Check for updates

Article

American Behavioral Scientist
2022, Vol. 66(4) 412–433
© 2021 SAGE Publications
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/00027642211021623
journals.sagepub.com/home/abs
SAGE

Survey Data Quality in Analyzing Harmonized Indicators of Protest Behavior: A Survey Data Recycling Approach

Kazimierz M. Slomczynski[1,2], Irina Tomescu-Dubrow[1,2], and Ilona Wysmulek[1]

Abstract
This article proposes a new approach to analyze protest participation measured in surveys of uneven quality. Because single international survey projects cover only a fraction of the world's nations in specific periods, researchers increasingly turn to ex-post harmonization of different survey data sets not a priori designed as comparable. However, very few scholars systematically examine the impact of the survey data quality on substantive results. We argue that the variation in source data, especially deviations from standards of survey documentation, data processing, and computer files—proposed by methodologists of Total Survey Error, Survey Quality Monitoring, and Fitness for Intended Use—is important for analyzing protest behavior. In particular, we apply the Survey Data Recycling framework to investigate the extent to which indicators of attending demonstrations and signing petitions in 1,184 national survey projects are associated with measures of data quality, controlling for variability in the questionnaire items. We demonstrate that the null hypothesis of no impact of measures of survey quality on indicators of protest participation must be rejected. Measures of survey documentation, data processing, and computer records, taken together, explain over 5% of the intersurvey variance in the proportions of the populations attending demonstrations or signing petitions.

[1]Polish Academy of Sciences, Institute of Philosophy and Sociology, Warsaw, Mazowieckie, Poland
[2]Department of Sociology, The Ohio State University, Columbus, OH, USA

Corresponding Author:
Kazimierz M. Slomczynski, Institute of Philosophy and Sociology, Polish Academy of Science, Nowy Swiat 72, Warsaw, Mazowieckie 00-330, Poland.
Email: slomczynski.1@osu.edu

Keywords
survey data quality, data harmonization, political protest, demonstrations, signing petitions

Introduction

Repeated cross-national public opinion surveys appended with theoretically relevant country-year variables, such as democracy level, economic development, and levels of economic inequality, among others, remain a widespread and powerful tool to study political protest in comparative perspective (Claassen, 2019; Dalton et al., 2010; Dubrow et al., 2008; Foa & Mounk, 2016, 2017; Norris, 2017; Solt, 2008). Nonetheless, individual projects, even purportedly worldwide ones like the World Value Survey (WVS), provide only partial coverage of countries over time. To broaden the scope of comparison without fielding more surveys, social scientists increasingly turn to reprocessing information from different cross-national data sets into new, integrated databases. In doing so, researchers rely on ex-post harmonization methods (Granda & Blasczyk, 2016; Granda et al., 2010; Günther, 2003; Slomczynski et al., 2016) to strengthen the comparability of answers from respondents interviewed about the same issue but in projects whose methodology can vary considerably (Blasius & Thiessen, 2012; Oleksiyenko et al., 2018; Thiessen & Blasius, 2016).

In this context, there is growing research on how to deal with differences in properties of the survey items measuring the same concept (e.g., Gummer & Roßmann, 2013; Saris & Gallhofer, 2014). However, there is hardly any discussion of whether, and to what extent, ex-post harmonization should take into account intersurvey variability stemming from an uneven implementation of standards for data collection established in the specialized methodological literature (for an exception, see Slomczynski & Tomescu-Dubrow, 2018; also, Tomescu-Dubrow & Slomczynski, 2016).

Our article proposes a new approach in which the impact of survey data quality on substantive variables, such as attending demonstrations and signing petitions, is explicitly tested.[1] This approach stems from the Survey Data Recycling (SDR) project, a large-scale ex-post harmonization effort to create a multicountry multiyear database for cross-national research on political participation, social capital, and subjective well-being (asc.ohio-state.edu/dataharmonization). Drawing from the frameworks of Total Survey Error (e.g., Biemer, 2010, 2016; Groves, 1989; Groves & Lyberg, 2010; Lyberg & Weisberg, 2016; Smith, 2011; Weisberg, 2005), Survey Quality Monitoring (e.g., Billiet et al., 2004; Lyberg & Biemer, 2008; Lyberg & Stukel, 2010; Morganstein & Marker, 1997) and Fitness for Intended Use (e.g., Biemer & Lyberg, 2003; Juran & Gryna, 1980), SDR develops methodological indicators for quality of (a) data as reflected in the survey documentation, (b) survey data processing, and (c) computer survey records. We present a subset of the SDR database v.1.0-55 source files comprising 1,184 national surveys from 19 international projects containing harmonized measures of attending demonstrations and signing petitions, together with variables describing intersurvey differences in the survey question wording. Then, we test whether the measures of the survey quality are related to indicators of protest behavior. Finally, in the concluding part of the article, we discuss the different strategies that could be used for such analysis of survey data that explicitly take into account aspects of data quality.

The Need for Survey Harmonization Control in Studying Protest Behavior

Democratization research consistently theorizes about the relevance that protesting against the establishment has for social processes, such as democratic consolidation or democratic backsliding (e.g., Claassen, 2019; Foa & Mounk, 2016; Mechkova et al., 2017; Mounk, 2018; Norris, 2017; Waldner & Lust, 2018). Yet, worldwide there is marked variation in protest across nations (e.g., Claassen, 2019; Dalton et al., 2010). The questions of why this is so, and with what consequences for the fate of democracy, has been at the heart of comparative social science research since *Political Action: An Eight Nation Study, 1973-1976* (Marsh & Kaase, 1979). Evidently, social structure places constraints on individual resources and decision making (e.g., Dubrow, 2015; Marien et al., 2010; Solt, 2008; Vrablikova, 2013), and on opportunity conditions for political protest (e.g., Kitschelt, 1986; Koopmans & Statham, 2000; Meyer, 2004). At the same time, however, agency is a major determinant of social phenomena (e.g., Elder et al., 2003) and contributes to the transformation of social structure (e.g., Archer, 2003; Rosenstone & Hansen, 1993; Tomescu-Dubrow & Slomczynski, 2014). As Quaranta (2016) notes,

> Political protest is part of contemporary democracies and it has often worked as a vehicle of change. Until the 1960s protest activities were considered irrational, dangerous, disruptive [. . .] Nowadays, the activities related to protest politics are spread in democracies. (p. 2)

Yet, protest occurs also in nondemocracies and partial democracies, without necessarily leading to greater democratization, but rather to repressive responses by the state and even conflict escalation.

For studying how individual determinants and features of the social context interact to enhance or suppress protest, data sets with information at both the person- and the country-year levels that cover regimes in different democratization and economic development stages are the main prerequisite. In the absence of cross-national panel survey data on political protest, international survey projects with a repeated cross-sectional design, such as the WVS, International Social Survey Programme, European Social Survey, the Eurobarometer and its regional counterparts, among others, constitute a powerful tool to analyze peoples' reported protest behavior, together with macrolevel variables that survey data are frequently appended with (e.g., Claassen, 2019; Dubrow et al., 2008; Solt, 2008; Stockemer, 2014). However, any single project, even WVS and International Social Survey Programme, is limited in how many countries—from historically marginalized regions, especially—it covers and for how many time points. This situation also holds for regional coverage, including Europe (Kołczyńska, 2014; Slomczynski & Tomescu-Dubrow, 2006; Wysmułek, 2018).

