EUROPEAN MASTER IN OFFICIAL STATISTICS
EMOS Webinar

The mode effect in mixed-mode surveys

Claudia De Vitiis <devitiis@istat.it>
Francesca Inglese <fringles@istat.it>

May 26 2020

OUTLINE

A. The mode effect: definitions, assesement and adjustment

1. Mixed mode survey designs
2. Mode effect in mixed mode surveys
3. Mode effect formalization
4. Mode effect assessment
5. Methods to disentangle selection and measurement effects
6. Methods to adjust mode effect
7. Focus on Propensity score method

B. Case study: an Experimental design with parallel independent surveys

A. The mode effect: definitions, assesment and adjustment

1. Mixed mode survey designs

Definition: The use of different data collection thecniques in the same survey

Thecniques for questionnaire administration

P.A.P. Paper and Pencil (postal)

P.A.P.I. Paper and Pencil Interviewing (with interviewer, face to face)

C.A.P.I. Computer Assisted Personal Interviews (with interviewer, face to face)

C.A.T.I. Computer Assisted Telephone Interviews

C.A.W.I. Computer Assisted Web Interviews

The choice of technique must be made according to the objectives of the survey and the characteristics of the population, in order to maximize quality and limit burden on respondents and costs.

Advantages of Mixed mode

- Contrast declining response and coverage rates (the sampling units can be contacted in the most suitable way for each of them).
- Reduce the cost of the surveys, for example introducing web mode.

Mixed mode designs

The combination of different data collection mode in the same survey

- **Concurrent:** the modes are assigned ex ante to the unit of the sample (different modes start at the same time)
    a) randomly
    b) on the basis on a priori known characteristics, often on contact variables; separate samples (generally independent)
- **Sequential**: the same mode is proposed to all the sampling units and then the non-respondents are re-approached using a different mode; single sample

Complexities of MM designs

- The estimates must be **consistent and comparable** with the analogue ones obtained in the previous survey editions, for ensuring that any changes in the time series are exclusively due to real changes of the observed phenomenon.

- **In the design phase** actions should be taken to reduce the risk of introducing non-sampling errors in the survey

  Need to define the best **data collection instruments** to reduce the risk of measurement errors (Uni-mode or Mode-specific questionnaires)

- **In the estimation phase** analyses and possibly adjustments of the errors introduced by the different techniques are needed to ensure the **accuracy of the estimates**

  Need of accuracy in the inferential process

2. Mode effect in mixed mode surveys

Mode effect refers to the introduction of **bias effects** on the survey estimates due to observational and non observational errors.

- **Selection effect**: different coverage errors and total non response in each technique (non observational errors), that determine a difference in the composition of the observed samples (desirable aspect of MM strategy).
- **Measurement effect**: different measurement errors due to the modes of survey administration (observational errors), that introduces systematic differences in data. **Measurement error** refers to the influence of a survey mode on the answers of the respondents, such that one person would give different answers in different modes (***interviewer effect*** and ***social desirability***, ***primacy*** *and* ***recency effects***, recall bias, etc.).

The assessment of **mode effect is complex** as the **two effects are confused**. Consequence of this problem is that a **discrepancy in the estimates** calculated on respondents can be caused by the different composition of the samples or by differences in measurement errors (de Leeuw, 2005; Weisberg, 2005).

Need to **disentangle the two effects** for the estimation of measurement errors

3. Mode effect formalisation

Selection and measurement effects are contained in the **relationship between the survey variable and the mode**, but they act in a dissimilar way: the measurement effect produces differences in the estimated means on the respondents, while the selection effect, influencing the correlation structure of the data, produces measures not equivalent among the respondents (Hox et al., 2015).

