



3MC International Conference 2016

Chicago, July 27, 2016

Survey data harmonization and the quality of data documentation in cross-national surveys

Marta Kołczyńska

The Ohio State University, Polish Academy of Sciences

Matthew Schoene

Davidson College



Evaluating Data Quality

Quality indicators based on:

- Data
- Data–documentation (in)consistencies
- **Documentation**



Purpose

Describe and analyze the variation in quality of cross-national surveys.

Survey selection criterion for users of the harmonized data.

Motivate improvement in survey documentation.



Standards for Describing the Survey Process

Crucial stages of the survey process:

- Sample and sampling,
- Response,
- Preparation of the questionnaire,
- Testing of the questionnaire,
- Post-hoc control of fieldwork.

(De Leeuw, Hox, and Dillman 2008; Groves et al. 2009; Biemer 2010)



Basic Quality Indicators

Availability of sufficient information in the available survey documentation on:

- Sampling
- Response rate
- Questionnaire translation method
- Pretest
- Fieldwork control

Describe survey quality as reflected in documentation.



Basic Quality Indicators

Set of 5 dichotomous variables:

{ 1 if sufficient information is available
0 otherwise

Unit of analysis:

national survey = project*country/territory*year



What is documentation?

Documentation = all we could find on the internet about how the survey was carried out:

study descriptions, technical reports, methodological reports, codebooks, inside questionnaires, on-line text.



What is sufficient?

Sampling: possible reconstruction of the sampling process and selection at all stages.

Response rate: value or information necessary to calculate.

Translation: indicated method other than translation by the survey team.

Pretest: indication of pretesting.

Fieldwork control: indication of control.



Results

Variation in quality indicators in three not entirely independent dimensions:

Project

Country

Year

Project	N national surveys	Sampling	Response rate	Translation	Pretest	Fieldwork control
		% of surveys with sufficient information				
ABS	30	70	33	13	27	43
AFB	66	100	77	0	0	0
AMB	92	97	0	0	0	0
ARB	16	100	6	0	0	0
ASES	18	94	0	0	0	6
CB	12	100	25	0	0	0
CDCEE	27	11	0	0	0	0
CNEP	8	0	13	0	0	0
EB	152	100	0	77	0	0
EQLS	93	100	99	100	99	70
ESS	146	99	100	100	100	100
EVS	128	95	62	48	49	56
ISJP	21	81	81	5	0	0
ISSP	363	92	89	70	32	73
LB	260	90	0	0	0	2
LITS	64	55	100	55	100	0
NBB	18	100	61	0	0	22
PA2	3	100	0	0	0	0
PA8NS	8	100	25	0	0	0
PPE7N	7	100	71	0	0	0
VPCPCE	5	100	0	100	100	0
WVS	184	70	41	20	29	44
Total	1721	89	51	44	32	38

Project	N waves	Sampling		Response rate		Translation		Pretest		Fieldwork control	
		first wave	last wave	first wave	last wave	first wave	last wave	first wave	last wave	first wave	last wave
ABS	3	100	89	88	0	0	44	0	89	50	78
AFB	4	100	100	50	95	0	0	0	0	0	0
AMB	5	100	100	0	0	0	0	0	0	0	0
ARB	2	100	100	17	0	0	0	0	0	0	0
CB	4	100	100	0	100	0	0	0	0	0	0
CDCEE	2	25	0	0	0	0	0	0	0	0	0
EB	7	100	100	0	0	0	100	0	0	0	0
EQLS	3	100	100	100	100	100	100	100	100	0	100
ESS	6	100	100	100	100	100	100	100	100	100	100
EVS	4	100	90	0	100	0	96	0	92	0	94
ISJP	2	86	71	93	57	14	0	0	0	0	0
ISSP	13	83	94	83	90	0	42	0	42	0	68
LB	15	100	95	0	0	0	0	0	0	0	0
LITS	2	0	100	100	100	0	100	100	100	0	0
NBB	6	100	100	100	0	0	0	0	0	33	33
WVS	5	50	93	0	59	0	35	0	50	0	57
Total	89	78	90	49	56.2	23	52	34	46	11	40



