



2018

European Conference
on Quality in Official Statistics

Assessing and adjusting bias deriving from mode effect in mixed mode social surveys

C. De Vitiis, A. Guandalini, F. Inglese and M.D. Terribili
(ISTAT, Italian National Statistical Institute)

29th June 2018

Session 29

Summary

1. The Mixed Mode in social surveys
2. The survey context
3. The analyses of mode effect treatment for the Aspect of Daily Life survey
4. Final considerations and future developments

1. Mixed mode in social surveys

Why mixed modes?

- ❑ To contrast declining response rates and coverage, reducing also the total cost of the surveys
 - ✓ The use of different data collection techniques helps in contacting different types of respondents in the most suitable way for each of them so allowing a gain in population coverage and response rate

Which drawbacks has this choice?

The difficulty of control over **mode effects** and the **confounding** between **selection** and **measurement** effect

- ❑ **Mode effect** refers strictly to **measurement error** differences due to the mode of survey administration
- ❑ A **selection effect** generally occurs, due to the differences in the distributions of the respondents to the alternative modes, even if this is a desirable aspect of MM strategy

How and when dealing with mode effect?

- ❑ Mainly in the planning of the survey (questionnaire and survey design) to limit measurement error as much as possible
- ❑ In the estimation phase to treat mainly the selection effect, while estimating the measurement effect

1. Mixed mode in social surveys

- ❑ The focus of the present work is the illustration of the experimentation plan for the treatment of mode effect in the web/PAPI “Multipurpose Survey on Aspect of daily life” (ADL) survey.
 - ✓ Through the linkage of survey data with administrative data we exploit the auxiliary variables to define mixed mode models.
 - ✓ The final goal is to make an assessment of the introduction of the mixed mode and define an estimation strategy for the future editions of the survey

2. The survey context

The sample survey “Multipurpose survey on households: Aspects of daily”

- ❑ Collects information about recreational and cultural activities in free time, such as sports, reading, cinema, music, the Internet, social relations, issues for the quality of life of people
- ❑ 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: **sequential web/PAPI** with a **control single mode** sample PAPI
- ❑ The selected **sample** of individuals was **linked to an administrative data base** (Archimede Project) through the individual code available from the selection frame to obtain external auxiliary variables

To treat mode effect the use of models is advisable and the availability of auxiliary variables is a crucial issue

- ❑ External sources: from registers or administrative data, socio-demographic and economic variables
- ❑ Survey variables: Mode insensitive socio-demographic variables; Mode preference (not yet introduced at ISTAT); Paradata (information about data collection phase)
- ❑ Auxiliary 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
 - ✓ Income class: 5 quintiles (€ 11.955, 20.892, 30.028, 46.119)
 - ✓ Citizenship: Italian/Foreign household

3. The analyses on ADL survey data

- ❑ The aim of the presented analyses is to
 - Evaluate first the impact on the estimates of the survey of the introduction of mixed mode design with respect to the previous single mode design (control SM sample)
 - To analyze in depth the reasons that determine significant differences in the estimates obtained with the two designs
- ❑ For this purpose, the study is developed on two main levels of analysis:
 - the first level is based on the comparison between the two samples SM and MM
 - tests were performed on the differences in the estimates calculated on the two sample, SM and MM
 - analyses were conducted to study the bias caused by the total nonresponse in the two samples
 - ✓ Total response rates and indicators of response representativeness were evaluated in order to identify differences (especially in terms of magnitude of the bias) that could explain the differences in the estimates of the survey produced with the SM and MM samples
 - the second level investigates the mode effect (selection and measurement) of the samples of respondents web and PAPI in the MM design
 - analysis of the mode effect in the MM sample was carried out using methods that make the samples of respondents web and PAPI comparable, as propensity score (Rosenbaum and Rubin, 1983), to study the selection effect and the measurement effect of some target variables of the survey