To broaden the scope of comparative research without conducting new cross-national surveys—a notoriously demanding endeavor financially and in terms of infrastructure needs—political participation studies (but also demography, epidemiology, and health studies), increasingly turn to reprocessing data from the trove of already available survey projects (Burkhauser & Lillard, 2005; Fortier et al., 2017; Slomczynski et al., 2016; Sobek et al., 2007). Researchers pool information on the same concept from sources not a priori designed as comparative, transform it to increase the comparability of answers from respondents interviewed in different populations and periods, and create a new data set whose coverage—of individuals, countries, and years—is much wider than that of its constituents. The literature refers to these methods as ex-post harmonization, to the original survey data sets and variables as source data sets and source variables, respectively, and to the harmonized, common, measures produced from the source variables, as target variables (Ehling & Rendtel, 2006; Granda & Blasczyk, 2016; Granda et al., 2010; Günther, 2003; Minkel, 2004). Recent examples of cross-national analyses of micro- and macrolevel determinants of protest participation using harmonized data sets include Slomczynski et al. (2016) and Kołczyńska (2020).

As the interdisciplinary field of survey data harmonization grows (e.g., Dubrow & Tomescu-Dubrow, 2016), so do efforts to strengthen and integrate the methodology of ex-post harmonization (e.g., Granda & Blasczyk, 2016; Kołczyńska & Schoene, 2018; Oleksiyenko et al., 2018; Slomczynski et al., 2016; Slomczynski & Tomescu-Dubrow, 2018; Wolf et al., 2016). An important area of development deals with methodological variability among international survey projects that differences in the properties of source variables measuring the same concept, introduce. Our solution, which we implement in this article, builds on insights from classic and recent comparability and equivalence studies (e.g., Billiet, 2003; Cheung, 2008; Cieciuch et al., 2016; Jowell, 1998; Kenneth & Yeates, 2001; Matsumoto & van de Vijver, 2010; Medina et al., 2009; Przeworski & Teune, 1972). The core of this solution relies on evaluating the target variables through *controls of specific harmonization procedures*. In short, important intersurvey differences in properties of the source items about question wording are identified and stored as indicators in the harmonized data set; and their impact on the target variable is assessed empirically (Kołczyńska & Slomczynski, 2018; Slomczynski & Tomescu-Dubrow, 2018). We discuss the harmonized political protest measures and their corresponding harmonization controls in the article's section on data on protest behavior.

The Need for Assessment of Survey Data Quality in Studies of Protest Behavior

An impressive body of research, subsumed under the quality assessment frameworks of Total Survey Error (e.g., Biemer, 2010; Groves, 1989; Groves & Lyberg, 2010; Lyberg & Weisberg, 2016; Smith, 2011), Survey Quality Monitoring (e.g., Lyberg &

Biemer, 2008; Lyberg & Stukel, 2010; Morganstein & Marker, 1997), and Fitness for Intended Use (e.g., Biemer & Lyberg, 2003; Juran & Gryna, 1980), has designed an extensive methodology for the survey production process. Data providers can use a wide array of methods during the stages of survey design, data collection, and data processing, to increase the accuracy of survey estimates and the survey's responsiveness to users' needs, that is, to produce good quality surveys. However, this accumulated knowledge is, so far, largely neglected in the harmonization of survey data stemming from international projects. As Thiessen and Blasius (2016) note, rarely does anybody examine deficiencies in publicly released cross-national survey data. Notable exceptions are Blasius and Thiessen (2012), Thiessen and Blasius (2016), and the SDR project (asc.ohio-state.edu/dataharmonization).

Drawing on the frameworks of Total Survey Error, Survey Quality Monitoring, and Fitness for Intended Use, we define survey quality in terms of *standards* available in the specialized methodological literature and promoted by professional organizations of public opinion and marketing research. The standards considered in this article pertain to three dimensions of *the survey life cycle*: (a) documentation on data collection, (b) data processing, and (c) computer files.[2] They range from reporting on the type of sample, through the preparation of labels for variables, to eliminating duplicated records. Adherence to the standards of this kind indicates high survey quality; methodological errors and biases that stem from deviations from these standards indicate low survey quality.

The definition of survey quality by adherence to certain standards has two important limitations. First, some "smart" fabrication of the data and the metadata (documentation), could declare the methodological standards' fulfillment as a deliberate part of fraudulent activity. Consequently, we could assign a "high quality" score to the surveys with cooked data or "enhanced" documentation; this could result in a *false positive*. Second, producers of a survey may strictly follow all high-quality data collecting requirements but neglect to report on fulfillments of these requirements in survey documentation. Thus, the real data may be of better quality than established by our analysis; this could produce *false negatives*.

Although we do not know of any method for the systematic detection of this kind of false positives and false negatives, we argue that they are relatively infrequent. On the one hand, it is unlikely that those involved in fraudulent data preparation could anticipate a full set of standards included in our approach. On the other hand, it is not in the interest of data producers to underreport the fulfillment of methodological requirements that they actually applied. Thus, we assume that a possible sporadic occurrence of some false positives and false negatives should not introduce substantial distortion to our analysis.

How to Study the Impact of Survey Data Quality?

The target variable (T), participation in demonstrations, or signing petitions, is a function of source variable(s) (S). The form of the relationship between T and S, $T = f(S)$ is determined by the recoding procedure that varies across survey projects and survey waves since source variables are not the same. We use two types of control variables, H and Q, in a linear manner:

$$T = b0 + b1\,H + b2\,Q + b3\,X + e$$

where H stands for item metadata (harmonization controls) accounting for variability in the questionnaire items; Q stands for data quality controls of survey documentation, data processing, and computer records; X refers to other substantive variable(s); and e denotes a random error.

Equation 1 is presented in a simplified way, without appropriate subscripts denoting different levels of measurement of Q and H. Some data quality controls are or could be defined on the level of national surveys or even entire international survey projects, while others are at the respondent level. Thus, the equation joining T with H and Q must include subscripts reflecting the complex structure of the data, where the survey itself represents a separate dimension in the multilevel data structure. Individuals are nested in the national surveys of projects' waves to which data quality indicators can be attached. Also, it must be noted that Equation 1 does not include possible interactions of Q and H with X, either. Bearing this in mind, we note the following:

If the random error is negligible and the estimated parameters are of the form $b1 = b2 = b3 = 0$, then, $b0 = f(S)$ corresponds to the assumed relationship between T and S. However, if $b1 \neq 0$ or $b2 \neq 0$, some intervention is needed to correct for biases and errors in T due to influences of H and Q. The intervention solutions we propose in the last section of this article range from eliminating "bad data" to the weighting of data, giving high multipliers for "good data," and low multipliers for "bad data." Between these extremes, there is a different solution. We can partial out the effects of Q and H on the relationships between T and other substantive variables X, using multiple regression. This procedure is analogous to computing partial correlation or partial covariance of T and X, controlling for Q and H.

Survey Data on Protest Behavior

The SDR database v.1.0 (Slomczynski, Jenkins, et al., 2017), constructed via ex-post harmonization of cross-national survey data in the SDR project, is particularly suitable for studying individual and contextual determinants of peoples' propensity to engage in political protest. The database provides harmonized measures of two types of conventional protest (Jenkins & Form, 2005). The first, *attending demonstrations*, is a good example of collectivistic behavior. The second form of protest, *signing a petition*, features both collectivistic and individualistic aspects: preparing a petition calls for cooperation and signing a petition is a personal act that can occur without public engagement (Dubrow et al., 2008). The distribution of these variables across international projects discussed in this article is provided in Table 1.

