

# Italian Methods and Practices Regarding the LFS: Case of Response Rate and Response Burden

Rita Lima<sup>1</sup> | ISTAT, Rome, Italy

## Abstract

Data users and stakeholders judge the success of a survey by its response rate. According to the measures of the American Association for Public Opinion Research (AAPOR, 2011) and following the Eurostat's standard (EUROSTAT, 2005), the Italian Institute of Statistics (ISTAT) applied a number of best practices to increase the households response rate – and, consequently, confidence in data quality - such as: planning a well-structured and a possibly short questionnaire, sending advance letters to sample households to inform them of the pending survey, giving standard survey introductions to well-trained interviewers, and monitoring data collection at all stages of the fieldwork. The focus on high response rates has many consequences. One is a high burden placed on respondents when a survey repeatedly contacts sample households, as ISTAT Labour Force Survey, and the other is the pressure on interviewing staff to spend time and resources to increase their response rates and not to accept a non-interview.

## Keywords

*Response rate, interview burden on respondents, panel surveys*

## JEL code

*C83, C80*

## INTRODUCTION

As the demand for statistics is growing and becoming more diversified, in response to social and economic developments, there is also an increasing pressure on National Statistical Institutes (NSIs) to reduce the burden on respondents related to statistical surveys. The need for moderating the burden on respondents is embodied in the principle 9 of the European Statistics Code of Practice: “*The reporting burden is proportionate to the needs of the users and is not excessive for respondents. The statistical authorities monitor the response burden and set targets for its reduction over time*” (Eurostat, 2005).

The NSIs most common way to measure the burden on respondents has so far been to estimate the response propensities (Fisher and Kydonieffs, 2001). Response rates are used as a measure of data quality and efficient data collection operations because conventional wisdom presumes that higher

<sup>1</sup> Italian National Institute of Statistics, Social and Environmental Statistics Department, Socio-economic Statistics Directorate. E-mail: lima@istat.it, phone: (+39)0646731. The finding and conclusions in this paper are those of the author and do not necessarily represent the views of the Italian Institute of Statistics.

response rate provides more accurate results (Babbie, 1990; Rea and Parker, 1997), and response rates are often used to evaluate survey data quality (Atrostic et al., 2001; Biemer and Lyberg, 2003). However, it is generally acknowledged now that high response rate does not automatically lead to lower nonresponse bias and, thus, better quality. Therefore, many survey methodologists question the use of the response rate as a measure of nonresponse bias (Groves, 2006; Groves and Peytcheva, 2008).

Often low levels of nonresponse cannot ensure high quality. On the contrary, high response rates can even mask measurement error, coverage error, and item nonresponse problems. Biemer (2003) notes that *“there has been much focus on response rates in surveys, and some surveys with low response rates have been discontinued on the basis of the response rate without evaluating the bias.”* In addition, insight in the differences in the characteristics of respondents and nonrespondents must accompany response rates to understand nonresponse bias and the impact of a given response rate on the accuracy of survey estimates. Studies indicate that the behaviour displayed in the recruitment for the survey influences the respondent’s behaviour in answering or participating in the survey (Groves and Couper, 1998; Groves and Heeringa 2006; Couper et al., 1997; Couper, 2000; Stoop, 2005). E.g. when people are reluctant to participate because they are not interested, they may adopt satisficing behaviour, instead of a conscientious attitude.

In cross-sectional surveys, obtaining a completed interview requires locating the respondent, successfully making contact, and achieving cooperation. Failure at any one of these tasks will result in inability to complete an interview. In panel surveys, the process is the same; however it must be repeated at each round. Lepkowski and Couper (2002) theorize that a respondent’s overall response propensity for any given round is a function of their propensities for location, contact and cooperation conditional upon respondent characteristics, survey design features, organizational efforts and experience during prior rounds of interviewing. Each of these tasks requires different kinds of effort on the part of the interviewer and NSI’s survey organization. Locating respondents who have moved can involve conducting internet or public record searches, contacting former neighbours and relatives, or knocking on doors of possible new residencies. Contacting effort may include phone calls, in-person visits. Furthermore, if a respondent has relocated, these efforts are wasted if it is not immediately apparent that the respondent has moved. Refusal conversion efforts can involve re-contacting respondents, providing monetary incentives, and sometimes sending interviewers with a specialty in refusal conversion from outside of the area. Reducing these efforts would reduce the likelihood that people participate in a survey and increase attrition in panel surveys (Merce, 2012).

In general, NSIs need to understand the costs and benefits associated with the response rates they achieve. Which level of efforts and respondent burden are associated with achieving those rates and do these extra efforts improve the quality of survey estimates? The paper outlines how ISTAT organizes, conducts and analyses response burden in the case of survey, repeatedly contacts sample households, like the Italian Labour Force Survey (IT LFS) where each household in the sample is surveyed over a limited period of time. It deals with ISTAT practices regarding the following: an appropriate sampling method, more intensive use of data on the process of asking and answering questions<sup>2</sup> (Couper and Lyberg; 2005; Kreuter et al, 2010), and a control of longitudinal burden on survey participation concerning the request for participation in a survey and the subsequent decision (e.g., if the household in question agrees to participate, the subsequent answering behaviour when filling in the questionnaire).

## 1 A BRIEF OVERVIEW ON THE IT LFS SAMPLING METHOD

A priority for NSIs is to reduce the response burden. One reason for this is the assumption that a too high response burden may result in lower data quality and higher costs for data collection.

