Comparing Two Measures of Education Harmonized *Ex Post*:
Levels of Education and Years of Schooling

Marta Kołczyńska, Irina Tomescu-Dubrow, Kazimierz M. Słomczyński

Institute of Philosophy and Sociology, Polish Academy of Sciences
Cross-National Studies: Interdisciplinary Research and Training Program (CONSIRT)

March 26, 2018

Comparative Survey Design and Implementation Workshop

Mary Immaculate College, Limerick, Ireland

Purpose

Everyone understands "education",

Yet in measures there's much variation.

Is it levels or years?

Is it grades, tests, or tears?

Our paper provides illustrations.

Education level and **years of schooling** in cross-national surveys

SDR Data

Survey Data Recycling (SDR) database, version 1.0

(available at Havard's Dataverse)

22 cross-national survey projects meeting the following criteria:

(1) contain measures of *political attitudes and behaviors,* and main individual-level correlates;

(2) non-commercial;

(3) designed as cross-national, and, preferably, multi-wave;

(4) national samples are intended as representative of the adult population;

(5) documentation (study description, codebook, questionnaire) is provided in English;
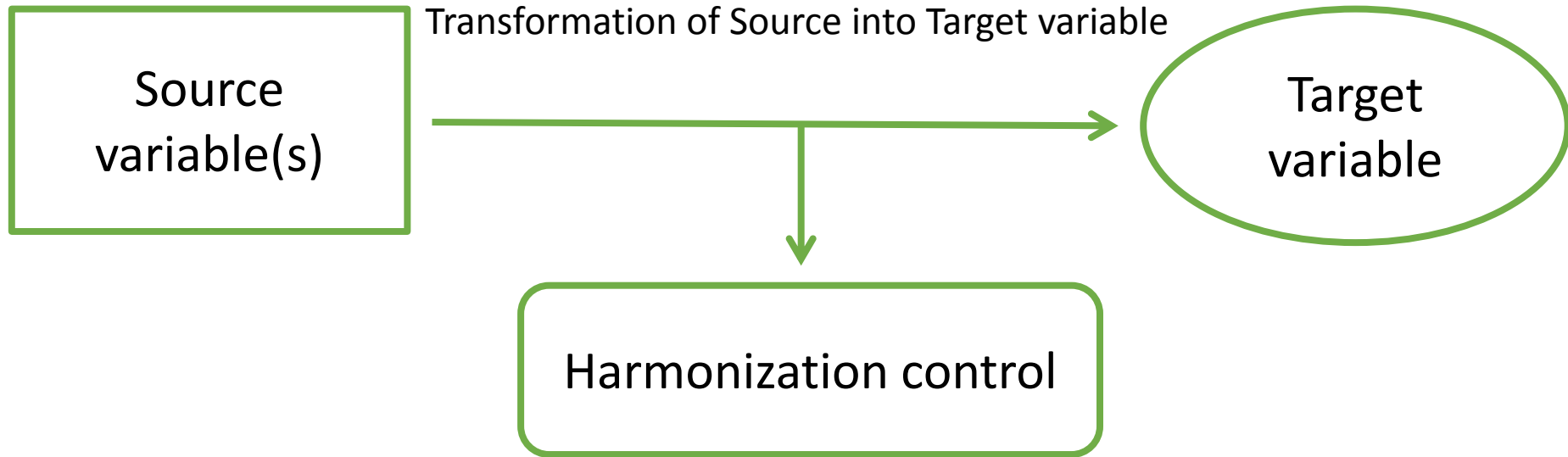
(6) projects are freely available in the public domain.

Survey Projects	Time span	# Waves	# Files	# Data Sets	# Cases
Asian Barometer	2001-2011	3	3	30	43691
Afrobarometer	1999-2009	4	4	66	98942
Americas Barometer	2004-2012	5	1	92	151341
Arab Barometer	2006-2011	2	2	16	19684
Asia Europe Survey	2000	1	1	18	18253
Caucasus Barometer	2009-2012	4	4	12	24621
Consolidation of Democracy (C./East 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/Political Change, Post-communist Europe	1993	1	5	5	4723
	1966-2013			1721	2,290,060

