40th Meeting of the Voorburg Group

Statistics Netherlands

The Hague, The Netherlands

21 - 25 September 2026

The use of Electronic Invoicing data for the Retail Trade Sector

Giuseppe Amato[(*)], Romina Ciavardini[(*)], Luigi Martone[(*)],
Mariagrazia Moschetta[(*)], Debora Pontesilli[(*)], Francesca Tuzi[(*)]

[(*)] Istat, Italian National Institute of Statistics. The paper is the result of joint research by the six authors. Introduction, sections 3, 4.1 and 5.1 by F. Tuzi, section 1 by G. Amato e M. Moschetta, section 2 by L. Martone, section 4.2 by D. Pontesilli, section 5.2 and conclusion by R. Ciavardini.
Statistical analyses, including data processing and the selection and application of statistical methods, were performed by F. Tuzi.

Contents

Introduction

1. Recent experiences with administrative data in short term business statistics

2. Current work: the Retail Trade Survey

2.1 The Current Survey Framework

2.2 Introducing Administrative Data into the Survey

3. Methodological Framework and Data sources

3.1 The methodological approach

3.2 Data sources and Comparability Assessment

4. Main results

4.1 Diagnostic approach

4.2 Graphical approach

5. Future strategy for Administrative-Data-Based Estimation

5.1 Potential impact of the transition to administrative-data-based estimation

5.2 Implications for the RTS sample design

Conclusion

Introduction

The use of administrative data (AD) for statistical purposes is a key part of the recent strategy adopted by Istat to modernise business statistics, based on information already available within public agencies. Besides improving the efficiency of statistical production, the administrative sources offer significant potential to reduce the statistical burden on enterprises, particularly micro and small businesses that are traditionally covered by direct surveys.

Over the last year, the regular provision of AD, electronic invoicing (EI) and electronic receipts (ER), by the Italian Revenue Agency, the institution responsible for collecting fiscal information, has become increasingly established. This has enabled Istat to increase the use of this administrative source in producing short-term business statistics. Following an extensive experimental phase within the Turnover Survey on Services (TSS), the use of AD was introduced for a selected group of sectors in January 2026. The decision was based on promising results in terms of data quality, coverage and comparability. At this initial stage, a cautious approach was adopted to ensure the stability of the statistical production process while assessing the contribution of AD in a production environment. This gradual introduction enabled the new source to be progressively integrated and its ability to meet the requirements of official statistics to be assessed over time.

Following the experience gained through the TSS survey, new research activities have been launched to assess the potential extension of AD to other fields of short-term business statistics. Current experimental work focuses on the Retail Trade Survey (RTS) and aims to evaluate the suitability of AD for producing turnover indices and to identify the conditions for their future integration into the statistical production system. This contribution presents the results of the ongoing assessment framework, with a focus on comparing survey-based estimates with administrative turnover measures, and on evaluating the potential of AD to support the production of turnover indices within the RTS, with the aim of reducing reliance on direct survey information over time.

The contribution is structured as follows. Section 1 describes the main features of the transition from survey-based information to AD within the TSS survey, with particular attention to its implications for reducing the response burden on enterprises. Section 2 provides an overview of the RTS, with particular attention to the response burden on enterprises and the practical implication of introducing AD in the production process. Section 3 describes the methodological framework adopted to assess the suitability of tax sources for producing short-term statistics on retail trade and the main characteristics of the data sources used for the simulations. Section 4 illustrates the most relevant results obtained from an analytical and graphical perspective. In the last case, with particular attention to some activity sectors and the main limitations and critical issues encountered. Section 5 explores the future strategy for AD-based estimation, focusing on its impact on the statistical burden on businesses and on the extension of its application to the full population within the scope of the survey. Finally, the contribution offers some concluding remarks.

1 – Recent experiences with administrative data in short term business statistics

The Italian monthly Turnover Survey on Services is a well-established direct survey that produces high-quality short-term indicators in compliance with Regulation (EU) 2019/2152 and within the required dissemination deadlines. The survey provides monthly turnover indices, in both value and volume, at the NACE Rev. 2 division/group level, based on data collected up to 2025 from a sample survey of approximately 25,600 enterprises operating in the main market services sectors. While the survey ensures the production of

reliable statistics, it also entails significant operational costs and response burden for both Istat and the sampled enterprises.

Last year, a first experimental result was presented on the use of AD data as a source for the compilation of monthly short-term economic statistics on services turnover. These results were made possible through an agreement between the Italian Revenue Agency and Istat, under which Istat receives the complete monthly dataset of EI. Since September 2025, the agreement has been extended to include ER, further expanding the potential of AD sources for the production of official short-term economic statistics.

Thanks to this collaboration between Istat and the Italian Revenue Agency, starting from the reference month of January 2026, the Italian monthly survey on services turnover has integrated AD from EI and ER. This innovation has enabled the replacement of direct survey data with administrative sources for 68 of the 277 sampling strata[1], mainly composed of small and very small enterprises (2 to 20 employees), reducing the survey sample by more than 43 per cent: from about 25,500 to 14,500 enterprises.

The 68 selected sampling strata were identified on the basis of two main criteria. First, they exhibited a high degree of consistency between the information collected through the direct survey and the corresponding AD derived from EI and ER. The **degree of agreement** between the two sources was evaluated through different **criteria**:

- **Graphical analysis** at different aggregation levels (e.g., scatter plots, short-term dynamics) to explore structural relationship.
- **Diagnostic tests** at "stratum" level including tests of equality between means, regression test on proportionality between TSS and AD turnover.
- **Accuracy indicators** of the degree of similarity of month-on-month variations estimated on the differences between TSS and AD at "stratum" level including correlations, simple differences, MAE, sign concordance.

Second, they were mainly composed of small enterprises, for which the statistical response burden is generally higher than larger firms.

For these strata, the integration strategy did not simply consist of replacing survey observations with the corresponding administrative records for the sampled units. Instead, the survey component was entirely replaced by AD covering the full target population of enterprises included in the Italian Statistical Business Register (Asia). As a result, the estimates for these strata are now based on approximately 820,000 enterprises rather than on a sample.

As shown in Table 1, the *Accommodation and Food Service Activities* experienced the most significant changes following the integration of AD. This sector recorded the largest reduction in response burden, with 72.7 per cent of the sampled enterprises (1,865 out of 2,566) replaced by AD, corresponding to 79.7 per cent of the total weight of the sector. *Real Estate Activities* also experienced a substantial reduction in response burden, with 68 per cent of the sampled enterprises (1,676 out of 2,463) replaced by AD, accounting for 49.9 per cent of the total weight of the sector. The *Wholesale and Retail Trade; Repair of Motor Vehicles and Motorcycles* recorded a 64.8 per cent reduction in the survey sample, corresponding to 5,162 out of 7,969 enterprises being replaced by AD. Overall, the innovation covers 41.2 per cent of the total weight of Section G, reaching 46.5 per cent for *Wholesale Trade* and 17.2 per cent for *Wholesale and Retail Trade and Repair*

[1] Sampling stratum refers to the breakdown of the reference population by economic activity (Group NACE Rev.2), turnover class and employee class.

of Motor Vehicles and Motorcycles. Different situation for *Transportation and Storage*; *Information and Communication*; *Professional, Scientific and Technical Activities*; and *Administrative and Support Service Activities* (Sections H, J, M and N) that continue to rely more extensively on direct data collection due to their structural characteristics. Nevertheless, the integration of AD still represents a significant share of the estimates, accounting for 24.5 per cent of the total weight in Section H, 22.1 per cent in Section M, 15.3 per cent in Section N and 8.1 per cent in Section J.

Table 1 – Services Turnover – percentage weights by NACE Section – enterprises and turnover.

NACE Section	TSS sample 2026			Composition % (enterprises)		Weights % (turnover)	Population level
	Target	Sample	AD	Sample	AD		
G – Wholesale and retail trade	***7,969***	***2,807***	***5,162***	***35.2***	***64.8***	***41.2***	***244,638***
G45 Wholesale and retail trade and repair of motor vehicles and motorcycles	2,160	855	1,305	39.6	60.4	17.2	35,570
G46 Wholesale trade, except of motor vehicles and motorcycles	5,809	1,952	3,857	33.6	66.4	46.4	209,068
H–N – Other services	***17,637***	***11,792***	***5,845***	***66.9***	***33.1***	***27.8***	***575,230***
H – Transportation and storage	3,850	3,027	823	78.6	21.4	24.4	25,343
I – Accommodation and food service activities	2,566	701	1,865	27.3	72.7	79.7	171,811
J – Publishing, broadcasting, and content production and distribution activities	2,094	1,969	125	94.0	6.0	8.1	11,109
L – Real estate activities	2,463	787	1,676	32.0	68.0	49.9	98,022
M – Professional, scientific and technical activities	3,103	2,562	541	82.6	17.4	22.1	212,152
N – Administrative and support service activities	3,561	2,746	815	77.1	22.9	15.3	56,793
G–N – Total services	**25,606**	**14,599**	**11,007**	**57.0**	**43.0**	**35.3**	**819,868**

Note: 'Population' refers to the average number of enterprises, identified through AD sources, for estimating the TSS index during the first five months of 2026.

