SURVEYS ON TRADE SECTOR: A COMPARISON BETWEEN QUALITATIVE AND QUANTITATIVE INDICATORS[1]

Maria Rita Ippoliti, Luigi Martone, Fabiana Sartor, Graziella Spera

1. Introduction

Qualitative surveys are extremely relevant for the short-term economic analysis and exceptionally useful in the building process of cyclical indicators of the economic development as they gather information quickly and explore topics that are not investigated by quantitative surveys. Therefore, these variables are often considered complementary to official macroeconomic data, and they are both used for analysing the short-term economic development. This approach underlines the important role of assessments and expectations for the economic decision makers. As qualitative series are so important for policy makers, it is necessary to assess the reliability of results of qualitative surveys, measuring the distance with official quantitative series and estimating the similarity of signals coming from qualitative surveys and short-term movement in the economic activity (Koopmans, 1947; Zarnowitz, 1992). There has been much debate on this topic in the economic literature, mainly with regard to the manufacturing sector, while the discussion remained poor concerning the other economic sectors, such as services, trade and construction (Crosilla and Leproux, 2007; Crosilla *et al.*, 2009; Martelli and Rocchetti, 2007). This paper investigates the trade sector, comparing the evolution of business confidence in retail trade with two quantitative indicators, which measure the different sectors of section G: retail trade (Division 47) and turnover in services (Division 45) from 2010 until 2019. Business confidence in trade sector involves enterprises classified in the NACE Rev. 2 section G[2], apart from the wholesale trade classified in Division 46, the itinerant retailers classified in Group 47.8 and non-store retailers, stalls and markets classified in the 47.9. The aim of the project is to assess the relationship between qualitative and quantitative

---

[1] Though the article is the result of a joint work, the single paragraphs are attributed as follows: paragraph 1 and 2.2 to Fabiana Sartor; paragraph 2 to Graziella Spera; paragraph 3 and 4 to Luigi Martone; paragraph 5 and 6 to Maria Rita Ippoliti. The published articles are exclusively expressing the authors' opinions; Istat shares no responsibility for the published contents.

[2] Section G includes Division 45 (Wholesales trade, retail trade and maintenance of motor vehicles), Division 46 (Wholesales trade, except wholesales trade of motor vehicles) and Division 47 (Wholesales trade, except wholesales trade of motor vehicles).

components concerning the retail trade sector: analysing the trend in time series and testing the ability of the qualitative index to timely catch the economic development. Therefore, the present paper examines and compares official qualitative retail trade to test the reliability of qualitative series. In order to make the comparison more effective we purified the quantitative series from the long-term trend. This leaded us to use the seasonal difference of the logarithm of the quantitative series as a transformation to remove seasonality and trend from the quantitative series. However available data are insufficient to guarantee strong scientific evidence: considered time series are indeed short (Division 45 of turnover has been surveyed since 2010) and survey samples include different types of enterprises depending on the specific features of each survey (i.e. Retail trade survey struggles to keep track of store closures and openings). This paper proposes the development of a new quantitative index, based on appropriate transformations of the two quantitative indices and including all activity sectors covered by the qualitative indicator. This new ad hoc index interestingly leads to a higher correlation with the qualitative index and the application of ADL models and Granger causality test for a comparative analysis confirms the reliability of the new index. This study is structured as follows: paragraph 1 presents the background to the study, introducing the qualitative and quantitative surveys involved; paragraph 2 provides a description of methods and tools used, it presents a first graphic analysis comparing quantitative and qualitative indicators and it introduces a new "ad hoc" indicator covering all sectors involved in the business confidence. Paragraph 3 shows the cyclical analysis, focusing on turning points and cross-correlation. Paragraph 4 shows a further comparative analysis of the involved indicators, using ADL models and Granger causality test. Finally, the last paragraph presents our conclusions.