Availability of education variables

Education level	Schooling years (incl. age / year of completion)		Total surveys
	Yes	No	
Yes	1,189	221	1,410
No	305	6	311
Total surveys	1,494	227	1,721

Project	Level	Years	Both	Both
		Proportion of surveys		N surveys
Asian Barometer	1	0.9	0.9	27
Caucasus Barometer	1	1	1	12
Consolidation of Democracy in CE Europe	0.926	0.519	0.519	14
Eurobarometer	0.007	1	0.007	1
European Quality of Life Survey	1	0.634	0.634	59
European Social Survey	1	1	1	146
European Values Study	0.641	1	0.641	82
International Social Survey Programme	0.997	0.986	0.983	357
Latinobarometro	0.996	1	0.996	259
Life in Transition Survey	1	1	1	64
New Baltic Barometer	1	0.556	0.556	10
Political Action - An Eight Nation Study	1	1	1	8
Political Participation and Equality in 7 Nations	1	0.143	0.143	1
Values and Political Change in Postcomm. Europe	1	1	1	5
World Values Survey	0.886	0.783	0.685	126
Overall (among 22 projects, 1721 surveys)	**0.819**	**0.868**	**0.691**	**1189**

Harmonization control variables

Source
variable(s)

Transformation of Source into Target variable

Target
variable

Harmonization control

Variables recording properties of source survey items that would be lost in the process of harmonization.

They accompany the harmonizied (target) variables, in our case,the two measures of education.

They are constructed separately for each target variable.

Education level

Values of the harmonized (target) variable:

0   Less than primary
1   Primary
2   Lower secondary
3   Upper secondary
4   Post-secondary non-tertiary

5   Short cycle tertiary
6   Bachelor's or equivalent
7   Master's or equivalent
8   Doctorate or equivalent
9   Not elsewhere classified

Harmonization control (HC) variables:

(a) incomplete level,

(b) completed and higher level,

(c) education included a vocational component.

Schooling years

Values of the harmonized (target) variable:

0 (no education) … 18 (higher values top-coded)

Harmonization control (HC) variables:

(a) the target value was derived from age or year of completion of full-time education,

(b) the lowest category of the source-response scale was "open-ended," representing the given value together with lower ones,

(c) the highest category of response was „open-ended",

(d) the respondent was still in school at the time of the survey,

(e) the source value exceeds 18 years and was top-coded,

(f) the number of distinct values the schooling years variable takes in the source data.
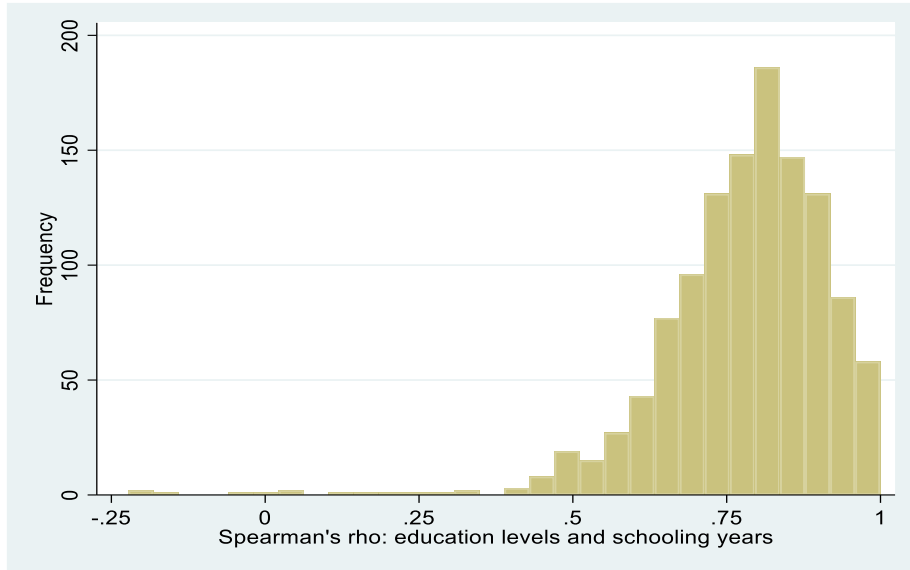
Overall covariation

Expectation: if two variables measure the same concept, they should be highly correlated.

