Dealing with under-coverage bias via Dual/Multiple Frame designs: a simulation study for telephone surveys

Errori di stima in situazioni di under-coverage e campionamento da più liste: uno studio di simulazione per le indagini telefoniche

Abstract Multiple-frame surveys are commonly used to deal with under-coverage bias. The use of more than one frame introduces the possibility that frames overlap leading to increased inclusion probabilities of units that appear in multiple lists. Following the guide example of a dual frame set-up (DF) in telephone surveys, this contribution presents an extensive simulation study where different types of screenings to deal with the overlap issue are compared with the proper DF approach. We empirically show that the efforts of screening do not guarantee estimators more efficient than the DF estimators that do not need screening. Moreover simulation results show that screening at sample level does not correct for the increased inclusion probability of units in both frames produced by the DF set-up nor improve efficiency. The different estimation options will be compared with regards to survey costs, amount of information required and statistical properties of the final estimates.
Abstract *Il campionamento da più liste (multiple frame) è un metodo utilizzato quando la lista completa di tutte le unità della popolazione da cui campionare non esiste o è difficile da ricostruire. Il lavoro presentato contiene un ampio studio di simulazione in cui, a partire dallesempio delle indagini telefoniche e al variare della copertura delle due liste, vengono simulati diversi metodi di screening della popolazione o del campione. Nel caso di screening a livello della popolazione, ovvero di ricostruzione di una lista unica, si mostrerà come lo sforzo in termini di costi e tempi non sia corrisposto da un effettivo guadagno di efficienza rispetto allapproccio multiple frame. Nel caso invece di screening a livello di campione si mostrerà che le stime multiple frame classiche sono necessarie rispetto a stimatori HT e sono comunque preferibili ai metodi di post-screening.*

Key words: Complex survey, Multiple frame survey, Post-screening designs, Coverage bias

1 Introduction and Motivational Example

Modern telephone surveys based on the traditional landline list are tipically affected by under-coverage bias due to the decrease in landline phone use. A natural and popular device for improving population coverage in telephone surveys consists of supplementing the traditional landline frame with a list of cellular numbers. A telephone survey using both landline and cell phone numbers configures a Dual Frame survey.

In addition to help improving population coverage, Dual/Multiple Frame surveys may be convenient for allowing to blend different sampling designs and different data collection modes and can be cost effective especially in case of different cost per unit across frames, proving to be more flexible compared to the traditional designs.

Despite these advantages the involved frames often overlap leading to challenges in estimation given by multiple opportunities of sampling the same individual in the final sample and by the increased inclusion probability of units appearing in more than one frame. In order to avoid such an additional difficulty at the estimation stage while still acknowledging the under-coverage issue at the selection stage, screenings can be implemented for removing the overlaps between frames and hence eliminate any need of adjusting at the estimation stage for the issues mentioned above. If units appearing in both frames could be identified at population level *prior* to sample selection, frames' overlap could be eliminated by screening out such duplications. In this case a unique list is recovered and conventional single-frame estimators can be used for either a simple or a stratified design. However, *pre*-screening at population level is a costly and error prone activity and it usually requires extra-information available to actually be implemented.

In the practice of telephone surveys using both landline and cell phones, it appears to be not uncommon the performance of some sort of less demanding screening at sample level, particularly *interim*-screening during interviewing or else *post* screening of duplicate sampled units, if any, in the final data. Both *interim* and *post* screenings for being limited to sampled units appear less resource consuming than a complete *pre*-screening at population level. It is speculated that the difference with *pre*-screening might be assumed negligible and the Dual/Multiple Frame structure actually underlying screened sample data might be eluded, hence leading to the employment of standard single-frame Horvitz-Thompson (HT)-type estimation as well.

