DO THE RESPONSE RATES FOR SHORT-TERM BUSINESS SURVEYS HAVE A SEASONALITY?[1]

Katia Bontempi, Claudio Ceccarelli, Francesca Rossetti

Abstract. As with many statistical outputs produced by the Italian National Statistics Institute (ISTAT), short-term business statistics are subject to specific European regulations. These regulations define the structure of questionnaires and surveys, the scope of observation, the reference population, the sampling methodology, and the precision of the estimates. Within the specific context of enterprises, especially in the Italian production landscape, where small businesses are predominant, regulatory criteria and constraints necessitate the continuous inclusion of large enterprises in survey samples over time.

Through the time-series analysis of business response rates, this paper aims to verify the presence of seasonality in the response behaviour of businesses over time. Should seasonal factors emerge, it would be possible to consistently provide specific interventions modulated according to the phase in which business participation is most lacking. This would allow an optimization of the resources deployed and a reduction of the burden in periods when participation is highest.

The data analysed in this study come from the *The Business Statistical Portal*, i.e. the system that collectively manages all short-term business surveys. The information used, properly normalized and standardized, covers all the survey units involved in short term business surveys in the period 2016 to 2024.

1. Introduction

Short-term business statistics are governed by specific European regulations that define the structure of questionnaires and surveys, the scope of observations, the reference population, sampling methodology, and the accuracy of estimates. In the context of enterprises—particularly within Italy's production sector, which is predominantly composed of small businesses—regulatory requirements necessitate the repeated inclusion of large enterprises in survey samples over time.

The analysis presented in this article is aimed at investigating characteristics of enterprise participation in surveys conducted by the Italian National Institute of

[1] The article was only possible thanks to the joint work of the authors. In particular, Claudio Ceccarelli wrote section 1 and 6, Francesca Rossetti wrote section 2 and 4, Katia Bontempi wrote sections 3 and 5.

Statistics (ISTAT) and complements previous work conducted using the survival data analysis approach (Binci *et al.*, 2025).

The analysis carried out in this paper aims to verify the presence of seasonal components in the trend of response rates by applying time-series analysis and checking whether different behaviours can be observed between surveys with differing characteristics. The data used cover all the survey units involved in two specific short-term surveys, Employment in large enterprises and Retail trade, over the period 2016-2024.

Section 2 details the characteristics of the short-term business surveys and describes the characteristics of the platform used for the data collection, "The Business Statistical Portal" (BSP). Section 3 outlines the characteristics of the business surveys being analysed. Section 4 presents the methodology underlying the work. The principal results are illustrated in Section 5, while Section 6 provides a summary of the findings and discusses future research directions.

2. Centralized Management of Short-Term Business Surveys

Short-term Business Statistics (STS) are economic indicators collected on a monthly or quarterly basis, essential for tracking business cycles and supporting monetary and fiscal policies. These indicators are regulated by Regulation EC No. 2152/2019 amending Regulation EC No. 1158/2005 and all further implementing and amending regulations. Furthermore, STS form part of the National Statistical Programme, which establishes the rules for data collection and management.

The introduction of a centralized access system through the Business Statistical Portal (BSP) and a dedicated Contact Centre has optimized enterprise support by providing standardised responses, reducing the burden on respondents, and improving communication management (Fazio *et al.* 2013). This innovation has fostered a virtuous cycle in business collaboration, resulting in significant increases in response rates for the surveys considered. In particular, the BSP allows for an integrated and harmonized management of all survey phases, enabling shared access to information based on user profiles.

To ensure high-quality and timely data, communications with enterprises have been improved through a revision of the ISTAT information letter, focusing on deadlines, along with a notification system that includes certified emails (PEC), regular emails, and telephone reminders and a compliance penalty system has been introduced (Binci *et al.*, 2025).