The equivalence of mode measurement => **reference concept for mode effect analysis**

- Can be illustrated by introducing a simple measurement model (Klausch, 2014)

$$E(y_{m_1}) = E(y) + MB_{m_1}$$

$MB_{m_1} = E(y_{m_1}) - E(y)$ systematic measurement error of m_1 mode

The equivalence of measurement with two modes is achieved when the measurement errors associated with them are the same:

$$MB_{m_1} = MB_{m_2} \iff E(y_{m_2}) - E(y_{m_1}) = 0$$

3. Mode effect formalisation (2)

For the estimate of the average of a generic variable y, starting from the decomposition of the total bias, we can define **the estimator of the total mode effect** and its components

- We define the average obtained as a combination of data collected with modes m_1 and m_2

$$\bar{y}_{mm} = \pi_{m_1}\bar{y}_{m_1} + \pi_{m_2}\bar{y}_{m_2}$$

π_m proportion of respondents and $\bar{y}_m$ sample mean for *m* (*m*=1,2)

- We define the expected value of the sample mean estimator for respondents with mode m_1

$$E\left(\hat{\bar{y}}_{m_1}\right) = E\left(y_{m_1} \middle| R_{m_1} = 1\right)$$

3. Mode effect formalisation (3)

The **total bias of estimator** $\hat{\bar{y}}_{m_1}$ can be expressed (the same for mode m_2)

$$TB\left(\hat{\bar{y}}_{m_1}\right) = E\left(y_{m_1} \middle| R_{m_1} = 1\right) - E(y) =$$

$$= E\left(y_{m_1} \middle| R_{m_1} = 1\right) - E\left(y \middle| R_{m_1} = 1\right) + E\left(y \middle| R_{m_1} = 1\right) - E(y) =$$

$$= MB_{m_1} + SB_{m_1}$$

Measurement error conditional on response

Selection error with respect to the population mean

➢ The impact of the mixing of modes on the total survey error of an estimate depends on how the specific errors associated with each modes operate

3. Mode effect formalisation (4)

Mode effect in a concurrent mixed mode design

We consider **the estimator of the total mode effect** due to m_1 e m_2 (Klausch, 2014)

$$\boxed{\hat{T}_{SM} = \hat{\bar{y}}_{m_2} - \hat{\bar{y}}_{m_1}}$$

with expected value

$$E\left(\hat{T}_{SM}\right) = \left(E\left(y_{m_2} \middle| R_{m_2} = 1\right) - E(y)\right) - \left(E\left(y_{m_1} \middle| R_{m_1} = 1\right) - E(y)\right) =$$

$$= \left(TB_{m_2} - TB_{m_1}\right) = \left(MB_{m_2} - MB_{m_1}\right) + \left(SB_{m_2} - SB_{m_1}\right)$$

ME, measurement effect

SE, Selection effect

MB_m Measurement error of mode *m*

SB_m Selection error of mode *m*

3. Mode effect formalisation (5)

Mode effect in a concurrent mixed mode design

If we consider the **true value** to be the measurement obtained with a **reference mode** (or benchmark), m_1 for example, we can write

$$ME = E\left(y_{m_2} \middle| R_{m_2} = 1\right) - E\left(y_{m_1} \middle| R_{m_2} = 1\right) = ME_{m_2}$$

$$SE = E\left(y_{m_1} \middle| R_{m_2} = 1\right) - E\left(y_{m_1} \middle| R_{m_1} = 1\right) = SE\left(y_{m_1}\right)$$

- Measurement effect ME_{m_2} conditional on respondents with m_2 can be view as a variation of bias due to measurement error in m_2.
- Selection effect $SE\left(y_{m_1}\right)$ respect to the variable measured with m_1 is a variation of bias due to the selection error generated by the use of the mode m_2, instead of the m_1 mode for measuring y_{m_1}
- $E\left(y_{m_1} \middle| R_{m_2} = 1\right)$ is a "**counterfactual**" quantity, also called "potential result", which in reality is not observed, but which under certain conditions can be estimated (Klausch, 2014).

