

Polish Academy of Science
Institute of Philosophy and Sociology
Graduate School for Social Research

Inter-survey variability of corruption measures: Implications of harmonization procedures

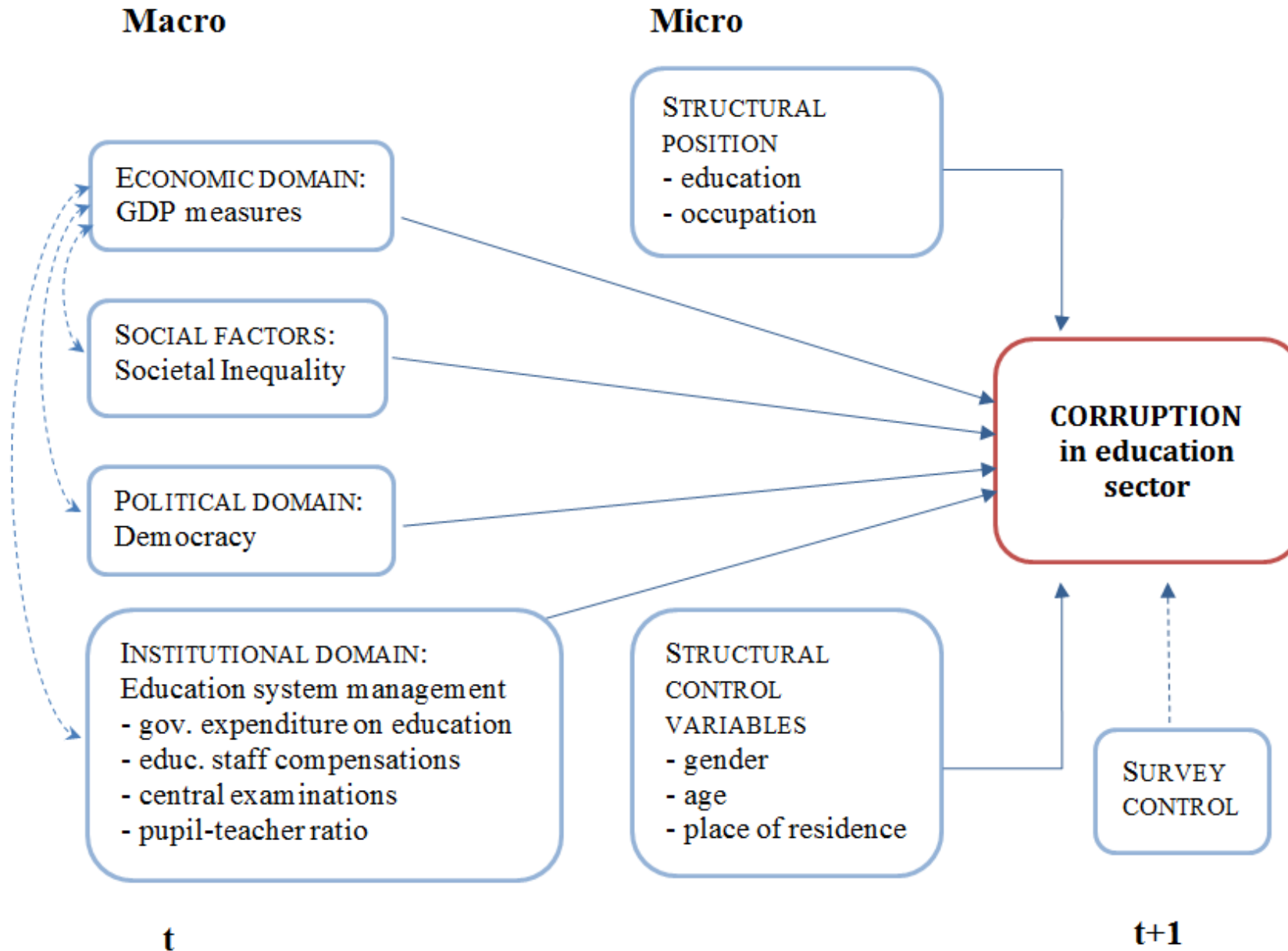
Ilona Wysmułek
ilona.wysmulek@ifispan.waw.pl

