14th Workshop on Labour Force Survey Methodology
Budapest, 16–17 May 2019

Integrating LFS and HBS from the interviewers' network

Manuela Morricone, Alessandro Ortenzi, Rita Ranaldi, Emanuela Vergura
morricon@istat.it, alortenz@istat.it, ranaldi@istat.it, emvergur@istat.it,

ISTAT – Italian National Statistical Institute

1. Introduction

Italian National Statistical Institute (Istat) does not own an interviewers' network, neither for CAPI mode nor for CATI mode. So it usually entrusts data collection phase of each households survey to different external companies through calls for tenders[1].

For small surveys that present a high dispersion of interviews on the Italian territory, sometimes also characterized by a complicated orographic situation, the coverage of the entire territory is rather onerous even for external companies specially recruited for the management of the fieldwork.

This paper presents the Istat strategy adopted in the last two years in order to solve the problem with reference to two of the main household surveys, Labour Force Survey (LFS) and Household Budget Survey (HBS). This innovation introduced and presented in the next paragraphs is part of the integration project of social surveys to which Istat is currently moving with the new IESS Framework regulation just adopted.

2. Italian LFS and HBS background

Italian LFS is a mixed mode CAPI-CATI[2] survey involving every year about 280,000 households spread in about 1,100 municipalities. Almost half of the households sample is interviewed using CAPI mode. The reference period is the week, so the fieldwork data collection takes place all the days of the year. In particular, the interview is carried out usually within the first two weeks, at most within four weeks, after the reference week and an interview is considered complete only when all the household members have been interviewed.

Italian HBS[3] is a multimode CAPI-CADI[4] survey involving every year about 20,000 households residing in about 500 municipalities. The reference period is the month and, as for LFS, the fieldwork data collection occurs all the days of the year. In particular, for HBS a first part of the interview takes place within two weeks before the reference period of the diary (first phase), the self-compilation of the diary within two weeks of the reference period (second phase) and a second part of the interview takes place within two weeks following the reference month (third phase); an interview is considered complete when all parts of the interview and the diary have been completed.

[1] Only in a number limited of surveys data collection phase is entrusted to an interviewers' network specifically recruited by sample municipalities.

[2] CAPI mode is usually used for the first wave, whereas CATI mode is usually used for later waves. Households without a telephone and non-Italian households are interviewed always by CAPI mode.

[3] It includes also a module dedicated to Trips and Holidays survey.

[4] Data collection consists in an interview (in two different moments) carried out by an interviewer using CAPI mode and in a self-administration by a family member of a paper diary that is computer assisted data entered by the interviewer within the end of the fieldwork period.

The smallness of HBS sample size distributed through the year at territory level has determined over time a problem of coverage. Sometimes for private companies data collection is not convenient in municipalities not easily accessible and with a low number of assigned interviews, with consequent low response rates. The percentage of complete interviews on the total expected theoretical interviews in 2016 was sometimes below 50%, a figure considered too low for the dissemination of quality estimates. So, considering the peculiarities of both the surveys, with particular reference to the continuity of data collection, Istat has though to take advantage of the pre-existing and consolidated LFS network and at the end of 2016 it entrusted to a private company, for a period of 42 months[5], the management of the interviewers' network and CAPI data collection fieldwork for both LFS and HBS.
In order to achieve this goal, Istat preventively revised: 1) the HBS sampling design in order to maximise the overlapping of the sample units at the first stage (municipalities) with those of the LFS for about 80%, with the aim to optimize the territorial distribution of the sample units at the second stage (households), and 2) the interviewers training program which, although different for the two surveys, has included transversal training cycles.