The main purpose of this paper is to produce empirical evidence of the actual effects of such an assumption over the statistical properties of the final estimate, namely bias and stability. By simulating a telephone survey involving both landline and cellular numbers, strategies based on screenings at sample level (*interim* and *post*) will be compared with a proper Dual Frame strategy, with no screening at the selection stage and adjusting weighting at the estimation stage. The case of complete *pre*-screening at population level will be simulated as well, with the purpose of investigating the actual gains in the accurancy of the final estimate deriving from the extra-costs incurred in order to purposively avoid the use of a proper Dual Frame estimation.

2 Dual frame stategy versus screenings in telephone surveys using both landline and cell numbers

According to the guide-example of a telephone survey using both landline and cell phones, in this section we formalize the theoretical framework based upon the traditional representation of a Dual Frame survey in terms of overlapping sets.

2.1 Dual Frame selection and estimation

In this paragraph the DF set-up is presented based on the two frames L for the landline phones and C for that of cellular phones. In figure 1 two independent samples are selected under a given fixed-size sampling design: a sample s_L of size n_L of landline numbers and a sample s_C of size n_C of cell numbers providing an overall sample of fixed size $n = n_L + n_C$. Dual/Multiple Frame estimation is based on the post-classification of the frame-samples into domain samples, i.e. landline owners only $s_{l(L)}$, individuals interviewed via landline phone who possess also a mobile number $s_{lc(L)}$, mobile owners only $s_{c(C)}$, and units interviewed via cell phone who possess a landline too $s_{lc(C)}$.

Among the several Dual/Multiple Frame estimators proposed in the literature so far (see [4] and [3] for a review), we focus on three unbiased estimators: the (simplest) multiplicity estimator $\hat{Y}_M$, the (more complex) Kalton-Anderson estimator $\hat{Y}_{KA}$ and the (theoretically) optimal estimator $\hat{Y}_{opt}$.
The multiplicity estimator in a Dual Frame telephone survey as described in figure

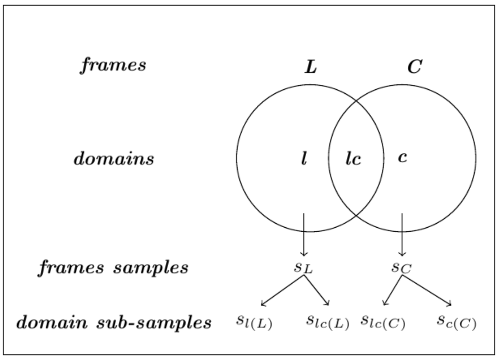

Fig. 1 DF set-up in telephone survey

1, is given by:

$$\hat{\hat{Y}}_M = N^{-1}\left(\sum_{i\in s_C} y_i m_i^{-1}\pi_{i(C)}^{-1} + \sum_{i\in s_L} y_i m_i^{-1}\pi_{i(L)}^{-1}\right) \tag{1}$$

where the the inclusion π_i are corrected by the unit multiplicity m_i, i.e. the number of frames unit i belongs to. Note that in the DF design, m_i is either equal to 2, when the unit belongs to both frames and to 1 otherwise.

The Kalton-Anderson estimator is slightly more complex and is given by:

$$\hat{\hat{Y}}_{KA} = N^{-1}\left(\sum_{i\in s_C} y_i \pi^*_{i(C)} + \sum_{i\in s_L} y_i \pi^*_{i(L)}\right) \tag{2}$$

where the (adjusted) *-weights are given by

$$\pi^*_{i(C)} = \begin{cases} \pi_{i(C)}^{-1} & \text{if } i \in s_c \\ (\pi_{i(C)} + \pi_{i(L)})^{-1} & \text{if } i \in s_{lc(C)} \end{cases}$$

and

$$\pi^*_{i(L)} = \begin{cases} \pi_{i(L)}^{-1} & \text{if } i \in s_l \\ (\pi_{i(C)} + \pi_{i(L)})^{-1} & \text{if } i \in s_{lc(L)} \end{cases}$$