This methodological innovation provides several important advantages. In addition to eliminating non-response, which is one of the main sources of bias in business surveys, it virtually removes sampling error for the affected strata, since the estimates are based on exhaustive administrative information. Consequently, the integration of AD is expected to improve the accuracy and robustness of the turnover indicators while substantially reducing the response burden for businesses and the operational costs associated with data collection and represents the first step towards a broader integration AD into the production of short-term business statistics.

2 – Current work: the Retail Trade Survey

2.1 – The Current Survey Framework

The Monthly Retail Trade Survey is one of ISTAT's main short-term business surveys for monitoring developments in household goods consumption and represents a key component of Italy's system of short-term economic statistics.

The indicators produced measure monthly changes in the value of retail sales generated by retail enterprises and provide an essential tool for analysing the business cycle, both at the national level and within the European Statistical System. The survey complies with the requirements laid down in the **Regulation (EU) 2019/2152 on European Business Statistics (EBS)**, contributing to the production of harmonised indicators that meet high standards of quality, reliability and timeliness.

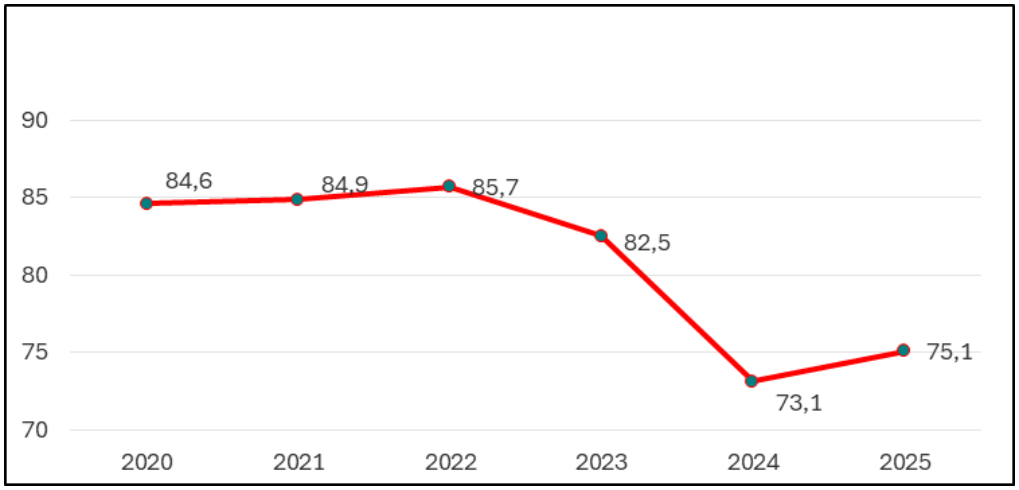
The data are disseminated in three forms—unadjusted, seasonally adjusted and calendar-adjusted—according to a predefined release schedule. Preliminary estimates are published **38 days** after the end of the reference month, while final estimates are released after **60 days**. In addition, revised data are disseminated on **a half-yearly basis**.

The survey covers enterprises classified under **Division 47 of NACE Rev. 2**, corresponding to retail trade. It distinguishes between **specialised enterprises**, whose activity is concentrated on a specific product category, and **non-specialised enterprises**, which sell a broader range of goods. The target population includes both traditional store-based retailing and retail activities carried out through alternative sales channels, such as e-commerce and other forms of non-store retailing. In addition to overall retail sales indices, the survey produces indicators by type of retail outlet and by major commodity group, together with **volume indices**, which are obtained by deflating current value indices.

Data are collected monthly through an electronic questionnaire completed via the **Business Statistical Portal** by a stratified sample of approximately **8,000 enterprises**, selected from the Asia Register. The sample design is stratified by economic activity and enterprise size class in order to ensure representativeness across the main estimation domains. Large enterprises, which account for a substantial share of total sector turnover, are included with certainty, whereas the remaining units are selected using probability sampling with the periodic rotation of part of the sample. The collected data undergo a comprehensive production process, including validation, consistency checks, outlier detection, treatment of non-response and imputation, with the aim of ensuring the completeness, comparability and overall quality of the estimates.

One of the survey's main strengths is the high quality of the indicators it produces. This is ensured not only by the sampling design and estimation methods adopted, but also by the continuous monitoring of both data collection and data quality, which guarantees that the information used in the statistical production process is complete, consistent and available in time to meet the dissemination schedule. Up to 2023, this monitoring system enabled the survey to maintain a stable response rate of around **84 per cent**.

Figure 1 – Overall response rate for the years 2020 –2025 (*percentage values*)



In recent years, however, enterprise participation has progressively declined despite the systematic use of reminder procedures and respondent support services. Smaller enterprises increasingly encounter difficulties in completing the questionnaire because of limited staff resources and the growing reporting burden associated with participation in multiple statistical surveys. As a result, from 2024 onwards the response rate declined markedly (Figure 1), suggesting that the traditional measures adopted to encourage participation may no longer be sufficient to ensure the high-quality standards required for official statistics. These developments highlight the need to explore alternative approaches capable of preserving both the quality and the timeliness of the indicators while reducing the reporting burden placed on enterprises. Against this background, the use of AD represents a promising opportunity.

2.2 – Introducing Administrative Data into the Survey

The availability of EI and ER data provides an important opportunity to rethink the traditional data collection model. Both sources are updated with a frequency consistent with the requirements of short-term economic statistics and provide a high level of detail at the enterprise level. These characteristics make it possible to redesign part of the statistical production process by developing an integrated system in which survey data and administrative sources are used in a complementary manner. At this stage, the objective is not to replace the direct survey entirely, but rather to identify the estimation domains in which AD can provide estimates of equal or higher quality than those obtained through direct data collection. The integration of administrative sources should therefore be regarded as a process of survey redesign aimed at preserving the quality of official statistics while, at the same time, reducing the reporting burden placed on enterprises.

The use of AD, however, requires several preliminary considerations. Data derived from EI and ER do not provide uniform coverage of the entire retail trade sector. Under the current legislation (Article 2 of Presidential Decree No. 696 of 21 December 1996, as subsequently amended, together with the Ministerial Decrees of 13 February 2015 and 27 October 2015), certain activities are exempt from the obligation to transmit ER. These include e-commerce, mail-order sales, tobacco products, motor fuels, and the sale of newspapers and periodicals. In addition, specific reporting arrangements apply to some activities. For example, in parts of the second-hand goods market (NACE Rev. 2 code 47.79) only the commission charged on sales made on behalf of third parties is recorded through the electronic receipt system, while direct sales are subject to the **Margin Scheme** (which is used to avoid double taxation by calculating VAT only on the profit—that is, on the difference between the selling price and the purchase price—rather than on the entire amount). A further challenge concerns **VAT Groups**, where ER are transmitted centrally by the representative entity of the group, potentially limiting the attribution of data to individual enterprises.

Furthermore, the information content of the AD differs from that collected through the survey. While the survey collects, for non-specialised enterprises, information on the distribution of sales by commodity group, the administrative sources mainly provide the total turnover of the enterprise, classified according to the fiscal characteristics of the transactions. This difference in information content represents an important aspect when assessing the substitutability of AD. The pilot study was therefore carried out considering the different characteristics of the estimation domains. Specialised retail activities exhibit a closer correspondence between the principal economic activity and the goods sold, whereas for non-specialised enterprises the absence of a commodity breakdown requires a more careful assessment of the suitability of AD.

The analysis was conducted at the level of **sampling strata**, defined by the combination of economic activity and enterprise size class, in order to identify the contexts in which the use of AD is most appropriate[2]. The assessment combines the results of statistical quality indicators computed for each sampling stratum with graphical analyses performed at the highest level of data disaggregation (firm-period). The findings are expected to support the identification of those sectors in which AD can provide a statistically reliable alternative to direct data collection.

Table 2 – Expected evolution of the statistical production process towards the use of administrative data.

Current process	**Future scenario**
All sampling strata are covered through the survey questionnaire.	Only those sampling strata that cannot be replaced continue to be surveyed directly.
Direct data collection from approximately 8,000 enterprises.	Reduced direct data collection through the replacement of suitable sampling strata.
Statistical production relies primarily on survey data.	Integrated statistical production based on survey data, electronic invoicing and electronic receipt data.
The reporting burden is spread across the entire sample.	The reporting burden is concentrated exclusively in the domains where direct data collection remains necessary.