<sup>2</sup> For example, how long the interviews took, how many times there were contacts with each interviewee or attempts to contact the interviewee, the reluctance of the interviewee, and the mode of communication (Couper, 2000).

The first practices of NSIs in the area of statistical burden reduction are an appropriate sampling method and a better control of sample size. The IT LFS complements traditional survey data on labour market activity and income with an additional dimension: the changes experienced by individuals over time. It is a sample survey with a cross-sectional design and a longitudinal follow-up. Its sample design is a two stage sampling with stratification of the primary units. The primary sampling units (PSU) are the municipalities and the final sampling units (FSU) are the households, randomly selected from the registry offices in all the municipalities drawn at the first stage. Stratification of PSUs is carried out in each NUTS-3 domain and it is based on the population of the municipalities, so in each NUTS 3 domain, large municipalities, with population over a given threshold (also called auto-representative municipalities), are always included in the sample; smaller municipalities (not auto-representative) are grouped in strata, then one municipality in each stratum is selected with probability proportional to the population. With reference to the sampling method of FSUs, households are randomly selected from the registry offices in all the municipalities drawn at the first stage. As shown in figure 1, FSUs are rotated according to a 2-(2)-2 rotation scheme where households participate for two consecutive quarters, and then they temporally exit for the following two quarters, and finally come back to the sample for other two consecutive quarters (Trivellato and Torelli, 1989).

**Figure 1** Rotating panel design of the IT LFS

Year	Q	Mounth	Rotation groups and wave					
2011	4	October	A4	B3	E2	F1		
2012	1	January		B4	C3	F2	G1	
2012	2	April			C4	D3	G2	H1
2012	3	July				D4	E3	H2 I1
2012	4	October					E4	F3 I2 L1
2013	1	January						F4 G3 L2 M1

Source: Trivellato and Torelli (1989)

Therefore, 50% of the households, interviewed in a quarter, are re-interviewed after three months, 50% after twelve months, 25% after nine and fifteen months. According to the rotation scheme, all the members of the households, who were interviewed in different time periods, can be matched in order to obtain longitudinal data. In particular, individual records can be matched to produce 12-month and 3-month longitudinal data which is almost 50% of the total sample. As a result, each household is included in four waves of the survey in a period of 15 months. Every year about 71 000 households are drawn from the registry offices of about 1 100 municipalities and this new sample is gradually introduced as the first wave rotation group in each quarter. Similar data are also available from other labour force surveys in developed countries (in fact, the scheme adopted in Italy closely resembles the Current Population Survey carried out in the U.S.) and data like these have actually been used to estimate transition models to employment.

Although a statistical burden is implicit from the sample design with its panel structure, this rotating panel design is a way that potentially reduces its burden on respondents. As the size of the sample has

a direct impact on the statistical burden, a new sample of PSUs, instead, is available as long as a new stratification is introduced in the sample design. In 2004, stratification for the new continuous LFS was introduced.

The list of municipalities has always been the same from one year to another except for the fact that every year some replacements were necessary. The main reason for PSU's substitution is due to the depletion of the list of households in the registry office: almost all the FSUs were already selected in the previous years and an additional inclusion of the PSU in the sample would have meant an increased statistical burden on FSUs and consequently a greater risk of nonresponse. The criterion used until 2011 for substitution was to identify municipalities with high risk (more than 75%) of drawing a household already selected in the previous occasions. In 2012 a new stratification of the municipalities was made, to take into account the updated information on their population. Consequently, a new selection of the PSUs has been done and newly selected municipalities entered in the sample in the third quarter of 2012.

Since 2013 a random rotation of the sample municipalities has been introduced in order to avoid any discretion in the replacement of PSUs. This rotation concerns in particular the strata with the smallest municipalities, in particular where more than 50% of the municipalities in the stratum do not have sufficient demographic size for the selection, or the strata where the maximum number of the registered households is below a threshold or where the minimum number of the registered households is below a threshold.

## 2 USE OF DATA IN THE PROCESS OF ASKING AND ANSWERING QUESTIONS

There is no consensus on how to define response burden, although it is quite well known that one can distinguish between "*perceived response burden*" and "*actual response burden*" (Jones, 2011; Stoop, 2005). The first one is the subjective burden differing from one individual to another and the second should be the real burden which in principle could be measured. Clearly, the measurement of perceived burden is out of scope of official statistics. Thus, only the actual burden can be the basis for the quantification of response burden. The actual burden can be measured by the time required to perform the task of completing a survey. Other factors such as questionnaire length, density of sampling, cognitive load required by completing the survey, and the layout and interface of the reporting format have been suggested to affect the strain on the respondent unit.

In literature, similar concepts are respondent burden, response fatigue, and subject burden. Increased response burden has been proposed to result in lower response rates, reduced completion, and reduced data quality. Response burden may be particularly problematic in demographic groups such as the severely ill, older individuals, and children. A strong focus has been put on questionnaire length, and, consequently, potential response burden is frequently a rationale for reducing the number of items in existing questionnaires and is also driving development of questionnaires with a minimum of items.

Heberlein and Baumgartner (1978) present several factors that were found to influence response rates to mailed questionnaires. They have found that response burden, approximated by the number of pages (or questionnaire length), shows a significant influence on response rate. However, they do not further differentiate the response burden by accounting for the complexity of the posed questions. They also find that other factors, such as the saliency of the survey content and incentives given to the respondents, all have an influence on the response rate (for a description of the so called leverage-saliency theory, see Groves et al., 2000). However, the degree of this effect and its influence on the actual response rates that are the subject of discussion in the present paper are difficult to quantify.