Test of differences between estimates

- ❑ To evaluate the differences between the estimates of the main parameters of interest of the survey, obtained with the mixed and the single mode samples, hypothesis tests were carried out (Martin and Lynn, 2011).
- ❑ Test of the differences in proportions through t-test, while the independence between the distributions were evaluated as a whole through the Chi-square test
- ❑ The hypothesis tests concerned the following estimates:
 - Satisfaction for life (Satisfaction)
 - Health conditions (Health)
 - Valuation of the economic situation compared to the previous year (EcoSit)
 - Reading books in the last 12 months (Books)
 - Frequency of seeing friends (Friends)
 - Habit to smoke (Smoke)
- The difference for **Satisfaction, Books and Friends** resulted significant

3. The analyses on ADL survey data

Test of differences between estimates

Table 1. Estimates for “Reading books” in SM and MM samples

	No	Yes	NR
MM	54.8	41.6	3.6
SM	57.5	39.9	2.6
t-test	<.0001	0.0004	<.0001
χ^2	<.0001		

Table 2. Estimates for “Seeing friends” in SM and MM samples

	Everyday	Sometimes a week	Once a week	Sometimes a month	Sometimes a year	Never	No friends	NR
MM	14.8	26.0	20.4	19.4	10.8	5.2	2.0	1.5
SM	17.3	27.2	20.8	18.6	8.3	5.0	1.6	1.1
t-test	<.0001	0.0018	0.3271	0.0303	<.0001	0.5372	0.0029	0.0027
χ^2	<.0001							

Test of differences between response rates

- ❑ To assess whether the response rate distributions are independent from the individual structural variables, the hypothesis of independence between the response and the variables was tested, in the two samples
 - The structural variables influence the response in both samples

Table 3. Independence test between response and auxiliary variables in SM and MM samples

Sample	Auxiliary variables	DF	χ^2	p-value
SM sample (PAPI)	Geographical area	4	131.1118	<.0001
	Municipal typology	5	293,713	<.0001
	Household typology by number of components and age	6	295.3983	<.0001
	Income class	4	270.174	<.0001
	Nationality	2	567.6386	<.0001
MM sample (web/PAPI)	Geographical area	4	91.9192	<.0001
	Municipal typology	5	268,3902	<.0001
	Household typology by number of components and age	6	142.5824	<.0001
	Income class	4	127.8876	<.0001
	Nationality	2	168.9341	<.0001

3. The analyses on ADL survey data

Test of differences between response rates

- ❑ Analysis is also carried out on the PAPI component of the two samples
 - The result shows that the distribution by geographical area and income class of the respondents to PAPI is not independent from whether they were selected for the PAPI or web/PAPI samples

Table 4. Independence test between auxiliary variables in PAPI respondent of SM and MM samples

Variable	DF	χ^2	p-value
Geographical area	4	186.5848	<.0001
Municipal typology	5	17,3572	0,0039
Household typology by number of components and age	6	6.2375	0.3971
Income class	4	144.5565	<.0001
Nationality	2	6.2907	0.0431

Analysis of total nonresponse bias

- R-indicators 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

$$R(\rho_X) = 1 - 2S(\rho_X)$$

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

- MM sample of respondents deviates less from the representative response with respect to the SM sample

Table 5. R-indicators in SM and MM samples

R_Indicator	SM sample	MM sample
$R(\rho_X)$	0.812	0.852
$\hat{R}(\hat{\rho}_X)$	0.814	0.854

3. The analyses on ADL survey data

Analysis of total nonresponse bias

- ❑ R-indicator calculated on the basis of the estimated response propensity through response models defined for each geographical area (North, Center, South and Islands)
 - For the North the values of the R-indicators are similar for the two samples, for the other geographical areas they are very different

Table 6. R-indicators in SM and MM samples by geographical area

R_Indicator	SM sample	MM sample
	$R(\rho_X)$	$R(\rho_X)$
North	0.847	0.840
Center	0.752	0.842
South and Islands	0.840	0.907

The analysis of mode effect in the MM sample based on propensity score

- ❑ To get an assessment of selection and measurement effect in the MM sample, a Propensity score stratification adjustment methods was used (Rosenbaum and Rubin, 1983)
- ❑ Propensity score (PS) approach is adopted in observational studies by achieving a balance of covariates between comparison groups, while in MM surveys can be interpreted as the probability of mode assignment conditional on observed covariates.
- ❑ With adjustments based on PS, the confounding effects of the selection mechanism are mitigated
- Propensity score model is defined at household level, as the choice of the survey mode depends on household; $P(M = \text{web}|\mathbf{X})$ for the case of ignorability, 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

