Electronic invoicing: a promising lever for public statistics

François Meunier[1]

Paper presented at the 18th conference of the National Accounting Association, June 20 and 21, 2024.

§§§

Electronic invoicing (E-invoicing or *e-I*) is becoming increasingly widespread throughout the world. In France, it is already mandatory for all sales contracts with the government. It will be mandatory from 2026 for large companies and SMIs, and one year later for SMEs, for sales to companies subject to VAT. The EU, since a 2014 Directive, has been pushing such a move, and the 2022 VAT in the Digital Age (ViDA) proposal foresees the mandatory adoption of *e-I* for all cross-border transactions within the Union. The changes to be expected in the business world and for macroeconomic monitoring cannot be overlooked.

What is electronic invoicing?

Given the rapid progress being made in digitization, E-invoicing is the order of things. It is in the natural interest of companies, and in fact, large corporations have largely accepted or imposed it: paper or pdf invoices are more expensive than *e-I*, payment delays potentially shorter, and the digitization reduces the cost of managing corporate data. But the acceleration of the project is above all of fiscal origin: the aim is to monitor VAT collection in order to reduce the significant fraud recently estimated by INSEE at between €20 and €26 billion a year in France. For public policies, this means a much more detailed monitoring of the economy. More of that in a moment.

The electronic invoice is still an invoice, i.e. a legal document attesting to a sales contract and a request for payment that the buyer must honor. It includes the same information as before: name and address of the parties, date of the sale or provision of services, quantity and precise designation of the products or services with their unit price and any discount granted, payment due date and penalties for late payment, applicable VAT rate. Since July 2021, any legal conformity guarantees have also been included.

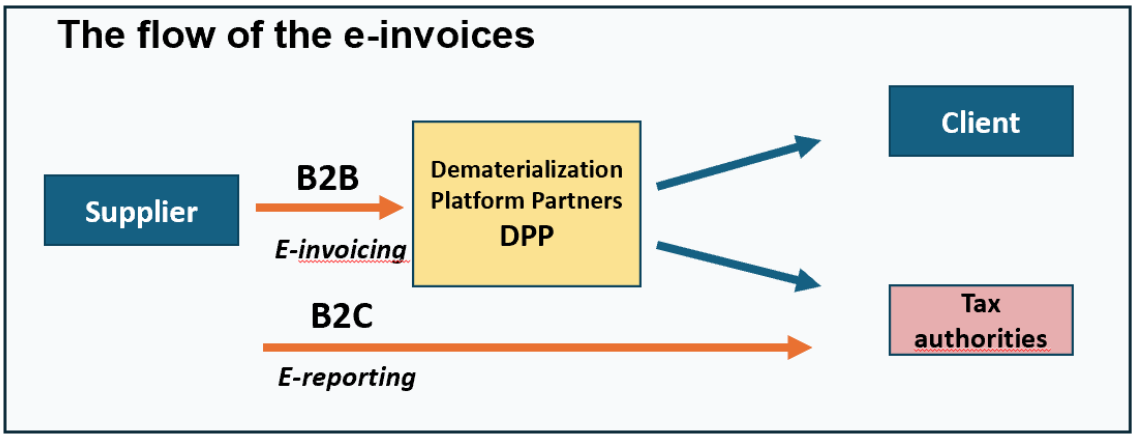
Its scope of application also remains the same as before: it applies to all exchanges of goods and services between VAT-registered companies established in France. Excluded are the health and education sectors, certain real estate transactions, and the banking and insurance sectors. Business-to-business (BtoB) flows are the specific focus of *e-I*, also known officially as "*e-invoicing*". However, companies also invoice entities not subject to VAT, mainly private individuals, without necessarily issuing them an invoice, let alone an e-

[1] Affiliated professor at ENSAE-IPP.

invoice (which must be digitally authenticated by both parties). Companies are required by law to declare transactions with private individuals (BtoC) in what is known officially as "*e-reporting*".

The core of the project remains the declaration to the tax authorities. To do this, companies submit their incoming and outgoing invoices to competitive operators, the so-called DPPs or D*ematerialization Platform Partners* (of the administration), the administration itself having an operator open to all, Chorus Pro. The DPPs send the invoices to the counterparty company, and to the tax authorities. They are likely to offer additional services, such as native integration with the accounting software used, dashboards, accounts receivable and payable management, software connection with cash management, sales management, etc. For companies, this undoubtedly represents a revolution in the internal processing of information flows.