The use of the BSP has standardised the data collection processes, allowing the comparability of enterprises' response behaviour. At the same time, the standardisation of the STS makes a long series of data available for each of the surveys (Bellini *et al.* 2019).

3. The Short-Term Business Surveys selected for the analysis

The analysis focuses on two surveys that fall within the scope of the STS and share similar methodological characteristics: Monthly Survey on Employment, Working Hours, Wages, and Labour Costs in Large Enterprisesand and Monthly Retail Trade survey. These surveys were selected because they allow for a comparative evaluation of response rates across two different sampling designs (see Table 1).

Table 1 – *Selected short-term business surveys: main characteristics.*

Survey	Observation field	Sampling design	Statistical unit
Employment, Working Hours, Wages, and Labour Costs in Large Enterprises (OCC)	Enterprises with at least 500 employees	Census survey for enterprises with at least 500 employees	Functional unit[2]
Retail trade (DETT)	Enterprises with main economic activity in sec G of the Nace Rev. 2 classification	Stratified random sampling for enterprises with less than 50 employees - Census survey for enterprises with at least 50 employees	Enterprises

Istat

3.1 The Monthly Survey on Employment, Working Hours, Wages, and Labour Costs in Large Enterprises Survey

The Monthly Survey on Employment, Working Hours, Wages, and Labour Costs in Large Enterprises (OCC) is a critical statistical tool designed to monitor labour demand in enterprises with employees.

The OCC survey is part of the National Statistical Program (IST-00050), and subject to Legislative Decree No. 322/1989, which mandates enterprises to provide accurate labour data. The results are processed to generate monthly index numbers,

[2] Functional unit: the unit within an enterprise that groups together all parts contributing to the conduct of an economic activity at the class level (four digits) of the NACE Rev.1 classification. It is an entity corresponding to a system of information that allows at least the value of production, intermediate consumption, staff expenses, operating results, employment, and gross fixed investments to be provided or calculated for each economic activity unit.

published in the IstatData[3] database, while aggregated annual data appear in the Italian Statistical Yearbook (ASI). The survey targets large enterprises in industry and services, specifically those classified under Ateco 2007 sectors B to S, excluding temporary employment agencies, with at least 500 employees; the enterprises surveyed number is around 1,600. Key variables recorded include employee positions, workforce inflow and outflow, hours worked, paid but non-worked hours, unpaid strike hours, and wage guarantee fund hours. Data on solidarity contract hours, detailed wage components, social charges, and job vacancies at the end of each quarter are also collected. These indicators are analysed by qualification categories: clerical staff, manual workers, and executives.

The OCC survey is part of an integrated system alongside the Quarterly Survey on Job Vacancies and Hours Worked and the Quarterly Survey on Employment, Wages, and Social Charges, ensuring comprehensive coverage of employment trends. By integrating data coming from different sources, these surveys help policymakers track employment dynamics, wage trends, and labour costs, contributing to informed economic decisions.

3.2 Retail trade survey

The monthly retail trade survey (DETT) is a sampling-based statistical survey aimed at monitoring the performance of the retail sector across Italy. The survey is conducted in strict adherence to European regulatory frameworks, notably the EU Regulation 2019/2152, issued by the European Parliament and Council on November 27, 2019, replacing the earlier Regulation (EC) No. 1165/1998 and subsequent updates and is part of the National Statistical Program (IST-00151).

The survey is designed to capture essential economic trends in consumer purchasing behavior, providing data-driven insights for policymakers, businesses, and researchers. The methodology ensures statistical representativeness through a cut-off sampling strategy: for enterprises with fewer than 50 employees, a random selection is used, while enterprises with 50 or more employees are included in the survey by census.

The survey focuses on approximately 8,000 enterprises engaged in retail commerce, forming a longitudinal study where certain units overlap over time to maintain consistency and comparability. The largest enterprises, defined as those with more than 50 employees, are sourced from the Statistical Archive of Active Enterprises (ASIA) to ensure comprehensive sector coverage. By adopting a strategic sampling approach, ISTAT aims to create a dataset that accurately represents the full spectrum of retail businesses, from small independent retailers to large-scale commercial enterprises.

