The use of a multi-mode data collection system and of a web management system for the Italian Population Census: lessons learnt and future challenges (working paper)

Novella Cecconi[1], Donatella Zindato[2]

Draft

Abstract

The paper discusses the impact of the use of a mixed-mode data collection system on the organization of the Italian Population Census fieldwork, on dissemination timeliness and on data quality.

The multi-mode data collection system was first introduced on the occasion of the 2011 census in order to increase the spontaneous response rate. Mailed out paper questionnaires were collected by a multi-mode system, with web collection being introduced for the first time within a concurrent design. The need to manage information coming from different sources (on-line questionnaires, Post Offices monitoring system, MCCs and the information required to address the households' list under coverage) required a web IT system constantly updated, enabling census staff to follow the status of every questionnaire over time and allowing directing enumerators only to non-responding households and potential under coverage addresses.

Thanks to the use of questionnaires mail out and of a multimode data collection system, the front-office staff was dramatically reduced in comparison to past censuses (about 40%) and a greater flexibility was allowed to respondents. Indeed, the so-called 'spontaneous return' went very well and web return was much higher than expected. On the contrary, the logistics of the mail out/mail back process proved very complex and with many points of failure, calling for a totally paperless census, implemented in the combined design of the Permanent Population and Housing Census (PPHC).

The evolution of the multi-mode data collection system and of the web management system are described from its first use for the 2011 census, through the first and second cycle of the PPHC. Critical issues and lessons learnt will be discussed, on which the continuous planning of the PPHC data collection is based.

1. The use of a mixed-mode data collection system in the 2011 Population Census

The mixed-mode data collection system was first introduced at Istat for the 2011 Population Census in order to reduce municipalities' workload by increasing self-response rate. The planning of 2011 census had been based on the need to improve dissemination timeliness by at the same time reducing the statistical burden on respondents, enforcing Municipalities Census Offices (that in Italy are in charge of the census fieldwork) and containing costs.

The once "door-to-door" census thus became a *register-supported* census, implemented by means of questionnaires' *mail out* to the households enrolled in the municipal population registers. Self-completed questionnaires were then collected by a mixed-mode data collection system allowing households to choose the way in which they preferred to complete and return the questionnaire, with web collection being introduced for the first time within a concurrent design, together with the possibility to return the paper questionnaire at any post office in Italy or to Municipal Collection Centres. A further mode i.e. the targeted recovery of non-response by enumerators was foreseen in the second phase of the enumeration, as concomitant to the 'spontaneous' ones which were still available until the end of the enumeration.

Thanks to the use of questionnaires *mail out* (instead of enumerators' delivery) and of a multimode data collection system (where enumerators were just one of the possible return modes, and hierarchically the last one), the front office staff was dramatically reduced in comparison to past censuses (about 40%) and a greater flexibility was allowed to respondents. Indeed, the self-response or so-called 'spontaneous return'

---

[1] Istat, Data Collection Department
[2] Istat, Population and Housing Census Division

went very well and web return was much higher than expected, accounting for one third of the total completed questionnaires. Indeed, although the communication campaign did not place particular emphasis on the web return mode (Istat's "preferred" return mode for obvious reasons of timeliness and data quality), and notwithstanding Italy being at the time the 4th country in Europe by number of people who had never had access to the Internet (37,2% of the population against the average 22,4%), around 8.5 million households chose to complete the online questionnaires. CAWI was therefore the most used channel (Figure 1), followed by the delivery of the questionnaire to the Municipal Collection Centres (31.7 percent). The percentage of households who chose to deliver the questionnaire to Post Offices was also significant (22.6 per cent), while only 12.3 per cent of the questionnaires were collected door to door by enumerators.

Figure 1 - Mixed mode data collection at the 2011 Population Census (percentage values)
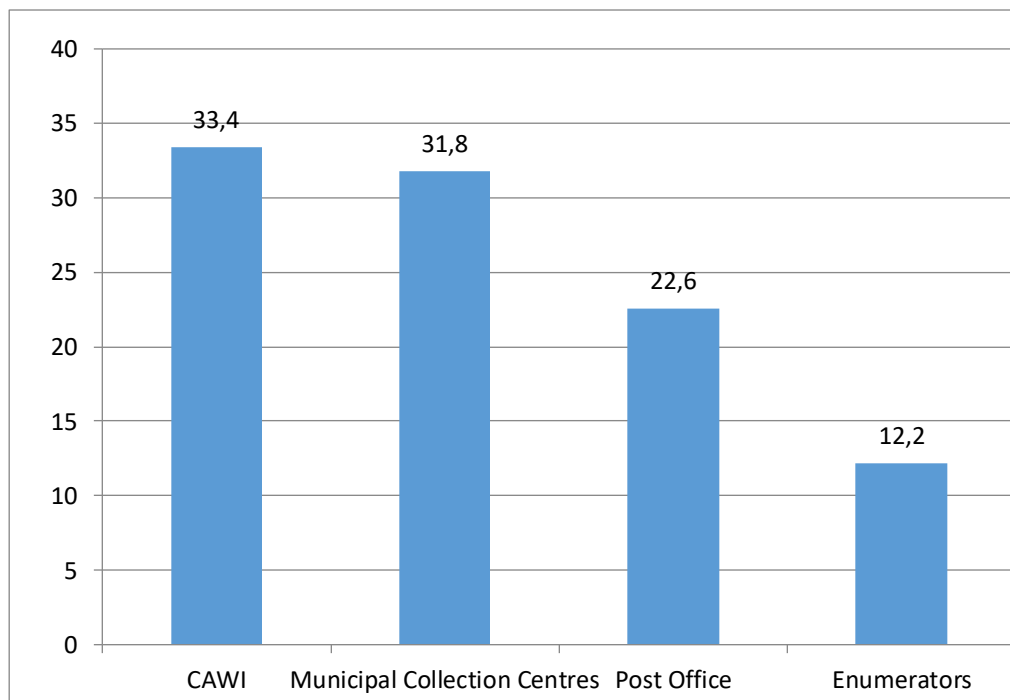
If the decision to include the CAWI mode among the possible return options proved to be successful, more generally the flexibility allowed by the adoption of a multi-mode data collection system seemed to be the winning choice. In fact, the percentage of self-response altogether was very high, equal to 82.7 per cent[3].

