

Conference and Workshop on Survey Data Harmonization

Warsaw, 20 December 2013

Survey Data Quality

Matthew Schoene

Marta Kołczyńska

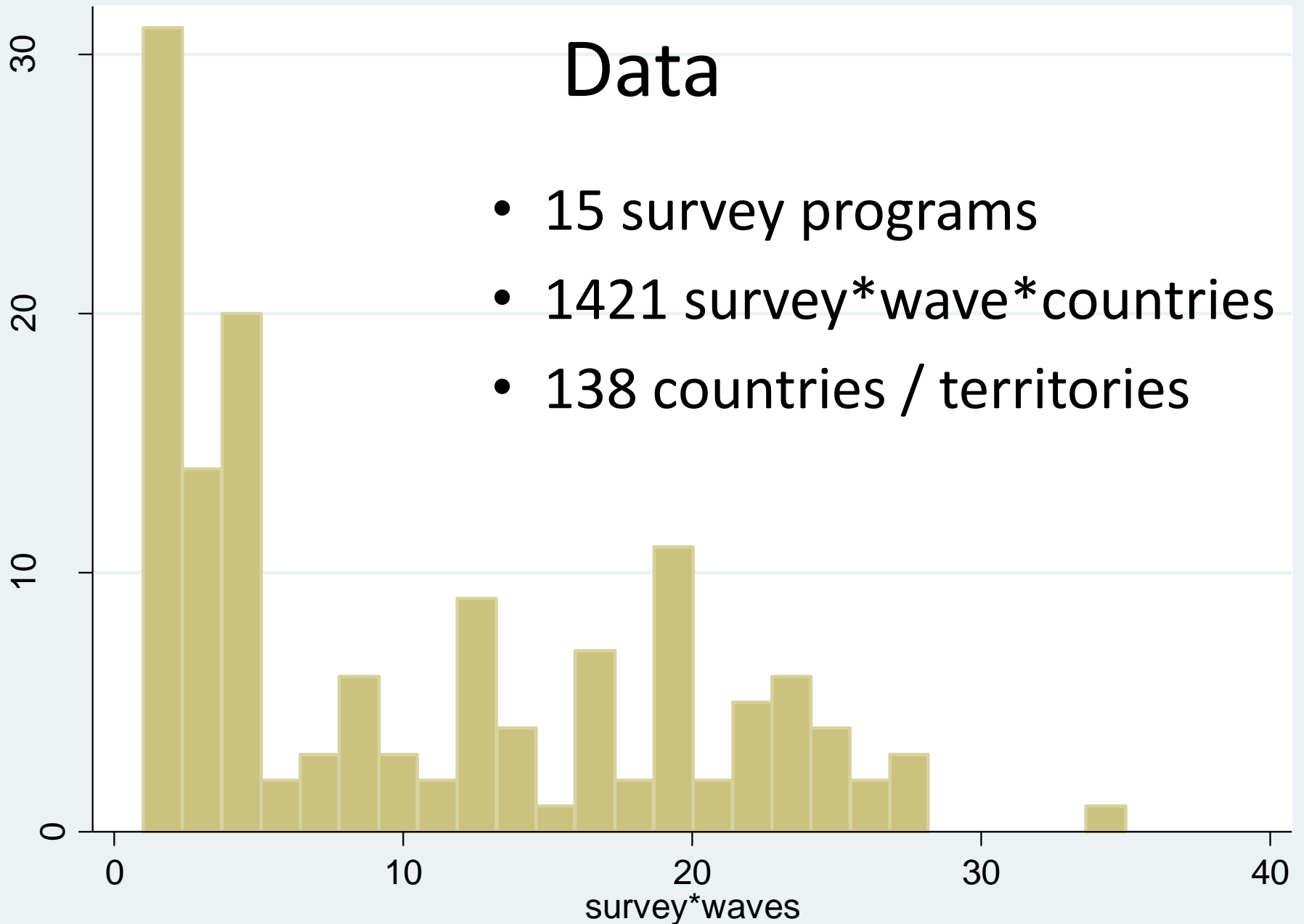
The Ohio State University

Data

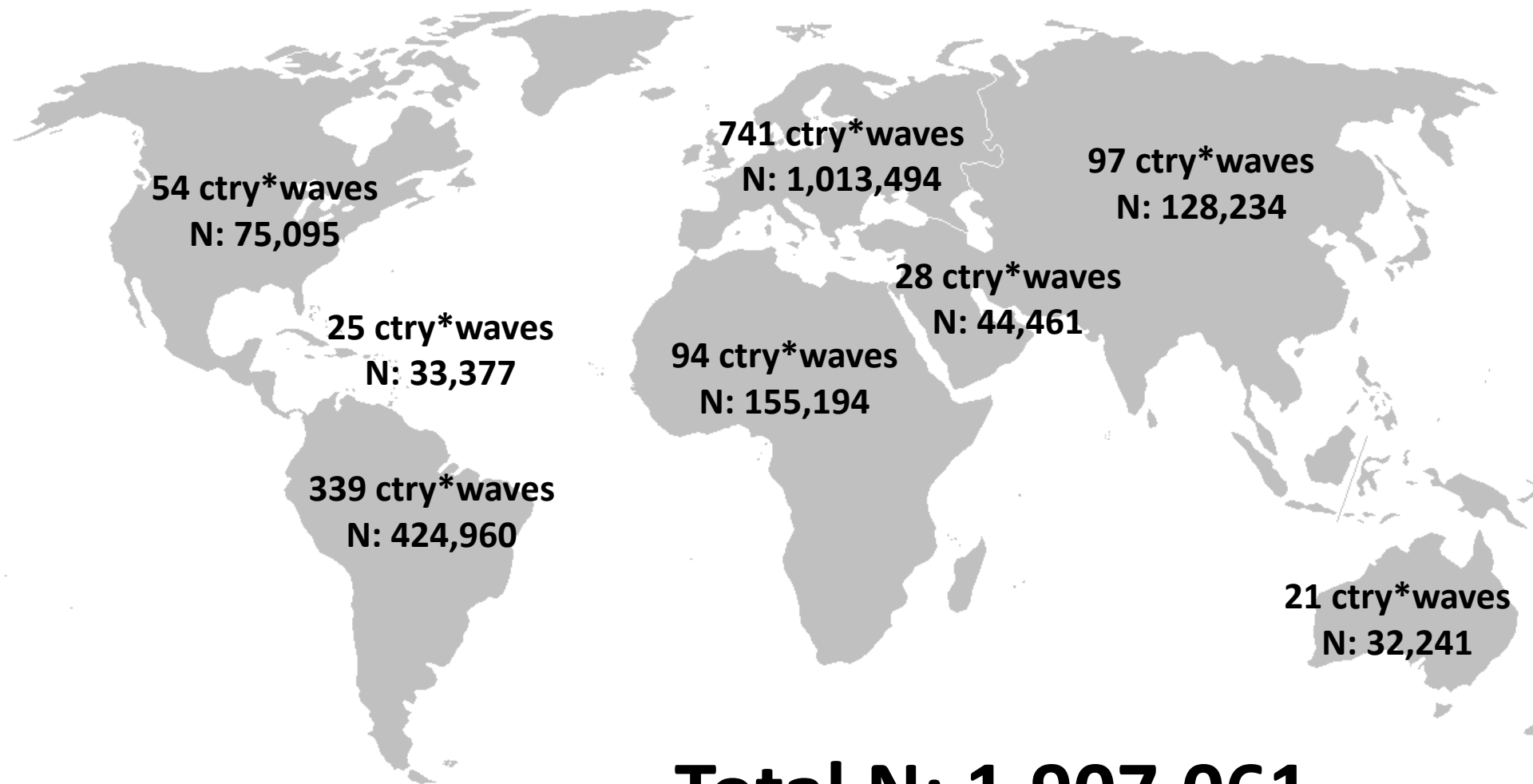
- AFB - Afrobarometer
- AMB - AmericasBarometer
- ARB - Arab Barometer
- ASB - Asian Barometer
- ASES - Asia Europe Survey
- CB - Caucasus Barometer
- CDCEE - Consolidation of Democracy in Central and Eastern Europe
- EB - Eurobarometer
- EQLS - European Quality of Life Survey
- ESS - European Social Survey
- EVS - European Values Study
- FSU - Panel Study of Political Values in the Former Soviet Union
- ISSP - International Social Survey Programme
- LB - Latinobarometer
- WVS - World Values Survey

