



3MC International Conference 2016

Chicago, July 27, 2016

Item-specific metadata in ex-post harmonization of international survey projects

Marta Kołczyńska

The Ohio State University, Polish Academy of Sciences

Kazimierz M. Słomczynski

Polish Academy of Sciences, The Ohio State University



Controlling the Harmonization Process

Problem: Lack of standard procedures of dealing with methodological differences between surveys that would allow to jointly analyze data from different surveys

Proposed solution: **set of Harmonization Control Variables (item-specific metadata)** to capture features of the item that may affect the target variable

Future: construction of a synthetic „goodness of measurement” indicator to be included in analyses with harmonized survey data



Harmonization controls

1. Properties of the item with regard to:
 - 1a. Wording of the question
 - 1b. Formal properties of the item
 - 1c. Placement in questionnaire
2. Characteristics of data collection (item non-response)
3. Transformation of source into target variables:
 - 3a. Number of source variables
 - 3b. Source variable response scales
4. Reliability of the target variable



Examples

	Trust in parliament	Participation in demonstrations
N projects	18	19
N waves	67	59
N surveys	1314	1148
N countries/ territories	137	136
N individuals (unweighted)	1,676,289	1,513,332



1a. Wording

Example: Participation in demonstrations

*Taken part in a **lawful** public demonstration (ESS)*

Attended a protest demonstration or rally (ISJP)

*Participated in **illegal** protest activities (EB/62.2/2004)₅*



1b. Formal properties

Example: Trust in parliament, Poland

Survey Quality Predictor (SQP) takes into account over 70 properties of the survey item to estimate:

reliability ($r^2 = 1 - \text{random error}$)

validity ($v^2 = 1 - \text{method effect}$)

quality ($r^2 * v^2$)

Prediction	EB 74.1	EB 77.3	EQLS 2012	ESS 2012	EVS 2008	ISSP 2008
Reliability (r^2)	0.714	0.710	0.715	0.770	0.740	0.725
Validity (v^2)	0.887	0.959	0.802	0.939	0.957	0.917
Quality (q^2)	0.633	0.680	0.573	0.723	0.708	0.665



1c. Position in the questionnaire

Example: Trust in parliament

Based on master/core questionnaire for each wave

Range: 6 (ARB 1) – 320 (EVS 1)

Mean: 87.5

Quartiles: 23; 62; 136



2. Characteristics of data collection (item non-response)

Example: Trust in parliament

Project	Number of surveys	% item non-response		
		Mean	Lowest	Highest
ESS	147	3.17	0.29	16.24
EQLS	93	3.41	0.40	13.30
EVS	128	4.13	0	14.08
LB	260	4.55	0	17.80
AMB	92	4.63	0	19.97
ISSP	363	4.72	0.42	17.59
WVS	182	4.98	0	24.74
VPCPCE	5	5.02	0.92	13.15
ARB	16	5.71	1.15	10.23
LITS	64	6.11	0.57	31.22
CNEP	8	6.46	5.05	7.86
EB	152	6.50	1.39	17.95
ASB	30	6.75	0.07	22.20
NBB	18	7.09	0	17.08
AFB	66	7.60	0.83	24.30
ASES	18	8.47	3.25	18.10
CB	12	8.49	3.99	18.59
CDCEE	27	10.26	2.57	27.60
Total		5.00	0	31.22



3a. Number of source variables

Example: Participation in demonstrations

985 surveys: single question

163 surveys: more than one question

ASES: Attend a protest, march or demonstration

ISSP/1996:

Attended a public meeting organised to protest against the government

Gone on a protest march or demonstration



3b. Source variable response scales

Example: Trust in Parliament

Length of scale	Direction of scale	
	Descending	Ascending
11		CNEP, ESS
10		EQLS
7		AMB, NBB (5, 6)
5	ISSP, VPCPEE	CB, LITS
4	ARB, ASB, ASES, CDCEE, EVS, LB, NBB (1, 3), WVS	AFB
2	EB	



