

# Mixed-Mode Surveys: An Overview of Design, Estimation and Adjustment Methods and Empirical Applications

---

Z. Tuba Suzer-Gurtekin, University of Michigan

Edith D. de Leeuw, Utrecht University

Richard Valliant, University of Michigan

Steven G. Heeringa, University of Michigan

Ana Villar, City University London

# Outline

---

- Background
  - Data
  - Literature on mode effects specific to subjective well-being
- Methods
- Conclusions

# Background

---

- The International Social Survey Programme (ISSP) is a good example of cross-national collaboration (Bandilla & Bosnjak, 2003; Lauer & Yodanis, 2004).
  - A cross-national survey where there is a recommended set of methodologies to increase comparability
  - Variations given the individual budget and resource constraints
- ISSP 2011 surveys include modules on evaluation of the health care system in the country, personal health, and on health insurance
- Core questionnaire and country-specific questionnaires are available online as are data from 32 countries
- Nine out of 32 countries reported a mixed-mode strategy for data collection

	Freq	%
<b>Belgium (n=3083)</b>		
SC, arrives with I, drops-off, mailed back by R	1,210	39.25
SC, mailed to, mailed back by R	1,873	60.75
<b>Denmark (n=1388)</b>		
SC, mailed to, mailed back by R	24	1.73
SC, web questionnaire	1,297	93.44
CATI	67	4.83
<b>Germany (n=1681)</b>		
CAPI, no visuals	272	16.18
SC, CASI	1,409	83.82
<b>Taiwan (n=2199)</b>		
CAPI, interpreter or translator, no visuals	5	0.23
CAPI, no visuals	1,532	69.67
CAPI, R reading questionnaire (paper or on monitor)	1	0.05
CAPI, visuals	311	14.14
PAPI, no visuals	314	14.28
PAPI, visuals	30	1.36
PAPI, interpreter or translator, no visuals	6	0.27
<b>UK (n=936)</b>		
SC, arrives with I, drops-off, mailed back by R	99	10.58
SC, arrives with I, drops-off, picks up later	307	32.80
SC, arrives with I, I attending	530	56.62
<b>US (n=1550)</b>		
CAPI, phone	311	20.06

# Web-Mail Surveys

---

	Freq	%
<b>Finland (n=1340)</b>		
SC, mailed to, mailed back by R	612	45.67
SC, web questionnaire	728	54.33
<b>Total</b>	<b>1,340</b>	<b>100.00</b>
<b>Italy (n=1186)</b>		
SC, arrives with I, drops-off, mailed back by R	968	81.62
SC, web questionnaire	218	18.38
<b>Total</b>	<b>1,186</b>	<b>100.00</b>
<b>Norway (n=1834)</b>		
SC, mailed to, mailed back by R	1,345	73.34
SC, web questionnaire	489	26.66
<b>Total</b>	<b>1,834</b>	<b>100.00</b>

# Data

---

- For Italy the survey population is defined as 16 years and older, for Finland it is between 15 and 74, and for Norway 18 and older
- Italy applied a four stage stratified sampling design
  - Respondent level weighting includes poststratification based on age (18 groups), place of residence (14 geo-types), and gender
- Finland selected persons using systematic sampling from the population register sorted by domicile code and birthdate.
  - 95% of the interviews were completed in Finnish and the remaining in Swedish
  - Weights include calibration by gender, age, region and type of community
- Norway applied a random sample of 3,800 individuals from the Central Register of Persons, born 1934 - 1993.
  - No weighting

# Operationalization

---

- In general, would you say your health is  
(PLEASE TICK ONE BOX ONLY)
  - 1  Excellent
  - 2  Very good
  - 3  Good
  - 4  Fair
  - 5  Poor
  - 8  Can't choose
  - 9  No answer.
- In this study, we quantify well-being as the proportion of excellent and very good in the general health question.