2. Trade surveys

2.1. *Business Confidence Survey in Retail Trade*

Business Confidence Survey in Retail Trade is part of a joint project harmonised at European level and coordinated by the European Commission, it allows to have information on the economic evolution of retail trade (NACE Division G, except for Division 46 - Wholesale trade, except of motor vehicles and motorcycles and for Group 47.9 - Retail trade not in stores, stalls or markets including retail sales via mail order or via Internet).[3] The survey asks enterprises to express their opinions (judgements and expectations over the following 3 months)

---

[3] Divisions of NACE Section G involved in the survey are Division 45 (Wholesale and retail trade and repair of motor vehicles and motorcycles) and Division 47 (Retail trade, except of motor vehicles and motorcycles).

about the main economic variables (orders, prices, employment), giving therefore an updated overview on the evolution of the sector. Respondents are requested to state their consideration on their total sales in the last three months, on their current volume of stock and on prices charged by their suppliers. They are also invited to express how they expect the volume of orders, the employment, the prices they charge and total sales to change in the next three months. Information about enterprises of the Business Confidence Survey in Retail Trade are taken from a panel of approximately 1.000 commercial enterprises. The theoretical sample is stratified by enterprise employment size class (1-2 employees, 3-5; 6-999; at least 1.000 employees), by geographical area (North-West, North-East, Centre, South and the Islands) and by main activity (45.1 sales of motor vehicles; 45.2-45.4 maintenance of motor vehicles and sales of accessories; 47.1, 47.2 retail sales of food, drinks and tobacco; 47.3 retail sales of automotive fuel; 47.4-47.7 retail sales of other goods). The sampling scheme depends upon randomisation for enterprises with less than 1.000 employees and upon a census placement for all units with 1.000 employees and above. The data processing method sets out the estimate of the frequency percentages of each reply option relating to each item of the questionnaire. For this purpose, the processing of the micro data is based on a double weighting system: a) the frequencies of each reply option are firstly weighted using the number of employees declared by the enterprise at the time of the interview (internal weight); b) subsequently fixed weights reflecting the distribution of the added value of the reference sector (external weight) are used. Since March 2015, the aggregation procedure uses an external weighting structure derived from the added value at factor cost referred to 2012. Each variable is measured calculating balances as percentage differences between favourable and unfavourable responses. Weighted balances are seasonally adjusted if needed. The Index of Business Confidence in Retail Trade is calculated as the arithmetic mean of seasonally adjusted balances based upon opinions and expectations on sales and upon judgments on volume of stocks (the above-mentioned values have inverse signs).

2.2. Retail Trade Survey

Monthly Retail Trade Survey collects data from enterprises that mainly operate in the retail trade sector (except for sale of motor vehicles and motorcycles and sale of automotive fuel). Therefore, the survey covers the retail trade sector only partially (NACE Rev. 2, G 47 - Retail trade, except of motor vehicles and motorcycles not including automotive fuel)[4]. Data here considered include Nace

---

[4] According to NACE Rev. 2, Retail trade (Division G47) is first classified by type of sale outlet (retail trade in stores: groups 47.1 to 47.7; retail trade not in stores: groups 47.8 and 47.9). For retail

