Cross-national and cross-contextual analysis: controlling for sample composition and survey characteristics in comparative analysis

Ahu Alanya (Ipsos MORI)
Femke De Keulenaer (Ipsos MORI)

2nd 3MC Conference, Chicago
25-29 July 2016
Country differences in survey outcomes are not only caused by **context effects**, but also by **composition effects**.
Example 1:
Cross-country dual frame telephone survey

Potential causes of composition effects:

- Fieldwork teams applying different call scheduling (e.g. not enough evening calls for landline numbers leading to an underrepresentation of young people)

- Differences across countries in mobile phone usage/mobile phone “etiquette” (e.g. likelihood to answer when contacted by an unknown number, participating in a survey when in a public place etc.)
Example 2: Cross-country online survey (access panels)

Potential causes of composition effects:

**Differential coverage**

**Differences in selection bias**

- Not being able to use the same panel provider across countries; differences in recruitment strategies, differences in quality procedures etc.

- Differences across countries in internet use ("webographics")
Cross-country and cross-contextual comparisons

- Controlling for composition effects is typically done by adding a small set of control variables, together with a set as predictor variables, in an ordinary (OLS) regression model.

- An alternative method to compare survey outcomes across countries, while controlling for composition effects, is by using propensity score methods.

- What makes propensity score analysis appealing?
  - The propensity score model is estimated first, and thus independently from the final analysis model.
  - Allows researchers to include many control variables for more sources of composition effects (propensity score model can be complex with multiple interactions and higher order terms).
Our examples

Variables to be included in propensity score model

Cross-country dual frame telephone survey
- Demographic variables (age, activity status etc.)
- Variables explaining mobile phone usage

Cross-country online survey (access panels)
- Demographic variables (age, level of education etc.)
- Variables explaining differences in internet usage (“webographics”)
Data: Global @dvisor

- Ipsos online panels
- conducted every month in 20+ countries
- approximately 1,000 respondents in each country
- weights are provided to balance demographics to reflect the adult populations

- This example focuses on four countries with different Internet penetration: France, Germany, Turkey and the UK
Differences in coverage

% of HH with internet access, Eurostat

- UK: 91%
- Germany: 90%
- France: 83%
- Turkey: 70%
Differences in internet use

Frequency of internet use, Global @dvisor

<table>
<thead>
<tr>
<th>Country</th>
<th>Several times a day</th>
<th>Everyday or almost everyday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>80%</td>
<td>17%</td>
</tr>
<tr>
<td>UK</td>
<td>78%</td>
<td>19%</td>
</tr>
<tr>
<td>Turkey</td>
<td>75%</td>
<td>20%</td>
</tr>
<tr>
<td>France</td>
<td>73%</td>
<td>21%</td>
</tr>
</tbody>
</table>

- **Green** represents internet use several times a day.
- **Red** represents internet use everyday or almost everyday.
Differences in internet use

Visiting social network websites, Global @dvisor

- Everyday or almost everyday:
  - France: 21%
  - Germany: 18%
  - UK: 23%
  - Turkey: 26%

- Several times a day:
  - France: 31%
  - Germany: 33%
  - UK: 38%
  - Turkey: 59%
Propensity scores for multiple treatments

The `mnps` function in `twang` package

- `mnps` is part of the Toolkit for Weighting and Analysis of Nonequivalent Groups (`twang`) written by Burgette, Griffin & McCaffrey (April, 2016)
- Extends propensity score matching to more than two groups.
- `twang` methods rely on tree-based regression models that are built in an iterative fashion.
- Estimates propensity score for multiple groups (treatments) with interactions.
- `Mnps` provides useful tools (e.g. tables, graphics) to explore group differences.
Propensity score matching

Multinomial propensity scores with small set of countries

Step 1
Compare countries on socio-demographic and webographic vars

Step 2
Estimate propensity scores using GBM
(taking into account interactions btw demographic & webographic vars)

Step 3
Calculate matching weights (ATE/ATT). Assess overlap and balance across countries.

Step 4
Evaluate matched and unmatched country differences on survey outcomes
Variables used in matching

- Age
- Gender
- Marital status
- University graduate

- Internet use
- Use of social network sites
- Frequency of watching TV
- Frequency of watching online TV
- Frequency of reading printed news source
- Frequency of reading news online
- Frequency of using Twitter

main effects + interactions
Case study 1: comparisons across low/high internet penetration countries
There is not enough overlap between the countries in terms of socio-demographic and webographic vars.

Turkey data set is used as reference in matching (ATT).

Inference on country differences is limited to a subpopulation that matches the Turkish web sample.
Balance in demographic & webographic variables

Balance for UK versus unweighted TURKEY

Balance for US versus unweighted TURKEY
Survey outcomes after matching

Looked at several survey outcomes.

1. Satisfaction with government
2. Satisfaction with country
3. Rating of the current status of economy
4. Happiness
5. Right direction/wrong track
6. Subjective income
Unmatched population comparisons, weighted by socio-demographics

The way the government is running the country? Overall, how satisfied or dissatisfied?

Survey outcomes before and after matching

Source: Global Advisor 2015
Survey outcomes before and after matching

Matched comparisons, (socio-demographics+ webographics)

The way the government is running the country? Overall, how satisfied or dissatisfied

Source: Global Advisor 2015

Ipsos Public Affairs
Case study 2: comparisons across high internet penetration countries
Propensity score distributions

- There is enough overlap between the countries in terms of socio-demographic and webographic variables.
- A pseudo-population (population of all countries) used as reference in matching (ATE).
- Inference on country differences refer to the population of all countries combined.
Balance in demographic & webographic variables

Balance of UK versus FRANCE

Balance of UK versus GERMANY

Balance of FRANCE versus GERMANY
Survey outcomes before and after matching

Now, thinking about your country, overall, are you satisfied or dissatisfied with the way things are going in your country today?

Unmatched comparisons, weighted by socio-demographics

<table>
<thead>
<tr>
<th>Country</th>
<th>Satisfied</th>
<th>Dissatisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>69%</td>
<td>31%</td>
</tr>
<tr>
<td>FR</td>
<td>85%</td>
<td>15%</td>
</tr>
<tr>
<td>DE</td>
<td>56%</td>
<td>44%</td>
</tr>
</tbody>
</table>

Matched comparisons, (socio-demographics + webographics)

<table>
<thead>
<tr>
<th>Country</th>
<th>Satisfied</th>
<th>Dissatisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>71%</td>
<td>29%</td>
</tr>
<tr>
<td>FR</td>
<td>84%</td>
<td>16%</td>
</tr>
<tr>
<td>DE</td>
<td>51%</td>
<td>49%</td>
</tr>
</tbody>
</table>

Source: Global Advisor 2015
Take-home message

- Different types of composition effects confound estimates of cross-country and cross-contextual comparisons.
- In addition to controlling for socio-demographics, “webographics” can be used to assess and adjust differences in survey estimates across countries.
- An effective way to compare contexts is by using multinomial matching (note: comparison limited to a few countries).
- “Inference populations” need to be discussed while presenting results, especially when overlap between populations is low.
- Main problem remains the selection of variables to be included in the multinomial matching (what really explains the selection/composition effects?)
For more information / Questions:

Femke De Keulenaer
femke.dekeulenaer@ipsos.com