

E-invoicing as essential administrative source for the future of official business statistics: The case of short term statistics

Salvatore Filiberti, Giuseppe Amato, Danilo Orsini,
Giancarlo Bruno, Fabio Massimo Rapiti
Istat, Rome, Italy

1. Introduction

In a world that nowadays is almost completely and - in perspective - will totally be digitised, the use, for statistical purposes, of data acquired by central administrations will represent the standard model for the production, by the National Statistical Institutes (NSI), of official quantitative information on the main economic variables. In particular, the integration of survey and administrative data makes it possible to build a robust economic statistics system capable of returning very detailed and timely information to users, while keeping the statistical burden on respondents at an acceptable level.

The recent introduction of the rules on electronic invoicing has opened up new scenarios in Italy, also with regard to the compilation of official economic statistics. Data acquired by the Revenue Agency (RA)¹ and used for statistical purposes by the NSI could represent one of the main sources of administrative quantitative information on economic variables of enterprises. The practice of using administrative data is also envisaged and strongly suggested by the European legislation on official statistics with the aim to reduce the statistical burden on businesses, especially for small economic units.

There are many concrete perspectives of using administrative information about e-invoicing for statistical purposes, for example as up-to-date information to establish the activity status of an enterprise and thus to update business registers, or as information on the turnover levels of enterprises, which is necessary to produce short-term turnover indices on a monthly basis.

E-invoicing may also be used in surveys, as an auxiliary variable, both when defining the sample design, but also in the estimation procedures, according to the calibration estimators approach, as an additional reference variable in order to obtain more accurate and detailed final estimates of the variables of interest.

Moreover, turnover from e-invoices may represent an interest variable in the frame of structural business statistics and then used to produce the extended business statistics registers. It is also possible to explore flows of goods/services between the transferor/provider and the transferee/customer in order to identify the production chains and quantify/qualify trade from and to abroad, etc.

The focus of this paper is to show first experimental results concerning the use of electronic invoicing data as a source for the compilation of monthly short-term economic statistics on companies' turnover. In Section 2 some basic information on main short-term business statistics sources on turnover and the tax administrative source of e-invoicing is provided. The preliminary analyses conducted on survey and administrative data and main results obtained are shortly shown in Section 3. Section 4 summarises some considerations regarding the prospect of using e-invoicing data for short-term business statistics.

¹ Special thanks go to the Revenue Agency, Resources Division, Central Directorate for Technology and Innovation, Database Sector, for the cooperation and willingness to provide the necessary data for the experimentation.

2. Short term business statistics and e-invoicing

Istat's system of short-term economic statistics is based on well-established system of direct statistical surveys that provide quality results with reference to the main indicators, also in compliance with Regulation (EU) 2019/2152 and within the strict deadlines provided for. The production processes of the surveys, however, are not without operational costs, especially for the contacted enterprises, which have to spend time and resources to provide the required information.

The recent introduction of the law on the obligation of companies to notify the RA of the details of e-invoicing transactions has made it possible to set up a continuous data flow and a database with a huge amount of information. Such information, always respecting the confidentiality of the data, can be available for the institutional purposes of other administrations. Istat, in particular, could use this fiscal administrative source for various purposes including the preparation of short-term turnover statistics.

The experimentation conducted at Istat within the Directorate of business statistics aims at assessing use of administrative e-invoicing data for their progressive use in production.

Short-term business statistics surveys on industrial turnover *Fat* (Istat, *Fatturato dell'industria*, Statistiche flash) and turnover in services *Fas* (Istat, *Fatturato dei servizi*, Statistiche flash) are involved in the experimentation. These direct surveys produce, on a regular monthly basis, turnover indices (unadjusted, seasonal adjusted and calendar adjusted indices) within 60 days from the reference period at a Nace rev.2 economic activity sector detail (sections B, C) for *Fat* survey and at a Nace rev.2 division/group economic activity sector detail (sections G, H, I, J, M, N) for *Fas* survey. In total, about 28,000 enterprises are involved in the surveys and then requested to provide information via an e-questionnaire about their total monthly turnover.

Electronic invoice data from businesses are acquired via the Interchange System (SdI) according to a well-defined record layout including a large amount of information with respect to transferor and transferee. E-invoice data contain also information about the date of issue of the invoice, the country of the assignee, the type of counterparty (B2B, B2C, etc.), the type of document (invoice, simplified invoice, advance on invoice, credit/debit note, etc.).

Since the entry into force of e-invoicing (January 2019), there have been and are still ongoing evolutions of the regulations. In addition, the fiscal regimes and consequently the number of companies concerned by the e-invoicing regulations is growing. All information on the evolution of the regulations is available online.