Group G 47.3, Retail sale of automotive fuel, which ISTAT does not disseminate at national level.[5] Estimates of Retail Trade Survey provide useful information on consumer spending. Monthly indices on retail trade are released at national level, consistently with the European Union Regulations concerning short-term statistics (see European Regulations n. 1165/98 and n. 1158/2005)[6]. Retail trade data are collected from a sample of about 8.000 enterprises having at least a legal unit that operates in Italy. The sample is stratified considering the following variables: main activity according to NACE Rev. 2 and employment size class (1-5, 6-49 and at least 50 employees). According to the sampling scheme, enterprises with less than 50 persons employed are selected at random, while any enterprise with equal or more than 50 persons employed is included in the survey. All the enterprises employing 50 or more persons add up to more than 1.100 units. The sampling design of the survey rotates some units out and rotates new units in each year (belonging to employment size classes 1-5 and 6-49 only) to share burden and refresh the sample. This aspect gains relevance when comparing retail trade indicators with qualitative data, as the Retail Trade Survey struggles to keep track of quick evolution of stores' closures and openings. According to their distribution channel, enterprises in the retail trade sample can be classified into large-scale distribution, small-scale distribution, internet sales and non-store sales. Within the weighting structure of Monthly Retail Trade Index (base=2015), large scale-distribution accounts for 46.4% of total turnover, while small-scale distribution reaches 48.0% of total turnover. Retail trade indices are calculated as weighted means of the sub-indices of each stratum. Concerning the calculation method of the indicator, the synthetic index numbers are constructed as weighted averages of indices related to the domains identified by the intersection of the stratification variables (main activity and employment size). To calculate aggregate indices up to the retail trade total, the Laspeyres index is used. The weights are based on turnover data from SBS of the year 2015. Value of sales indices measures the retail trade turnover over time at current prices and, therefore, incorporates the effects in changes of quantity sold and prices. In order to determine estimates on the volume of sales, value of sales indices is processed to allow removing price effects on turnover, using the Harmonised index of consumer prices (HICP). Monthly data are first revised in the following month after publication (which occurs 38 days past the reference time period). Estimates are then subject to a second revision, which occurs on annual basis and replaces the provisional estimates with the final indices.

---

sale in stores, there exists a further distinction between specialised retail sale (groups 47.2 to 47.7) and non-specialised retail sale (group 47.1).

[5] See Annex 1 for a list of involved surveys and covered Nace sectors.

[6] See https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1998R1165:20120621:EN:PDF

2.3. Turnover in Services Survey

Quarterly Turnover in Services Survey focuses on short-term dynamics of value of services sold by enterprises, which operate in the services field as a main economic activity. These dynamics incorporate growth rates of volumes and prices, giving a prompt information. The methodology, the breakdown and the frequency of releases are defined by the European Regulations on short-term statistics (see European Regulations n. 1165/98 and n. 1158/2005 and footnote 9). This is a sample survey, enterprises are the units of observation, while turnover and average number of persons employed by the economic unit in the reference quarter are the variables of interest. Quarterly indicators on turnover in services are calculated for each economic sector, setting 2015 as the base year; these indices are then aggregated according to the Laspeyres formula that uses a weighting structure reflecting the proportion of turnover by economic sector in the base year 2015. The survey does not include retail trade; however it includes retail trade of motor vehicles and motorcycles (NACE Rev. 2 Division G 45)[7]. The latter, which is the sector analysed in this study, hereafter will be referred to as FAS45; the sample for this survey includes 1.627 enterprises, starting from 2010. For this economic sector, the method applied to build domain indices first calculates the quarterly turnover levels using weights to estimate the population value from the sample. Then the growth rates are calculated. The weights used as expansion factors are obtained as the ratio between the total reference population (taken from Asia archive) and the sample size for each stratum.

3. Data analysis: tools and methods

Our first exercise investigates the relationship between FIDCOM and the two quantitative indicators VEN and FAS45. Results show that FIDCOM and VEN series diverged considerably in almost the entire period except for the last part, around 2019, when the two series appear to be aligned. Interestingly, FIDCOM and FAS45 show a slightly better alignment throughout the whole period. Dissimilarity between the two indices can be caused by the different sectoral coverage, by the different sampling scheme of the three surveys and by the different features of the seasonal component, which is higher in the Retail Trade Index. The comparison becomes more effective when the cyclical component of the reference economic variable is used; consequently, the quantitative series should be purified from the long-term trend. Although different detrending methods can present a good

---

[7] Sectors covered by the survey are: section G - Wholesale and retail trade and repair of motor vehicles and motorcycles (except division G47 - Retail trade), section H – Transportation and storage, section I – Accommodation and food service activities, J – Information and communication, L – Real estate activities, M – Professional, scientific and technical activities, N – Administrative and support service activities.