4. Reliability of the target variable

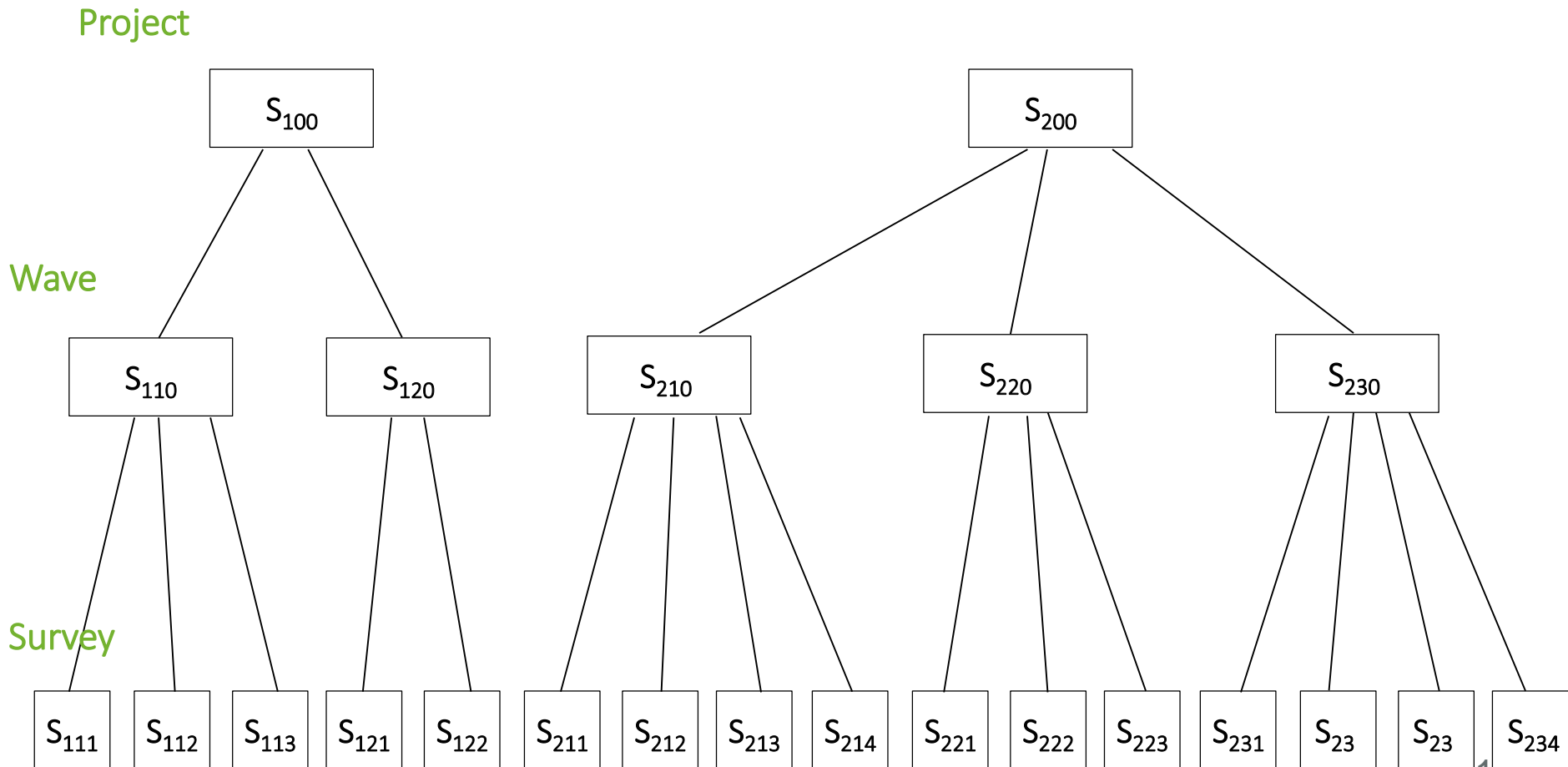
Example: Trust in state institutions

Items	Mean	SD	Factor loadings		
			All surveys	Survey with highest value	Survey with lowest value
Trust in parliament	4.272	2.453	0.776	0.903 (ESS/6/CZ)	0.376 (ABS/1/MN)
Trust in legal system	4.646	2.488	0.670	0.884 (LITS/1/MD)	0.272 (WVS/3/AZ)
Trust in political parties	3.614	2.307	0.710	0.887 (LB/1996/PA)	0.355 (ABS/1/MN)
Eigenvalue			1.554	2.230 (LITS/1/MD)	0.452 (ABS/1/MN)
% of expl. variance			51.81	74.33 (LITS/1/MD)	15.07 (ABS/1MN)

N=1,009,126 individuals; 740 national surveys; original scales 4, 5, 7, 10, 11 points.



Data structure: National survey, Wave, Project





Measurement Levels

Wording, response scales – typically project level, but sometimes wave or survey

Number of source variables, position/context of question in questionnaire – wave level

Properties of item that are language specific, Item non-response – national survey level



Concluding remarks

The presented classification of item-specific metadata and corresponding harmonization control variables should be treated as a guideline.

Every target variable is different and may require a different set of controls.

Ultimately, the goal is to construct a synthetic „goodness of measurement” indicator.




“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