The Large Number of Duplicate Records in International Survey Projects: The Need for Data Quality Control

Kazimierz M. Slomczynski, *The Ohio State University, Columbus, OH, USA, and Polish Academy of Sciences, Warsaw, Poland*

Przemek Powałko, *Polish Academy of Sciences, Warsaw, Poland*

Tadeusz Krauze, *Hofstra University, Hempstead, NY, USA*

All authors contributed equally to this work

Correspondence: slomczynski.1@osu.edu

Abstract: We report a curious and unexplained fact about responses in a large set of social science survey projects. A probabilistic model shows that the occurrence of identical answers to all questionnaire items in an average sample is exceedingly small, which compels us to treat duplicate records as erroneous or at least suspicious. We analyzed the set of 1,721 national surveys in 22 projects from 142 countries covering 2.3 million respondents. We found a surprisingly large number (5,893) of duplicate records in 162 national surveys from 17 projects, covering 80 countries. Although the number of duplicate records is a small fraction of all records, they diminish the data quality and potentially have undesirable effects on the results of statistical analyses. We demonstrate that identifying these records allows researchers to examine the consequences of their existence in the data files. We argue that such erroneous or suspicious records should be flagged in all publically available data archives.

Introduction

Comparative social sciences rely, to a great extent, on data from international survey projects, usually covering at least a few countries. Survey methodology specialists in comparative studies produce a large and increasing number of publications on data quality (*1*). However, one aspect of data quality has been largely neglected: the occurrence of non-unique respondents' answers across all questions in a given national survey. Although in some books and papers on survey quality "duplicate cases" are referred to as "errors," no assessment of the spread of these errors has been made in a systematic manner (*2*).

We assume that a record is erroneous, or suspicious, if it is not unique -- that is, when the set of all answers of a given respondent is identical to that of another respondent. The commonly used organization of survey data, with cases in rows and variables in columns, facilitates comparing the records of interviewees' responses. The number of rows equals the sample size and the number of columns is equal to the number of variables provided in a given survey. Thus the search for duplicate cases involves the comparison of all rows for those columns that the researcher takes into account.

Are non-unique records for respondents' answers undoubtedly suspicious? The answer to this query depends on the probability of at least two respondents providing the same answers to all questions. For a given number of questions, this probability is determined by the number of

response options, the dependence among answers to different questions and the number of respondents. In calculating the probability of duplicates, we consider these three factors referring to specific survey projects.

Data and Results

We analyzed the collection of 1,721 national surveys in 22 projects from 142 countries covering 2.3 million respondents. Information about the selection of these projects is given in supplementary materials; the homepages of survey projects and links to source data files are provided in Tables S1 and S2. Table 1 lists 17 projects in which we found duplicate records.

Table 1. International Survey Projects with Non-unique Records

Survey project*	Number of surveys	Number of countries	Average number of questions	Average sample size	Number of cases	Number of non-unique records	Number of affected	
							surveys	countries
ABS	30	13	174	1456	43691	12	3	3
AFB	66	20	210	1499	98942	28	4	4
AMB	92	24	178	1645	151341	48	12	10
ASES	18	18	193	1014	18253	8	1	1
CB	12	3	275	2052	24621	2	1	1
CDCEE	27	16	299	1071	28926	168	3	3
EB†	152	37	342	913	138753	797	11	8
EQLS	93	35	167	1135	105527	40	8	7
ESS	146	32	223	1928	281496	14	5	5
EVS	128	50	347	1301	166502	570	5	5
ISJP	21	14	205	1229	25805	2	1	1
ISSP†	363	53	88	1359	493243	923	31	19
LB	260	19	251	1134	294965	1225	32	13
LITS	64	35	636	1060	67866	32	7	7
NBB	18	3	172	1200	21601	2	1	1
PPE7N	7	7	299	2360	16522	52	1	1
WVS	184	89	221	1394	256582	1970	36	31
All projects	**1681**	**137**	**228**	**1329**	**2234636**	**5893**	**162**	**80**

* Data were downloaded at the turn of 2013/2014. For details see Table S2.

† For Eurobarometer and International Social Survey Programme only selected survey editions were used.

Abbreviations: Asian Barometer (ABS), Afrobarometer (AFB), Americas Barometer (AMB), Asia Europe Survey (ASES), Caucasus Barometer (CB), Consolidation of Democracy in Central and Eastern Europe (CDCEE), Eurobarometer (EB), European Quality of Life Survey (EQLS), European Social Survey (ESS), European Values Study (EVS), International Social Justice Project (ISJP), International Social Survey Programme (ISSP), Latinobarometro (LB), Life in Transition Survey (LITS), New Baltic Barometer (NBB), Political Participation and Equality in Seven Nations (PPE7N), World Values Survey (WVS).

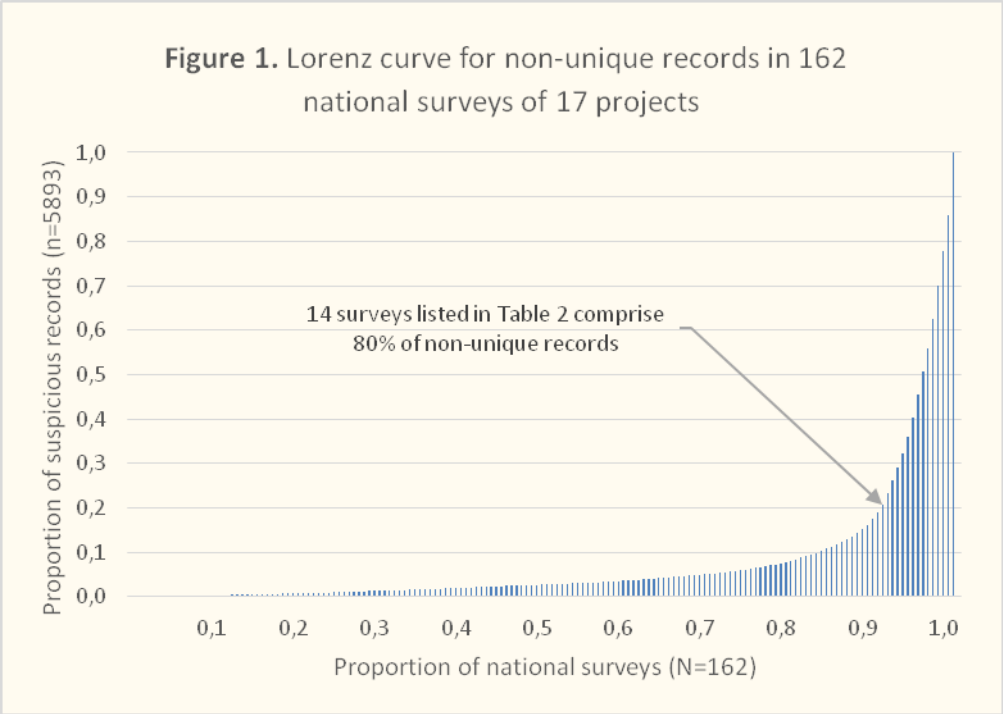
Survey questionnaires in these projects contain on average 228 questions addressed to respondents, with a range of 88 to 636 (*3*). To estimate the probability of duplicate records for such a number of questions, we assume dichotomous answers (binary choices) and the statistical independence of answers to 1/3 of the questions. The assumption about dichotomous answers provides the basis for a conservative estimate, since in practice respondents' answers are coded in multiple categories, which makes a duplicate record much less probable. The assumption of independence for a subset of questionnaire items is supported empirically: the usual pattern of statistically significant correlations of respondents' answers for a typical survey suggests that violations of postulated independence for 1/3 of items occur only rarely (*4*).

The uniqueness of records under the above assumptions is considered in terms of the classical birthday problem: how probable it is that among a given number of persons there will be a pair with the same birthday (*5*). The birthday problem is modified by replacing the number of days in a year by the number of possible sets of answers. In this model, an appropriate calculation shows that for 76 independent binary variables (1/3 of the average number of questions per national survey) one would need $3.90*10^{10}$ respondents in order to find a pair of identical sets of answers with the probability 0.01. For the 1/3 of the lower and upper bounds of the number of questionnaire items, the numbers of respondents needed for a duplicate are respectively 3,285 and $1.15*10^{31}$ (with the same probability 0.01). In other words, even for such a small number of cases as 3,285 we still need 100 samples of this size to expect a single duplicate. If duplicates occur in surveys with an average sample size (per survey project, see Table 1) ranging from 913 to 2,360 respondents and data administrators do not comment on their occurrence, such records are undoubtedly suspicious.

For all international survey projects listed in Table 1, we found 5,893 non-unique records (*6*). These records are concentrated in 162 surveys (9.4% of the total), unequally distributed. For example, they appear in 19.6% of surveys of the World Values Survey (the highest value) and 3.4% of surveys of the European Social Survey (the lowest value). Within each project, there are differences with respect to the number of countries in which surveys have non-unique records. In the extreme case, surveys in 13 out of 19 countries included in Latinobarometro contain non-unique records. The analogous numbers for the Americas Barometer are 10 in 24, and for the International Social Survey Programme 19 in 53; for other projects see Table 1. Overall, national surveys in 80 out of 142 countries have non-unique records. Thus, these results, with numbers defying the odds, show that suspicious records are common and universal.

We treat national surveys with non-unique records as imperfect since data administrators have not reported such records and relevant explanations are absent from data documentation. In 52% of the imperfect surveys a single duplicate record was found. Although a single duplicate in a survey has negligible consequences for statistical analysis, it may decrease confidence in the data. In the remaining 48% of imperfect surveys we found several patterns of non-unique records, such as multiple doublets or records repeated three, four, or even more times, often in combination. For example, in a survey conducted in Ecuador (Latinobarometro, 2000), 733 non-unique records (i.e. 272 doublets and 63 triplets) are present in the sample of 1,200, which means that over 60% of records are suspicious. In another survey in Norway (International Social Survey Programme, 2009) there are 54 records in 27 doublets, 36 in 12 triplets, 24 in 6 quadruplets, 25 in 5 quintuplets, 6 in 1 sextuplet, 7 in 1 septuplet, and 8 in 1 octuplet, with a total of 160 suspicious records in the sample of 1,456 (11.0%).

Distribution of non-unique records is provided in Figure 1.



The share of duplicate records is very uneven among national surveys: in 148 surveys there are 20% of all duplicated records and in 14 surveys there are 80% of all duplicated records. Commenting on the first group of surveys, we note that the proportion of duplicates in the respective sample usually does not exceed 1%. Are such rare occurrences troublesome for any statistical analyses? The answer to such a question depends on the kind of estimates the researchers are looking for. Even one forged duplicate record may significantly influence the results . For example, assume that the researcher is interested in finding out how many people live in the largest households in different countries. A forged duplicate record giving the misleading high maximum number may change the placement of this country among others with respect to the analyzed variable.

One duplicate record may change such a characteristic as the range of the distribution of the variable. However, what is particularly important for duplicate cases is a *pattern of values on all variables* taken into account in the analysis. A particular pattern of a single duplicated case may constitute a "deviant" case, influencing some taxonomic procedures in which respondents are clustered in multidimensional space.

In our view, even single duplicates should be subject to data quality control. Of course, a large number of duplicates in a particular survey is especially troublesome. Our analyses show that a large number of duplicates occur in the projects of Eurobarometer, European Values Study, World Values Surveys, International Social Survey Programme, Latinobarometro, and Consolidation of Democracy in Central and Eastern Europe. In 14 national surveys, included in these projects, there are more than 10% of non-unique records (see Table 2). Such a large number of non-unique records must influence the results of statistical analyses in which

variances of variables are involved, thus contaminating estimated regression models often used by researchers exploiting data from international survey projects. Analyses based on data including these imperfect surveys were published in influential books and prestigious journals.