Last, Hartley's optimal estimator $\hat{\hat{Y}}_{opt}$ is derived under an optimal approach, i.e. by minimizing the estimator variance. In the Dual Frame set-up considering the telephone survey it is defined as:

$$\begin{aligned}\hat{\hat{Y}}_{opt} = N^{-1} &\left[\sum_{i\in s_{l(L)}} y_i\pi_{i(L)} + \alpha_{opt}\sum_{i\in s_{lc(L)}} y_i\pi_{i(L)} + (1-\alpha_{opt})\sum_{i\in s_{lc(C)}} y_i\pi_{i(C)} + \sum_{i\in s_{c(C)}} y_i\pi_{i(C)}\right] \\ &= N^{-1}\left[\hat{Y}_{l(L)} + \alpha_{opt}\hat{Y}_{lc(L)} + (1-\alpha_{opt})\hat{Y}_{lc(C)} + \hat{Y}_{c(C)}\right] \qquad (3)\end{aligned}$$

where the coefficient $\alpha_{opt} \in [0,1]$ is given by:

$$\alpha_{opt} = \frac{V\left(\hat{Y}_{lc(C)}\right) + Cov\left(\hat{Y}_{c(C)}, \hat{Y}_{lc(C)}\right) - Cov\left(\hat{Y}_{l(L)}, \hat{Y}_{lc(L)}\right)}{V\left(\hat{Y}_{lc(L)}\right) + V\left(\hat{Y}_{lc(C)}\right)}$$

2.2 *Complete* pre-screening *at population level*

If all population units belonging into the frames overlap can be correctly identified *prior* to sample selection the DF structure can be avoided, leading to the recovering of a unique frame that ensures complete coverage (figure 2). A conventional single-frame sample selection or a conventional stratified sampling can hence be performed and traditional Horvitz-Thompson estimators apply.

2.3 **Interim**-*screening and* **post**-*screening at sample level*

Interim and *post*-screening are implemented at sample level so that the sample selection would actually be performed in the DF setup. In the case of *interim*-screening duplicated units are discarded and replaced during the interview while with the *post*-screening, possibly duplicate sampled units are removed after sample selection. Notice that *interim*-screening leaves the size of the overall final sample unaltered to fixed n, as duplicated units in the sample are replaced by new units, while a *post*-screening may return a sample of reduced size $\tilde{n} \leq n$.

As both screenings provide sample data immune from both under-coverage and duplications, it is speculated that this might induce the temptation to assume as negligible the difference with sample data provided by complete *pre-screening* and thus consider HT-type estimation still appropiate or with negligible bias.

In the case of *post*-screening Bankier's estimator [1] will be considered as well being an unbiased DF estimator that is based on a post-screening activity.

3 Simulation Study

The simulation study is performed with the aim of producing empirical evidence with respect to two main objectives:

- Assuming it is possible to recover a unique frame through *pre*-screening, conventional single-frame HT-type estimators are computed. It is expected that the

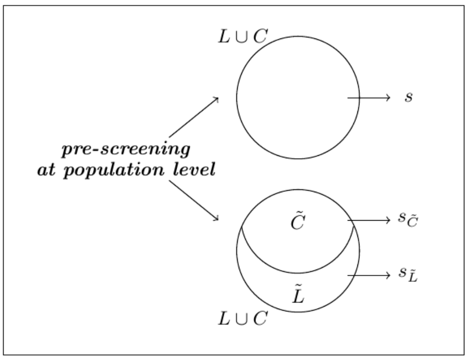

Fig. 2 Recovering of a unique frame (single frame set-up) via *pre*-screening at population level in telephone surveys

extra costs of *pre*-screening with respect to proper Dual Frame selection and estimation would be balanced by significant gains in the accuracy and stability of the final estimates.

