

# Comparative Surveys Fit for Analysis?

Draft

Presentation prepared for the 11<sup>th</sup> CSDI

March 21-23, 2013, Stockholm

Peter Ph. Mohler

Mannheim University & COMPASS



# Overview

The title of this paper is admittedly somewhat exaggerating, but in what sense? What does it mean for a published comparative survey to be fit for analysis? The fitness for use quality concept always takes a user/purpose perspective. Thus one reading of the title could "are published comparative surveys fit for analysis in a user perspective?". However, another reading could be "are users fit for using published comparative surveys?"

The answer is: in general yes, but there are major pitfalls. Some of them are due to the fact that producers up to now did not clearly state user profiles, i.e. the type of users for whom a survey is published and documented.

Consequently producers do not state the level of know how needed for users to benefit from the data a document that is standard for any university course. In turn users are not offered consultation and training focused on their analytical needs.

Finally, there seems to be a sort of user "self medication" tendency: because surveys are sold over the counter, one can take them like Aspirin, stomach ache included.

# Overview ctd.

- Are published comparative surveys fit for users' analysis?
  - Over the top background variables
  - Buried in documentation alarm signals
  - Unclear recommendations

# Overview ctd.

- Are users' of comparative surveys fit for data analysis?
  - Ignorance of basics in comparative research
    - no quota, please
    - weight, please
  - Aspirinitis
    - ignore all warnings, just do what you want

# Are published comparative surveys fit for users' analysis?

## Over the top background variables

- Firstly some good news
  - ESS has now Age of respondent and Education in simplified form
- Background variables
  - Education vars are lost in a flood of country specifics
  - Current paid working status (ESS)
    - Over precise definition “last seven days”
    - No simplified variable available
  - Where is Socio Economic Status (SES)?
    - Income + ISCO + Education is not “fit for use”

Are published comparative surveys fit for users' analysis?

## Alarms buried in documentation

- Mixed mode in ESS?
  - Standard should be
    - Face-to-face with differences between countries using various data collection techniques (PAPI or CASI)
- Mixed mode in ESS!
  - Supplementary questionnaire (Value items, MTMM experiments etc.)
  - Question I 1: *How was the supplementary questionnaire administered?\**

\*Documentation Round 1; Administrative Variables II

# Are published comparative surveys fit for users' analysis?

## Alarms buried in documentation ctd.

### How was the supplementary questionnaire administered?

Variable name and label: SUPQADM Administration of supplementary questionnaire

Values and categories

**1** Face-to-face interview

**2** Completed by the respondent while you were present

**3** Left with respondent, to be collected by you

**4** Left with respondent, to be returned by post

**9** No answer

**Comment:** From the Interviewer's questionnaire

\* ESS Documentation Round 1; Administrative Variables II

# Are published comparative surveys fit for users' analysis?

## Alarms buried in documentation ctd.

Yes

### **Survey Documentation**

---

ESS1-2002 Documentation Report  
App1: Population statistics and other...  
App2: Classifications and coding standards  
App3: Variables and questions  
App4: Variable lists  
Sampling Plans  
Sample data file documentation  
Reliability and validity of questions  
Design weights  
Weighting ESS Data  
Data Protocol  
More...

### **Deviations and Fieldwork Summary**

---

Overview of anomalies in data, fieldwork dates, response rates and sample sizes.

### **Classifications and indices**

---

Human Values Scale by Shalom Schwarz  
Social Class by Håkon Leiulfstrud

### **Fieldwork Documents**

---

Main questionnaire  
Supplementary questionnaire  
Split ballot design for...  
Showcards  
Fieldwork instructions  
Contact forms  
Country-specific information  
More...

### **Data Download**

---

ESS1 Integrated file [SAS] [SPSS]  
More files...

