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

Ilona Wysmułek
ilonawysmulek@ifispawn.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

**Macro**
- Economic Domain: GDP measures
- Social Factors: Societal Inequality
- Political Domain: Democracy
- Institutional Domain: Education system management
  - Gov. expenditure on education
  - Educ. staff compensations
  - Central examinations
  - Pupil-teacher ratio

**Micro**
- Structural Position:
  - Education
  - Occupation
- Structural Control Variables:
  - Gender
  - Age
  - Place of residence

**Survey Control**

**t**

**t+1**
Corruption: From theory to measurement (and back)

Theoretical construct

Corruption in education: is an abuse of public power for private gains (Rose-Ackerman 1999)

Operational definition

Corruption behavior:
- is a subjective measure of corruption that captures the direct participation of a citizen in a corrupt act in the form of giving a bribe to a public official in a recent past.

Corruption perception:
- is a subjective measure of corruption that captures the amount of corruption that respondents believe to exist in a specific sector or in a country.

Survey measurement

- QoG
- GCB
- LITS

Survey measurement

- QoG
- GCB
- LITS
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

\[ r = 0.75 \]

without AZ&UK: \[ r = 0.71 \]
Correlating harmonized indicators by Europe country groups

[Graph showing correlation between corruption perception in schools and bribe-giving experience in schools, categorized by post-soviet, post-socialist, and western countries.]
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
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EB_corr</td>
<td>Eurobarometer Corruption Themed</td>
<td>2005-2013</td>
<td>5</td>
<td>5</td>
<td>283</td>
</tr>
<tr>
<td>GCB</td>
<td>Global Corruption Barometer</td>
<td>2003-2013</td>
<td>8</td>
<td>1</td>
<td>349</td>
</tr>
<tr>
<td>LITS</td>
<td>Life in Transition Survey</td>
<td>2006-2010</td>
<td>2</td>
<td>2</td>
<td>43</td>
</tr>
<tr>
<td>ESS</td>
<td>European Social Survey *</td>
<td>2004-2010</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>EVS</td>
<td>European Values Study *</td>
<td>1990-2008</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>ISSP</td>
<td>International Social Survey Programme</td>
<td>2004-2009</td>
<td>3</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>WVSS</td>
<td>World Values Survey*</td>
<td>1989-2005</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>ASES</td>
<td>Asia Europe Survey</td>
<td>2000</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>CSES</td>
<td>Comparative Study of Electoral Systems</td>
<td>2001</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>QoG</td>
<td>European Quality of Government Survey</td>
<td>2010-2013</td>
<td>2</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>EB</td>
<td>General Eurobarometer</td>
<td>1997-2012</td>
<td>7</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>ISJP</td>
<td>International Social Justice Project*</td>
<td>1991-1996</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>PEW</td>
<td>Pew Global Attitudes Project</td>
<td>2002-2012</td>
<td>4</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>CCEB</td>
<td>Candidate Countries Eurobarometer</td>
<td>2003</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>CB</td>
<td>Caucasus Barometer</td>
<td>2009-2012</td>
<td>4</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>CDCEE</td>
<td>Consolidation of Democracy in CEE</td>
<td>1990-1998</td>
<td>2</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>NBB</td>
<td>New Baltic Barometer*</td>
<td>1993-2004</td>
<td>6</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>VPCPCE</td>
<td>Values and Political Change in PostcomEurope*</td>
<td>1993</td>
<td>1</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1989-2013</strong></td>
<td><strong>63</strong></td>
<td><strong>45</strong></td>
<td><strong>895</strong></td>
<td></td>
</tr>
<tr>
<td>Variable Label</td>
<td>Value labels</td>
<td>Source variables</td>
<td>Target variables</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>Corruption experience in education</td>
<td>[see Table 4.4]</td>
<td>1 = gave bribe/inf. payment</td>
<td>BRIBE-EXP</td>
<td>0.06</td>
<td>0.24</td>
</tr>
<tr>
<td>Corruption perception in education</td>
<td>[see Table 4.5]</td>
<td>1 = corruption is prevalent</td>
<td>CORR_PERC</td>
<td>0.23</td>
<td>0.42</td>
</tr>
<tr>
<td>Gender</td>
<td>LITS: 1 = male 2 = female</td>
<td>1 = female</td>
<td>female</td>
<td>0.58</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>QoG: 0 = male 1 = female</td>
<td>0 = male</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GCB: 1 = male 2 = female</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Place of residence</td>
<td>LITS: 1 = urban; 2 = rural; 3 = metropolitan</td>
<td>1 = rural</td>
<td>rural</td>
<td>0.35</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>QoG: 1 = Less than 10,000 (rural); 2 = 10,000-100,000; 3 = 100,000-1,000,000; 4 = Greater than 1,000,000</td>
<td>0 = other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GCB: 1 = rural; 2 = urban</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>LITS: age in years</td>
<td>1 = 18-29 years</td>
<td>age1</td>
<td>0.26</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>QoG: 1 = 18-29; 2 = 30-49; 3 = 50-64; 4 = 65+</td>
<td>2 = 30-49 years</td>
<td>age2</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>GCB: 1 = under 30; 2 = 30-50; 3 = 51-65; 4 = 65+</td>
<td>3 = 50 years and older</td>
<td>age3</td>
<td>0.24</td>
<td>0.42</td>
</tr>
<tr>
<td>Level of education</td>
<td>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</td>
<td>1 = Primary or less</td>
<td>edu1</td>
<td>0.23</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>QoG: 1 = lower secondary or less; 2 = medium education (higher secondary or post-secondary non tertiary) 3 = higher education</td>
<td>2 = Secondary</td>
<td>edu2</td>
<td>0.51</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>GCB: 1 = no education/basic education; 2 = secondary school; 3 = higher level education (university)</td>
<td>3 = Tertiary</td>
<td>edu3</td>
<td>0.26</td>
<td>0.44</td>
</tr>
</tbody>
</table>

**CONTROL VARIABLE**

Survey project [constructed variable] GCB_2010 GCB 0.36 0.48

**TECHNICAL VARIABLES**

Country Survey country codes Standardized ISO country codes t_cntr

Case identifier [constructed variable] t_id

Composite weights [constructed variable] t_wght