# INSTANT MESSAGING TOOLS IN OFFICIAL STATISTICS: A USAGE MODEL IN THE 7° ITALIAN CENSUS OF AGRICULTURE

Claudia Fabi

**Abstract.** The data collection of the Italian 7th General Census of Agriculture took place from January to July 2021, through a synchronous multi-technique (CAPI, CATI, CAWI) design. During the survey, innovative contact tools were offered to respondents, alongside traditional communication channels. Through SMS and WhatsApp it was possible to collect the requests of over 2,000 farms, facilitating subsequent contact with the interviewers. The following work intends to show the results obtained and the propensity to use innovative contact channels. It proposes a model that allows correlating this use to some representative indicators of the territory characteristics, demographics, and the diffusion of the so-called "Digital culture".

## 1. The reference context

Between January and July 2021, the survey of the seventh General Census of Agriculture took place in Italy, the last one before the advent, also in this sector, of annual Permanent Censuses, on a sample basis.

For the first time, in an Italian census survey, a synchronous multi-technique design (CAWI, CAPI, CATI) was adopted, which allowed respondents a large amount of flexibility in choosing how to fill in the questionnaire (De Gaetano *et al.*, 2021 and De Gaetano *et al.*, 2022).

Important technological innovations, necessary to ensure the fluidity of data collection operations conducted by different data collection networks, have been supported by new strategies of contact with respondents, capable of responding more promptly to user requests if compared to more traditional tools such as the toll-free number or the email address of the census.

Starting from May 2021, a new contact channel for respondents was introduced, through a mobile phone number to which users could send SMS or WhatsApp messages, to request an appointment to subsequently complete the interview by phone. The main results obtained will be illustrated below, examining in detail the diffusion and propensity to use innovative contact tools on the national territory,

compared with the reference population of the Census. Finally, a correlation model will be proposed among the propensity to use innovative tools and territorial indicators, such as sectoral and socio-demographic context indicators.

## 2. Characteristics of the SMS and WhatsApp service in supporting respondents

The communication and assistance channels dedicated to the respondents of a survey, whether a census or a sample one, can be classified as synchronous or asynchronous[1]. The former, usually managed by operators, trained in advance for the purpose, are represented in most cases by free toll-free numbers, which users can contact to request general information on the survey, on how to participate, or ask for assistance in completing it on specific thematic issues.

However, even a synchronous tool such as the toll-free number has potential critical issues in its use, especially in some time slots of the day or in periods of intense inbound traffic, concentrated in the weeks immediately following the massive sending of information letters to the population involved in the survey. In this sense, SMS and WhatsApp take the form of a hybrid channel of support for respondents, being able to guarantee both synchronous – when managed by operators – and asynchronous support.

In the case of the Agricultural Census, the mobile telephone number, available for sending SMS and WhatsApp by the respondents, was managed by operators on the same days and times as the CATI survey and remained available as an asynchronous channel for the rest of the time, including holidays[2]. Users became aware of the existence of this innovative channel by calling the toll-free number or, for some of them, through the memorandum sent to the units still not responding in June 2021, by paper letter. The service was activated with the sole aim of offering respondents a means of requesting an appointment to complete the questionnaire with the CATI technique, and not as a complete information channel. It should be remembered that the number to send SMS and WhatsApp messages was activated just later about the start of the survey, and in particular only for the last three months of fieldwork: from May to July 2021.

[1] A communication channel is defined asynchronous when allows the contextual interaction between user and operator (e.g. Toll Free Number, Contact Center, chat), while is defined as asynchronous the communication channel through which the user forwards a request that will be processed at a later time by the operators who supervise it (e.g. email, forum, etc.).

[2] The telephone survey was carried out from Monday to Friday, from 9.00 a.m. to 9.00 p.m. and on Saturday from 10.00 a.m. to 7.00 p.m., with the exception of public holidays.