3. Early experiments and main results

The main objective of the experiment was to compare the turnover levels of enterprises obtained via the RA source (e-invoices) with those obtained by the statistical data source (*Fat* and *Fas* surveys) and to assess their coherence and quality. Thus, turnover indices can be constructed according to the currently adopted methodologies².

An assessment was made of the correspondence, in both sources in question, between the survey/analysis units and between the variable recorded (turnover of the enterprise) in the time unit. Both the unit of survey/analysis (legal-economic unit) and the measured variable (turnover) definitions are compatible as far as definitions are concerned. In particular, turnover comprises the total amounts invoiced by the observation unit during the reference period and corresponds to the value of market sales of goods or services provided to third parties.

For experimental purposes, a subset of companies was selected in order to meet both Istat's requirements of representativeness in terms of sectors of economic activity and size (in term of number of persons employed), and the constraints indicated by the RA in order to be able to

² The interesting results obtained from an initial work with this objective (Amato et al., 2021) demonstrated the coherence between the two sources at the aggregate Nace rev.2 division level.

manage data in an efficient way. In fact, each economic unit can correspond to a very large number of invoices.

It was therefore necessary to build a database that was as consistent as possible with the information of interest in terms of economic units involved, variables, reference period, and completeness of data during the year, etc.

For this purpose, some preliminary steps were carried out: acquiring the data and metadata files from the Agency; reading of e-invoices data files and standardisation of formats of variables; study of the invoice items included in the transmission path. Then the turnover variable was obtained for each company, according to the indications provided in the technical notes of the RA, and it was linked to those of the Istat short-term turnover surveys.

It was essential to link also the records so obtained to the business register (Istat, *Registro statistico delle imprese attive*) in order to retrieve structural information for each economic unit (Nace rev.2 economic activity classification code, number of persons employed, etc.). The key for the links was the fiscal code of the economic unit.

In order to make an initial assessment of the suitability of the tax source as an auxiliary for business statistics, some comparisons at micro level were made.

As a first analysis, turnover levels of e-invoicing administrative data were compared with the turnover data coming from the surveys.

A wide time span was chosen (from 2019 to 2022) so that the correspondence between the turnover values recorded by the survey and those acquired through the administrative source can be better assessed over time.

Table 1 shows the percentage of enterprises in the e-invoicing out of the total enterprises in the samples of the industry and services turnover surveys by Nace rev.2 Sections. The overall coverage of 84.3% seems to guarantee excellent results in terms of coverage of the units included in the samples. In addition, it should be considered that regulatory developments have provided for and continue to provide for a progressive increase in the number of companies and operations subject to the e-invoicing obligation, so that the coverage of administrative data is already even more extensive today.

Table 1

Per cent of enterprises with e-invoicing data in the Fat and Fas survey samples by sector of economic activity (year 2021)	
Nace rev.2 section	Enterprises (%)
Total	84.3
B - MINING AND QUARRYING	98.3
C - MANUFACTURING	92.3
G* - WHOLESALE AND RETAIL TRADE; REPAIR OF MOTOR VEHICLES AND M	87.1
H - TRANSPORTATION AND STORAGE	91.1
I - ACCOMMODATION AND FOOD SERVICE ACTIVITIES	73.1
J - INFORMATION AND COMMUNICATION	92.6
L - REAL ESTATE ACTIVITIES	37.3
M - PROFESSIONAL, SCIENTIFIC AND TECHNICAL ACTIVITIES	72.8
N - ADMINISTRATIVE AND SUPPORT SERVICE ACTIVITIES	88.0

(*) excluding G47 - Retail trade, except of motor vehicles and motorcycles

The units considered for the experiment were chosen among those included in the Istat short-term statistics surveys on industry and service turnover, in total about 17.7K units (14.5K units for services sector and 3.2K units for industry) without missing data during the year of reference.

Initial experiments are aimed at understanding whether and to what extent tax data can be used as an auxiliary source and/or as in substitution of survey data. There may be cases in which, however, the tax data must be evaluated in the most appropriate manner (large enterprises, units belonging to specific estimation domains with a low number of enterprises, etc.) so that it may be

preferable to use the survey data.

Preliminary analyses show, as expected, a very good correlation between turnover from e-invoice and turnover from surveys. These analyses were conducted both with respect to different estimation domains and with reference to different periods.

For greater clarity of presentation, it was chosen to show, here, only the graphs relating to some of the numerous subpopulations of units belonging to the domains of interest. In particular, in Figure 1 and Figure 2 the micro comparisons are made between the turnover of service and industry sectors enterprises for 2020, as recorded by the Fas and Fat surveys and as acquired from the tax source. A high correlation seems to be evident; the graphs also shows some cases of enterprises for which the surveyed data and the administrative data are not well aligned. These are precisely the cases to be investigated in order to identify the reasons, which may stem from various causes, e.g. accounting, etc.