Quite different trends over time could be observed by mode. The absolute peak of CAWI occurred in the first week, while another peak was recorded in the last week of the spontaneous return phase, when the peak for the MCCs was also recorded. A more regular trend over time can be observed for questionnaires recovered by enumerators, all among the second phase until the end of December, to become predominant among the modes used in the first few months of 2012, when (as envisaged by the census plan) the enumeration was still ongoing only in the largest Municipalities.

More generally, the following elements emerged by the analysis of the time trends by mode:
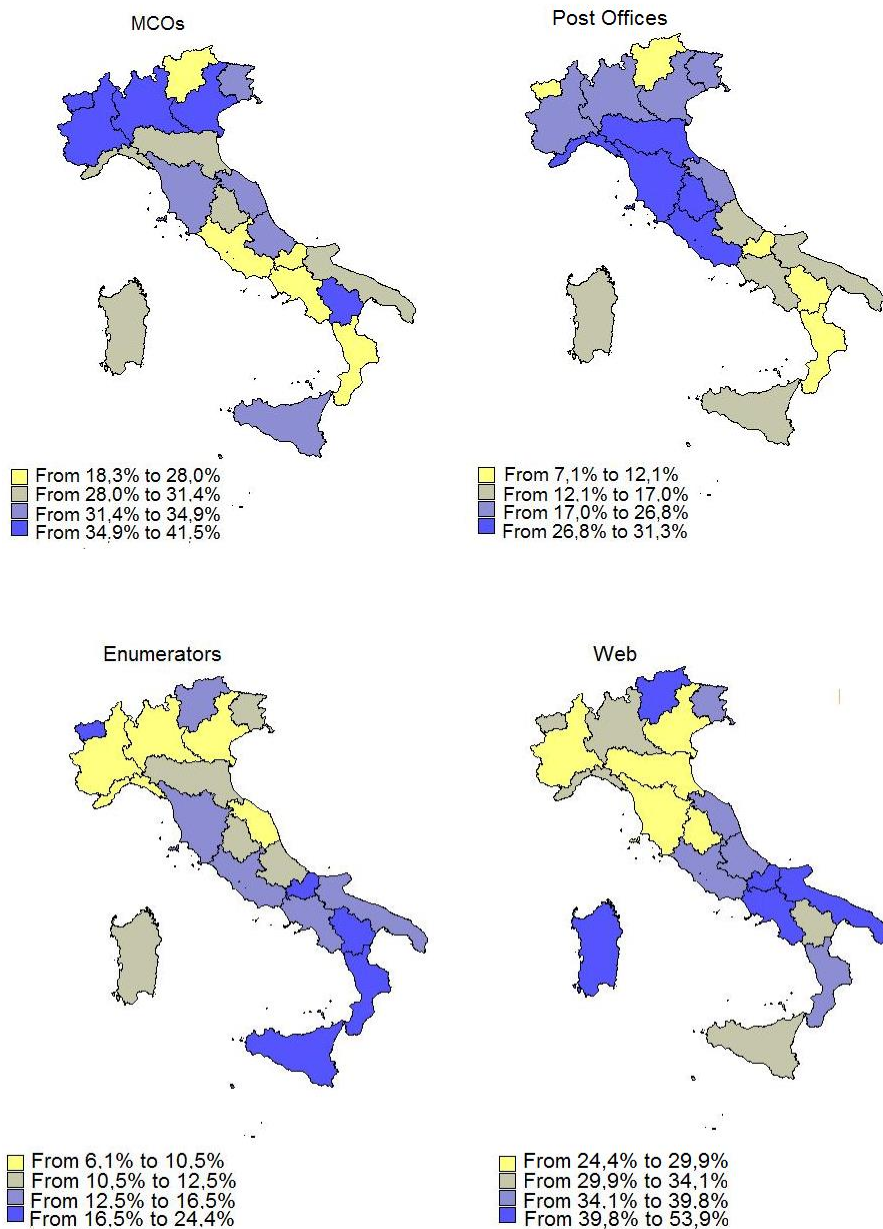- the Cawi was mainly used at the beginning. Those who chose to reply via the internet did it in the aftermath of the arrival of the letter from Istat;
- the delivery of questionnaires at MCCs followed a similar time trend to that of CAWI return, therefore it appeared as the "spontaneous" alternative for those who did not have the possibility to use CAWI (i.e.

---

[3] The approximately 4.2 million questionnaires (equal to 17.3 per cent of the total returned questionnaires) counted as "not returned spontaneously" included: questionnaires delivered by enumerators (i.e. questionnaires which, for different reasons, was not possible to mail out); questionnaires for which at least one attempt of contact by the enumerator was recorded in the monitoring system; questionnaires actually collected by the enumerators.

didn't have an internet connection or were not familiar enough with internet use as to choose to fill in the electronic questionnaire);

- the analysis of the trend of the questionnaires collected by enumerators showed that many Municipalities has chosen to send them to the field before the official start of the non-response follow-up phase.

Figure 2 - Share of questionnaires by return mode (Municipal Collection Centres, Post Offices, Enumerators, CAWI)

MCOs

Post Offices

From 18,3% to 28,0%
From 28,0% to 31,4%
From 31,4% to 34,9%
From 34,9% to 41,5%

From 7,1% to 12,1%
From 12,1% to 17,0%
From 17,0% to 26,8%
From 26,8% to 31,3%

Enumerators

Web

From 6,1% to 10,5%
From 10,5% to 12,5%
From 12,5% to 16,5%
From 16,5% to 24,4%

From 24,4% to 29,9%
From 29,9% to 34,1%
From 34,1% to 39,8%
From 39,8% to 53,9%

As expected on the basis of the data concerning households ICT access/usage in the different areas of the country, a great regional variability could be observed in the breakdown by return mode (see figure 2). In general, households living in the South and in the Centre of Italy favoured using the web to complete the questionnaire (chosen, respectively, by 38.3 per cent and 32.7 per cent of all households), while in the regions of the North, the preferred mode was the delivery at the MCCs (35.7 percent). More precisely, the percentage

of CAWI ranged from the 53.9 per cent in Trentino-Alto Adige[4] to the 24.5 per cent in Tuscany[5]. Vice versa, Tuscany and Trentino-Alto Adige were the regions with, respectively, the highest (31.3 per cent) and the lowest (7.1 per cent) percentage of questionnaires delivered to Post Offices. Valle d'Aosta was the region with the highest percentage of questionnaires delivered to the MCCs and Lazio the one with the lowest percentage (18.3 per cent). Finally, Calabria was the region with the highest share of questionnaires collected by the enumerators (24.4 per cent) and Lombardy the one with the lowest share (6.1 per cent).