3 – Methodological Framework and Data sources

3.1 – The methodological approach

In order to evaluate the feasibility of replacing survey data on RTS with AD, a **multidimensional framework** was adopted. The approach assessed the comparability of data sources across two complementary dimensions: **structural** and **dynamic**. Each dimension should capture distinct aspects of the differences between the survey data and the AD. The measurement of structural indicators is intended to provide a numerical representation of deviations in level, inconsistencies in scale and distributional biases. Conversely, dynamic indicators, which emphasise short-term movements, are expected to capture temporal discrepancies, smoothing effects and volatility patterns. Although the dynamic component is likely to be more relevant in the context of short-term statistics, the structural component may reveal interesting cases of approximation failures. Due to the multidimensional nature of the evaluation problem, the structural and dynamic components are obtained by aggregating a set of statistical measures. Specifically, for the structural component a Lin's concordance correlation coefficient (LCC) is used as a joint measure of precision and accuracy[3]; scale consistency is evaluated using total ratios between AD and survey data (*Ratio*); distributional differences are measured using Kolmogorov–Smirnov distance statistic (KS), while the mean absolute percentage error (sMAPE) summarises the overall magnitude of discrepancies[4].

[2] The strata are defined based on a combination of economic activity sector (at the five-digit level of the ATECO classification system) and firm size (defined as having '1–5', '6–49' or '50+' employees).

[3] Differently from the Pearson correlation coefficient, which is used to evaluate linear association, the Lin's concordance correlation provides a composite measure of precision and accuracy by explicitly penalising deviations from the bisector line.

[4] A symmetrised MAPE version (sMAPE) is used to measure average relative pointwise deviations between survey and administrative estimates. Its symmetric formulation avoids selecting a source as the reference, treating both series equally in the comparison.

The dynamic component combines three complementary measures of short-term similarities: the correlation between monthly growth rates (*Corr*) shows how closely changes move together over time; directional agreement is measured as the fraction of periods in which the two series exhibit changes with the same sign (*Sign*), capturing consistency in local dynamics; finally, volatility (*Vol*) measures the extent to which the two series show similar variability over time.

Each similarity measure is first transformed into a normalised indicator ranging from 0 to 1, in order to ensure comparability across the different dimensions. The structural measures are then combined into the **Structural Comparability Index (SCI)** using a multiplicative aggregation rule:

$$SCI = \prod_{k=1}^{4} S_k^{w_k}$$

where S_k denotes the k-th normalised structural similarity measures and w_k the associated weights, satisfying $\sum_k w_k = 1$. Specifically, the concordance measure (S_{CCC}) captures deviation from the bisector line; the scale (level) equivalence ratio (S_{Ratio}) measures proportional bias in aggregate levels; the distributional similarity measure (S_{KS}) captures differences in the empirical distribution functions; and the symmetrised MAPE measure (S_{sMAPE}) summarises overall comparability.

Similarly, the dynamic measures are aggregated into the **Dynamic Comparability Index (DCI)** defined as:

$$DCI = \prod_{j=1}^{3} D_j^{v_j}$$

where D_j denotes the j-th normalised dynamic similarity measures and v_j the associated weights, satisfying $\sum_j v_j = 1$. These include a co-movement indicator based on monthly growth rates (S_{Corr}), a measure of directional agreement (S_{Sign}) and a measure of volatility (S_{Vol})[5].

A weighted geometric aggregation is used to ensure no compensatory effects, meaning that a poor performance in a component cannot be fully offset by a strong performance in another. This form of aggregation is preserved in the construction of the overall **Composite-Comparability Index (CCI)**, defined as follows:

$$CCI = \sqrt{SCI \cdot DCI}$$

which combines structural consistency and dynamic consistency into a single indicator, ranging from 0 to 1.

3.2 – Data sources and Comparability Assessment

The database used for this experiment was constructed by integrating data from the RTS with the AD provided by the Italian Revenue Agency. The AD combines information from EI and ER. EI component includes both business-to-business (B2B) transactions and business-to-consumer (B2C) transactions documented through invoices, while ER capture B2C sales recorded through the ER system. Therefore, the AD enables two

[5] The weighting scheme is defined as follows, for the structural component, weights of 0.30 are assigned to S_{CCC} and S_{Ratio}, while weights of 0.20 are assigned to S_{KS} and S_{sMAPE}. For the dynamic component, the corresponding weighs are 0.4 for S_{Corr}, 0.35 for S_{Vol} and 0.25 for S_{sign}.

alternative turnover measures to be constructed: **total turnover (B2B + B2C)** and **turnover generated by B2C transactions only**.

To provide a comprehensive assessment of comparability, the analysis was carried out at different levels of data disaggregation. Firstly, retail activities were analysed separately for **specialised** and **non-specialised retail sectors**, reflecting their differing customer compositions and business models. Secondly, the comparability of each sector was evaluated using both the **B2C turnover generated by transaction** and the **total turnover (B2B + B2C)**. This distinction is particularly relevant in retail trade because the proportion of turnover generated by final consumers can vary considerably between sectors. While some activities are almost exclusively consumer-oriented, others have a substantial proportion of sales to businesses. Comparing survey estimates using these two definitions helps to determine whether discrepancies between survey and AD are driven by the transaction mix or by differences between the two data sources themselves. These differences are expected to be more pronounced in non-specialised retail sectors.

Integration of the two data sources (RTS and AD) was carried out at the most disaggregated level of information: the transferor's tax code. The experimental analysis was conducted on a monthly basis over the period from January 2025 to May 2026. A “processed” version of the original AD was used, for which preliminary consistency checks were performed on the data used to construct the turnover variable and the data was grouped by tax code.

Data from the survey and administrative sources were matched using the most recent version of the Asia Business Register (BR) when the last RTS samples were drawn (for the 2023 reference year).

The empirical analysis was conducted at two complementary levels. The first level focused on the **survey sample**, where RTS observations were directly matched with the corresponding administrative records, allowing a detailed comparison to be made between the two sources for the same reporting units. This analysis provides the basis for evaluating the feasibility of replacing survey data with administrative information in the short-term and addresses the issue of survey non-response. The second level extended the analysis to the **target population** of the RTS. All enterprises within the survey scope were identified through the BR and classified into the same strata as those of the survey design.

Due to the nearly complete coverage of the AD source, the analysis at the sampling unit level shows not only the suitability of the source for replacing survey-based estimates, but also its potential to reduce or minimise sampling error.

The experiment involved the subset of companies simultaneously present in the RTS survey and in the AD, for each month between January 2025 and May 2026. The AD data is composed only by B2C component and B2B+B2C. In many cases this differentiation has practically no effect, but in some sectors the difference, as expected, is significant. Overall, the simulation was conducted on approximately 130,200 “firm-period” records from the RTS sample, of which around 85 per cent relates to specialised enterprises. Of these, 92,200 were respondents in the considered period.

The total turnover by period and enterprise, derived from the two sources, was aggregated according to the classification criteria adopted in the survey (strata). Specifically, enterprises were classified by **size class and at the 5-digit level of economic activity sector** (**ATECO**).

A preliminary outlier detection procedure was applied to identify and remove anomalous observations, mainly arising from potential reporting or compilation errors. This step aimed at improving the consistency of the turnover measures and ensure that the assessment of comparability between the two sources was

unaffected by data quality issues unrelated to the underlying economics. For the different levels of data aggregation, complementary approaches were adopted:

1. **Diagnostic approach**. The Composite-Comparability Index (CCI) was used to assess comparability, capturing both the structural and dynamic dimensions of agreement.
2. **Graphical approach.** Scatter plots were examined to evaluate the relationship between administrative and survey-based measures at **the sampling unit level**. A qualitative assessment was carried out of the consistency of their trends by analysing the month-on-month (m-o-m) changes derived from the two sources at stratum level, as well as identifying any short-term discrepancies.

With respect to short-term changes, m-o-m were estimated at both the sample and population levels. As for the second level of analysis, a comparison was carried out between all the companies within the scope of the RTS survey, according to the Asia 2023 BR. The AD retail trade totals, at the "firm-period" level, were aggregated according to the stratification criteria adopted by RTS survey.

Overall, short-terms changes at the population level were estimated for approximately 380,000 firms. The results of the analysis are presented at an aggregate level of size class and activity sector. However, they are also available for each of the 211 strata defined by the RTS survey.

4 – Main results

4.1 – Diagnostic approach

Table 3 shows the fraction of strata within each of the five comparability ranges defined by the **Composite Comparability Index (CCI)** and its two components: the **Dynamic Comparability Index (DCI)** and the **Structural Comparability Index (SCI)**. To simplify the interpretation of the results, the numerical values of the indicators were converted into qualitative comparability categories. These categories were defined according to the following thresholds: *Very High* (> 0.90), *High* (0.75-0.90), *Moderate* (0.60-0.75), *Low* (0.40-0.60) and *Very Low* (≤ 0.40).

To consider the effects of the different components of EI and ER, diagnostics were estimated based on three types of turnover definitions:

- **Type A**: total turnover B2B + B2C (EI) and turnover generated by B2C transactions only (ER).
- **Type B**: total turnover B2C (EI) and turnover generated by B2C transactions only (ER).
- **Type C**: turnover generated by B2C transactions only (ER).

