EVALUATING COMPUTER-ASSISTED QUESTIONNAIRE USABILITY: THE CASE OF PERMANENT CENSUS OF THE POPULATION AND HOUSING[1]

Sabrina Barcherini, Katia Bontempi, Manuela Bussola, Barbara Maria Rosa Lorè, Simona Rosati

Abstract. This paper examines the optimization of a questionnaire, aiming to simplifying the respondent's task and improving data quality. Various methods, including pre-testing techniques and paradata analysis, are utilized to evaluate the survey tool's content, usability, and functionality. Focusing on the Census of the Population and Housing questionnaire, the study analyzes data from 2019 and 2021 to assess the impact of usability optimization on reducing completion difficulties. Findings demonstrate the effectiveness of the usability solutions implemented in alleviating critical issues and addressing respondents' difficulties in completing the questionnaire. An analysis based on regression methods has been used to classify households according to their dependence on external help to complete the questionnaire.

1. Introduction

Usability is a fundamental concept in user-centered design. It is a measure of how easy something is to use and refers to the interaction between product and user. The International Organization for Standardization (ISO) defines usability as the "*extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use*" (ISO 9241-11:2018). In this perspective, a product is usable when users can fulfil their task with minimal effort, stress and errors, while feeling satisfied with their interaction with it. This necessitates the design of interfaces that are intuitive, easy to use, and capable of supporting users in their tasks (Nielsen, 1993).

In the context of questionnaire design, usability focuses on creating a smooth and effortless user experience. A well-designed questionnaire should function as a virtual

[1] This article is the result of the collaboration between the authors. In particular: paragraphs 1 and 7 have been written by Barbara M. R. Lorè, paragraph 2 has been written by Manuela Bussola, paragraphs 3 and 4 have been written by Simona Rosati, paragraph 5 has been written by Katia Bontempi, paragraph 6 has been written by Sabrina Barcherini.

assistant, interacting with respondents, performing some operations on their behalf, reminding them of their progress, and providing instructions for the next steps. Taking usability principles into account when designing a questionnaire will result in an easy and stress-free completion experience, leading to low burden, promoting respondent fidelity, and ultimately yielding satisfactory response rates and accurate data. As usability is a key factor in creating a successful product, it is important to place the user at the centre of the product design and development process. When designing a questionnaire, this entails understanding what features of the instrument might make users uncomfortable. A number of methods are available to the researcher to explore the respondents' difficulties with the questionnaire. Some of these are specifically designed to gather information on critical issues, while others serve as secondary sources. The first category encompasses cognitive interviewing, respondent feedback, interviewer debriefing, pilot surveys and experiments, while paradata and help desk tickets analysis fall into the second group (Barcherini et al, 2022).

This paper discusses the results of respondent feedback, interviewer debriefing and helpdesk ticket analysis from the census of the population and housing. In 2018, when the Italian census became permanent, a brand-new electronic questionnaire was developed, for both CAPI (computer-assisted personal interviewing) and CAWI (computer-assisted web interviewing). The questionnaire consists of three parts: the list of the household members, pre-filled with administrative data and requiring verification and potential edits by the respondents; an individual questionnaire for each family member; and a questionnaire for the household as a whole, collecting information on the dwelling. During the initial two years, evidence of respondents' difficulties with the new questionnaire was reported both by field staff and respondents themselves. Consequently, in 2020 layout and functionality changes were implemented to address these issues.

In the following paragraphs, a comparison between the new and original versions of the questionnaire is provided to evaluate the extent of improvement.

2. Critical points in the use of the electronic questionnaire and the need for improvement

In 2018, a debriefing with the Heads of the Municipal Census Offices (HMCOs) was conducted through an electronic questionnaire a few weeks after the conclusion of the census survey. It provided a valuable opportunity to gather feedback from the managers. A specific section of the debriefing questionnaire aimed to collect the managers' opinions on the usability and navigation of the electronic questionnaire. The managers' views were based on their interactions with households that had

received assistance from to Census Office (UCC) and on the experiences reported by fieldwork staff who conducted face-to-face or phone interviews the households using the computer-assisted (CA) questionnaire.