A significant role was also played by the municipality size (see Figure 3). The web was the preferred mode by households living in small municipalities (37 per cent), especially in the South (46 per cent), while households living in medium-sized and small-medium municipalities mostly opted for returning the questionnaire to Municipal Collection Centres (respectively, 40.2 per cent of households living in municipalities between 5,000 and 20,000 inhabitants and 35.1 per cent of households living in municipalities between 20,000 and 50,000 inhabitants), with higher percentages in northern municipalities (respectively 44, 5 percent and 37.9 percent). Finally, in the big cities, the delivery of the questionnaire to Post Offices was by far the preferred option (41.6 per cent of households living in municipalities with at least 100,000 inhabitants), probably due to their widespread territorial distribution which made them the most "sustainable" mode in the largest municipalities.

However, the remarkable differences in the territorial distribution by return mode have also to be explained by taking into account the strong influence of the field-work organization put in place by the different Municipal Census Offices (MCOs), since they had a large autonomy in promoting one or the other of the return modes, as completed questionnaires were payed differently according to the return mode[6].

As far as the "choice" of the return mode is concerned, the difference between Italian and foreign households also appears to be relevant (with less than 24 per cent of households whose members are all foreigners completing the questionnaire on the web, while the share of questionnaires collected by enumerators is much higher than that recorded for Italian households - 28.7 per cent versus 11.1 per cent - most likely precisely because of linguistic difficulties)[7].

Analysing the breakdown by return mode and presence of elderly household members, a higher use of MCCs can be noted by multi-person households whose members are all elderly (36,3 per cent) and a lesser use of the 'enumerator mode' (8 per cent).

Despite the higher propensity of youngsters to use the internet, a share of CAWI returns just 3 percent higher than the average is recorded for households with at least one member aged below 30, probably because completing the census questionnaire is considered a task reserved to adults.

Instead, the presence of at least one high educational qualification in the household seems to have an impact on the choice of the CAWI mode (chosen in the 39 per cent of the cases). The difference is even more relevant in the North (37.1 per cent against 28.9 per cent), where the overall percentage of households who chose the CAWI mode was lower than the national average.

---

[4] In this region a different enumeration strategy was adopted, i.e. paper questionnaires were only available on request to the enumerator.

[5] The regions where Cawi share was higher than the national average were, in order, Sardinia (44.9 per cent), Molise (41.3 per cent), Puglia (40.6 per cent), Campania (40.5 percent), Calabria (39.9 percent), Friuli-Venezia Giulia (39.2 percent), Abruzzo (38.4 percent), Lazio (38.1 percent), Marche (35.7 percent) and Basilicata (34.2 percent).

[6] Questionnaires were paid differently according to the return mode: questionnaires returned to MCCs or to enumerators were paid 6 euros while internet questionnaires were paid 4 euros. In order to encourage promoting the CAWI mode by MCOs, Istat would pay 5,50 euros instead of 4 per each questionnaire returned via the internet if the overall CAWI percentage in the municipality would be above the 25% of the completed questionnaires. Therefore, in some small municipalities respondents who chose to return the questionnaire to the MCC were invited to fill in the online questionnaire, with the assistance of MCCs operators.

[7] Despite specific actions targeting foreign citizens usually resident in Italy (Gallo et al., 2014)and the availability of translations of facsimiles of the questionnaire in 17 foreign languages, the insufficient linguistic mastery represented an obstacle to the autonomous compilation (and therefore to the spontaneous return) of the questionnaire. In fact, for logistical reasons, the electronic questionnaire could only be completed in Italian (or in German or Slovenian, languages of the linguistic minorities protected by the law).

These preliminary evidence was confirmed also by the analysis at the micro level, carried out to identify the household profiles with a greater propensity for web responses[8].

Figure 3 - Breakdown by return mode and municipality size
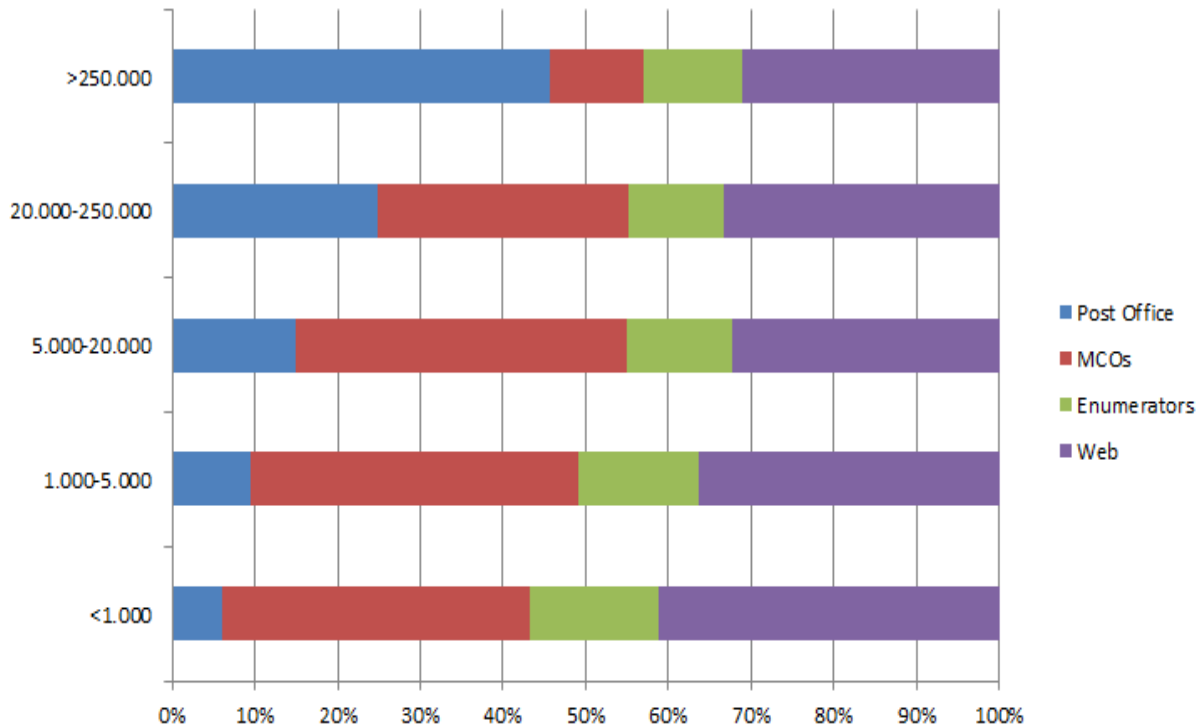
2. Response management in the 2011 census

