Measurement of attitudes towards homosexuality: testing measurement invariance across European countries

Jan Gromadzki^{1,2}

Published online: 25 April 2019
© The Author(s) 2019

Abstract

The statement regarding the approval of homosexuality from the European Social Survey has been used in several studies of attitudes towards homosexuality. I use data from the 8th wave of the ESS to evaluate the validity of the approval measure. The 8th wave included an additional statement regarding the support for LGBT adoption rights, which allows me to test the relevance and the measurement invariance of the approval measure. I find that the approval measure is strongly correlated with the measure of the support for LGBT rights, but there exist fundamental differences in the understanding of the approval statement between European countries.

Keywords Construct validity · Measurement invariance · Homosexuality · Public attitudes · European Social Survey

1 Introduction

Over the last 2 decades, several researchers attempted to compare attitudes towards homosexuality between countries. It is crucial to ensure the validity of measures used in the research, because using invalid constructs may lead to biased results. In this study, I focus on the measure of the approval of homosexuality based on a statement included in the European Social Survey. I contribute to the literature by testing the measurement validity of the approval measure, which is a prerequisite for any meaningful comparisons between groups (Davidov et al. 2014; van de Schoot et al. 2012; Vandenberg and Lance 2000). Although several studies have used the approval measure to analyse the attitudes towards homosexuality in Europe, it had been impossible to test the measurement validity before the release of data from the most recent wave of the ESS. I evaluate the validity of the approval measure in two ways. First, I test how precisely the approval measure predicts the support for LGBT adoption rights. Second, I check whether the interpretation of the approval statement differs between countries.

✉ Jan Gromadzki
jgromad@sgh.waw.pl

¹ Warsaw School of Economics, aleja Niepodległości 162, 02-554 Warsaw, Poland

² Institute for Structural Research, Warsaw, Poland

The European Social Survey has been frequently used to analyse the attitudes towards homosexual persons. The first wave of the ESS was conducted in 2002 and it has been repeated every 2 years since then. 36 countries participated in at least one of the waves of the survey. 15 countries participated in all of the 8 waves. The following ESS question has been used to measure attitudes towards homosexuality: ‘Using this card, please say to what extent you agree or disagree with each of the following statements: Gay men and lesbians should be free to live their own life as they wish’. The response is measured on the 5-point Likert scale (‘Agree strongly’, ‘Agree’, ‘Neither agree nor disagree’, ‘Disagree’, ‘Disagree strongly’). The statement was included in all of the waves of the ESS. Several studies found a significant relationship between approval of homosexuality and gender, age, education level, religiosity, denomination, institutional characteristics (van den van den Akker et al. 2013; Donaldson et al. 2017; Hooghe and Meeusen 2013; Kuntz et al. 2014; Röder and Lubbers 2015).

All of the mentioned authors use the same measure of attitudes towards homosexual persons. They build a measure of the disapproval of homosexuality where 1 is ‘Agree strongly’ and 5 is ‘Disagree strongly’ (Kuntz et al. 2014) revert the scale, and van den van den Akker et al. (2013) shift it to 0–4).

The validity of a measure in cross-country comparison studies requires satisfying two fundamental conditions. First of all, the construct should measure what it claims it measures. Second, the construct should be measured in the same way in each country (Davidov et al. 2014). A limitation, often mentioned by the authors is that the attitudes towards homosexuality were measured with only one item. It is not clear how closely the response to the approval statement is related to the support for LGBT rights. Moreover, the individual interpretation of the questionnaire statement may differ both within and between countries. Freedom to live life “as they wish” may be interpreted as being as free as heterosexual persons. It may be also understood as being free within the current legal system. Finally, some respondents may interpret it as a freedom without limits. A statement regarding the support for LGBT adoption rights, which was added to the ESS questionnaire in 2016 allows me to test the validity of the approval measure. I analyse how closely the approval measure is related to the measure of the support for LGBT adoption rights and I test for the measurement invariance of the approval measure. The remainder of the article is structured as follows. In the next section, I introduce the definitions of the measures of attitudes towards homosexual persons and describe the dataset. Next, I discuss methods used to examine the equivalence of the measures and measurement invariance of the approval measure. Then, I present the results of the tests. The last section concludes.

2 Data

I use individual data of the 8th wave of the ESS, which was conducted in 23 countries in 2016. The appropriate sample weights are applied.¹ The measure of attitudes towards homosexual persons is the approval of homosexuality as defined by Kuntz et al. (2014). It equals 5 when the response is ‘Agree strongly’, 4—‘Agree’, 3—‘Neither agree nor

¹ I apply design weights for examining the data within a country and I combine design weights with population size weights for the cross-country analysis.

disagree', 2—'Disagree', 1—'Disagree strongly'. At the country level, it is weighted average of the individual responses.