4. Mode effect assessment

Different contexts and approaches

Design type	Objective
Experimental	
Parallel independent surveys (single mode and mixed mode)	Mode assessment Mode adjustment
Re-interview study - repeated measurement designs	
Other (Embedded experiments, Split sample designs)	
Non-experimental	
Observational studies (Mixed-mode design only)	Control for selection effects through weighting or regression-based inference methods Adjusting for measurement effect

5. Methods to disentangle selection and measurement effects (1)

Method	Analysis	Conditions	Context
Weighting **- Propensity score (PS)** **- Calibration** **- Post-stratification**	Analysis based on response model to control for respondent characteristics (comparable samples)	MAR assumption Mode-insensitive auxiliary variables Balancing assumption in PS	Observational studies
Regression model (Kolenikov, Kennedy, 2014)	Model analysis to estimate measurement and selection errors	Mode-insensitive auxiliary variables to control selection effect	Observational studies
Other methods - Use of outcome regression with a propensity score model	Model to estimate causal effect	Appropriate statistical models	Observational studies
Instrumental variable approach (Vannieuwenhuyze et al., 2010)	Analysis based on benchmark single-mode design	Validity of comparability and representativity assumptions	Parallel independent surveys
Re-interview (Biemer, 2001) ***Re-interview data combined with administrative data and paradata.***	Estimate **Measurement effect** - as remaining difference between modes. Estimate **Selection effect** - using mix of re-interview data, administrative data and paradata.	Re-interview does not affect measurement behavior of respondent. Nonresponse to re-interview is unrelated to variables of interest given administrative data and paradata.	Re-interview of subset of mixed-mode respondents

5. Methods to disentangle selection and measurement effects (2)

Theoretical framework in observational studies

- From an inferential point of view the **selection and measurement effects** need to be investigated separately, to obtain a correct formulation of the total non-sampling error and to apply methods to adjust the estimates of the parameters of interest for the bias effects.
- The problem of the **confounding** between the two effects is the central theme of the theory of **causal inference** (*Pearl, 2009*).
  - The **measurement error** is conceptualized as a **causal effect** of the mode on the survey variable, while the **selection effect** is seen as a spurious correlation between the target variable and the mode.
  - For the estimation of the two effects causal inference is used according to a **counterfactual perspective**: the existence of a potential result not really observed (the value that the respondent would have provided with the other mode) is hypothesized.

5. Methods to disentangle selection and measurement effects (3)

Requirement for observational studies

- **Covariates that explain the selection mechanisms**. If available, differences between mode groups are attributed to measurement differences, conditional on the covariates. Validating this assumption can be achieved when variables that are observed without error are available, potentially obtainable from external data sources (frame data and administrative data and paradata from the contact and participation processes).
- Types of **auxiliary variables (causal inference)** (Vannieuwenhuyze et al., 2014):
  - ✓ **Confounding variables** (mode-insensitive), which that explain the selection effect as a common cause of the survey variable and mode, are control variables (back-door model).
  - ✓ **Intermediate variables**, which explain the measurement error as an intermediate variable between survey variable and mode, have the function of clarifying the nature of the relationship between the independent variable and the dependent variable of a model (front-door model).

6. Methods to adjust mode effect

Method	Aim	Conditions
Weighting **- Propensity score** **- Calibration** **- Post-stratification**	To equate samples To correct selection effect	Ignorability of selection mechanism (MAR) Mode-insensitive auxiliary variables Measurement error negligible
Mode calibration **(Buelens and van den Brakel, 2015, 2017)**	To stabilize the selection effect in repeated surveys	Independence of measurement and selection error Time-stability of measurement error
Regression **(Kolenikov and Kennedy, 2014)**	To estimate measurement and selection effects To correct measurement error	Appropriate statistical models
Other methods -Use of outcome regression with a propensity score model	To estimate causal effect To correct measurement error	Appropriate statistical models
Multiple imputation		
Multiple (standard) imputation (Rubin, 1987)	To predict counterfactual data (potential outcomes) To correct measurement error	Choice of benchmark mode MAR assumption
Multiple imputation with response and selection models proposed by Suzer-Gurtekin et al. (2012)		Choice of benchmark mode Sequential design and two modes (Possibility – non-ignorability of selection mechanism)
Fractional multiple imputation proposed by Park et al. (2016)		Sequential design and more than two modes Possibility – non-ignorability of selection mechanism

Methods for adjusting selection effect - Weighting methods

- Propensity score, calibration of weights modified through the correction factors (*Rosenbaum and Rubin, 1983 - Vandenplas et al., 2016*)
- Standard calibration on demographic totals
- Calibration on fixed levels of mode proportions (method proposed by *Buelens and Van den Brakel, 2015*), to stabilize the selection effect in repeated surveys, assuming the invariance of measurement effect, with the aim to obtain reliable changes over time

✓ Assuming the hypothesis of ignorability of the selection effect and absence/stability of measurement effect

Are there any questions on what has been shown so far?