The results show that the dynamic dimension exhibits a higher degree of agreement than the structural one. According to the DCI, around 50 per cent of the strata fall into the *Very High* category for all three aggregation types, while only a small fraction is classified as *Low* or *Very Low*. This suggests that the administrative sources are generally able to approximate the short-term dynamics of the survey estimates, regardless of the adopted turnover definition.

On the other hand, the structural dimension exhibits a more heterogeneous pattern. Type A achieves the highest structural comparability, with over 60 per cent of the strata classified as *High* or *Very High*, and only around 11 per cent in the *Very Low* category. When limiting the analysis to the B2C component (Type B), a decrease in structural agreement is observed, mainly due to the proportion of *Very Low* cases which has almost doubled. The decline is particularly marked for Type C, which relies exclusively on ER: almost a third of the strata fall into the *Very Low* category, whilst less than a quarter achieve a *Very High* level of comparability.

Table 3 – Fraction of strata by range of comparability indexes.

	Composite Comparability Index			Dynamic Comparability Index			Structural Comparability Index		
	Type A	Type B	Type C	Type A	Type B	Type C	Type A	Type B	Type C
Very High	23.7	30.3	28.0	50.2	48.3	43.1	14.2	25.6	23.7
High	46.4	26.5	19.0	30.8	32.2	26.5	46.4	20.9	18.0
Moderate	13.3	19.0	15.6	12.8	12.8	16.6	18.0	19.4	11.4
Low	11.4	11.4	14.2	5.2	5.7	9.5	10.0	13.3	14.2
Very Low	5.2	12.8	23.2	0.9	0.9	4.3	11.4	20.9	32.7

Note: (Type A) EI (B2B+B2C) and ER; (Type B) EI (only B2C) and ER; (Type C) only ER.

The findings suggest that excluding invoicing data (EI) primarily affects the structural consistency of the two aggregates, whilst the ability to capture short-term dynamics remains relatively robust.

The overall CCI, which combines the structural and dynamic dimensions through non-compensatory aggregation, reflects this behaviour. Type A achieves the best performance overall, with approximately 70 per cent of the strata classified as *High* or *Very High*. This proportion decreases for Types B and C, along with a corresponding increase in the share of strata with *Very Low* comparability. These findings confirm that the inclusion of comprehensive AD, particularly EI data, substantially improves the overall suitability of AD as a replacement for survey estimates, primarily due to improvements in structural comparability rather than dynamic comparability.

Table 4 reports for the CCI, the distribution of strata by size class and NACE group, highlighting the differences in behaviour across economic sectors and levels of aggregation.

At the size class level, dimension '50+' shows the highest level of comparability, with approximately 70 per cent of strata classified as *High* or *Very High* across all aggregation types. Size class '1-5' also performs well, although comparability declines from Type A to Type C.

Table 4 – Fraction of strata by size class/NACE Group and range of comparability indexes (CCI).

Agg.	N. of strata	Number of records			Composite-Comparability Index								
					Very High/High			Moderate			Low/Very Low		
		Type A	Type B	Type C	Type A	Type B	Type C	Type A	Type B	Type C	Type A	Type B	Type C
							Size class						
1-5	73	58,782	58,788	52,818	76.7	53.4	42.5	9.6	24.7	20.5	13.7	21.9	37.0
6-49	78	12,599	12,601	11,278	61.5	46.2	34.6	15.4	20.5	16.7	23.1	33.3	48.7
50+	60	22,005	22,019	20,940	73.3	75.0	68.3	15.0	10.0	8.3	11.7	15.0	23.3
							NACE group						
47.1	19	16,277	16,277	16,161	63.2	68.4	68.4	10.5	15.8	15.8	26.3	15.8	15.8
47.2	29	9,757	9,757	9,589	55.2	41.4	41.4	24.1	20.7	20.7	20.7	37.9	37.9
47.4	8	6,118	6,118	5,653	37.5	50.0	37.5	12.5	12.5	25.0	50.0	37.5	37.5
47.5	43	10,145	10,145	9,292	83.7	60.5	34.9	7.0	18.6	14.0	9.3	20.9	51.2
47.6	18	12,626	12,635	11,663	77.8	66.7	61.1	11.1	11.1	16.7	11.1	22.2	22.2
47.7	73	27,528	27,534	25,744	75.3	64.4	56.2	12.3	19.2	15.1	12.3	16.4	28.8
47.8	6	4,572	4,574	4,479	50.0	50.0	50.0	33.3	33.3	33.3	16.7	16.7	16.7
47.9	15	6,363	6,368	2,455	60.0	20.0	6.7	13.3	26.7	0.0	26.7	53.3	93.3

Note: 'Size class' refers to number of employees. 'NACE group' refers to the NACE Rev. 2 system at the three-digit level.

By contrast, size class '6-49' exhibits the weakest performance, with a marked reduction in highly comparable strata and a corresponding increase in the share of *Low* or *Very Low* comparability, particularly for Type C.
Overall, the results suggest that larger units are characterised by a higher degree of comparability between survey and AD. Smaller units, particularly those in size class '6-49', exhibit greater discrepancies. Furthermore, comparability systematically decreases from Type A to Type C across all size classes, suggesting that the exclusion of invoicing information reduces the AD's ability to accurately reproduce the survey estimates. As shown by the indicators reported in Table 5, these differences are mainly driven by the structural component of comparability, while the dynamic component remains comparatively stable across size classes.
Results by NACE group reveal substantial heterogeneity across economic sectors. NACE groups such as 47.5, 47.6 and 47.7 show high comparability under Type A, indicating a strong alignment between administrative and survey data. In contrast, sectors such as 47.4 and 47.8 exhibit lower comparability.
It is worth noting that, in some cases, the closest agreement with the retail sales survey is obtained when considering only the B2C component, regardless of whether it is derived solely from ER or from the broader B2C component. This likely reflects the fact that including B2B transactions introduces operations that are not fully consistent with the retail sales concept measured by the survey. For instance, in non-specialised retail trade (NACE group 47.1), invoiced transactions may include B2B operations rather than purely retail sales to final consumers. The same pattern is observed for the largest size class, which is characterised by a higher concentration of enterprises operating in non-specialised retail trade.

Table 5 – Fraction of strata by size class/NACE Group and range of comparability indexes (SCI and DCI).

Agg.	Structural Comparability Index						Dynamic Comparability Index					
	Very High/High			Low/Very Low			Very High/High			Low/Very Low		
	Type A	Type B	Type C	Type A	Type B	Type C	Type A	Type B	Type C	Type A	Type B	Type C
						Size class						
1-5	69,9	42,5	37,0	19,2	35,6	47,9	87,7	79,5	71,2	2,7	5,5	9,6
6-49	51,3	37,2	30,8	26,9	44,9	61,5	71,8	74,4	60,3	10,3	11,5	24,4
50+	61,7	63,3	61,7	16,7	18,3	26,7	85,0	90,0	80,0	5,0	1,7	5,0
						NACE group						
47.1	57,9	47,4	47,4	21,1	21,1	21,1	68,4	78,9	78,9	15,8	10,5	10,5
47.2	27,6	27,6	31,0	31,0	51,7	51,7	82,8	79,3	79,3	3,4	0,0	0,0
47.4	25,0	25,0	25,0	50,0	37,5	37,5	62,5	62,5	62,5	12,5	37,5	37,5
47.5	74,4	51,2	32,6	11,6	34,9	62,8	88,4	83,7	60,5	4,7	7,0	23,3
47.6	77,8	55,6	55,6	11,1	27,8	33,3	77,8	72,2	72,2	16,7	16,7	22,2
47.7	71,2	57,5	54,8	19,2	24,7	37,0	83,6	84,9	75,3	1,4	2,7	5,5
47.8	50,0	50,0	50,0	50,0	50,0	50,0	66,7	66,7	66,7	16,7	16,7	16,7
47.9	40,0	13,3	6,7	26,7	60,0	93,3	80,0	80,0	40,0	6,7	0,0	33,3

Note: 'Size class' refers to number of employees. 'NACE group' refers to the NACE Rev. 2 system at the three-digit level.

Overall, the level of agreement remains statistically significant, even in cases where the structural component is weaker, and AD are still able to capture the main characteristics of the survey estimates. This aspect is particularly relevant in the context of short-term statistics, where the ability to capture short-term dynamics is often more critical than an exact agreement in levels. Data comparability remains high across sectors, confirming the AD's ability to capture short-term fluctuations despite some persistent structural differences.

In summary, the evidence from Tables 4 and 5 suggests that, although structural comparability remains an important issue when using AD, dynamic coherence represents a key strength and a crucial dimension for assessing their suitability for statistical production.

The availability of complete AD, notably EI data, has been shown to lead to a substantial improvement in the agreement between administrative aggregates and survey-based estimates in terms of levels, although the dynamics in the series appear to be predominantly preserved across different data combinations.

