AN ANALYSIS OF THE RESPONSE BEHAVIOUR OF FARMS IN THE FARM STRUCTURE SURVEY[1]

Ornella Mobilia, Annalisa Pallotti, Francesca Rossetti

Abstract. The 2023 edition of the Farm Structure Survey is a sample survey that includes approximately 100,000 agricultural and livestock units registered in the Farm Register. This survey was conducted in accordance with Regulation (EC) No 1091/2018 of the European Parliament and of the Council, following an organizational model based on collaboration among several private entities.
Thanks to an agreement with Istat, the Agricultural Assistance Centras (CAA) played a crucial role in the survey activities, acting as intermediaries in the data collection for various agricultural surveys. Operators affiliated with CAAs have a privileged relationship with farms due to their ongoing activities and possess specialized training in agricultural matters. Data collection was carried out using a combination of two survey techniques to optimize costs and minimize the statistical burden on respondents. Farms were given the option to complete the questionnaire online via CAWI (Computer Assisted Web Interviewing). Non-responding units were subsequently contacted by an operator who administered the questionnaire using CAPI (Computer Assisted Personal Interviewing).
The objective of this work is to analyse the possible links between the characteristics of farms and their response behaviour, also in relation to the possible effects of actions taken to encourage participation.

1. Introduction

In the context of the 2023 edition of the Farm Structure Survey (Eurostat, 2022), which covered around 100,000 agricultural and livestock holdings from the National Farm Register, this study examines the effectiveness of collaborative data collection strategies in the agricultural sector. Conducted under an European Regulation (EC), the survey employed an innovative organisational model involving multiple private actors. Central to this model were the Agricultural Assistance Centres (CAAs) (Poti B., 2011), which acted as trusted intermediaries in the collection of farm data through a formal agreement with Italian National Institute of Statistics (Istat). Thanks to their specialised training and long-standing relationships with local farms, CAA operators

[1] The article was only possible thanks to the joint work of the authors. In particular, Annalisa Pallotti wrote sections 2, 4 and 6, Francesca Rossetti wrote sections 1 and 5 Ornella Mobilia wrote section 3.

played a pivotal role in ensuring effective engagement with respondents. This article explores the relationship between farm characteristics and survey response behaviour, paying particular attention to the impact of targeted actions designed to increase participation. We propose a statistical analysis based on the logistic model to assess the propensity of farms to cooperate in the survey in relation to the survey characteristics of the targeted actions.

The document is structured as follows: Sections 2 and 3 outline the survey objectives and design, with a particular focus on the survey network and data collection methods; section 4 presents a detailed analysis of response rate trends over the course of the data collection period; section 5 describes the results of the application of the logistical model; finally, section 6 provides concluding remarks.

2. Farm Structure Survey

Farm Structure Survey represents one of the main structural statistical sources on the Italian agricultural system. It is a sample survey carried out every three years, conducted in the intermediate years between the decennial general agricultural censuses. This periodicity allows for dynamic monitoring of the evolution of farms between two full censuses, ensuring continuity of statistical information.

The 2023 edition involved 109,229 agricultural and livestock holdings, selected from the Farm Register, and was conducted in compliance with Regulation (EU) No. 2018/1091 and its related Implementing Regulation of the European Commission (No 2024/2914), as part of the system of harmonized surveys established at the EU level. The data collected are crucial for supporting impact analyses of the Common Agricultural Policy (CAP), environmental sustainability, and rural development, providing analytical insights in the inter-census period.

The main objectives of the survey are to analyze the structural evolution of farms, through the collection of information on their economic and physical size, production type, labor force, use of natural resources, and agro-environmental practices. The variables collected were designed to provide a multidimensional reading of the agricultural holding, covering both structural aspects (land area, livestock, labor, machinery) and farming techniques (crop types, fertilization methods, livestock systems).

The methodological design included a stratified and regionally representative sampling plan, ensuring sufficient coverage for disaggregated data analysis. Particular attention was paid to improving the quality of the questionnaire and simplifying the electronic data collection interface. The questions were adapted linguistically and functionally to make them easier for respondents to understand and reduce non-response rates.

In this context, a well-structured and target-specific data collection strategy proved essential to ensure the timeliness of information, alignment with current sector dynamics, and the statistical robustness required for analytical and policy-making purposes.