7. Focus on estimating mode effects : the Propensity score method

Used for disentangling selection and measurement effects

Weigthing method (Rosenbaum and Rubin, 1983; Rubin, 2006, Vandenplas et al., 2016)

- Aim: to make comparable the respondents to the different modes
- Assumption: differences between comparable groups are due only to measurement errors

Method

- **Equate the distributions of the samples of respondents with different modes, conditionally to a set of auxiliary variables** explaining the selection effect (confounding variables).

Hypothesis

1. **Ignorability** of selection effect (MAR);
2. **Invariance of the measurement error** determined by the mix of **modes over time** (not very sustainable in sequential mixed mode survey, as the composition of respondents by mode may change in subsequent editions of the survey).

7. Propensity score method (2)

In mixed mode survey

- The method is part of the **causal inference in a counterfactual perspective** such as an application of the **back-door model** (confounding auxiliary variables that are mode insensitive).

- Propensity score subclassification (Rosenbaum and Rubin, 1983**)** used **to disentangle selection and measurement effects**, assuming ignorability of selection effect

- Propensity score can be interpreted as the **probability of assigning** (or choosing) **a mode conditionally on the observed covariates**.

- The propensity to respond to one of the modes is modeled using a **logistic model** and the propensity score is obtained as a predicted probability.

- Predicted probabilities are used to **create groups of comparable respondents to modes** m_1 (reference mode) and m_2:
  - **Matching** (among respondents to modes)
  - **Stratification** (percentiles of distribution)

Equivalence - comparability

The answers provided with the m_2 mode must be made equivalent to those that would have been provided by the m_1 mode (counterfactual data).

- m_1 is considered as **reference mode**.
- **Balancing condition within the strata** (in each stratum, the independence between each auxiliary variable and mode must be verified with an appropriate statistical test (for example, the chi-square test for categorical auxiliary variables).
- In each balanced stratum *k*, the **weight is calculated** which makes the weighted proportion of respondents to mode m_1 equivalent to the proportion of respondents to mode m_2 in the same stratum:

$$w_k = \frac{n_{k,m_2}}{n_{m_2}} \Big/ \frac{n_{k,m_1}}{n_{m_1}}$$

7. Propensity score method (4)

With **weight**, w_k , in balanced strata, it is possible to estimate the two components of mode effect - selection and measurement effects (Vandenplas *et al.*, 2016).

Selection effect - Difference between weighted and unweighted estimates of respondents to mode m_1

$$SE_{m_1}\left(\bar{Y}\right)=\frac{\sum_{i=1}^{n_{m_1}} y_{i,m_1}}{n_{m_1}}-\frac{\sum_{i=1}^{n_{m_1}} w_{k,i}\, y_{i,m_1}}{n_{m_1}}$$

Measurement effect - Difference between the weighted estimate of the respondents to mode m_1 and the unweighted estimate of the respondents to mode m_2

$$ME_{m_1}\left(\bar{Y}\right)=\frac{\sum_{i=1}^{n_{m_1}} w_{k,i}\, y_{i,m_1}}{n_{m_1}}-\frac{\sum_{i=1}^{n_{m_2}} y_{i,m_2}}{n_{m_2}}$$

B. CASE STUDY: an experimental design with parallel independent samples

CASE STUDY: an experimental design

Objective:

- To evaluate the impact of the introduction of mixed mode design in a social sample survey, **"Aspects of Daily Life, 2017"** annual survey, traditionally PAPI

Type of experiment:

- Parallel independent samples
  - ✓ **Single mode (SM)** - PAPI
  - ✓ **Mixed mode (MM)** - sequential web/PAPI)