CSDI Workshop, Limerick, Ireland, March 26-28, 2018

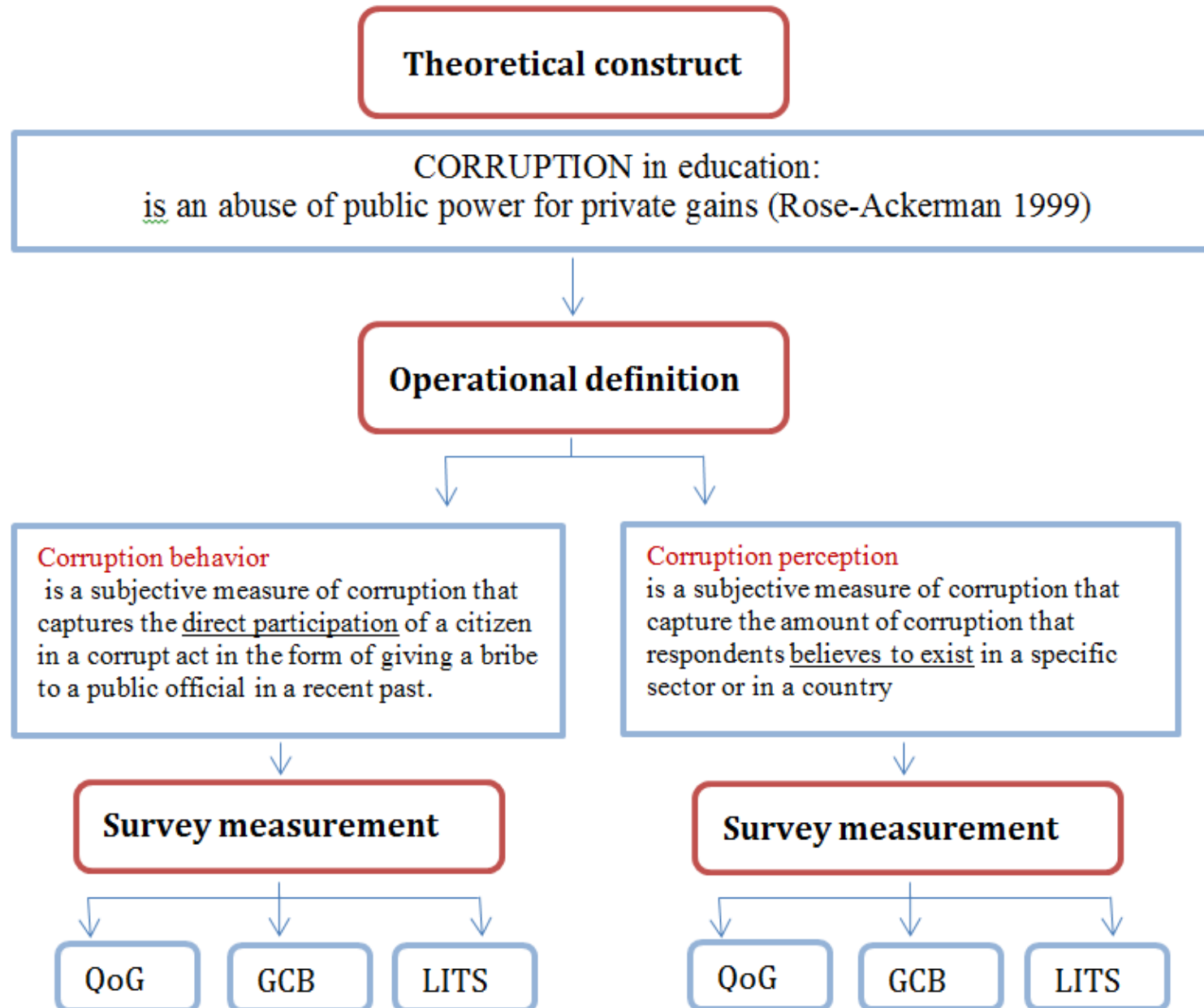
Dissertation research in brief

- “Informal Payments in Public Schools: Determinants of Corruption Perception and Behavior in Europe”
- Supervisors:
 - Prof. dr hab. Kazimierz M. Słomczyński, prof. IFiS PAN, prof. em. OSU
 - Dr hab. Zbigniew Sawinski, prof. IFiS PAN
- Research aim: analysis of the sources of corruption behavior (such as giving bribes, informal payments and expensive gifts) and perceptions of corruption prevalence in the education sector across Europe

Main relationships between variables tested in the research



Corruption: From theory to measurement (and back)



Methodological approach:

Multi-level framework with indicators harmonized ex-post

- **Survey data harmonization** is the procedure that allows to combine different sources into an integrated dataset with comparable indicators.