# Background

---

- Dolan & Kavestos (2012) found higher reporting of subjective well-being in telephone interviews compared to in-person interviews.
- Furthermore, cross-cultural survey research has studied the impact of response styles in subjective well-being (Jurges, 2007)
- Research in response styles by survey mode has not been widely integrated in the cross-cultural response styles literature (Weijters, Schillewaert, & Geuens, 2008; Ye, Fulton, & Tourangeau, 2011)
- So the empirical evidence on possible response style differences by mode in cross-cultural surveys for the subjective well-being is scarce to the best of our knowledge
- Yet, differences in response styles may threaten the data comparability.
- Here we focus on subjective well-being measured by the overall health question, as this measure may be sensitive to mode effects in a mixed-mode Web-Mail design (Messer & Dillman, 2011; Millar & Dillman, 2011)
- Both web and mail (paper and pencil) are self-administered data collection modes and use the visual communication channel. Therefore, the impact of possible sources of mode effects, presence of interviewer and differences in communication channel, could be expected to be minimal (De Leeuw, & Hox, 2011; Hox et al., 2017)

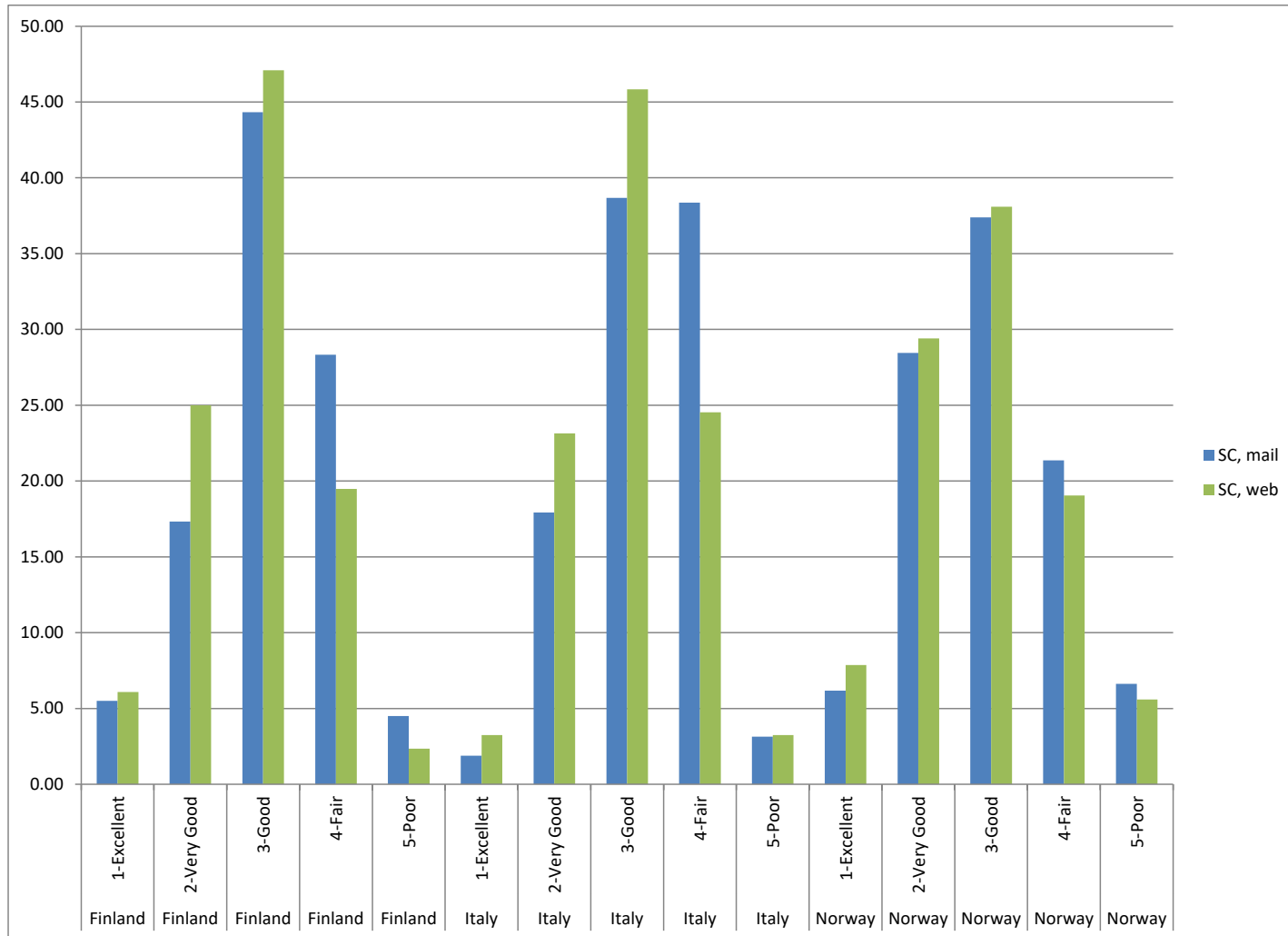


# Analysis Plan

---

- 1) Descriptives
- 2) Unadjusted self-reported health status proportions
- 3) Propensity score matching of web and mail respondents given a set of covariates to evaluate mode effects
- 4) Calibration of web and mail respondent proportions across countries in self-reported health status ratings
- 5) Counterfactual (potential) outcome imputation to adjust for mode effects in proportion of persons reporting Excellent or Very Good health status