> **Analyses and Models**
>
> **Some results**

Case Study, an experimental design

The sample survey "Multipurpose survey on households: Aspects of daily life"

- Collects yearly information about recreational and cultural activities in free time, such as sports, reading, cinema, music, the Internet, social relations, issues for the quality of people life
- Based on a sample of about 24.000 households, selected through a two stage sample design (municipalities/households) from the centralized municipal register (LAC)
- Mixed technique: sequential web-PAPI
  - ✓ A self-compiled questionnaire (web) proposed in the inviting letter sent by ISTAT and after, on non respondent households, direct interview with a questionnaire on paper with an interviewer (PAPI)
- In 2017 **experimental set up**: **sequential web/PAPI (MM)** with a **control single mode** (**SM**) sample PAPI
- The selected **sample** of individuals was **linked to an administrative data base** through the individual code available from the selection frame to obtain external auxiliary variables

Case Study, an experimental design

Analysis framework: Summary scheme of the experimental context and analyses

Parallel independent samples (SM/MM)	Mixed-mode: Sequential web-PAPI; Control sample: Single mode (PAPI)
Main goal of the analyses	– Evaluation of the impact of the switching from single to mixed mode – Evaluation of total non-sampling error components (measurement)
Theoretical context	Instrumental/Counterfactual approaches
Available auxiliary information	Register demo-social covariates
Phases of the analyses (target variables)	– Comparison between the SM and MM samples ▪ tests on the differences in the estimates SM and MM ▪ study of the total nonresponse bias – Analyses on the univariate distributions and multivariate structure of data – Assessment of the mode effect, disentangling selection and mesurement (propensity score and instrumental variable)
Phases of the adjustment	– Adjusting for selection effect in the MM design through weighting (standard calibration, fixed mode proportions and propensity score)

Experimental design: the impact on response rate

Response rates by geographical area

Geographical area	Response rates		
	SINGLE MODE/PAPI	MIXED MODE	
		web	final
North West	65.9%	32.5%	71.2%
North East	70.2%	36.0%	73.6%
Center	68.6%	27.8%	70.2%
South	79.3%	17.7%	79.4%
Islands	71.3%	17.3%	74.2%
ITALY	**71.0%**	**26.8%**	**74.0%**

Experimental design: auxiliary variables for the analyses

The auxiliary variables available for the following analyses and models

Auxiliary mode-insensitive variables in ADL survey at household level:

- ✓ Household type: one-component under 55, one-component over 54, couple with children at least one under 25, couple with children without under 25, couple without children, one parent at least one under 25, one parent without under 25, other types
- ✓ Higher education level: below/equal/above high school diploma
- ✓ Occupation type: Prevalence of: employed, self employed, not in labor age, mixed types
- ✓ Municipal type: Metropolitan cities, metropolitan area, other municipalities <2000, 2000-10000, 10000-50000, >50000
- ✓ Geographical area (North, Center, South and Islands)
- ✓ Income class: 5 quintiles (€ 11.955, 20.892, 30.028, 46.119)
- ✓ Citizenship (nationality): Italian/Foreign household

Experimental design: the impact on total nonresponse bias

Analysis of total nonresponse bias – *R*-indicators

- *R-indicators* (*Schouten et al., 2011*) are based on a measure of the variability of the response propensity and describe how the sample of respondents to a survey reflects the population of interest with respect to certain characteristics.

$$\hat{R}(\hat{\rho}_x) = 1 - 2\hat{S}(\hat{\rho}_x)$$

  - $\hat{\rho}_x$ is the response propensity estimated thought a logistic regression model
  - $\hat{S}(\hat{\rho}_x)$ is the estimate of standard deviation of $\hat{\rho}_x$

R-indicators in SM and MM samples

response models defined at national level

response models defined for each geographical area

*R*_Indicator	SM sample	MM sample
Italy	0.812	0.852
North	0.847	0.840
Center	0.752	0.842
South and Islands	0.840	0.907

- MM sample is more representative respect to the SM sample

Experimental design: the impact on univariate and multivariate distributions (1)