2.1. LFS and HBS sampling design and positive coordination of the PSUs

A two stages sampling design is used for both LFS and HBS: municipalities are the primary sampling units (PSUs) and households are the final sampling units (FSUs). In each NUTS 3 domain for LFS and in each NUTS 2 domain for HBS, PSUs are stratified according to the demographic size (and in HBS for municipal type too – metropolitan cities, metropolitan area, municipalities with more than 50,000 inhabitants and other municipalities up to 50,000 inhabitants). Large municipalities, with population over a given threshold (also called self-representative municipalities - SR), are always included in the sample; smaller municipalities (not self-representative - NSR) are grouped in strata, then one municipality for LFS and three municipalities for HBS in each stratum are selected with probability proportional to their population.
At the second stage, households are randomly selected from the population registers in all the municipalities drawn at the first stage. All persons living in private households selected at the second stage are included in the sample. In addition for LFS, FSUs are rotated according to a 2-(2)-2 rotation scheme.
For both the surveys the sample is evenly spread over the year, in particular for LFS over the 52 weeks of the year and for HBS over the 12 months of the year. For HBS the monthly FSUs are further divided into two subsets, distinguishing the FSUs that compile the diary on the expenses in the first half of the month from those that fill it in the second part of the month.
So for both the surveys, NSR PSUs are not always involved in the survey, but just one week in each month for LFS and one month in each quarter for HBS. SR PSUs are involved every month in HBS and, instead, in LFS one part every week of the year and another part only one week a month.
The number of sample municipalities for LFS is approximately 1,115 and, in order to reduce the statistical burden in municipalities with a small number of residents, every year 10-13 percent of PSUs are rotated in sampling strata composed by municipalities whose demographic size is small. The list of the PSUs is available within May and it is effective in the fieldwork since July of each year (Figure 1).
The PSUs sample size for HBS is smaller than LFS with 485 municipalities and the list is defined starting from June (after the availability of the LFS list) and it is effective in the fieldwork since the second half of December of each year. During the definition of the HBS list, a specific algorithm defined by Istat methodological directorate guarantees a good positive coordination of the PSUs with those of LFS.

[5] 36 effective months of fieldwork.

Considering the different timing in the definition of the PSUs list for both the surveys, the overlap of the sample municipalities can be appreciated in particular with reference to the first 6 months of each year.