The following tables display the characteristics of service users, taking into account the two available demographic factors: age and geographic origin.

**Table 1** – *Channel used by respondents (SMS or WhatsApp).*

| Channel | Users | |
|---|---|---|
| | A.V. | % |
| SMS | 195 | 9.3 |
| WhatsApp | 1,908 | 90.7 |
| TOTAL | 2,103 | 100.0 |

Table 1 shows that, out of the approximately 2,000 received requests, over 90% are represented by WhatsApp messages, while only 9.3% are SMS. It appears that WhatsApp has almost completely replaced the messaging services offered by mobile network operators, in a widespread and established manner.

**Table 2** – *SMS and WhatsApp users by age group*[3].

| Age group | SMS or WA users | | Census List | |
|---|---|---|---|---|
| | V.A. | % | V.A. | % |
| Up to 30 years | 15 | 0.7 | 27,816 | 1.7 |
| From 31 to 40 years | 63 | 3.1 | 82,882 | 5.2 |
| From 41 to 50 years | 210 | 10.3 | 177,750 | 11.1 |
| From 51 to 60 years | 381 | 18.8 | 320,700 | 20.1 |
| From 61 to 70 years | 581 | 28.6 | 370,022 | 23.2 |
| From 71 to 80 years | 438 | 21.6 | 339,186 | 21.2 |
| Over 80 years | 343 | 16.9 | 279,555 | 17.5 |
| TOTAL | 2,031 | 100.0 | 1,597,911 | 100.0 |

As Table 2 shows, the age distribution of service users reveals an interesting phenomenon. Service users are not concentrated in the younger age groups, as one would expect if instant messaging services were more widely used by age groups that are more familiar with digital tools. When comparing the age distribution of service users with that of the census list, it becomes evident that it is within the age group of 61 to 70 years that the percentage of users is higher, compared to the reference population (28.6% vs. 23.2%). Even the younger age groups, up to 40 years old, have a usage profile that is lower than the incidence of the same age groups in the census population.

[3] Missing any other source of information, the distribution by age was obtained starting from the year of birth present in the tax ID code of the units in the census list. Where this information was not available, either in the census list or for SMS and WhatsApp users, the record was excluded from the calculation of the frequency distribution.

**Table 3 –** *SMS and WhatsApp users by region compared to the reference population.*

| Region | SMS or WA users | | Census List | |
|---|---|---|---|---|
| | V.A. | % | V.A. | % |
| Piedmont | 115 | 5.5 | 78,492 | 4.6 |
| Valle d'Aosta | 4 | 0.2 | 4,265 | 0.3 |
| Lombardy | 102 | 4.9 | 75,205 | 4.4 |
| Bolzano/Bozen | 7 | 0.3 | 27,350 | 1.6 |
| Trento | 9 | 0.4 | 19,374 | 1.1 |
| Veneto | 94 | 4.5 | 112,553 | 6.6 |
| Friuli-Venezia Giulia | 32 | 1.5 | 27,059 | 1.6 |
| Liguria | 49 | 2.3 | 23,765 | 1.4 |
| Emilia-Romagna | 92 | 4.4 | 78,642 | 4.6 |
| Tuscany | 94 | 4.5 | 81,350 | 4.8 |
| Umbria | 56 | 2.6 | 41,897 | 2.5 |
| Marche | 60 | 2.9 | 51,219 | 3.0 |
| Lazio | 187 | 8.9 | 117,963 | 7.0 |
| Abruzzo | 95 | 4.5 | 66,212 | 3.9 |
| Molise | 17 | 0.8 | 28,600 | 1.7 |
| Campania | 103 | 4.9 | 134,413 | 7.9 |
| Puglia | 454 | 21.5 | 266,195 | 15.7 |
| Basilicata | 45 | 2.1 | 49,766 | 2.9 |
| Calabria | 151 | 7.2 | 132,553 | 7.8 |
| Sicily | 277 | 13.2 | 211,179 | 12.4 |
| Sardinia | 60 | 2.9 | 71,890 | 4.2 |
| TOTAL | 2,103 | 100.0 | 1,699,942 | 100.0 |