(Slomczynski et al. 2016, Granda and Blasczyk 2016)

Methodological approach:

Multi-level framework with indicators harmonized ex-post

- **Survey data harmonization** is the procedure that allows to combine different sources into an integrated dataset with comparable indicators.

(!) BUT HOW?

The developments in the field of survey data harmonization result in “accumulated practicalities, and not with the coordination or institutional apparatus one would expect from a 30 year effort”

(Dubrow, Tomescu-Dubrow 2015)

Methodological approach:

Multi-level framework with indicators harmonized ex-post

- **Survey data harmonization** is the procedure that allows to combine different sources into an integrated dataset with comparable indicators.

(!) WAIT BUT WHY?

Ex-post harmonization increases “the sample sizes (..), improves the generalizability of results, helps ensure the validity of comparative research, encourages more efficient secondary usage of existing data, and provides opportunities for collaborative and multi-centre research”

(Doiron et al. 2012)

Integrated Dataset with indicators harmonized ex-post

- Micro-level data:
 - 3 survey projects:
 - (1) Global Corruption Barometer [GCB],
 - (2) Life in Transition Survey [LITS] and
 - (3) Quality of Government survey [QoG]
 - 69 national surveys conducted in 2010 in 30 European countries
 - 31,578 respondents
- Macro-level data: country level indicators and education system characteristics
 - Sources: the World Bank Education Statistics, the Varieties of Democracy, the Quality of Government Standard Dataset and UNESCO Institute of Statistics Education Indicators

Implications of harmonization procedures: Strategies and compromises

Corruption perception:

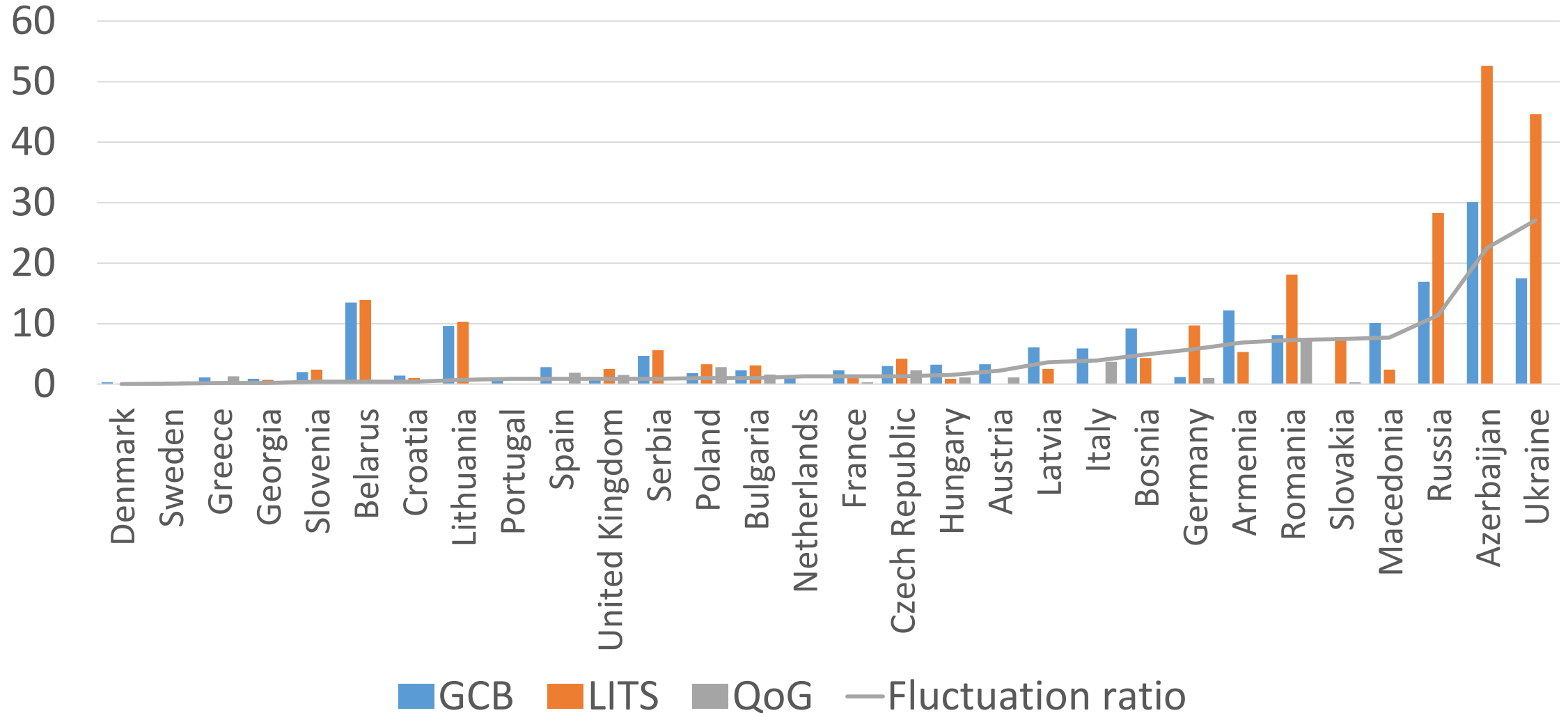
Recode 5- and 11- point scales to binary:

- overcomes the issue of comparability of scales
- informs about affirmative responses, but not their strength
- explanatory power is lower

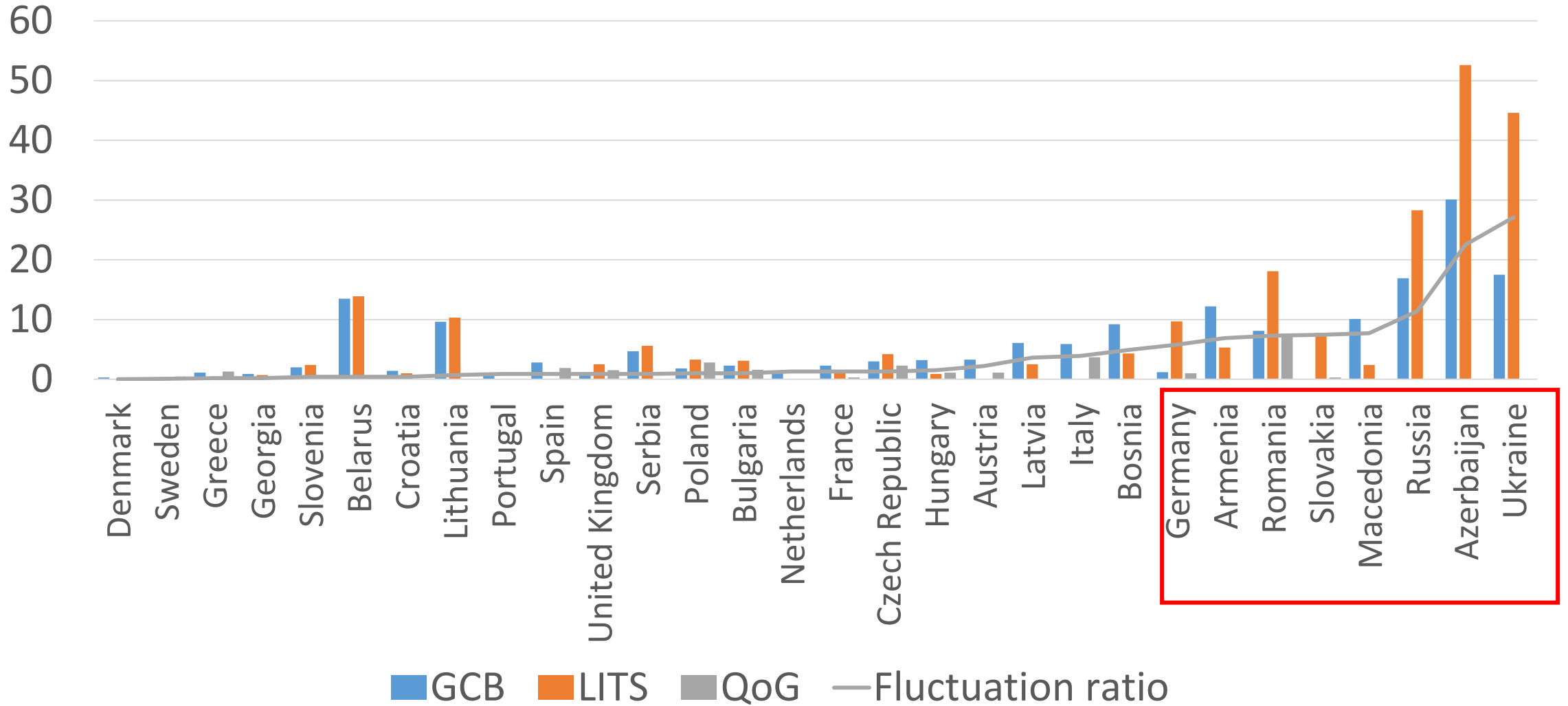
Corruption experience:

- informs about affirmative responses
- limitations of additional analysis for ,don't know's'

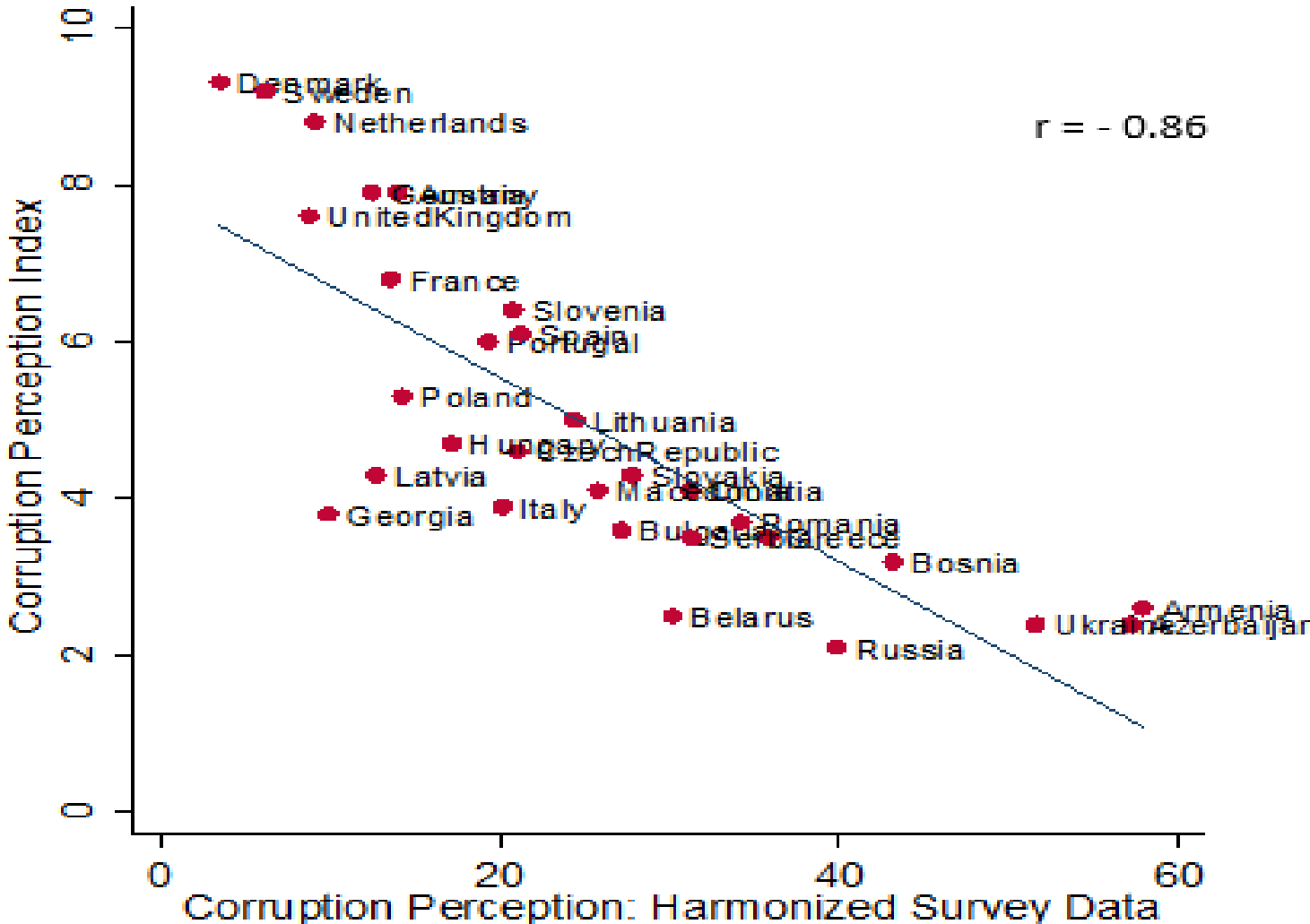
Inter-survey variability of corruption experience in public schools