Quality Index

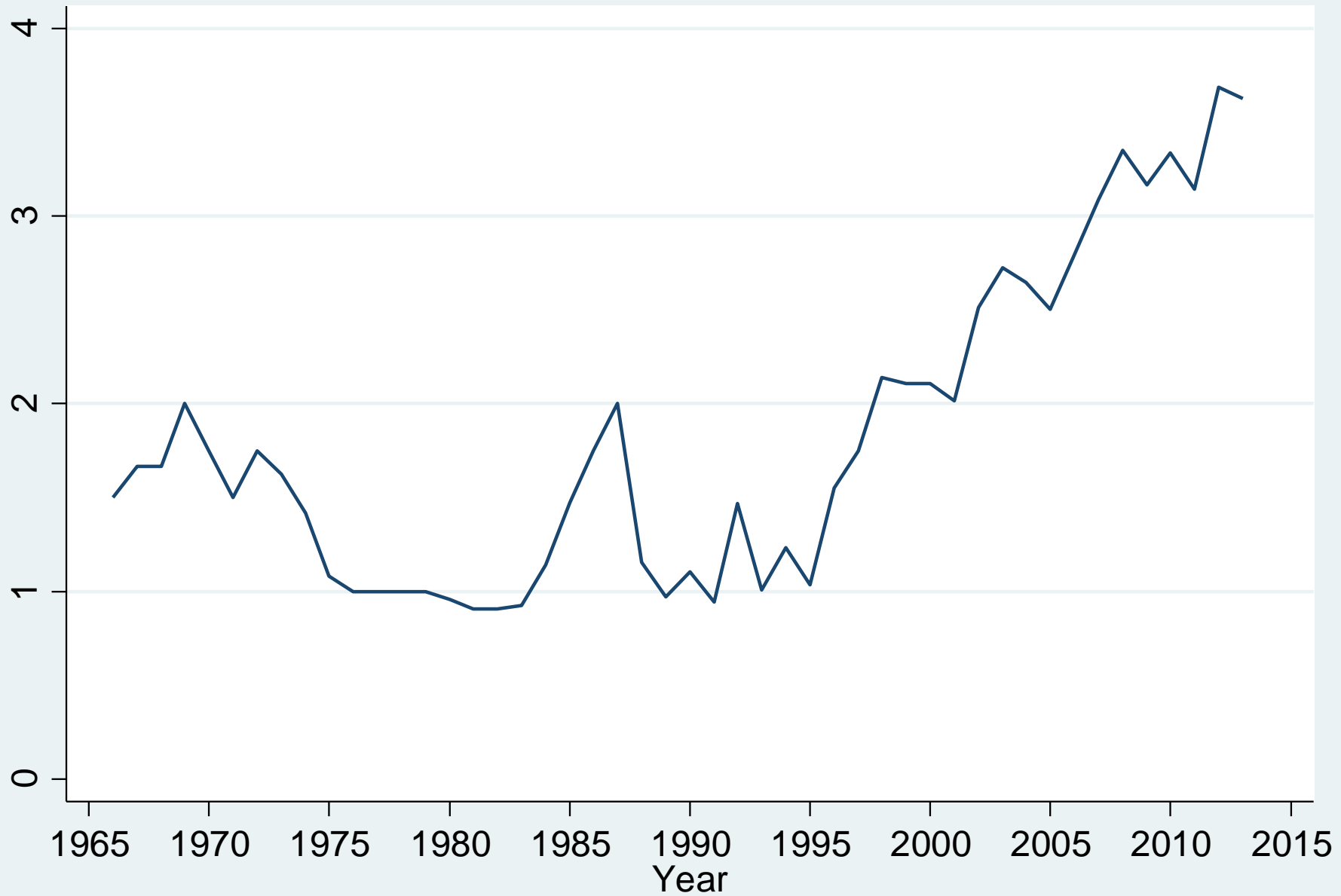
Unweighted sum of 5 dichotmous indicators (0-5)