The distribution of service users across Italian regions also highlights contrasting trends. While internet penetration among households still faces a gap between Northern and Southern Italy, favoring the northern regions, the usage of instant messaging services in the Agricultural Census is higher in the southern regions, such as Puglia and Sicily, where the agricultural sector plays a significant role in the local economy. In Puglia and Sicily, over one-third of service users are concentrated, accounting for 34.7%, a proportion higher than the presence of these regions in the census list (approximately 28%).

## 3. An explanatory model of the use of SMS and WhatsApp

With these premises, it seemed reasonable to delve deeper into the analysis by comparing a series of indicators selected to represent the cross-sectional dimension of "*smart farmers*" and the use of SMS and WhatsApp in the census survey.

In particular, the indicators by Italian region reported in Table 4 were considered to identify and aggregate, through a composite index, two main dimensions:

- the propensity of individuals to use digital communication channels.

This is the case of indicators Ind_062, AVQ_1, AVQ_2, and AVQ_3 which represent both the presence, in households, of connectivity devices and technology necessary for the use of digital communication tools, and their use particularly in contacts with the Public Administration (PA);

- the development and entrepreneurship in the Agricultural Sector.
  This dimension is represented by the indicators Ind_460, Ind_250, and Ind_031 which report information capable of representing the agricultural vocation of a territory and the growth trend of the agricultural sector in the local economic context. These indicators should therefore adequately capture the phenomenon highlighted in the regional distribution of Table 3, emphasizing those territories where agriculture represents a leading sector.

**Table 4** – *Indicators selected for the composite index*.

| Indicator | Content | Year |
|---|---|---|
| Ind_460 | Employment rate in rural areas (15-64 years) | 2021 |
| Ind_062 | Percentage of diffusion of the Internet in families | 2021 |
| Ind_250 | The growth rate of agriculture | 2021 |
| Ind_031 | Agricultural land productivity | 2021 |
| AVQ_1 | People aged 6 and over who use the Internet | 2021 |
| AVQ_2 | Possession of at least 1 mobile phone in the family | 2021 |
| AVQ_3 | People who have used the Internet to request information from PA Entities in the last 12 months | 2021 |
| AVQ_4 | Use of social networks in the last 3 months | 2020 |
| ISTR_CP | People with at least a high school diploma | 2021 |

Based on these hypotheses, these dimensions could represent the two main drivers to explain the inclination towards the use of innovative communication channels, even in fulfilling obligations towards the Public Administration, such as completing the Census questionnaire.

Furthermore, in the choice, only the indicators for which was available the annual data for 2021 or, in its absence, for 2020 at the regional level, were considered to represent phenomena contemporary to the data collection period of the Census.

The elaboration of a composite index was computed following the guidelines suggested by Mazziotta-Pareto for the synthesis of non-compensatory composite indexes (Mazziotta and Pareto, 2016) using the normalization formula that follows:

$$z_{ij} = 100 \pm \frac{(x_{ij} - M_{x_j})}{S_{x_j}} 10 \qquad (1)$$

where $M_{x_j}$ e $S_{x_j}$ is, respectively, the mean and the standard deviation of the indicator *j*, and the polarity, in the case of the subset of indicators chosen, is the positive one.

The composite index was computed following the formula:

$$MPI_i = M_{z_i} - S_{z_i} cv_{z_i} \quad (2)$$

where $cv_{zi} = S_{zi} / M_{zi}$ is the coefficient of variation for the unit *i*.

The following Table 5 shows MPI and the correlation with the percentage of users of SMS and WhatsApp, by region, computed concerning the total number of units in the census list. The Autonomous Provinces of Trento and Bolzano values have been computed separately.