Table 2. Fourteen National Surveys with the Largest Number of Non-unique Records

Project/wave	Country	Number of cases	Number of non--unique records	Proportion of non-unique records
CDCEE 1	Romania	1234	154	0.125
EB 19	Belgium	1038	148	0.143
EB 21	Belgium	1018	344	0.338
EB 31	Belgium	1002	220	0.219
EVS 1	United States	2325	528	0.227
ISSP 1989	Austria	1997	374	0.174
ISSP 1998	Bulgaria	1102	133	0.102
ISSP 2009	Norway	1456	160	0.110
LB 1996	Panama	1005	316	0.314
LB 2000	Ecuador	1200	733	0.611
WVS 1	Japan	1204	195	0.162
WVS 3	Mexico	2364	537	0.227
WVS 5	Ethiopia	1500	539	0.359
WVS 5	South Korea	1200	354	0.295

Discussion and conclusion

This study has implications for (a) already published work using international survey projects with duplicate records, (b) future research using these data sets, and (c) for survey methodology. Regarding the first, the estimated number of publications relying on data from analyzed projects differs depending on the source: based on information from the projects' web pages it is 11,000+, according to Google Scholar – 25,000+, and according to the Web of Science Core Collection – 2000+ publications and almost 20,000 citations (for details see Table S3 in supplementary materials). In the spirit of good science, authors may want to consider replication of their analyses with the goal of eliminating non-unique records or controlling for their presence (*8*).

To facilitate sound analyses in the future, we suggest that all institutions providing survey data check them for non-unique records and flag these records (i.e., create dedicated variables in data sets) but should not remove them from the data files. This suggestion is additionally motivated by the need for preserving original data in order to assess the effect of repeated records. It bears directly on the field of survey methodology.

Our next recommendation is that organizations conducting surveys check whether the data files contain repeated records and investigate possible sources of their occurrence. Theoretically, for any pair of identical records there are three possibilities: a) both records correspond to real respondents, b) one record corresponds to a real respondent and another one is its duplicate, and c) both records are fakes. The first possibility, as a miracle (*9*) or improbable coincidence (*10*), should be rejected on statistical grounds. For the two remaining possibilities, one could investigate whether the errors were caused by interviewers, data coders, or data processing staff.

Recent developments in the usage of paradata, the data about the process of generating survey data (*11*), provide tools for identifying the sources of non-unique records.

Our search for identical records in national surveys covers all questionnaire items that correspond to respondents' answers. In past research, investigation of duplicate records was limited to a subset of questionnaire items, with the claim that duplicates refer to "fake" cases (*12, 13*). The approach we propose is flexible with respect to the choice of variables in the investigated records, requiring that suspicious records be identified by a separate binary variable (suspicious = 1, otherwise = 0). Substantive analyses should take into account duplications as a kind of measurement error. Errors of this kind, shown to be voluminous in some national surveys administered by well-known international projects, need to be controlled for, since they reduce confidence in data and/or their effects potentially distort the results of substantive research.

References and notes

1. See, for example, L. Lyberg, *Survey Measurement and Process Quality* (New York: Wiley 1997); J. Harkness, F. J. R. van de Vijver, and P. Ph. Mohler (eds.) *Cross-cultural Survey Methods* (Hoboken, NJ: Wiley 2003); P. P. Biemer, L. Lyberg, *Introducton to Survey Quality* (Hoboken, NJ: Wiley 2003); L. Gideon (ed.) *Handbook of Survey Methodology for the Social Sciences* (New York: Springer 2012); D.E. McNabb, *Nonsampling Error in Social Surveys* (Thousand Oakes, Calif.: SAGE 2014). For a review of criteria of assessing the quality of cross-national surveys, with references to fitness for intended use, total survey error, and survey process quality, see Survey Research Center, *Guidelines for Best Practice in Cross-Cultural Surveys* (Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan 2010). http://www.ccsg.isr.umich.edu Retrieved: March 12, 2015.

2. We acknowledge here the earlier work on non-unique records, especially by J. Blasius, V. Thiessen, *Assessing the Quality of Survey Data* (London: SAGE 2012), pp. 64-67, and N. Kuriakose, M. Robbins, *Falsification in Surveys: Detecting Near Duplicate Observations* (Social Science Research Network, 2015. Retrieved March 30, 2015 from http://ssrn.com/abstract=2580502).

3. In order to obtain records of variables corresponding to questionnaire items, the following types of variables have been excluded from original cases: a) technical variables (i.e., variables created at the administrative level, e.g. population/post-stratification weights, geographical regions, size/type of community), b) variables containing interviewers' remarks (e.g. interview details, level of respondent's cooperation, respondent's race), c) variables derived from respondent's answers (e.g. BMI, classified education/occupational levels), and d) all variables which can be derived from sample characteristics or from construction of the sample (e.g. respondent's age and gender, information about household members).

4. This empirical evidence gives only intuitive support for our assumption since even zero-correlations do not imply statistical independence.

5. W. Feller, *An Introduction To Probability Theory and Its Applications* (New York: Wiley, 3rd ed 1968), vol.1, p. 33.

6. Among the non-unique records, only 67 are clearly lacking the respondents' answers as if the relevant interviews had been interrupted or not even begun. For the complete list of suspicious records see Table S4 in supplementary materials.

7. For books, see e.g.: R. Inglehard, *Silent Revolution. Changing Values and Political Styles in 43 Western Publics* (Princeton, NJ: Princeton University Press 1977) and R. Inglehard, *Modernism and Postmodernism. Cultural, Economic, and Political Change in 43 Societies* (Princeton, NJ: Princeton University Press 1997), each extensively cited (cf. Google Scholar, citation indices, retrieved May 30, 2015). For articles, cf. indexes in *Am. Poli Sci. Rev., Am. Socio. Rev.* and other leading journals in the social sciences, and a recent article by N. Fuchs-Schündeln, M. Schündeln, Political economy. On the Endogeneity of Political Preferences: Evidence from Individual Experience with Democracy. *Science* 347 (2015) pp. 1145-1148.

8. G. King, Replication, Replication, *PS: Poli. Sci. and Politics* 28 (1995), pp. 443-49.

9. J. Kruskal, Miracles and Statistics: the Casual Assumption of Independence. *J. Am. Stat. Assoc.* 83 (1988), pp. 929-940.

10. P. Diaconis, F. Mosteller, Method of Studying Coincidences. *J. Am. Stat. Assoc.* 84 (1989), pp. 853-861.

11. F. Kreuter, Ed., *Improving Surveys with Paradata* (Haboken, NJ: Wiley 2013), pp. 2-3.

12. J. Blasius, V. Thiessen, *op. cit.*

13. N. Kuriakose, M. Robbins, *op. cit*.

Acknowledgments: This work was supported by a grant "Democratic Values and Protest Behavior: Data Harmonization, Measurement Comparability, and Multi-Level Modeling in Cross-National Perspective" from the (Polish) National Science Centre (2012/06/M/HS6/00322). The data used in this paper come from publically available sources provided in the supplementary materials, Table S1.

Supplementary Materials

Originally, we have chosen 22 international projects according to the following criteria: the projects are non-commercial; they were designed as cross-national, and – preferably – multi-wave; the samples were intended to represent the adult population of a given country or territory; questionnaires contain questions about political attitudes and behaviors; data are freely available; survey documentation (study description, codebook and/or questionnaire) is provided in English.

Table S1. Homepages of the 22 International Survey Projects*

Project	Official name of project	Homepage
AFB	Afrobarometer	http://afrobarometer.org
AMB	Americas Barometer	http://www.vanderbilt.edu/lapop
ARB	Arab Barometer	http://www.arabbarometer.org
ABS	Asian Barometer	http://www.asianbarometer.org
ASES	Asia Europe Survey	http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/22324?q=asia+europe+survey
CB	Caucasus Barometer	http://www.crrccenters.org
CDCEE	Consolidation of Democracy in Central and Eastern Europe	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=4054
CNEP	Comparative National Elections Project	http://www.cnep.ics.ul.pt
EB	Eurobarometer	http://zacat.gesis.org/webview/main.jsp?object=http://zacat.gesis.org/obj/fCatalog/Catalog57
EQLS	European Quality of Life Survey	http://discover.ukdataservice.ac.uk/Catalogue/?sn=7348
ESS	European Social Survey	http://www.europeansocialsurvey.org
EVS	European Values Study	http://www.europeanvaluesstudy.eu
ISJP	International Social Justice Project	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=3522
ISSP	International Social Survey Programme	http://www.issp.org
LB	Latinobarometro	http://www.latinobarometro.org
LITS	Life in Transition Survey	http://www.ebrd.com/what-we-do/economic-research-and-data/data/lits.html
NBB	New Baltic Barometer	http://discover.ukdataservice.ac.uk/catalogue/?sn=6510
PA2	Political Action II	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=1188
PA8NS	Political Action - An Eight Nation Study	http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/07777
PPE7N	Political Participation and Equality in Seven Nations	http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/07768
VPCPCE	Values and Political Change in Postcommunist Europe	http://discover.ukdataservice.ac.uk/catalogue/?sn=4129
WVS	World Values Survey	http://www.worldvaluessurvey.org

* For some projects that do not have their own web pages, the archiving organization's web page was used as a source.

Table S2. Information on Source Data Files*

Target dataset name	Original data file name	Source link	Time downloaded
AFB_1	merged_r1_data.sav	http://afrobarometer.org	2014-02-12
AFB_2	merged_r2_data.sav	http://afrobarometer.org	2014-04-07
AFB_3	merged_r3_data.sav	http://afrobarometer.org	2014-04-07
AFB_4	merged_r4_data.sav	http://afrobarometer.org	2014-04-07
AMB_1_5	AmericasBarometer Merged 2004-2012 Rev1.5_FREE.sav	http://www.vanderbilt.edu/lapop	2014-04-07
ARB_1	ABI_datafilefinal_1_1.sav	http://www.arabbarometer.org	2014-02-12
ARB_2	ADBII Merged Data_file English FINAL_0.sav[1]	http://www.arabbarometer.org	2014-04-07
ASB_1	ABS-w1-merge-v4.2.sav	Directly from the project authors	2013-10-24
ASB_2	ABS-w2_merge13_v20100810.sav	Directly from the project authors	2013-10-24
ASB_3	ABS3w.IN.sav[2]	Directly from the project authors	2013-10-24
	ABS3w.KR.sav[2]		2013-10-24
	ABS3w.MA.sav[2]		2013-10-24
	ABS3w.MN.sav[2]		2013-10-24
	ABS3w.PH.sav[2]		2013-10-24
	ABS3w.SG.sav[2]		2013-10-24
	ABS3w.TH.sav[2]		2013-10-24
	ABS3w.TW.sav[2]		2013-10-24
	ABS3w.VN.sav[2]		2013-10-24
ASES	22324-0001-Data.sav	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/22324	2013-10-18
CB_2009	CB2009_Regional_only_responses_012011.dta[3]	http://caucasusbarometer.org	2014-04-13
CB_2010	CB2010_Regional_only_responses_032011.dta[4]	http://caucasusbarometer.org	2014-04-13
CB_2011	CB2011_Regional_only_responses_02012.dta[5]	http://caucasusbarometer.org	2014-04-13
CB_2012	CB2012_Regional_only_responses_032013.sav[6]	http://caucasusbarometer.org	2014-04-13
CDCEE_1_2	ZA4054.sav	https://dbk.gesis.org/dbksearch/sdesc2sp?no=4054	2013-10-14
CNEP_3_ES	CN3Spain2004Mod.sav	http://www.cnep.ics.ul.pt	2014-01-20
CNEP_3_HU	CN3Hungary2006Mod.sav	http://www.cnep.ics.ul.pt	2014-01-20
CNEP_3_MX	CN3Mexico2006Mod.sav	http://www.cnep.ics.ul.pt	2014-01-20
CNEP_3_MZ	CN3Mozambique2005Mod.sav	http://www.cnep.ics.ul.pt	2014-01-20
CNEP_3_PT	CN3Portugal2005Mod.sav	http://www.cnep.ics.ul.pt	2014-01-20
CNEP_3_TW	CN3Taiwan2004Mod.sav	http://www.cnep.ics.ul.pt	2014-01-20
CNEP_3_UY	CN3Uruguay2004Mod.sav	http://www.cnep.ics.ul.pt	2014-01-20
CNEP_3_ZA	CN3SouthAfrica2004Mod.sav	http://www.cnep.ics.ul.pt	2014-01-20