The questions on actual participation in demonstrations appear in 19 international survey projects (1,148 national surveys). This yields data on 1,560,943 respondents

Table 1. Selected International Survey Projects With Questions on Attending Demonstrations and Signing Petitions, 1966-2013.

Abbreviation	Survey project	Time span[a]	Attending demonstrations: Number of surveys	Attending demonstrations: Number of respondents	Signing petitions: Number of surveys	Signing petitions: Number of respondents
ABS	Asian Barometer Survey	2001-2011	28	41,842	19	24,527
AFB	Afrobarometer	1999-2009	65	96,671	—	—
AMB	Americas Barometer	2004-2012	74	119,291	—	—
ARB	Arab Barometer	2006-2012	14	17,269	14	17,269
ASES	Asia Europe Survey	2000	17	17,251	17	17,251
CNEP	Comparative National Elections Project[b]	2004-2006	7	12,339	—	—
EB	Eurobarometer[c]	1983-2004	64	58,222	42	38,686
EQLS	European Quality of Life Survey	2007-2012	65	79,270	34	43,636
ESS	European Social Survey	2002-2013	146	281,496	146	281,496
EVS	European Values Study	1981-2009	127	165,399	127	165,399
ISJP	International Social Justice Project	1991-1997	19	23,769	19	23,769
ISSP	International Social Survey Programme[c]	1995-2006	64	85,345	73	97,749
LB	Latinobarometro	1995-2008	204	229,345	110	126,426
LITS	Life in Transition Survey	2006; 2010	62	65,866	62	65,866
PA2	Political Action II[d]	1979-1981	3	4,057	3	4,057
PA8NS	Political Action; An Eight Nation Study	1973-1976	8	12,588	8	12,588
PPE7N	Political Participation in Seven Nations	1966-1971	4	8,693	1	1,743
VPCP	Values in Postcommunist Europe[b]	1993	5	4,723	5	4,723
WVS	World Value Survey[e]	1981-2009	172	237,499	172	239,217
Total		1966-2013	1,148[f]	1,560,935	852[f]	1,164,402

[a]EB, EVS, ISJP, ISSP, and LITS survey projects have no continuity in time span; the gap(s) between survey waves in these projects is larger than 2 years. [b]Only postelection samples. [c]Selected waves (editions). [d]Only cross-sectional samples. [e]WVS/4 Morocco has two samples (country sets) for the same survey wave. Both WVS/4 (2001) Morocco Data Set 1 and Data Set 2 have questions on demonstrations and petitions. [f]Combined number of surveys is 1,184.

interviewed at various times from 1966 to 2013 in 136 countries and territories. In 332 national surveys (out of 1,148), there is a question about participation in demonstrations but no question about signing petitions. Thus, in the SDR database, most national surveys (816) include questions on both forms of protests. The question only on signing petitions appears in 36 national surveys, making the total number of surveys with a question on this form of protest equal to 852, stemming from 16 projects and covering 1,164,402 respondents.

While all source variables of interest for this article aim to capture individuals' *actual* experience of attending demonstrations, the ways of asking the question differ among surveys. Here are two examples:

> 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, whether you have actually done any of these things, whether you might do it or would never, under any circumstances, do it. Attending lawful demonstrations. Response categories: 1—Have done, 2—Might do, 3—Would never do, −1—Don't know, −2—No answer. (Source: WVS/1-5).

> 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 . . . taken part in a lawful public demonstration? Response categories: 1—Yes, 2—No, 7—Refusal, 8—Don't know, 9—No answer. (Source: European Social Survey/1-5).

Based on the source variable(s), we created the target variable *attending demonstrations*, which deals with whether the respondent took part in demonstrations. This variable takes the value: 1 = yes, means that respondent took part in a demonstration (in source variables, the answers were as follows: *yes*, or *once*, or *more than once*, or *several times*, or *have done*), or 0 = no, means that respondent did not declare having participated in demonstrations (in source variables, the answers were *no*, or *I have never participated*, or *never done but might do*, etc.).[3]

Variation in questions on signing petitions is similar to that about attending demonstrations. For both variables, we created the same control variables, capturing the variation in time span and the breath of the question (Kołczyńska & Slomczynski, 2018). The variables of H-type are

1. *Time span* during which respondent might have taken action, coded in terms of years from 1 to 10, with 11 for "ever."
2. *Extended meaning*, whether the question extends its scope by adding different words to "demonstrations" (e.g., street protest) or "petitions," (e.g., protest letter) coded 1 if yes, and 0 otherwise.

In addition, in the case of demonstrations, we include two other control variables:

3. *Illegality*, whether the wording of the question is atypical by making reference to the illegality of the demonstration, coded 1 if yes, and 0 otherwise.

4. *Set of questions*, whether a given survey contained a set of specific questions about the demonstration, coded 1 if yes, and 0 if only one question.

In the case of petitions, we include one other control variable:

5. *Issue*, whether the question wording specifies a petition cause (e.g., environmental issues), coded 1 if yes and 0 otherwise.

Table 2 provides basic information on the distribution and effects of harmonization controls. Survey questions about protest behavior are asked in various time frames, most commonly in terms of "last year" (34.1% of surveys), and 10 years and "ever" which constitute 61.3% of surveys. Most surveys follow the standard in which attending a demonstration is not lumped with other forms of protest, illegality is not mentioned, and the single question is applied.

In the case of signing petitions, the distribution of time frames offered to respondents is similar to attending demonstrations: last year (30.2% of surveys), and 10 years or "ever" (61.5% of surveys). The vast majority of surveys used standard wording restricted to signing petitions and not specifying what issue the petition addressed.

Harmonization controls of questionnaire items of both attending demonstrations and signing petitions account for relatively small and insignificant differences in answers' proportions, with one exception: time span. For time span, the difference between categories "1 year" and "10 years and ever" equals 9% for attending demonstrations and 11% for signing petitions, which in both cases is statistically significant. However, we do not claim that all other harmonization controls could be ignored. They could affect protest behavior variables in combination with the survey quality measures. We will demonstrate that this a case.

The Impact of Survey Data Quality on Harmonized Indicators of Protest Behavior

Measurement of Survey Data Quality. To systematically evaluate the quality of all the national surveys and construct metadata of survey quality, we applied rules developed within the SDR analytic framework. SDR identifies three relevant dimensions of survey quality, conceptualized as *Q*-type variables:

1. The quality of surveys as reflected in the general survey documentation. We analyzed documents containing information on how the survey was conducted, such as study descriptions or technical reports. This material was taken from archiving institutions and project websites. Our evaluation yielded metadata about the following elements of survey implementation: details of the sample; response rate; control of the quality of the questionnaire translation; questionnaire pretesting; and fieldwork control.[4] Panel I in Table 3 provides information on these indicators.

Table 2. Distribution and Effects of Harmonization Controls for International Survey Projects for Two Target Variables: Attending Demonstrations and Signing Petitions.