The application of the Propensity Score Subclassification (1)

- ❑ Propensity Score Subclassification steps:
 - Estimate of the propensity score model (choice model) parameters
 - Definition of strata of respondents (web and PAPI) based on quintiles and deciles of the propensity score distribution
 - Validation of the balancing assumption in each stratum (independence of all X from the mode)
 - The calculus of weighs, for each balanced group k , that equate the weighted proportion of web respondent households with the proportion of PAPI respondent households in the same stratum

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

The application of the Propensity Score Subclassification (2)

- A global evaluation of mode effects using the weights in the balanced groups (Vandenplas et al, 2016)
 - **Selection effect**, by the difference between weighted and not weighted estimates on web respondents

$$S_{web}(y) = \frac{\sum_{i=1}^{n_{web}} y_{i,web}}{n_{web}} - \frac{\sum_{i=1}^{n_{web}} w_{k,i} y_{i,web}}{n_{web}}$$

- **Measurement effect**, measured by the differences between weighted web and not weighted PAPI estimates

$$M_{web}(y) = \frac{\sum_{i=1}^{n_{web}} w_{k,i} y_{i,web}}{n_{web}} - \frac{\sum_{i=1}^{n_{papi}} y_{i,papi}}{n_{papi}}$$

- An evaluation of measurement error within the balanced groups is obtained through the test of the independence between mode and Y

$$d_{k,web} = (F^*_{k,web} - F_{k,web}) / n_{k,web} , \quad d_{k,papi} = (F^*_{k,papi} - F_{k,papi}) / n_{k,papi}$$

3. The analyses on ADL survey data

Selection and measurement effect estimated through PS

Table 7. Evaluation of selection and measurement effect in MM sample

Target variable	Web mean	Weighted Web mean	PAPI mean	Selection effect	Measurement effect
Reading books (last 12 months)					
No	0.479	0.397	0.602	-0.082	-0.123
Yes	0.417	0.522	0.320	0.205	0.202
NR	0.045	0.037	0.031	-0.008	0.014
Frequency of seeing friends					
Everyday	0.102	0.082	0.190	-0.020	-0.089
Sometimes a week	0.235	0.243	0.255	0.008	-0.020
Once a week	0.190	0.202	0.179	0.012	0.011
Sometimes a month	0.189	0.209	0.166	0.020	0.023
Sometimes a year	0.131	0.139	0.081	0.009	0.058
Never	0.052	0.042	0.053	-0.009	-0.002
No friends	0.017	0.017	0.018	-0.001	-0.001
NR	0.025	0.021	0.010	-0.004	0.014

4. Final considerations and future developments

- ❑ From a strictly methodological point of view, the analyses need to taken into account the sampling variability, by means of a simulation to assess significance of the evaluated mode effects
- ❑ For the Aspect of Daily Life survey
 - ✓ SM and MM, produce generally different estimates
 - ✓ The MM sample is more “representative” of the population with respect to the main sociodemographic variables, but produces both selection and measurement effect for some of the considered target variables
 - ✓ The strategy for the future editions of the survey, which is carried out yearly with MM design, should be based both on prevention of measurement effect and use of calibration method, such as those proposed by Buelens et al. (2015), which seems to allow to stabilise the mode effect in repeated surveys

Selected references

- ✓ Buelens, B., and Van den Brakel, J. A. (2015). Measurement error calibration in mixed-mode. *Sociological methods & Research*, 2015, Vol. 44(83) 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.
- ✓ 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.
- ✓ 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.
- ✓ Schouten, B., Shlomo, N. and Skinner, C. (2011). Indicators for Monitoring and Improving Representativity of Response. *Journal of Official Statistics* 27: 231–253.
- ✓ 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>

Assessing and adjusting bias deriving from mode effect in mixed mode social surveys

Thanks

Claudia De Vitiis, ISTAT, devitiis@istat.it