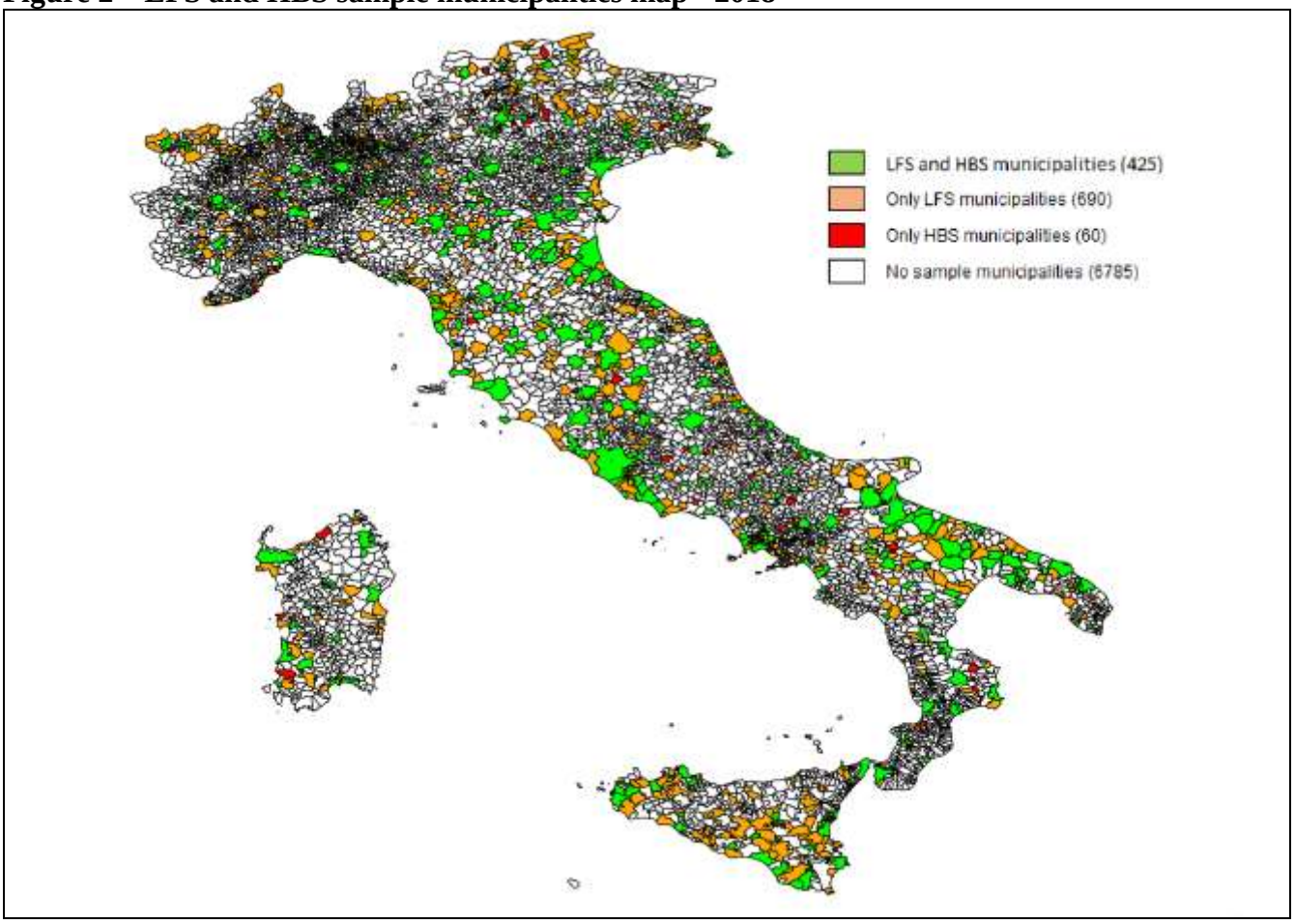
Figure 1 – LFS and HBS PSUs selection

	Year T												Year T+1											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
LFS																								
HBS																								

PSUs selection
Effectivity on the fieldwork
MAXIMUM OVERLAP OF MUNICIPALITIES

At the beginning of 2017 (the first year of implementation of the positive coordination of the PSUs) 82% of the municipalities in the HBS list was present also in the LFS PSUs list; in the second part of the year this percentage decreased to 79% by virtue of the update made in the LFS list.
In the geographic map shown in figure 2 we can have an overview of the Italian municipalities' map for both LFS and HBS, at the beginning of 2018, 425 municipalities (in green colour) were common both for HBS (87% of its total PSUs) and for LFS (38% of its total PSUs); only 60 municipalities (in red colour) were exclusively involved in HBS.

Figure 2 – LFS and HBS sample municipalities map - 2018



Having a high overlap of the municipalities of HBS with those of LFS, the company awarded the tender notice would certainly have been more inclined to set up a single network, exploiting the existing LFS interviewers' network, thus minimizing the interviewers' movements on the territory.

2.2. Interviewers training program

The initial interviewers training was scheduled at different times for the two surveys, before for LFS and after for HBS. Survey specific modules and transversal modules have been planned. Survey specific modules concern characteristics, contents and aims of the survey, definitions, questionnaire structure and meaning of particular questions; they include also more technical aspects connected to software functioning of the electronic questionnaire. Transversal modules include aspects related to effective communication and techniques to reduce refusals, the tasks of the interviewer and, from a thematic point of view, the codification of the occupation.

So the training program would have lasted 3.5 days for all the interviewers involved in LFS; the subsequent training courses for HBS would have lasted 3 days for interviewers who did not attend the previous LFS training cycle and 2.5 days for interviewers already trained and working in the LFS fieldwork.

Training and briefings of the interviewers and supervisors are organized at the expenses of the private company. The costs of the location, as well as the travel and accommodation costs of all the staff involved in the training are borne by the company. Therefore, with the use of an interviewers' integrated network, the company, that would have been awarded the tender, would have obtained considerable savings both in terms of recruitment costs and expenses for training briefings.

3. The integration of the interviewers' network and a comparison with the past

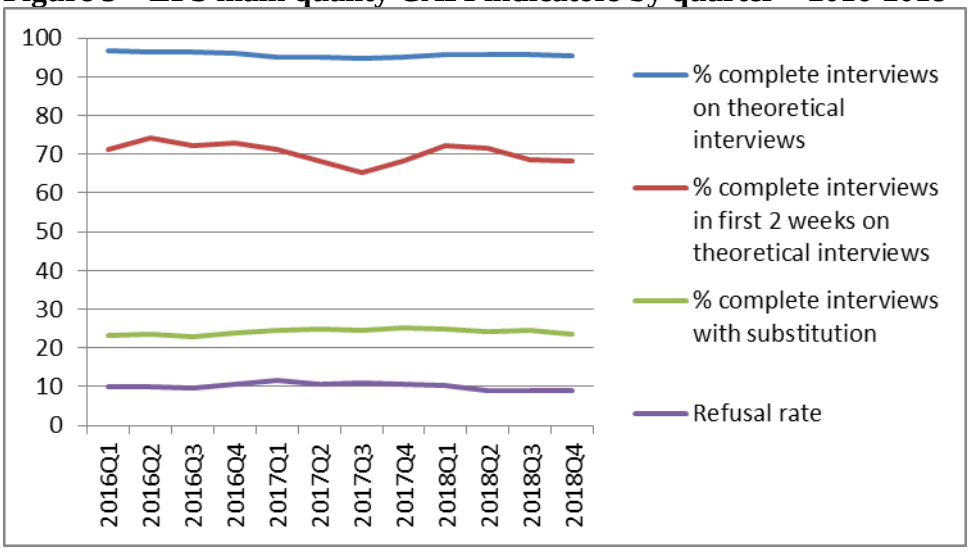
In the background previously outlined, for the private company that won the call for tender, it has been convenient to use for both surveys a single national interviewers' network, that one already existing for the LFS, making only a few additions. In particular, the pre-existing LFS interviewers' network was composed of about 310 interviewers, whereas the interviewers' network integrated for LFS and HBS is composed of about 350 interviewers spread across the entire territory: almost all of them (91.3%) are carrying out the interviews for both LFS and HBS, and only a residual part exclusively for a single survey (5.9% are involved only in LFS fieldwork and 2.8% only in HBS fieldwork). In particular, with reference to the interviewer's workload, the strategy adopted by the private company has been to assign to each HBS interview about four interviews for LFS.