## 3 A DIFFERENT DATA COLLECTION MODE

IT LFS data are collected by the combination of both computer assisted personal (CAPI) and telephone (CATI) interviewing. CAPI is usually used for the first interview, whereas CATI for subsequent interviews. Households without a telephone and non-Italian households are interviewed always by CAPI. Mode switching can be an effective way of enhancing response, although its actual effect depends on the initial single mode design. For instance, sometime NSIs have been unsuccessful to date

in enhancing the response of previously single mode CAPI or CATI surveys by adding the web as the first mode in a consecutive mixed mode design (Wetzels et al., 2007). Computer assisted interviewing reduces the interview burden for the respondents. It presents interviewers with screen of questions to be asked, with the software guiding the interviewer and respondent thought the interview. Inapplicable questions are skipped automatically based on prior response patterns and preload information. Wording for probes are suggested when a respondent provided a response that was out of range for a given item (Bergamasco et al., 2004; ISTAT, 2006). Various performance indicators are also estimated to measure process quality and data collection effort in the CAPI/CATI system. However, the difference between telephone and face to face modes highlight the need for different performance indicators. Thus, CATI data collection indicators are often not directly applicable to face to face interviews. To this end, ISTAT has fixed for both, CAPI and CATI, a number of rules on the interview management, transmission and execution, which have to be strictly observed. Figure 2 details some IT LFS performance indicators based on CAPI/CATI system for the final sample units (FSU).

**Figure 2** The performance indicators in the CAPI/CATI system

Indicators
% of assigned of FSU on theoretical sample
% of CAPI FSUs on total assigned FSUs
% of CAPI FSUs on theoretical sample
% of CAPI validated FSUs on assigned FSUs
% of CAPI validated FSUs on theoretical sample
% of CAPI validated FSUs on assigned FSUs
% of CAPI validated FSUs on validated FSUs
% of CAPI distributed FSUs on assigned FSUs
% of CAPI distributed FSUs on validated FSUs
% of CAPI distributed FSUs on theoretical sample
% of CAPI FSUs with at least one contact attempt on distributed FSUs
% of CAPI FSUs with at least one contact attempt on theoretical sample
% of CATI FSUs on total assigned FSUs
% of CATI FSUs on theoretical sample
% of CATI FSUs with at least one contact attempt on assigned FSUs
% of CATI FSUs with at least one contact attempt on theoretical sample
% of CAPI FSUs on distributed FSUs
% of CAPI interviews (base households) on distributed FSUs
% of CAPI interviews (replacing households) on distributed FSUs
% of CATI interviews on assigned FSUs
% of CATI interviews (base households) on assigned FSUs
% of CATI interviews (replacing households) on assigned FSUs
% of CAPI and CATI interviews on theoretical sample (completeness rate)
% of CAPI and CATI interviews (base households) on theoretical sample (fidelity rate)
% of CAPI and CATI interviews (replacing households) on theoretical sample (infidelity rate)

Source: ISTAT (2006)

In LFS, ISTAT calculate and monitor mode-specific response rates (e.g., Lima and Ranaldi, 2012). Given the sequential nature of IT LFS data collection, the response rates in the early mode (CAPI) drive the subsequent mode (CATI and CAPI). The interviewing success in both modes defines the survey's non-response universe (Table 1). Hence, this paper focuses on the response rates and burden associated with CAPI/CATI system.

**Table 1** Rates of nonresponse by survey mode (Annual average – IT LFS 2011)

Survey Mode	Response	Non-response
CAPI	80.4	19.6
CATI	93.5	6.5
Total	88.9	11.1

Source: Author's computation

ISTAT managers analyse not only the survey's response rates across the country and regularly monitor mode-specific response rates, but also the efforts to maintain these rates. This paper focuses on some additional metrics that can help us understand the burden placed on survey respondents in the data collection process.

#### 4 THE ADVANCE LETTER

In order to motivate householders to participate in IT LFS, ISTAT has sent an advance letter (a letter to announce a forthcoming survey) by its President and explained to them that the survey is of national importance and that from the information collected will ultimately benefit all Italians. The idea is that respondents who understand the purpose of the statistics are more motivated to respond (De Leeuw et al., 2007). In general, NSIs consider the letter as an opportunity to explain what the survey is about, why it is important for the respondent to participate, and how and why the respondent was selected, underlining the survey organization's authority (Cialdini, 2001). Moreover, it gives interviewers a point of reference, and is a source of confidence for them in the initial contact (Becks, 2008). Advance letters might be able to increase response rates – given a response burden – because they intend to make participation more attractive. In surveys that could be more burdensome for respondents (e.g. involving diary keeping) the respondent believes that answering is important for the society and, thus, decides to participate (Groves et al., 2000). On the contrary, experiments with advance letters have shown that minor adjustments to the content of the letter can affect the response rate (e.g., Luiten, 2013; Campanelli, Klaassen and Beukenhorst, 2008; White, Martin, Benneth and Freeth, 1997).

Kanuk and Berenson (1975) also found an improvement in response rate when special delivery was implemented. However, they have found no significant relationship between personalization of advance letters and response rate, and that a cover letter may reduce response rates if the respondents disagree with the aims of the survey. In this sense, an advance letter could be a vehicle for applying socio-psychological theories in order to be as persuasive as possible towards respondents (Cialdini, 2001). Analysis of the effect of advance letters on response rates by survey mode and waves on IT LFS data (see Table 2 for the 1<sup>st</sup> quarter 2011) shows a significant difference between households in terms of either response rate between those who had received a letter and those who had not (89.9% vs. 80.8%, respectively).