In this framework, the assessment and selection of AD sources will consider both their capacity to capture dynamics over time and their structural comparability with survey-based estimates. However, particular emphasis will be placed on the dynamic dimension, given its crucial role in ensuring that AD adequately reflect economic developments over time.

4.2 – Graphical approach: focus on some economic activity sectors

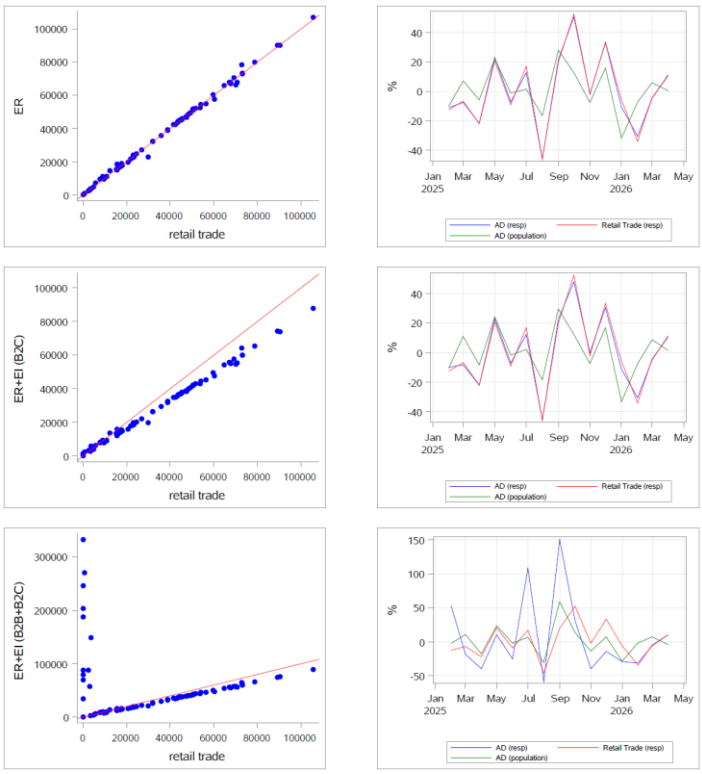
Although the analysis was conducted on all 211 sample strata, a few selected case studies are presented below to highlight specific characteristics of both the AD and the data derived from the survey.

The examples illustrate both situations where there is a high degree of consistency between the two sources (A, B, C) and cases where differences arise due to sector-specific characteristics, regulatory aspects, or variations in the coverage of the administrative sources under consideration (D, E, F).

A) Virtuous case 1: *Retail sale of yarn and haberdashery*

The first scatter plot shows a good correlation between the survey data and that derived solely from ER. The addition of EI (B2C component) slightly weakens this correlation, but it is primarily with the inclusion of EI (B2B component) that a deterioration in this correlation is observed.

Figure 2 - Scatter plots, m-o-m changes, diagnostics – ***Yarn and haberdashery*** stratum (2_47512)



Indicator	Type A	Type B	Type C
Share of ER on total	64.1	99.0	100.0
SCI	-	**0.922**	**0.950**
CCC	-0.059	0.998	0.941
Ratio	1.821	1.015	1.005
KS dist	0.203	0.114	0.128
MAPE sym	0.340	0.268	0.031
DCI	**0.651**	**0.984**	**0.993**
Corr	0.520	0.994	0.996
Vol	2.265	0.958	0.981
Sign conc.	0.875	1.000	1.000
CCI	-	**0.952**	**0.971**

Notes: Type A - EI (B2B+B2C) and ER; Type B - EI (only B2C) and ER; Type C - only ER. SCI and DCI in normalised form; their components in the original formulation.

It can therefore be inferred that businesses report to the survey only the portion of revenue related to retail trade, in accordance with the instructions provided to users in the guide for completing the retail sales questionnaire.

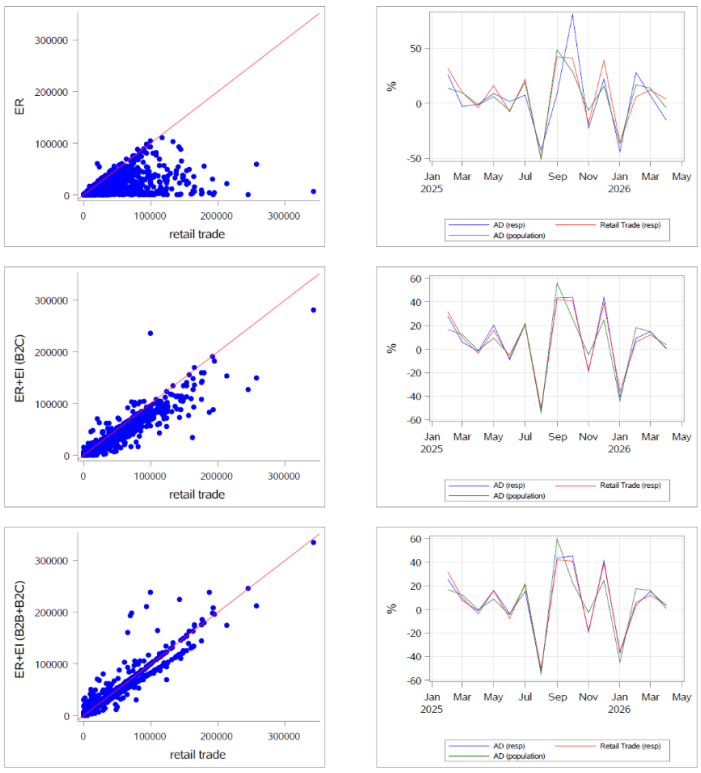
The results are confirmed by the month-on-month variations and diagnostics, which highlight a slight deterioration in the alignment between survey data and AD when the B2C component is included (CCI decreases from 0.971 to 0.952). The relationship is no longer optimal when the share of B2B transactions is included, as also shown by the diagnostics on Type A, where the CCI cannot be estimated due to the value of a SCI component.

It is interesting to note that the month-on-month variations estimated at population level follow the same pattern as the variations calculated from the survey data, but with the peaks in the series being more smoothed out.

B) Virtuous case 2: *Retail sale of household furniture*

This is a sector where EI plays a significant role compared to ER alone, about 70 per cent if B2B transactions are included, 65 per cent considering only B2C transactions. In fact, when purchasing furniture, an invoice is generally issued as a tax document. This allows customers to claim the tax credits under Italian law and helps to manage down payments and deliveries. The final assembly of items is often postponed beyond the order date.

Figure 3 - Scatter plots, m-o-m changes, diagnostics – **Household furniture** stratum (1_47591)



Indicator	Type A	Type B	Type C
Share of ER on total	30.4	36.1	100.0
SCI	**0.889**	**0.927**	**0.394**
CCC	0.927	0.937	0.274
Ratio	1.154	0.972	0.348
KS dist	0.061	0.023	0.306
MAPE sym	0.199	0.215	0.824
DCI	**0.990**	**0.976**	**0.879**
Corr	0.993	0.995	0.817
Vol	1.025	1.068	1.116
Sign conc.	1.000	1.000	0.813
CCI	**0.938**	**0.951**	**0.589**

Notes: Type A - EI (B2B+B2C) and ER; Type B - EI (only B2C) and ER; Type C - only ER. SCI and DCI in normalised form; their components in the original formulation.

The comparison charts between AD and survey data by responding units, as well as the diagnostic analyses of the CCI indicator—which confirm the alignment between directly collected data and AD—show a low level of correspondence for the structural component (SCI) when considering only ER (Type C). This component reaches high values when EI is included.

The dynamic component (DCI) related to month-on-month variations is maximized when the total revenue of companies in the sector (ER + B2C + B2B) is considered. When EI is included, the month-on-month changes estimated at the sample level, using both survey and AD, closely match those computed for the entire population, indicating a high degree of consistency in the underlying dynamics.