The result, somewhat surprising, is a clear absence of correlation between the propensity to use innovative channels of communication with the PA and the PMI index calculated as detailed above. The two determinants of the phenomenon (digitalization and development of the agricultural sector) do not seem to have both an impact and a role in explaining the use of SMS or WhatsApp to communicate with institutions.

However, this is not entirely true. In fact, by calculating the simple correlations between each indicator and the percentage of SMS and WhatsApp users in the census list, a slight correlation has emerged with indicators Ind_031, AVQ_4, and Ind_250 (respectively the use of social networks, the productivity of agricultural land and the growth rate of agriculture), as shown in Table 6.

The same drivers are also highlighted in the *biplot* (Figure 1) among regions and the matrix of indicators, obtained with a principal component analysis, on the set of indicators listed in Table 4.

There is therefore a slight "*smart farmers*" phenomenon, as the correlations seem to suggest, even if this phenomenon does not significantly affect the propensity to use SMS and WhatsApp among the reference population.

**Table 5** – *MPI Index by Region.*

| Region | MPI | % Users |
|---|---|---|
| Piedmont | 98.8 | 0.15 |
| Valle d'Aosta | 102.4 | 0.09 |
| Lombardy | 105.2 | 0.14 |
| Province of Bolzano/Bozen | 106.5 | 0.03 |
| Province of Trento | 103.1 | 0.05 |
| Veneto | 99.3 | 0.08 |
| Friuli-Venezia Giulia | 101.4 | 0.12 |
| Liguria | 109.1 | 0.21 |
| Emilia-Romagna | 104.1 | 0.12 |
| Tuscany | 104.2 | 0.12 |
| Umbria | 98.0 | 0.13 |
| Marche | 98.9 | 0.12 |
| Lazio | 105.7 | 0.16 |
| Abruzzo | 98.8 | 0.14 |
| Molise | 91.8 | 0.06 |
| Campania | 96.5 | 0.08 |
| Puglia | 93.2 | 0.17 |
| Basilicata | 91.1 | 0.09 |
| Calabria | 88.1 | 0.11 |
| Sicily | 89.2 | 0.13 |
| Sardinia | 100.0 | 0.08 |
| *Coefficient of correlation* | | *0.098* |

**Table 6** – *Coefficient of correlations between SMS and WA user percentage and regional indicators.*

| Indicator | Content | Coefficient of correlation |
|---|---|---|
| Ind_460 | Employment rate in rural areas (15-64 years) | -0.115 |
| Ind_062 | Percentage of diffusion of the Internet in families | -0.069 |
| Ind_250 | The growth rate of agriculture | 0.177 |
| Ind_031 | Agricultural land productivity | 0.250 |
| AVQ_1 | People aged 6 and over who use the Internet | -0.063 |
| AVQ_2 | Possession of at least 1 mobile phone in the family | 0.121 |
| AVQ_3 | People who have used the Internet to request information from PA Entities in the last 12 months | -0.111 |
| AVQ_4 | Use of social networks in the last 3 months | 0.313 |
| ISTR_CP | People with at least a high school diploma | 0.029 |