The need to manage information coming from different sources (on-line questionnaires, Post Offices monitoring system, MCCs and the information required to address the households' list under coverage) required a web IT system (SGR, according to the Italian acronym) constantly updated, enabling census staff to follow the status of every questionnaire over time and allowing directing enumerators only to non-responding households and potential under coverage addresses.

Indeed, the management of a modular and flexible strategy represented a big challenge. On one hand it helped solving problems that traditionally had a great impact on census process, negatively affecting the timeliness of data dissemination. On the other hand, the introduction of this new strategy implied a higher level of complexity and a multiplication of risk factors[9].

Such a web management system was in fact crucial to the performing and success of the entire census, being a complete instrument that guided and supported census operators during all the survey phases. It was designed to provide the different users of the system with: (i) up-to-date information at different aggregation levels, including single questionnaire level; (ii) a tool for cooperative working, guided through a forced workflow of questionnaire life-cycle.

Accessible online to all of the different levels of census staff, it enabled the status of every individual questionnaire to be followed in almost real time, thus allowing the targeted recovery of missing questionnaires. The availability of constantly updated information on the status of each questionnaire

---

[8] See Zindato, 2017.
[9] See Benassi et. al., 2013.

enabled enumerators to be directed only to households to which the questionnaire had been sent but not yet returned.

Furthermore, it was designed to automate back-office work and to guarantee flexibility to fieldwork organization within each Municipal Census Office. Municipal Census Offices managers had to assign an organisational role and a system profile to every user and allocate enumeration areas to enumerators. Each census office could thus freely decide how to distribute work in terms of assignment of enumeration areas to enumerators and back office work to operators. A hierarchical organisation could also be defined by setting dependency relationships between staff with a coordinator role and other staff and of enumerators to co-ordinators.

Finally, being as well a monitoring system, SGR also allowed to produce census progress reports. The census web based management and monitoring system was part of a general strategy aiming at minimising errors, reducing organizational workload and holding down costs.

3. Towards a totally paperless census: the 15th Population Census Pilot Survey

As already mentioned, the need of reducing municipalities' workload and the burden on respondents had called for the use of new data collection techniques and new territorial instruments meant to improve coverage and quality of the enumeration. However, the innovations designed for 2011 Italian census were not sufficient to achieve a stable and enduring balance between census costs and benefits. In fact, costs remained high and too concentrated in time, while the use of administrative data was not up to the potential offered by the Italian context. Moreover, the supply of highly detailed geographic data remained only decennial so census data continued becoming quickly outdated.

Furthermore, the logistics of the *mail out/mail back* process (managed in outsourcing) proved very complex and with many points of failure, calling for a *totally paperless* census. Among the main problems have to be mentioned the long times necessary for moving around and the huge spaces needed for the storage of tens of millions of paper questionnaires, the high number of addresses from the municipal population registers which had not been recognised, thus making it necessary to deliver about 2 million questionnaires via enumerators. Furthermore, the monitoring system was not updated in real time on the questionnaires delivered to the Post Offices (which amounted to almost a quarter of the total completed questionnaires); as a result, the management of the fieldwork was problematic, as enumerators would be sent to collect questionnaires already returned.

These critical issues added up to the ones above mentioned concerning the ever least sustainability of the traditional census (intended as universal and simultaneous field-enumeration). For all of these reasons, the development of a completely different approach seemed necessary, based on a sequence of operations and surveys designed ad hoc in such a way to build a complete information system producing specified census output results at given times. The new census strategy would be "rolling" (later it would become the Permanent Census) and join a greater use of administrative sources with sample surveys rotating through a multi-year period of time, so to avoid big "one shot" activities and sunk costs. More precisely, two ad hoc sample surveys would be conducted annually:
- a C-sample short-form only survey designed as a statistical test on the entity of the coverage error of municipal population registers to determine the usual resident population
- a D-sample rolling survey designed to collect on the field information for variables non replaceable by administrative data, in order to produce the hypercube required by EU Regulation on Population and Housing Censuses.

A Pilot Survey has been conducted in 2015 in order to test the data collection techniques and the fieldwork organization to be adopted in the new Permanent Census. As to the data collection techniques, the main objective of the Pilot Survey was to test the sustainability of a *totally paperless* enumeration, i.e. a mixed-mode data collection strategy based only on the use of electronic questionnaires.

The C-sample was based on a capture-recapture methodology with the first capture being represented by the population registers and the second capture being conducted as a door-to-door enumeration with CAPI technique, while the D-sample would be conducted with a mixed-mode technique. More precisely, the C-sample survey was conducted by interviewers with hand-held devices such as tablet and laptop computers, which were used both for accessing enumeration areas maps and address lists, and as a means of data capture. On the other hand, the D-sample survey was conducted via a multimode data collection system where several paperless returning modes were concurrently offered to respondents: self-response by CAWI, the possibility of contacting a toll-free number and completing the questionnaire via telephone interview or to go to Municipal Survey Centres and fill out the questionnaire by using an internet station or ask for a CAPI interview. So, CAWI was again used within a concurrent design, in order to avoid the use of paper questionnaires (related to many critical issues of the 2011 strategy) and at the same time provide respondents with a choice of spontaneous return modes which would minimize recourse to enumerators (thus minimizing costs and municipalities'' workload). Differently from the 2011 census, no paper questionnaire would be mailed out to households, which would receive only a letter with the questionnaire login details.

The main mode innovation with respect to the 2011 mixed-mode system concerned the possibility for the households to call the contact centre not only for information and for clarification but also to ask to be interviewed (so-called inbound CATI[10]). During the first phase, the households could complete the questionnaires by themselves, or go the Municipal Collection Centres or call the contact centre. During the second phase, the non-response follow up would be conducted by phone (as far as possible[11]) or in the field by enumerators Four different combinations of modes were tested, and some of these included also the non-response follow-up by enumerators, who would carry out the interview on the field by using a portable device (tablet or laptop.