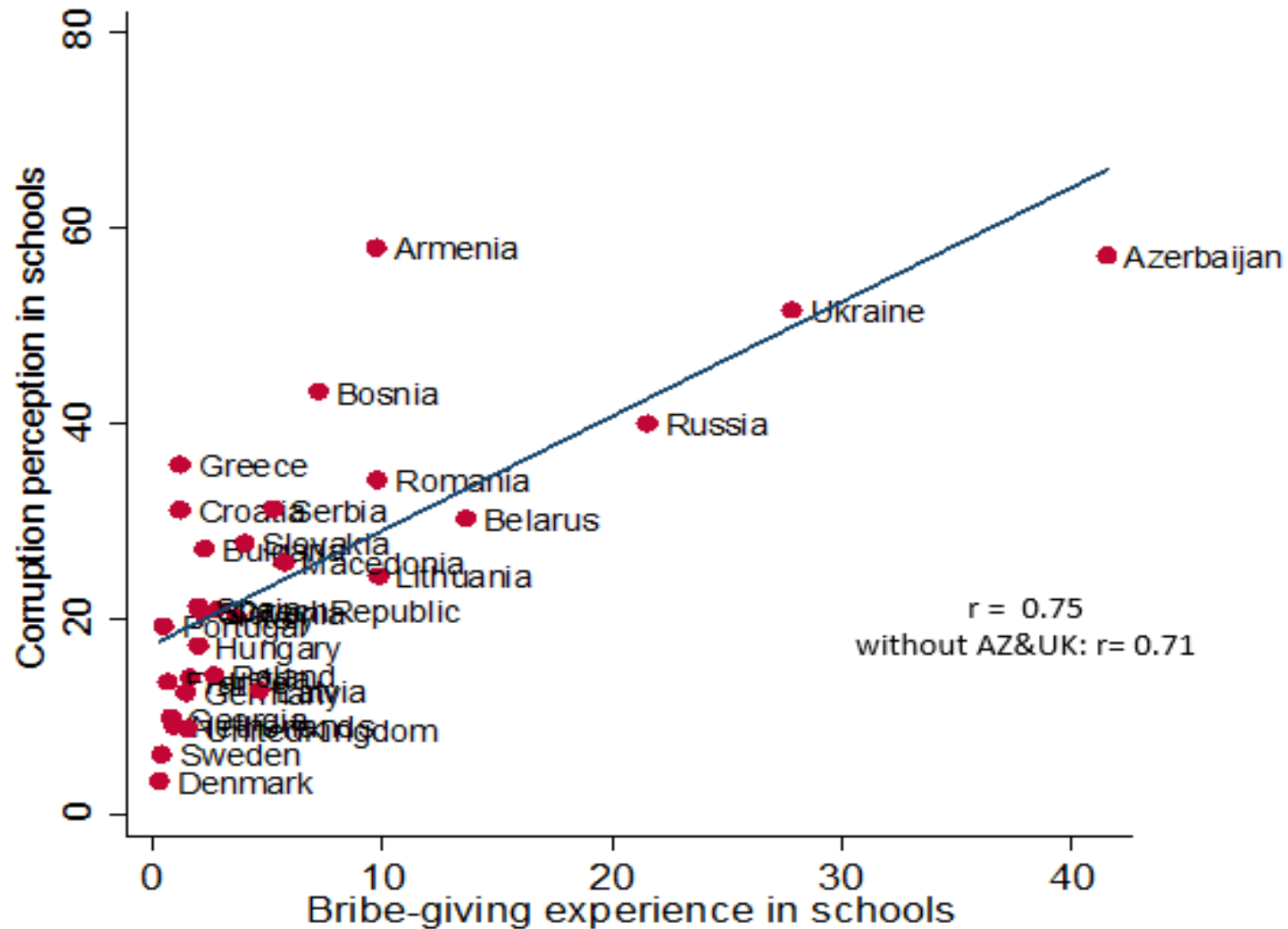
Inter-survey variability of corruption experience in public schools