approximation, they do not estimate the same cyclical components (Canova, 1999), causing issues for an accurate estimate at the end of the reference time period. In the light of the above, in this exercise we used the seasonal difference of the logarithm of the quantitative series as a transformation to remove seasonality and trend from the quantitative series. Although this transformation can be seen as a rather simple method to remove these components from quantitative series, it also allows elimination of estimation problems at the end of the period; moreover, this approach represents the most used method in forecasting models (i.e. Bruno and Lupi, 2004). The different sectoral coverage emerged in the previous analyses made it necessary to create an ad hoc index (also referred to as ADHOC) involving both the surveyed sectors (Divisions 45 and 47), calculated as a weighted average of the two indices, where weights reflect the distribution of turnover across the different economic sectors (see Annex 1). The graphical analysis of FIDCOM and the cyclical component of the ADOCH index (Figure 1) reveals that the two series tend to be significantly aligned keeping a small gap between them starting from year 2017.

Figure 1 – *Comparison between FIDCOM and ADHOC (quarterly data, seasonally adjusted, year 2010-2019)*
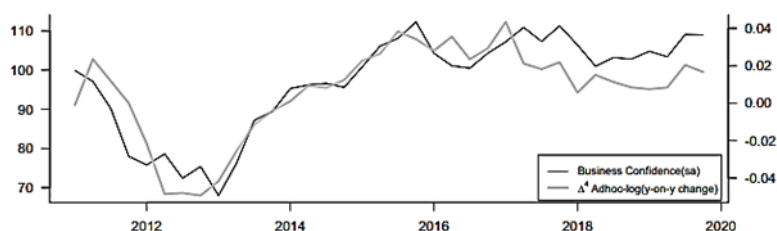
Source: Elaboration on ISTAT data.

In order to verify previous results, we calculated the correlation between FIDCOM and the logarithmic transformations of the three considered quantitative indices. Values in table 1 show that the correlation between business confidence and the new ad hoc index is higher (0.87) than the correlation between the indicators of each quantitative survey and Business Confidence Survey (VEN 0.43 and FAS45 0.79).

4. Cyclical analysis

For a more in-depth analysis of the comparison between the qualitative indicator of the survey on business confidence and the quantitative reference indicator, it is also useful to consider the behaviour of the series at turning points and the ability of the qualitative series to timely detect and sometimes anticipate

the turning points in the cyclical movement of the quantitative series. The Bry-Boschan method (1971) was used to detect turning points and to identify expansion and recession phases, catching the moments when the fluctuation of the business cycle changes in direction, from recessive to expansive (peak) and from expansive to recessive (trough). According to this approach, a business cycle is completed when it goes through a single expansion (minimum-maximum) and a single contraction (maximum-minimum) in sequence, therefore it is defined as the period between two troughs. The applied procedure searches for turning points observing the following constraints: each complete cycle lasts at least five quarters and each single cyclical phase (including alternation of expansive and recessive phases) lasts at least two quarters [8] Turning points that are too close to the extreme values of the series are excluded by the procedure. Table 1 shows the history of turning points detected by the procedure.

Table 1 – *History of turning points: VEN FAS45 and ADHOC (cyclical component) FIDCOM.*

	Δ ^{4}LogVen	Δ ^{4}LogFas	Δ ^{4}LogAdhoc	Confidence
Number of cycles	2	1	2	2
Number of turning points	6	3	5	5
N. of common turning points with Δ ^{4}LogVen	-	-	-	5
N. of common turning points with Δ ^{4}LogFas	-	-	-	3
N. of common turning points with Δ ^{4}LogAdhoc	-	-	-	5
Turning points				
Max				
Min	201301	201203	201204	201301
Max	201503	201602	201503	201504
Min	201603		201603	201603
Max	201701		201701	201704
Min	201801	201901	201901	201802
Max	201803			
Average lag(-)/lead(+)(in quarters)				
Compared to Δ ^{4}LogVen	-	-	-	-0.8
Compared to Δ ^{4}LogFas	-	-	-	2.3
Compared to Δ ^{4}LogAdhoc	-	-	-	-0.4
Agreement index				
Compared to Δ ^{4}LogVen	-	-	-	0.72
Compared to Δ ^{4}LogFas	-	-	-	0.67
Compared to Δ ^{4}LogAdhoc	-	-	-	0.78

Source: Elaboration on ISTAT data.