The HMCOs highlighted some recurring issues. The majority of them reported that both the field staff (88.9%) and the households (79.5%) had experienced difficulties, either occasionally or frequently, in using the questionnaire. Only a residual percentage of managers stated that neither the households (20.5%) nor the fieldworkers (15.1%) had found any difficulties. The HMCOs were then asked to clarify the nature of these problems. In addition to issues related to the application itself, such as network slowness and access problems, 38.1% of HMCOs received reports from households of difficulties during the final submission stage. Fieldworkers also encountered similar problems: in this case, the percentage of HMCOs who collected at least one report of this type increased to 41.1%. Furthermore, 13.3% of HMCOs reported feedback from households about difficulties in navigating the electronic questionnaire. However, this percentage decreased to 9.7% when it was the fieldwork staff who reported the problem to their HMCO. These results highlighted the existence of some critical issues in the original version of the electronic questionnaire that required attention. Improvements needed to be made in order to enhance the efficiency and usability of the system and to ensure a better overall experience for respondents during the data collection process.

In the second year of the census (2019), requests for assistance from households received by the dedicated Contact Center were analysed. A total of 2,276 assistance requests out of approximately 43,000, were specifically related to the navigation and usability of the electronic questionnaire (Istat, 2021). The results show that three out of four requests were related to difficulties encountered during the final submission of the questionnaire. This high percentage highlights the importance of focusing on the difficulties encountered by users in completing and submitting the electronic questionnaire. It is therefore essential to allocate resources and efforts towards improving this phase of the census process, to help users submit the questionnaire correctly and timely. Less common but still noteworthy, were difficulties in accessing individual questionnaires, which accounted for 13.2% of the requests for assistance. Similarly, 5.7% of the assistance requests were related to adding family members to the initial list or completing editable fields. Although these figures are lower than the previous ones, they still require attention in order to guide subsequent improvement measures. Finally, 4.4% of the requests for assistance revealed problems related to the partial saving of data. This issue is relevant as inefficient partial saving can lead to data loss and to an overall increase in the burden on respondents.

In conclusion, the analysis of the requests for assistance, which focused on the navigation and usability of the electronic questionnaire, confirmed the HMCO's

opinions. The main issues identified were related to the final submission process, and to a lesser extent, to the access to the individual questionnaires, the use of editable fields, and the partial data saving.

3. Usability improvements implemented in the computer-assisted questionnaire

Based on the issues identified in 2018 and 2019, and taking advantage of the interruption of the census in 2020, a number of improvements have been implemented for the 2021 edition of the computer-assisted questionnaire (CA questionnaire) in order to address the identified problems and eliminate, or at least reduce, their impact.

As mentioned above, it is important to consider the usability of a CA questionnaire within the broader context of websites usability. Following the definition of usability as "the extent to which specified users achieve specific goals with effectiveness, efficiency, and satisfaction in a given environment" (ISO, 2018), it is evident that in the context of a CA survey, users are represented by respondents in CAWI surveys and interviewers in CAPI surveys, while the website is represented by the questionnaire itself (Gabbiadini, Mari, Volpato, 2011).

According to Gabbiadini, Mari, and Volpato, it is possible to apply the five qualitative components that define web usability to the essential attributes that a CA questionnaire should possess to be deemed usable. Adhering to these principles involves creating a CA questionnaire that "is appropriately interpretable, thus reducing the cognitive load on respondents and minimizing errors stemming from the inherent design characteristics of the web artefact" (Gabbiadini, Mari, Volpato, 2011, p. 250). These attributes include: a) Learnability. The ease with which users can successfully complete all required operations; b) Efficiency. The speed at which users can perform various response operations to the items; c) Memorability. The ease with which users can acquire and remember the necessary operations for subsequent completion sessions; d) Errors. The ability to proactively reduce the likelihood of errors in responding or to guide respondents in resolving any potential completion errors (reducing the number of errors, dropout rates, and partial completions); e) Satisfaction. The extent to which the system is perceived as enjoyable and reduces the cognitive load on respondents.

Based on these principles, improvements have been made to several aspects of the census CA questionnaire. Specifically, interventions have been conducted in the following three main areas: 1) graphic restyling; 2) simplified navigation; 3) guided completion.

Graphic restyling. Screen reading has been proven to take longer than paper reading (Nielsen, 2000). Therefore, web pages should be designed to facilitate reading and responding to questions. In general, in adherence to usability principles, text should be concise, appealing, and visually light.