3. Network and Data Collection Techniques

The organizational structure of the survey was based on an integrated collaboration model between ISTAT and the Agricultural Assistance Centers, formalized through the signing of an operational agreement that assigned the CAA the role of official data collection bodies on behalf of the Institute. The CAA, entities recognized by the Ministry of Agriculture and accredited by the National Agricultural Information System (SIAN), act as certified intermediaries between the State and farms for the management of the farm register and access to key instruments of the Common Agricultural Policy (MASAF, 2023).

Their role within the data collection network is not only logistical but also methodological: thanks to their familiarity with local farming practices, in-depth knowledge of regulatory definitions, and technical expertise, the CAA contribute to coherent, standardized, and highly reliable data collection.

For the first time, during the 2020 Agricultural Census, the network of Agricultural Assistance Centers (CAAs) was employed as the primary data collection channel, producing positive outcomes in terms of territorial coverage and data quality (Istat, 2023). Owing to their longstanding relationships with farms and specialized expertise in agricultural matters, CAA operators served as qualified intermediaries, enabling comprehensive, detailed, and accurate data gathering across the national territory. In the aftermath of the COVID-19 emergency, in fact, reliance on administrative records had resulted in an over coverage of the farm register, thus underscoring the necessity of field verification to determine the actual number of active agricultural holdings. In this context, the CAA network—characterized by its technical competence and in-depth sectoral knowledge—proved instrumental in achieving the cognitive goals of the Census, further consolidating its collaboration with Istat.

In the 2023 edition of the survey, their involvement enabled the combination of operational efficiency and territorial coverage, ensuring high data quality nationwide. The network deployed included 6 lead CAA organizations, 1,819 local offices, and 2,290 trained operators, who maintained trusted relationships with local farms. The 6 lead CAAs each coordinated a group of affiliated CAAs, bringing the total number of national CAAs involved in the network to 28.

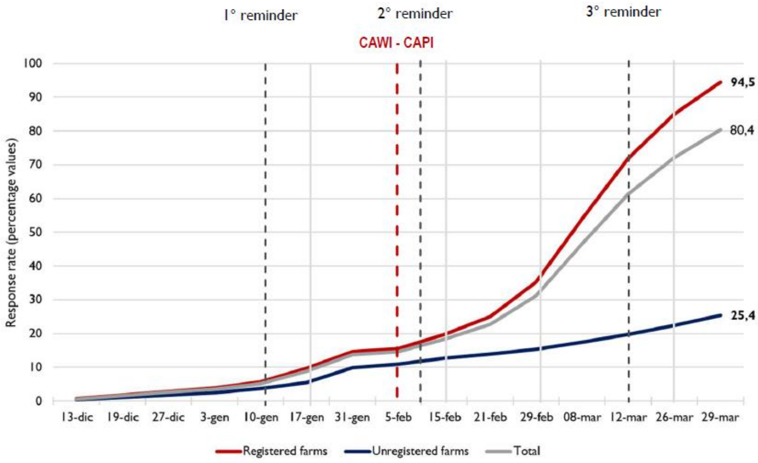
The sample of approximately 109,229 holdings was divided into two groups: farms with an active farm register, hereinafter "registered" (86,995 units registered with a participating CAA) and farms without an active register, hereinafter "unregistered", (22,304 units either not associated with or not recorded in a participating CAA). The survey, conducted between December 2023 and March 2024, utilized two data collection modes: Computer-Assisted Web Interviewing (CAWI), for self-completion of the electronic questionnaire, and Computer Assisted Personal interviewing (CAPI) conducted by CAA operators. While the first group had access to both modes sequentially, the second group could only use the CAWI mode. The initial communication inviting units to participate in the survey also includes a description of the technique(s) through which they can contribute. This enables registered units to decide whether to wait for the CAA support available during the second part of the survey period.

This multimodal approach made it possible to optimize participation, reduce respondent burden, and ensure the quality of the data collected.

4. Trends in Survey Response Rates During Fieldwork

During the field data collection period, farm response rates were continuously monitored. Figure 1 illustrates the response rate trends in relation to actions taken to encourage survey participation, distinguishing between the two groups defined by the survey design: registered and non-registered units.