Because of the shortness of the series, the procedure detected only a small

---

[8] With regard to monthly series, constraints impose that each cycle lasts at least fifteen months and each cyclical phase lasts at least five months.

number of cycles (2 for VEN, ADHOC and FIDCOM and 1 for FAS45). Specifically, two complete FIDCOM cycles last approximately eleven quarters and they are composed by eight quarters of expansion and three quarters of contraction. Now and then, the FIDCOM cycle reproduces the reference series' turning points with a slight lag: averagely one-quarter lag when compared to VEN and two-quarter lead when compared to FAS45. Remarkably, this result suggests that, despite the little amount of available data, the correlation between the composite indicator and the qualitative indicator (having the same sectoral coverage) is higher than the correlation between the indicators of each quantitative survey and the qualitative indicator.

5. Econometric exercise

5.1. ADL MODELS

In order to compare qualitative and quantitative data, besides correlation method, it is possible to apply ADL models, forecasting models with autoregressive distributed lag such as

$$y_t = \alpha + \Phi(L)y_{t-1} + \beta(L)x_t + \varepsilon_t \tag{1}$$

where x_t is the indicator deriving from qualitative survey (confidence) and y_t is the quantitative variable (VEN, FAS45 or the ADHOC composite indicator). These models allow analysing the dynamic features of $\beta(L)$ and $\Phi(L)$, the effectiveness of the model, the structural stability (considering shortness of the available time series) and the forecasting ability of treated indicators. Table 2 shows the results of the three quantitative indicators while Figures 2 graphically represent recursive residuals.

Table 2 – *ADL models for a comparison between qualitative and quantitative data.*

| ADL model	Coef	P-value	R²	F test	Box-Pierce test	Granger-test
VEN ~ L(VEN) + FIDCOM			0.77	0.00	0.58	0.10
VEN	0.9	0.00				
FIDCOM	0.01	0.14				
FAS45 ~ L(FAS45) + FIDCOM			0.97	0,00	0.26	0.00
FAS45	0.89	0.00				
FIDCOM	0.20	0.00				
ADHOC ~ L(ADHOC) + FIDCOM			0.97	0,00	0.16	0.00
ADHOC	0.93	0.00				
FIDCOM	0.05	0.00				

Source: Elaboration on ISTAT data.

ADL models allow examining the dynamic features of the coefficients related to the indicators involved in the model, therefore these models allow the analysis of the impact on quantitative indicator when qualitative indicator changes by one unit.

Given (1), then:

$$y_t = \frac{\beta(L)}{1-\Phi(L)}\, x_t \qquad\qquad (2)$$

where β(L) is the immediate effect (how much the quantitative indicator changes when the confidence indicator changes by one unit) and 1-Φ(L) is the measure of how fast the long-run effect emerges, which is given by the ratio between the two components. The latter depends on both the confidence indicator and VEN / FAS45 / ADHOC indicators at time t-1. Our results in Table 4 shows that FAS45 is the fastest indicator, as its immediate effect score is 0.20 while the score for VEN is 0.01 and the score for ADHOC is 0.06. Moreover, the long-run effect score for FAS45 is higher than the scores of other indicators (1.66 versus 0.11 for VEN and 0.75 for ADHOC). The analysis of the ability of models to accurately fit data reveals that the FAS45-FIDCOM model and the ADHOC-FIDCOM model fit better than the VEN-FIDCOM model (R2 0.97 versus 0.77). F test results are significant for all models. Furthermore, the surveys' structural stability was examined, acknowledging shortness of available time series, through recursive residuals test (see Figure 2).