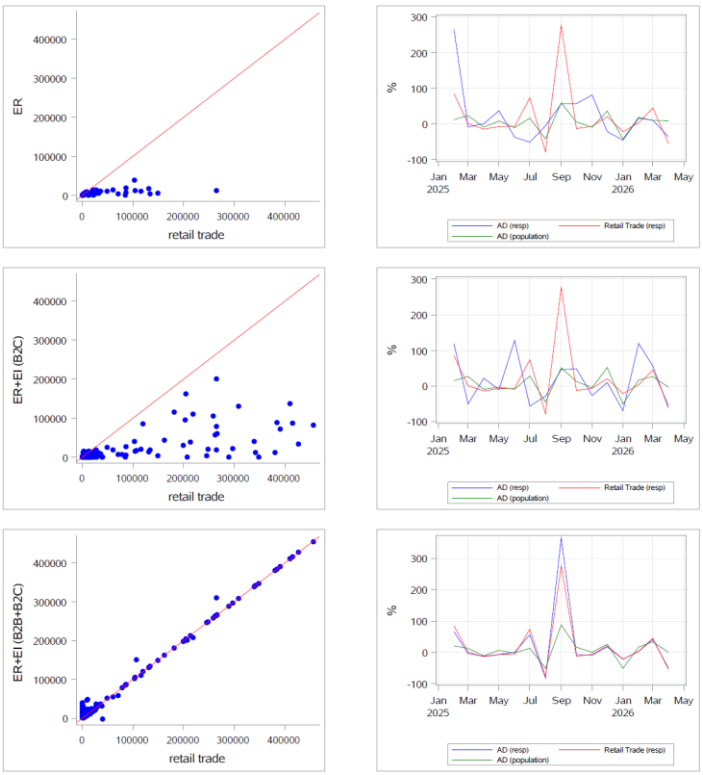
C) *Virtuous case 3*: *Retail sale of office furniture*

This sector exhibits the same pattern observed for stratum "1_47591", with B2B component of EI accounting for a larger share of transactions. ER account for only a negligible share of total transactions (2.8 per cent), indicating that invoices are the primary tax documents in the office furniture retail sector, with B2B invoices representing the majority of sales. Although this activity is classified as retail trade, transactions are primarily documented through invoices issued to VAT-registered units rather than to final consumers. This reflects the fact that, although ATECO code 47.78.1 refers to the retail sale of furniture and should include sales for personal consumption, businesses are classified under retail trade whenever the type of customer served (business or final consumer) cannot be distinguished.

The comparison charts between survey data and AD, along with the analysis of the structural component (SCI), confirm that the use of ER alone does not cover the office furniture sector and that, when reporting data to the retail sales survey, businesses include their total revenue—which encompasses invoices issued to other businesses as well as those related to transactions with end consumers. In contrast, the dynamic component performs very well across all data estimates.

It is interesting to note that using census data would make it possible to obtain a more consistent time series that is less affected by sampling variability, thereby smoothing out the sudden spikes and drops found in the series derived from survey data.

Figure 4 - Scatter plots, m-o-m changes, diagnostics – **Office furniture** stratum (1_47781)



Indicator	Type A	Type B	Type C
Share of ER on total	2.8	13.3	100.0
SCI	**0.809**	**0.378**	**0.080**
CCC	0.973	0.359	0.017
Ratio	1.243	0.277	0.033
KS dist	0.174	0.286	0.426
MAPE sym	0.516	1.112	0.934
DCI	**0.911**	**0.713**	**0.724**
Corr	0.985	0.286	0.332
Vol	1.235	0.824	0.954
Sign conc.	0.938	0.688	0.563
CCI	**0.859**	**0.519**	**0.241**

Notes: Type A - EI (B2B+B2C) and ER; Type B - EI (only B2C) and ER; Type C - only ER. SCI and DCI in normalised form; their components in the original formulation.

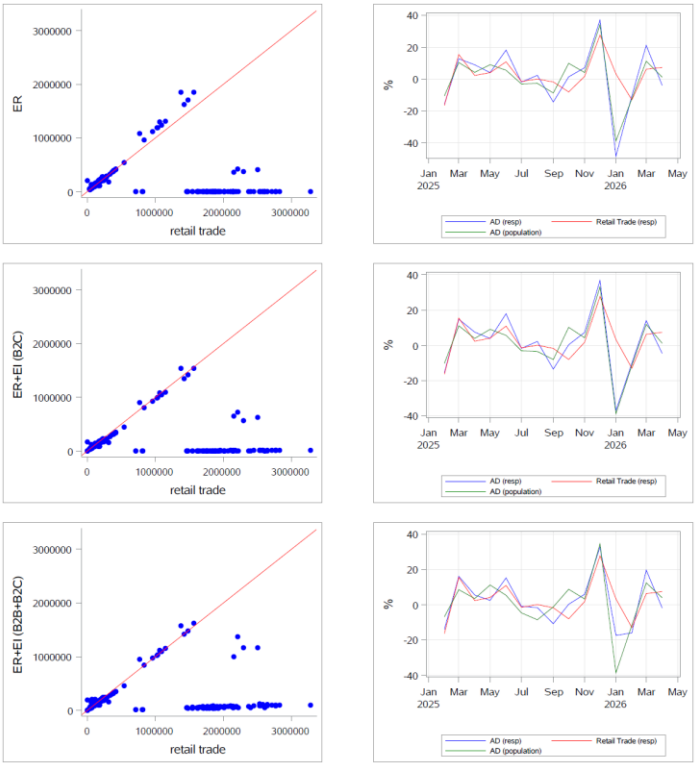
D) Sector-specific issues: *Non-specialised retail sale*

An example where AD will not be used to replace survey data at the first stage is the *Department stores* sector. It falls within the non-specialised retail sector (predominantly non-food related), where AD do not allow the type of products sold to be determined as in the survey (there is no direct correspondence between the ATECO code and the product sold).

The comparative diagnostics between AD and direct survey data are not satisfactory. The structural SCI component significantly lowers the overall CCI indicator, despite the dynamic component relating to month-on-month variations is quite good. Furthermore, the use of AD accentuates certain peaks in the series, unlike when using survey data alone.

The scatter plots comparing the survey data with AD, in fact, show several cases with ER values of zero, as opposed to positive values reported in the direct survey. The approximation between the data does not improve when EI —both B2C and B2B—are considered in addition to ER. These cases could be due to various factors, such as, for example, the sale or lease of a business to another legal entity, or the management of a department within another sales organization that issues the sales receipt.

Figure 5 - Scatter plots, m-o-m changes, diagnostics – ***Department stores*** stratum (2_47191)



Indicator	Type A	Type B	Type C
Share of ER on total	75.8	94.6	100.0
SCI	**0,157**	**-**	**-**
CCC	0,018	-0,035	-0,054
Ratio	0,254	0,202	0,192
KS dist	0,362	0,366	0,373
MAPE sym	0,828	0,820	0,829
DCI	**0,817**	**0,756**	**0,706**
Corr	0,819	0,679	0,595
Vol	1,301	1,573	1,802
Sign conc.	0,750	0,813	0,813
CCI	**0,358**	**-**	**-**

Notes: Type A - EI (B2B+B2C) and ER; Type B - EI (only B2C) and ER; Type C - only ER. SCI and DCI in normalised form; their components in the original formulation.

E) Limitations of AD substitution: *Retail sale of pharmaceutical products subject to medical prescription*

The comparability of data in this sector is limited by specific regulatory features relating to the National Health System. Due to this issue, this sector is not currently slated to be replaced with AD.

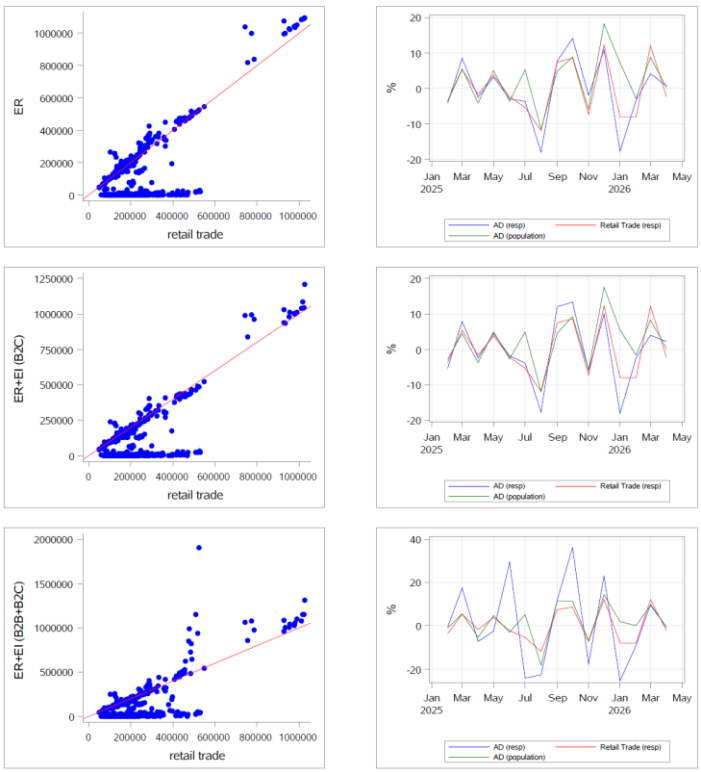
The diagnostic analysis highlights limited comparability between AD and survey data across the various levels of aggregation considered. The CCI indicator falls predominantly into the *Low* or *Moderate* classes, with a structural component (SCI) representing the main critical factor in the non-alignment between the two sources.

The graphical comparison between AD and survey data—using the same sample of firms from the survey—reveals many cases with zero ER values, as opposed to positive values reported in the retail sales survey. The frequent occurrence of zero values in the AD appears to be attributable to the sector's specific regulatory and operational characteristics, which result in partial coverage of transactions by the sources considered[6].

The inclusion of EI component —particularly the B2B component— leads to increased volatility in the administrative time series, as evidenced by spikes in monthly changes.

At the population level, an interesting offsetting effect emerges in the dynamics: the positive and negative deviations observed at sample level tend to offset each other, resulting in greater stability in the aggregate trend.