- *Interim* and *post*-screening at sample level, for being limited to sampled units might affect to an uncontrollable extent conventional HT estimates computed under the assumption of a negligible difference with the complete *pre*-screening at population level. Empirical evidence will be produced of both the possible bias and efficiency losses of HT-type estimates as compared with proper Dual Frame estimation

3.1 Simulated scenarios

A number of different Dual Frame (DF) scenarios have been produced, letting the overlap pattern vary between a slight overlap and a total overlap when L is nested into C. Table 1 shows details of a selection of 3 scenarios, included the special case of L nested into C (scenario 3).

Table 1 Simulated scenarios

	Frame	sizes	Coverage		Overlap		Strata	sizes
Scenario	N_L	N_C	$\frac{N_L}{N}$	$\frac{N_C}{N}$	$\frac{N_{cl}}{N}$	$\frac{N_{cl}}{N_L}$	$N_{\tilde{C}}$	$N_{\tilde{L}}$
1	2717	2753	54%	55%	9%	17%	2283	2717
2	2648	4041	53%	81%	34%	64%	2352	2648
3	2781	5000	56%	100%	56%	100%	2219	2781

Both a quantitative and a dicotomous variable (y) were considered. Samples have been selected under different sample sizes and different sampling designs including constant probability sampling designs (SRS) and probability proportional to size sampling designs (PPS). Moreover, a mixed sampling design with SRS in one frame and PPS in the other has been implemented which applies to a proper Dual Frame selection only.

3.2 Results

Assuming a *pre*-screening is performed, one of the simulation's focuses is to investigate whether and to what extent the extra costs employed into the screening were balanced by a gain in accurancy as compared to the DF estimates. Both cases of a simple and stratified single-frame selection are simulated. The performance of conventional HT estimators is hence compared with a proper DF strategy as discussed in 2.1, with no screening of possibly duplicated units. Simulation results suggest that a DF strategy with no screening may be convenient even if a conventional single-frame strategy would exactly apply after screening for producing estimate with very similar efficiency in all scenarios and for allowing to blend differet sampling designs. The third and fourth column of table 2 show results for the three scenarios described in table 1, in the case of simple random sampling and proportionate sampling factions (5%), where the population parameter to estimate is $\bar{Y} = 0.6$. Comparison has been facilitated by calculating *(i)* the percentage relative bias $\%RB = E_{MC}[\hat{\bar{Y}} - \bar{Y}]/\bar{Y} \times 100$ *(ii)* percentage relative root mean squared error $\%RRMSE = \sqrt{E_{MC}[\hat{\bar{Y}} - \bar{Y}]^2}/\bar{Y} \times 100$ *(iii)* when appropriate, a MC efficiency ratio showing the loss/gain compared to the expected most efficient estimator $eff = (1 - \%RRMSE/\%RRMSE_{strHT}) \times 100$.

Table 2 *Pre* and *interim*-screening vs Dual Frame

		Pre-screening		Interim-screening	
Scenario	Est.	%RRMSE	eff vs $\hat{\bar{Y}}_{HTstr}$	%RBias	%RMSE
1	$\hat{\bar{Y}}_{HT}$	4.7	-33%	18.01	18.52
	$\hat{\bar{Y}}_{HTstr}$	3.55	-	12.67	13.06
	$\hat{\bar{Y}}_{KA}$	3.77	-6%	~ 0	3.52
	$\hat{\bar{Y}}_{M}$	3.77	-6%	~ 0	3.42
	$\hat{\bar{Y}}_{opt}$	3.63	-2%	~ 0	3.41
2	$\hat{\bar{Y}}_{HT}$	4.83	-30%	20.37	20.58
	$\hat{\bar{Y}}_{HTstr}$	3.7	-	49.22	49.4
	$\hat{\bar{Y}}_{KA}$	3.81	-2%	~ 0	3.62
	$\hat{\bar{Y}}_{M}$	3.81	-2%	~ 0	3.42
	$\hat{\bar{Y}}_{opt}$	3.49	+6%	~ 0	3.39
3	$\hat{\bar{Y}}_{HT}$	4.93	-27%	22.75	22.9
	$\hat{\bar{Y}}_{HTstr}$	3.89	-	79.07	79.22
	$\hat{\bar{Y}}_{KA}$	3.54	8%	~ 0	3.34
	$\hat{\bar{Y}}_{M}$	3.54	8%	~ 0	3.2
	$\hat{\bar{Y}}_{opt}$	3.48	11%	~ 0	3.23

