Scaling Up to the Cloud: Cloud Technology Use and Growth Rates in Small and Large Firms

Bernardo Caldarola[1,2] and Luca Fontanelli[3,4,*]

[1]European Commission, Joint Research Center, Seville, Spain
[2]UNU-MERIT and Maastricht University, Maastricht, the Netherlands
[3]Università di Brescia, Italy
[4]RFF-CMCC European Institute on Economics and the Environment
[*]Corresponding author: luca.fontanelli@unibs.it

March 20, 2025

Abstract

Recent empirical evidence shows that investments in ICT disproportionately improve the performance of larger firms versus smaller ones. However, ICT may not be all alike, as they differ in their impact on firms' organisational structure. We investigate the effect of the use of cloud services on the long run size growth rate of French firms. We find that cloud services positively impact firms' growth rates, with smaller firms experiencing more significant benefits compared to larger firms. Our findings suggest cloud technologies help reduce barriers to digitalisation, which affect especially smaller firms. By lowering these barriers, cloud adoption enhances scalability and unlocks untapped growth potential.

Keywords: cloud, ICT, firm growth rate, firm performance, concentration.
JEL Codes: L20, L25, O33

Acknowledgments

We thank Flavio Calvino, Marco Grazzi, Kristina McElheran, and participants in the 2024 DRUID, 2024 EARIE conferences, 2024 SBEJ Conference for Young Economists and 2025 TPRI's Works-in-Progress Seminar Series for insightful comments. A previous draft of this work had the title "Cloud technologies, firm growth and industry concentration: Evidence from France". Access to French data benefited from the use of Centre d'accès sécurisé aux données (CASD), which is part of the "Investissements d'Avenir" program (reference: ANR-10-EQPX-17) and supported by a public grant overseen by the French National Research Agency. Luca Fontanelli gratefully acknowledges funding from the European Union's Horizon 2020 research and innovation program under the ERC project grant agreement No. 853487 (2D4D). Usual caveats apply.

1 Introduction

Recent studies have established a positive association between digitalisation and intangible assets on one side, and industry concentration on the other (see among others, Bessen, 2020; Bajgar et al., 2025; Lashkari et al., 2024), showing that the use of information and communication technologies (ICT) benefits larger firms to a greater extent (Brynjolfsson et al., 2023; Babina et al., 2024). However, ICTs may not be all alike. In particular, cloud computing services can replace costly and fixed ICT investments (Bloom and Pierri, 2018; DeStefano et al., 2023a). Furthermore, cloud technologies allow flexibility in scaling firms' operations (up and back) with fewer risks (see Jin and McElheran, 2024), without requiring large capital investments (Brynjolfsson et al., 2008) and producing positive effects on both size and productivity (Gal et al., 2019; Duso and Schiersch, 2025; Jin and Bai, 2022). This positive link has been found to be more pronounced for younger firms (Bloom and Pierri, 2018; DeStefano et al., 2023b; Jin and McElheran, 2024), suggesting that these firms disproportionately benefit from the use of cloud services by gaining access to ICT assets that would otherwise be inaccessible or risky to buy.

Yet, to our knowledge, there is limited evidence[1] exploring whether also smaller firms enjoy higher benefits from cloud use than larger firms, conditional on age. Investigating the impact of cloud adoption on small firms' growth is relevant for several reasons. First, the costs of ICT adoption and complementary intangible investments are large (De Ridder, 2024), in line with the positive relation between size and digital technologies adoption (see Zolas et al., 2020; Calvino and Fontanelli, 2023b; Cirillo et al., 2023). Furthermore, investments in ICT assets are irreversible. Cloud technologies represent a crucial service for small firms as they lower the fixed costs and the risk of digitalisation, allowing to store data and files, operate software, and undertake computationally intensive activities without owning the underlying physical IT facilities. This helps firms re-organise operations and increase their intangible capital, which is crucial to scale up (Coad et al., 2024). Second, notwithstanding the significant decline of cloud service prices in the 2010s (Byrne et al., 2018; Coyle and Nguyen, 2018), the purchase of cloud services in a market dominated by few large providers (for instance, Amazon, Microsoft, and Google – see discussion in Crémer et al., 2024) implies the existence of positive markups on the cost of using providers' ICT assets via cloud service. This suggests that when the need for ICT assets is high – as it is the case for large firms (Lashkari et al., 2024) – the purchase of cloud services may not be cost-effective.