Data

- 15 survey programs
- 1421 survey*wave*countries
- 138 countries / territories

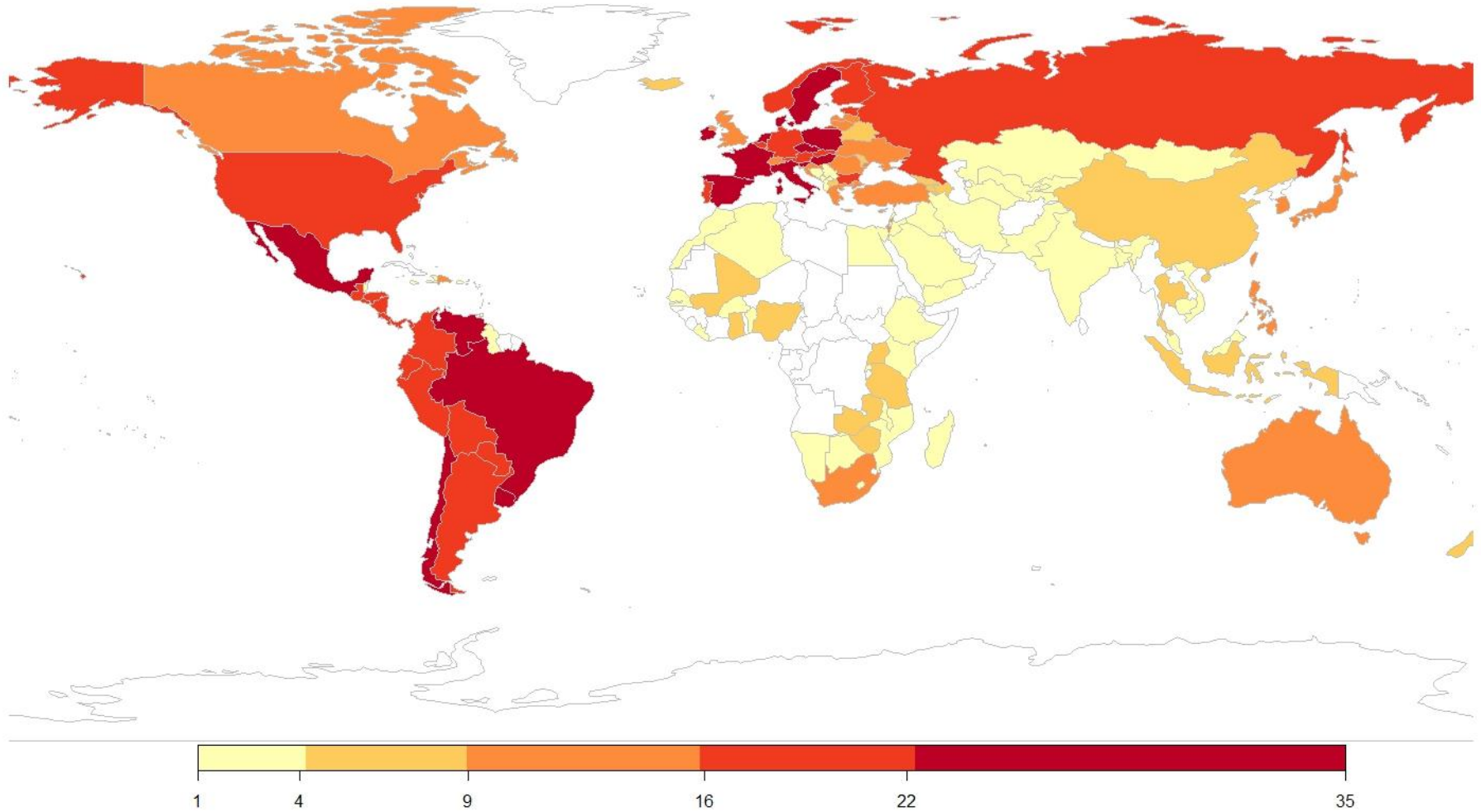


Data Coverage