Users' interest is generally the relations among variables, studied through statistical models

- **What is the impact of data collection design on distributions and/or associative structure of the variables? (*Martin and Lynn, 2011*)**

Univariate analysis - impact of mixed-mode design (SM/MM) on the distributions of ADL variables

- Regression models, with the survey variable as the dependent variable and a dummy variable "survey design" as the independent variable
  - ✓ appropriate statistical models and tests to evaluate if the distributions are significantly different

Multivariate analysis - impact of mixed-mode design (SM/MM) on the estimation of models

- Regression models, with interaction effects between "survey design" and auxiliary socio-demographic variables to estimate the association
  - ✓ appropriate statistical models and tests to evaluate the statistically significance of the interaction effects

- **Significant interaction effects would show different relations among structural and target variable depending on the survey design**

Experimental design: the impact on univariate and multivariate distributions (2)

Results – regression model with independent variable "survey design"

		coefficient		p-value		ANOVA
VARIABLE	Category	Intercept	Survey design	Intercept	Survey design	p-value
Frequency of seeing friends (Everyday)	Sometimes a week	0,500	-0,050	0,000	0,111	
	Once a week	0,204	-0,039	0,000	0,241	
	Sometimes a month	0,175	-0,139	0,000	0,000	
	Sometimes a year	-0,407	-0,317	0,000	0,000	
	Never	-1,114	-0,165	0,000	0,002	
	No friends	-2,167	-0,281	0,000	0,001	
	NR	-2,411	-0,458	0,000	0,000	0,000
Performing physical activity (NO)	Sometimes a week	-0,828	0,189	0,000	0,000	
	Sometimes a month	-1,643	-0,124	0,000	0,006	
	Sometimes a year	-1,527	-0,248	0,000	0,000	
	NR	-2,588	-0,025	0,000	0,702	0,000
Playing sports, with continuity (NO)	Yes	-1,117	0,097	0,000	0,000	
	NR	-3,835	-0,134	0,000	0,117	0,000
Playing sports, occasionally (NO)	Yes	-1,926	0,013	0,000	0,719	
	NR	-3,312	-0,097	0,000	0,168	0,345
Hospitalized, in last 3 months (NO)	Yes	-3,427	-0,009	0,000	0,871	
	NR	-3,923	-0,184	0,000	0,020	0,061

Experimental design: the impact on univariate and multivariate distributions (3)

Results – regression model with interaction effects between "survey design" and auxiliary variables

VARIABLE	Category	Single effect
Performing physical activity (NO)	Sometimes a week	Sex, Age class, Educational level, Income class, Occupation type, Geographical Area, Municipal type
	Sometimes a month	Age class, Educational level, Income class, Occupation type, Geographical Area, Municipal type
	Sometimes a year	Survey design, Age class, Citizenship, Educational level, Income class, Occupation type, Geographical Area, Municipal type
	NR	Age class, Citizenship, Educational level, Municipal type
Frequency of seeing friends (Everyday)	Sometimes a week	Sex, Age class, Educational level, Occupation type, Geographical Area, Municipal type
	Once a week	Survey design, Sex, Age class, Educational level, Income class, Occupation type, Geographical Area, Municipal type
	Sometimes a month	Survey design, Sex, Age class, Citizenship, Educational level, Income class, Occupation type, Geographical Area, Municipal type
	Sometimes a year	Sex, Age class, Citizenship, Educational level, Income class, Occupation type, Geographical Area, Municipal type
	Never	Sex, Age class, Citizenship, Income class, Geographical Area, Municipal type
	No friends	Sex, Age class, Citizenship, Income class, Occupation type, Geographical Area, Municipal type
	NR	Sex, Age class, Educational level, Occupation type, Geographical Area, Municipal type
Playing sports, with continuity (NO)	Yes	Survey design, Sex, Age class, Citizenship, Educational level, Income class, Occupation type, Geographical Area, Municipal type
	NR	Survey design, Sex, Age class, Citizenship, Income class, Occupation type, Municipal type

Experimental design: the impact on univariate and multivariate distributions (4)