In this perspective, the colour scheme of introductory text and completion instructions has been revised to align with the broader communication standards of the population and housing census (Gabbiadini, Mari, Volpato, 2011). For instance, the colour red has been removed from the text and reserved only for structural components of the screen (e.g., headings) and institutional titles, since red is the distinctive colour of the population and housing census. Additionally, certain keywords have been highlighted in a brighter colour to guide reading and focus the users' attention on the most relevant concepts for navigation and completion support.

Simplified questionnaire navigation. Providing users with a tool that guides and orients them throughout the completion process is crucial in the design of a CA questionnaire. For instance, the presence of a progress indicator allows users to track their progress and know their current position within the web system. To achieve this, the navigation menu has been simplified. Its colour scheme has been revised to better differentiate completed sections (indicated in blue) and sections that are yet to be filled out (indicated in gray). Moreover, efforts have been made to reduce call-to-action items and to make the buttons more intuitive. For example, the labels on the buttons that enables partial saving and advancement to subsequent screens have been made more descriptive. This is important to reassure users that they can progressively save their answers, in respect of the effort made up to that point (Nielsen, 2000; Polillo, 2006).

Guided completion. Reducing the cognitive load on respondents is a priority that a usable web system should ensure. In this regard, interventions have been made in three aspects of the functionality of the census CA questionnaire.

The first aspect addressed the verification and modification of the list of family members. This operation was initially confusing and redundant, requiring multiple actions within the same table, such as "edit," "confirm," and "complete." To streamline the process, pre-filled fields have been made immediately editable by simply placing the mouse cursor in the field, eliminating the need to click the "edit" button. The graphical display of the list has been simplified, and now shows only the data of the family members that the respondent must verify or update if necessary. The "confirm list" button has been made more prominent, placed outside the table, and visually standardized with other call-to-action buttons.

The second improvement focused on optimizing the operations required to access individual questionnaires. The number of steps has been reduced, and clear labels identify the required actions, which are arranged sequentially within the screen. Access to individual profiles has been emphasized with a dedicated button that only

appears displayed once the list of family members has been confirmed. Previously, access to individual profiles was embedded in the list of family members.

The third aspect addressed a significant issue, which was the difficulty in completing the final submission of the questionnaire. A number of measures have been taken to overcome this challenge. The screen has been visually streamlined to highlight the submit button. The screen has been visually streamlined to highlight the submission button. Graphics have been aligned with previous screens, and text has been simplified to eliminate redundancy.

4. The interviewers debriefing and help desk tickets analysis to evaluate the impact of the changes to the questionnaire

With the completion of the first cycle of the census (2018-2021), the municipal fieldwork network was invited to respond to a debriefing questionnaire, administered using CAWI technique. The purpose was to gather opinions and suggestions on the entire census process. Approximately 50% of the 23,524 operators responded to the questionnaire.

Specific questions in the consultation questionnaire were aimed at gathering feedback on any difficulties related to the completion of the CA census questionnaire. A total of 9,852 network operators who conducted field interviews responded to these questions, representing 80.4% of those who completed the consultation questionnaire. The results were overwhelmingly positive. According to the majority of network operators, the CA questionnaire works very well. Over 90% of operators reported no difficulties for every aspect of functionality and usability considered: filling in the family members list (98.3%); access to the individual forms (98.5%); navigation menu (97.3%); questionnaire navigation (95.3%); warning prompts (96.8%); tooltip use (98%); visualization of the sections summary (98.6%); access to the preview of the questionnaire (98.5%); final submission (97.9%).

This result is further supported by the significant decrease in the number of tickets received by the Contact Center in 2021 regarding the functioning of the CA questionnaire. Only 500 out of the 100,000 tickets collected were associated with completion difficulties.

In conclusion, these data demonstrate the effectiveness of the developed CA questionnaire and the progressive reduction of problems encountered during the completion process by the end users: respondents and interviewers.

5. The analysis of respondents' feedback to detect users' difficulties

The census questionnaire includes a final section of standardized questions aimed at collecting feedback on the respondents' completion experience. This feedback encompasses aspects such as the mode of completion, the difficulties encountered, requests for assistance, and more. The analysis of this feedback can help improve the questionnaire design and, consequently, make it more user-friendly.