This entailed the need for the Municipal Census Offices to guarantee assistance to the respondents at the MCCs for the entire duration of the survey and the need for the coordinators/enumerators/back office operators to be constantly updated through the monitoring system. It was therefore necessary to implement an integrated data acquisition and production process management system, within a Bring Your Own Device strategy. The management system was in fact the evolution of the management system developed for the 2011 Census.

The D-sample Pilot Survey was conducted on a sample of 148 municipalities including those where at the 2011 Census the lowest percentage of total self-response and of CAWI return had been recorded. As to the results, they were partially different from the 2011 ones. The CAWI was still the preferred mode but to a much higher extent accounting for almost half of the completed questionnaires (49.2 per cent). Of the total CAWI questionnaires, almost the 75% were self-completed by the households without any assistance, while the 17.7 per cent had requested the help of relatives or friends, and the 7.7 per cent of the Municipal Collection Centres.

The breakdown by mode was only partially comparable, as the Post Office and the inbound CATI mode were respectively present only in 2011 and in 2015 (see figure 4). However, a remarkable increase of the CAWI percentage can be noted (even when considering the increase in the use of the Internet by the households occurred since the Census), to be considered positively, especially considering the absence of the massive communication campaign put in place for the Census[12]. On the other hand, the percentage of questionnaires completed via the MCCs had decreased (though, given the fact that it was a sample survey, it was difficult for the municipalities to organize widespread centres on the territory). Finally, the share of questionnaires collected by enumerators had slightly grown (13.5 per cent versus 11.8 per cent).

---

[10] The term CATI is used improperly, for the sake of brevity, with reference to the telephone interviews carried out by operators of the contact centre managed by Istat or by the MCCs operators of a few municipalities involved in the Pilot Survey. In fact, it was not implemented a specific CATI acquisition system but, in order to maximize the spontaneous return rate, operators were trained to carry out telephone interviews using the application developed for CAWI.

[11] Only about the 20 per cent of landline phone numbers were available at Istat.

[12] For the Pilot Survey only local communication events were foreseen, to be organized by the municipalities.

As for the new return mode (the inbound CATI), it was used to complete the 10 per cent of the total completed questionnaires. Focus groups were organized among the Istat personnel who had been providing the contact centre service, from which useful indications emerged both on the use of the toll-free number as a possible return mode and on the usability of the electronic questionnaire. In particular, the toll-free number seemed to be a viable return mode especially for the elderly and for those who were not familiar with new technologies (as it was confirmed also by the analysis performed on the characteristics of the respondents, which will be briefly reported further down in the text).

Figure 4 – Breakdown by return mode at 2011 Census, 2015 Pilot Survey, and at 2018 and 2022 PPHC waves
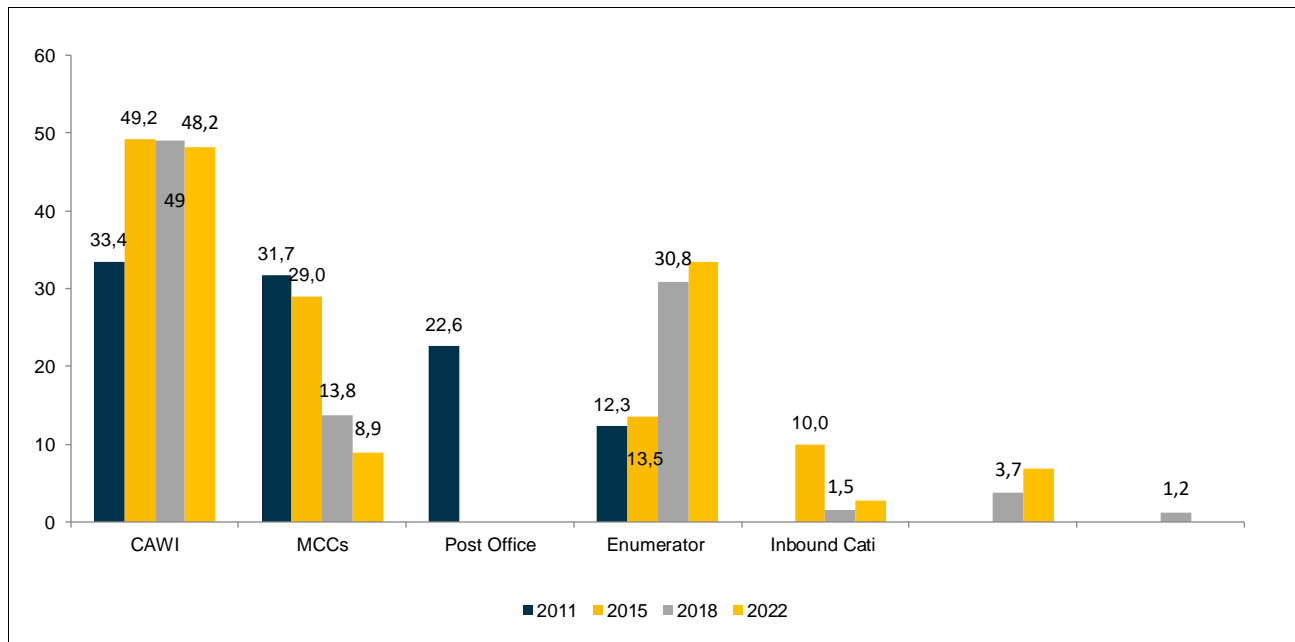
The territorial profile of the web respondents appeared significantly different from that observed in 2011. In fact, differently from the 2011 Census, during the Pilot Survey the greatest use of the CAWI mode was recorded in the municipalities of the North, and especially of the North-East (which chose CAWI in the 32.6 percent of cases, compared to a national average of 23.1 percent; while in the Island the percentage of households who used CAWI was just 17 percent). In general, the highest response rates via the web were recorded in the regions where the use of new technologies was more widespread[13], even though the differences in the total response rate by municipality as well as the strategy (combination of modes) assigned to each municipality should be taken into account when evaluating the results.

