Applying Survey Data Harmonization in Corruption Research

Ilona Wysmułek
ilonawysmulek@ifispan.waw.pl

GESIS SDM team meeting, Mannheim, March 14, 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
Dissertation research in brief

Peculiarities:

→ Cross-national yet sector specific approach

→ Testing same set of hypothesis with two different measures of corruption:
  • perception of corruption in education (surveys on public opinion)
  • experience of bribe-giving (crime victimization surveys)

→ Analytical model: multi-level framework with harmonized indicators
  Accounts for:
  • characteristics of individuals,
  • contextual characteristics (including educational systems characteristics),
  • and (!) a survey project effect within the same analytical framework
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.

Survey measurement
- QoG
- GCB
- LITS

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
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.
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.

Three types of sources:

1. original sources that match,
2. sources created with intent to match (prospectively harmonized and/or harmonized *ex-ante*), and
3. secondary sources adjusted to match (harmonized *ex-post*)
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, p.1).
Survey Harmonization ex-post: WORKFLOW

Creating harmonized measures *ex-post strongly relies* on released survey data and documentation linked to it.

0. Time zero: data and documentation downloaded
1. Theoretical model and concept specification
2. Source variables: first selection; item availability template
3. Define target variable based on source variables; final source variable list
4. Cross-walk coding and additional source variables check
5. Creating harmonization control variables
6. Quality control variables for items, documentation and data
First Step:
Data Availability on Corruption in Cross-national Surveys

Aim: a systematic review of questionnaires and codebooks of international public opinion surveys in search for questions on corruption

Criteria:
- at least one question on corruption
- cover European countries [1989 – 2013]
- designed as cross-national, with representative samples
- freely available in public domain, documented in English

Sources:
→ Data Harmonization Project
→ Consortium for Political and Social Research (ICPSR)
→ GESIS Data Archive for the Social Sciences
→ ROPER Public Opinion Research Archive

+ Literature review
+ Academic consultations
<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>WVS</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 PostcommEurope*</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>
</tbody>
</table>
Second step:
Integrated dataset with indicators harmonized ex-post

• Micro-level data:
  • 3 survey projects:
    (1) Global Corruption Barometer,
    (2) Life in Transition Survey and
    (3) Quality of Government survey
  • 69 national surveys conducted in 2010
  • 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
Multi-level analytical framework with harmonized indicators

\[ \text{Prob}(BRIBE-EXP_{ij}=1|\beta_j) = \phi_{ij} \]
\[ \log[\phi_{ij}/(1 - \phi_{ij})] = \eta_{ij} \]

\[ \text{BRIBE-EXP-log}_{ij} = \gamma_{00} + \gamma_{10} \cdot \text{female}_{ij} + \gamma_{20} \cdot \text{rural}_{ij} + \gamma_{30} \cdot \text{age}_i + \gamma_{40} \cdot \text{edu}_{ij} \]
\[ + \gamma_{50} \cdot \text{GCB}_{ij} + \gamma_{60} \cdot \text{LITS}_{ij} \]
\[ + \gamma_{01j} \cdot \text{GDP}_j + u_{0j} \]
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
  ➔ corruption behavior in education sector
  ➔ and perception of corruption prevalence in education sector
Concluding remarks

Questions answered/questions asked

• New possibilities (both methodological and substantive) opened by harmonization; yet how to fully use them remains opened.

• Systematic survey data overview on corruption:
  - item level analysis: inventory of questions on corruption
  - analysis of the overlapping/rarely covered issues
  - country representation bias in corruption research

• Integrated dataset with indicators harmonized ex-post
  - analysis of the inter-survey variability of corruption measures
  - assessing and (!) increasing comparability of corruption measures (single-item indicator’s problem)
Thank you!

Ilona Wysmulek
iwysmulek@ifispan.waw.pl