Target dataset name	Original data file name	Source link	Time downloaded
EB_1983	ZA1318_v1-0-1.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=1318	2013-12-01
EB_1984	ZA1320_v1-0-1.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=1320	2013-12-01
EB_1989	ZA1750_v1-0-1.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=1750	2013-12-01
EB_2000	ZA3387_v1-1-0.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=3387	2013-12-01
EB_2004	ZA4231_v1-1-0.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=4231	2013-12-01
EB_2010	ZA5234_v2-0-1.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=5234	2013-12-01
EB_2012	ZA5612_v1-0-0.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=5612	2013-12-01
EQLS_1_3	eqls_final_integrated_dataset_2003-2012.sav[7]	http://discover.ukdataservice.ac.uk/Catalogue/?sn=7348	2014-02-03
ESS_1_5	ESS1-5_cumulative_e01_1.sav[8]	http://www.europeansocialsurvey.org	2013-11-08
	ESS4IL.sav[9]	http://www.europeansocialsurvey.org	2014-07-23
	ESS4SK.sav[9]	http://www.europeansocialsurvey.org	2014-07-23
	ESS4TR.sav[9]	http://www.europeansocialsurvey.org	2014-07-23
	ESS5DK.sav[10]	http://www.europeansocialsurvey.org	2014-07-23
	ESS5IL.sav[10]	http://www.europeansocialsurvey.org	2014-07-23
	ESS5SK.sav[10]	http://www.europeansocialsurvey.org	2014-07-23
ESS_6	ESS6e01.2_F1.sav[11]	http://nesstar.ess.nsd.uib.no/webview/velocity?mode=download&analysismodetable&v=2&study=http%3A%2F%2F19.177.90.83%3A-1%2Fobj%2FfStudy%2FESS6e01.2	2014-07-22
	ESS6PT.sav[12]	http://www.europeansocialsurvey.org	2014-07-23
ISJP_1_2	06705-0001-Data.sav	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/6705	2014-01-21
ISSP_1985	ZA1490.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=1490	2013-09-12
ISSP_1989	ZA1840_2006-04-27.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=1840	2013-09-12
ISSP_1990	ZA1950_2006-04-12.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=1950	2013-09-12
ISSP_1991	ZA2150_2006-04-24.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=2150	2013-09-12
ISSP_1996	ZA2900_2006-04-12.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=2900	2013-09-12
ISSP_1998	ZA3190_2006-04-27.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=3190	2013-09-12
ISSP_2004	ZA3950_v1-3-0.sav	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=3950	2013-09-12
ISSP_2006	ZA4700.sav	https://dbk.gesis.org/dbksearch/sdesc2	2013-09-12

Target dataset name	Original data file name	Source link	Time downloaded
		sp?no=4700	
ISSP_2007	ZA4850_v2-0-0.sav	https://dbk.gesis.org/dbksearch/sdesc2.sp?no=4850	2013-09-12
ISSP_2008	ZA4950_v2-2-0.sav	https://dbk.gesis.org/dbksearch/sdesc2.sp?no=4950	2013-09-12
ISSP_2009	ZA5400_v3-0-0.sav	https://dbk.gesis.org/dbksearch/sdesc2.sp?no=5400	2013-09-12
ISSP_2010	ZA5500_v2-0-0.sav	https://dbk.gesis.org/dbksearch/sdesc2.sp?no=5500	2013-09-12
ISSP_2011	ZA5800_v2-0-0.sav	https://dbk.gesis.org/dbksearch/sdesc2.sp?no=5800	2013-09-12
IVS_1_9	ZA4804_v2-0-0.sav	https://dbk.gesis.org/dbksearch/SDESC2.asp?no=4804	2013-12-03
	ZA4804_v2-0-0_2012_09-07_patch_1.sps[13]	https://dbk.gesis.org/dbksearch/SDESC2.asp?no=4804	2013-12-03
	ZA4804_v2-0-0_UNI_code_patch.sps[14]	https://dbk.gesis.org/dbksearch/SDESC2.asp?no=4804	2013-12-03
	wvs1981_2008_v20090914_spss.sav	http://www.worldvaluessurvey.org	2013-12-03
	MergeSyntax_EVS_WVS_v2-0-0.sps[15]	https://info1.gesis.org/dbkedit18/download.asp?id=49440	2013-12-03
LB_1995	latinobarometro1995_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_1996	latinobarometro1996_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_1997	latinobarometro1997_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_1998	latinobarometro1998_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_2000	latinobarometro2000_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_2001	latinobarometro2001_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_2002	latinobarometro2002_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_2003	latinobarometro2003_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_2004	latinobarometro2004_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_2005	Latinobarometro_2005_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_2006	Latinobarometro_2006_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_2007	Latinobarometro_2007_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_2008	Latinobarometro_2008_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_2009	Latinobarometro_2009_eng.sav	http://www.latinobarometro.org	2013-10-14
LB_2010	Latinobarometro_2010_eng.sav	http://www.latinobarometro.org	2013-10-14
LITS_1	LITS 2006 data.dta	http://www.ebrd.com/what-we-do/economic-research-and-data/data/lits.html	2014-02-20
LITS_2	lits2.dta	http://www.ebrd.com/what-we-do/economic-research-and-data/data/lits.html	2014-02-20
NBB_1_6	mbltrv.sav[16]	http://discover.ukdataservice.ac.uk/catalogue/?sn=6510	2014-02-04
	mblt-weights.sps[17]	http://discover.ukdataservice.ac.uk/catalogue/?sn=6510	2014-02-04

Target dataset name	Original data file name	Source link	Time downloaded
PA2	ZA1188.sav	https://dbk.gesis.org/dbksearch/sdesc2sp?no=1188	2014-02-03
PA8NS	07777-0001-Data.por	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07777	2014-02-02
PPE7N_AT	07768-0001-Data.txt	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07768	2014-02-02
	07768-0001-Setup.sps	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07768	2014-02-02
PPE7N_IN	07768-0002-Data.txt	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07768	2014-02-02
	07768-0002-Setup.sps	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07768	2014-02-02
PPE7N_JP	07768-0003-Data.txt	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07768	2014-02-02
	07768-0003-Setup.sps	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07768	2014-02-02
PPE7N_NG	07768-0005-Data.txt	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07768	2014-02-02
	07768-0005-Setup.sps	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07768	2014-02-02
PPE7N_NL	07768-0004-Data.txt	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07768	2014-02-02
	07768-0004-Setup.sps	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07768	2014-02-02
PPE7N_US	07015-0001-Data.txt	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07015	2014-02-02
	07015-0001-Setup.sps	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07015	2014-02-02
PPE7N_YU	07768-0006-Data.txt	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07768	2014-02-02
	07768-0006-Setup.sps	http://www.icpsr.umich.edu/icpsrweb/CPSR/studies/07768	2014-02-02
VPCPCE_CZ	C-PUBLIC.POR	http://discover.ukdataservice.ac.uk/catalogue/?sn=4129	2014-02-04
VPCPCE_HU	H-PUBLIC.POR	http://discover.ukdataservice.ac.uk/catalogue/?sn=4129	2014-02-04
VPCPCE_RU	R-PUBLIC.POR	http://discover.ukdataservice.ac.uk/catalogue/?sn=4129	2014-02-04
VPCPCE_SK	S-PUBLIC.POR	http://discover.ukdataservice.ac.uk/catalogue/?sn=4129	2014-02-04
VPCPCE_UA	U-PUBLIC.POR	http://discover.ukdataservice.ac.uk/catalogue/?sn=4129	2014-02-04

* Data files were downloaded at the turn of 2013/2014. Whenever possible patches and corrections published through the end of March 2014 were applied.

[1] The data file was changed (without warning) sometime between the first and last access; we are using the later version. [2] All nine files have an identical structure; we joined them into one integrated file. [3] On first access another file was available: CB_2009_Final_22.10.2010_Regional_Without_NR.sav. [4] On first access another file was available: CB_2010_regional_weighted_only_responses_03212011.sav. [5] On first access another file was available: CB2011_Regional_Weighted_12.07.12.sav. [6] On first access another file was available: CB2012_Regional_Weighted_01.03.2013.sav. [7] 2014-01-21 a new version appeared with this time signature. [8] In this file, wave 4 is in edition 4.1, wave 5 in edition 3.0. [9] Wave 4 edition 4.2 published 2014-04-29; an auxiliary file for a patch. [10] Wave 5 edition 3.1 published 2014-04-29; an auxiliary file for a patch. [11] Wave 6 edition 1.2 published 2014-01-13. [12] Wave 6 edition 2.0 published 2014-05-14; an auxiliary file for a patch. [13] EVS data file has to be patched before merging with WVS. [14] This may also be done before the merge. [15] Merge syntax file. [16] First edition: 2010-07-29. [17] The relevance of this patch file has not been determined.

Table S3. Estimated number of publications that used data from international survey projects

Project	Number of publications listed in			Number of Citations in Web of Science[†]
	homepages[*]	Google Scholar[#]	Web of Science[†]	
AFB[1]	428	1307 (5230)	55	204
AMB[2]	312	251 (502)	13	27
ARB[3]	30	174 (348)	3	6
ABS[4]	322	177 (354)	4	2
ASES[5]	1	37 (74)	2	0
CB[6]	96	66 (164)	0	0
CDCEE[7]	1	81 (163)	0	0
CNEP[8]	65	49 (326)	3	1
EB[9]	825	1167 (40000)	409	4992
EQLS[10]	70	915 (1830)	27	116
ESS[11]	1362	4600 (13800)	590	3637
EVS[12]	1384	3293 (9878)	175	1397
ISJP[13]	2	230 (461)	20	518
ISSP[14]	6569	1443 (9660)	283	3281
LB[15]	54	1437 (4600)	21	156
LITS[16]		195 (391)	7	1
NBB[17]	27	118 (237)	2	3
PA2[18]	12	46 (93)	0	0
PA8NS[19]	50	78 (156)	0	0
PPE7N[20]	8	23 (47)	0	0
VPCPCE[21]		30 (60)	1	0
WVS[22]	128	9334 (28003)	472	5385
Total	**11746**	**25051 (116377)**	**2087**	**19726**

[*] Data gathered on 2015-02-06

[#] Data gathered on 2015-03-19. For the total number of items found on Google Scholar for a given project (provided in parentheses), we estimated the number of publications that refer to the project data in two steps: first, we decreased the total number of items proportionally to the number of relevant waves (e.g. for Eurobarometer we took 7 waves out of 80, i.e. 40,000 * 0.0875); second, for large projects with the total number of items over 3000, we divide this number by 3; for the remaining projects we divide this number by 2.

[†] Data gathered on 2015-03-31

The following expressions have been used for searches: [1] "afrobarometer" OR "afro-barometer" OR "afro barometer" [2] "americas barometer" [3] "arab barometer" [4] "asian barometer survey" [5] "asia europe survey" [6] "caucasus barometer" [7] "consolidation of democracy in central and eastern europe" [8] "comparative national elections project" OR "comparative national election project" [9] "eurobarometer" [10] "european quality of life survey" [11] "european social survey" [12] "european values study" OR "european value study" OR "european values survey" OR "european value survey" [13] "international social justice project" [14] "international social survey programme" OR "international social survey program" [15] "latinobarometro" OR "latino barometro" OR "latino barometer" OR "latinobarometer" [16] "life in transition survey" [17] "new baltic barometer" [18] "political action ii" [19] "political action" "eight nation study" [20] "political participation

and equality" "verba" [21] "values and political change in post communist europe" [22] "world values survey" OR "world value survey" OR "world values study" OR "world value study"

Table S4. List of IDs of suspicious records.