Through a multivariate analysis (analysis of multiple correspondences and cluster analysis), it was possible to characterize the households who had used the toll-free number to fill in the questionnaire. The micro-level analysis of the characteristics of the households who had chosen the CAWI option (instead of the CATI interview or of the MCC) confirmed to some extent the results of the 2011 analyses, especially with regard to education as a crucial factor for the internet choice. More generally, households showing a higher propensity to the use of the web mode for completing the questionnaire were those with at least 3 members, at least an elder, at least a member with a medium/high degree of education and a reference person normally using the web.

In summary, the results of the Pilot Survey did confirm the feasibility of the paperless choice and at the same time the need of a mixed-mode system offering a variety of possible return modes in order to reach different households profiles. In fact, the share of households choosing the CAWI option (both in an independent and supported way), albeit growing, did not allow foreseeing the CAWI as the sole return mode. Inter alia, the

---

[13] See Eurostat, 2006.

so-called inbound CATI i.e. the mode allowing households to call the toll free number and be interviewed, appeared as a viable alternative (if financially sustainable)[14] for a variegated universe of households (mainly socio-economically disadvantaged households but also well-off families made up of both elderly and young people).

4. From the 'door to door' field enumeration to the Permanent Population Census

After the first Pilot Survey conducted in 2015 and a second one conducted in 2017, the rolling strategy of the new Population census was further tuned, in accordance with Istat modernization strategy, which places the integrated system of statistical registers at the core of statistical production. The role of field surveys in this system is to support registers, in the broad sense of assessing their quality and to add information that is missing, incomplete or of insufficient quality. The Population Census thus became the Permanent Population and Housing Census (PPHC), at the core of which is the Statistical Population Register (RBI according to the Italian acronym), whose main sources are the local population registers of Italian municipalities (administrative local registers). Together with the Register of Addresses and with the thematic registers on education and employment, RBI provides the basis to produce population census results while two sample surveys (Area survey and List survey) are conducted annually in self-representative municipalities and every 4 years, according to a rotation scheme, in smaller ones, to evaluate and correct the coverage errors of RBI and to collect data for variables not available (or only partially available) from the registers[15].

The Area survey, still being conducted as a door-to-door enumeration with CAPI technique as the previous C-sample, is currently undergoing a redesign after the 2020 move to a fully register-based count estimation. No more used to measure and correct the Population register, it will be aimed at providing a measure of the error of such estimation (while at the beginning of the first cycle of the PPHC it was used within a capture-recapture model for direct estimates of the coverage errors of RBI). The redesign will concern both the survey methodology and sampling frame and the data collection techniques. The List survey, which represents the evolution of the D-sample survey, is based on a sample of households drawn from the population register and is conducted through a mixed-mode data collection system. The mixed mode system includes in the first phase (spontaneous response-only) self-response by CAWI, the possibility to go to the MCCs either to use an internet station or to be interviewed (CAPI) and the inbound CATI mode. The second phase includes the sme return options plus the non-response follow up by enumerators. The CAWI mode has seen a slight increase during the first cycle of the PPHC (2018-2021).

The management of such a complex and diverse enumeration strategy entails the need of a very flexible web management system, guiding and supporting census operators during all the survey phases, which has been developed as an evolution of the management system developed for the 2011 census, and generalized to become the management system of Istat social surveys.

Even though the breakdown by mode is only partially comparable with the past mixed-modes systems, we can observe that the share of households choosing to use the CAWI option is still close to half of the total responding households (see figure . We also note a further decreases the percentage of households going to the MCCs either for receiving assistance to complete the electronic questionnaire/use the MCC internet station or to receive a CAPI interview.

In Figure 4 the distribution by return mode at the 2011 census is compared with the ones at the 2015 Pilot Survey and at the first years of respectively the first (2018) and the second (2022) cycle. Altogether, if the CAWI share is more or less stable, is the self-response (i.e. independent response not needing a field follow-up) that shows a setback, therefore requiring a stronger effort of front-office fieldwork in order to keep the required high response rate (the share of questionnaires completed by CAPI interview performed by enumerators has therefore raised from the 12.3 per cent of the 2011 census to the 33,4 per cent of the 2022

---

[14] Given the experimental purpose of the survey, the service was provided internally to Istat, by non-professional operators and only during working hours. Therefore its financial sustainability should be assessed in the case of a professional service, contracted in outsourcing for longer hours (as it should likely be in the case of the actual census).

[15] A detailed description of the surveys is provided in Falorsi (2017).

List survey). Furthermore, if we look at the CAWI share along the four years series[16] since the start of the PPHC (see figure 5), after slightly increasing during the first 3 years (but not as much as it could be expected in 2021 given the large use of digital technologies imposed by the pandemic), it clearly decreases in 2022. Furthermore, a relevant decrease can be observed as to the share of households going to MCCs.

Figure 5 - Breakdown by return mode and PPHC wave

RETURN MODE	2018	2019	2021	2022
CAWI	49,0	50,0	51,9	48,2
MCCs	13,8	13,0	12,1	8,9
INBOUND CATI[17]	1,5	1,3	2,3	2,8
CATI	3,7	2,8	7,6	6,8
ENUMERATOR	30,8	32,4	26,2	33,4
OTHER[18]	1,2	0,6		
TOTAL	100,0	100,0	100,0	100,0

This trend reversal should be related to a number of factors among which the lack of a massive communication campaign as those traditionally put in place for traditional censuses. Even in the PPHC most resources were invested from a communication point of view in the launch of the new strategy in 2018, with progressively less funds available along the other years of the cycle as the census becomes one of many household sample surveys). Another major difference between the 2011 'traditional' census and the yearly sample survey of the PHHC is related to the scope of the survey, involving only a sample of households in each municipality, thus not justifying for the Municipal Census Offices the organization of the widespread network of Municipal Collection Centres which played a quite significant role in order to reach the high share of autonomous response achieved in 2011. Finally, it should be mentioned the ever-growing respondents' burden and respondents' fatigue, which represent a challenge for official statistics, as keeping a high response rate is crucial to quality.

