: Rand McNally College Publishing Company.
- Mokken, R.J. (1971), *A Theory and Procedure of Scale Analysis*, Berlin: De Gruyter.
- Morales, L. (2009), *Joining Political Organizations*, Colchester: ECPR Press.
- Norris, P. (2002), *Democratic Phoenix: Reinventing Political Activism*, Cambridge: Cambridge University Press.
- (2009), 'The globalization of comparative public opinion research', in T. Landman and N. Robinson (eds), *The Sage Handbook of Comparative Politics*, London: Sage, pp. 522–540.
- Parry, G., G. Moyser and N. Day (1992), *Political Participation and Democracy in Britain*, Cambridge: Cambridge University Press.

- Raghunathan, T.E., J.M. Lepkowski, J. Van Hoewyk and P. Solenberger (2001), 'A multivariate technique for multiply imputing missing values using a sequence of regression models', *Survey Methodology* 27(1): 85–95.
- Rubin, D.B. (1987), *Multiple Imputation for Nonresponse in Surveys*, New York: Wiley.
- Rucht, D. (2007), 'The spread of protest politics', in R.J. Dalton and H.-D. Klingemann (eds), *The Oxford Handbook of Political Behavior*, Oxford: Oxford University Press, pp. 708–723.
- Sartori, G. (1970), 'Concept misformation in comparative politics', *American Political Science Review* 64(4): 1033–1053.
- (1984), 'Guidelines for concept analysis', in G. Sartori (ed.), *Social Science Concepts: A Systematic Analysis*, Beverly Hills: Sage, pp. 15–85.
- Sartori, G., F. Riggs and H. Teune (eds) (1975), *The Tower of Babel*, University of Pittsburgh: International Studies Association, Occasional Paper no. 6.
- Sijtsma, K. (1998), 'Methodology review: Non-parametric IRT approaches to the analysis of dichotomous item scores', *Applied Psychological Measurement* 22(1): 3–31.
- Sijtsma, K. and I.W. Molenaar (1987), 'Reliability of test scores in nonparametric item response theory', *Psychometrika* 52(1): 79–97.
- (2002), *Introduction to Nonparametric Item Response Theory*, Thousand Oaks: Sage.
- Sijtsma, K., P. Debets and I.W. Molenaar (1990), 'Mokken Scale Analysis for polychotomous items: theory, a computer program and an empirical application', *Quality and Quantity* 24(2): 173–188.
- Sijtsma, K., R.R. Meijer and A.L. Van der Ark (2011), 'Mokken Scale Analysis as time goes by: an update for scaling practitioners', *Personality and Individual Differences* 50(1): 31–37.
- Steenbergen, M.R. and B.S. Jones (2002), 'Modeling multilevel data structures', *American Journal of Political Science* 46(1): 218–237.
- Steenkamp, J.-B.E.M. and H. Baumgartner (1998), 'Assessing measurement invariance in cross-national consumer research', *Journal of Consumer Research* 25(1): 78–107.
- Stegmuller, D. (2011), 'Apples and oranges? The problem of equivalence in comparative research', *Political Analysis* 19(4): 471–487.
- Teorell, J., M. Torcal and J.R. Montero (2007), 'Political participation. Mapping the terrain', in J.W. Van Deth, J.R. Montero and A. Westholm (eds.), *Citizenship and Involvement in European Democracies. A Comparative Analysis*, London and New York: Routledge, pp. 334–357.
- Tilly, C. and S. Tarrow (2006), *Contentious Politics*, Lanham: Paragon Press.
- Van de Vijver, F.J.R. and N.K. Tanzer (2004), 'Bias and equivalence in cross-cultural assessment: an overview', *European Review of Applied Psychology* 54(2): 119–135.
- Van der Ark, L.A. (2007), 'Mokken Scale Analysis in R', *Journal of Statistical Software* 20(11): 1–19.
- Van der Meer, T.W.G., M.T. Grotenhuis and P.L.H. Scheepers (2009), 'Three types of voluntary associations in comparative perspective: the importance of studying associational involvement through a typology of associations in 21 European countries', *Journal of Civil Society* 5(3): 227–241.
- Van Deth, J.W. (1986), 'A note on measuring political participation in comparative research', *Quality and Quantity* 20(2): 261–272.
- (ed.) (1998a), *Comparative Politics: The Problem of Equivalence*, London and New York: Routledge.
- (1998b), 'Equivalence in comparative research', in J.W. Van Deth (ed.), *Comparative Politics: The Problem of Equivalence*, London and New York: Routledge, pp. 1–19.
- Van Deth, J.W., J.R. Montero and A. Westholm (eds) (2007), *Citizenship and Involvement in European Democracies. A Comparative Analysis*, London and New York: Routledge.
- Van Schuur, W.H. (2003), 'Mokken Scale Analysis: between the Guttman Scale and parametric Item Response Theory', *Political Analysis* 11(2): 139–163.
- Vandenberg, R.J. and C.E. Lance (2000), 'A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research', *Organizational Research Methods* 3(1): 4–70.
- Verba, S. and N.H. Nie (1972), *Participation in America: Political Democracy and Social Equality*, New York: Harper and Row.
- Welzel, C. and F. Deutsch (2012), 'Emancipative values and non-violent protest: the importance of 'ecological' effects', *British Journal of Political Science* 42(2): 465–479.