Figure 1 – *Response rates by reminder and technical changes.*



Source: our elaboration on data from Farm Structure Survey 2023.

The response curve showed a steady upward trend from the beginning of data collection, with a marked acceleration following the first scheduled reminder sent on 10 January. By 17 January, the overall response rate had increased by 5.3 percentage points, and by 5.8 points among registered farms. The reminder stated that registered farms could respond via CAWI mode until 21 January[2]. Following the reminder, a notable increase was observed across all series, particularly among registered farms.

From 5 February, with the launch of the CAPI phase, the implementation of the mixed CAWI-CAPI technique for registered farms marked a qualitative leap in data collection. As of 15 February, the response rate stood at 19.8% for registered units and 12.8% for non-registered ones.

The second reminder, sent on 12 February, positively influenced survey performance—especially for registered farms: by 29 February, the observed response rates had reached 35.2% for registered farms and 15.3% for non-registered farms.

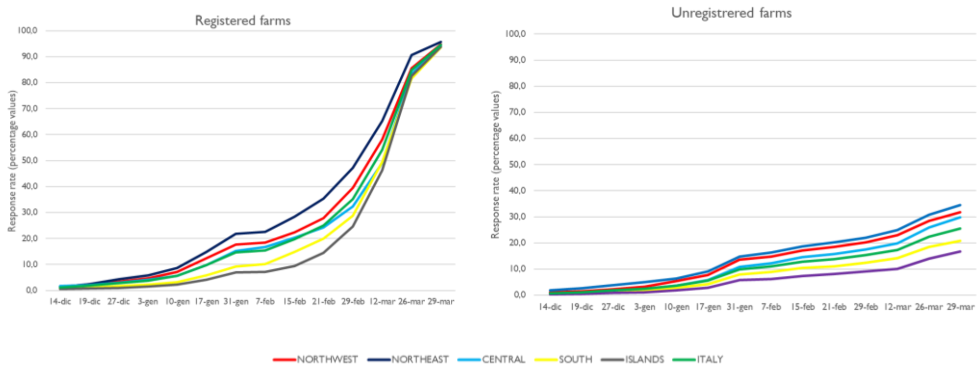
The final reminder, issued on 12 March, was instrumental in maximizing coverage, particularly among registered farms. The ability of CAAs to rapidly mobilize local structures contributed to the final increase, in line with the closure target set for the end of March.

By 29 March, the end of the data collection period, the overall response rate had reached 80.4%, with a peak of 94.5% among registered farms and a final response rate of 25.4% among non-registered farms.

The trend in response rates over time, revealing consistent patterns across different geographic areas, but with higher rates in the North than in the South. This pattern is evident in both the registered and unregistered groups, which, as previously noted, follow distinct trajectories. A comparison of the two sample segments shows that the introduction of the CAPI phase for registered farms, during the second part of data collection, marks a turning point in the trend lines, as the gap in response rates between geographic areas gradually narrows over time. By the end of the data collection period, response rates among registered participants had converged, with differences reduced to within 1.2 percentage points. The intervention of CAA operators helped offset the lower survey participation rates among firms in the South and the Islands, contributing to improved data quality in these areas. In contrast, among unregistered participants, disparities in response propensity persisted throughout the period, with a gap of 15.1 percentage points observed between the North-East and the final phase (Figure 2).

[2] The start of the CAPI phase, initially scheduled for 22 January, was postponed to 5 February.

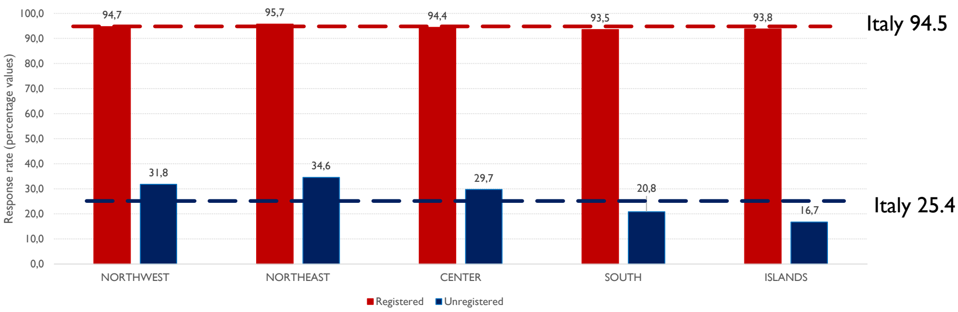
Figure 2 – *Response rates by subsample and geographic area*.