				ABS				
				AFB				
		AMB		EB				
		ARB		ISJP				
		ASES		NBB	EVS			
		LB	CB	PPE7N	LITS			
CNEP	CDCEE	PA2	PA8NS	WVS	VPCPCE	ISSP	EQLS	ESS
0.00	0.1	0.9-1.1	1.25	1.7-2.0	3.0-3.1	3.6	4.7	4.99



	5+ surveys 90 countries	7-15 year coverage N = 21 countries	16-30 N = 50 countries	31-45 N = 19 countries
Highest scores	GB: 4.89	GB: 4.89	CY: 3.94	CH: 4.05
	BE-FLA: 4.83	BE-FLA: 4.83	HR: 3.88	FI: 3.48
	DE: 4.45	DE: 4.45	PT: 3.6	DK: 3.37
	CH: 4.05	IL-JEW: 3.4	UA: 3.53	AT: 3.25
	CY: 3.94	AL: 3.2	SI: 3.5	BE: 3.2
Lowest scores	CO: 0.82	BO: 0.94	CO: 0.82	MX: 1.6
	EC: 0.84	ZW: 1.4	EC: 0.84	GB-NIR: 1.69
	GT: 0.85	VE: 1.52	GT: 0.85	NG: 1.88
	SV: 0.85	UG: 1.6	SV: 0.85	AU: 1.93
	CR: 0.89	TH: 1.8	CR: 0.89	IN: 2.2

Changes in Quality over Time





Conclusions

The quality of surveys as reflected in survey documentation varies significantly across projects, countries and time.

Differences between survey projects are especially large.

There has been clear improvement over time: both overall, and within survey projects.




“Democratic Values and Protest Behavior: Data Harmonization, Measurement Comparability, and Multi-Level Modeling in Cross-National Perspective.”

Institute of Philosophy and Sociology, Polish Academy of Sciences

The Ohio State University

Cross-National Studies: Interdisciplinary Research and Training Program (CONSIRT)

Funded by:  NATIONAL SCIENCE CENTRE
POLAND
(grant number 2012/06/M/HS6/00322)

dataharmonization.org

dataharmonization.org/newsletter/





The Harmonization Project, 2013-2016

A joint venture of The Polish Academy of Sciences and The Ohio State University:
Democratic Values and Protest Behavior: Data Harmonization, Measurement Comparability, and Multi-Level Modeling

Substantive focus: democracy and protest behavior worldwide

Core of the project: harmonization of survey data from 22 international survey projects, 142 countries/territories, 1721 national surveys, a time span of 47 years, over 2.2 million respondents

Survey Project	Time span	Waves	Files	Samples	Cases
Afrobarometer	1999-2009	4	4	66	98942
Americas Barometer	2004-2012	5	1	92	151341
Arab Barometer	2006-2011	2	2	16	19684
Asian Barometer	2001-2011	3	3	30	43691
Asia Europe Survey	2000	1	1	18	18253
Caucasus Barometer	2009-2012	4	4	12	24621
Consolidation of Democracy in Central and Eastern Europe	1990-2001	2	1	27	28926
Comparative National Elections Project	2004-2006	1	8	8	13372
Eurobarometer	1983-2012	7	7	152	138753
European Quality of Life Survey	2003-2012	3	1	93	105527
European Social Survey	2002-2013	6	2	146	281496
European Values Study/World Values Survey	1981-2009	9	1	312	423084
International Social Justice Project	1991-1996	2	1	21	25805
International Social Survey Programme	1985-2013	13	13	363	493243
Latinobarometro	1995-2010	15	15	260	294965
Life in Transition Survey	2006-2010	2	2	64	67866
New Baltic Barometer	1993-2004	6	1	18	21601
Political Action II	1979-1981	1	1	3	4057
Political Action – 8 Nation Study	1973-1976	1	1	8	12588
Political Participation and Equality in 7 Nations	1966-1971	1	7	7	16522
Values and Political Change in Postcommunist Europe	1993	1	5	5	4723
TOTAL	1966-2013	89	81	1721	2289060