In the last two years, the fieldwork has been the object of a specific quality control. The concern was in particular for LFS because the previous period was characterized by elevated performance and the new interviewers' integrated network could have destabilized the situation.

We have observed across time specific performance indicators at different levels of detail (NUTS2 and NUTS3 level, municipality level, interviewer level, etc.), in order to monitor the quality of fieldwork and the impact of introduced changes for HBS and in particular for LFS. In the graphs in Figures 3 and 4 we report some specific performance indicators respectively for LFS and HBS with reference to the 2016-2018 three-year period during which it is possible to distinguish two periods characterized by differences in data collection: the first one, up to the fourth quarter 2016, in which two different private companies managed the interviewers for the realization respectively of LFS and HBS; the second one, starting from the first quarter 2017, in which there is a single private company for conducting both surveys, with a single interviewers' network. The analysis aims to highlight the trends of the indicators in the two surveys, to identify advantages and critical issues deriving from the changes that have occurred.

The LFS trend of the percentage of complete interviews and the percentage of complete interviews realized in the first two weeks on the total expected theoretical interviews, the percentage of complete interviews with substitution and the refusal rate are shown in the graph in Figure 3.
The percentage of complete interviews on the total expected theoretical interviews has remained substantially constant throughout the period at around 95.7% with fluctuations that have not exceeded one percentage point. The percentage of complete interviews in the first two weeks on the total expected theoretical interviews, which represents an indicator of timeliness in terms of distance of interviews from the reference week, on the other hand, always above 71% in 2016, during 2017 also fell by 6 percentage points, then it rise again in 2018. The remaining two indicators, the percentage of complete interviews with substitution and the refusal rate, have showed an imperceptible increase only during 2017. The observed trend is presumably due, at the beginning, to the introduction of an interviewers' integrated network, and then to its stabilization during 2018. In the first year of its life, the new network had greater difficulty in ensuring a certain timeliness in the execution of interviews without any loss in terms of expected responses. In particular, this last result is also partially the effect of the introduction in the contract with the private company of the payment of some penalties if this does not reach a certain threshold in the percentage of complete interviews on those expected and a certain threshold in the percentage of complete interviews carried out in the first two weeks; on the other hand, penalties are not envisaged for aspects connected to substitution of the units.