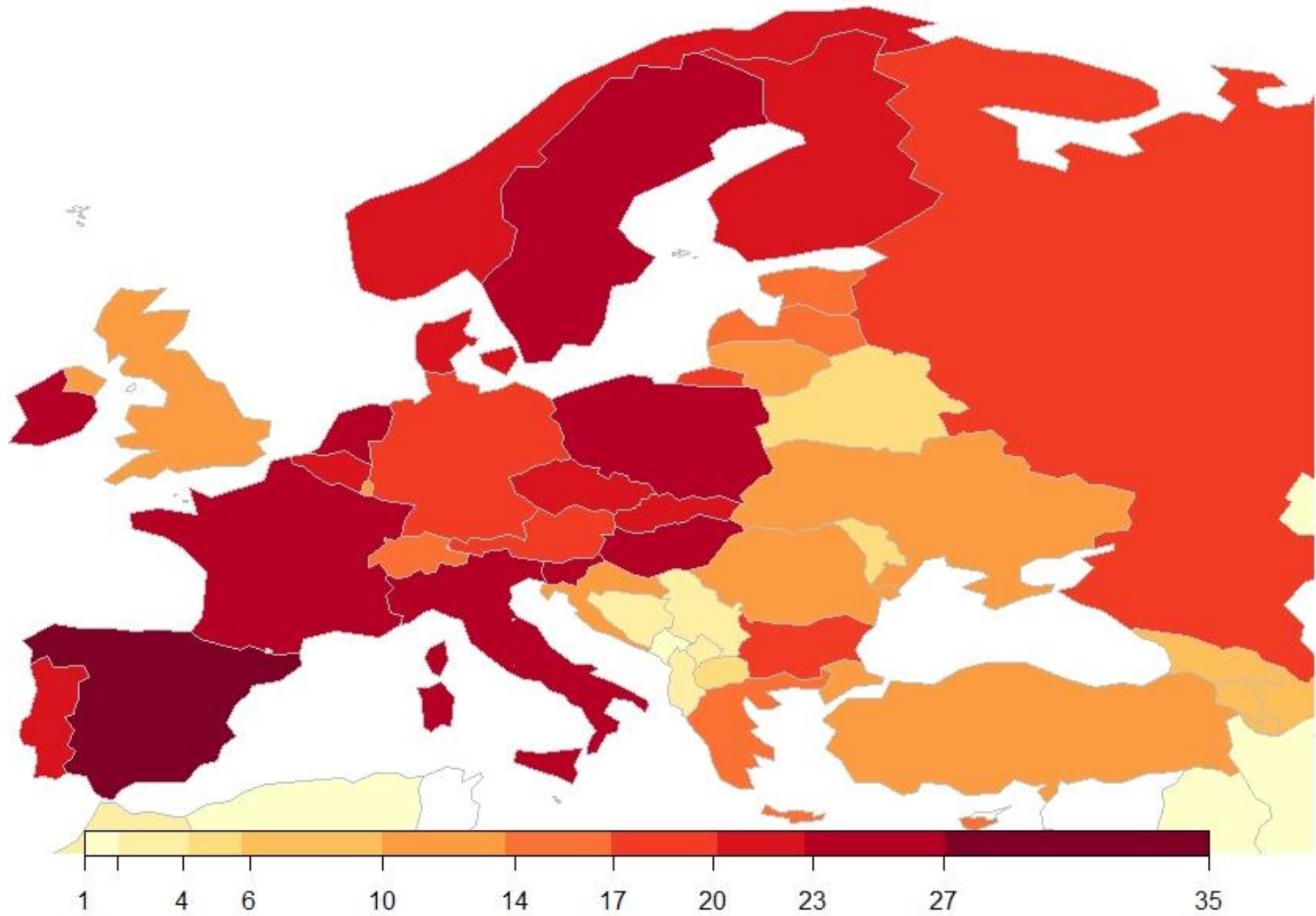


Total N: 1,907,061

Data Coverage



Data Coverage



Documentation

Where to find information about survey and data quality?