As to the territorial differences, a distribution by return mode similar to the one observed at the 2015 Pilot Survey has been recorded all along the PPHC 4-year series, with huge differences in the share of CAWI among Italy's regions (see Figure 6). More precisely, the percentage of households having self-completed the questionnaire online ranges from being much above the national average (as high as the 62.5 per cent registered in Lombardy in 2022) to reaching a low of half the national average (with the 24.5% recorded in Calabria in the same year). Furthermore, it is worth noting that the distance between the two extremes is increasing (it was equal respectively to 60.8 per cent and 29.3 per cent in 2018). These data should also be analysed in the light of the total response rate, which, although still quite high, in some regions is decreasing more than in others. These dynamics should be further analysed at the micro level but some studies already performed at the micro level confirm education as having a crucial impact on the choice of the return mode, as long as variables such as the household composition by citizenship or the type of municipality where the household lives (according to the 'inner area' variable)[19].

---

[16] The break in the series is due to the Covid-19 pandemic and the subsequent withdrawal of all household surveys to be held in 2020.
[17] In the PPHC this is not an official return mode i.e. it is not advertised as a possible return option, but in exceptional cases (households unable to access the CAWI and move to reach the MCCs) households are being interviewed on the phone.
[18] 'Other' is referred to a small share of cases when, due to poor functioning of the devices or of the internet connection, enumerators were allowed/obliged to write down the answers on paper and later report them into an electronic questionnaire.
[19] See Bussola M. et alii, 2023.

Figure 6 – Breakdown by region and return mode at the 2018 and 2022 PPHC waves

REGION	2018							2022					
	Spontaneous response				Non-response follow up			Spontaneous response				Non-response	
	CAWI	CAWI at MCCs	CATI at MCCs	INBOUND CATI	CATI	ENUMERATOR	OTHER	CAWI	CAWI at MCCs	CATI at MCCs	INBOUND CATI	CATI	ENUMERATOR
Italia	**49,0**	**2,4**	**11,5**	**1,5**	**3,7**	**30,8**	**1,2**	**48,2**	**0,5**	**8,4**	**2,8**	**6,8**	**33,4**
Piemonte	55,0	4,4	14,6	2,0	4,3	19,1	0,8	53,7	1,0	11,8	3,1	8,5	21,9
Valle d'Aosta	52,7	3,4	9,0	0,7	5,4	28,5	0,3	51,4	1,9	7,2	1,9	8,4	29,2
Lombardia	60,8	2,5	12,6	1,4	3,6	18,3	0,8	62,5	0,5	11,3	3,1	7,8	14,8
Provincia autonoma di Bolzano	53,7	6,4	12,6	0,9	2,4	23,2	0,8	58,0	1,7	8,4	1,2	5,1	25,6
Provincia autonoma di Trento	58,9	1,2	5,0	0,6	1,2	32,8	0,2	53,2	0,5	4,5	1,4	5,4	35,0
Veneto	57,3	0,9	8,1	1,1	2,5	29,4	0,7	56,9	0,4	7,0	2,7	6,5	26,5
Friuli-Venezia Giulia	55,9	1,6	11,8	1,3	2,1	26,8	0,6	56,8	0,5	7,4	2,6	5,5	27,3
Liguria	54,1	2,6	10,0	1,6	3,9	26,5	1,3	51,5	0,6	7,6	3,7	6,7	29,9
Emilia-Romagna	54,2	1,1	9,4	1,7	3,6	27,9	2,1	54,5	0,3	6,1	4,1	8,7	26,3
Toscana	53,4	2,1	8,2	1,6	4,5	26,4	3,8	51,4	0,3	5,4	3,4	8,4	31,1
Umbria	52,1	3,0	12,7	2,3	6,1	22,2	1,6	51,8	0,3	6,6	4,3	10,4	26,5
Marche	51,6	1,8	13,0	1,7	3,5	27,3	1,0	49,7	0,4	9,0	3,8	7,6	29,5
Lazio	52,1	2,1	12,5	2,4	3,7	26,6	0,6	50,6	0,8	8,5	3,7	6,3	30,2
Abruzzo	46,4	1,9	10,6	1,1	3,5	35,7	0,8	44,1	0,8	9,4	1,8	5,6	38,2
Molise	40,5	2,8	11,1	0,7	2,9	41,5	0,4	34,9	0,4	6,1	1,5	5,6	51,4
Campania	34,3	2,1	15,6	1,5	3,8	42,0	0,6	34,0	0,5	9,3	1,7	4,1	50,3
Puglia	42,6	3,7	8,5	0,9	3,5	39,8	1,0	40,6	0,5	6,1	2,0	4,9	45,9
Basilicata	38,3	0,4	10,4	0,8	2,4	47,4	0,4	38,2	1,0	9,9	1,5	6,0	43,4
Calabria	29,3	3,0	12,2	1,2	3,7	49,3	1,3	24,5	0,7	7,9	1,8	5,1	59,9
Sicilia	32,7	3,0	12,7	1,7	4,4	44,6	0,9	28,3	0,2	7,9	2,3	5,9	55,3
Sardegna	41,8	1,3	15,0	1,2	4,1	35,7	0,8	39,3	0,4	9,3	2,2	4,9	43,9

5. Lessons learned

The need for budget, timely and accurate census data along with the changes in technology and in society has guided the transition from the traditional 'door-to-door' census to the PPHC. This transformation reflects the extent to which digital data are changing the routines of production and use of statistics[20].

The sustainability of a totally paperless strategy was tested through the experimental surveys conducted in 2015. Unsurprisingly, if the web spontaneous return rate was growing (accounting for almost half of the total completed questionnaires), it was still far from ensuring a successful enumeration if not complemented by other return modes. A similar mixed-mode system has been implemented for the List survey data collection, within a totally new census strategy based on the integration of administrative and survey data (the PPHC).

From a respondent's point of view, the mixed-mode system is certainly welcome (and more and more expected) as it allows a greater flexibility, but it entails organizational challenges and needs a continuous technical support to the different field-work levels and an accurate training strategy, in order to reduce as much as possible non-sampling errors. In the context of the PPHC, the training organization plays a crucial role and has undergone dramatic changes due both to Istat modernization (and its transformation from the stovepipe to the matrix organizational model) and to the break-in of the digital technologies and to their potential (we refer e.g. to the use of distant learning versus face-to-face training).

The new census strategy allows a significant reduction of census costs, of respondents' burden and of the organizational impact on municipalities (that are responsible for the census fieldwork) but to further achieve these goals, tailored communication strategies need to be put in place in order to raise the CAWI response rate.