# Differences in Subjective Well-being by Mode



# Respondent Characteristics by Mode

---

- A web respondent is more likely to be a male, younger, married, employed and have higher education level compared to a mail respondent
- This demographic profiling is consistent across three countries
- Web and mail respondent characteristics seem to differ by characteristics that could be also related to subjective well-being and mode effects
- That is, there seems to be substantive selection effects that should be adjusted to determine whether the mode effects are significantly different given a set of covariates

# Item Missing Percentages

	SC,mail	SC,web	SC,mail	SC,web
	%	%	N	N
Country specific personal income: Finland				
Missing	9.97	4.4	61	32
0-1150 Euro	23.37	21.84	143	159
1151-2000 Euro	26.47	23.08	162	168
2001-2999 Euro	20.92	21.02	128	153
>=3000 Euro	19.28	29.67	118	216
Country specific household income: Finland				
Missing	42.81	31.04	262	226
0-3000 Euro	17.81	16.21	109	118
3001-4400 Euro	14.71	14.29	90	104
4401-6100 Euro	15.52	17.03	95	124
>=6200 Euro	9.15	21.43	56	156

# Propensity Score Matching

---

- 1:1 Nearest Neighbor Matching (Parsons, 2005; Stuart, 2010)
- Diagnostics based on subclasses
- Effect size (difference between Web and Mail responses on the average after matching) is not significant
- Balance samples

# Calibration Estimator

---

- Buelens and Van den Brakel (2011, 2015) extended the classical GREG estimator for the mean of Y using arbitrary population controls for the number persons choosing each mode
- Method aims to cancel any differential measurement errors when differences are estimated so that they reflect the true differences
- If same mode control values are used for surveys at different time periods, intent is that mode effect will subtract out when taking differences of estimates for two time periods
- Buelens and Van den Brakel (2011) chose the controls to be equal to the reference survey response mode proportions conducted in their example application
- They also acknowledge a more appropriate method would be based on experimental designs

# Counterfactual (potential) Outcome Imputation to Adjust for Mode Effects in Proportion of Persons Reporting Excellent or Very Good Health Status.

---

$Y_1$ (Web)	$Y_2$ (Mail)
$R_W$	$NR_M$
$NR_W$	$R_M$

- A special case of a missing data problem
- Impute data for each phase through a series of multiple imputation models as if all units had reported in that particular mode
- Impute nonrespondent data for Telephone and In-person phases via multiple imputation models
- X covariates in the models are combination of personal and residential data (such as age, gender, etc.)

# Counterfactual (potential) Outcome Imputation to Adjust for Mode Effects in Proportion of Persons Reporting Excellent or Very Good Health Status.

---

$Y_1$ (Web)	$Y_2$ (Mail)
$R_W$	$NR_M$
$NR_W$	$R_M$

$$\bar{y}_W^* = \frac{1}{N} \left[ \sum_{j \in U_W} y_j + \sum_{j \in U_M} y_{jW}^* \right]$$

$$\bar{y}_M^* = \frac{1}{N} \left[ \sum_{j \in U_M} y_j + \sum_{j \in U_W} y_{jM}^* \right]$$