Clear differences emerged also in terms of wave: while for later waves the difference in response rates between who has received the advance letter and who has not, is quite moderate. In the first wave this

**Table 2** Response Rate with and without advance letter by different survey mode and wave in IT LFS (1<sup>st</sup> quarter 2011)

Survey mode	Advance letter	Wave 1	Wave 2-4	Total
Total	Yes	87.3	91.2	89.9
	No	67.0	89.6	80.8
CAPI	Yes	88.9	91.5	90.1
	No	67.5	91.1	76.5
CATI	Yes	86.7	95.3	92.8
	No	66.5	89.2	82.7

Source: Author's computation

discrepancy is on average larger showing more than 20 percentage points of difference (87.3% vs 67.0% for the total).

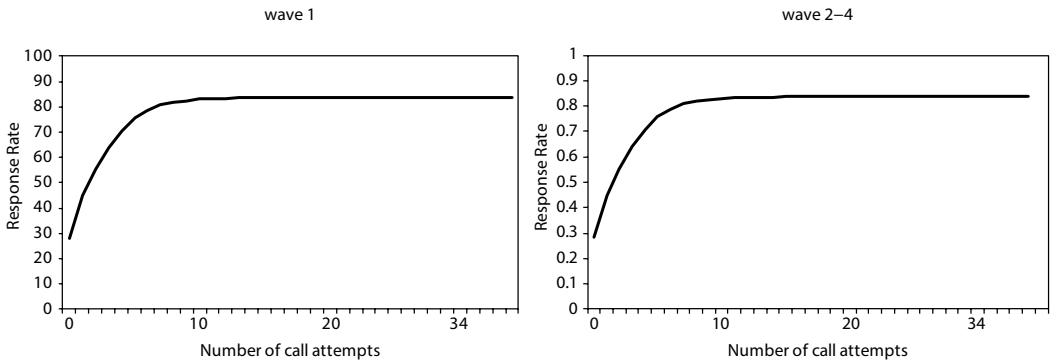
**5 THE OPTIMAL NUMBER OF CALL ATTEMPTS**

The NSIs' focus on response rates has many consequences. One is an increased burden placed on survey respondents when they have to complete a questionnaire several times, as in IT LFS. The sequential mixed mode survey may include receipt of numerous contact attempts by phone and in-person visits. In telephone surveys, double-digit call attempts are common in an effort to increase response rates. Relentless pursuit of a respondent through repeated personal visit attempts is not unusual in some surveys. Another consequence is the pressure on interviewing staff to spend time and money to increase their response rates and not to accept a non-interview. This pressure to obtain high response rates can result in interviewer's error and even falsification. Such behaviour of the interviewer might be perceived by a household respondent as stalking (Poe, 2011). Having considered the number of contact attempts as a sign of respondent reluctance, a good recommendation is to collect and analyse detailed contact record data. Such information could help to understand reasons for nonresponse and whether certain interventions might be appropriate. The IT LFS contact history instrument (CHI) provides a resource to understand the potential respondent burden associated with the response rates achieved, including information such as number of contacts, interim outcomes, reasons for refusals, date and time of contact attempt, and demographics of refusal and contact households. In addition, it includes additional detail about the reasons for noncontacts and non interviews by contact attempt.

It is easy to derive from CHI alternative statistics like contact rates, cooperation rates, and refusal rates from the full breakout of final outcomes (see for detailed descriptions and definitions AAPOR, 2011). All of these metrics could tell us much more than the response rate alone.

Figure 3 shows the relationship between the response rates obtained in the 2011 IT LFS operation and the associated number of contact attempts by modelling the "time to resolution" of each sampling units and the number of attempts until a response see, for more details, Lima and Ranaldi, 2012). In the 1<sup>st</sup> quarter 2011, the final IT LFS CATI response rates have consistently been around 94%.

The response rates in Figure 3 are estimated as ratios of the number of completed interviews after each round of contact attempts relative to the total eligible universe (RR2 in AAPOR 2011). They focus on the ratio of completed interviews to eligible cases. A useful tool to review the full distribution of interviewer outcomes is to include the proportion of cases classified as ineligible. A sample unit is eligible if the name corresponds to the selected household and if it is a private household having usual residence in the municipality. On the contrary, the eligibility is unknown when it is not possible to collect sufficient information for a proper classification: for example, in CAPI no contacts are made, no one at home and no other information available is collected during the fieldwork period; in CATI a sample unit is unreachable due to telephone number.

**Figure 3** Response rate by number of call attempts for wave 1 and waves 2–4 – CATI IT LFS (1<sup>st</sup> quarter 2011)

Source: Author's estimation

To estimate the eligibility rate in IT LFS a continued-contacting approach is used, based on the CHI and the number of different attempts of contact through an automatic procedure, regardless of whether contact is made.

This makes call record data particularly useful for nonresponse analyses and nonresponse adjustments.

The “*contact attempts*” in Figure 3 comprise any form of attempt to contact the sample household. This could include a successful contact leading to an interview, a refusal, or a request for a call back - until each unit is resolved leading to a response or to a nonresponse – or many forms of noncontact. In different attempts to contact the sampling units (the household is the basic sampling unit), all the telephone numbers are dialled with a justifiable “*random censoring*” calling rules of call attempts for a given phone number that is the end of the fieldwork period.