### **Online Analysis**

---

ESS1 Integrated file - Open in Nesstar

No

# Are published comparative surveys fit for users' analysis?

## Alarms buried in documentation ctd.

**Country \* Administration of supplementary questionnaire Crosstabulation**

% within Country

		Administration of supplementary questionnaire				Total
		Face-to-face interview	Completed by respondent, you present	Left with respondent, collected by you	Left by respondent, returned by post	
Country	Austria	100,0%				100,0%
	Belgium	100,0%				100,0%
	Switzerland	100,0%				100,0%
	Czech Republic	95,7%	2,9%	1,4%		100,0%
	Germany	8,6%	86,9%	4,4%		100,0%
	Denmark	100,0%				100,0%
	Spain	58,4%	37,8%	3,7%	,1%	100,0%
	Finland				100,0%	100,0%
	France	59,9%	40,1%			100,0%
	United Kingdom	20,0%	52,1%	16,9%	10,9%	100,0%
	Greece	95,6%	4,2%	,1%	,0%	100,0%
	Ireland	66,7%	23,8%	8,6%	,9%	100,0%
	Israel	90,9%	6,8%	2,3%	,0%	100,0%
	Italy	74,0%	26,0%			100,0%
	Netherlands	21,8%	75,2%	1,4%	1,6%	100,0%
	Norway	45,8%	8,9%	1,8%	43,6%	100,0%
	Poland	98,3%	1,5%	,2%		100,0%
	Portugal	100,0%				100,0%
	Sweden	1,8%	2,6%	,5%	95,1%	100,0%
	Slovenia	100,0%				100,0%
Total		64,6%	20,3%	2,2%	12,9%	100,0%

# Are published comparative surveys fit for users' analysis?

## Unclear Recommendations

- Weighting required\*

- Design weights: on page 1 “Almost yes” (?)
- After 4 pages discussion:

*For multivariate analyses the same weighting rules apply. **The data should always be weighted:** the decision between using design weights alone or design weights plus population weights depends upon the purpose of the analysis.*

\* *ESS Round 5, Weighting PDF no date, no author given*

Are published comparative surveys fit for users' analysis?

## Weighting clarification

- Interpretation PM (only?)
  - Three types of weights
    - Design
    - Population
    - Bias/non-response
  - Design weights must be applied always
  - Populations weights must be applied for pooled analysis
  - Bias/non-response weights are not mandatory (and are not provided by ESS)

# Are users' of comparative surveys fit for data analysis?

## Ignorance of basics in comparative research

- No Quota please
  - Probability sampling up to date only known method that allows to compare multiple populations (3M)
  - Non-probability samples (quota) destroy comparability

\* Kish's assumption about multi-population sampling, *Gabler, S., Haeder, S., and Lahiri, P. (1999), A model-based justification of Kish's formula for design effects for weighting and clustering, Survey Methodology, 25, 105-106..*

# Are users' of comparative surveys fit for data analysis?

## Ignorance of basics in comparative research

- Recap basic definition of probability sampling
  - For each unit (respondent) the selection probability is known and not zero
- Quota samples are non-probabilistic because:
  - For some units the selection probability is zero (quota cell already full)
- Stratification is standard in probability sampling
  - Not to be confused with quota

Are users' of comparative surveys fit for data analysis?

## Quota

- Examples are WWS & EVS
- WVS Example Argentina

Within each household, the respondent is selected at random, although **quotas by gender and age** have been settled in order to guarantee that the sample is representative of the population, in terms of these both variables

- Others are United Kingdom, Finland, USA (Panel Quota), France

<http://www.wvsevsdb.com/wvs/WVSTechnical.jsp>

# Are users' of comparative surveys fit for data analysis?

## Quota

### Recent research

Current literature by leading scientists ignores the quota problem....

[\[PDF\] World values surveys and European values surveys, 1981-1984, 1990-1993, and 1995-1997](#) [\[PDF\] von faith-health.org](#)

[R Inglehart, A de Estudios Sociales - 2000 - faith-health.org](#)

... SAMPLING: Both national random and **quota** sampling were used. ... For access to the complete **surveys** for earlier waves, users should consult **WORLD VALUES SURVEY, 1981-1983** (ICPSR 9309) and **WORLD VALUES SURVEY, 1981-1984 AND 1990- 1993** (ICPSR 6160). ...