The following chart describes the flow of invoices from the emitting company.



A major step forward for official statistics

We need to measure what is now available to central government and, we hope, to INSEE in particular: for billions of domestic transactions, we will feed a database showing the buyer, the seller (and therefore their sector of activity), the good or service concerned by economic nature and its price; and by aggregation, the sales and purchases of all companies subject to VAT. This will be usually available on a monthly basis with a time delay of one month. This database will be a kind of *Big Brother* looking over the economy, at any rate over all sectors subject to VAT. Indicators of economic activity at a very detailed level would be compiled, enabling the current state of the economy to be monitored in-time and step by step: activity, investment, producer prices, etc.

The stakes for national accounting are also high.

Historically, national accounts have followed specific routes, differing from one country to another depending on the state of its statistical system and its specific needs. Today, standardization at global level is about to be complete. In the early post-war years in France, let's say up to the 1970s, the dominant part of the national accounts was the supply

and use of goods and services, which, according to a product nomenclature, gave the basic equilibrium of Production + Imports = Intermediate consumption + Public and private demand + Exports. The resulting GDP was the sum of added values, obtained as production minus intermediate consumption. The Supply and Use Table provided a complete synthesis, presenting the resources and uses of each product, including the intermediate consumption of good or service *i* consumed in the production of product *j* (see summarized figure below).

Supply and Use Table (commodity x commodity)

	Product A	Product B	Product C	Final Demand	Σ
Product A	-	40	200	10	250
Product B	50	-	50	150	250
Product C	30	70	-	400	500
Labor	120	100	150		
Capital	50	40	100		
Σ	250	250	500	560	1000

This "product" section proves nonetheless complex to produce, due to the lack of statistical sources for completing it in full as regards to intermediate consumptions (grey part of the square table above). These are difficult to estimate on a regular basis. There is a considerable delay in updating them. As a result, some countries, such as Germany, still produce their SUT ex-post, while others (such as the UK) had not produced any SUT for a very long time.

In addition, statistics have long had access to important sources from the "institutional sectors": companies, households, the State, etc. In France, corporate tax data have been used since the early 1970s. Such data provide their income accounts, i.e. their resources and expenses. This was a major help in estimating GDP, since the latter is also the sum of income distributed within the country: essentially wages and profits. Based on the 1976 French national accounting standards (SECN), institutional sector accounts (the *Tableau économique d'ensemble* or TEE) were promoted, covering the income and expenses of economic agents, while including details of intermediate consumption by sector of activity, i.e. the SUT.

The statistical source that will emerge from the collection of inter-company flows will undoubtedly revitalize the product approach in the national accounts. It will be possible to know the intermediate consumption of good *i* by sector of activity *j*, on a monthly or quarterly basis and at an extremely fine-grained level of nomenclature, and even, of course

in theory, at the finest level which is the company (see the following figure, which we may call the *Big Brother SUT,* but which in figure shows only three of them).

Big Brother SUT

	Company A	Company B	Company C
Company A	-	40	200
Company B	50	-	50
Company C	30	70	-
Labor	120	100	150
Capital	50	40	100
Σ	**250**	**250**	**500**

The matrix of intermediate consumption per unit of production, known as the *Leontief matrix*, will be calculated, if not by simple aggregation, at least by simple econometric adjustments, following, for example, the Divay-Meunier (1980) methodology described in Appendix 1. Knowing the recipient of the invoice, whether foreign or domestic, we can also estimate the value-added included in exports (pending that trading partner countries do the same for the buyer country's imports, which will be possible - we're still "in theory" - for flows within the Union. For other countries, commercial pressure will probably be exerted by large resident importers to get their suppliers to digitize their invoices according to European invoicing standards. This will improve understanding of value chains.

Knowing that capital goods are subject to VAT, it will be possible to track investment (gross fixed capital formation) rapidly and exhaustively, and thus better estimate GFCF, which we know is not directly readable in companies' IFRS accounts. This will enable us to draw up an input-output table by product for capital goods, which is not done today in France. Incidentally, the distinction between capital goods and intermediate goods only makes sense in terms of a certain timeframe. If my oven lasts five years and I keep five-yearly rather than annual accounts, my oven will have been entirely consumed in the production process. And if I use quarterly rather than annual accounts, the wheat planted in the first quarter will be produced in the third. It has become an investment good. It was probably a conceptual error - but a proof of pragmatism - of the great founding pioneers of pre-war national accounting, including Stone and Léontief, to have classified gross fixed capital formation as final demand (Barro, 2021). The use of capital goods is part of value chains and spreads throughout the economy.