	Distribution of harmonization controls			Effects of harmonization controls		
	Number of surveys	Number of respondents	Percentage of respondents	Average proportion of respondents declaring involvement in a given activity	Standard error	Significance of the difference (z test)
A. Attending demonstrations						
Time span						
1 Year	335	530,808	34.0	0.08	0.015	$z = -3.90$[a],
2-4 Years	66	86,287	5.5	0.10	0.037	$p < .001$
5-8 Years	43	56,657	3.6	0.18	0.059	
10 Years and ever	704	887,183	56.8	0.17	0.014	
Total	1,148	1,560,935	100.00	0.14	0.010	
Extended meaning						
1-Yes	457	591,444	37.9	0.13	0.016	$z = -0.49$,
0-No	691	969,491	62.1	0.14	0.013	$p = .631$
Total	1,148	1,560,935	100.0	0.14	0.010	
Illegality						
1-Yes	127	138,828	8.9	0.14	0.031	$z = 0$,
0-No	1,021	1,422,107	91.1	0.14	0.011	$p = 1.0$
Total	1,148	1,560,935	100.0	0.14	0.010	
Set of questions						
1-Yes	163	181,413	11.6	0.16	0.031	$z = 1.04$,
0-No	985	1,379,522	88.4	0.13	0.011	$p = .298$
Total	1,148	1,560,935	100.00	0.14	0.010	
B. Signing petitions						
Time span						
1 Year	209	352,140	30.2	0.19	0.027	$z = -3.06$[a],
2-4 Years	38	46,519	4.0	0.16	0.059	$p = .002$
5-8 Years	37	49,421	4.2	0.14	0.057	
10 Years and ever	568	716,322	61.5	0.30	0.019	
Total	852	1,164,402	100.0	0.25	0.015	
Extended meaning						
1-Yes	67	85,432	7.3	0.16	0.045	$z = -1.80$,
0-No	785	1,078,970	92.7	0.26	0.016	$p = .073$
Total	852	1,164,402	100.0	0.25	0.015	
Issue						
1-Yes	34	45,199	3.9	0.14	0.059	$z = -1.57$,
0-No	818	1,119,203	96.1	0.26	0.015	$p = .116$
Total	852	1,164,402	100.0	0.25	0.015	

[a]The difference between proportions for (a) 1 year and (b) 10 years and ever.

Table 3. Measurement of Quality of Data for International Survey Projects: Indicators and Summary Indexes.

Indicators	Codes
I. General Survey Documentation: How were the data collected?	
(a) Does the survey documentation describe accurately the sampling procedure?	0 = No, 1 = Yes
(b) Does the survey documentation provide information on the response rate?	0 = No, 1 = Yes
(c) Was the quality of questionnaire translation checked in some systematic way?	0 = No, 1 = Yes
(d) Is there any evidence that the instrument (questionnaire) was pretested?	0 = No, 1 = Yes
(e) Is there evidence of fieldwork control?	0 = No, 1 = Yes
Summary index of Items from a to e, Q_1	From 0 to 5
II. Specific Data Description: How were the data defined?	
(a) Is the description of the data consistently provided in English?	No = 0, Yes = 1
(b) Are all variable values within the legitimate range?	No = 0, Yes = 1
(c) Are variable labels the same in the codebook and in the data file?	No = 0, Yes = 1
(d) Are value labels the same in the codebook and in the data file?	No = 0, Yes = 1
(e) Are different types of missing values explained?	No = 0, Yes = 1
Summary measure of Items a-e standardized for number of questions/ variables included in the analysis, Q_2	From 0 to 1.33
III. Computer Data Records: Were the data formally correct?	
(a) Are the data free of nonunique records?	No = 0, Yes = 1
(b) Are design/poststratification weights free of formal errors?	No = 0, Yes = 1
(c) Is the proportion of missing values for gender and age below 5%?	No = 0, Yes = 1
(d) Do survey cases (respondents) have unique identification numbers (IDs)?	No = 0, Yes = 1
Summary index of items (a-d), Q_3	From 0 to 4

2. The degree of consistency between the official description of the data with the actual data records in the computer files. A typology of processing errors used in the SDR framework and applied in this article includes the following: illegitimate variable values, misleading variable values, contradictory variable values, variable values discrepancy, and lack of variable value labels. These were checked for demographic variables (gender, age, year of birth), social background variables (educational level, years of schooling), behavioral variables (attending demonstrations), and attitudinal variables (trust in parliament). Indicators that refer to inconsistencies between the description of the data and the data in the computer files are provided in Panel II of Table 3.
3. The quality of the survey computer records. We constructed control variables for nonunique records (Slomczynski et al., 2017), departures from formal properties of sampling weights (Zieliński et al., 2018), the amount of missing

Table 4. Distribution Properties of Summary Measures of Survey Quality for 1,184 National Surveys Included in the Analysis[a].

Distribution properties	Quality of documentation (Q_1)	Quality of data processing (Q_2)	Quality of computer files (Q_3)
Min. and max values, max = no errors	[0, 5]	[0, 1]	[0, 4]
Proportion of surveys with no errors	0.25	0.40	0.47
Mean value	2.53	0.85	2.37
Standard deviation	1.77	0.22	0.67
Kurtosis	1.52	6.24	3.03
Skewness	0.30	−1.89	−0.73

[a]See Table 1.

data on basic sociodemographic variables, and errors in respondents IDs (this includes missing or duplicated IDs); see Panel III in Table 3.

Table 3 also includes summary indicators that we use in this article. In the case of the quality of documentation (Panel I, Items a-e) and computer files (Panel III, Items a-d), we use indexes that are sums of binary variables. Methodological literature does not rate the importance of our indicators by saying, for example, that field control is more (or less) important than pretesting or that duplicates are less (or more) harmful than excess missing data. Nor does it seem that building a latent construct is viable since these indicators are weakly correlated with each other. Thus, we rely on the unweighted summation of the items. The simplest interpretation of the indexes for survey documentation and the quality of computer files is that they provide counts of whether the established standards have been followed. For the first index, the minimal value is 0 (none of the standards adhered to) and the maximal value 5 (all standards have been adhered to). For the index of quality of survey files, the corresponding values are 0 and 4.

The index pertaining to data processing (Panel II, Items a-e, was computed differently). Coding of the material involved counting the number of errors, which were summed up and divided by the number of variables. Since we measure the quality of surveys positively, the scale was reversed and standardized, with a minimal value of 0 (none of the standards adhered) and a maximal value of 1 (all standards have been adhered to).

For all indexes we use in this article, the higher their value, the higher the quality of the survey. Table 4 provides distribution properties. The highest value is achieved by 47% in the case of computer files, then, 39% in the case of data processing. In the case of the quality of survey documentation, only 25% of national surveys belong to this "no error" category. All variables have moderate standard deviations, at least as compared with the means. The shape of the three variables' distribution differs: the quality of documentation is platykurtic with positive skew, while the quality of data processing is leptokurtic with negative skew; kurtosis of the quality of computer files corresponds to that of normal distribution, but skewness is on a negative side. Although not ideal, these variables can be used for examining linear relationships among them and with other variables.

Our measures of the quality of survey documentation are moderately positively correlated with the quality of data processing and computer files, $r_{Q1,Q2} = 0.114$, $r_{Q1,Q3} = 0.266$, respectively, with $p < .01$. The correlation between the quality of data processing and computer files is close to zero ($r_{Q2,Q3} = -0.014$). Thus, each variable constitutes a distinguishable dimension and could have an independent impact on substantive variables, such as on specific forms of protest behavior. This observation is important for the next section since the effect of quality variables can be analyzed separately and in a combined additive manner in multiple regression.