In order to determine the impact of the usability improvements introduced in 2021 on reducing respondents' completion difficulties, the respondent feedback collected in 2019 and 2021 was analysed[2]. In 2019, CAWI was chosen by 51.4% of respondents, which increased to 53.1% in 2021. The percentage of questionnaires filled in by a member of the household rose from 86.9% in 2019 to 87.7% in 2021. Furthermore, the percentage of households that did not need any help to complete the questionnaire (not from friends or relatives, not from the help desk, not from the municipal census offices, etc.) increased from 78.2% to 79.9%.

When examining households where all members are elderly (65 years or older), the percentage of CAWI usage increased from 44.5% in 2019 to 45.7% in 2021. The percentage of questionnaires filled in by a household member also rose from 57.9% to 60.5%, while the percentage of households not needing any assistance increased from 50% to 52%. Therefore, the analysis of respondents' feedback seems to confirm that the usability improvements have succeeded in simplifying the completion of the questionnaire, especially for elderly users.

In 2021, a new question asking for the reasons of the requests for help has been introduced. Figures show that about one household out of four reported experiencing usability difficulties such as navigation or submission problems. Among elderly households, this percentage increases to 28.2%, showing that this is a subpopulation still needing attention as for the questionnaire design.

6. A classifications of households who need assistance in filling out the questionnaire

In order to better identify households experiencing difficulties, it has been necessary to classify households according to the degree of difficulty they encountered in completing the questionnaire. Based on the respondents' feedback, a simple way to assess the level of difficulty is to calculate the number of different types of assistance required by respondents.

[2] Unweighted raw data have been used.

Some of the indicators calculated for the respondents have been used in a tree regression model to classify households according to their requirement for assistance in completing the questionnaire, whether it be from friends or relatives, the help desk, the municipal census offices, or others. The regression tree allows for finding a relationship between a quantitative variable and a set of independent variables. The dependent variable was the number of different types of help requested, while the independent variables were household socio-demographic indicators: geographical distribution of residence, level of education (represented by the highest educational title within the household), number of household members, presence of elderly members (determined by an indicator that distinguishes households where all members are 65 years or older from those with at least one member below 65), employment status (households with at least one employed person versus those without), citizenship (households with all foreigners versus households with at least one Italian citizen), and ownership or tenancy of the dwelling.

A regression tree is an iterative binary recursive partitioning method that divides data into groups, which are then divided into smaller groups (Breiman, 1984).

The tree regression model has been applied to census data[3] from 2019 and 2021. The results of the classification method are interesting, and a very high goodness-of-fit of the model is observed, as evidenced by a high chi-square value.

According to the model, in 2019 nearly 22% of the respondents needed help to fill out the questionnaire, and the first variable that characterized the analysis was the level of education (Figure 1). The only variable that did not help explain the model was the one related to rental or ownership of the dwelling in which the family lives.

The propensity to require assistance decreases as the level of education increases. For 11 percent of households with no or very low levels of education (nodes 3 and 4) the percentage requesting help is more than 63%. Request for assistance drops below 15 percent for the most educated households, which account for three-quarters of all households (nodes 1 and 2).

The next partition is determined by employment status: the request for assistance is higher when there is no employed member in the household. In node 13, characterized by households with a primary education level and no employed member (9% of all households), the demand for help is 64.2%, compared to 7.3% observed in node 6 represented by graduated households with at least one employed member (25% of all households).

The absence of foreign members in the household also contributes to a further decrease in the need for help. In nodes 16 and 20, representing 63% of households, the absence of foreign members reduces the need for assistance to 7% when there is

[3] Unweighted raw data have been used.

at least one employed and graduated member in the household, and to 11.2% when there is at least one member under 65 with an upper secondary education.

Finally, the geographic breakdown of residence is another variable that explains a high propensity to request assistance in filling out the questionnaire. In node 28, represented by the households of low-educated elderly people residing in Southern Italy or in the Islands, the percentage of requesting assistance is close to 70%.