[3] https://esploradati.istat.it/

The information gathered includes total sales value, broken down by product category and sales method (such as brick-and-mortar stores, e-commerce platforms, door-to-door sales, and vending machines). Additionally, enterprises report the number of retail locations open to the public and provide year-over-year comparisons of sales performance for the corresponding months. The collected data undergoes extensive processing to generate monthly index numbers, which serve as a benchmark for analysing retail activity fluctuations. These results are subsequently published in the ISTAT press release, "Retail Trade," and made accessible via the IstatData database, ensuring transparency and availability for further analysis.

4. Applying ARIMA Models to Identify Seasonal Patterns

The application of seasonal adjustment techniques using an additive approach is a widely recognized method in time series analysis, aimed at eliminating seasonal fluctuations to isolate the underlying structure of a dataset. The seasonal adjustment of time series data with intra-annual frequency is based on the assumption that such series can be represented as a combination of distinct orthogonal components: the cycle-trend (CT) component, which reflects the medium- to long-term trend of the time series, unaffected by short-term fluctuations; the seasonal component (S), which recurs throughout the year and captures fluctuations attributable to meteorological, customary, or legislative factors and the irregular component (I), which results from erratic factors.

Under the additive model assumption, the data series (Y) can be expressed as:

$$Y = CT + S + I \tag{1}$$

Once the individual components are identified using appropriate statistical techniques (seasonal adjustment procedures), the seasonally adjusted series is obtained by subtracting the seasonal component from the original time series:

$$Dest(Y) = Y - S = CT + I \tag{2}$$

This method allows for the isolation of the underlying trend and irregular fluctuations, ensuring a more accurate representation of the fundamental dynamics within the dataset. In this study, the first step—identification of seasonality—was implemented to assess whether seasonal fluctuations were present. The procedure was intentionally halted before the second phase, as the goal was not full seasonal adjustment but rather verification of the existence of seasonal influences in the dataset. To fulfil this purpose, an autoregressive integrated moving average

(ARIMA) model was employed. This parametric approach utilizes filter-based techniques to process time series data, capturing underlying statistical dependencies and preparing the dataset for further modelling. The ARIMA model consists of three distinct components: AR (Autoregressive), which captures dependency on past observations within the time series; I (Integrated), which determines the number of differencing steps required to transform the series into a stationary format; and MA (Moving Average), which models dependencies based on past error terms (Dagum *et al.* 1999).

The key preprocessing step involved the application of first-order differencing (LAG(1)), which serves multiple statistical purposes: stabilizing the mean of the series over time, eliminating non-stationary trends or systematic components, and preparing the dataset for subsequent ARIMA modelling. These steps are critical, as stationarity is a fundamental requirement for the efficacy of ARIMA models in forecasting and trend analysis.

$$LAG(1) = Y_m - Y_{m-1} \tag{3}$$

A central objective of this study was to determine whether the removal of the trend component would reveal residual seasonal structures in the autocorrelation analysis. This approach is particularly useful in cases where apparent seasonal patterns are suspected but not explicitly confirmed. By the application of first-order differencing in order to eliminate the trend component and to analyse residual autocorrelation, researchers can assess whether additional adjustments or modelling refinements are necessary to capture persistent seasonal effects in the dataset.

In what follows, we propose applying a time series approach to analyse response rates observed over time (Ceccarelli *et al.*, 2016). These rates pertain to the DETT and the OCC surveys; to ensure comparability between the data collected in the two surveys, enterprises are used as the statistical units in this analysis. In the OCC survey, where the reference units are functional units, an occurrence of non-response is defined as the situation in which no functional unit has provided a response.