The “*maximum number of planned contact attempts reached*” concerns the cases where all attempts foreseen were made, but without managing to contact the household and, thus, to collect the necessary information to establish if the unit is eligible. So, a case with 15 call attempts resulting in “*no answer*” or “*busy*” outcomes could be classified as non-contacted unit after 15 attempts. In the case of a telephone number attempted 16 times with the first 15 attempts resulting in “*no answer*” or “*busy*” outcomes, the 16th call results in a refusal.

Here a broad definition of “*respondent burden*” is used, i.e. the sum of all contact attempts. It is hard to know which forms of attempted contacts a respondent would perceive as burdensome. In fact, a core issue is the definition of the optimal number of contact attempts that the interviewer may make with a sample unit: abandoning contact attempts prematurely (people who are hard to contact for the LFS are workers and travellers, see also Lynn et al., 2002), or allowing a relatively high noncontact rate will introduce bias in the estimation of the response rate and of the eligibility rate. Recently, high respondent burden and repeated contact attempts to obtain high response rates have been criticized (Martinsson and Riedel, 2014). NSIs’ fieldwork organizations generally prescribe a minimum number of contact attempts before writing off a survey unit as a non-respondent unit. This number will be lower for CA-PI than for CATI surveys, because of differences in the costs of contact attempts in the various modes. Sometimes in CATI, no rules to determine the excess of attempts are set and the contact attempts have to continue up to the end of the data collection period.



Often, CATI fieldwork organizations do not permit unlimited contact attempts, because at some point the costs of additional approaches will no longer outweigh the benefits in terms of bias reduction of marginal improvement in response. Moreover, the number of visits determines not only the contact rate, but also the quality of the data: there are differences between the people you reach immediately and those you find only after several visits (Luiten, 2013).

Given the large sample size in the IT LFS, even a low percentage of cases with high numbers of contact attempts can represent a notable universe of respondents that may feel harassed by the survey’s repeated contact efforts. To give you a sense of scale, look at the universe of sample units in the 2011 CATI operation that are ultimately determined to be survey-eligible units. This universe includes sample cases that are able to be interviewed and those that could not be interviewed. Table 3 shows that interviewers resolved about 43% of these units after 1 or 2 CATI contact attempts.

**Table 3** CATI IT LFS Contact Attempts (1<sup>st</sup> quarter 2011)

Total CATI Contact Attempts	Percent of eligible*	Estimated respondent units
1 or 2	43.38	17.92
3 or 5	24.06	9.94
6 or 9	11.71	4.84
10 or 15	6.13	2.53
more 16	5.86	2.42

\* The eligibility rate is estimated by the SAM method in Brick et al. (1997, 2002).

Source: Author’s computation

In contrast, we estimate<sup>3</sup> that just over 6% (about 2 500 individuals) involved 10 or 15 CATI contact attempts.

In the waves following the first one, there are the highest numbers of call attempts. This is not surprising, given the greater number of call-backs that are normally for the CATI IT LFS. The bulk of cases (about 80% of sample units) are a respondent household at the 10th attempt and after the 20th call attempt there are only 2% of unsolved units. As in earlier studies we found no association between the difficulty of making contact and response rate (see for example Lepowski et al., 2002; Stoop, 2005). The highest households’ response propensity in IT LFS, may not consist exclusively of cases with the highest number of call attempts. So other ways are necessary to try to identify cumulative respondent burden, considering earlier contact attempts in other modes. Our future research intends to look more closely at how to acknowledge this burden in any metric and how to utilize previous burden in developing CATI stopping rules.

**7 THE TIME SPENT ON THE INTERVIEW**

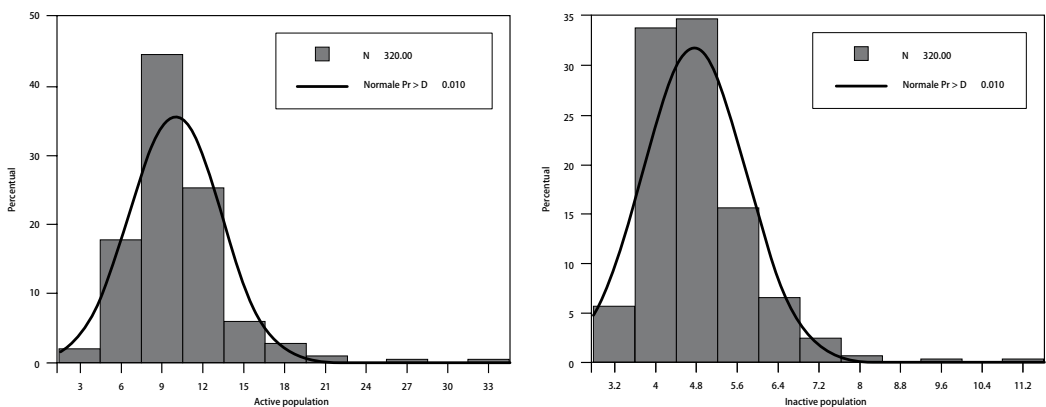
Computer-aided collection is an important factor in reducing NSIs’ interview times: automatically filtering questions, according to the interviewee’s previous responses or known characteristics, means that the interviewer will not waste time by asking non-applicable questions or questions that have already been

<sup>3</sup> See Lima and Ranaldi (2012).

answered in the previous waves. In fact, dependent interviewing in the successive waves of the survey enables to avoid re-asking questions about the individual's characteristics and to just identify changes (for example about educational levels or qualifications). The European Council Regulation No. 1897/00 allows to conduct interviews following the first one asking the sampling person to confirm the information given in the previous interview. Consequently, when the labour status and/or other characteristics of the respondent have not changed substantially, we are able to reduce the time of interview.