Regarding screenings at sample level (section 2.3) simulation results show significant biases and efficiency losses in the HT estimators when the DF structure underlying sample selection is eluded at the estimation stage. Moreover the Bankier's estimator's strategy, which involves a possibly costly *post*-screening, appears to not improve the accuracy of the final estimate since it results sligtly less efficient or almost equally efficient than any of the considered DF estimators with no screenings. Table 3 and columns 5-6 of table 2 show a selection of results in the case of disproportionate sampling ($f_L = 10\%$ and $f_C = 5\%$) and simple random sampling, where the population parameter to estimate is $\bar{Y} = 0.6$.

In conclusion, simulation results clearly show that resorting to HT estimation, either simple or stratified, after screening at sample level would lead to severe biases

Table 3 *Post*-screening vs Dual Frame

Scenario	Est.	%RBias	%RRMSE	eff vs $\hat{\hat{Y}}_B$
1	$\hat{\hat{Y}}_{HT}$	17.64	17.87	-
	$\hat{\hat{Y}}_{HTstr}$	12.55	13.14	-
	$\hat{\hat{Y}}_{B}$	~ 0	3.46	-
	$\hat{\hat{Y}}_{M}$	~ 0	3.55	-3%
	$\hat{\hat{Y}}_{KA}$	~ 0	3.44	+1%
	$\hat{\hat{Y}}_{opt}$	~ 0	3.42	+1%
2	$\hat{\hat{Y}}_{HT}$	19.54	21.23	-
	$\hat{\hat{Y}}_{HTstr}$	32.50	47.02	-
	$\hat{\hat{Y}}_{KA}$	~ 0	3.33	-
	$\hat{\hat{Y}}_{M}$	~ 0	3.44	-3%
	$\hat{\hat{Y}}_{opt}$	~ 0	3.41	+1%
		~ 0	3.32	+1%
3	$\hat{\hat{Y}}_{HT}$	23.09	23.25	-
	$\hat{\hat{Y}}_{HTstr}$	72.07	72.19	-
	$\hat{\hat{Y}}_{KA}$	~ 0	3.19	-
	$\hat{\hat{Y}}_{M}$	~ 0	3.32	-4%
	$\hat{\hat{Y}}_{opt}$	~ 0	3.15	+1%
		~ 0	3.12	+1%

and efficiency losses in all scenarios simulated, with increasing gravity as the overlap between frames increases. They also suggest that *pre*-screening at population level as well as the use of Bankier's estimator does not produce gains in accuracy. Hence the use of DF designs may be convenient for its flexibiliy, its reduced costs and for guranteeing a very similar efficiency to screening desins without the need of screening.

References

1. Bankier, M. (1986). *Estimators based on several stratified samples with applications to multiple frame surveys*. Journal of the American Statistical Association, 81:10741079.
2. Brick, J. M., Cervantes, I. F., Lee, S., & Norman, G. (2011). *Nonsampling errors in dual frame telephone surveys*. Survey Methodology, 37(1), 1-12.
3. Lohr, S. (2009). *Handbook of Statistics Sample Surveys: Design, Methods and Applications*, Volume 29, Part A: 71-88.
4. Mecatti, F., Singh, A. C. (2014). *Estimation in multiple frame surveys: a simplified and unified review using the multipliciy approach*, Journal de la Socit Franaise de Statistique. Vol. 155 No. 4, 5169.
5. Singh, A. C., Mecatti, F (2011). *Generalized multiplicity-adjusted Horvitz-Thompson estimation as a unified approach to multiple frame surveys*, Journal of Official Statistics Vol.27 No.4, 633650.