- Online text.
- Study descriptions.
- Methods reports.
- Technical reports.
- Codebooks.
- Random places.

Data Quality

- What is data quality?

Basic Quality Indicators

- Response rate (unit non-response).
- Questionnaire translation
 - back-translation or other documented method.
 - Native language: English, Spanish?
- Pretest.
- Fieldwork control.
 - Interview supervised?
- Item non-response.

Ideal Type

- 1) Fully documented with detailed information about every stage of the survey process,
- 2) Carried out certain quality assurance and control measures.

Response Rate: Definitions (AAPOR)

- Response Rate = full interviews / full & partial interviews + non-interviews (refusal + break-off + non-contacts + others) + all cases of unknown eligibility.
- 4 more definitions of response rates
- 4 definitions of cooperation rates
- 3 definitions of refusal rates
- 3 definitions of contact rates

Response Rate: Problems

„For non-probability samples, response rate calculations make little sense, given the broader inferential concerns. Further, for many of these surveys, the denominator is unknown, making the calculation of response rates impossible”

Source: Standard Definitions report (7th edition, 2011), p. 32, aapor.org

Response Rate: Data Gaps

- 2047 survey*wave*countries
- 847 s*w*c with response rate info (41.4%)

Example: Asian Barometer, 3rd wave

9 countries, 9 country reports, 3 response rates

Survey	Total wave* countries	with response rates	with rr % total	average rr	stdev rr
ESS	155	155	100.00%	0.62	0.10
EQLS	93	91	97.80%	0.53	0.18
ISSP	354	328	92.70%	0.57	0.18
AFB	101	64	63.40%	0.80	0.12
EVS	127	72	56.70%	0.60	0.18
WVS	208	118	56.70%	0.68	0.21
ASB	29	12	41.40%	0.70	0.22
CB	21	6	28.60%	0.75	0.07
ARB	6	1	16.70%	0.95	0
AMB	109	0			
ASES	18	0			
CDCEE	26	0			
EB	504	0			
LB	284	0			
FSU	12	0			
TOTAL	2047	847	41.4%	0.62	0.182

Asian Barometer 3, Indonesia

„Respond Rate: The original sample size is 1550. Successful interviews were 1226 without substitution, and therefore the respond rate is 79%. The number of substitution is 324, and there were no respondent unsuccessful interviews.”

Asian Barometer 3, Singapore

„Response Rate

There were 1,000 successful interviews with 2427 households interviewed by interviewers, and therefore the response rate is 41.2%.„