The IT LFS computer-aided collection realises a "confirmation questionnaire", in which the confirmation questions are the fundamental nodes of interview flow. If the status of the respondent has not been changed, all information regarding a sub-flow of the questionnaire depends on certain confirmation questions that are automatically registered and skipped (Grassia et al., 2004). It turns out that the burden of completing a survey does not affect the propensity that a respondent will agree to participate in the survey at a later time point (Figures 4 and 5).

**Figure 4** Response times by participation of people to labour market. CAPI (Median in minutes, year 2011)

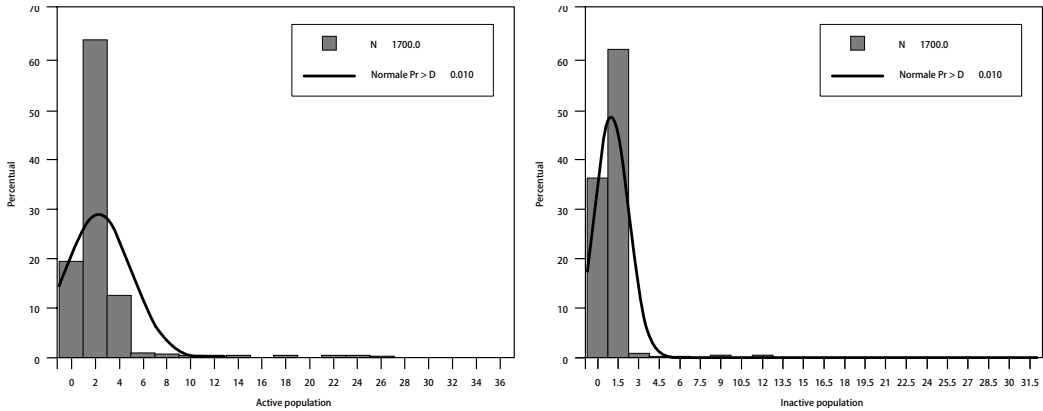


Source: Author's computation

Time passes quickly if the topic is interesting and the respondent feels that he/she is competent to answer the questions posed, while it feels longer if the respondent does not understand the point of the questions or feels incompetent (EUROSTAT, 2007). In 2011, for CAPI individual questionnaires the successful attempts (the outcome code indicates a response) have a length ranging from a few minutes to 33 minutes. The extreme cases should be excluded from further studies using other available paradata. For CATI, telephone attempts tend to be shorter and have lower variance in duration compared to personal visits: It is not surprising given that in general during the IT LFS data collection, CAPI is used for the first interviews while CATI is used for the confirmation interviews following the first wave. Moreover, the modalities of approaching the households, the rules that fix the assignment of the final outcome codes to the households and the management of the contacts with the households differ, depending on the survey mode and the family wave. So, since the first interview is long, the CAPI visits take a long time. The CATI interviews are only to pick up a confirm of information collected in the previous wave. Therefore they are much shorter. The big difference between CAPI and CATI in time to respond is between "inactive" and "active" (Figure 4 and 5).

The slowest questionnaire to fill in is for the active people: about 60% has a median of 10 minutes in CAPI, while the bulk of CATI interview takes less than 6 minutes.

**Figure 5** Response times by participation of people to labour market. CATI (Median in minutes, year 2011)



Source: Author's computation

For the inactive people, about 80% of interviews take under 6 minutes in CAPI and less than 1.5 minute in CATI. Based on the above analyses, it also becomes clear that, in this way, the NSIs traditional response burden indicator “*time spent*” (not the time use itself but the perception of time and effort that is likely to affect response quality) may not be the best response burden indicator in IT LFS.

**8 LONGITUDINAL SURVEY PARTICIPATION**

In the IT LFS questionnaire, the interviewer, after finishing the household interview, has to answer the following two questions in Figure 6. The aim is to obtain an evaluation of the longitudinal component of response burden, approximated by the willingness to respond. Although clever sampling designs and dependent interviewing reduce the time the respondent has to spend on the survey, it is possible that: a) the households’ perception of a low response burden may give rise to an atmosphere

**Figure 6** Questions about Perceived Response Burden in the IT LFS

**Question 1. How was the availability of the household to the interview?**

- Always little 1|\_
- Little at the beginning, better then 2|\_
- Good at the beginning, little then 3|\_
- Enough 4|\_
- Good 5|\_
- Excellent 6|\_

**Question 2. Is the household available for further interviews in the coming months?**

- Not at all available 1|\_
- Little available 2|\_
- Indifferent 3|\_
- Enough available 4|\_
- Very available 5|\_

Source: ISTAT Contact history form

of motivation and willingness to respond, as well as to the quality of their response, and b) the lower response rate in subsequent interviews in the coming months could be explained by monitoring facets of survey climate over time and during the previous interview.

When analysing the answers to the above questions, ISTAT monitor changes over the “*life cycle*” phase of households (the final sample unit) or, in other cases, measures differences on a matter of a suitable group of households (the reference population or cohort of individuals). Here it is proposed a measurement in longitudinal evaluation of burden on survey participation of Italian households on a cohort of individuals who differ with respect to the perception of availability to the interview and for recall waves.

The group consists of about 18 000 households, who share the common characteristic of having been interviewed in the first quarter 2011 for the first time and respond to the re-interviews at the subsequent waves until the second quarter 2012. The aim is to see whether they develop the outcome of interest (that is, the availability for further interviews). The longitudinal response burden is based on two questions filled by the interviewer, as soon as closed the IT LFS interview (Figure 6), to evaluate the cohort’s respondent availability to be contacted again for possible future surveys.