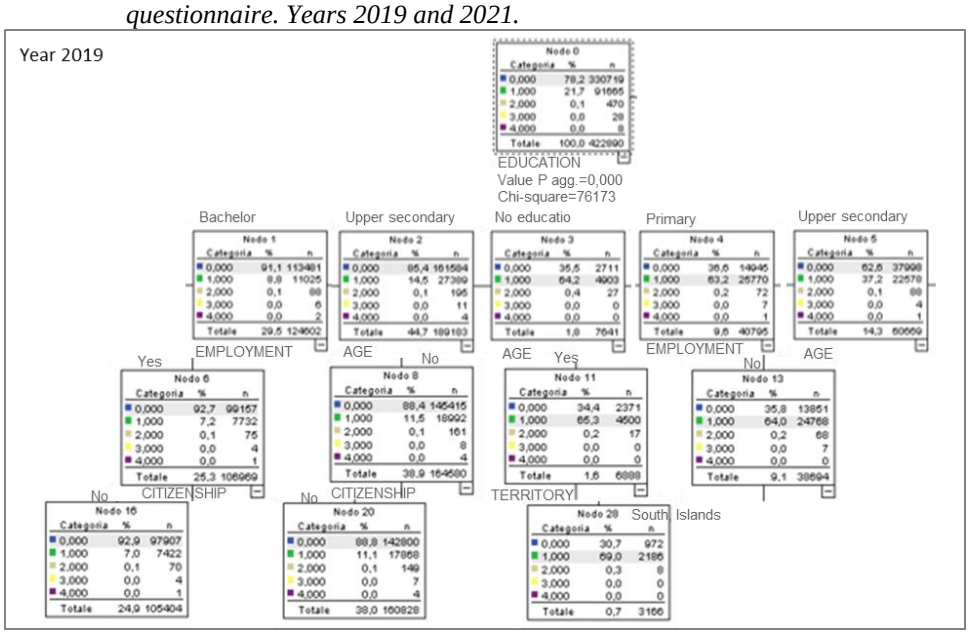
The model applied in 2021 also shows a good fit to the data (Figure 1). The percentage of people demanding assistance is 20%, lower than in 2019. The level of education continues to be the most representative variable, but the geographical distribution is no longer significant. As in 2019, being a homeowner or a renter does not contribute to the model.

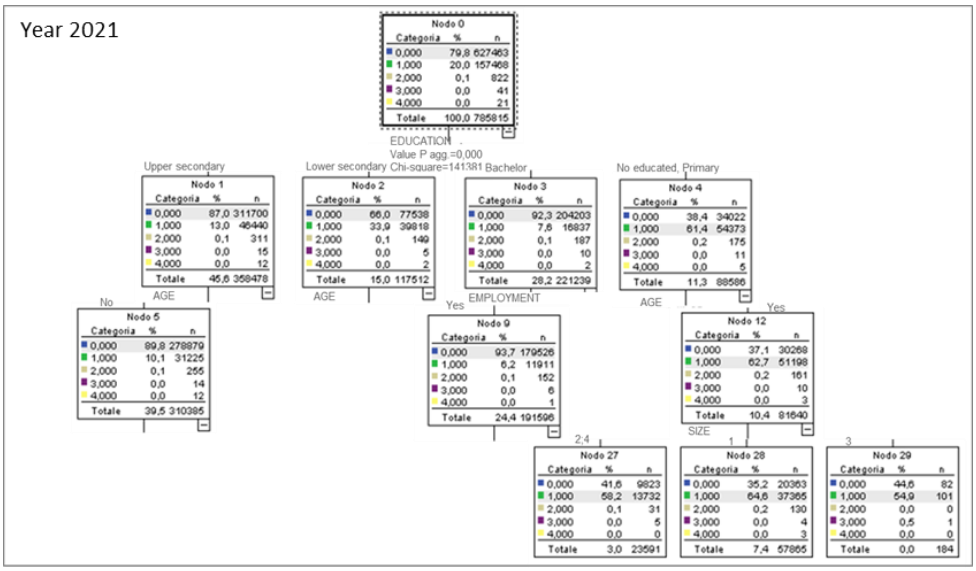
More than 45% of households are represented by node 1, which comprises respondents who have reached a high school level of education (upper secondary education), for whom the need for assistance is only 13%. The need for assistance drops to 7.6% at node 3 (28.2% of households), where there is at least one member who has attained a bachelor's degree. The highest rates of assistance continue to be observed among those with the lowest levels of education: in 2021, the absence of education or primary education (node 4), is represented by 11.3% of households, who need assistance in 61.4% of cases (percentage still lower than in 2019).

In 2021, the next partition is determined by age: the demand for assistance increases with age. Having an upper secondary education, accompanied by the absence of elderly members, reduces the demand for assistance to 10%, as observed at node 5, which includes about 40% of households. Conversely, the request for assistance reaches almost 63% when the lack of education is combined with the absence of members younger than 65 (node 12 represented by 10.4% of households).