The questionnaire of the 8th wave of the ESS included a new statement regarding attitudes towards homosexuality. The statement is 'Gay male and lesbian couples should have the same rights to adopt children as straight couples'. I build a measure of the support for LGBT adoption rights in the same way as in the case of the approval statement (using 5-point scale, where 5 is 'Agree strongly' and 1 is 'Disagree strongly').

Figure 1 shows the country scores on the measure of approval of homosexuality and the support for LGBT rights in 2016. There is a clear positive relationship between both measures. The highest levels of both the approval and the support were observed in Iceland and the lowest in Russia. In all of the countries, the support for LGBT adoption rights was lower than the approval of homosexuality. There are two countries that clearly stand out. Israel and Poland are characterised by similar levels of the approval measure. However, they differ significantly in terms of support for LGBT adoption rights. In Poland, the level of support for LGBT rights is similar to that of Russia and Lithuania. By contrast, the level of support for LGBT rights in Israel is relatively high—it places Israel among countries such as Ireland, Germany and Great Britain. This indicates that letting homosexual persons live their life 'as they wish' may mean something substantially different in Poland than in Israel.

3 Measurement validity

3.1 Equivalence testing

In the first step, I test the similarity of the measure of the approval of homosexual persons and the measure of the support for LGBT adoption rights. Techniques to examine construct equivalence are summarised by de Vijver (1997). I use two coefficients of similarity that are not influenced by multiplications: Tucker's coefficient and Pearson's correlation coefficient.

Some rules of thumb are used to evaluate the values of the coefficients. A Tucker's coefficient value in the range of 0.85–0.94 is a sign of a fair similarity of the factors, while a value higher than 0.95 means that both factors can be considered equal (Lorenzo-Seva and ten Berge 2008). According to Evans (1996) the Pearson's correlation coefficient of 0.60–0.79 is a sign of the strong correlation, and a value of the coefficient greater than 0.80 can be interpreted as a very strong correlation.

Apart from the two coefficients, I use the instrumental variable approach to test the similarity of the measures. The advantage of the instrumental variable approach is that it can be used both to analyse the similarity of two constructs, and to test for measurement invariance. The instrumental variable approach is a common method used to eliminate bias caused by endogenous variables (Wooldridge 2002). The simple OLS estimation of an equation with endogenous variables leads to a bias. The instrumental variable method may solve the endogeneity issue in the case of reverse causality or omitted variable bias. An instrument is a variable that is strongly correlated with the endogenous variable of interest and it is exogenous in the initial equation. The application of the instrumental variable

Fig. 1 Approval of homosexuality and support for LGBT rights in Europe (2016). Note: Figure shows average approval of homosexuality and support for LGBT rights in 23 countries. Approval of homosexuality and support for LGBT rights are measured on 5-point scale, where 1 is 'Disagree strongly' and 5 is 'Agree strongly'. The statements are 'Gay men and lesbians should be free to live their own life as they wish' and 'Gay male and lesbian couples should have the same rights to adopt children as straight couples', respectively. Data: European Social Survey

approach to the equivalence testing requires only the first condition to be satisfied—the instrument must be strong, i.e. it must accurately predict the variable of interest. This condition is often called the relevance condition.

In my analysis, I treat the approval of homosexual persons as an instrument for the support for LGBT adoption rights. At the country level, I test the relevance of the instrument by estimating the following equation

$$s_i = \alpha_0 + \beta a_i + \epsilon_i \quad (1)$$

where s_i is the measure of support for LGBT adoption rights and a_i is the measure of approval of homosexual persons. The equation is estimated by the standard OLS estimator. The F-test is used to test the relevance of the instrument. Again, the rule of thumb is applied—an instrument is not weak if F-statistic exceeds 10 (Stock et al. 2002). I estimate the same equation for the approval measure at the individual level.

3.2 Cross-country measurement invariance

After establishing the similarity between constructs, it is crucial to test the measurement invariance. It is not clear whether the interpretation of the approval statement is the same between countries. The example of Israel and Poland shows that countries with similar levels of the approval measure may differ greatly in the support for LGBT rights. Measurement invariance of a latent variable measured with multiple items is usually assessed using CFA models (Davidov et al. 2014; van de Schoot et al. 2013). Before 2016, however, there was only one statement regarding attitudes towards homosexuality in the ESS questionnaire. Therefore, it is impossible to use any measure that would be based on multiple items to analyse the attitudes towards homosexuality before 2016. It would be convenient if we could treat the approval measure as an instrument for the support for LGBT rights, which is based on a more direct statement than the approval measure. Apart from testing for the relevance of the instrument I have to ensure that the approval measure is invariant across countries. The presence of substantial differences in the interpretation of the statement between countries would bias cross-country comparisons and analyses at the individual level.