Table 4 displays estimated odds ratios and relative risks with their 95% confidence intervals: There is a positive association between availability and response but it is not statistically significant because all 95% confidence intervals cover unity.

With respect to the patterns of attrition, this analysis has assumed that rules governing the level of effort expended do not differ by wave. In fact, wave 1 exhibits the highest risk of attrition, followed by a levelling off and gradual reduction in risk between wave 2 and 4. However, given the disproportionate level of nonresponse that occurs in wave 1, it may be worthwhile to investigate the patterns of attrition for wave 1 and subsequent waves since rules governing the level of effort differ by waves.

**Table 4** Odds ratios, relative risk and 95% confidence intervals (1<sup>st</sup> 2011–2<sup>nd</sup> 2012)

Event by wave	Odd Ratio	95% CI	Relative Risk	95% CI
Interview in wave 2 and availability at wave 1	1.85	1.57–2.18	1.07*	1.05–1.10
Interview in wave 3 and availability at wave 2	2.42	1.97–2.97	1.10*	1.07–1.14
Interview in wave 4 and availability at wave 3	0.26	0.21–0.31	0.95*	1.97–2.97

Source: Author’s computation

## CONCLUSION

As all NSIs, ISTAT has features that managed response burdens including sending an advance letter, using a mixed-mode collection data and measuring how long the interviews were. Three new proposals how to further lower response burden have recently been tested for IT LFS. The first proposal, which is the rotating panel design and the control of sample selection with the PSUs’ substitution, proved to be an effective way to reduce the response burden of households. The estimation of the optimal number of call attempts with a continued-contacting approach according to the Survival Analysis Method, which is to define a threshold along the continuum of contact attempts, also reduces the response burden. The second proposal is the analysis of interview duration as a measure of the pressure on interviewing staff to spend time and resources to increase their response rates and not to accept a non-interview, which could be useful to include the interviewer in the model, as the duration depends on both interviewer and respondent.

The last proposal concerning the evaluation of the longitudinal burden on survey participation to the interview and for recall waves, also reduces the response burden, but it came with some risks.

It is based on the questions about the interviewing process filled by the interviewer. It has not been decided yet whether this proposal will be implemented in the survey production.

In the future, the next step will be improved tools to identify new performance measures related to the measure of burden on respondents for example, a strategy for optimizing the sampling and collection processes at the household level, which is the actual survey respondent, by minimizing overlap of LFS respondents with other household surveys. In this way, we propose a monitoring of burden on respondents from a wider point view, and not limited to a specific survey.

## References

- AGANS, R. P., KALSBECK, W. D. *Nationwide Study of Female Sexual Health*. Chapel Hill, NC: University of North Carolina Survey Research Unit, 2005.
- ATROSTIC, B. K., BATES, N., BURT, G., SILBERSTEIN, A. Nonresponse in US Government household surveys: Consistent measures, recent trends, and new insights. *Journal of Official Statistics*, 2001, 117, pp. 209–226.
- BABBIE, E. R. *Survey Research Methods*. Belmont, CA: Wadsworth, 1990.
- BERGAMASCO, S., GAZZELLONI, S., QUATTROCIOCCI, L., RANALDI, R., TOMA, A., TRIOLO, V. New strategies to improve quality of ISTAT new CAPI/CATI Labour Force Survey, *European Conference on Quality and Methodology in Official Statistics*, Mainz, 24–26 May, 2014.
- BIEMER, P. P., LYBERG, L. E. Introduction to Survey Quality. New York: John Wiley & Sons, Inc. Council of American Survey Research Organizations (1982). *On the Definition of Response Rates*, Port Jefferson, NY: CASRO, 2003.
- BRICK, M., MONTAQUILA, J., SCHEUREN, F. *Estimating the Residency Rates for Undetermined Telephone Numbers*, *Public Opinion Quarterly*, 2002, Vol. 66, pp. 18–39.
- CIALDINI, R. B. *Influence. Science and practice*. Boston, MA: Allyn and Bacon, 2001.
- COUPER, M., HANSEN, S., SADOSKY, S. Evaluating interviewer performance in a CAPI survey. In: LYBERG, L., BIEMER, P., COLLINS, M., DELEEUW, E., DIPPO, C., SCHWARZ, N., TREWIN, D. (Eds.) *Survey Measurement and Process Quality*, New York: John Wiley, 2007, pp. 267–285.
- COUPER, M. *Usability evaluation of computer-assisted survey instruments*. *Social Science Computer Review*, 2000, 18(4), pp. 384–396.
- COUPER, M., LYBERG, L. The use of paradata in survey research. In: *Proceedings of the 55<sup>th</sup> Session of the International Statistical Institute*, Sydney, Australia, 2005.
- DE LEEUW, E., HOX, J., KORENDIJK, E., LENSVELT-MULDERS, G., CALLEGARO, M. The Influence of Advance Letters on Response in Telephone Surveys: A Meta-Analysis. *Public Opinion Quarterly*, 2007, 71(3), pp. 413C443.
- EUROSTAT. *European Statistics Code of Practice* [online]. 2005. <[http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/code\\_of\\_practice](http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/code_of_practice)>.
- EUROSTAT. *Handbook for Monitoring and Evaluating Business Survey Response Burdens* [online]. 2007. <[http://epp.eurostat.ec.europa.eu/portal/page/portal/research\\_methodology/documents/HANDBOOK\\_FOR\\_MONITORING.pdf](http://epp.eurostat.ec.europa.eu/portal/page/portal/research_methodology/documents/HANDBOOK_FOR_MONITORING.pdf)>.
- FISHER, J., KYDONIEFS, L. *Using a Theoretical Model of Respondent Burden to Identify Sources of Burden in Surveys*. Paper presented at the 12<sup>th</sup> International Workshop on Household Survey Nonresponse, Oslo, Groves, 2001.
- GRASSIA, M. G., PINTALDI, F., QUATTROCIOCCI, L. *The electronic questionnaire in ISTAT's new CAPI/CATI Labour Force Survey*. European Conference on Quality and Methodology in Official Statistics, 2004.
- GROVES, R. M., COUPER, M. P. *Nonresponse in household interview surveys*. New York: Wiley, 1998.
- GROVES, R. M., CIALDINI, R. B., COUPER, M. P. Understanding the Decision to Participate in a Survey. *Public Opinion Quarterly*, 1992, 56, pp. 475–495.
- GROVES, R. M., SINGER, E. CORNING, A. Leverage-Saliency Theory of Survey Participation. Description and an Illustration. *Public Opinion Quarterly*, 2000, Vol. 64, pp. 299–308.
- GROVES, R. M., HEERINGA, S. Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society, Series A: Statistics in Society*, 2006, 169, Part 3, pp. 439–457.
- GROVES, R. M. Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, 2006, Vol. 70(5), pp. 646–675.
- GROVES, R. M., PEYTICHEVA, E. The Impact of Nonresponse Rates on Nonresponse Bias. A Meta-analysis. *Public Opinion Quarterly*, 2008, 72 (2), pp. 167–189.
- HEBERLEIN, T. A., BAUMGARTNER, R., Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature. *American Sociological Review*, 1978, 43(4), pp. 447–462.
- ISTAT. *La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione* [online]. *Metodi e norme*, 2006, No. 32. <[http://www.istat.it/dati/catalogo/20060830\\_00](http://www.istat.it/dati/catalogo/20060830_00)>.