$$E_M E_I [\bar{y} - \bar{y}_W^*] = \bar{B}_W$$

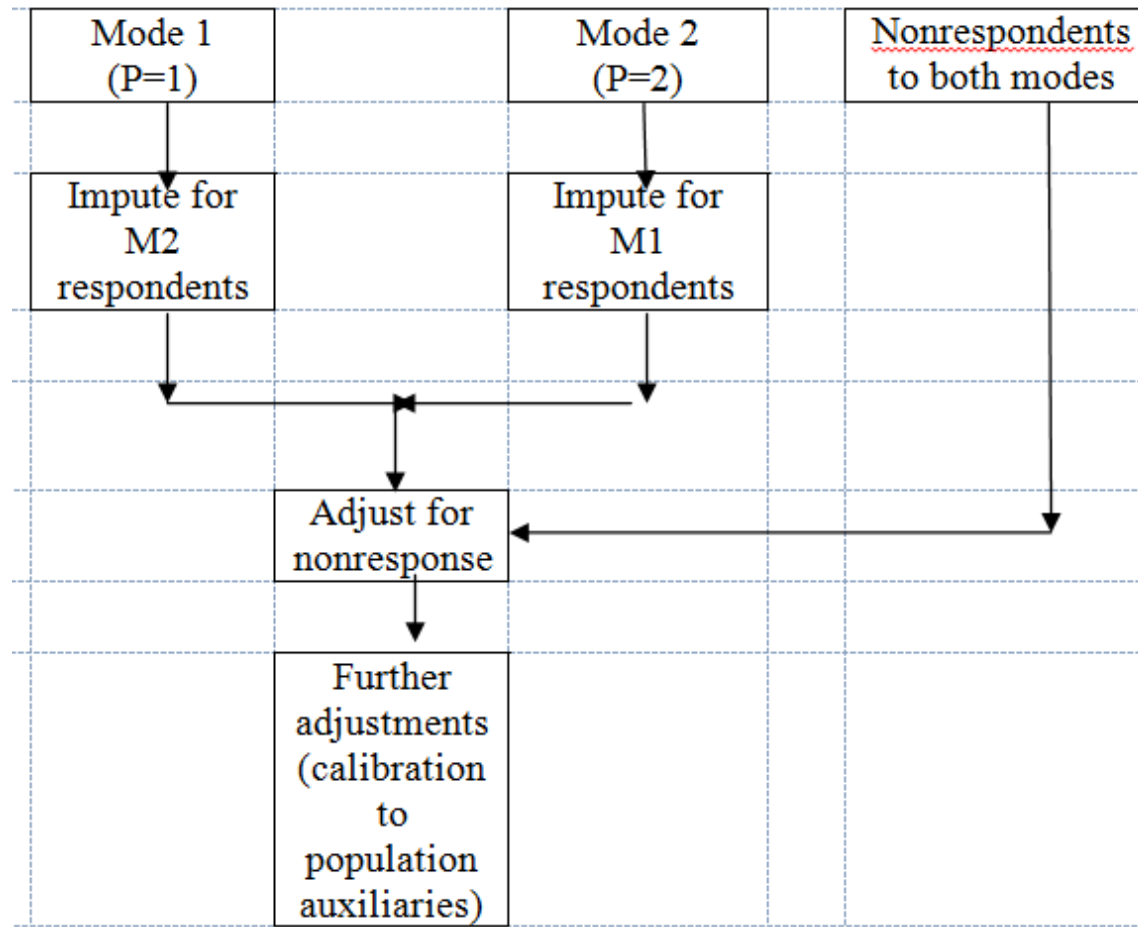
$$E_M E_I [\bar{y} - \bar{y}_M^*] = \bar{B}_M$$

## Combine mode specific estimates

$$\bar{y} = \alpha \bar{y}_W^* + (1 - \alpha) \bar{y}_M^* \text{ where } 0 \leq \alpha \leq 1$$



# Schematic Chart for the Imputation Method



# Estimated Proportion of Persons Reporting Excellent or Very Good and Standard Errors using Unadjusted data, Calibration and Multiple Imputation Methods, International Social Survey Programme (ISSP) 2011

	Unadjusted		Adjusted (Calibration)		Adjusted (Multiple Imputation)	
	Proportion	SE	Proportion	SE	Proportion	SE
<b>Finland</b>	0.28	0.013	0.27	0.012	0.29	0.024
<b>Norway</b>	0.35	0.011	0.36	0.013	0.33	0.039
<b>Italy</b>	0.21	0.012	0.23	0.017	0.21	0.024

# Conclusions

---

- Diagnosis and adjustment are related to the ignorability of mode effects in the context of mixed-mode surveys.
- There are two general paths to follow:
  - 1) Explicit data collection (De Leeuw, Hox, and Dillman, 2008; Vannieuwenhuyze, Loosveldt and Molenberghs, 2012; Klausch, 2014)
  - 2) Use available data. In this short course, we present some methods that use available data
- Variable specific
- Conditioned on data
- Model assumptions

Comments? Questions?

[tsuzer@umich.edu](mailto:tsuzer@umich.edu)