With the price of the product in the invoices, it will be possible to refine considerably the producer and user price indices for the various aggregates of goods, and consequently improve the calculation of the volume of GDP and its components.

Unexpected advances

Knowledge of inter-company flows will make it possible to track the trajectory of any product in the economy. We immediately think of goods whose production generates greenhouse gas emissions, mainly CO_2. This is what Chile's national accountants are already doing. Chile is well ahead of EU countries when it comes to *e-I*: it was made compulsory by a 2014 law with gradual and now complete application, and centralization by the tax office.

There, accountants now make an exhaustive request for a list of these goods (fuel, cement, gas...) from the tax office's central database. They therefore know who sells and who buys these goods, so that they have the direct emissions (also known as *scope 1* emissions) of each sector of activity at a detailed level. By using the Leontief matrix, these direct emissions can be used to determine the indirect emissions (known as *scope 2* and *upstream scope* 3) of all sectors of activity, according to the nomenclature used. This is the methodology used by Eurostat as part of the Figaro project (see Appendix 2).

The statistical innovation of electronic invoicing is not without its difficulties. A centralized VAT database can easily identify companies by their ID number, thus cross-referencing the data of a buyer with that of a seller - an essential point in limiting VAT fraud. But it's more difficult to identify products. For example, Chilean accountants come up against the fact that there is no perfectly established nomenclature for products at a very detailed level - with their various specifications, there are hundreds of millions of them! - nor are they necessarily correctly denominated by the vendor. For the fossil fuel products mentioned above, an AI program helps accountants to reclassify and order them according to a pre-established nomenclature. In any case, here lies a major logistical problem when it comes to making full use of this database. Initially, enquiries will probably be limited to a pre-established list of easily identifiable products, in much the same way as INSEE in France uses checkout data from major retailers to compile its consumer price index.

If we can track individual products, we can also retrace the "trajectory " of a specific product, directly and then indirectly, through the entire value chain, from the moment the good enters the country. In other words, the methodology used for carbon can be generalized to a product with polluting potential, such as plastics. Or, knowing the number of employees in each sector, we can approach what 19th-century economists called the *labor value* of goods and services.

Having an electronic medium for transactions will ultimately enrich the information that companies wish to transmit to each other. For example, given the major challenge of climate change, it would be conceivable for companies to declare on their invoices the direct and, progressively, indirect CO_2 emissions of products sold. We would move from a procedure where, in the establishment of carbon footprints, the onus of calculation is

currently borne by the purchaser, to a procedure where it would be the supplier who would provide the service of downstream data transmission. "*I sell you this product and, in addition to the price, I kindly indicate the amount of carbon emissions that its production has caused."* This cascading mechanism, very similar to the VAT flows, is virtuous because it allows us to produce important information at little cost. It could be integrated into company accounting systems in the same way as VAT.

The e-invoicing opens up a gigantic field for statistics on goods and services, capable of enriching or even renewing macroeconomic monitoring and the methods used to compile national accounts.

§§§

Appendices

1/ Determining the coefficients of the Leontief input-output matrix.

If we call x_i^k the input *i* of company *k*, y_j^k , j =1… J, the *J* products of the same company, and finally u_i^k,a statistical noise, we can write and thus estimate the technical coefficients on a large sample of companies *k*:

$$x_i^k = a_{i1}y_1^k + \cdots + a_{iJ}y_J^k + u_i^k.$$

Details of the calculation method can be found in Divay-Meunier (1980).

2/ Determining indirect carbon emissions from direct emissions

Statistical institutes are directly involved in projects to measure the carbon content of goods. They do so using the Leontief input-output methodology. A few initiatives - *e-Liabilities* in the US and the UK (Kaplan-Ramanna, 2021); Carbones-sur-Factures in France (Cazes, 2024) - are based on a decentralized procedure quite similar, in terms of information collection, to VAT.