The additional statement, which was included in the 8th wave of the ESS, allows me to identify differences in the interpretation of the approval statement. The statement regarding LGBT adoption rights is much more precise and I assume that its interpretation is the same in all countries. I estimate an extended version of the model (1) to formally verify the significance of differences in regression coefficients between countries. Adding interaction terms between group dummy variables and an independent variable is a widely used method to compare differences in regression coefficients (see for example studies on moderators of intervention effectiveness by Gardner et al. 2009; Wang and Ware 2013). The model is given by

$$s_i = \alpha_0 + \beta a_i + \sum_k \gamma_k c_i^k + \sum_k \lambda_k (a_i * c_i^k) + \epsilon_i \quad (2)$$

where s_i is the measure of support for LGBT adoption rights, a_i is the measure of approval of homosexual persons, c_i^k is the set of dummy country variables, and $a_i * c_i^k$ is a set of interactions of dummy country variables and the approval measure. The hypothesis of the lack of significant differences in regression coefficients β between countries requires coefficients λ_k to be zero. The parameter λ_k significantly different than zero would imply that the

Table 1 Coefficients of similarity

	Correlation	Tucker's	F-statistic
Country level	0.899	0.993	957.37
Individual level	0.650	0.931	151908.82

Results represent coefficients of similarity of the approval measure and the measure of the support for LGBT adoption rights. At the individual level both constructs are measured on the 5-point scale and at the country level the measures are averages of individual responses.

Data: European Social Survey

marginal effect of the approval of homosexuality on the support of LGBT adoption rights differs between countries. Such a situation would mean the violation of metric invariance. Therefore, I estimate a restricted version of the model with $\lambda_k = 0$ and the unrestricted version. I perform ordered logit regressions and use the likelihood ratio test to examine the joint significance of parameters λ_k .

The parameters λ_k and γ_k from the proposed measurement test based on the instrumental variable approach are closely related to the measurement invariance testing based on the standard CFA methods. Parameters λ_k refer to testing for the equality of factor loadings in the CFA-based measurement invariance test. Parameters γ_k refer to testing for the equality of factor intercepts in the the CFA-based measurement invariance test (see Steenkamp and Baumgartner 1998). The highest level of measurement invariance is the scalar invariance. In this case, both λ_k and γ_k have to be zero. γ_k is a parameter that captures changes in regression intercepts α_0 between countries. The values of the γ_k significantly greater than zero would mean that the instrument estimates are biased upwards or downwards by the same value in a country. There can be a situation that metric invariance would be established, but the scalar invariance would not. It would mean that although the interpretation of the statement is the same between countries, there exist some country-specific characteristics that shift the measurement in parallel upwards or downwards. The changes in the instrument over time could be still interpreted, but differences in aggregated constructs between countries would not reflect the real differences (Steenkamp and Baumgartner 1998). To test the scalar invariance I estimate the restricted version of the model with $\lambda_k = 0$ and $\gamma_k = 0$ and compare the fit of this model to the fit of the unrestricted model.

4 Results

At the country level, the measure of the approval of homosexuality and the support for LGBT adoption rights exhibit high degree of similarity (correlation coefficient of 0.9 and Tucker's coefficient of 0.99, see Table 1). The country-level measure of approval is also a strong instrument for the support for LGBT rights. At the individual level, the Tucker's coefficient is also high, and the correlation is strong (0.65). High F-test statistics show that the measure of approval of homosexuality is a strong instrument for the support of LGBT adoption rights both at the individual and at the country level.

Table 2 Measurement invariance: likelihood ratio test

Variable	Metric invariance		Scalar invariance	
	Test statistic	<i>p</i> value	Test statistic	<i>p</i> value
Full sample	926.0	0.0000	6290.0	0.0000
Limited sample	14.0	0.1730	782.0	0.0000

Results represent test statistics and *p* values of likelihood ratio test for full sample of 23 countries and a limited sample of 11 countries. The unrestricted model includes the approval measure, country dummy variables and interactions of country dummy variables with the approval measure. In case of metric invariance the restricted model includes the approval measure and country dummy variables. In case of scalar invariance the restricted model includes only the approval measure.