Figure 2 – *Recursive residuals VEN, FAS45 and ADHOC.*
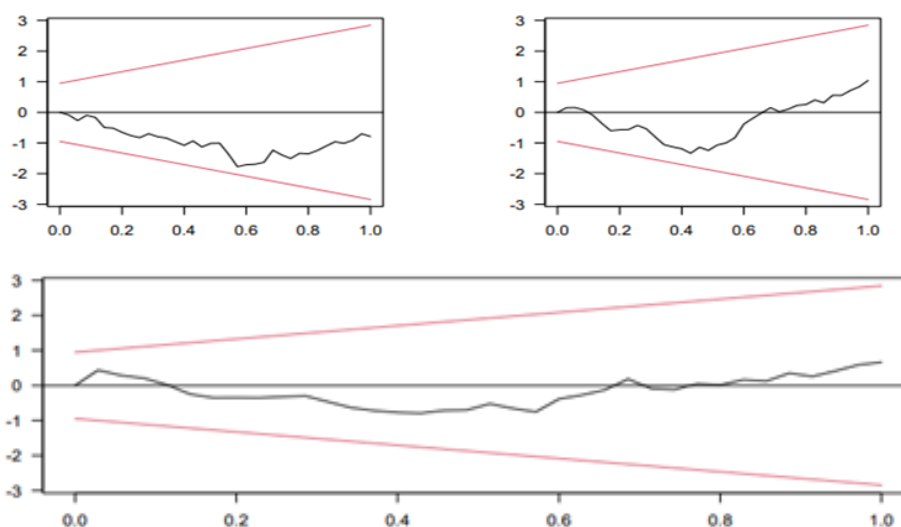
Source: Elaboration on ISTAT data

Graphs show that all indicators remain within the bands defined by the critical values (level α=0.05), ensuring the consistency of the parameters in the reference period. Finally, we analysed the ability of the qualitative series to forecast the quantitative series. In fact, if a series x_t causes series y_t, then past values of the first series should contain information that helps predict the second series. Results of this exercise show that time series of confidence in trade sector helps to predict future values of the FAS45 indicator and the ADHOC combination (p-value 0.0 for both of them), whereas it does not help in determining values for VEN (p-value 0.9). Acknowledging the purpose of this study focused on evaluating the performance of the new ad hoc indicator, despite the appreciable concordance between FIDCOM and FAS45, the result of the above-mentioned analysis emphasizes the validity of the ad hoc index.

5.2 Prediction of consumption in National Accounts

After analysing ADL models, it is useful to verify the forecasting ability of the treated indicators. We carried out a performance evaluation for VEN and ADHOC indicators referring to data on Household final consumption expenditure provided by National Accounts (CONS). Despite the shortness of the available time series, Granger causality test confirms that in trade sector the ADHOC indicator is successful in forecasting the retail trade indicator (*p-value* 0.03 versus 0.07). This finding validates the usefulness of a composite index, which performs better than the individual indicators and has a good predictive ability.

6. Conclusions and perspectives

Confidence surveys provide timely information on the short-term economic evolution for different economic sectors; they also are exceptionally useful in the building process of cyclical indicators of the economic development as they gather information quickly and explore topics that are not investigated by quantitative surveys. The study focuses on the comparison between the confidence index in trade sector and the equivalent quantitative indices: the Retail Trade Index and the Turnover in Services Index (Division 45 only, Wholesale and retail trade and repair of motor vehicles and motorcycles) concerning years 2010-2019. The analysis of the relationship between FIDCOM and the two quantitative indicators and the comparison between FIDCOM and the cyclical components of the quantitative indicators highlighted a good alignment among the different series, which diverged towards the end of the reference period. The diverse sectoral coverage made it necessary to create an ad hoc index (ADHOC) involving both the surveyed sectors (Divisions 45 and 47), calculated as a weighted average of the two indices, where weights reflect the distribution of turnover. The analysis revealed that the