ISSP 2010, Israel

„Interviews: 1023

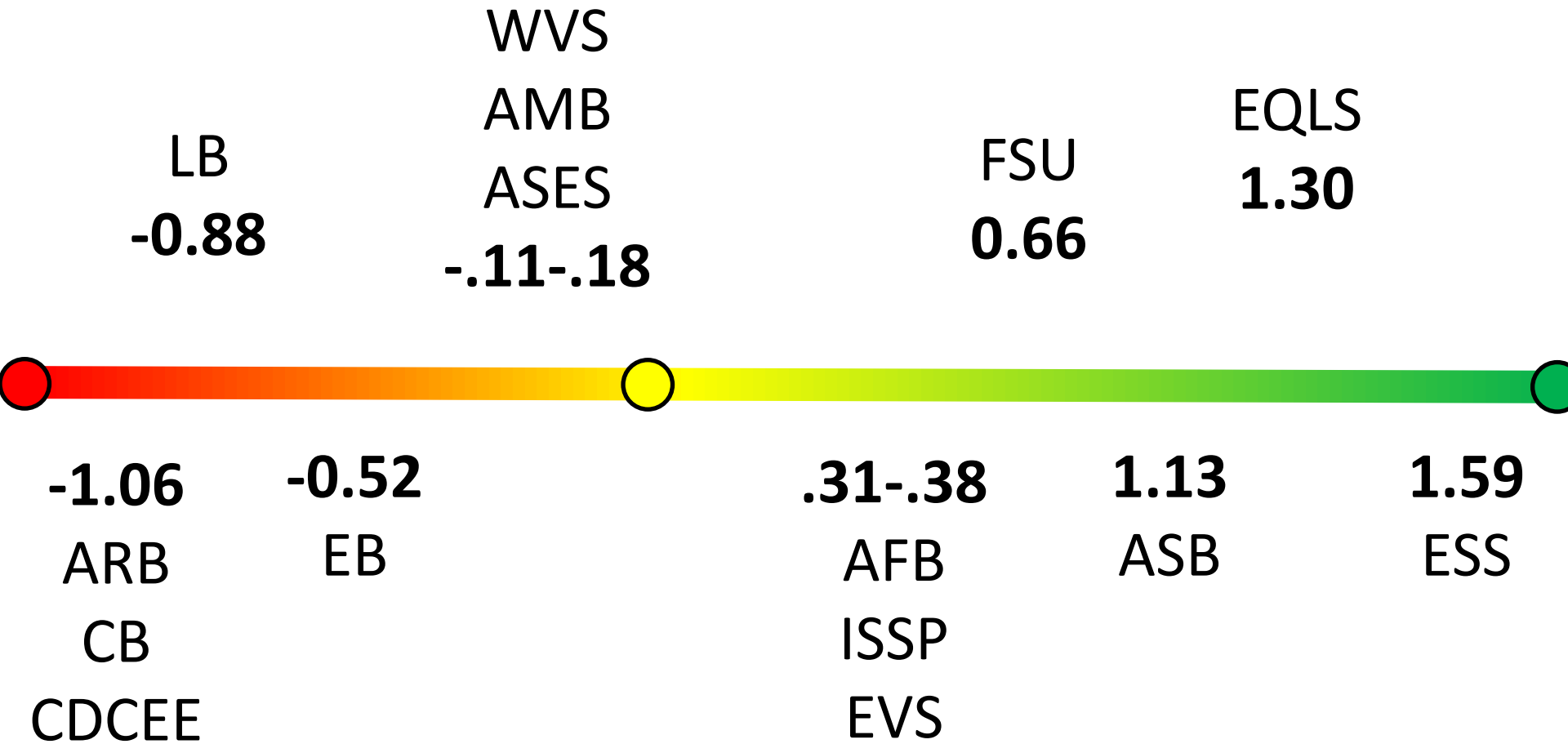
These figures pertain to interviews in Jewish and Mixed (Jewish-Arab) communities. In the case of additional 193 interviews conducted in small Arab communities there was no sampling list and we have no information on response rates”

Survey Process Quality

- Questionnaire backtranslation
- Pretest
- Fieldwork control
- 1: implementation documented
- 0: non-implementation documented + no information

Survey	Total wave* countries	Back translation	Pretest	Fieldwork control	average (bt+pt+fc)
ESS	155	100%	100%	100%	3.00
EQLS	93	100%	99%	70%	2.69
ASB	29	100%	52%	100%	2.52
FSU	12	100%	100%	0%	2.00
AfB	101	54%	54%	54%	1.63
ISSP	354	61%	29%	72%	1.62
EVS	127	46%	46%	63%	1.55
ASES	18	100%	11%	0%	1.11
WVS	208	17%	35%	53%	1.05
ANB	109	0%	100%	0%	1.00
EB	504	69%	0%	0%	0.69
LB	284	0%	0%	20%	0.2
ARB	6	0%	0%	0%	0
CB	21	0%	0%	0%	0
CDCEE	26	0%	0%	0%	0
TOTAL	2047	48%	35%	38%	1.22