Source: our elaboration on data from Farm Structure Survey 2023

Farms with a company file achieve the highest response rates in all geographical areas, with clear peaks in the north-east of the country. These areas therefore appear to be characterised by a greater formalisation of agricultural businesses. Companies without a registered file, on the other hand, show lower response rates everywhere, with the greatest critical issues being highlighted in the Islands and the South. This data could reflect structural differences in farms, with fewer formalised companies in these areas (Figure 3).

Figure 3 – *Territorial distribution of response rate.*



Source: our elaboration on data from Farm Structure Survey 2023.

5. Exploring Predictors of Participation through Logistic Modelling

Analysis of the response rates revealed varying levels of participation in the survey among registered farms and the rest of the sample. This characteristic also determined assignment to one of two survey designs, using either dual CAWI-CAPI or CAWI-only techniques.

To further analyse the determinants of farm participation in the survey, a logistic model was applied (Long et al., 2014; Agresti, 2013). In the model the dependent variable is having or not having collaborated in the survey. The explanatory variables used in the model are the survey technique and the variables that provide a measure of the farms, such as labour units per year (ULA), standard farm production value in euros (SO), utilised agricultural area in hectares (UAA) and Standard livestock units (UBA). The model also considers the geographical area and membership of one of the CAA involved in the convention. The values entered the model by the variables are shown in the Table 1.

Table 1 – *Explanatory variables and mode encoding applied*[3].

Explanatory variables	Code
Survey techniques	ind_tecnica=0 for CAWI; ind_tecnica =1 for CAPI
Labour Units per Year – ULA	ULA_TOT_ind1= until 1; ULA_TOT_ind2=from 1 to 9; ULA_TOT_ind3= over 10
Value of the company's standard output - SO in euros	SO_ITA_ind1= Up to 2000; SO_ITA_ind2=From 2.000 to 15.000; SO_ITA_ind3=from 15.000 to 100.000; SO_ITA_ind4=from 100.000 to 1.000.000; SO_ITA_ind5=over 1.000.000
Utilised Agricultural Area – SAU	SAU_ind1= Until 0.99; SAU_ind2=From 1 to 1.99; SAU_ind3=From 2 to 4.99; SAU_ind4=From 5 to 19.99; SAU_ind5=over 20
Standard livestock units – UBA	UBA_ind1= Until 1.99; UBA_ind2=From 2 to 4.99; UBA_ind3=From 5 to 9.99; UBA_ind4=From 10 to 19.99; UBA_ind5=From 20 to 49.99; UBA_ind6=over 50
Geographical area	ripa_NO=North-West; ripa_NE=North-East; ripa_CE=Centre; ripa_SU=South; ripa_IS=Islands for Geographical area
CAA of membership	caa_1=1; caa_2=2; caa_3=3; caa_4=4; caa_5=5; caa_6=6

Source: our elaboration on data from Farm Structure Survey 2023

[3] The CAAs that contributed to the data collection process, listed in alphabetical order, are: CAA CAF Agri, CAA CIA, CAA Coldiretti, CAA Confagricoltura, CAA Degli Agricoltori, CAA Delle Venezie. To safeguard the confidentiality of individual performance—which is not pertinent to the objectives of this paper—we have opted to refer to them using random numerical codes.

Table 2 reports the Chi-square statistics along with the corresponding p-values, highlighting variables that are not statistically significant in the model. It also includes the odds ratios, which represent the logistic regression coefficients for the explanatory variables[4].