Figure 6 - Scatter plots, m-o-m changes, diagnostics – ***Pharmaceutical products*** stratum (2_47731)



Indicator	Type A	Type B	Type C
Share of ER on total	79.2	93.8	100.0
SCI	**0.440**	**0.407**	**0.392**
CCC	0.571	0.568	0.548
Ratio	0.460	0.385	0.361
KS dist	0.739	0.748	0.748
MAPE sym	1.365	1.460	1.500
DCI	**0.665**	**0.893**	**0.903**
Corr	0.765	0.867	0.862
Vol	2.526	1.219	1.177
Sign conc.	0.875	0.938	0.938
CCI	**0.541**	**0.603**	**0.595**

Notes: Type A - EI (B2B+B2C) and ER; Type B - EI (only B2C) and ER; Type C - only ER. SCI and DCI in normalised form; their components in the original formulation.

F) Regulatory exemptions*: Retail sale of newspapers and other periodical publications*

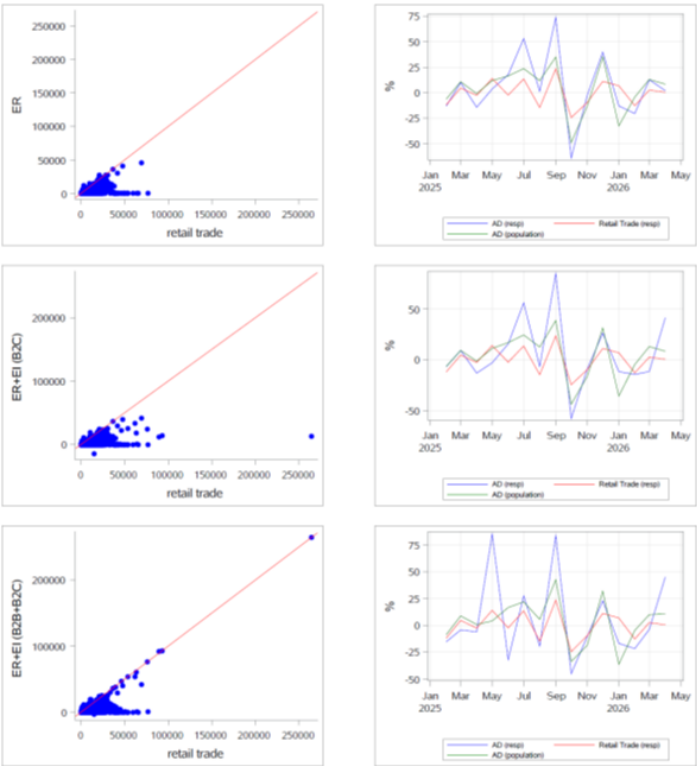
This stratum includes sales of newspapers and other periodical publications that, under national law, are not subject to the requirement for electronic storage and electronic submission of daily sales data.

The analysis reveals limited consistency between the estimates derived from direct data collection and the information obtainable from the administrative sources considered. The high prevalence of zero administrative values appears to be attributable to the partial coverage of ER data, due to the sector-specific regulatory exemptions already described. Integrating the ER component with the B2B and B2C EI data results in only a marginal improvement in data coverage, without producing a significant increase in explanatory

[6] This is likely because the electronic submission of payments and healthcare expenses for pharmacies is carried out through electronic cash registers connected to the Health Card System. It is therefore necessary to verify the operation of this system with the Italian Revenue Agency. In addition, there is the time it takes for the Italian National Health System to reimburse pharmacies for prescriptions dispensed during the month, which can range from 30 to 60 days.

power compared to the baseline administrative source (sales receipts). Consistently with this, the statistical indicators fall predominantly within the *Low*-to-*Moderate* range, highlighting that the integration of additional administrative sources yields limited benefits in terms of estimation accuracy.

Figure 7 - Scatter plots, m-o-m changes, diagnostics – **Newspapers** stratum (1_47621)



Indicator	Type A	Type B	Type C
Share of ER on total	64.1	90.1	100.0
SCI	**0.396**	**0.241**	**0.222**
CCC	0.526	0.158	0.122
Ratio	0.290	0.205	0.185
KS dist	0.564	0.611	0.594
MAPE sym	1.245	1.353	1.242
DCI	**0.618**	**0.631**	**0.662**
Corr	0.800	0.773	0.830
Vol	3.026	2.646	2.534
Sign conc.	0.813	0.750	0.813
CCI	**0.495**	**0.390**	**0.384**

Notes: Type A - EI (B2B+B2C) and ER; Type B - EI (only B2C) and ER; Type C - only ER. SCI and DCI in normalised form; their components in the original formulation.

5 – Future strategy for Administrative-Data-Based Estimation

The results presented in the previous sections provide evidence of the potential role of AD in supporting the future development of the RTS. The assessment carried out through both the diagnostic and graphical approaches has allowed the identification of the domains in which administrative sources show a satisfactory level of consistency with survey estimates, while also highlighting the limitations that currently prevent their application to the entire retail sector.

The future strategy is therefore based on a gradual and selective integration of AD into the current estimation framework. This approach allows preserving the reliability and continuity of official statistics while reducing the statistical burden on businesses and improving the efficiency of the production process.

5.1 – Potential impact of the transition to administrative-data-based estimation

Table 6 provides an indication of the potential reduction in statistical burden for firms in the RTS sample, based on *High* and *Very High* scores of the CCI. It also shows the potential increase in the number of businesses covered by the statistics, resulting from the transition from a sample-based approach to a population-based framework within the scope of the survey.

The results show that over 70 per cent of the enterprises currently included in the sample could be replaced by AD. This highlights the significant potential of moving towards an administrative-data-based estimation

framework, both in terms of reducing direct data collection and improving the coverage of economic indicators.

A relevant aspect is the strong concentration of the target population among micro-enterprises. Businesses with '1-5' employees account for around 89 per cent of the total population considered for the transition to AD, corresponding to around 80 per cent of the selected population. This confirms that the main benefits of the transition are expected to arise from improving the coverage of the smallest businesses, which are typically the most difficult units to observe through direct surveys.

Table 6 – Composition of the survey sample and target business population eligible for AD substitution, by size class.

Size class	Total Sample		Selected Sample		Total Population		Selected Population	
	n.	frac.	n.	frac.	n.	frac.	n.	frac.
1-5	5.285	65,8	4.317	53,8	339.033	89,0	303.042	79,6
6-49	1.095	13,6	746	9,3	39.964	10,5	25.673	6,7
50+	1.647	20,5	727	9,1	1.748	0,5	738	0,2
Total	**8.027**	**100,0**	**5.790**	**72,1**	**380.745**	**100,0**	**329.452**	**86,5**

Note: 'Total sample' refers to the RTS sample for the year 2026. 'Selected sample' refers to enterprises belonging to strata with a *Very High*/*High* CCI range.

Conversely, larger businesses ('50+' employees) represent a relatively small fraction of the population (0.5 per cent of the total population and 0.2 per cent of the selected population), although they account for a substantially larger proportion of the survey sample (approximately 20 per cent of the total sample). This finding is consistent with the different sampling strategy traditionally adopted for large enterprises, which are often selected with higher probabilities due to their economic relevance.

Overall, the distribution of units considered for the transition to AD highlights the potential benefits of moving towards a population-based measurement framework. Introducing administrative sources into the current estimation process would primarily affect the largest proportion of the business population, namely micro-enterprises, which are traditionally characterised by a lower response rate and a greater reporting burden. Furthermore, the transition would substantially reduce the number of units subject to direct surveys, particularly in those strata where the sampling efforts is relatively high compared with their representativeness within the business population. In addition to reducing survey's burden, using AD could improve the coverage and timeliness of economic indicators by providing information on all the enterprises covered by the survey at a population level. This would allow for a more comprehensive representation of economic dynamics, including demographic and structural changes that are difficult to capture in a sample-based framework. The findings therefore support a gradual, sector-specific adoption of AD, driven by the ability of administrative sources to ensure consistency with existing estimates and to effectively represent economic dynamics.