Impact of Quality Variables on Aggregate Measures of Protest Behavior. The question that we pose is the following: What proportion of variance in the estimates of the frequencies of attending demonstrations and signing petitions could be attributed to the quality control variables? Could we reject the null hypothesis positing that control variables have no impact on estimates of reported protest behavior?

Table 5 shows the extent to which the indexes of data quality are associated with frequencies of attending demonstrations and signing petitions with and without harmonization controls. We present both correlations and standardized coefficients (beta weights) from a multiple regression.

Consider the percentage of persons who declare attending demonstrations in the past year (Table 5, Panel A1). The higher the quality of documentation (Q_1) and the higher the quality of data processing (Q_2), the lower percentage of persons who declare having attended demonstrations. Correlations and beta coefficients are not negligible in terms of their size and they are statistically significant at $p < .01$ level. The effect of the quality of computer files (Q_3) goes in the opposite direction with beta coefficient statistically significant. All these three variables (Q_1, Q_2, and Q_3) explain more than 10% of the variance in the dependent variable. Adding harmonization controls increases the relative impact of Q_1 and Q_3.

We do not have a good explanation for why the effects of the same data quality variables, Q_2 and Q_3, have opposite signs in the case of attending demonstrations in the past year as contrasted with 10 years or ever (Table 5, Panel A2 vs. A1). However, their effect is statistically significant. All three variables (Q_1, Q_2, Q_3) explain over 5% of the dependent variable.

The effects of data quality on reported frequencies of signing petitions (Table 5, Panels B1 and B2) are also statistically significant for both the past year and past 10 year and ever versions of the question. While the direction is likewise variable and displays no simple pattern, the conclusion is that studies of attending demonstration and signing petitions should take into account the quality of data. Definitely, we reject the null hypothesis stipulating no relationship between measures of survey quality and national survey estimates of protest behavior.

Table 5. Relationship Between Survey Quality and Aggregate Measures of Protest Behavior for National Surveys of International Projects.

	Without harmonization controls		With harmonization controls	
Indexes[a]	Correlation	Beta	Partial correlation	Beta
A.1. Percentage of persons who declare attending demonstrations in the past year				
Quality of documentation, Q_1	−0.175**	−0.166**	−0.123*	−0.279**
Quality of data processing, Q_2	−0.298**	−0.293**	−0.128*	−0.215**
Quality of computer file, Q_3	0.052	0.167**	0.109*	0.227**
R^2 Adjusted	—	.117	—	.157
Number of national surveys	335	335	335	335
A.2. Percentage of persons who declare attending demonstrations in past 10 years or ever				
Quality of documentation, Q_1	−0.004	0.019	0.002	0.022
Quality of data processing, Q_2	0.207**	0.190**	0.212**	0.197**
Quality of computer files, Q_3	−0.140**	−0.114**	−0.125**	−0.097*
R^2Adj	—	.051	—	.054
Number of national surveys	704	704	704	704
B.1. Percentage of persons who declare signing petition in the past year				
Quality of documentation, Q_1	0.157*	0.143*	0.186**	0.148*
Quality of data processing, Q_2	0.012	0.045	−0.075	−0.027
Quality of computer files, Q_3	−0.226**	−0.204**	−0.203**	−0.177*
R^2 Adjusted	—	.056	—	.064
Number of national surveys	209	209	209	209
B.2. Percentage of persons who declare signing petition in past 10 years or ever				
Quality of documentation, Q_1	0.005	0.046	—[b]	—[b]
Quality of data processing, Q_2	0.298**	0.336**	—[b]	—[b]
Quality of computer files, Q_3	0.003	0.102*	—[b]	—[b]
R^2Adj	—	.096	—[b]	—[b]
Number of national surveys	568	568	568	568

[a]Quality of documentation from 0 = lowest quality to 5 = highest quality; quality of data processing from 0 = lowest quality to 1.33 = highest quality; quality of computer files from 0 = lowest quality to 4 = highest quality. [b]Not calculated due to too small variation in harmonization controls.
$^*p < .05$. $^{**}p < .01$.

Strategies of Including Information on Quality of Data Into Substantive Analyses

Researchers can use quality controls in substantive analyses in three ways. First, scholars could use the quality-control variables as "filters," that is, to select those data sets that best fit their research and data requirements. To illustrate this option, assume that you are interested in establishing a correlation between attending the demonstration in the past 10 years or longer and education measured by years of schooling. You could select the best national surveys according to data documentation, data processing, and data records. It appears that this correlation, averaged for all "good" surveys, is 0.168 (with standard deviation = 0.073. If the same correlation is computed for all surveys, its value is lower, $r = .135$; with standard deviation = 0.079, min = −0.116 and max = 0.380).

The second option is partialling out the effects of data quality controls, thus "recycling" rather than throwing out low-quality data. Consider a partialling out procedure in the context of simple regression. The procedure's basic idea is to remove some variables' effects and then run a simple regression that returns β of the interest. Assume, for example, that you are interested in the impact of the attending demonstration, denoted by X, on some measurable outcome Y. We know that X depends on the quality of surveys, measured by Q_1, Q_2, and Q_3. The strategy of partialling out of the effects of the quality controls, in the regression framework, is to estimate the equation $X = \alpha_0 + \alpha_1 Q_1 + \alpha_2 Q_2 + \alpha_3 Q_3 + e$, and calculate residuals $R(X)$, called here R. Regression of Y on R, $Y = \lambda_0 + \lambda_1 R + c$, provides a "pure" effect of X on Y.

From the Frisch–Waugh–Lovell theorem (Hayashi, 2000: 18-19), it follows that the resulting slope λ_1 is always equal to β_1 in the equation: $Y = \beta_0 + \beta_1 X + \beta_2 Q_1 + \beta_3 Q_2 + \beta_4 Q_3 + u$. Since β_1 could be obtained by a regression of Y on R, we deal with the variation unique to X, that is, after the effects of quality controls have been "partialled out." This is the essence of the technique.

An extension of this technique could be applied to a dependent variable by removing quality controls' effects. Thus, using the SDR data, scholars could construct variance-covariance matrices partialling out the effects of data quality controls and applying these matrices to further regression analysis. For example, when the effects of quality control variables are removed, the correlation between interest in politics and attending demonstrations during the past 10 years or ever is 0.189. Without such control, the result is 0.149. For a sample of aggregated data ($N = 505$), the difference of 0.040 could occur by chance in circa 50% of trials. That the correlation increases when the effects of quality control are removed supports the importance of this control.

The third option is to use data quality control variables to construct weights for survey quality. Researchers could weight source surveys according to their overall quality so that information from the high-quality surveys will carry more impact on the substantive results than "problematic" surveys. For example, the data of the highest quality could be weighted 1, the lowest quality 0.25, and in the middle 0.75. In this case, the correlation between interest in politics and attending demonstrations in the past 10 years or ever for combined surveys is higher than without weights (0.186 as compared with 0.149), as expected.