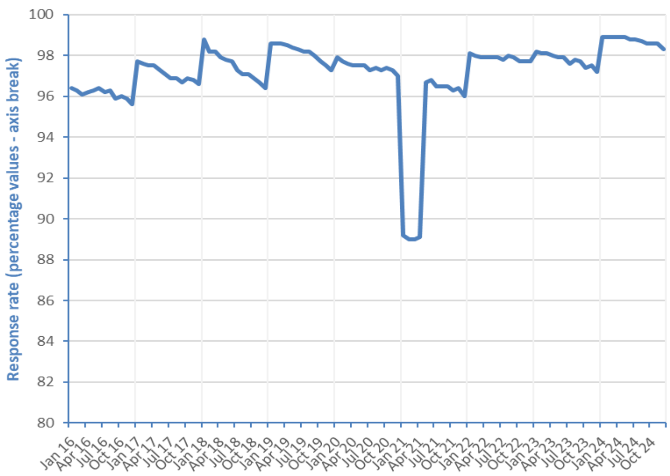
In every survey occasion, response rates are calculated as the ratio between the number of responding units and the number of units included in the sampling list. All analyses were undertaken using SAS Version 9.4.

5. Results

Both surveys considered in this study are characterised by a time series consisting of 108 observations, which represent the response rates recorded at each survey occasion between 2016 and 2024.

Figure 1 displays the time series of response rates for the OCC survey. The y-axis shows the response rate values on a scale starting at 80%, which better capture the fluctuations in the series. This survey has very high response rates, with a mean value of 97.2%, a minimum value of 89.0% (in March 2021) and a maximum value of 98.9% (from January to May 2024)[4].

Figure 1 – *Time series of response rates for the OCC survey. Years 2016-2024.*



Despite the evident stability in the response behaviour of the enterprises involved in the survey, a stepped profile can still be recognised, with a peak in January followed by a slight decrease in subsequent months each year[5].

The time series obtained by applying first-order differencing to the original time series is shown in Figure 2: the new series appears to be stationary, with values centred around zero, except for the peaks corresponding to January.

Table 2 shows the values of the autocorrelations of the observations with past data, from lag 1 to lag 24, for both the time series: the observed series and the first-order differenced series. Positive autocorrelations are coloured green, with the intensity of the colour decreasing as the autocorrelation value decreases. The observed series of response rates for the OCC survey shows relevant positive autocorrelation for lags up to eight, indicating that the series is non-stationary,

[4] The March 2021 drop in response rates is linked to COVID-19 emergency management, as it coincided with the final deadline for submitting previously suspended data, possibly impacting business participation.

[5] Please note that immediately prior to the data collection phase in January, the enterprises included in the sample receive an information letter from ISTAT, informing them of their involvement in the survey and providing all the details necessary for completing it.

whereas the autocorrelations between observations for the new time series confirm that the series is stationary and that a seasonality component is no longer present.

Figure 2 – *Time series obtained by applying first-order differencing to the response rates time series for the OCC survey. Years 2016-2024.*

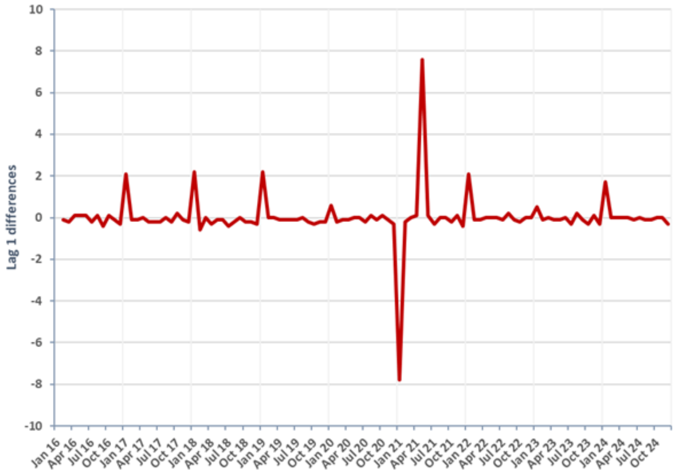


Table 2 – *Autocorrelations of the response rates and the first-order differenced time series for the OCC survey (lags 1 to 24).*