Table 2 – *Explanatory variables, Chi-square statistics with p-value associated.*

	Parameter	odds ratio	Pr > Chi-sq
	Intercept	105,0	<.0001
Survey techniques	ind_tecnica	1,2	<.0001
Work Units per Year – ULA	ULA_TOT_ind2	1,8	<.0001
	ULA_TOT_ind3	1,0	<.0001
Value of the company's standard output – SO	SO_ITA_ind2	1,2	0,3441
	SO_ITA_ind3	1,3	0,0092
	SO_ITA_ind4	2,3	<.0001
	SO_ITA_ind5	1,1	<.0001
Utilised Agricultural Area – SAU	SAU_ind2	1,0	0,2619
	SAU_ind3	1,2	0,8967
	SAU_ind4	1,4	<.0001
	SAU_ind5	0,8	<.0001
Standard livestock units – UBA	UBA_ind2	0,6	0,0021
	UBA_ind3	0,6	<.0001
	UBA_ind4	0,6	<.0001
	UBA_ind5	0,7	<.0001
	UBA_ind6	2,3	<.0001
Geographical area	ripa_NO	4,4	<.0001
	ripa_NE	2,2	<.0001
	ripa_CE	1,3	<.0001
	ripa_SU	105,0	<.0001
CAA of membership	caa_1	1,2	<.0001
	caa_2	1,8	<.0001
	caa_3	1,0	<.0001
	caa_4	1,2	<.0001
	caa_6	1,3	<.0001

Source: our elaboration on data from Farm Structure Survey 2023.

According to the observation of the p-values, it can be inferred that nearly all the independent variables considered in the analysis are statistically significant. Based on the multicollinearity analysis - which excluded the presence of critical collinearity among predictors - and the negligible difference observed in goodness-of-fit indicators when excluding non-significant variables, the model was retained in its

[4] The baseline values of each predictor are ULA_TOT_ind1 for ULA; SO_ITA_ind1 for SO; SAU_ind1 for SAU; UBA_ind1for UBA; Islands for Geographical area; caa_5 for CAA.

entirety. This choice ensures thematic coherence and structural consistency throughout the overall reading of the article. The overall fit of the logistic regression model was deemed acceptable. The model's effectiveness was confirmed by multiple concordance measures, each showing strong results (Hosmer et al., 2013). Detailed information is provided in the Appendix.

The model confirms that the assigned data collection method is the primary driver of survey participation: this variable is determined by the survey design and is closely linked to a farm's registration with a CAA involved in the survey network. In fact, the observed odds for CAA membership generally indicate a higher likelihood of participation among registered units compared to unregistered ones, with varying propensities depending on the specific CAA grouping to which they belong. The odds ratios associated with farm characteristics are all greater than 1, indicating that the propensity to participate increases with the size of the farm. The only exception is found in relation to the number of standard livestock units, for which the odds ratios are less than 1: here, the likelihood of participation decreases as the size class increases. Regarding geographic area, units located outside the island regions show a significantly higher propensity to participate—up to more than four times greater in the eastern regions.

6. Conclusion

The success of the survey strategy relies heavily on the unique strengths of the Agricultural Assistance Centres. Their direct knowledge of farms and widespread territorial presence play a critical role in fostering participation among statistical units. The professional network they represent—often composed of agronomists with deep expertise in the subject matter—enhances the quality of microdata collected from registered holdings. However, the inclusion of unregistered farms, often less inclined to engage due to their weaker links with CAAs, required a deliberate oversampling approach to mitigate risks of under-coverage and self-selection bias. The adoption of the CAWI technique further ensured broad accessibility, enabling participation from all farms regardless of registration status. Notably, the communication of the assigned survey technique from the outset can influence response behavior, particularly for CAPI among CAA-affiliated units. Ultimately, the use of a technically equipped and territorially embedded network remains essential for high-quality, inclusive data collection across the agricultural sector.

Appendix

Assessment and key indicators of the deployed logistic model

Multicollinearity analysis.

The correlation matrix computed for the variables listed in Table 1 shows coefficients ranging from –0.58 to 0.46, indicating the absence of strong linear associations among the variables. Furthermore, the assessment of multicollinearity through the Variance Inflation Factor (VIF) reveals values below the commonly accepted threshold of 5 for the majority of variables. Only a limited number of exceptions were observed, with VIF values falling within the moderate range of 6 to 8 (SO_ITA_ind3 = 6.09; SO_ITA_ind = 7.78; SAU_ind5 = 6.53), which do not raise substantial concerns regarding multicollinearity (see Table A1).

Table A1 – *Variance Inflation Factor values for all variables included in the model*.