The carbon emission *e* caused by the production of a good or service *j*, also called the footprint of product j, e_j in the following, is the sum of the footprints of intermediate consumption and direct emissions to produce it, named d_j:

(1) $\quad e_j = a_{1j} \times e_1 + \cdots + a_{nj} \times e_n + d_j,$

where a_{ij} is the quantity (in constant euros or physical units) of product *i* needed to produce one unit of product *j*. This coefficient is immediately derived from the national accounts input-output table, by dividing the amount of intermediate consumption of good *i* needed to produce good *j* by the output of that good.

This gives *n* equations of this type for the *n* products/industries of the economy, which can be written more conveniently in matrix form, i.e.:

$e' = e'A + d'$ ou encore : $d' = e'(I - A)$, where *I* is the unit matrix, *e* the vector of product footprints and the apostrophe signifying vector transposition. The *A* matrix is called the Leontief matrix and is often used in macroeconomic analysis. See Meunier (2023) for more details.

The *(I - A)* matrix is generally invertible. It is so if the Leontief matrix has positive or zero coefficients - which is natural - and if, in a broad sense, the economy is "productive", i.e. if the level of output is greater than the intermediate consumption used to produce it (for example, it takes less wheat to sow than is harvested). In this way, the footprints of all products are obtained as follows:

(2) $e' = d'(I - A)^{-1}$.

This equation indicates that if we know the direct emissions of all the products in the economy, we can derive the footprints. It is this relationship that is used to form the table above. The equation also legitimizes the iterative approach imagined earlier in our thought experiment. It consists of accumulating direct emissions over time. At iteration 1, direct emissions are equal to d'. At the next iteration in the exchanges, companies additionally have and communicate the direct emissions of Tier 1 suppliers, i.e. $d'A$; and thus in total; $d' + d'A$. In the next round, those of Tier 2 suppliers are added, i.e. $d'A^2$. And so on. So, as with any convergent series, here in matrix form:

(3) $e' = d' + d'A + d'A^2 + \cdots + d'A^t + \cdots$, which converges to: $d'(I - A)^{-1}$.

If the algorithm described in equation (3) leads to the solution of the problem, the approach adopted by carbon accounting converges even faster, since at each step an estimate close to the end point is introduced.

An important dual relationship

Let's call *q* the vector of gross production and *f* the vector of final demand for all products in the economy. Input-output analysis also indicates that production is equal to intermediate consumption and final demand. Formally, at the macroeconomic level and in vector terms:

(4) $q = Aq + f$ ou $q = (I - A)^{-1}f$.

This relationship will clarify an important point. Under the validity conditions of equation (2), we show that footprints are greater than direct emissions: $e \geq d$. Consequently, the sum of footprints is greater than all the carbon that is emitted into the atmosphere. This means that footprints cannot be added up without care, since their aggregate level exceeds what is emitted. If good *A* is used to produce good *B*, then its footprint is an indirect emission from product *B*'s point of view. Adding the two footprints together results in double counting. Hence the need for caution when reasoning about the footprint of a set of products or companies, as we sometimes do too quickly when we want to know the footprint of a financial portfolio.

However, we find equality, not for the economy's total goods, but for final demand goods, mainly household consumption and exports.

In fact, by premultiplying the two terms in (4) by *d*, we obtain, given equation (2):

$$(5) \quad d'q = d'(I - A)^{-1}f = e'f.$$

Total direct emissions from gross production equal total footprints from final demand. Let's take the case of a yoghurt, whose demand increases by one unit. The *f* vector bears a 1 for this product and zeros everywhere else. Yogurt's footprint is indeed the sum of the direct emissions caused by the marginal production of the goods used to make it. We have: $e_{\text{yogurt}} = d'q_{\text{yogurt}}$.

§§§

References

Barro, Robert J., 2021, *Double Counting of Investment*, The Economic Journal, Volume 131, Issue 638, August.

Cazes, Jérôme, 2024, *Les TPE et les PME peuvent tous tracer leur contribution à la décarbonation*, Le Monde, January 26.

Divay, Jean-François and François Meunier, 1980, *Deux méthodes de confection du tableau entrées-sorties*, Annales de l'INSEE, n°37.

Kalckreuth (von), Ulf, 2022, *Pulling ourselves up by our bootstraps: the greenhouse gas value of products, enterprises and industries*, Discussion Paper, Deutsche Bundesbank No 23.

Kaplan, Robert S., and Karthik Ramanna, 2021, *Accounting for Climate Change*, Harvard Business Review, December.

Meunier, François, 2023, *Towards a generalized carbon accounting system. Tracing carbon at product level*, Institute Messine.