Quality Index



Quality vs. GNP pc by Region

	Number of countries ever covered	Av GNPpc of countries covered when covered	Quality1 Sum(bt pt fc)	Quality2 Factor(bt pt fc)
Africa	36	3,305	1.6 (1.4)	0.36 (1.26)
Asia	18	10,344	1.69 (1.09)	0.43 (0.98)
Caribbean	5	6,976	0.94 (0.81)	-0.22 (0.73)
Europe	46	18,662	1.36 (1.1)	0.11 (0.99)
Latin America	19	6,565	0.51 (0.71)	-0.6 (0.65)
Middle East	10	14,337	1.47 (1.19)	0.23 (1.07)
North America	3	20,671	0.92 (1.01)	-0.24 (0.92)
Oceania	2	21,247	0.64 (0.64)	-0.48 (0.58)
TOTAL	139	12,764	1.22 (1.11)	0 (1)

Data: World Bank: GNI per capita based on purchasing power parity (PPP)

Item Non-response

Around 30 different labels:

Not sure - Can't choose - Don't know

Do not understand the question - Haven't thought much about it

Decline to answer

Break off

Interviewer error

Not asked in survey - Not applicable - Legal skip

No answer - Missing - Missing; Unknown

22 different values:

-9, -7, -5, -4, -3, -2, -1, 0, 3, 5, 6, 7, 8, 9,

77, 88, 98, 99, 998, 888888, 988888, 999999

Trust in Parliament

Survey	% non-response	No non-response categories
EQLS*	1.3	2
ESS	3.2	3
AMB	4.6	3
LB*	4.6	2/3/4
ISSP	5.4	2
ARB*	6.5	2/3
ASB*	7.1	3
AFB	7.4	3
CB	8.8	3/4/6
IVS	9.5	5
EB	9.9	2
FSU	10.0	3
CDCEE	10.7	2
ASES	13.5	3
TOTAL	6.1	2-6

*Completely Missing Data

Trust in parliament:

EQLS 2: 1297 in 35634, spread across countries

ASB 2: 1 in 12217

ASB 3: 10 in 19798

Age:

ARB_2006_2008: 28 in 5625

ASB_2004_2008: 7 in 19798

ASB_2010_2011: 15 in 11676

LB_1997: 9 in 20243

Gender: 44 in total

Little accidents

- Human error!
 - Slovakia vs. Slovenia.
 - Denmark: DNK or DEN?
 - Territories: North Cyprus, Germany(ies), Kosovo...

ISSP 2010, Argentina

„All questions were asked as prescribed. But, unfortunately, we had a «little accident». The field work was carried out both for Social Inequality IV and Environment III at the same time. ... By mistake, instructions in the background part of the questionnaires asked interviewers to ask background questions 21a and 21b to the Environment interviewees, not to the Social Inequality ones !! However, results showed that they were not – as suspected productive questions. For 21a, only 19.3% informed about a money range (most were non-owners, or answered DK, or refused). For 21b, only 3.5% informed about a money range.”

Other Quality Measures

- Sample description.
- Sampling method.
- National Coverage.
- ?

Survey Data and Documentation Links

[European Social Survey \(ESS\)](#)

[European Values Study \(EVS\)](#)

[World Values Survey \(WVS\)](#)

[International Social Survey Program \(ISSP\)](#)

[Eurobarometer \(EB\)](#)

[Asian Barometer \(ABS\)](#)

[Afrobarometer \(AFB\)](#)

[Arab Barometer \(ARB\)](#)

[Latinobarometer \(LB\)](#)

[Consolidation of Democracy in Central and Eastern Europe 1990-2001](#)

[European Quality of Life Survey \(EQLS\)](#)

[New Baltic Barometer \(data\)](#)

[East Asian Social Survey \(EASS\)](#)

[Caucasus Barometer](#)

[Americas Barometer](#)

[Asia Europe Survey \(ASES\)](#)

[Harmonia project website: dataharmonization.wordpress.com](http://dataharmonization.wordpress.com)