Among these options, the first one is most restrictive because it takes into account only part of the available data, eliminating the rest as unsuitable. In this option, the decision where to put the threshold is somewhat arbitrary. As used in the third option, weights also contain some arbitrariness. Generally, we advocate the second option since it "recycles" all data, and it is based on the statistical method of removing the unwanted effects. However, we also see good reasons for using this option after eliminating surveys with a large number of "bad cases," by which we mean excessive number of nonunique records. In another analysis, we discovered that 13 national surveys in the SDR1 data set, coming from a respectable international project, have more 10% duplicates (Slomczynski, Powałko, et al., 2017); they could be disregarded for further analyses.

Conclusion

Very few scholars have systematically examined the variation in the quality of source data, conceptualized in terms of deviations from standards established in the methodological literature. The SDR framework pays special attention to survey data quality and provides the indexes for the quality of documentation, data processing, and computer files. In this article, we applied the SDR framework to measure survey quality in major international projects and analyze its impact on estimates of attending demonstrations and signing petitions. We show that indicators of the quality of documentation, data processing, and computer files are associated with the estimates of declared protest behavior. We urge scientists who use survey data on protest behavior in cross-national analysis to pay attention to intersurvey methodological differences. Much like differences in survey question wording, these aspects need to be addressed in substantive analyses of these data. We suggest how the effects of survey quality measures on the variable of substantive interest can be minimized by partialling out their influence.

Further development of this line of research could focus on elaborating on the mechanism through which the survey quality measures impact estimates of political protest frequencies. For example, how does it happen that a large number of errors in survey documentation is significantly associated with overestimating of the proportion of people who attended demonstrations past year? One could speculate that lack of fieldwork control, reflected in survey documentation, relaxes standards of interviewing and prompts interviewers to seek positive answers as more valuable. An equally plausible explanation is that neglecting rigorous control of the questionnaire translation increases the chances of such item formulations that lead to the "prodemonstration" bias. However, we need different explanations when the relationship between survey quality measures and estimates of protest behavior goes in the opposite direction, as in the case of demonstrations in the past 10 years or ever. Discovering that the errors in survey documentation, data processing, and computer files are not randomly distributed with respect to estimates of protest behavior identifies the problem. Explanations of "why is so" require new research. In it, the separation of the causal effect from other effects should be carefully executed. Errors, or lack of them, are phenomena with their own determinants.

Authors' Note

The data analyzed in this article are available at Harvard Dataverse, a free data repository (https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/VWGF5Q).

Acknowledgments

We thank all research assistantse who took part in the data preparation, Marta Kołczyńska, Olena Oleksiyenko, Marcin Zieliński, and Przemek Powałko, in particular. We also thank anonymous reviewers and Craig Jenkins for their comments on earlier drafts.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This material is based on work supported by the National Science Foundation under Grant No. PTE Federal award 1738502; the data were gathered in the project supported by the National Research Center, Poland, under Grant No. 23/B/HS6/03916.

ORCID iD

Kazimierz M. Slomczynski https://orcid.org/0000-0001-7704-0885

Notes

1. Our decision of choosing attending demonstrations and signing petitions was motivated by both theoretical arguments (van Deth, 2014) and empirical considerations since both of these variables are good indicators of protest behavior (Slomczynski et al., 2016).
2. Relying on the concept of survey life cycling (Survey Research Center, 2016), we cover all these dimensions that significantly affect the end product of survey data production (Tomescu-Dubrow & Slomczynski, 2016). The full survey life cycle extends to issues arising prior to survey design (tenders, bids, and contracts) and after data dissemination (statistical analysis). Of course, in each dimension, some specific indicators can be added to our list. However, ours are the most frequently mentioned in the literature of the Total Survey Error (e.g., Biemer, 2010; Groves, 1989; Groves & Lyberg, 2010; Lyberg & Weisberg, 2016; Smith, 2011), Survey Quality Monitoring (e.g., Lyberg & Biemer, 2008; Lyberg & Stukel, 2010; Morganstein & Marker, 1997), and Fitness for Intended Use (e.g., Biemer & Lyberg, 2003; Juran & Gryna, 1980).
3. We do not take into account the questions that refer to the general attitude to demonstrations as a form of expressing an opinion in democratic societies nor do we analyze the source questions about fear of taking part in demonstrations. However, the SDR data set contains a variable dealing with the potential of taking part in demonstrations that is not part of this article.
4. The interest of the survey management calls for showing that the fieldwork is performed according to the methodological standards; thus, there is a reason to believe that a lack of information on fulfilling a given requirement reflects a real deficiency of the fieldwork. However, just poor management in preparing documentation could be involved. Either way, the quality problem exists.

References

Archer, M. (2003). *Structure, agency and the internal conversation*. Cambridge University Press.

Biemer, P. P. (2010). Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly*, *74*(5), 817-848. https://doi.org/10.1093/poq/nfq058

Biemer, P. P. (2016). Total survey error paradigm: Theory and practice. In C. Wolf, D. Joye, T. W. Smith, & Y. Fu (Eds.), *The Sage handbook of survey methodology* (pp. 122-142). Sage.

Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to survey quality*. John Wiley & Sons.

Billiet, J. (2003). Cross-cultural equivalence with structural equation modeling. In J. A. Harkness, F. J. R. van de Vijver, & P. Ph. Mohler (Eds.), *Cross-cultural survey methods* (pp. 247-263). John Wiley & Sons.

Billiet, J., Carton, A., & Loosveldt, G. (2004). Assessment of survey data quality: A pragmatic approach focused on interviewer tasks. *International Journal of Market Research*, *46*(1), 65-82. https://doi.org/10.1177/147078530404600101

Blasius, J., & Thiessen, V. (2012). *Assessing the quality of survey data*. Sage.

Burkhauser, R. V., & Lillard, D. R. (2005). The contribution and potential of data harmonization for cross-national comparative research. *Journal of Comparative Policy Analysis: Research and Practice*, *7*(4), 313-330. https://doi.org/10.1080/13876980500319436

Cheung, G. W. (2008). Testing equivalence in the structure, means, and variances of higher-order constructs with structural equation modeling. *Organizational Research Methods*, *11*(3), 593-613. https://doi.org/10.1177/1094428106298973

Cieciuch, J., Davidov, E., Schmidt, P., & Algesheimer, R. (2016). Assessment of cross-cultural comparability. In C. Wolf, D. Joye, T. W. Smith, & Y. Fu (Eds.), *The Sage handbook of survey methodology* (pp. 630-648). Sage.

Claassen, Ch. (2019). Does public support help democracy survive? *American Journal of Political Science*, *64*(1), 118-134. https://doi.org/10.1111/ajps.12452

Dalton, R., Van Sickle, A., & Weldon, S. (2010). Individual–institutional nexus of protest behaviour. *British Journal of Political Science*, *40*(1), 51-73. https://doi.org/10.1017/S000712340999038X

Dubrow, J. K. (2015). Political inequality is international, interdisciplinary and intersectional. *Sociology Compass*, *9*(6), 477-486. https://doi.org/10.1111/soc4.12270

Dubrow, J. K., Slomczynski, K. M., & Tomescu-Dubrow, I. (2008). Effects of democracy and inequality on soft political-protest in Europe: Exploring the European Social Survey data. *International Journal of Sociology*, *38*(3), 36-51. https://doi.org/10.2753/IJS0020-7659380302

Dubrow, J. K., & Tomescu-Dubrow, I. (2016). The rise of cross-national survey data harmonization in the social sciences: Emergence of an interdisciplinary methodological field. *Quality & Quantity*, *50*(4), 1449-1467. https://doi.org/10.1007/s11135-015-0215-z

Ehling, M., & Rendtel, U. (2006). *Research results of Chintex: Summary and conclusions*. Statistisches Bundesamt.