qualitative series and the new ad hoc quantitative series were more aligned, keeping a smaller gap between them towards the end of the period. Moreover, the correlation found between business confidence in retail sector and the new ad hoc index was higher than the correlation between business confidence and the indicators of each quantitative survey. In order to investigate the ability of the qualitative indicator to timely detect movements or to predict fluctuations of the economic evolution, a turning points analysis was carried out. The results clearly show that expansive and recessive phases of the economic cycle are rather coincident between confidence index and the quantitative series. On the whole, FIDCOM appears to have a higher correlation with the cyclical movement of the ADHOC indicator. This evidence suggests that the correlation between a composite indicator and a qualitative indicator with the same sectoral coverage is higher than the correlation between a qualitative indicator and the indicators of each quantitative survey. In order to compare qualitative and quantitative data, besides correlation method, we applied ADL models and results show a good effectiveness of the considered models and a reliable structural stability of the surveys. Moreover, the Granger causality test revealed the ability of the trade confidence time series to predict future values of both FAS45 indicator and ADHOC combination, whereas it does not help in determining values for VEN. As a further econometric exercise, we assessed the ability of VEN and ADHOC indicators to predict data on Household final consumption expenditure provided by National Accounts. Our findings confirmed the ADHOC composite indicator to be successful in forecasting the retail trade series and to perform better than each individual indicator. Therefore, results proved that the ADHOC indicator helps to predict future values of data referred to our country's economic activity.

References

BRUNO G., LUPI C. 2004. Forecasting industrial production and the early detection of turning points, *Empirical Economics*, Vol. 29, No. 3, pp. 647-671.

BRY, G., BOSCHAN, C. 1971. Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*, NBER Technical Papers*, National Bureau of Economic Research.

CANOVA, F. 1999. Does detrending matter for the determination of the reference cycle and the selection of turning points?, *The Economic Journal*, Vol. 109, No. 452, pp. 126–150.

CROSILLA L., LEPROUX S. 2007. Leading Indicators on construction and retail trade sectors based on ISAE survey data, *ISAE working papers,* Istituto di Studi e Analisi Economica.

CROSILLA L., LEPROUX S., MALGARINI M. e SPINELLI F. 2009. Model-based Confidence Indicators for the Italian Economy, *ISAE working papers*, Istituto di Studi e Analisi Economica.

KOOPMANS T. 1947. Measurement without theory. *Review of Economic Statistics*. Vol. 29, No. 3, pp. 161–172.

MARTELLI B.M. e ROCCHETTI G. 2007. Cyclical features of the ISAE business services series, *ISAE working papers*, Istituto di Studi e Analisi Economica.

MOORE, G.H. e SHISKIN J. 1967, Indicator of business expansion and contractions, *NBER Occasional Paper,* National Bureau of Economic Research.

ZARNOWITZ V. 1992. *Business Cycles, Theory, History, Indicators, and Forecasting*. Chicago*:* The University of Chicago Press.

SUMMARY

Data from qualitative surveys are often considered complementary to quantitative data when analysing the short-term economic evolution. This paper aims at investigating the retail trade sector, comparing data stemming from the different kinds of surveys conducted by Istat, providing qualitative and quantitative data on monthly and quarterly basis. This study compares the Retail Trade Confidence Index (NACE rev.2, G45 and G47), a business sentiment indicator, along with quantitative data from Retail Trade Survey (NACE rev.2 G47) and Turnover in Services Survey (NACE rev.2 G45) concerning years from 2010 through 2019. Despite the shortness of the available time series and the different features of the samples involved, interesting findings emerged when observing the performance of a new 'ad hoc' index (which was built including both the economic activities examined by the quantitative surveys) and when comparing the above-mentioned indicators applying ADL models.

_____

Maria Rita IPPOLITI, Istat, marippoliti@istat.it
Luigi MARTONE, Istat, martone@istat.it
Fabiana SARTOR, Istat, sartor@istat.it
Graziella SPERA Istat, grspera@istat.it