Appendix

Table A1. *H* coefficients, standard errors, *Z* statistics, items' *H* coefficients for the political protest scale constructed with three items for each country and for the pooled sample using the EVS data

	<i>H</i>	SE	<i>Z</i>	Items <i>H</i>			Ordering		
				P	B	D			
AT	0.70	0.03	21.7	0.80	0.68	0.65	P	D	B
BE	0.65	0.03	19.3	0.67	0.66	0.61	P	D	B
CH	0.62	0.03	17.8	0.67	0.61	0.60	P	D	B
CY	0.59	0.04	18.3	0.59	0.65	0.58	D	P	B
DE	0.66	0.02	22.7	0.68	0.67	0.63	P	D	B
DK	0.54	0.03	18.0	0.56	0.57	0.50	P	D	B
ES	0.54	0.03	23.0	0.50	0.82	0.50	P	D	B
FI	0.62	0.03	19.5	0.74	0.58	0.56	P	B	D
FR	0.75	0.02	23.8	0.75	0.81	0.71	P	D	B
GB	0.70	0.03	21.8	0.89	0.64	0.64	P	D	B
GR	0.61	0.03	26.9	0.58	0.68	0.60	D	P	B
IE	0.73	0.03	18.1	0.87	0.66	0.67	P	D	B
IS	0.58	0.03	19.6	0.64	0.55	0.56	P	B	D
IT	0.64	0.02	24.4	0.61	0.78	0.60	P	D	B
LU	0.64	0.03	21.8	0.65	0.68	0.60	P	D	B
MT	0.71	0.03	27.2	0.70	0.78	0.67	P	D	B
NL	0.67	0.03	22.8	0.74	0.66	0.62	P	D	B
NO	0.61	0.03	16.2	0.74	0.57	0.57	P	D	B
PT	0.66	0.04	25.7	0.69	0.62	0.65	P	D	B
SE	0.57	0.03	17.1	0.72	0.57	0.51	P	B	D
Pooled	0.63	0.01	94.6	0.68	0.65	0.59	P	B	D

EVS = European Values Study; AT = Austria; BE = Belgium; CH = Switzerland; CY = Cyprus; DE = Germany; DK = Denmark; ES = Spain; FI = Finland; FR = France; GB = Great Britain; GR = Greece; IE = Ireland; IS = Iceland; IT = Italy; LU = Luxembourg; MT = Malta, NL = the Netherlands; NO = Norway; PT = Portugal; SE = Sweden; P = signing a petition; B = joining in boycotts; D = attending lawful demonstrations.