Dataset	**Country** *	**Filtering variables**	**ID variable**
AFB_1	ZW		refnumb= (1479 1645)
AFB_1	ML		refnumb= (13416 13436)
AFB_1	NG		refnumb= (16940 16941) (13553 15734) (16309 16311) (16285 17057) (14052 15860) (16647 16648) (13552 15733) (13551 15732)
AFB_1	ZA		refnumb= (8459 8460) (8339 8898) (7318 8686) (7484 7485)
AMB_1_5	BO	year=2004 pais=10	idnum= (10092 10080)[1]
AMB_1_5	NI	year=2006 pais=5	idnum= (852 515) (513 850) (508 855)
AMB_1_5	PA	year=2006 pais=7	idnum= (694 692)
AMB_1_5	GY	year=2008 pais=24	idnum= (125 126)
AMB_1_5	NI	year=2008 pais=5	idnum= (487 479)
AMB_1_5	CR	year=2008 pais=6	idnum= (473 472)
AMB_1_5	PY	year=2010 pais=12	idnum= (1082 853)
AMB_1_5	GT	year=2010 pais=2	idnum= (987 988) (922 921) (977 978) (928 929) (924 925) (1501 1499) (798 799) (718 719) (1490 1491) (481 479)
AMB_1_5	DO	year=2010 pais=21	idnum= (622 196) (199 625)
AMB_1_5	EC	year=2010 pais=9	idnum= (1740 1744)
AMB_1_5	BR	year=2012 pais=15	idnum= (1317 1443)
AMB_1_5	EC	year=2012 pais=9	idnum= (1311 1397)
ASB_2	CN	country=4	idnumber= (4533 4568 4572)
ASB_3	SG	country=10	idnumber= (152 991) (518 998) (519 994 999)
ASB_3	VN	country=11	idnumber= (731 732)
ASES	PT		v0001= (800495 801013) (800466 801014) (800587 801011) (800489 801012)
CB_2009	AM		id= (191803714 191804614)

Dataset	Country *	Filtering variables	ID variable
CDCEE_1_2	RO		id= (22024 22034 22044) (22395 22396) (22456 22466 22571 22581) (21993 21994 21998) (21732 22340 22350 22360) (21492 21493) (22515 22516) (22333 22343 22353 22363) (22011 22012) (22436 22441 22445 22449 22453) (22675 22676) (22462 22567 22577 22587) (22019 22029 22039 22049) (22464 22569 22579 22589) (21733 22341 22351 22361) (22439 22443 22447 22451) (22524 22525 22526) (22437 22442 22446 22450) (22458 22563 22573 22583) (22018 22028 22038 22048) (21730 22338 22348 22358) (22022 22032 22042 22670) (22459 22564 22574 22584) (22457 22562) (22463 22568 22578 22588) (22016 22026 22036 22046) (22335 22345 22355 22365) (22020 22030 22040 22050) (21731 22339 22349 22359) (22017 22027 22037 22047) (21729 22337 22347 22357) (21505 21506) (22332 22342 22352 22362) (22455 22465 22570 22580) (22334 22344 22354 22364) (22336 22346 22356 22366) (22440 22444 22448 22452) (22021 22031 22041 22051) (21832 21833) (22460 22565 22575 22585) (22461 22566 22576 22586) (22023 22033 22043 22672) (22015 22025 22035 22045)
CDCEE_1_2	RU-KRA		id= (25742 26617) (25734 26204) (25738 26854) (25722 26160) (25728 26051) (25736 26022)
CDCEE_1_2	BG		id= (2378 2519)
EB_1983	BE		v4= (20913 21019) (20624 21022) (20623 21021) (20444 20487) (20445 20488) (20733 20983) (20448 20479) (20900 21008) (20926 21029) (20440 20484) (20897 21005) (20887 20998) (20889 21000) (20452 20483) (20796 20986) (20442 20485) (20797 20987) (20893 21001) (20885 20996) (20920 21023) (20730 20980) (20799 20989) (20886 20997) (20940 20977) (20937 21038) (20924 21027) (20955 21033) (20957 21035) (20941 20978) (20446 20489) (20958 21036) (20449 20480) (20921 21024) (20956 21034) (20922 21025) (20899 21007) (20954 21032) (20936 21037) (20450 20481) (20942 20979) (20943 20984) (20802 20992) (20794 20985) (20923 21026) (20622 21020) (20805 20995) (20903 21011) (20908 21016) (20804 20994) (20927 21030) (20905 21013) (20443 20486) (20953 21031) (20909 21017) (20906 21014) (20904 21012) (20888 20999) (20902 21010) (20898 21006) (20732 20982) (20801 20991) (20925 21028) (20803

Dataset	Country *	Filtering variables	ID variable
			20993) (20910 21018) (20798 20988) (20894 21002) (20447 20490) (20895 21003) (20901 21009) (20896 21004) (20907 21015) (20731 20981) (20800 20990) (20451 20482)
EB_1983	DE-W		v4= (41061 41070) (41059 41072) (41049 41083) (41052 41724) (41054 41726) (41055 41889) (41057 41069) (41060 41073) (41058 41071) (41048 41077) (41062 41739) (41056 41106)
EB_1983	IT		v4= (55416 55427)
EB_1984	BE		v4= (1310 1846) (1460 1924) (1461 1925) (1437 1901) (1833 2005) (1836 2008) (1446 1910) (1031 1840) (1785 1976) (1824 1996) (1425 1889) (1390 1859) (1406 1872) (1831 2003) (1398 1864) (1418 1882) (1450 1914) (1387 1856) (1443 1907) (1458 1922) (1383 1852) (1454 1918) (1444 1908) (1439 1903) (1445 1909) (1409 1875) (1718 1953) (1417 1881) (1028 1837) (1407 1873) (1401 1867) (1778 1969) (1063 1841) (1408 1874) (1442 1906) (1411 1877) (1780 1971) (1465 1929) (1313 1849) (1801 1979) (1030 1839) (1391 1860) (1659 1946) (1827 1999) (1441 1905) (1779 1970) (1435 1899) (1820 1992) (1385 1854) (1784 1975) (1834 2006) (1835 2007) (1656 1943) (1421 1885) (1384 1853) (1393 1862) (1448 1912) (1803 1981) (1447 1911) (1436 1900) (1429 1893) (1405 1871) (1029 1838) (1809 1987) (1457 1921) (1462 1926) (1314 1850) (1459 1923) (1315 1851) (1455 1919) (1781 1972) (1065 1843) (1311 1847) (1309 1845) (1432 1896) (1832 2004) (1388 1857) (1729 1964) (1389 1858) (1720 1955) (1456 1920) (1713 1948) (1711 1947) (1722 1957) (1808 1986) (1823 1995) (1813 1991) (1464 1928) (1822 1994) (1423 1887) (1805 1983) (1658 1945) (1400 1866) (1426 1890) (1066 1844) (1438 1902) (1428 1892) (1786 1977) (1825 1997) (1420 1884) (1452 1916) (1427 1891) (1730 1965) (1449 1913) (1424 1888) (1821 1993) (1802 1980) (1727

Dataset	Country *	Filtering variables	ID variable
			1962) (1064 1842) (1399 1865) (1721 1956) (1728 1963) (1714 1949) (1828 2000) (1386 1855) (1807 1985) (1402 1868) (1726 1961) (1312 1848) (1830 2002) (1829 2001) (1731 1966) (1403 1869) (1392 1861) (1787 1978) (1397 1863) (1463 1927) (1719 1954) (1434 1898) (1643 1935) (1641 1933) (1416 1880) (1404 1870) (1410 1876) (1433 1897) (1650 1942) (1768 1968) (1430 1894) (1642 1934) (1422 1886) (1646 1938) (1724 1959) (1783 1974) (1716 1951) (1806 1984) (1649 1941) (1782 1973) (1723 1958) (1612 1930) (1440 1904) (1715 1950) (1640 1932) (1732 1967) (1412 1878) (1811 1989) (1826 1998) (1812 1990) (1639 1931) (1645 1937) (1451 1915) (1453 1917) (1717 1952) (1810 1988) (1725 1960) (1644 1936) (1431 1895) (1415 1879) (1648 1940) (1419 1883) (1804 1982) (1657 1944) (1647 1939)
EB_1984	DK		v4= (2734 2736 2737)
EB_1989	PT		v4= (11400 11404) (10824 10890) (11131 11171) (11443 11549) (11417 11557) (11442 11548) (11438 11544) (11441 11547) (10889 10891) (11403 11405) (11439 11545) (11420 11556) (11445 11543) (11426 11558) (10990 10994) (10991 10999) (10828 10892) (10989 10993) (11440 11546) (11431 11555)
EB_1989	BE		v4= (1670 1740) (1658 1968) (1219 1590) (1240 1646) (1473 1579) (1207 1636) (1617 2006) (1361 1582) (1202 1587) (1477 1628) (1612 1846) (1672 1743) (1393 1633) (1291 1560) (1171 1667) (1588 1705) (1072 1648) (1611 1847) (1673 1963) (1652 1769) (1668 1715) (1266 1600) (1613 1845) (1564 1973) (1662 1809) (1506 1634) (1184 1637) (1654 1782) (1152 1594) (1659 1818) (1520 1585) (1457 1629) (1631 1989) (1206 1674) (1569 1848) (1625 2013) (1671 1742) (1661 1812) (1312 1596) (1621 1793) (1556 1630) (1074 1650) (1614 1823) (1638 1692) (1567 1852) (1586 1997) (1669 1719) (1456 1583) (1217 1595) (1602 1956) (1297 1591) (1610 1917) (1644 1955) (1665 1903) (1233 1647) (1653 1784) (1365 1603) (1657 1953) (1469 1620) (1664 1906) (1465 1604) (1608 1836) (1606 1843) (1470 1619) (1607 1837) (1574 1868) (1572 1879) (1107 1593) (1623 1895) (1605 1915) (1073 1649) (1655 1779) (1616 2005) (1639 1694) (1656 1961) (1584 1980) (1571 1883) (1570 1885) (1626 1986) (1490

Dataset	Country *	Filtering variables	ID variable
			1666) (1609 1822) (1663 1908) (1565 1867) (1601 1924) (1589 1709) (1642 1704) (1208 1635) (1277 1651) (1599 1679) (1392 1632) (1346 1581) (1660 1815) (1592 1960) (1562 1766) (1561 1755) (1566 1853) (1624 1889) (1421 1580) (1641 1703) (1577 2018) (1615 2004) (1573 1872) (1645 1964) (1640 1702) (1598 1676) (1568 1851) (1578 1984) (1622 1795) (1576 2025) (1575 2026)
EB_2000	LU		v5= (8931 8961)
EB_2004	PT		v5= (11358 11452)
EB_2004	AT		v5= (9749 10156) (9750 10157)
EB_2004	IE		v5= (6279 7073) (6273 7076) (6243 7075) (6285 7074)
EQLS_1_3	ES		uniqueid= (ES8240020 ES8240709)
EQLS_1_3	FR		uniqueid= (FR8080264 FR8081101) (FR8080121 FR8081053) (FR8080397 FR8080517) (FR8080932 FR8081091) (FR8080191 FR8080532)
EQLS_1_3	HU		uniqueid= (HU8122742 HU8129742)
EQLS_1_3	IE		uniqueid= (IE8130873 IE8130871)
EQLS_1_3	SE		uniqueid= (SE8250841 SE8250148) (SE8250708 SE8250991) (SE8250456 SE8250565) (SE8251091 SE8250903) (SE8250621 SE8251149) (SE8251021 SE8251084) (SE8250707 SE8250467) (SE8250149 SE8250499) (SE8250516 SE8251083)
EQLS_1_3	GB		uniqueid= (UK8102109 UK8102119)
EQLS_1_3	GB		uniqueid= (UK9034188 UK9035606)
EQLS_1_3	HR		uniqueid= (HR3001206 HR3001408)
ESS_1_5	DK		cseqno= (58959 59103)
ESS_1_5	PT		cseqno= (185687 186615)
ESS_1_5	DE		cseqno= (54437 54476) (54444 54526) (52568 53656)
ESS_1_5	HR		cseqno= (122226 122247)
ESS_1_5	BG		cseqno= (20411 20433)
ISJP_1_2	HU		caseid= (4960422 4960472)
ISSP_1989	IL		v2= (1200385 1200850)
ISSP_1989	AT		v2= (500622 501867) (501649 501982) (501539 501969) (500238 501824) (501173 501928) (501683 501986) (500942 501904) (501345 501943) (501573 501972) (500501 501857) (500439 501850) (501362 501946) (501178 501929) (501694 501987) (500040 501809) (501159 501925) (500420 501846) (501711 501988) (500341 501835) (500734 501878) (501467 501958) (500532 501862) (500946 501905) (501527 501964) (501440 501956) (501750 501991) (501534