The presence of employed members also contributes to a reduction in the demand for help: at node 9, represented by a quarter of the responding households, the percentage of assistance decreases to 6.3% for households with employed graduates. If focusing on the respondents facing the greatest difficulty in filling out the questionnaire, the most vulnerable group are uneducated elderly individuals living alone (7.4%), for whom the need for assistance reaches 65% (node 28).

Figure 1 – *Regression Trees of the households requests for assistance in completing the questionnaire. Years 2019 and 2021.*





Elaborations from unweighted raw census data.

7. Discussion

Completing an electronic questionnaire is a process that entails different skills and abilities. It is essentially a cognitive task, in which respondents find themselves alone facing an object they are unfamiliar with. It is a task that demands considerable cognitive effort, wherein both linguistic proficiency (knowledge and comprehension of the language used) and problem-solving skills come into play.

Hence, it is not surprising that education plays a pivotal role in determining respondents' capability to perform the task independently. Moreover, this phenomenon is not binary in nature. Respondents' autonomy seems to be associated with the amount of education they have received. When the level of education is lacking or at a minimal level, the reliance on external help is four times higher compared to more educated households.

The impact of education on questionnaire completion capability is reinforced by the presence of employed individuals within the family. Although this association may be attributed to the increased employability of educated individuals, resulting in a more favourable employment status of highly educated households, it may also be due, to some extent, to a greater problem-solving ability possessed by those who are employed. In fact, whatever the occupation, it undoubtedly requires a constant effort to analyze situations, identify problems, and seek and apply solutions. The parallels with the demands of an electronic questionnaire are strong.

However, human capital alone is not everything when the problem that needs to be solved is completing an electronic questionnaire. Social support also matters. The analysis reveals that the need for assistance intensifies among the most vulnerable segment of the population, namely uneducated elderly individuals living alone, who constitute the 7.4% of the population. Nonetheless, with appropriate support, these respondents still manage to complete and submit the questionnaire, demonstrating the significance of social capital. The good news is that in 2021 the territorial gap has reduced, probably as a consequence of the digitalisation process of the elderly during the pandemic. Thus, residing in the southern regions or on islands no longer seems to be a factor that worsens the autonomy of this target group.

As questionnaire designers, our focus must be on the needs of respondents, particularly those with limited resources. The comparison between 2019 and 2021 demonstrates that thoughtful actions yield prompt results. The crucial point is to find solutions that will allow the questionnaire to replicate as closely as possible the supportive behaviour provided to the respondents. To achieve this goal, we need to employ targeted methods, specifically designed for this population segment. These could entail a combination of usability testing and cognitive interviewing. Ultimately, this approach will enable us to offer respondents a stress- and frustration-free experience, which should serve as the guiding principle of our work.

References

BARCHERINI S., BONTEMPI K., FAZZI G., LIANI S., LORÈ B., PIETROPAOLI S., ROSATI S. 2022. The respondent as the focus of the questionnaire design. In *Proceedings of the UNECE Expert Meeting on Statistical Data Collection*, Rome.

BREIMAN L., FRIEDMAN J., OLSHEN R., STONE C. 1984. *Classification and regression trees*. CRC press.

GABBIADINI A., MARI S., VOLPATO C. 2011. Internet Come Strumento di Ricerca: Linee Guida per la Creazione di Web Survey, *Psicologia Sociale*, Vol. 2.

ISTAT. 2021. *La conduzione della raccolta dei dati del Censimento permanente della popolazione e delle abitazioni 2019.* Roma: Istituto Nazionale di Statistica.

NIELSEN J. 1993. *Usability engineering*. Burlington: Morgan Kaufmann.

NIELSEN J. 2000. *Design web usability: the practice of simplicity*. Indianapolis: New Riders Publishing.

POLILLO R. 2006. *Plasmare il web: road map per siti di qualità*. Milano: Apogeo.

Sabrina BARCHERINI, Istat, sabrina.barcherini@istat.it
Katia BONTEMPI, Istat, katia.bontempi@istat.it
Manuela BUSSOLA, Istat, bussola@istat.it
Barbara Maria Rosa LORÈ, Istat, lore@istat.it
Simona ROSATI, Istat, srosati@istat.it