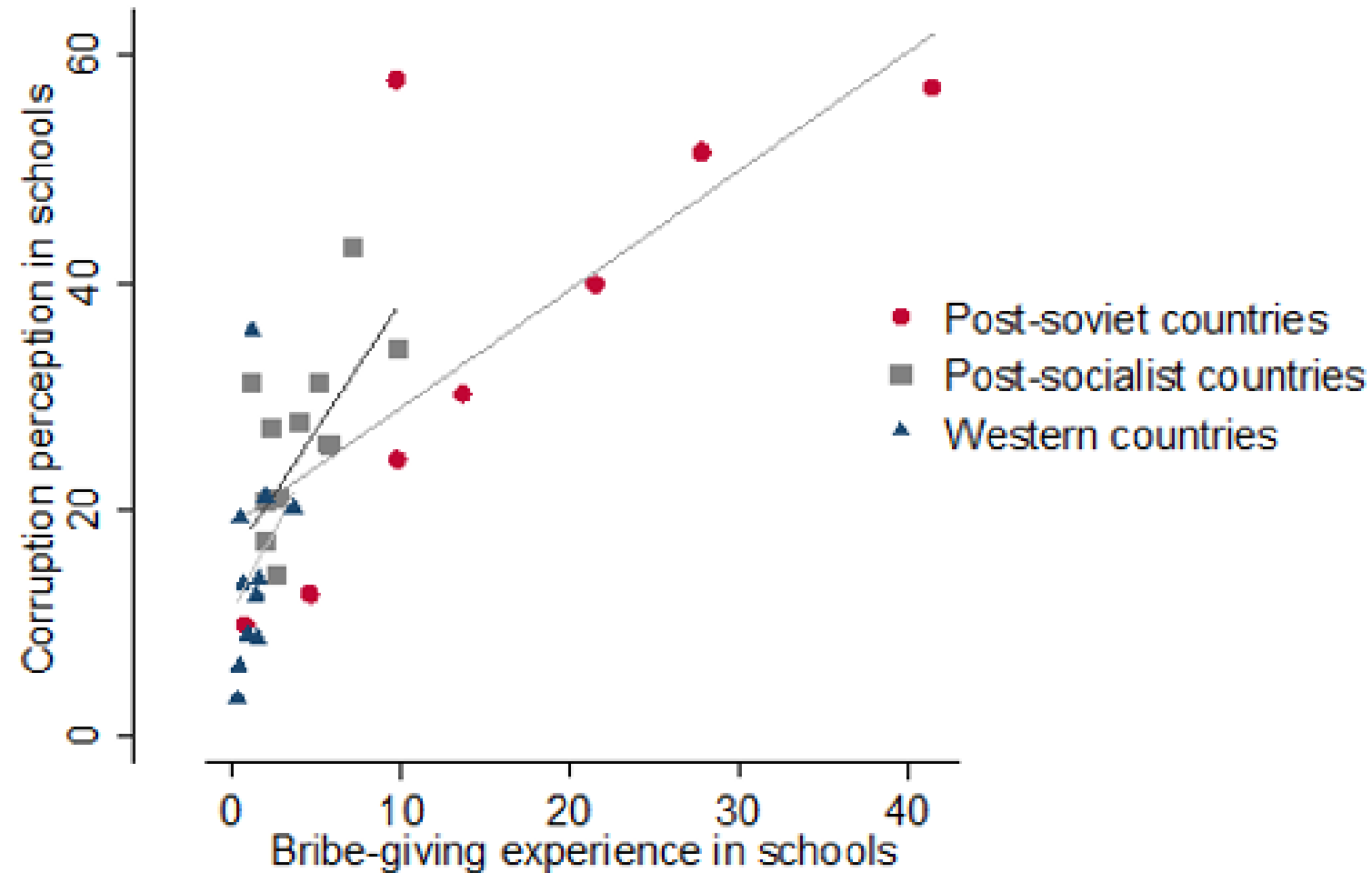
External validity of a harmonized indicator



Correlating harmonized indicators



Correlating harmonized indicators by Europe country groups



Concluding remarks

- Inter-survey variability = always present in ex-post harmonization (to different extent)
- Harmonization workflow helps to signalize problems (like great deviances in distributions), strategies how to deal with it can vary
- In case of corruption measures – despite survey noise, the relations for key indicators are as expected
- New possibilities (both methodological and substantive) offered by harmonization; yet how to fully use them remains opened

New research tools:

- Publicly available documentation of corruption variables available in cross-national projects
 - published and freely available at the Harvard Dataverse https://dataverse.harvard.edu/dataverse/survey_data_on_corruption
 - education specific information is summarized on the UNESCO ETICO Statistics page <http://etico.iiep.unesco.org/resources/statistics/>
- Integrated dataset with harmonized indicators: research tool for further analysis

Thank you!