Figure 3 – LFS main quality CAPI indicators by quarter – 2016-2018

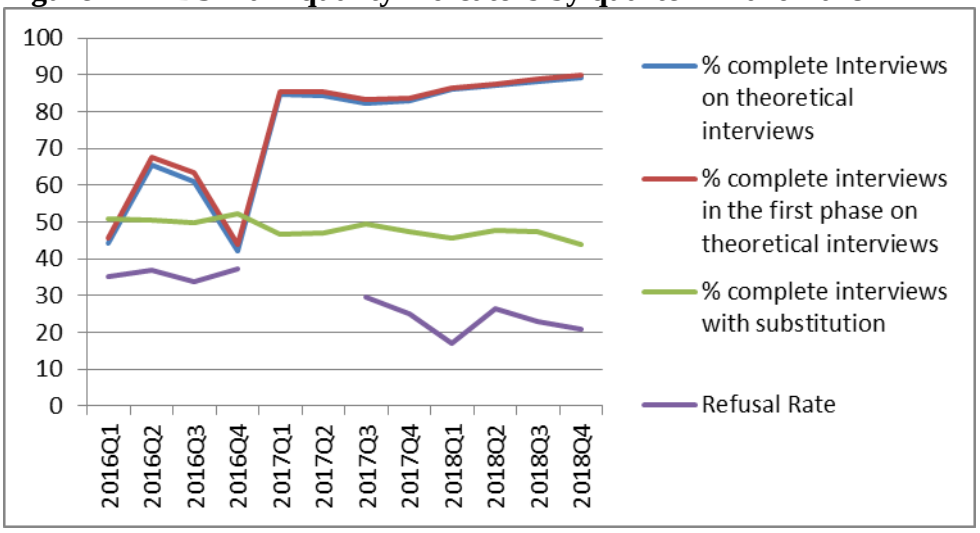


The HBS trend of the percentage of complete interviews and the percentage of complete interviews in the first phase on the total expected theoretical interviews, the percentage of complete interviews with substitution and the refusal rate are shown in the graph in Figure 4.
In the case of the HBS, the percentage of complete interviews on the total expected theoretical interviews shows a considerable increase, with the introduction of the interviewers' integrated network. In the four quarters of 2016, this indicator had reached levels slightly above 40% (in the first and fourth quarter 2016), remaining however below the 70% threshold (in the second and third quarter 2016). At the base of this figure there was a problem of coverage of the territory by the previous interviewers' network and a consequent impossibility of assigning all the interviews to be carried out[6]. From 2017,

[6] The percentage of assigned interviews on theoretical interviews from 2017 is always 100% also for HBS.

on the other hand, this percentage stands at levels that are always above 80%, up to almost 90% in the last quarter 2018[7].
The second indicator calculated for HBS, and considered to compare the time trend of the survey before and after the transition to the interviewers' integrated network, is the percentage of interviews completed in the first phase of the fieldwork, on the total expected theoretical interviews. The difference with the first indicator represents the "loss" of completeness, which occurs between the first and the last phase of data collection. With reference to this aspect, an average loss of almost two percentage points can be observed in 2016. In 2017, the new network improved this performance, keeping the loss of completeness below one percentage point. In 2018, this difference was further reduced.

Figure 4 – HBS main quality indicators by quarter – 2016-2018



The percentage of complete interviews with substitution, for this survey, is always high, in the whole period considered. Over time the indicator is above the line by 40% and below that of 55%. The particular survey methodology - with the three distinct moments for the collection of information at the single survey unit, and with time intervals that are fix on each phase - makes "physiological" the tendency to replace households, which are absent in coincidence with one (or more) specific predetermined time interval. Moreover, the survey is perceived as much more demanding than LFS, generating, in general, a greater number of refusals. However, this element also improves with the introduction of the integrated network, and the level of substitutions in the quarters of 2017 and 2018 is lower, compared to that recorded in the quarters of 2016.
The refusal rate also improves overall, while showing an irregular trend. In the third and fourth quarters of 2017[8], the indicator already shows improvements, compared to the corresponding quarters of the previous year. Moreover, in all the quarters of 2018 there is a further lowering of the refusal rate compared to the same quarters of the past period. This figures demonstrates the progressive investment made by the external company for taking charge of the survey, whose contents and collection methods required, by the interviewers, the development of new strategies for initial contact and, from part of the field coordinators, the refinement of techniques for monitoring and the continuous adjustment of

[7] To make this indicator comparable to the past and to the analogous one of LFS, the calculation of the indicator was carried out, simulating the same numerical relationship between theoretical and substitute interviews in each of the two surveys (three substitute households for each theoretical household).
[8] The refusal rate for the first and second quarter 2017 is not shown here. In fact, due to a problem in downloading data, the data was not sufficiently reliable.

workloads. Similarly to what argued for LFS, the contract with the private company provides for penalties for failure to achieve a certain level of the percentage of complete interviews on those expected, but it does not provide for penalties for aspects connected to substitution of the units.

4. Conclusions

LFS and HBS integration from the interviewers' network point of view turned out a good strategy that has allowed to maintain high LFS response rates and, at the same time, to increase significantly HBS response rates. On the one hand, Istat was able to lay the premises to promote the realization of this ambitious result, on the other hand, the company, which won the call for tender, was able to take up the challenge and did its best to achieve the result, for example by organizing the work in such a way as to minimize the interviewers' turn-over; this goal is also very important as it allows fostering better stability in the survey and increasing the specific professional skills.

Only in an initial period, which we can define as transition, some performance indicators for LFS were affected by the change and then returned to almost normal once the interviewers' network has gone to full speed. In summary, LFS peculiarities, in particular the continuity over time and the capillarity in the national territory, are suitable so that its network can be used to embed much smaller surveys but with continuity features, without losses in terms of expected responses. This innovation, already introduced in the LFS and HBS design, can be considered one of the first steps of the integration project of social surveys to which Istat is currently moving in the context of the new IESS Framework regulation.