- KREUTER, F., COUPER, M., LYBERG, L. The use of paradata to monitor and manage survey data collection. Section on Survey Research Methods – JSM 2010 [online at amstat.org]. Citing: COUPER, M. Measuring survey quality in a CASIC environment. In: *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 1998.
- JONES, J. *Response Burden: Introductory Overview Lecture*. Paper presented at the 2011 Joint Statistical Meeting, 30 July–4 August, 2011, Miami.
- LEPOWSKI, J. M., COUPER, M. P. Nonresponse in the second wave of longitudinal household surveys. In: GROVES, R. M., DILLMAN, D., ELTINGE, J. L., LITTLE, R. J. A. (Eds.) *Survey Nonresponse*. Chichester: Wiley, 2002.
- LIMA, R., RANALDI, R. *Calculating the eligibility rate of sampling units with unknown eligibility*. 7<sup>th</sup> Workshop on Labour Force Survey Methodology – Madrid, 10–11 May 2012.
- LUITEN, J. M. M. *Improving survey fieldwork with paradata* [online]. Sociale Wetenschappen Proefschriften, 2013, <<http://www.cbs.nl/NR/rdonlyres/1071A190-B552-4758-94C3-B9E29CD584DE/0/2013x11Luitenpub.pdf>>.
- LYNN, P., CLARKE, P., MARTIN, J., STURGIS, P. The effects of extended interviewer efforts on nonresponse bias. In: GROVES, R. M., DILLMAN, D. A., ELTINGE, A., LITTLE, R. J. A. (Eds.) *Survey nonresponse*. New York: Wiley, 2002, pp. 135–149.
- MARTINSSON, J., RIEDEL, K. *Long-term effects on data quality of different initial response rates in web panel recruitment*. LORE Methodological Notes, Göteborgs universitet, 2014(2).
- MERCER, A. *Using Paradata to Understand Effort and Attrition in a Panel Survey* [online], 2012. <[https://www.amstat.org/sections/srms/proceedings/y2012/files/304514\\_73249.pdf](https://www.amstat.org/sections/srms/proceedings/y2012/files/304514_73249.pdf)>.
- POE, T. *The Pros and Cons of Making the American Community Survey Voluntary*. Prepared Testimony of Ted Poe (TX-02) Before the House Oversight and Government Reform Committee's Subcommittee on Health Care, District of Columbia, Census, and National Archives, March 6, 2011.
- REA, L., PARKER, R. *Designing and conducting survey research* (2<sup>nd</sup> Ed.). San Francisco: Jossey-Bass, 1997.
- STOOP, I. *The Hunt for the Last Respondent. Nonresponse in sample surveys*. Social and Cultural Planning Office, 2005.
- THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* (7<sup>th</sup> Ed.) [online]. 2011. <<http://aapor.org/Content/NavigationMenu/AboutAAPOR/StandardsampEthics/StandardDefinitions/StandardDefinitions2011.pdf>>.
- TRIVELLATO, U., TORELLI, N. Analysis of Labor Force Dynamics from Rotating Panel Survey Data. *Bulletin of The International Statistical Institute*, 1989, Vol. LIII, Book 2, pp. 425–444.
- WETZELS, W., JANSSEN, B., SCHOOTEN, M. *Mixed-mode enquêtering bij personen en huishoudens 2005–2007: Een overzicht* (Mixed-mode surveying of persons and households 2005–2007: An overview). Internal report, Statistics Netherlands, Heerlen, 2007.
- WHITE, A., MARTIN, J., BENNETH, N., FREETHS. *Improving Advance Letters for Major Government Surveys*. ZUMA Spezialband 4, 1997.