Elder, G. H., Johnson, M., & Crosnow, R. (2003). The emergence and development of life course theory. In J. T. Mortimer & M. S. Shanahan (Eds.), *Handbook of the life course* (pp. 3-19). Kluwer.

Foa, R. S., & Mounk, Y. (2016). The danger of deconsolidation: The democratic disconnect. *Journal of Democracy*, *27*(3), 5-17. https://doi.org/10.1353/jod.2016.0049

Foa, R. S., & Mounk, Y. (2017). The signs of deconsolidation. *Journal of Democracy*, *28*(1), 5-16. https://doi.org/10.1353/jod.2017.0000

Fortier, I., Raina, P., Van den Heuvel, E. R., Griffith, L. E., Craig, C., Saliba, M., Doiron, D., Stolk, R. P., Knoppers, B. M., Ferretti, V., Granda, P., & Burton, P. (2017). Maelstrom Research guidelines for rigorous retrospective data harmonization. *International Journal of Epidemiology*, *46*(1), 103-105. https://doi.org/10.1093/ije/dyw075

Granda, P., & Blasczyk, E. (2016). *Cross-cultural survey guidelines: Data harmonization*. http://ccsg.isr.umich.edu/index.php/chapters/data-harmonization-chapter

Granda, P., Wolf, Ch., & Hadorn, R. (2010). Harmonizing survey data. In J. A. Harkness, M. Braun, B. Edwards, T. Johnson, L. Lyberg, P. Ph. Mohler, B-E. Pennell, & T. W. Smith (Eds.), *Survey methods in multinational, multicultural and multiregional contexts* (pp. 315-332). John Wiley & Sons.

Groves, R. M. (1989). *Survey errors and survey costs*. John Wiley & Sons.

Groves, R. M., & Lyberg, L. E. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly*, *74*(5), 849-879. https://doi.org/10.1093/poq/nfq065

Gummer, T., & Roßmann, J. (2013). Good questions, bad questions? A post-survey evaluation strategy based on item nonresponse. *Survey Methods: Insights From the Field*. http://surveyinsights.org/?p=2330

Günther, R. (2003). *Report on compiled information of the change from input harmonization to ex-post harmonization in national samples of the European Community Household Panel: Implications on data quality* (Working Paper #19). Statistisches Bundesamt.

Hayashi, F. (2000). *Econometrics*. Princeton University Press.

Jenkins, J. C., & Form, W. (2005). Social movements and social change. In T. Janoski, R. Alford, A. Hicks, & M. Schwartz (Eds.), *Handbook of political sociology* (pp. 1446-1528). Cambridge University Press.

Jowell, R. (1998). How comparative is comparative research? *American Behavioral Scientist*, *42*(2), 168-177. https://doi.org/10.1177/0002764298042002004

Juran, J., & Gryna, F. (1980). *Quality planning and analysis* (2nd ed.). McGraw-Hill.

Kenneth, P., & Yeates, N. (2001). Defining and constructing the research process. In P. Kennet (Ed.), *Comparative social policy* (pp. 40-61). Open University Press.

Kitschelt, H. P. (1986). Political opportunity structures and political protest: Anti-nuclear movements in four democracies. *British Journal of Political Science*, *16*(1), 57-85. https://doi.org/10.1017/S000712340000380X

Kołczyńska, M. (2014). Representation of Southeast European countries in international survey projects: Assessing data quality. *Ask: Research and Method*, *23*(1), 57-78. https://kb.osu.edu/bitstream/handle/1811/69609/ASK_2014_57_78.pdf

Kołczyńska, M. (2020). Micro- and macro-level determinants of participation in demonstrations: An analysis of cross-national survey data harmonized ex-post. *mda: methods, data, analyses*, *14*(1), 91-126. https://doi.org/10.12758/mda.2019.07

Kołczyńska, M., & Schoene, M. (2018). Survey data harmonization and the quality of data documentation in cross-national surveys. In T. P. Johnson, B.-E. Pennell, I. A. L. Stoop, & B. Dorer (Eds.), *Advances in comparative survey methods: Multinational, multiregional, and multicultural contexts* (pp. 963-984). Wiley.

Kołczyńska, M., & Slomczynski, K. M. (2018). Item metadata as controls for ex post harmonization of international survey projects. In T. P. Johnson, B.-E. Pennell, I. A. L. Stoop, & B. Dorer (Eds.), *Advances in comparative survey methods: Multinational, multiregional, and multicultural contexts* (pp. 1011-1033). Wiley.

Koopmans, R., & Statham, P. (2000). *Challenging immigration and ethnic relations politics: Comparative European perspectives*. Oxford University Press.

Lyberg, L. E., & Biemer, P. B. (2008). Quality assurance and quality control in surveys. In E. D. de Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 421-441). Psychology Press.

Lyberg, L. E., & Stukel, D. M. (2010). Quality assurance and quality control in cross-national comparative studies. In J. A. Harkness, M. Braun, B. Edwards, T. Johnson, L. Lyberg, P. Ph.

Mohler, B.-E. Pennell, & T. W. Smith (Eds.), *Survey methods in multinational, multicultural and multiregional contexts* (pp. 227-349). John Wiley & Sons.

Lyberg, L. E., & Weisberg, H. F. (2016). Total survey error paradigm in practice. In C. Wolf, D. Joye, T. W. Smith, & Y. Fu (Eds.), *The Sage handbook of survey methodology* (pp. 27-40). Sage.

Marien, S., Hooghe, M., & Quintelier, E. (2010). Inequalities in non-institutionalised forms of political participation: A multi-level analysis of 25 countries. *Political Studies*, *58*(1), 187-213. https://doi.org/10.1111/j.1467-9248.2009.00801.x

Marsh, A., & Kaase, M. (1979). Measuring political action. In S. H. Barnes & M. Kaase (Eds.), *Political action: Mass participation in five western democracies* (pp. 57-96). Sage.

Matsumoto, D., & van de Vijver, F. J. R. (Eds.). (2010). *Cross-cultural research methods in psychology*. Cambridge University Press.

Mechkova, V., Lührmann, A., & Lindberg, S. I. (2017). How much democratic backsliding? *Journal of Democracy*, *28*(4), 162-169. https://doi.org/10.1353/jod.2017.0075

Medina, T. R., Smith, S. H., & Long, J. S. (2009). Measurement models matter: Implicit assumptions and cross-national research. *International Journal of Public Opinion Research*, *21*(3), 333-361. https://doi.org/10.1093/ijpor/edp037

Meyer, D. S. (2004). Protest and political opportunities. *American Sociological Review*, *30*(1), 125-45. https://doi.org/10.1146/annurev.soc.30.012703.110545

Minkel, H. (2004). *Report on data conversion methodology of the change from input harmonization to ex-post harmonization in national samples of the European Community Household Panel: Implications on data quality* (CHINTEX Working Paper No. 20). Statistisches Bundesamt.

Morganstein, D., & Marker, D. A. (1997). Continuous quality improvement in statistical agencies. In L. E. Lyberg, P. P. Biemer, & M. Collins (Eds.), *Survey measurement and process quality* (pp. 475-500). John Wiley & Sons.

Mounk, Y. (2018). *The people vs. democracy: Why our freedom is in danger and how to save it*. Harvard University Press.