Variables	VIF
ULA_TOT_ind2	1,72
ULA_TOT_ind3	1,18
SO_ITA_ind2	3,06
SO_ITA_ind3	6,09
SO_ITA_ind4	7,78
SO_ITA_ind5	2,81
SAU_ind2	1,62
SAU_ind3	2,44
SAU_ind4	4,68
SAU_ind5	6,53
UBA_ind2	1,02
UBA_ind3	1,03
UBA_ind4	1,10
UBA_ind5	1,19
UBA_ind6	1,55
ripa_NO	1,94
ripa_NE	2,12
ripa_CE	1,97
ripa_SU	2,13
caa_103	1,91
caa_124	1,26
caa_107	1,46
caa_129	1,24
caa_105	1,59

Source: our elaboration on data from Farm Structure Survey 2023.

Logistic model diagnostics.

The Akaike Information Criterion (AIC = 39192) and the Bayesian Information Criterion (BIC = 39441) suggest a reasonable balance between model complexity and goodness of fit, with lower values indicating better parsimony. The –2 Log-Likelihood value (–2 Log L = 39140) further supports the adequacy of the model in capturing the observed data structure.

In terms of explanatory power, the model yielded a Coefficient of Determination (R^2) of 0.39, indicating that approximately 39% of the variance in the outcome variable is accounted for by the predictors. Notably, the Adjusted R-squared value of 0.68 suggests a substantial improvement in model fit when adjusting for the number of predictors, highlighting the robustness of the model specification (see Table A2).

Table A2 – *Logistic model diagnostics: comparison between models using significant variables vs. models using all variables, with percentage differences in indicators.*

Fit indicator	Model using significant variables	Model using all variables	Percentage difference
AIC - Akaike Information Criterion	39188	39192	0,01
SC (BIC) - Schwarz Criterion (Bayesian Information Criterion)	39409	39441	0,08
-2 Log L - Negative Two Times the Logarithm of the Likelihood Function	39142	39140	-0,01
R-squared - Coefficient of Determination	0,39	0,39	0,00
Adjusted R-squared Adjusted Coefficient of Determination	0,68	0,68	0,00

Source: our elaboration on data from Farm Structure Survey 2023.

Predictive performance.

The model demonstrates excellent discriminative ability, as indicated by an AUC/c-statistic of 0.959, which is considered outstanding. Concordance statistics—Somers' D of 0.917 and Gamma of 0.918—further confirm that the predicted probabilities are well aligned with the observed outcomes. The Tau-a coefficient, equal to 0.239, is comparatively lower, as expected, since it accounts for the total number of possible pairs (see Table A3).

Table A3 - *Diagnostic indicators assessing the discriminative power and ordinal association of the logistic model.*

Diagnostic indicators	Value
Percentage of concordant pairs	95,8
Percentage of discordant pairs	4,1
Percentage of tied pairs	0,1
Somers' D	0,917
Gamma	0,918
Kendall's Tau - Tau-a	0,239
Area Under the Curve - AUC/c-statistic	0,959

Source: our elaboration on data from Farm Structure Survey 2023.

References

AGRESTI A. 2013. *Categorical Data Analysis.(*3rd) Hoboken, Wiley.

EUROSTAT. 2022. Farm Structure Survey 2020 – *Main Results*. Luxembourg: Publications Office of the European Union

HOSMER D. W., LEMESHOW S., STURDIVANT R. X. 2013, *Applied Logistic Regression* (3rd ed.). Hoboken, Wiley.

ISTAT. 2023. 7th General Census of Agriculture. *Final Results*. Rome: ISTAT.

LONG J.S., FREESE J**.** 2014. *Regression Models for Categorical Dependent Variables Using Stata*. College Station: Stata Press.

MINISTRY OF AGRICULTURE, FOOD SOVEREIGNTY AND FORESTS (MASAF). 2023. *National List of Authorized Agricultural Assistance Centers (CAA)*. Rome: MASAF

POTI B. 2011. The Organization of Services for Agricultural Enterprises: Evolution and Role of Agricultural Assistance Centers (CAA). In: *QA Journal of the Rossi-Doria Association*, No. 1, pp. 103–122. Rome: Donzelli.

Ornella MOBILIA, Istat, ornella.mobilia@istat.it
Annalisa PALLOTTI, Istat, pallotti@istat.it
Francesca ROSSETTI, Istat, frrosset@istat.it