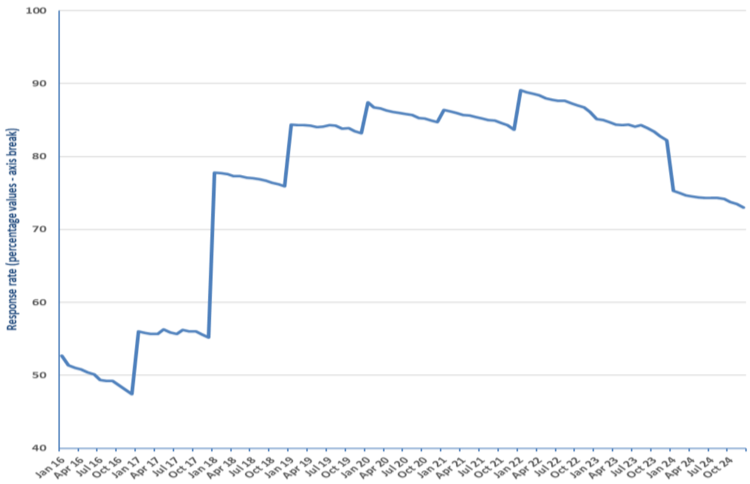
Time Series	Up to lag	Chi-sqr	DF	Pr > ChiSqr	Autocorrelations*					
					Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
Original time series	6	129.54	6	<.0001	0.789	0.577	0.366	0.168	0.148	0.141
	12	134.99	12	<.0001	0.135	0.132	0.086	0.038	0.000	-0.038
	18	140.11	18	<.0001	-0.047	-0.051	-0.060	-0.076	-0.102	-0.122
	24	162.41	24	<.0001	-0.131	-0.138	-0.157	-0.175	-0.188	-0.187
First-order differencing time series	6	20.52	6	0.0022	0.004	-0.011	-0.029	-0.423	-0.033	-0.004
	12	22.00	12	0.0375	-0.012	0.096	0.01	-0.022	0.003	-0.05
	18	22.27	18	0.2202	-0.016	0.007	0.017	0.025	-0.011	-0.027
	24	23.37	24	0.4983	-0.014	0.022	0.004	-0.012	-0.038	-0.075

Legend: 0.000-0.333 0.334-0.666 0.667-1.000

** Each column contains autocorrelations for lags multiples of 6 (e.g., "Lag 1" includes lags 1, 7, 13, and 19)*

The time series of response rates for the DETT survey is shown in Figure 3. The y-axis shows the response rate values on a scale starting at 40%, to better visualise the fluctuations in the series.

Figure 3 – *Time series of response rates for the DETT survey. Years 2016-2024.*

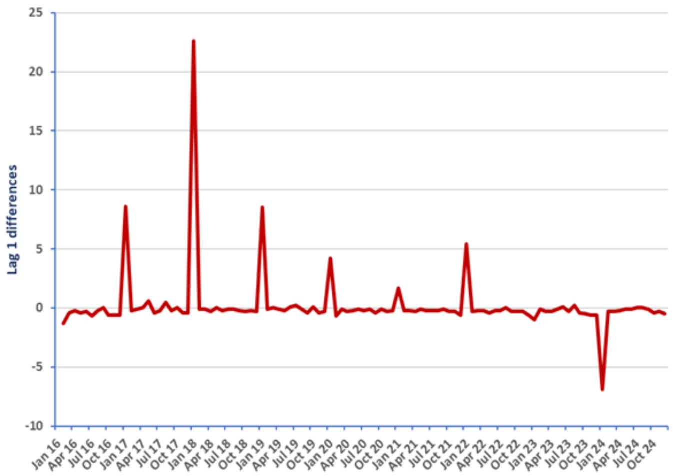


The series presents good response rates, especially since 2018, with an average value of 78.5%, a minimum of 47.5% in December 2016, and a maximum of 89.4% in January 2022. The series presents a stepped profile, more pronounced compared to the OCC response rates series. Once again, there is a peak in January[6], followed by a slight decrease in response rates in subsequent months each year.