Data: European Social Survey

The results of measurement invariance test using the full sample of 23 countries are presented in the first row of Table 2. I reject the null hypothesis about scalar invariance of the approval measure and I also reject the hypothesis about the metric invariance of the measure. It means that in at least one country the interpretation of the approval statement is different than in other countries. Thus, it is impossible to establish measurement invariance of the constructs using the full sample of 23 countries.

Nevertheless, removing countries with significant differences in the interpretation of the statement from the sample may result in establishing the measurement invariance. I reduce the sample to 11 countries, in which coefficients on interaction terms were not significant (Austria, Finland, France, Great Britain, Israel, Iceland, Norway, Portugal, Slovenia, Sweden and Switzerland, see the detailed regression results in Table 3 in Appendix). I perform measurement invariance test to check whether approval measure is invariant in the limited sample (see the second row of Table 2). I establish metric invariance of the approval measure, and I reject the hypothesis about the scalar invariance of the approval measure using the reduced sample. Country-specific characteristics shift the support for LGBT adoption right parallel upwards or downwards, but there are no significant differences in interpreting the approval statement between 11 selected countries.

5 Conclusions

The European Social Survey has been frequently used by researchers, who analysed differences in attitudes towards homosexuality between countries. Using the approval statement, which was the only statement regarding homosexuality before 2016, requires two strong assumptions. Responses to the statement have to reflect the attitudes towards homosexuality, and the interpretation of the statement should not differ between countries and across time. Before 2016, the attitudes towards homosexuality could be measured with only one item. I use a statement regarding the support for LGBT adoption rights, which was added in the 8th wave of the ESS. It can be assumed that the interpretation of the support statement was the same in all countries. I treat the approval measure as an instrument for the

support for LGBT rights and verify whether there were any significant differences in interpreting the approval statement between countries.

Both at the country level and the individual level the approval measure is a strong instrument for the support for LGBT adoption rights. Unfortunately, it is impossible to establish the measurement invariance for the full sample of 23 countries. A very effective way to establish measurement invariance is to restrict the sample to observations from countries with no significant differences in the interpretation of the statement. In a sample with observations from 11 countries I establish metric invariance of the approval measure. The study shows that there exist fundamental differences in the interpretation of the statement regarding the approval of homosexuality between European countries. Letting gays and lesbians live their life “as they wish” does not mean the same in all countries. Therefore, the comparative studies of attitudes towards homosexuality should probably restrict the sample to respondents from countries that exhibit a similar understanding of the statement. I show that the limited sample of 11 countries exhibits the high degree of measurement invariance. It can definitely be extended by other countries, but the selection of the countries should be verified with the measurement invariance test. It is still not clear whether the interpretation of the approval statement is stable over time. The future research should address also this dimension of measurement invariance. From 2016 onwards, it is possible to build a measure of attitudes towards homosexuality based on three items. Testing for measurement invariance of such construct using the standard multigroup CFA methods seems particularly necessary in the light of the results of this study.

The methods presented in this article can be applied to test the measurement invariance of constructs measured with only one item. The analysis of measurement invariance relies on the key assumption that the interpretation of the statement regarding the support for LGBT rights does not vary across countries. I argue that the support statement is very precise, so its interpretation should be the same in all countries. Unfortunately, it is impossible to formally test the assumption. Therefore, it is crucial to justify the choice of the statement used in the analysis.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

See Table 3.

Table 3 Measurement invariance: ordered logit regression results

	Full sample			Limited sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Approval	1.24***	1.12***	1.28***	1.13***	1.21***	1.19***
Approval × BE			0.22**			
Approval × CH			− 0.08			− 0.08
Approval × CZ			− 0.42***			
Approval × DE			0.17**			
Approval × EE			− 0.34*			
Approval × ES			0.46***			
Approval × FI			0.09			0.08
Approval × FR			0.05			0.05
Approval × GB			0.07			0.06
Approval × HU			− 0.20**			
Approval × IE			− 0.27**			
Approval × IL			− 0.03			− 0.03
Approval × IS			0.51			0.51
Approval × IT			− 0.46***			
Approval × LT			− 0.49***			
Approval × NL			0.25**			
Approval × NO			0.01			− 0.00
Approval × PL			− 0.57***			
Approval × PT			− 0.19*			− 0.18*
Approval × RU			− 0.36***			
Approval × SE			0.06			0.05
Approval × SI			− 0.13			− 0.11
Country dummy	No	Yes	Yes	No	Yes	Yes
Log likelihood	− 64,191	− 61,509	− 61,046	− 21,798	− 21,414	− 21,407
N	42,390	42,390	42,390	18,059	18,059	18,059