Ilona Wysmulek

iwysmulek@ifispan.waw.pl

Abbrev.	Survey Project	Time span	Waves	Files	Corr.
			Counts		
EB_corr	Eurobarometer Corruption Themed	2005-2013	5	5	283
GCB	Global Corruption Barometer	2003-2013	8	1	349
ICVS	International Crime Victims Survey*	1992-2005	4	1	108
LITS	Life in Transition Survey	2006-2010	2	2	43
ESS	European Social Survey *	2004-2010	2	2	5
EVS	European Values Study *	1990-2008	3	1	4
ISSP	International Social Survey Programme	2004-2009	3	3	7
WVS	World Values Survey*	1989-2005	4	1	5
ASES	Asia Europe Survey	2000	1	1	3
CSES	Comparative Study of Electoral Systems	2001	1	1	1
QoG	European Quality of Government Survey	2010-2013	2	2	20
EB	General Eurobarometer	1997-2012	7	7	12
ISJP	International Social Justice Project*	1991-1996	2	1	4
PEW	Pew Global Attitudes Project	2002-2012	4	4	9
CCEB	Candidate Countries Eurobarometer	2003	2	2	5
CB	Caucasus Barometer	2009-2012	4	4	10
CDCEE	Consolidation of Democracy in CEE	1990-1998	2	1	11
NBB	New Baltic Barometer*	1993-2004	6	1	14
VPCPCE	Values and Political Change in PostcomEurope*	1993	1	5	2
Total		1989-2013	63	45	895

Variable Label	Source variables		Target variables		
	Value labels	Value Labels	Name	Mean	SD
Corruption experience in education	[see Table 4.4]	1 = gave bribe/ inf. payment 0 = no/DK	BRIBE-EXP	0.06	0.24
Corruption perception in education	[see Table 4.5]	1 = corruption is prevalent 0 = other	CORR_PERC	0.23	0.42
Gender	LITS: 1 = male 2 = female QoG: 0 = male 1 = female GCB: 1 = male 2 = female	1 = female 0 = male	female	0.58	0.49
Place of residence	LITS: 1 = urban; 2 = rural; 3 = metropolitan QoG: 1 = Less than 10,000 (rural); 2 = 10,000-100,000; 3 = 100,000-1,000,000; 4 = Greater than 1,000,000 GCB: 1 = rural; 2 = urban	1 = rural 0 = other	rural	0.35	0.48
Age	LITS: age in years QoG: 1 = 18-29; 2 = 30-49; 3 = 50-64; 4 = 65+ GCB: 1 = under 30; 2 = 30-50; 3 = 51-65; 4 = 65+	1 = 18 - 29 years 2 = 30 - 49 years 3 = 50 years and older	age1 age2 age3	0.26 0.50 0.24	0.44 0.50 0.42
Level of education	LITS: 1 = no degree; 2 = primary; 3 = lower secondary; 4 = upper secondary 5 = post-secondary non tertiary; 6 = BA or more; 7 = MA or PhD QoG: 1 = lower secondary or less; 2 = medium education (higher secondary or post-secondary non tertiary) 3 = higher education GCB: 1 = no education/basic education; 2 = secondary school; 3 = higher level education (university)	1 = Primary or less 2 = Secondary 3 = Tertiary	edu1 edu2 edu3	0.23 0.51 0.26	0.42 0.50 0.44
CONTROL VARIABLE					
Survey project	[constructed variable]	GCB_2010 LITS_2010 QoG_2010	GCB LITS QoG	0.36 0.23 0.41	0.48 0.42 0.49
TECHNICAL VARIABLES					
Country	Survey country codes	Standardized ISO country codes	t_cntr		
Case identifier	[constructed variable]		t_id		
Composite weights	[constructed variable]		t_wght		