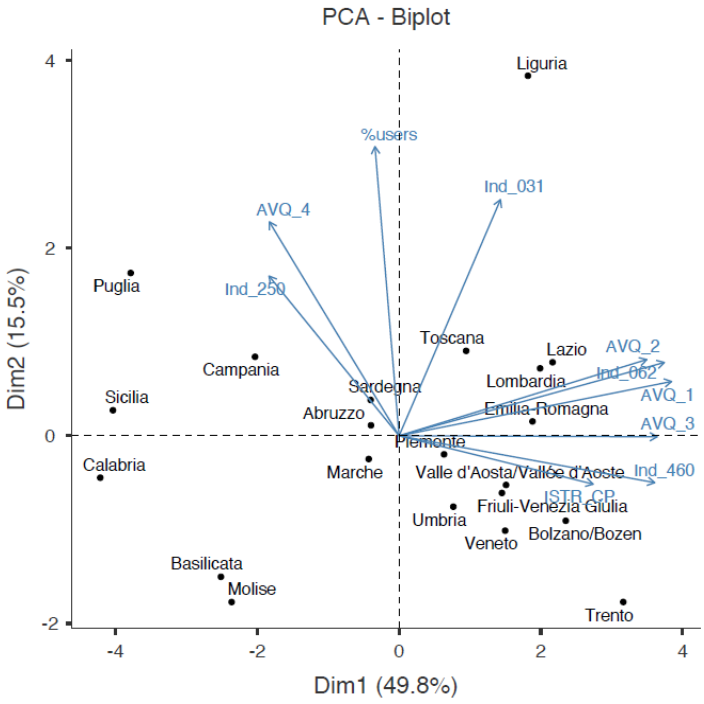
## 4. Conclusions

The correlation that emerges between individual/sectoral indicators and the use of SMS and WhatsApp, albeit slight, seems to suggest an expected tendency to favor these innovative channels exactly by people who possess dynamic characteristics

and a propensity to digitalization, even in the agricultural sector that is traditionally managed by small operators of high average age.

Furthermore, the absence of correlation with the PMI also suggests that investing in a digital communication channel such as WhatsApp to support respondents, is not an exclusive choice, in the literal sense of the term.

**Figure 1 –** *Biplot regions - indicators.*




It appears clear that there is not one (or more) specific socio-demographic profile that concentrates the majority of users of digital communication tools with the PA. This is probably due to the widespread diffusion of this tool among the population, which makes it a means of mass communication, just like we are used to considering the telephone (and consequently the Toll-Free Numbers). This makes SMS and above all WhatsApp, considering its user percentage, extremely interesting, right now, as an additional resource to complement the more usual contact strategies with respondents.

In the future, it is advisable to expand the use of these innovative contact services with the Public Administration, investing resources and conducting further analyses, not only in census surveys but also in sample surveys, particularly those concerning

households and individuals. Additionally, it will be crucial to invest in these innovative tools right from the beginning of field surveys, incorporating them clearly and prominently in the informative letters sent to sample units, to spread awareness and promote their usage.

Moreover, it is important to extend the use of these innovative channels not only for scheduling appointments for subsequent interviews but also as support tools for respondents, providing general information and assistance in completing online questionnaires. Digital assistance tools are in line with the dissemination and experimentation of Computer-Assisted Web Interviewing (CAWI) techniques in household surveys, allowing respondents to rely on quick and easily accessible assistance. Finally, it will be important to understand user satisfaction regarding the use of these instant communication services: an area of research that is still largely experimental in official statistics.

## References

DE GAETANO L., FABI C., TRIOLO V. 2021. Tecniche integrate di data collection per il Censimento Generale dell'Agricoltura 2020. In *Scientific Posters of the 14° National Conference of Statistics*, https://www.istat.it/storage/14-Conferenza-nazionale-statistica/poster/022.jpg.

DE GAETANO L., FABI C., TRIOLO V. 2022. 7th Agriculture Census: strategies, and methodological and technological innovations for the survey. In *Book of Abstracts of 7° Italian Conference on Survey Methodology*. https://www.dropbox.com/s/a9uuajjvb7gbyd9/BoA%20_itacosm2022.pdf?dl=0.

ISTAT 2023. Indicatori territoriali per le politiche di sviluppo - https://www.istat.it/it/archivio/16777.

MAZZIOTTA M., PARETO A. 2016. On a Generalized Non-compensatory Composite Index for Measuring Socio-economic Phenomena, *Social Indicators Research*, Vol. 127, pp. 983-1003, https://doi.org/10.1007/s11205-015-0998-2.

Claudia FABI, Istat – Istituto Nazionale di Statistica, claudia.fabi@istat.it