Dataset	Country *	Filtering variables	ID variable
			501966) (501502 501961) (500863 501894) (500897 501899) (500973 501907) (501579 501974) (500385 501840) (500048 501811) (501426 501953) (501138 501921) (501662 501984) (501218 501931) (500625 501869) (500811 501887) (501400 501950) (500725 501874) (501433 501954) (500274 501829) (500722 501873) (500314 501834) (501296 501937) (501115 501917) (501807 501997) (501165 501926) (500748 501881) (500709 501872) (501542 501970) (501518 501962) (501792 501996) (500735 501879) (500442 501851) (500820 501889) (500843 501892) (500727 501875) (501285 501935) (500786 501884) (500623 501868) (501381 501948) (500090 501817) (501404 501951) (501784 501993) (501538 501968) (500543 501863) (500054 501812) (500759 501882) (501317 501940) (501370 501947) (500877 501897) (500837 501891) (500512 501858) (500739 501880) (500730 501877) (500519 501861) (501241 501933) (500254 501826) (501064 501915) (500908 501900) (500517 501860) (501329 501941) (501301 501938) (501473 501959) (500372 501837) (501156 501924) (501679 501985) (501287 501936) (500121 501819) (500968 501906) (500421 501847) (500460 501852) (501790 501995) (501435 501955) (501358 501945) (500249 501825) (500499 501856) (500603 501865) (501310 501939) (501395 501949) (500057 501814) (501143 501922) (500438 501849) (501616 501978) (501110 501916) (501758 501992) (501137 501920) (500267 501828) (500933 501903) (501548 501971) (500415 501845) (500061 501815) (501459 501957) (500289 501831) (501647 501981) (501116 501918) (500055 501813) (501268 501934) (501578 501973) (500041 501810) (500729 501876) (500307 501833) (501580 501975) (500377 501838) (501409 501952) (501039 501913) (500812 501888) (500855 501893) (500995 501909) (500621 501866) (500403 501843) (501036 501912) (501653 501983) (501646 501980) (500779 501883) (500366 501836) (500492 501855) (500412 501844) (500378 501839) (500402 501842) (501743 501990) (500516 501859) (500931 501902) (500154 501820) (500835 501890) (500918 501901) (501526 501963) (501789 501994) (501049 501914) (500483 501854) (501357 501944) (501031 501911) (500283 501830) (500793 501885) (501535 501967) (501126

Dataset	Country *	Filtering variables	ID variable
			501919) (500871 501896) (500033 501808) (501638 501979) (500672 501871) (501029 501910) (500882 501898) (501715 501989) (500263 501827) (500232 501823) (500391 501841) (500806 501886) (500864 501895) (501585 501976) (500114 501818) (500984 501908) (501341 501942) (500586 501864) (500087 501816) (501589 501977) (501166 501927) (500430 501848) (500660 501870) (501150 501923) (500195 501822) (500194 501821) (501530 501965) (500302 501832) (501198 501930)
ISSP_1991	SI		v2= (1209996 1209999)
ISSP_1996	BG		v2= (1708109 1708110)
ISSP_1996	RU		v2= (1800994 1801002) (1801357 1801361)
ISSP_1996	LV		v2= (2600392 2600395)
ISSP_1998	BG		v2= (1704343) (1725151) (1734748) (1781314) (1720710) (1711314) (1742121) (1784949) (1741717) (1761717) (1702121) (1722121) (1744950) (1720707) (1710102) (1790709) (1717172) (1797375) (1773334) (1751717) (1762727) (1782728) (1724748) (1761920) (1772930) (1782930) (1712930) (1701920) (1722930) (1725363) (1770710) (1747171) (1753131) (1781920) (1754950) (1720102) (1753334) (1703334) (1755353) (1727678) (1727676) (1721314) (1733639) (1701314) (1740102) (1745363) (1742930) (1762930) (1752930) (1741314) (1792930) (1702930) (1735152) (1713335) (1737171) (1716770) (1735363) (1714748) (1725356) (1745356) (1755356) (1714950)[2]
ISSP_1998	RU		v2= (1800896 1800906) (1801039 1801085)
ISSP_1998	CY		v2= (2900094 2900671) (2900064 2900725) (2900089 2900676) (2900109 2900642) (2900108 2900643) (2900095 2900670) (2900070 2900720) (2900115 2900677) (2900082 2900739) (2900140 2900662) (2900001 2900724) (2900062 2900647) (2900130 2900681) (2900106 2900653) (2900078 2900733) (2900096 2900669) (2900065 2900726) (2900086 2900648) (2900150 2900668) (2900087 2900650) (2900152 2900686) (2900137 2900659) (2900084 2900742) (2900080 2900728) (2900120 2900688) (2900131 2900680) (2900139 2900661) (2900075 2900736) (2900138 2900639) (2900085 2900741) (2900060 2900740) (2900117 2900691) (2900077 2900734) (2900093 2900672) (2900118 2900690) (2900097 2900682) (2900114 2900678) (2900105 2900645) (2900091 2900674)

Dataset	Country *	Filtering variables	ID variable
ISSP_1998	PT		v2= (3000010 3001143) (3000005 3001138) (3000003 3001136) (3000009 3001142) (3000001 3001134) (3000007 3001140) (3000002 3001135) (3000008 3001141) (3000006 3001139) (3000004 3001137)
ISSP_2004	RU		v3= (18001772 18001782)
ISSP_2004	ES		v3= (25000421 25000455)
ISSP_2004	PT		v3= (30001899 30003014)
ISSP_2004	ZA		v3= (40006177 40006696)
ISSP_2006	DO	v3=214	v2= (998 999)
ISSP_2006	FR	v3=250	v2= (1495 1538)
ISSP_2006	AU	v3=36	v2= (11319 12337)
ISSP_2006	IL-JEW	v3=376.1	v2= (430 8430)
ISSP_2006	ZA	v3=710	v2= (877 1984 4616) (2492 4978 6733) (4995 4996 5429)
ISSP_2007	AU	v5=36	v3= (1247 1352 1801 2585)
ISSP_2007	IE	v5=372	v3= (5402 5421) (6513 6523) (6709 6909)
ISSP_2008	IE	v5=372	v3= (5402 5421) (6513 6523)
ISSP_2008	ES	v5=724	v3= (2697 2700) (3269 3272)
ISSP_2009	AU	v5=36	v3= (110380 112214 115102)
ISSP_2009	NO	v5=578	v3= (842 870 1117 1217 1314 1345) (936 1205 1383 1419) (580 1202 1335 1370 1387 1401 1411) (248 1281 1302 1328 1384) (490 1168 1271 1297) (371 530 823 1317 1329) (849 1113 1192) (780 814 1042 1445 1452) (1167 1288) (905 1194) (1208 1379 1447) (757 1464) (595 1393) (909 1193) (157 1215 1405) (543 1282 1366) (1179 1457) (391 1195 1221) (1293 1361) (362 415 553 906 1019 1164 1219 1319) (205 460) (272 1437) (35 404 1262 1336) (46 800) (1280 1309 1347) (845 1357) (900 1130) (1206 1290) (605 1211 1267 1389 1418) (482 1147 1223 1268 1285) (334 1227 1409) (1220 1299) (63 526) (886 1212) (948 999) (381 1269) (874 1162) (527 1102 1203 1374) (93 274 1181) (396 1304) (899 1013 1200) (1209 1451) (306 424 1146) (1170 1339 1386 1433) (175 969 1197) (54 1394 1448) (1250 1455) (931 1214) (12 1107) (934 1175 1342 1416) (33 367) (1187 1382) (368 428)
ISSP_2009	ZA	v5=710	v3= (2735 4691)
ISSP_2009	VE	v5=862	v3= (129 967) (831 994) (727 990) (79 965) (363 976) (779 992) (675 988) (805 993) (233 971) (493 981) (441 979) (389 977) (19 963) (415 978) (701 989) (337 975) (545 983) (311 974) (467 980) (935 998)

Dataset	Country *	Filtering variables	ID variable
			(104 966) (259 972) (909 997) (857 995) (156 968) (623 986) (961 999) (571 984) (597 985) (753 991)
ISSP_2010	DK	v4=208	caseid= (208046975796 208046976531)
ISSP_2010	ZA	v4=710	caseid= (20107100002462 20107100002575)
ISSP_2011	BE-WAL	v4=56	caseid= (20115602004157 20115602004090 20115602005331 20115601999901 20115602004823 20115602003991 20115602002093 20115602002190 20115602002235 20115602002518 20115602002819 20115602000006 20115602000089 20115602000106 20115602000121 20115602000243 20115602000350 20115602001219 20115602001255 20115602001989 20115602010895 20115602007389 20115602010953)
ISSP_2011	ZA	v4=710	caseid= (20110710003536 20110710003814) (20110710003394 20110710003728 20110710003962) (20110710003411 20110710006435)
IVS_1_9	US		s007= (19818400327 19818401457) (19818400416 19818402260) (19818400353 19818401658) (19818400299 19818401170) (19818400282 19818401052) (19818400220 19818401909) (19818400208 19818401718) (19818400285 19818401075) (19818400298 19818401168) (19818400412 19818402211) (19818400164 19818400785) (19818400269 19818400843) (19818400395 19818402009) (19818400338 19818401556) (19818400310 19818401312) (19818400305 19818401233) (19818400359 19818401745) (19818400186 19818401106) (19818400423 19818402307) (19818400216 19818401830) (19818400347 19818401582) (19818400233 19818402118) (19818400223 19818401974) (19818400372 19818401791) (19818400361 19818401755) (19818400185 19818401095) (19818400200 19818401438) (19818400368 19818401773) (19818400288 19818401084) (19818400258 19818400757) (19818400174 19818400851) (19818400253 19818400710) (19818400277 19818400967) (19818400202 19818401491) (19818400389 19818401964) (19818400256 19818400738) (19818400230 19818402108) (19818400240 19818402271) (19818400248 19818400672)

Dataset	Country *	Filtering variables	ID variable
			(19818400417 19818402261) (19818400326 19818401456) (19818400249 19818400674) (19818400366 19818401771) (19818400315 19818401329) (19818400161 19818400639) (19818400390 19818401975) (19818400380 19818401929) (19818400197 19818401390) (19818400245 19818400616) (19818400270 19818400877) (19818400313 19818401324) (19818400171 19818400840) (19818400199 19818401415) (19818400294 19818401112) (19818400213 19818401796) (19818400283 19818401065) (19818400196 19818401364) (19818400357 19818401723) (19818400409 19818402192) (19818400210 19818401750) (19818400275 19818400952) (19818400266 19818400810) (19818400422 19818402305) (19818400244 19818400615) (19818400328 19818401478) (19818400407 19818402172) (19818400322 19818401412) (19818400188 19818401119) (19818400281 19818401049) (19818400385 19818401960) (19818400239 19818402266) (19818400303 19818401215) (19818400195 19818401360) (19818400236 19818402217) (19818400316 19818401330) (19818400284 19818401074) (19818400309 19818401301) (19818400252 19818400696) (19818400304 19818401218) (19818400176 19818400904) (19818400397 19818402037) (19818400241 19818402285) (19818400182 19818401014) (19818400355 19818401676) (19818400172 19818400841) (19818400336 19818401546) (19818400184 19818401047) (19818400414 19818402232) (19818400168 19818400826) (19818400291 19818401104) (19818400183 19818401021) (19818400356 19818401677) (19818400388 19818401963) (19818400321 19818401356) (19818400408 19818402185) (19818400317 19818401334) (19818400369 19818401774) (19818400386 19818401961) (19818400387 19818401962) (19818400334 19818401536) (19818400379 19818401912) (19818400340 19818401560) (19818400323 19818401426) (19818400259 19818400761) (19818400362 19818401756)