Table A2. *H* coefficients, standard errors, *Z* statistics, items' *H* coefficients for the political protest scale constructed with three items for each country and for the pooled sample using the ISSP data

	<i>H</i>	SE	<i>Z</i>	Items <i>H</i>			Ordering		
				P	B	D			
AT	0.53	0.03	15.8	0.56	0.47	0.60	P	B	D
BE-FLA	0.42	0.03	14.5	0.58	0.37	0.36	P	D	B
CH	0.39	0.03	13.3	0.35	0.36	0.54	P	B	D
CY	0.47	0.05	11.7	0.52	0.43	0.46	D	P	B
DE-W	0.52	0.03	18.8	0.57	0.49	0.50	P	B	D
DK	0.42	0.03	14.9	0.47	0.38	0.43	P	B	D
ES	0.55	0.02	30.4	0.54	0.61	0.51	D	P	B
FI	0.52	0.03	20.9	0.51	0.49	0.60	P	B	D
FR	0.40	0.02	18.1	0.51	0.35	0.38	P	D	B
GB	0.70	0.04	14.8	0.73	0.68	0.72	P	B	D
IE	0.56	0.03	20.5	0.69	0.52	0.48	P	B	D
NL	0.56	0.02	24.2	0.67	0.49	0.53	P	B	D
NO	0.50	0.03	18.6	0.55	0.46	0.49	P	B	D
PT	0.48	0.02	28.1	0.50	0.51	0.43	P	D	B
SE	0.48	0.03	15.2	0.54	0.44	0.49	P	B	D
Pooled	0.38	0.01	66.1	0.49	0.38	0.30	P	B	D

ISSP = International Social Survey Programme; AT = Austria; BE-FLA = Belgian Flanders; CH = Switzerland; CY = Cyprus; DE-W = Germany-West; DK = Denmark; ES = Spain; FI = Finland; FR = France; GB = Great Britain; IE = Ireland; NL = the Netherlands; NO = Norway; PT = Portugal; SE = Sweden; P = signing a petition; B = joining in boycotts; D = attending lawful demonstrations.

Table A3. *H* coefficients, standard errors, *Z* statistics, items' *H* coefficients for the political protest scale constructed with three items for each country and for the pooled sample using the ESS data

	<i>H</i>	SE	<i>Z</i>	Items <i>H</i>			Ordering		
				P	B	D			
BE	0.44	0.03	18.4	0.52	0.42	0.38	P	B	D
CH	0.47	0.03	19.0	0.48	0.51	0.43	P	B	D
CY	0.36	0.06	16.5	0.34	0.43	0.34	B	P	D
DE	0.34	0.02	20.0	0.33	0.45	0.29	P	B	D
DK	0.26	0.03	10.9	0.29	0.32	0.19	P	B	D
ES	0.45	0.02	30.4	0.47	0.46	0.41	P	D	B
FI	0.32	0.02	14.9	0.30	0.64	0.29	P	B	D
FR	0.40	0.02	22.8	0.44	0.44	0.34	P	B	D
GB	0.48	0.03	19.7	0.50	0.57	0.43	P	B	D
GR	0.48	0.04	23.5	0.51	0.42	0.51	B	D	P
IE	0.38	0.03	18.9	0.45	0.34	0.34	P	B	D

Table A3. (Continued)

	<i>H</i>	SE	<i>Z</i>	Items <i>H</i>			Ordering		
				<i>P</i>	<i>B</i>	<i>D</i>			
NL	0.39	0.04	13.7	0.44	0.38	0.35	<i>P</i>	<i>B</i>	<i>D</i>
NO	0.42	0.03	15.1	0.45	0.49	0.35	<i>P</i>	<i>B</i>	<i>D</i>
PT	0.39	0.04	27.3	0.47	0.38	0.31	<i>P</i>	<i>D</i>	<i>B</i>
SE	0.38	0.03	14.7	0.38	0.51	0.33	<i>P</i>	<i>B</i>	<i>D</i>
Pooled	0.40	0.01	74.7	0.43	0.41	0.35	<i>P</i>	<i>B</i>	<i>D</i>

ESS = European Social Survey; BE = Belgian; CH = Switzerland; CY = Cyprus;
 DE = Germany; DK = Denmark; ES = Spain; FI = Finland; FR = France; GB = Great
 Britain; GR = Greece; IE = Ireland; NL = the Netherlands; NO = Norway; PT = Portugal;
 SE = Sweden; *P* = signing a petition; *B* = joining in boycotts; *D* = attending lawful
 demonstrations.