Figure 1:
E-invoicing vs Fas, total (year 2020, EUR)

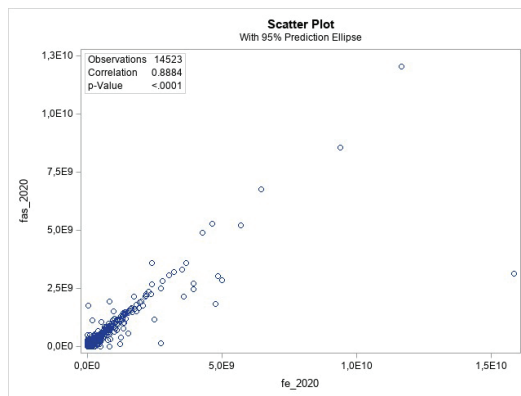
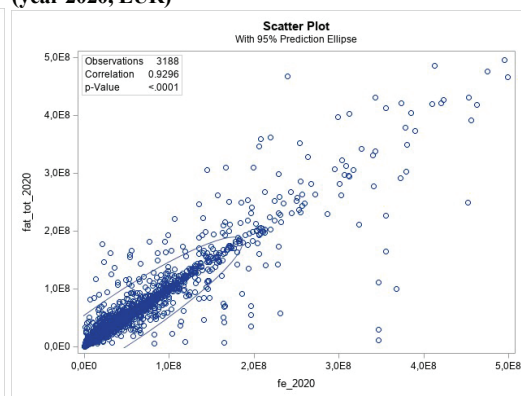


Figure 2:
E-invoicing vs Fat, enterprises < 500mln EUR turnover (year 2020, EUR)



Comparisons and analyses were also made at macro level, with reference to the following domains of interest: sector of economic activity, (Nace rev.2 divisions and groups) per quarter and total, as well as size class in terms of persons employed (1-9; 10-19; 20-49; 50-249; 250+), or turnover classes (< EUR 3 mln by size class; < EUR 1 mln; 1-5 mln; 5-10 mln; ≥ EUR 10 mln). Comparisons at the micro level, also per quarter were carried out³.

Below are the graphs of some representative cases showing the good degree of alignment between the data collected by surveys and the administrative data.

Figure 3:
E-invoicing vs Fas by Nace rev.2 group (year 2020, EUR)

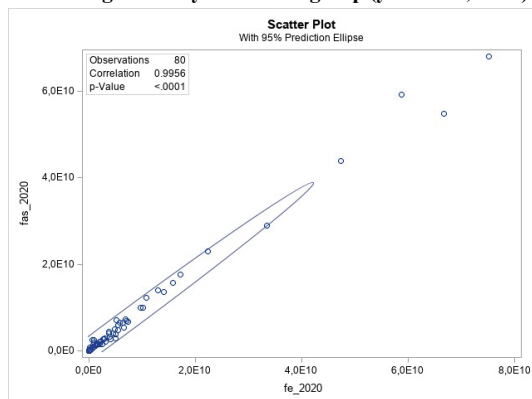
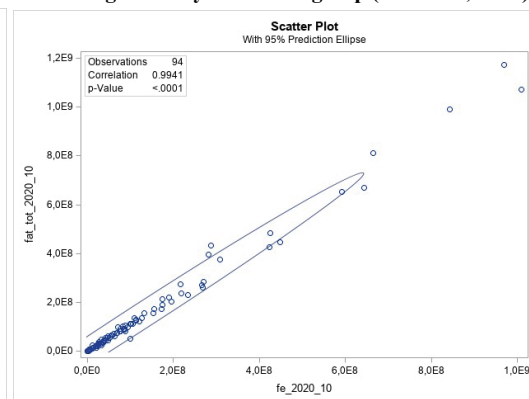


Figure 4:
E-invoicing vs Fat by Nace rev.2 group (Oct. 2020, EUR)



³ The results and graphs relating to the economic activity sectors not included in this work are possibly available upon request.

Figure 3 and Figure 4 show the high correlation between statistical and administrative data respectively at Nace rev.2 group-level for the service sector for 2020 and for the industry sector for October 2020.

Analyses of changes over time between surveyed turnover levels and e-invoice data were made for quarters 2021_1 to 2022_3 for each Nace rev.2 division.

Below are the graphs with the trend changes for G45 (Wholesale and retail trade and repair of motor vehicles and motorcycles) and G46 (Wholesale trade, except of motor vehicles and motorcycles) Nace rev.2 divisions, which together represent 56.0% of the entire services sector⁴, therefore with a greater relative weight, in terms of turnover, compared to the other Nace rev. 2 sections.