The PPHC first cycle started in 2018 and ended in 2022 (with a suspension of the field surveys in 2020 due to

---

[20] See Aragona and Zindato, 2016.

the pandemic). The second cycle, started in 2022, shows still high response rates, but a slight decrease of the web spontaneous return rate, notwithstanding the expectations of further increase due to the 'forced' digitalisation undergone by many during the pandemic period and the boost of remote and smart working.

Analyses performed on the process data show peculiar territorial patterns in the choice of the different return modes and the impact of several individual variables such as education or citizenship of the household members. These findings will be further investigated in order to design adaptive survey strategies and to tackle different population targets with tailored communication campaigns.

If the specific CAWI response rate could benefit of such tailored strategies addressing the respondents' segments potentially most prone to the use of digital technologies, different strategies could be put in place to enhance the overall spontaneous response rate. To this end, the feasibility of adopting the inbound CATI (as tested in the 2015 Pilot Survey) as a viable alternative return mode to the CAWI should be explored, as it could successfully reach some an important share of those left behind by the digital divide.

Finally, the need of designing the questionnaire in order to reduce the mode effect should be taken into account, as it has been designed based on the assumption of the self-administration (as it was, in fact, the case up to the 2011 Census), but due to the changes in the enumeration strategy is being more and more administered by enumerators.

References

Aragona, B., e D. Zindato. 2016. "Counting People in the Data Revolution Era: Challenges and Opportunities for Population Censuses." International Sociological Review 26(3).

Benassi F., L. Cassata, G. Sindoni and D. Zindato. 2013. "Tales from the 2011 Italian Population Census. The use of a multi-mode data collection system: lessons learnt and future challenges". Relazione presentata alla Conferenza: 5th Conference of the European Survey Research Association (ESRA). Ljubljana 15-19 July.

Benassi, F., Bruno, M., Giacummo, M., Silipo, M., Vaste, G., e D. Zindato. 2014. Managing census complexity through highly integrated web systems, Istat, Rivista di Statistica Ufficiale, n. 3. http://www.istat.it/it/files/2015/03/Art.3_Managing-census-complexity.pdf.

Bernardini, A., Chieppa, A., Cibella, N., Solari, F., Zindato D. 2022. Evolution of the Italian Permanent Population Census. Lessons learnt from the first cycle and the design of the Permanent Census beyond 2021, Twenty-fourth Meeting of the Group of Experts on Population and Housing Censuses, Geneva, Switzerland, 21-23 September 2022, https://unece.org/statistics/documents/2022/07/working-documents/evolution-italian-permanent-population-census.

Bussola M., Cecconi N., Donati E., Porciani L. 2023. Respondents and non respondents to population and housing census: some strategies for data collection design in the era of low response rate and high response burden, paper presented at the conference LIX Riunione Scientifica SIEDS, NAPOLI, 25th May.

Cassata, L. and M.T. Tamburrano. 2011. The 15th Population Census Pilot Survey: how the register driven census changes the enumerators role in: S. Migani and M. Costa (Eds.) "Statistics in the 150 years from Italian unification", Serie Ricerche n. 10, Università di Bologna, Dipartimento di Scienze Statistiche "Paolo Fortunati", Bologna, http://amsacta.unibo.it/3202/1/Quaderni_2011_10_SIS2011_BookofShortPaper.pdf (April 2013).

Cecconi, N., e F. Cecconi. 2016. "Il profilo delle famiglie intervistate da Numero Verde. Un'analisi macro e micro dei questionari compilati." Paper presented at the seminar "Verso il censimento permanente della popolazione: le rilevazioni sperimentali C-sample e D-sample 2015", Rome, Istat, 26 January.

Dillman, D.A., Smyth, Jolene D., e Leah Melani Christian. 2009. Internet, Mail and Mixed-Mode Surveys: The Tailored Design Method, 3rd edition. Hoboken, NJ: John Wiley.

Eurostat. 2006–2016. Annual survey on ICT (Information and Communication Technologies) usage in households and by individuals. http://ec.europa.eu/eurostat/web/digital-economy-andsociety/publications/news.

Falorsi, S.: Census and Social Surveys Integrated System, 19th Meeting of the Group of Experts on Population and Housing Censuses, Geneva, Switzerland, 4-6 October (2017), https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2017/Meeting-Geneva-Oct/WP23_ENG.pdf.

Gallo G., Paluzzi E., Benassi F., Ferrara R. 2014. "The 2011 Italian experience towards supported Census for measuring migration." Economic Commission for Europe, Experience from 2010 round censuses for measuring migration, Working paper 7, 1–9.

Gallo, G. and Zindato, D. (2018). Annex H. Italy case study, in UNECE, Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses, Geneva, https://unece.org/guidelines-use-registers-and-administrative-data-population-and-housing-censuses-0.

Gallo, G. and Zindato, D. (2021). Italy: The combined use of survey and register data for the Italian Permanent Population Census count in UNECE,, Guidelines for Assessing the Quality of Administrative Sources for Use in Censuses (endorsed by the 69th plenary session of the Conference of European Statisticians), https://unece.org/statistics/publications/CensusAdminQuality.

Istat. 2009. A new strategy for the 2011 Italian Population Census. Product innovations and the compliance with CES Recommendations, UNECE/CES Group of Experts on Population and Housing Censuses, Twelfth Meeting, Geneva, 28-30 October, http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2009/5.e.pdf (April 2013).

Istat. 2012. Lessons learned from use of registers and geocoded databases in population and housing census, Sixtieth plenary session, Paris, 6-8 June, http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2012/22-IP_Italy.pdf (April 2013).

Tininini, L., e T. Virgillito. 2013. "The Design of the Online Questionnaire of the Italian Population Census." NTTS - Conferences on New Techniques and Technologies for Statistics, Brussels, March 5–7.

Virgillito, A., e L. Tininini. 2012. "The Web-based Data Collection in the Italian Population and Housing Census." Meeting on the Management of Statistical Information Systems, MSIS. http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.50/2012/18_Italy.pdf.

Zindato D (ed.), 2017, Dalla rilevazione 'porta a porta' alla rilevazione *paperless*: il Censimento della Popolazione, in Istat, L'utilizzo della tecnica CAWI nelle indagini su individui e famiglie, https://www.istat.it/it/files/2017/09/Lutilizzo-della-tecnica-Cawi.pdf.