Overall covariation

Expectation: if two variables measure the same concept, they should be highly correlated.

Spearman's rank correlation for the pooled sample of

1,524,897 respondents: rho = 0.77

Correlations by survey
(N = 1189)

Histogram with y-axis labeled "Frequency" (0 to 200) and x-axis labeled "Spearman's rho: education levels and schooling years" ranging from -.25 to 1.

Low correlations:
36 surveys: rho < 0.5
4 surveys: rho < 0

High correlations:
78 surveys rho > 0.95

Schooling years for educational levels (no harmonization controls)

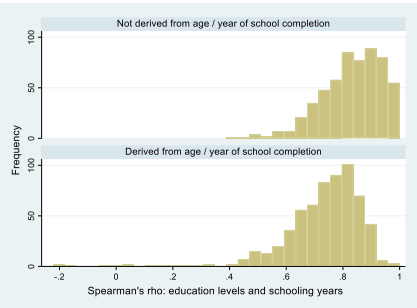| Level of education	Schooling years		
	Mean	**S.D.**	**N**
Less than primary	5.0	4.22	119,189
Primary	7.6	2.79	242,745
Lower secondary	10.2	2.22	166,871
Upper secondary	12.0	2.41	563,553
Post-secondary	13.5	2.37	99,921
Short-cycle tertiary	14.7	2.58	48,114
Bachelor's or equivalent	15.9	2.37	254,037
Master's or equivalent	16.1	1.86	30,442
Doctoral or equivalent	17.1	0.49	25
Not elsewhere classified	9.7	5.95	1,760
Total	**11.5**	**4.165**	**1,526,657**

Schooling years by educational levels with harmonization controls for education level

Level	Control variables	Mean	S.D.	N
Primary	Incomplete level	6.8	3.51	7986
	Reference	7.6	2.75	231732
	Completed and higher	7.3	2.83	2135
	Completed and higher, vocational training	10.0	2.37	892
Upper secondary	Incomplete level	11.1	2.70	85756
	Incomplete level, vocational training	10.5	2.64	45133
	Reference	12.3	2.14	346039
	Vocational training	12.5	2.50	86625
Bachelor's or equivalent	Incomplete level	14.5	2.61	43985
	Incomplete level, vocational training	14.4	2.80	499
	Reference	15.8	2.18	61828
	Vocational training	16.1	2.10	12140
	Completed and higher	16.3	2.19	134878
	Completed and higher, vocational training	15.2	2.60	707

Schooling years derived from age at completion of education

Level of education	Schooling years			
	Derived		Not derived	
	Mean	**S.D.**	**Mean**	**S.D.**
Less than primary	5.146	4.378	4.520	3.738
Primary	8.051	2.831	7.189	2.688
Lower secondary	11.020	2.687	10.073	2.145
Upper secondary	12.191	2.667	11.849	2.014
Post-secondary	14.340	2.918	13.376	2.293
Short-cycle tertiary	15.636	2.497	13.799	2.330
Bachelor's or equivalent	16.006	2.603	15.766	2.146
Master's or equivalent	17.178	1.791	16.813	1.865
Doctoral or equivalent	17.600	0.894	16.950	0.224
Not elsewhere classified	13.385	3.776	9.673	5.957
Total	**11.427**	**4.488**	**11.530**	**3.883**

Correlations by derived from age / year of completion of education

Summary

In 920 out of 1189 surveys (77%) with both measurements of education, the correlation (Spearman's rho) between them is greater than 0.7.

Some mismatch between education level and schooling years reflects real differences: repeated grades, returns to school after a break, multiple degrees at the same level, level completion ahead of schedule.

Surveys where the association is

(a) lower than expected for two operationalizations of the same concept or

(b) higher than expected for survey questions asked separately

require further investigation.

Acknowledgements

SDR Team at the Polish Academy of Sciences and The Ohio State University: Kazimierz M. Słomczyński (PI), Irina Tomescu-Dubrow (PI), J. Craig Jenkins (PI), Marta Kołczyńska, Ilona Wysmułek, Przemek Powałko, Nika Palaguta, Weronika Boruc, Denys Lavryk, Marcin W. Zieliński, Bashir Tofangsazi

dataharmonization.org
https://dataharmonization.org/newsletter/