Figure 4 displays the time series obtained by applying first-order differencing to the original one. The values in the new series fluctuate around zero, with very noticeable peaks occurring in January each year. This suggests that first-order differencing alone is insufficient to make the series stationary.

[6] Please note that immediately prior to the data collection phase in January, the enterprises included in the sample receive an information letter from ISTAT, informing them of their involvement in the survey and providing all the details necessary for completing it.

Figure 4 – *Time series obtained by applying first-order differencing to the response rates time series for the DETT survey. Years 2016-2024.*



As shown in Table 3, the original time series exhibits positive autocorrelations up to lag 24, confirming its non-stationary behavior. The autocorrelations of the transformed series still display significant values at lags 12 and 24, indicating the persistence of an annual seasonal component.

Table 3 – *Autocorrelations of the response rates and the first-order differenced time series for the DETT survey (lags 1 to 24).*

Time Series	Up to lag	Chi-sqr	DF	Pr > ChiSqr	Autocorrelations*					
					Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
Original time series	6	523.96	6	<.0001	0.965	0.929	0.893	0.858	0.822	0.785
	12	838.89	12	<.0001	0.749	0.713	0.676	0.639	0.603	0.567
	18	968.85	18	<.0001	0.519	0.472	0.426	0.381	0.338	0.293
	24	990.18	24	<.0001	0.250	0.208	0.166	0.124	0.084	0.045
First-order differencing time series	6	0.25	6	0.9997	-0.031	-0.023	-0.018	-0.007	-0.003	-0.02
	12	41.47	12	<.0001	-0.020	-0.005	-0.014	-0.025	-0.045	0.577
	18	42.07	18	0.0011	-0.037	-0.032	-0.035	-0.017	-0.005	-0.028
	24	50	24	0.0014	-0.021	-0.029	-0.018	-0.038	-0.070	0.220

Legend: 0.000-0.333 0.334-0.666 0.667-1.000

** Each column contains autocorrelations for lags multiples of 6 (e.g., "Lag 1" includes lags 1, 7, 13, and 19)*

6. Synthesis of Results and Research Outlook

The time series analysis approach applied to two different datasets yields seemingly different results. A previous study, which examined participation in

ISTAT short-term surveys using survival curve analysis (Binci *et al.*, 2025), observed that response burden acts as a deterrent to continued participation for small enterprises compared to larger ones, as the organizational structures of large enterprises allow them to bear such burden. Similarly, the results presented in this article suggest that the differences observed in the analysis of the two time series may be explained by differences in sampling and strategy design: the OCC survey involves only enterprises with more than 500 employees, whereas the DETT survey also includes smaller enterprises. However, the additional analysis presented in the appendix does not confirm this interpretation. Further research is needed—possibly extending the analysis to other surveys—to investigate the factors underlying the different seasonal patterns observed in response rates. Potential hypotheses to explore include characteristics of the types of enterprises in the samples, such as differences in the sectors in which they operate.

In the set of studies that Directorate for data collection **(**DCRD) has been carrying out for a number of years, the approach presented is aimed at bringing to light any respondent behaviour that, overall, may be repetitive over time regardless of the other characteristics that we have analysed in other works. Highlighting the non-stationarity of the series of respondents in the two surveys gives us the possibility to make targeted actions to support the survey and avoid costly and inefficient ones. It should always be stressed that the adjective 'costly' does not express a quantity whose unit of measurement is “euro” or “dollar”, but, in our case, “time”. This cost element enters predominantly into the surveys as it is the fundamental parameter in fieldwork and data collection. It is on this parameter that we must intervene to increase the quality of the information produced.