Norris, P. (2017). *Is Western democracy backsliding? Diagnosing the risks* (HKS Working Paper No. RWP17-012). https://www.hks.harvard.edu/publications/western-democracy-backsliding-diagnosing-risks

Oleksiyenko, O., Wysmulek, I., & Vangeli, A. (2018). Identification of processing errors in cross-national surveys. In T. P. Johnson, B.-E. Pennell, I. A. L. Stoop, & B. Dorer (Eds.), *Advances in comparative survey methods: Multinational, multiregional, and multicultural contexts* (pp. 985-1010). Wiley.

Przeworski, A., & Teune, H. (1972). *The logic of comparative social inquiry*. John Wiley & Sons.

Quaranta, M. (2016). Protesting in "hard times": Evidence from a comparative analysis of Europe, 2000–2014. *Current Sociology*, *64*(5), 736-756. https://doi.org/10.1177/0011392115602937

Rosenstone, S. J., & Hansen, J. M. (1993). *Mobilization, participation, and democracy in America*. Macmillan.

Saris, W. E., & Gallhofer, I. N. (2014). *Design, evaluation and analysis of questionnaires for survey research* (2nd ed.). John Wiley & Sons.

Slomczynski, K. M., Jenkins, J. C., Tomescu-Dubrow, I., Kołczyńska, M., Wysmułek, I., Oleksiyenko, O., Powałko, P., & Zieliński, M. W. (2017). *SDR Master Box* (V1). Harvard Dataverse. https://doi.org/10.7910/DVN/VWGF5Q

Slomczynski, K. M., Powałko, P., & Krauze, T. (2017). Non-unique records in International Survey Projects: The need for extending data quality control. *Survey Research Methods*, *11*(1), 1-17. https://doi.org/10.18148/srm/2017.v11i1.6557

Slomczynski, K. M., & Tomescu-Dubrow, I. (2006). Representation of European post-communist countries in cross-national public opinion surveys. *Problems of Post-Communism*, *53*(4), 42-52.

Slomczynski, K. M., & Tomescu-Dubrow, I. (2018). Basic principles of survey data recycling. In T. P. Johnson, B.-E. Pennell, I. A. L. Stoop, & B. Dorer (Eds.), *Advances in Comparative survey methods: Multinational, multiregional, and multicultural contexts* (pp. 937-962). Wiley.

Slomczynski, K. M., Tomescu-Dubrow, I., Jenkins, J. C., Kołczyńska, M., Powałko, P., Wysmułek, I., Oleksiyenko, O., Zieliński, M. W., & Dubrow, J. (2016). *Democratic values and protest behavior in cross-national perspective: Harmonization of data from international survey projects*. IFiS Publishers.

Smith, T. W. (2011). Refining the total survey error perspective. *International Journal of Public Opinion Research*, *23*(4), 464-484. https://doi.org/10.1093/ijpor/edq052

Sobek, M., Hindman, M., & Ruggles, S. (2007). *Using cyber-resources to build databases for social science research* (Working Paper No. 2007-01, Minnesota Population Center Working Paper Series). https://pop.umn.edu/sites/pop.umn.edu/files/wp-2007-1.pdf

Solt, F. (2008). Economic inequality and democratic political engagement. *American Journal of Political Science*, *52*(1), 48-60. https://doi.org/10.1111/j.1540-5907.2007.00298.x

Stockemer, D. (2014). What drives unconventional political participation? A two level study. *Social Science Journal*, *51*(2), 201-211. https://doi.org/10.1016/j.soscij.2013.10.012

Survey Research Center. (2016). *Guidelines for best practice in Cross-Cultural Surveys*. http://ccsg.isr.umich.edu

Thiessen, V., & Blasius, J. (2016). Another look at survey data quality. In C. Wolf, D. Joye, T. W. Smith, & Y. Fu (Eds.), *The Sage handbook of survey methodology* (pp. 613-630). Sage.

Tomescu-Dubrow, I., & Slomczynski, K. M. (2014). Democratic values and protest behavior: Data harmonization, measurement comparability, and multi-level modeling in cross-national perspective. *Ask: Research and Methods*, *23*(1), 103-114.

Tomescu-Dubrow, I., & Slomczynski, K. M. (2016). Harmonization of cross-national survey projects on political behavior: Developing the analytic framework of survey data recycling. *International Journal of Sociology*, *46*(1), 58-72. https://doi.org/10.1080/00207659.2016.1130424

van Deth, J. W. (2014). A conceptual map of political participation. *Acta Politica*, *49*(2), 349-367. https://doi.org/10.1057/ap.2014.6

Vrablikova, K. (2013). How context matters? Mobilization, political opportunity structures and nonelectoral political participation in old and new democracies. *Comparative Political Studies*, *47*(2), 203-229. https://doi.org/10.1177/0010414013488538

Waldner, D., & Lust, E. (2018). Unwelcome change: Coming to terms with democratic backsliding. *Annual Review of Political Science*, *21*(5), 1-5.21. https://doi.org/10.1146/annurev-polisci-050517-114628

Weisberg, H. F. (2005). *The total survey error approach: A guide to the new science of survey research*. University of Chicago Press.

Wolf, C., Schneider, S. L., Behr, D., & Joye, D. (2016). Harmonizing survey questions between cultures and over time. In C. Wolf, D. Joye, T. W. Smith, & Y. Fu (Eds.), *The Sage handbook of survey methodology* (pp. 503-524). Sage.

Wysmułek, I. (2018). Europe of uneven data: Country representation in international surveys on corruption, 1989–2017. *Ask: Research and Methods*, *27*(1), 87-104.

Zieliński, M. W., Powałko, P., & Kołczyńska, M. (2018). The past, present, and future of statistical weights in international survey projects: Implications for survey data harmonization. In T. P. Johnson, B-E. Pennell, I. A. L. Stoop, & B. Dorer (Eds.), *Advances in comparative survey methods: Multinational, multiregional, and multicultural contexts* (pp. 1035-1053). Wiley.

Author Biographies

Kazimierz M. Slomczynski is head of the section at the Institute of Philosophy and Sociology, the Polish Academy of Sciences (PAN), devoted to the comparative studies of social inequalities. He also directs CONSIRT – the *Cross-national Studies: Interdisciplinary Research and Training* program of The Ohio State University and PAN (consirt.osu.edu). His research interest includes social stratification and mobility, social structure and personality, and methodology of social sciences. He is a member of the Academia Europaea.

Irina Tomescu-Dubrow is a professor of sociology at the Institute of Philosophy and Sociology (IFiS), the Polish Academy of Sciences (PAN), and manages CONSIRT – the *Cross-national Studies: Interdisciplinary Research and Training* program of The Ohio State University and PAN (consirt.osu.edu). She is director of the Graduate School for Social Research (GSSR), IFiS PAN. Her research interests include social transformations, stratification, and comparative survey methodology, ex-post survey data harmonization methods especially. Irina co-edits the *Harmonization: Newsletter on Survey Data Harmonization in the Social Sciences*.

Ilona Wysmulek is an assistant professor of sociology at the Institute of Philosophy and Sociology of the Polish Academy of Sciences, Warsaw. She joined the Polish Panel Survey research team in 2013. Her main research interests cover topics of corruption, social inequality and meritocracy, as well as methodology of comparative empirical research, including ex-post harmonization methods and survey data quality.