Results – regression model with interaction effects between “survey design” and auxiliary variables

VARIABLE	Category	Interaction effects: survey design
Performing physical activity (NO)	Sometimes a week	Age class, Geographical Area
	Sometimes a month	Sex
	Sometimes a year	Citizenship
	NR	Geographical Area
Frequency of seeing friends (Everyday)	Sometimes a week	Sex, Educational level, Geographical area
	Once a week	Age class, Municipal type
	Sometimes a month	Age class
	Sometimes a year	Age class, Geographical area, Municipal type
	Never	Sex, Age class
	No friends	-
	NR	-
Playing sports, with continuity (NO)	Yes	Age class, Educational level, Geographical area
	NR	Age class, Educational level, Income class, Occupation type, Geographical area, Municipal type

- **relevant warning for the researchers which utilize survey data to go in depth in the data analyses of complex phenomena**

Experimental design: Selection and measurement effects – Propensity score

Assessing mode effect in the MM sample using propensity score sub-classification

- Propensity score subclassification (Rosenbaum and Rubin, 1983) used **to disentangle selection and measurement effects**, assuming ignorability of selection effect
- Propensity score (PS) approach in MM surveys can be interpreted as the probability of mode assignment conditional on observed covariates

➢ Propensity score model is defined at household level, as the choice of the survey mode depends on household; $P(M = \text{web}|\mathbf{X})$ is a binomial logistic model at household level

Survey mode ~ geo area + municipal type + household type + household income class + higher education level + occupation type + citizenship

➢ Calculus of weighs, for each group k defined on the deciles of the propensity score distribution, that equate the weighted proportion of web respondent with the proportion of PAPI respondent in the same group

$$w_k = \frac{n_{k,papi}/n_{papi}}{n_{k,web}/n_{web}}$$

Experimental design: Selection and measurement effects – Propensity score

Results Estimates of selection and measurement effects for some target variables – Propensity score method

Variable	Category	Estimates			Effects	
		Web mean	Weighted Web mean	PAPI mean	Selection	Measurement
Reading books (last 12 months)	No	0.451	0.485	0.618	0.034	-0.132
	Yes	0.508	0.432	0.347	-0.075	0.085
	NR	0.041	0.043	0.035	0.002	0.007
Internet access	No	0.163	0.197	0.361	-0.045	-0.120
	Yes	0.804	0.765	0.612	0.046	0.110
	NR	0.033	0.038	0.027	-0.008	0.016
Use of Personal Computer	Yes, in the last 3 months	0.623	0.577	0.423	0.055	0.109
	Yes, from 3 months to 1 year-ago	0.032	0.034	0.025	-0.002	0.010
	Yes, more than 1 year-ago	0.059	0.058	0.049	0.001	0.010
	Never	0.249	0.291	0.471	-0.055	-0.134
	NR	0.037	0.041	0.032	-0.005	0.011
Use of internet	Yes, in the last 3 months	0.698	0.662	0.548	0.042	0.077
	Yes, from 3 months and 1 year ago	0.026	0.026	0.020	(-0.001)	0.008
	Yes, more than 1 year ago	0.049	0.049	0.030	(0.000)	0.018
	Never	0.190	0.221	0.372	-0.040	-0.109
	NR	0.037	0.042	0.031	-0.006	0.013
Life satisfaction	0-2	0.022	0.023	0.019	-0.001	0.006
	3-5	0.126	0.138	0.149	-0.015	0.009
	6-7	0.417	0.417	0.423	-0.005	0.002
	8-10	0.392	0.374	0.374	0.017	-0.024
	NR	0.043	0.047	0.034	-0.006	0.016
Trust in other	In the majority of people	0.244	0.223	0.171	0.022	0.025
	You have to be careful	0.713	0.730	0.796	-0.026	-0.033
	NR	0.043	0.047	0.033	-0.006	0.018