Appendix

We present the results of an additional analysis carried out following a suggestion from the referee, whom we thank. The aim is to assess whether the different seasonal patterns in response rates observed in the OCC and DETT surveys are influenced by differences in sampling design. In this follow-up, the analysis for the DETT survey refers to the same target group of enterprises involved in the OCC survey, focusing exclusively on enterprises with more than 500 employees. Table A1 displays the autocorrelation values for both the response rate series and the series obtained through first-order differencing, highlighting patterns similar to those calculated for the full sample (see Table 3); when assessing the robustness of the analysis, note that limiting the focus to enterprises with over 500 employees reduces observations from about 8,000 to 160.

Table A1 – *Autocorrelations of the response rates and the first-order differenced time series for the DETT survey – Only enterprises with 500+ employees (lags 1 to 24).*

Time Series	Up to lag	Chi-sqr	DF	Pr > ChiSqr	Autocorrelations*					
					Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
Original time series	6	491.07	6	<.0001	0.954	0.913	0.871	0.829	0.779	0.732
	12	741.04	12	<.0001	0.690	0.643	0.596	0.561	0.527	0.494
	18	856.00	18	<.0001	0.458	0.428	0.392	0.362	0.342	0.323
	24	911.44	24	<.0001	0.297	0.282	0.266	0.253	0.237	0.216
First-order differencing time series	6	4.840	6	0.565	-0.120	-0.072	-0.040	0.109	-0.092	-0.041
	12	24.560	12	0.017	-0.061	0.063	-0.179	-0.033	-0.039	0.346
	18	28.110	18	0.060	-0.109	0.039	-0.084	-0.022	-0.069	0.050
	24	47.100	24	0.003	-0.105	0.033	-0.091	0.035	0.005	0.339

Legend: 0.000-0.333 | 0.334-0.666 | 0.667-1.000

** Each column contains autocorrelations for lags multiples of 6 (e.g., "Lag 1" includes lags 1, 7, 13, and 19)*

References

BELLINI G., BINCI S., DE GAETANO L., MONETTI F., PAPA P. 2019. Effects of Istat CDC (Centralised Data Collection) approach on the reduction of the Total Survey Error: experiences in data collection implementation field. In *Workshop on Statistical Data Collection 'New Sources and New Technologies'*.

BINCI S., BONTEMPI K., CECCARELLI C., MONETTI F., ROSSETTI F. 2025. Burden analysis for short-term business surveys respondents. *Rivista Italiana di Economia Demografia e Statistica* , Vol. LXXIX – No. 3 Luglio – Settembre 2025 pp. 251-262.

CECCARELLI C., GRAZIANI C., LORIGA S., SALVATORE M.A., SPIZZICHINO A. 2016. Gli errori campionari nelle serie storiche destagionalizzate: alcune analisi sui dati mensili della rilevazione sulle forze lavoro. *Rivista Italiana di Economia Demografia e Statistica*, Vol. LXX – No. 4 Ottobre-Dicembre 2016 pp. 55-66.

DAGUM E, *et al.*1999. X11ARIMA version 2000: Foundations and User's Manual.

FAZIO N.R., MURGIA .M, NUNNARI A. 2013. The Business statistical Portal: a new way of organizing and managing data collection processes for business surveys in Istat. *Working Papers 14*, Seminar on Statistical Data Collection. Unece - Conference of European statisticians, 25-27 September 2013, Geneva, Switzerland.

Katia BONTEMPI, Italian National Institute of Statistics, katia.bontempi@istat.it
Claudio CECCARELLI, Italian National Institute of Statistics, clceccar@istat.it
Francesca ROSSETTI, Italian National Institute of Statistics, frrosset@istat.it;