Dataset	Country *	Filtering variables	ID variable
			(19818400257 19818400749) (19818400237 19818402238) (19818400260 19818400779) (19818400217 19818401863) (19818400221 19818401945) (19818400384 19818401957) (19818400166 19818400821) (19818400418 19818402278) (19818400225 19818402002) (19818400392 19818401977) (19818400232 19818402114) (19818400293 19818401111) (19818400180 19818400938) (19818400330 19818401512) (19818400211 19818401783) (19818400273 19818400919) (19818400167 19818400824) (19818400170 19818400839) (19818400242 19818402306) (19818400365 19818401768) (19818400374 19818401838) (19818400342 19818401564) (19818400411 19818402208) (19818400373 19818401828) (19818400250 19818400692) (19818400350 19818401630) (19818400209 19818401721) (19818400163 19818400750) (19818400401 19818402123) (19818400344 19818401566) (19818400404 19818402138) (19818400192 19818401342) (19818400274 19818400934) (19818400325 19818401454) (19818400364 19818401766) (19818400421 19818402304) (19818400290 19818401102) (19818400399 19818402098) (19818400165 19818400789) (19818400371 19818401787) (19818400393 19818401988) (19818400287 19818401080) (19818400201 19818401490) (19818400160 19818400637) (19818400251 19818400695) (19818400234 19818402139) (19818400410 19818402207) (19818400335 19818401544) (19818400329 19818401508) (19818400175 19818400895) (19818400295 19818401143) (19818400265 19818400808) (19818400179 19818400930) (19818400178 19818400920) (19818400383 19818401955) (19818400300 19818401182) (19818400214 19818401800) (19818400382 19818401954) (19818400306 19818401250) (19818400348 19818401587) (19818400405 19818402141) (19818400204 19818401584) (19818400324 19818401436) (19818400181 19818400959) (19818400235 19818402164) (19818400312 19818401323)

Dataset	Country *	Filtering variables	ID variable
			(19818400419 19818402287) (19818400278 19818401037) (19818400367 19818401772) (19818400292 19818401109) (19818400318 19818401347) (19818400255 19818400724) (19818400296 19818401146) (19818400219 19818401906) (19818400222 19818401966) (19818400396 19818402030) (19818400218 19818401884) (19818400370 19818401776) (19818400243 19818402312) (19818400264 19818400805) (19818400391 19818401976) (19818400402 19818402125) (19818400207 19818401708) (19818400403 19818402134) (19818400363 19818401765) (19818400203 19818401532) (19818400360 19818401746) (19818400289 19818401085) (19818400351 19818401631) (19818400345 19818401567) (19818400169 19818400837) (19818400346 19818401569) (19818400406 19818402162) (19818400301 19818401196) (19818400376 19818401878) (19818400262 19818400795) (19818400377 19818401885) (19818400247 19818400657) (19818400352 19818401634) (19818400320 19818401354) (19818400231 19818402113) (19818400319 19818401348) (19818400271 19818400880) (19818400400 19818402099) (19818400193 19818401345) (19818400276 19818400954) (19818400339 19818401559) (19818400228 19818402043) (19818400378 19818401905) (19818400297 19818401160) (19818400226 19818402008) (19818400263 19818400804) (19818400268 19818400832) (19818400375 19818401839) (19818400333 19818401530) (19818400254 19818400712) (19818400358 19818401737) (19818400337 19818401555) (19818400314 19818401327) (19818400302 19818401204) (19818400343 19818401565) (19818400215 19818401801) (19818400238 19818402241) (19818400415 19818402242) (19818400272 19818400889) (19818400332 19818401529) (19818400189 19818401195) (19818400229 19818402104) (19818400307 19818401293) (19818400280 19818401046) (19818400354 19818401674) (19818400308 19818401299)

Dataset	Country *	Filtering variables	ID variable
			(19818400279 19818401040) (19818400331 19818401525) (19818400190 19818401269) (19818400267 19818400831) (19818400413 19818402220) (19818400206 19818401637) (19818400311 19818401314) (19818400261 19818400782) (19818400341 19818401563) (19818400394 19818402007) (19818400246 19818400640) (19818400381 19818401940) (19818400398 19818402058) (19818400177 19818400907) (19818400227 19818402012) (19818400286 19818401078) (19818400173 19818400850) (19818400420 19818402294) (19818400205 19818401586) (19818400212 19818401790) (19818400224 19818401981) (19818400349 19818401621) (19818400194 19818401353) (19818400162 19818400643) (19818400198 19818401404) (19818400191 19818401276) (19818400187 19818401108)
IVS_1_9	BE		s007= (199005601684 199005602527)
IVS_1_9	DE-W		s007= (199027600162 199027603045)
IVS_1_9	BG		s007= (200810001230 200810001324)
IVS_1_9	LV		s007= (200842800009 200842801286) (200842800235 200842801422) (200842800005 200842801266) (200842800022 200842801307) (200842800014 200842801291) (200842800016 200842801293) (200842800234 200842801421) (200842800239 200842801426) (200842800002 200842801263) (200842800015 200842801292) (200842800004 200842801265) (200842800013 200842801290) (200842800334 200842801434) (200842800012 200842801289) (200842800001 200842801262) (200842800006 200842801267) (200842800199 200842801416) (200842800003 200842801264)

Dataset	Country *	Filtering variables	ID variable
IVS_1_9	JP		s007= (5875 5876) (5214 5215) (4782 4783) (4784 4785 4786) (5564 5565) (5316 5317) (5551 5552) (4872 4873) (5251 5252 5253) (5134 5135) (4705 4706) (5354 5355) (5817 5818) (5306 5307) (5111 5112 5113) (5156 5157) (5675 5676) (4754 4755) (4790 4791) (5233 5234) (4709 4710) (5841 5842) (5847 5848) (5593 5594) (5672 5673) (5324 5325) (5483 5484) (5812 5813 5814) (5240 5241) (4711 4712) (4716 4717) (4895 4896 4897) (5652 5653 5654) (4874 4875) (5128 5129) (4719 4720) (4757 4758) (5598 5599) (4855 4856) (4780 4781) (4866 4867) (5130 5131) (5193 5194) (4799 4800) (4770 4771) (4931 4932) (5203 5204) (5531 5532) (5879 5880) (4707 4708) (5249 5250) (5330 5331 5332) (5717 5718) (5798 5799) (5310 5311) (4774 4775) (5032 5033) (4703 4704) (5121 5122) (5036 5037) (5100 5101) (5270 5271) (5826 5827) (5628 5629) (4787 4788) (5103 5104) (5206 5207) (5072 5073 5074) (5006 5007) (5900 5901 5902) (5066 5067 5068) (4994 4995 4996) (4721 4722) (5479 5480) (5394 5395) (5605 5606 5607) (5210 5211) (5372 5373) (5548 5549) (4759 4760) (5418 5419 5420) (5230 5231) (5384 5385) (5298 5299) (5402 5403) (5769 5770 5771) (4980 4981) (5575 5576) (4963 4964 4965) (5641 5642)
IVS_1_9	NG		s007= (24448 24824)
IVS_1_9	BG		s007= (46570 46572) (45737 45738 45739)
IVS_1_9	BY		s007= (47454 48074) (47453 48073) (47450 48070) (47455 48075) (47449 48069) (47451 48071) (47452 48072)
IVS_1_9	AZ		s007= (36201 37855)

Dataset	Country *	Filtering variables	ID variable
IVS_1_9	MX		s007= (73812 75322) (73871 75381) (73801 75311) (73784 75294) (73646 75156) (75681 75686) (73739 75249) (73766 75276) (73765 75275) (73764 75274) (73778 75288) (73638 75148) (73825 75335) (73992 75502) (73798 75308) (73585 75093) (73683 75193) (73411 74922) (73738 75248) (73767 75277) (73409 74920) (73840 75350) (73568 75101) (73581 75089) (73532 75042) (73802 75312) (73639 75149) (74048 75558) (73691 75201) (73693 75203) (73813 75323) (73464 74974) (73668 75178) (73702 75212) (73586 75094) (73804 75314) (73811 75321) (74998 73488) (73771 75281) (73864 75374) (73822 75332) (73446 74956) (73496 75006) (73841 75351) (73622 75132) (73590 75098) (73728 75238) (73439 74929) (73775 75285) (73589 75097) (73815 75325) (73623 75133) (73547 75057) (74126 75636) (74084 75594) (73376 74886) (73882 75392) (73578 75086) (73826 75336) (74144 75654) (74996 73486) (73498 75008) (73452 74962) (75000 73490) (74990 73480) (73791 75301) (73497 75007) (73618 75128) (73526 75036) (73679 75189) (73897 75407) (75682 75687) (74091 75601) (73657 75167) (73838 75348) (74076 75586) (73521 75031) (73371 74881) (73927 75437) (73943 75453) (73455 74965) (73456 74966) (74121 75631) (74988 73478) (73528 75038) (73705 75215) (73669 75179) (74981 73471) (74980 73470) (73633 75143) (73959 75469) (73629 75139) (73494 75004) (73792 75302) (73570 75078) (73640 75150) (73626 75136) (73695 75205) (73712 75222) (73751 75261) (73760 75270) (73777 75287) (73789 75299) (73463 74973) (73501 75011) (73799 75309) (73580 75088) (73461 74971) (73392 74902) (73418 74928) (73503 75013) (73403 74914) (73630 75140) (73708 75218) (75377 73867) (74170 75680 75685) (73981 75491) (74131 75641) (73506 75016) (74993 73483) (73685 75195) (74982 73472) (75679 75684) (73400 74911) (73696 75206) (73690 75200) (73687 75197) (73507 75017) (73508 75018) (73502 75012) (73401 74912) (73654 75164) (73621 75131) (73467 74977) (74991 73481) (74145 75655) (73445 74955) (73870 75380) (73793 75303) (74086 75596) (74071 75581) (74128 75638) (73525 75035) (73587 75095) (74123 75633) (74097 75607) (73527 75037) (73816 75326) (73539 75049) (73965 75475) (75678 75683) (74163 75673) (73584 75092)