Zitiert durch: 326 Ähnliche Artikel Alle 43 Versionen Zitieren Mehr▼

[Measuring Political Culture in Multiethnic Societies Reaggregating the World Values Survey](#)

[\[PDF\] von uh.edu](#)

[BD Silver, KM Dowley - Comparative Political Studies, 2000 - cps.sagepub.com](#)

... Some countries used **quota** samples; others favored certain regions or the urban population. ...  
Silver, Dowley / REAGGREGATING THE **WORLD VALUES SURVEY** 523 Table 1 Distribution on New Ethnic Variable for 16 **World Values Survey** Societies ...

Zitiert durch: 48 Ähnliche Artikel Alle 3 Versionen Zitieren

[Subjective well-being and age: An international analysis](#)

Neither the data publishers nor prominent data users take the Quota issue seriously

Are users' of comparative surveys fit for data analysis?

## Weighting

- Weight, please
  - Recall weighting recommendations
    - (Design and Population weights are mandatory)
  - ESS Conference Cyprus
    - Keynote speaker reported pooled analysis
    - No comments on weighting
    - Similarly other presentations
  - Good news
    - Recent serious literature indicates that weighting was done

Are users' of comparative surveys fit for data analysis?

## Aspirinitis

- As documentation can be overwhelming (see buried in documentation)
- One can be tempted to avoid headache by ignoring all warnings and just do what you want
- But beware of stomach ache, do read the red flags in the variable deviation table

# Are users' of comparative surveys fit for data analysis?

## Aspirinitis

Greece**	06.05.11-05.07.11	2 715	65.6	
Hungary***	19.10.10-10.12.10	1 561	49.2	F6 (RSHPSTS), F11 (MARSTS), MARITALB), F53 (JBSPVP), G33 (LSINTJB), G86 (ICB1964), G87 (AGERTR), G88 (PLNCHLD), H1 (TESTD1) to H36 (TESTC39), Fieldwork
Ireland	20.09.11-31.01.12	2 576	65.2	F6 (RHIPSTS), F7 (LVGPTNEA), F11 (MARSTS), Interview time (INWSHH, INWSMM, INWEHH, INWEMM, INWTM)
Israel	09.01.11-13.06.11	2 294	72.9	
Lithuania	21.04.11-20.08.11	1 677	39.41	G88 (PLNCHLD)
Netherlands	27.09.10-02.04.11	1 829	60.0	
Norway	09.09.10-15.02.11	1 548	58.0	D17 (PLCEXDC), Item non-response
Poland	01.10.10-06.02.11	1 751	70.3	Item non-response
Portugal	11.10.10-23.03.11	2 150	67.1	F41 (HINCTNTA), G85 (WNTRTR), (SUPQDD), (SUPQMM), (SUPQYR)
Russian Federation	24.12.10-14.05.11	2 595	66.6	F29 (WKHCT), Item non-response
Slovakia	29.10.10-28.02.11	1 856	74.7	F57 (WKHTOTP)
Slovenia	20.10.10-31.01.11	1 403	64.4	C10 (CRVCTEF)
Spain	11.04.11-24.07.11	1 885	68.5	F41 (HINCTNTA), F63 (OCCF14B), F69 (OCCM14B), G86 (ICB1964), G87 (AGERTR), G88 (PLNCHLD)
Sweden	27.09.10-01.03.11	1 497	51.0	C18 (RLGDNM), C20 (RLGDNME), G86 (ICB1964), G87 (AGERTR), G88 (PLNCHLD), (INTAGEA), (INTGNDR)



# Conclusions

- Fit for use/purpose of comparative surveys
  - Partially
- Due to
  - Missing definition of “intended user”
  - Missing know how of users about the “intended use”
- Silver Linings

*“If clouds are blocking the sun, there will always be a silver lining that reminds me to keep on trying.”*