Results represent coefficients obtained from ordered logit estimations of support for LGBT rights measured on the 5-point scale. Full sample includes 23 countries and the limited sample includes 11 countries.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Data: European Social Survey

References

- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., Billiet, J.: Measurement equivalence in cross-national research. *Ann. Rev. Sociol.* **40**, 55–75 (2014). <https://doi.org/10.1146/annurev-soc-071913-043137>
- de Vijver, F.V.: *Methods and Data Analysis of Comparative Research*. Allyn & Bacon, Boston (1997)

- Donaldson, C.D., Handren, L., Lac, A.: Applying multilevel modeling to understand individual and cross-cultural variations in attitudes toward homosexual people across 28 European countries. *J. Cross-Cult. Psychol.* **48**(1), 93–112 (2017). <https://doi.org/10.1177/0022022116672488>
- ESS: European social survey round 8 data. data file edition 2.0. In: NSD—Norwegian Centre for Research Data, Norway—Data Archive and distributor of ESS data for ESS ERIC (2016). <http://www.europeansocialsurvey.org/>
- Evans, J.D.: *Straightforward Statistics for the Behavioral Sciences*. Cole Publishing, Three Lakes (1996)
- Gardner, F., Connell, A., Trentacosta, C.J., Shaw, D.S., Dishion, T.J., Wilson, M.N.: Moderators of outcome in a brief family-centered intervention for preventing early problem behavior. *J. Consult. Clin. Psychol.* **77**(3), 543–553 (2009). <https://doi.org/10.1037/a0015622>
- Hooghe, M., Meeusen, C.: Is same-sex marriage legislation related to attitudes toward homosexuality? *Sex. Res. Soc. Policy* **10**(4), 258–268 (2013). <https://doi.org/10.1007/s13178-013-0125-6>
- Kuntz, A., Davidov, E., Schwartz, S.H., Schmidt, P.: Human values, legal regulation, and approval of homosexuality in Europe: a cross-country comparison. *Eur. J. Soc. Psychol.* **45**(1), 120–134 (2014). <https://doi.org/10.1002/ejsp.2068>
- Lorenzo-Seva, U., ten Berge, J.M.F.: Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodol. Eur. J. Res. Methods Behav. Soc. Sci.* **2**, 57–64 (2008). <https://doi.org/10.1027/1614-2241.2.2.57>
- Röder, A., Lubbers, M.: Attitudes towards homosexuality amongst recent Polish migrants in Western Europe: migrant selectivity and attitude change. *J. Ethnic Migr. Stud.* **41**(11), 1858–1884 (2015). <https://doi.org/10.1080/1369183X.2015.1023183>
- Steenkamp, J.-B.E.M., Baumgartner, H.: Assessing measurement invariance in cross-national consumer research. *J. Consum. Res.* **25**(1), 78–107 (1998). <https://doi.org/10.1086/209528>
- Stock, J.H., Wright, J.H., Yogo, M.: A survey of weak instruments and weak identification in generalized method of moments. *J. Bus. Econ. Stat.* **20**(4), 518–529 (2002). <https://doi.org/10.1198/07350102288618658>
- van de Schoot, R., Lugtig, P., Hox, J.: A checklist for testing measurement invariance. *Eur. J. Dev. Psychol.* **9**(4), 486–492 (2012). <https://doi.org/10.1080/17405629.2012.686740>
- van de Schoot, R., Kluytmans, A., Tummers, L., Lugtig, P., Hox, J., Muthén, B.: Facing off with Scylla and Charybdis: a comparison of scalar, partial, and the novel possibility of approximate measurement invariance. *Front. Psychol.* **4**, 770 (2013). <https://doi.org/10.3389/fpsyg.2013.00770>
- van den Akker, H., van der Ploeg, R., Scheepers, P.: Disapproval of homosexuality: comparative research on individual and national determinants of disapproval of homosexuality in 20 European countries. *Int. J. Public Opin. Res.* **25**(1), 64–86 (2013). <https://doi.org/10.1080/00224499.2013.858306>
- Vandenberg, R.J., Lance, C.E.: A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organ. Res. Methods* **3**(1), 4–70 (2000). <https://doi.org/10.1177/109442810031002>
- Wang, R., Ware, J.H.: Detecting moderator effects using subgroup analyses. *Prevent. Sci. Off. J. Soc. Prev. Res.* **14**(2), 111–120 (2013). <https://doi.org/10.1007/s1121-011-0221-x>
- Wooldridge, J.M.: *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA (2002)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.