Dataset	Country *	Filtering variables	ID variable
			(73828 75338) (73842 75352) (74078 75588) (73398 74909) (73892 75402) (73499 75009) (73579 75087) (74100 75610) (73555 75065) (74983 73473) (73511 75021) (73711 75221) (73904 75414) (74136 75646) (73713 75223) (73672 75182) (73947 75457) (73583 75091) (73545 75055) (73717 75227) (73520 75030) (74094 75604) (73536 75046) (73512 75023) (73866 75376) (73675 75185) (74987 73477) (73857 75367) (73852 75362) (73845 75355) (73951 75461) (73863 75373) (73569 75102) (73862 75372) (73843 75353) (73849 75359) (73865 75375) (74154 75664) (74062 75572) (73688 75198) (73844 75354) (73805 75315) (73848 75358) (73859 75369) (73850 75360) (74108 75618) (73551 75061) (74103 75613) (74182 75697) (73855 75365) (73846 75356) (73567 75077) (73535 75045) (74099 75609) (74073 75583) (73915 75425) (73673 75183) (73559 75069) (73896 75406) (74984 73474) (73723 75233) (74074 75584) (73413 74924) (74117 75627) (74107 75617) (73427 74942) (74137 75647) (73720 75230) (74135 75645) (74069 75579) (73741 75251) (73714 75224) (73754 75264) (74159 75669) (74162 75672) (73735 75245) (73747 75257) (73750 75260) (73749 75259) (73954 75464) (74118 75628) (73552 75062) (73957 75467) (73875 75385) (74130 75640) (73441 74931) (73710 75220) (73678 75188) (73557 75067) (73649 75159) (74183 75698) (74157 75667) (73406 74917) (73645 75155) (73752 75262) (75002 73492) (73454 74964) (74166 75676) (74155 75665) (73562 75072) (73692 75202) (73436 74951) (73531 75041) (74072 75582) (74111 75621) (73662 75172) (73648 75158) (73549 75059) (73419 74934) (73908 75418) (74181 75696) (73650 75160) (74122 75632) (73753 75263) (73429 74944)
IVS_1_9	BD		s007= (41911 41923)
IVS_1_9	AM		s007= (44310 44316) (44049 44340) (44044 44303) (44041 44051)
IVS_1_9	NZ		s007= (77108 77748)
IVS_1_9	RO		s007= (87012 87036) (87007 87009) (87014 87041) (87013 87033)
IVS_1_9	SI		s007= (91548 91616) (91546 91614) (91547 91615)
IVS_1_9	BR		s007= (44819 45648) (45627 45641) (45268 45655) (45365 45659) (44758 45642) (44521 44527) (44522 44528) (45619 45640) (45072 45652) (45528 45664) (44946 45662) (45585 45639) (45393 45660) (45274

Dataset	Country *	Filtering variables	ID variable
			45656) (45404 45661) (44542 45643) (44546 45644) (44716 45666) (44578 45665) (45261 45654) (45350 45658) (45316 45657) (44962 45651) (45574 45667) (44763 45647) (45189 45653) (45522 45638) (45198 45637) (44519 44525) (44520 44526) (45058 45636) (44631 45645) (44524 44530) (44958 45650) (44915 45649) (44523 44529) (45438 45663) (44732 45646) (44810 45635)
IVS_1_9	UA		s007= (102639 102651)
IVS_1_9	DZ		s007= (115263 115264) (115405 115549) (115395 115593) (115598 115704) (115402 115579) (115408 115543) (115406 115548) (115044 115045) (115379 115525) (115586 115706) (115378 115594) (115587 115705) (115397 115589)
IVS_1_9	CL		s007= (121219 121316) (120933 121375)
IVS_1_9	IN		s007= (124663 124696) (124116 124188) (124714 124737) (124698 124738) (124713 124736) (124117 124189) (124162 124187)
IVS_1_9	ID		s007= (124939 125140 125341 125542)
IVS_1_9	IQ		s007= (129940 129942) (129879 129880) (129939 129941) (129977 129979) (129980 129981) (129911 129913) (129912 129914)
IVS_1_9	KR		s007= (134757 134776)
IVS_1_9	BD		s007= (118005 118107) (118006 118108) (118007 118109)
IVS_1_9	NG		s007= (141880 141882) (142289 142291) (141979 141981) (142169 142171)
IVS_1_9	PH		s007= (147924 147942)
IVS_1_9	SA		s007= (149058 149395)
IVS_1_9	SG		s007= (150564 150778)
IVS_1_9	AL		s007= (113566 113567)
IVS_1_9	TZ		s007= (168068 168069)
IVS_1_9	ME		s007= (172696 172695)
IVS_1_9	CY		s007= (190497 191002) (190499 191004) (190494 190998) (190505 191009) (190492 190996) (190509 191013) (190507 191011) (190502 191007) (190495 191000) (190508 191012) (190506 191010) (190501 191006) (190504 191008) (190496 191001) (190498 191003) (190493 190997) (190500 191005)

Dataset	Country *	Filtering variables	ID variable
IVS_1_9	ET		s007= (191747 192408) (191750 192411) (192489 192497) (191388 192327) (191422 192511) (191708 192369) (191753 192414) (191707 192368) (191749 192410) (191717 192378) (192066 192469) (191711 192372) (191712 192373) (191722 192383) (191703 192364) (192063 192466) (192029 192458) (191856 192435) (192032 192461) (191780 192427) (192061 192464) (191714 192375) (191036 192229) (191694 192355) (191719 192380) (191032 192225) (191792 192431) (191790 192429) (192026 192455) (192028 192457) (192027 192456) (191857 192436) (191858 192437) (191860 192439) (191855 192434) (191699 192360) (191342 192309) (192030 192459) (191794 192433) (191793 192432) (191791 192430) (191859 192438) (191415 192339 192504) (192031 192460) (191347 192314) (192025 192454) (192208 192480 192481) (192125 192471) (191809 192498) (191345 192312) (191349 192316) (191344 192311) (192207 192478 192479) (191263 192254) (191268 192259) (191723 192384) (191701 192362) (191351 192318) (191715 192376) (191288 192279) (191283 192274) (191022 192216) (191419 192508) (191706 192367) (192064 192467) (191746 192407) (191287 192278) (191704 192365) (191696 192357) (191339 192306) (191383 192322) (191716 192377) (191724 192385) (191310 192296 192495) (191015 192210) (191418 192507) (191890 192440) (191698 192359) (191346 192313) (191343 192310) (191423 192512) (192062 192465) (191034 192227) (191303 192293) (191020 192214) (191730 192391) (192067 192470) (191390 192329) (191039 192232) (191697 192358) (191720 192381) (191033 192226) (191279 192270) (191354 192321) (191350 192317) (191556 191677) (191348 192315) (191353 192320) (191352 192319) (191019 192213) (191713 192374) (191017 192211) (191014 192209) (191021 192215) (191732 192393) (191298 192288) (191745 192406) (191300 192290) (191743 192404) (191710 192371) (191285 192276) (191294 192285) (191278 192269) (191725 192386) (191742 192403) (191270 192261) (191047 192239) (191702 192363) (191059 192251) (192205 192476 192477) (191737 192398) (191340 192307) (192021 192450) (191311 192297) (191335 192303) (191264 192255) (192022 192451) (191314 192299) (191312 192298)

Dataset	Country *	Filtering variables	ID variable
			(192065 192468) (191297 192287) (191291 192282) (191315 192300) (191744 192405) (191050 192242) (191309 192494) (191700 192361) (192033 192462) (191040 191041 192233) (192018 192447) (191734 192395) (191755 192416) (191729 192390) (191045 192237) (191048 192240) (191428 192344) (191043 192235) (191695 192356) (192019 192448) (191023 192217) (191728 192389) (191266 192257) (191271 192262) (191046 192238) (191748 192409) (192131 192483) (192034 192463) (191265 192256) (191049 192241) (192023 192452) (191299 192289) (191692 192353) (192132 192484) (191733 192394) (191037 192230) (191386 192325) (191304 192294) (191269 192260) (191413 192337) (191411 192335) (191296 192286) (191731 192392) (191029 192222) (191693 192354) (192340 192505) (191025 192219) (191409 192333) (191410 192334) (191408 192332) (192020 192449) (191338 192305) (191727 192388) (191781 192428) (191044 192236) (191042 192234) (191739 192400) (191759 192420) (191274 192265) (191306 192491) (191520 192351) (191334 192302) (191754 192415) (191018 192212) (191307 192492) (191978 192444 192445) (192135 192487) (192133 192485) (191709 192370) (191293 192284) (191414 192338 192503) (192130 192482) (191429 192345) (191038 192231) (191433 192349) (191056 192248) (191035 192228) (191277 192268) (191336 192304) (191779 192426) (191774 192422) (191718 192379) (191341 192308) (191421 192510) (191775 192423) (191052 192244) (191055 192247) (191387 192326) (191942 192441) (191389 192328) (191756 192417) (191305 192295) (191431 192347) (191057 192249) (191430 192346) (191051 192243) (191135 192490) (192024 192453) (191024 192218) (191030 192223) (191031 192224) (191432 192348) (191308 192493) (191741 192402) (191027 192221) (191427 192343) (191058 192250) (191752 192413) (191757 192418) (191751 192412) (191262 192253) (191417 192341 192506) (191735 192396) (191273 192264) (191301 192291) (191302 192292) (191384 192323) (191276 192267) (192204 192474 192475) (191740 192401) (191738 192399) (191284 192275) (191275 192266) (191267 192258) (191758 192419) (191726 192387) (191721 192382) (191333 192301) (191426 192342) (191973

Dataset	Country *	Filtering variables	ID variable
			192443) (192201 192472 192473) (191280 192271) (191282 192273) (191286 192277) (191289 192280)
IVS_1_9	IN		s007= (200030 200947) (200044 200821) (200895 200896) (200875 200876) (200835 200836) (200833 200834) (200793 200794) (200872 200873) (201026 201027) (200804 200805)
IVS_1_9	ID		s007= (202384 202386) (202382 202387) (203068 203071) (203372 203374) (203377 203379)

Dataset	Country *	Filtering variables	ID variable
IVS_1_9	KR		s007= (212550 213166) (213002 213422) (212977 213453 213468 213486) (212923 213170) (212920 213209) (212612 213248) (212544 213297) (212637 213157) (212649 213163) (213207 213238) (212757 213304) (213397 213430) (212643 213086) (212674 213175 213296) (213494 213546) (212519 212700) (213329 213351) (213063 213160) (212992 213456 213488) (212530 212715 213208 213311) (212940 213126) (213445 213459 213473 213474 213489 213500) (212552 212713 213308) (213444 213498) (212561 213135) (212703 212750) (213357 213628) (213240 213255) (213319 213325) (212743 213224) (212996 213457) (212557 213087) (212840 213055) (212777 213049) (212809 213052) (213239 213253) (212540 213204) (212722 213215) (213000 213458 213472) (212820 213046) (212985 213455) (213450 213550) (213322 213324) (212580 213301) (213461 213478 213520) (213517 213540) (213479 213535) (212657 213252) (212678 213245) (213477 213509) (212582 213130) (212658 213176) (212527 212716) (212775 213048) (213059 213684) (212749 213314) (212932 213200) (213226 213241) (213343 213616) (213109 213169) (213079 213188) (213118 213164) (213332 213342) (212593 213300) (213355 213631) (213114 213223) (213346 213563) (212981 213487) (212652 213155) (212949 213089) (212975 213483 213484) (212606 213217) (212578 213090) (212588 212711) (213376 213385) (213106 213210) (213491 213507) (213081 213138) (212576 213294) (213066 213313) (213133 213144) (212789 213032) (213356 213629) (212702 212756 213154) (213060 213697) (213124 213184) (212974 213439 213452 213467) (212988 213441) (212699 213178) (212676 213134) (213009 213423) (213446 213475 213502) (213080 213139) (213258 213287) (212925 213201) (212514 212615) (213405 213433) (212978 213437 213469 213470) (212592 213290) (213399 213432) (213462 213522 213541) (213019 213424) (213443 213497) (212520 212730) (212668 213128) (213094 213198) (212575 212712 213299) (213061 213702) (213493 213518) (212754 213165) (212565 213084) (213097 213243) (212584 213291) (212639 213156) (212581 213292) (213396 213429) (212665 212705) (212600 213218) (212682 213127) (212973 213482) (212632

Dataset	Country *	Filtering variables	ID variable
			213158) (213447 213463 213523) (213367 213426) (212744 213213) (213115 213168) (212669 213083) (212803 213051) (212709 213162) (212651 212707 213307) (213464 213525) (212671 213246) (213490 213506) (212618 212710) (212787 213050) (213481 213555) (212771 213047) (212515 212617 213289) (212608 213315) (213088 213149) (213460 213510) (212541 212714) (212638 213256) (212917 213177) (213476 213504) (213117 213167) (212976 213485) (212609 213173) (212620 213129) (213492 213512) (212660 212706) (212885 213360) (212729 213316) (212562 213205) (212746 213244) (212517 212677) (212733 213085) (212521 212737) (213451 213480 213554) (212991 213442) (213334 213341) (213340 213350) (213495 213549) (212731 213171) (212922 213180)
IVS_1_9	RW		s007= (231035 231088)
IVS_1_9	VN		s007= (232104 232105)
IVS_1_9	TH		s007= (240224 240242) (240234 240252) (240215 240267) (240221 240239) (240216 240268) (240225 240243) (240203 240255) (240229 240247) (240235 240253) (240213 240265) (240226 240244) (240227 240245) (240210 240262) (240218 240236) (240205