Final considerations on the experimental design

- For the Aspect of Daily Life survey
  - ✓ the introduction of mixed mode has an **important impact** both on the composition of the sample (and its **representativeness**) and on several indicators, whose quality seems to be affected by **measurement effect** which cannot be always easily assessed
  - ✓ MM seems to have an impact on simple and complex analyses as well
  - ✓ the application of all the presented methods is subject to the **validity of the hypotheses** underlying all these methods and that need to be verified by the researcher as far as possible
- The set of the analyses presented and applied in a specific survey context can be considered as a **possible checklist**, a **sequence of steps** usable by researchers of other NSIs to carry out an assessment of mode effect in similar situations
- Generally the underlying effort is hardly compatible with the usual resources and the timing of a statistical process: in general situations an accurate planning of the data collection phase is more advisable, in order to limit as far as possible ex-ante the measurement effect, which is the main drawback of the mixed mode

References

- *Biemer, P. (2001). Nonresponse bias and measurement bias in a comparison of face to face and telephone interviewing. Journal of Official Statistics 17:295-320.*
- *Buelens, B., and Van den Brakel, J. A. (2015). Measurement error calibration in mixed-mode. Sociological methods & Research, 2015, Vol. 4483) 391-426.*
- *de Leeuw, E. 2005. To mix or not to Mix Data Collection Modes in Surveys. Journal of Official Statistics 21(2), 233-55.*
- *Hox, J., de Leeuw, E. D. and Klausch, T. (2015). "Mixed Mode Research: Issues in Design and Analysis." Invited paper presented at the International Conference on Total survey error: improving quality in the era of big data. Baltimore, 19-22 September.*
- *Klausch, T. 2014. Informed Design of Mixed-Mode Surveys - Evaluating mode effects on measurement and selection error. Ipskamp Drukkers, En-schede.*
- *Kolenicov, S. and C. Kennedy. (2014). Evaluating three approaches to statisti-cally adjust for mode effects. Journal of Survey Statistics and Methodology 2 126–158.*
- *Martin, P. and Lynn, P. (2011). The effects of mixed mode survey designs on simple and complex analyses. Centre for Comparative Social Surveys. Working Paper Series. Paper n.04, November 2011.*
- *Park, S., J. K. Kim and S. Park. 2016. An imputation approach for handling mixed-mode surveys. The annals of Applied Statistics. Vol. 10, No. 2. 1063-1085. 2016.*
- *Pearl, J. 2009. Causality: Models, Reasoning and Inference. New York: Cambridge University Press, (2nd edition).*

References (2)

- *Roberts, C., Vandenplas, C., (2017). Estimating Components of Mean Squared Error to Evaluate the Benefits of Mixing Data Collection Modes, Journal of Official Statistics, Vol. 33, No. 2, 2017, pp. 303–334*
- *Rosenbaum, P. R. and Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika, Vol. 70, No. 1. (Apr., 1983), pp. 41-55.*
- *Rubin, D. B. 2006. Matching sampling for causal effects. New York, NY: Cambridge University Press.*
- *Rubin, D.B. 1987. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons, Inc.*
- *Schouten, B., Shlomo, N. and Skinner, C. (2011). Indicators for Monitoring and Improving Representativity of Response. Journal of Official Statistics 27: 231–253.*
- *Suzer-Gurtekin, Z. T., S. G. Heeringa and R. Valliant. 2012. Investigating the bias of alternative statistical inference methods in sequential mixed-mode surveys. Section on Survey Research Methods – JSM 2012.*
- *Vandenplas, C., Loosveldt, G., and Vannieuwenhuyze, J. T. A. (2016). Assessing the use of mode preference as a covariate for the estimation of measurement effects between modes. A sequential mixed mode experiment. Method, data, Analyses. Vol. 10(2), 2016, pp. 119-142.*
- *Vannieuwenhuyze, J. T. A., Loosveldt, G. and Molenberghs, G. (2010). A Method for Evaluating Mode Effects in Mixed-mode Surveys. Public Opinion Quarterly, Volume 74, Issue 5, 1 January 2010, Pages 1027–1045,https://doi.org/10.1093/poq/nfq059.*
- *Weisberg, Herbert F. 2005. The Total Survey Error Approach: A Guide to the New Science of Survey Research. Chicago: University of Chicago.*

Thank you for your attention !