With reference to each quarter, shown in the abscissa, the changes were calculated at the stratum level (constructed by crossing economic activity and class of employees), considering the set of microdata available for e-invoicing and Fas survey respectively (uncommon observations are also included). Observations with a trend change outside the 2nd-98th percentile range of the distribution of changes in the reference quarter are not included. The stratum indices were subsequently aggregated using the weight structure of the services turnover survey.

Figure 5:
G45 Wholesale and retail trade and repair of motor vehicles and motorcycles (trend change %)

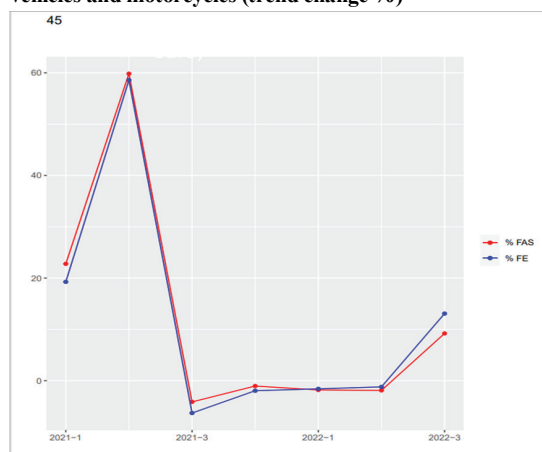
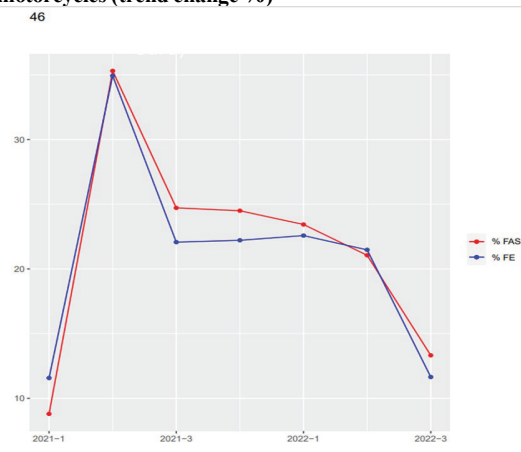


Figure 6:
G46 Wholesale trade, except of motor vehicles and motorcycles (trend change %)



Also this analysis, which is of an exploratory nature and should not be considered exhaustive, confirms on the whole the very good approximation of the trend changes based on administrative data compared to those obtained through survey data.

4. Some remarks

First results from the comparison of company turnover data with e-invoicing data seem to be extremely positive and provide with some relevant indications for the use, when operational, of the tax source for official statistics.

The evaluation of data quality, made possible through the analysis of the correlation coefficients between the levels of turnover from administrative data and survey data, allows to hypothesize that the use of tax data on company turnover would represent an unprecedented opportunity that should be exploited in the best possible way. It would also represent an investment of strategic importance for Istat, which offers prospects of absolute interest in the field of production of official economic statistics. The scenarios are in fact considerably interesting

⁴ Excluding Nace rev.2 division 47 - Retail trade, except of motor vehicles and motorcycles, that is covered by Retail Trade survey.

both with regard to the improvement in the quality of the statistics produced (timeliness, coverage, etc.) and with regard to the considerable reduction or even elimination of the costs incurred by companies to participate in direct surveys.

The quality of short-term statistics on the turnover of companies is supposed to improve with the massive use of fiscal data. Indeed, a reduction of both sampling error and bias and thus a reduction of the Mean Squared Error (MSE) is expected. Accuracy, which is relative to the coverage of enterprises belonging to the economic sectors of interest, appears to be very good looking at the results of initial analyses and in perspective, as data availability is extended to all units in the field of observation. Timeliness and punctuality of acquisition are also very good for the purposes of short-term statistics. When operational, e-invoice data will be available 20-25 days after the end of the reference period, therefore with a very good timing considering the production process deadlines of short-term statistics and so improving their overall efficiency.

Obviously, the positive perspectives outlined above will then have to be verified once the project is actually implemented. In fact, some potential critical issues related to the acquisition, management, maintenance and utilisation of administrative data flows and files may occur, as they must be compatible and integrated into Istat production processes.

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> Accessed 30 November 2021.
- Istat, *Fatturato dell'industria*, Statistiche flash, <https://www.istat.it/it/archivio/fatturato+industria>. Accessed 15 November 2023.
- Istat, *Fatturato dei servizi*, Statistiche flash, <https://www.istat.it/it/archivio/fatturato+servizi>. Accessed 15 November 2023.
- Istat, *Registro statistico delle imprese attive*, <https://www.istat.it/it/archivio/263692>. Accessed 23 November 2021.
- Regulation (EU) 2019/2152 of the European Parliament and of the Council - <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019R2152&qid=1700157163629>. Accessed 01 January 2022.