Dataset	Country *	Filtering variables	ID variable
			240257) (240206 240258) (240228 240246)
IVS_1_9	ZM		s007= (254178 254298) (253112 254296) (253109 254293) (253113 254297) (253209 254304) (253111 254295) (253208 254303) (254180 254300) (253110 254294) (254182 254302) (254179 254181 254299 254301)
LB_1995	AR	pais=1	enc= (648 737)
LB_1996	PA	pais=13	enc= (177 820) (599 890) (196 819) (707 892) (193 816) (307 791) (712 897) (195 818) (299 783) (202 805) (527 932) (861 977) (837 953) (852 969) (189 812) (203 806) (503 908) (200 803) (859 975) (870 986) (65 885) (204 807) (598 889) (716 901) (729 902) (511 916) (709 894) (526 931) (504 909) (515 920) (190 813) (731 904) (708 893) (732 905) (298 782) (305 789) (518 923) (514 919) (827 943) (862 978) (192 815) (830 946) (512 917) (600 891) (715 900) (199 802) (529 934) (730 903) (520 925) (507 912) (509 914) (825 941) (532 937) (849 966) (523 928) (531 936) (302 786) (191 814) (857 974) (303 787) (848 965) (835 951) (597 888) (519 924) (710 895) (38 858) (525 930) (871 987) (596 887) (201 804) (860 976) (524 929) (522 927) (508 913) (194 817) (864 980) (517 922) (513 918) (63 883) (856 973) (62 882) (301 785) (846 962) (855 972) (844 961) (306 790) (839 955) (528 933) (506 911) (300 784) (823 939) (873 989) (533 938) (834 950) (829 945) (206 809) (309 793) (826 942) (521 926) (66 886) (313 797) (865 981) (866 982) (64 884) (836 952) (311 795) (845 960) (843 959) (316 800) (833 949) (61 881) (842 958) (831 947) (381 821) (872 988) (297 781) (867 983) (838 954) (505 910) (308 792) (847 963) (832 948) (840 956) (510 915) (711 896) (841 957) (733 906) (874 990) (304 788) (314 798) (714 899) (312 796) (187 810) (58 878) (60 880) (516 921) (828 944) (499 822) (188 811) (530 935) (59 879) (57 877) (850 967) (868 984) (869 985) (851 968) (315 799) (205 808) (875 991) (713 898) (824 940) (876 992) (853 970) (854 971) (310 794) (175 801) (734 907) (863 979)
LB_1996	VE	pais=17	enc= (738 739) (1024 1025) (932 940)
LB_1996	EC	pais=7	enc= (385 390)

Dataset	**Country** *	**Filtering variables**	**ID variable**
LB_1996	SV	pais=8	enc= (56 57)
LB_1997	PA	idenpa=13	numentre= (168 413) (267 690)
LB_1997	CR	idenpa=5	numentre= (882 891) (880 889)
LB_1997	EC	idenpa=7	numentre= (971 1018) (755 760) (1016 1020) (1007 1019)
LB_1998	EC	idenpa=7	numentre= (693 1163) (692 1162) (694 1164) (556 560) (554 558) (696 1166) (700 1170) (699 1169) (697 1167)
LB_2000	HN	idenpa=10	numentre= (565 566)
LB_2000	NI	idenpa=12	numentre= (20 326) (137 138) (104 594) (113 598)
LB_2000	BO	idenpa=2	numentre= (3028 5028)
LB_2000	EC	idenpa=7	numentre= (231 1004) (186 959) (128 901) (240 1013) (95 868) (152 925) (149 922) (31 804 1166) (357 1130) (305 1078) (261 1034) (279 1052) (203 976) (197 970) (289 1062) (80 853) (6 779 1141) (88 861) (290 1063) (294 1067) (13 786 1148) (67 840) (165 938) (28 801 1163) (61 834 1196) (87 860) (177 950) (56 829 1191) (162 935) (160 933) (166 939) (275 1048) (48 821 1183) (103 876) (142 915) (148 921) (342 1115) (340 1113) (41 814 1176) (141 914) (145 918) (150 923) (129 902) (64 837 1199) (341 1114) (299 1072) (163 936) (78 851) (343 1116) (346 1119) (359 1132) (194 967) (259 1032) (205 978) (43 816 1178) (138 911) (207 980) (242 1015) (86 859) (24 797 1159) (125 898) (126 899) (271 1044) (111 884) (153 926) (263 1036) (284 1057) (58 831 1193) (30 803 1165) (59 832 1194) (316 1089) (303 1076) (29 802 1164) (35 808 1170) (50 823 1185) (27 800 1162) (224 997) (287 1060) (19 792 1154) (22 795 1157) (310 1083) (312 1085) (307 1080) (350 1123) (3 776 1138) (52 825 1187) (123 896) (114 887) (347 1120) (20 793 1155) (221 994) (360 1133) (790 1152) (278 1051) (200 973) (151 924) (143 916) (112 885) (140 913) (65 838 1200) (68 841) (358 1131) (329 1102) (84 857) (250 1023) (16 789 1151) (201 974) (119 892) (118 891) (247 1020) (208 981) (55 828 1190) (336 1109) (245 1018) (265 1038) (146 919) (144 917) (39 812 1174) (45 818 1180) (44 817 1179) (318 1091) (309 1082) (215 988) (345 1118) (291 1064) (191 964) (130 903) (285 1058) (63 836 1198) (277 1050) (154 927) (176 949) (25 798 1160) (262 1035) (139 912) (288 1061) (117 890) (137 910) (281 1054) (228 1001) (156 929) (333 1106) (253 1026) (36 809 1171) (83 856) (15 788 1150) (175 948) (206

Dataset	Country *	Filtering variables	ID variable
			979) (314 1087) (348 1121) (302 1075) (109 882) (108 881) (353 1126) (244 1017) (254 1027) (37 810 1172) (344 1117) (209 982) (164 937) (79 852) (81 854) (246 1019) (74 847) (106 879) (258 1031) (252 1025) (355 1128) (264 1037) (326 1099) (34 807 1169) (72 845) (11 784 1146) (272 1045) (18 791 1153) (349 1122) (38 811 1173) (181 954) (196 969) (26 799 1161) (14 787 1149) (76 849) (21 794 1156) (135 908) (60 833 1195) (330 1103) (356 1129) (256 1029) (51 824 1186) (193 966) (260 1033) (73 846) (157 930) (213 986) (270 1043) (8 781 1143) (85 858) (331 1104) (225 998) (116 889) (354 1127) (33 806 1168) (161 934) (132 905) (77 850) (62 835 1197) (66 839) (71 844) (202 975) (23 796 1158) (120 893) (97 870) (179 952) (217 990) (282 1055) (75 848) (82 855) (121 894) (49 822 1184) (319 1092) (304 1077) (311 1084) (32 805 1167) (286 1059) (292 1065) (94 867) (214 987) (255 1028) (237 1010) (198 971) (227 1000) (5 778 1140) (298 1071) (47 820 1182) (134 907) (313 1086) (293 1066) (7 780 1142) (96 869) (297 1070) (155 928) (170 943) (171 944) (199 972) (104 877) (296 1069) (1 774 1136) (337 1110) (147 920) (308 1081) (320 1093) (216 989) (322 1095) (57 830 1192) (159 932) (232 1005) (249 1022) (110 883) (12 785 1147) (124 897) (219 992) (325 1098) (295 1068) (361 1134) (127 900) (91 864) (69 842) (113 886) (105 878) (321 1094) (243 1016) (274 1047) (4 777 1139) (210 983) (268 1041) (107 880) (351 1124) (323 1096) (93 866) (283 1056) (220 993) (241 1014) (212 985) (324 1097) (54 827 1189) (167 940) (172 945) (195 968) (9 782 1144) (338 1111) (335 1108) (332 1105) (46 819 1181) (276 1049) (315 1088) (239 1012) (238 1011) (334 1107) (42 815 1177) (115 888) (92 865) (133 906) (98 871) (248 1021) (266 1039) (257 1030) (280 1053) (273 1046) (40 813 1175) (173 946) (158 931) (218 991) (70 843) (234 1007) (10 783 1145) (251 1024) (101 874) (178 951) (100 873) (306 1079) (99 872) (2 775 1137) (362 1135) (136 909) (269 1042) (226 999) (222 995) (211 984) (204 977) (339 1112) (301 1074) (300 1073) (102 875) (223 996) (328 1101)
LB_2001	CR	idenpa=5	numentre= (247 395)
LB_2001	EC	idenpa=7	numentre= (1179 1180)
LB_2001	GT	idenpa=9	numentre= (11 1079)

Dataset	Country *	Filtering variables	ID variable
LB_2003	PA	idenpa=13	numentre= (990 1023)
LB_2003	BO	idenpa=2	numentre= (634 713)
LB_2003	CR	idenpa=5	numentre= (893 1008)
LB_2003	CL	idenpa=6	numentre= (200 584)
LB_2004	AR	idenpa=1	numentre= (17 680) (10 673) (15 678) (19 682) (13 676) (12 675) (11 674) (14 677) (20 683) (16 679) (18 681) (22 1199) (9 672) (21 684)
LB_2004	PA	idenpa=13	numentre= (948 1029)
LB_2004	SV	idenpa=8	numentre= (809 1040)
LB_2005	SV	idenpa=8	numentre= (321 502)
LB_2005	GT	idenpa=9	numentre= (403 430)
LB_2006	UY	idenpa=16	numentre= (593 596)
LB_2006	VE	idenpa=17	numentre= (20 1000) (491 997)
LB_2006	BO	idenpa=2	numentre= (179 718)
LB_2008	AR	idenpa=1	numentre= (99 613)
LB_2008	NI	idenpa=12	numentre= (695 825) (507 767) (692 822) (437 557) (698 828) (509 769) (693 823) (503 763) (436 556) (508 768) (697 827) (696 826) (700 830) (439 559) (510 770) (502 762) (505 765) (504 764)
LB_2008	UY	idenpa=16	numentre= (331 436) (518 576) (225 320) (226 319) (612 658) (411 1184) (227 318) (522 523) (7 106) (981 984)
LB_2009	MX	idenpa=11	numentre= (641 671)
LITS_1	SK		id= (13791 13801)
LITS_1	TR		id= (15564 15729) (15563 15615)
LITS_1	AM		id= (17003 17008) (17764 17778)
LITS_1	KG		id= (22542 22802) (22820 22922) (22683 22817) (22819 22921) (22543 22803) (22684 22818) (22881 22979)
LITS_1	TJ		id= (27262 28002) (27278 28001)
LITS_1	MD		id= (8989 8993)
LITS_2	GE		serialid= (63040011 63040001)
NBB_1_6	EE		(863 864)[3]
PPE7N	AT		(1568 1762) (483 1736) (1033 1748) (535 1737) (1510 1758) (438 1734) (152 1722) (1299 1753) (1509 1757) (693 1741) (237 1728) (1539 1761) (365 1733) (552 1739) (914 1744) (325 1730) (54 1720) (1313 1754) (1513 1759) (548 1738) (1531 1760) (182 1725) (181 1724) (179 1723) (1317 1755) (1218 1751)

* Country codes according to the list provided by the International Organization for Standardization (http://www.iso.org/iso/country_codes.htm)

[1] ID 10080 occurs twice in the data set, but the other case is not a duplicate of the mentioned pair. [2] All mentioned values of variable v2 for ISSP 1998, Bulgaria are non-unique and indicate both suspicious cases . [3] There is no original variable assigning respondent/case ID; we use an SPSS system variable $CASENUM instead.