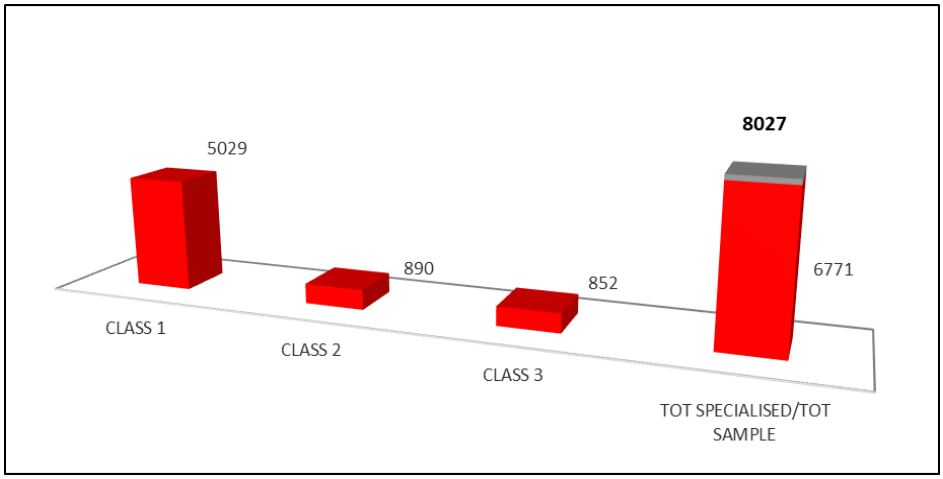
5.2 – Implications for the RTS sample design

The transition towards AD-based estimation would have important implications for the current design of the RTS. As discussed in the previous sections, the first implementation phase is expected to focus on specialised retail activities, where the correspondence between the predominant economic activity of the enterprise and the type of products ensures greater consistency between survey and administrative information.

In terms of the composition of the RTS sample by employee category and sector, the preliminary substitution could potentially affect a substantial proportion of specialised enterprises, which would have a significant

impact on the current sample. Considering the 2026 sample, for example, out of 8,027 surveyed enterprises, 6,771 are specialised enterprises, corresponding to 84.4 per cent of the total sample. Among these, 5,029 belong to 'Class 1' (74.3 per cent), 892 to 'Class 2' (13.1 per cent) and 852 to 'Class 3' (12.6 per cent).

Figure 8 – Specialised enterprises by size class (*absolute value*)

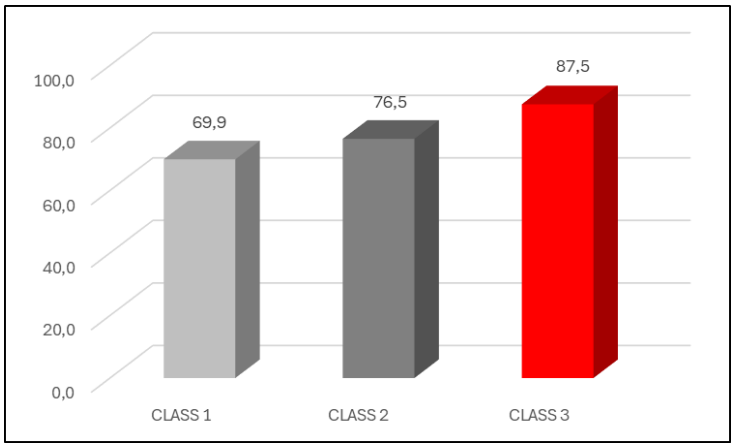


Note: Values refer to the RTS sample for the year 2026. 'Class 1' (1-5 empl.), 'Class 2' (5-49 empl.), 'Class 3' (50+ empl.)

The replacement strategy would initially concern mainly small and medium-sized specialised enterprises, corresponding to approximately 5,900 units[7]. This represents a significant change in the current survey framework, as specialised enterprises currently constitute the main component of the sample.

An additional relevant aspect concerns the characteristics of the enterprises that would be affected by the transition. The units identified as potential candidates for replacement are also those that generally experience the greatest difficulties in direct data collection. Small enterprises often have limited organisational resources, lack dedicated staff for statistical reporting and may face difficulties in completing electronic questionnaires. These factors contribute to lower response rates and a greater reliance on estimation procedures for non-response treatment.

Figure 9 – Average response rate by employee size class (*percentage values*)



Note: 'Class' refers to categories of employees in the 2025 survey sample, 'Class 1' (1-5), 'Class 2' (5-49), 'Class 3' (50+)

[7] The number of sampled units differs from that shown in Table 6, as it takes into account not only cases classified as *High* and *Very high*, but also cases falling within the lower ranges of the CCI indicator. For assessments that go beyond statistical quality, it is considered appropriate the use AD.

The analysis of response rates by size class (Figure 9) confirms that smaller enterprises show the highest difficulties in participating in the survey, while larger enterprises generally achieve higher response levels. Consequently, the use of AD for these units could not only reduce reporting burden but also improve the stability and quality of the estimation process by reducing dependence on imputation procedures. Overall, the integration of AD into the RTS production process represents a structural redesign opportunity. A progressive transition, starting from the domains where administrative sources show the highest level of reliability, would allow maintaining the quality of official indicators while improving efficiency, reducing burden on businesses and moving towards a more comprehensive measurement framework.

Conclusion

The analysis on the comparability between survey-based measures and administrative turnover estimates, provides positive evidence of a gradual transition towards the use of administrative sources also for the Retail Trade Survey. This experiment is part of a broader process, started in year 2025, that has already led to the integration of administrative data within the Turnover Survey on Services and that will continue in the coming years with a progressive extension to other sectors.

The integration of electronic invoicing and electronic receipt data represents an important opportunity for improving the efficiency of statistical production processes, while preserving the quality and reliability of statistics. It significantly reduces the statistical burden on enterprises, particularly those identified as potential candidates for initial replacement, which generally experience the greatest difficulties with direct data collection. Moreover, the extension of administrative data to the entire target population has the potential to further improve the quality of statistical estimates by moving from information collected at sampling level to the whole population. This broader coverage can reduce sampling variability and provide more complete and consistent information to produce statistics, together with the almost total elimination of sampling error.

Building on these promising results, the efforts will continue in the future, with a gradual extension of the use of administrative sources to other short-term statistics. This will further contribute to the reduction of response burden while strengthening the efficiency and sustainability of the statistical production process.

References

Amato, G., Ciammola, A., Filiberti, S., Margani, P., Orsini, D., (2021). "Indagini congiunturali sul fatturato e fatturazione elettronica a confronto: prime evidenze e prospettive". Quattordicesima conferenza nazionale di statistica. https://www.istat.it/storage/14-Conferenza-nazionale-statistica/poster/008.jpg

Amato, G., Moschetta, M., Tuzi, F., (2025). "Introducing Administrative Data within the Survey on Turnover in Services". UNECE, United Nations Economic Commission for Europe Conference of European Statisticians Workshop 2025. Lisbon, 2-4 June 2025, Lisbon. https://unece.org/sites/default/files/2025-05/DC2025_Italy_Moschetta_Tuzi_Poster.pdf

Bruno, G., Filiberti, S., Tuzi, F., (2025). "Integrating Electronic Invoicing Data in the monthly Business Statistics on Services Turnover: from the experimental phase to the start of full production". Contribution presented at the EESW2025, The European Establishment Statistics Workshop 2025. Rome, 5-7 November 2025. https://drive.google.com/file/d/1WDRsVaNekijQ3hJmSj2KBUh1K4HFIPRv/view

Corea C., Bruno G., Ciavardini R., Martone L., Pontesilli D., Parise B., Sartor F. (2023). "I nuovi indici del commercio al dettaglio". Ebook, ISBN: 978-88-458-2095-3. https://www.istat.it/wp-content/uploads/2023/03/Nuovi-indici-commercio-al-dettaglio.pdf

Filiberti S., Amato G., Orsini D., Bruno G., Rapiti F.M., (2023). "E-Invoicing as Essential Administrative Source for the Future of Official Business Statistics: the Case of Short Term Statistics", paper presented at ASA 2023 Conference, 6-8 September, Bologna. https://www.sa-ijas.org/ojs_asaproc/35_3_2023/asaproc_0039.pdf

Filiberti, S., First assessments of advantages deriving from use of e-Invoicing data in short term business statistics. *United Nations Economic Commission for Europe Conference of European Statisticians – Expert Meeting on Statistical Data Collection and Sources*. 22-24 May 2024, Geneva, Switzerland. https://unece.org/sites/default/files/2024-04/DC2024_S2_Italy_Filiberti_A.pdf

Filiberti S., (2024). "Electronic invoicing as administrative source that simplifies and improves the production process of short-term business statistics surveys." Contribution presented at ASA 2024 Conference, 18-20 September, 2024, Rome. https://cess2022.dss.uniroma1.it/event/10/contributions/464/contribution.pdf

Rapiti F.M., Filiberti S., Bruno G., Amato A., Orsini D., (2023). "E-invoices as a new timely administrative (and big) data source for short term business statistics", EESW2023, The European Establishment Statistics Workshop 2023. Lisbon, 19-22 September 2023. https://drive.google.com/file/d/18oNcYXm5Ea-Az0LrczFgIOKfaugsHWlm/view

Rapiti, F. M., Filiberti, S., Bruno, G., Causo, M.S., Moscufo, M., Amato, G., Tuzi, F., (2026). "Use of e-invoicing data for official business statistics: a new paradigm for reducing statistical burden, cutting costs and producing better and new timely statistics". IAOS, The International Association for Official Statistics Conference 2026. Vilnius, 12-14 May 2026. https://www.isi-next.org/abstracts/submission/4705/view/

"Nota metodologica nel Comunicato stampa mensile del commercio al dettaglio". https://www.istat.it/tag/commercio-al-dettaglio/