In this work, we explore the heterogeneous impact of cloud technologies on the size growth

[1] While the relationship between IT outsourcing and firm size is partially addressed in Jin and McElheran (2024), the authors' cannot single out the adoption of cloud technologies in the data, and focus primarily on firms' age.

rate of firms based on a unique combination of four sources of micro data on French firms between 2005 and 2022 – French ICT surveys (2016, 2018 and 2020), administrative data from French firms' balance sheets (2005–2019), French matched employer-employee data (2005–2022), and the French business register (2005–2018). We focus on long run growth rates, in line with the idea that the effects of digital technologies may take time to materialise due to large and complex organisational changes characterised by uncertainty and implementation lags (Brynjolfsson and Hitt, 2003; Tambe and Hitt, 2012; Brynjolfsson et al., 2018; Acemoglu and Restrepo, 2020; Babina et al., 2024).

We find that cloud has a positive relationship with the growth rates of firms, and that this relationship is less pronounced for larger firms. We address potential endogeneity issues by adopting a causal identification strategy based on an endogenous treatment model (ET henceforth, see Heckman, 1976, 1978; Maddala, 1983; Vella and Verbeek, 1999), where the purchase of cloud services is our endogenous treatment variable. This latent variable model is widely used in research (some noteworthy examples are Shaver, 1998; King and Tucci, 2002; Campa and Kedia, 2002) as it addresses the issue of self-selection of firms into treatment (Hamilton and Nickerson, 2003; Clougherty et al., 2016). Furthermore, this approach allows for the inclusion of interaction terms between the endogenous treatment (cloud) and control variables (size) in a more parsimonious way compared to Two-Stage Least Squares (TSLS) (Wooldridge, 2015).

We employ lightning strikes density at the municipality (French commune) level, a source of spatial exogenous variation associated to investments in IT infrastructure (also see Andersen et al., 2012; Manacorda and Tesei, 2020; Guriev et al., 2021; Chiplunkar and Goldberg, 2022; Caldarola et al., 2023), to exogenously predict the adoption of cloud technologies by firms. In order to adopt cloud technologies, firms need to have access to stable and fast internet connection (Nicoletti et al., 2020; Garrison et al., 2015; Ohnemus and Niebel, 2016; DeStefano et al., 2023b). However, by causing energy spikes and dips, lightning strikes increase the maintenance costs of IT infrastructure, slowing down their diffusion (Andersen et al., 2012). Furthermore, lightning strikes lower the quality associated with broadband internet services, producing a four times larger frequency of broadband network failures during thunderstorms, if not adequately mitigated with additional and costly equipment (Schulman and Spring, 2011). Our instrument based on lightning strikes density reflects the trade-off faced by internet providers who will have to balance the costs of expanding the broadband network and the potential benefits that can be harvested by doing so.

Results from the endogenous treatment models show that the purchase of cloud technologies has a positive effect on firms' long run growth rates. However, the effect decreases for larger firms, suggesting that the diffusion of cloud services may help smaller firms scale up,

unlocking untapped growth potential. This finding holds across several robustness checks, including the change in the period wherein growth rates are computed, the change in the sample (the ICT survey collected during 2021 and referring to 2020, when COVID-19 boosted the digitalisation of firms – see Calvino et al. 2024), two additional types of econometric specification (long differences and Two-Stages Least Squares), and sample balancing methods (CEM).

We then extend our analysis by providing evidence on three possible mechanisms through which the purchase of cloud services may generate different firm-level growth rates across firms of different size. First, we leverage a unique feature of the French ICT surveys, which provide information on the type of cloud service purchased by firms, each of which represents a distinct function of cloud technologies within firms, and which are linked to different bundles of digital technologies. We distinguish firms purchasing cloud services for storing data and file, for running office, administrative or Customer Relationship Management (CRM) software applications, and for acquiring computing power. We find that only cloud technologies for software applications are linked to a higher growth performance of smaller firms, showing that cloud technologies can address small firms' barriers associated with the use of software solutions in office, administration and CRM. These are particularly pronounced for smaller firms, as they involve high fixed costs of IT adoption and complex changes to the organisational structure.

Second, we show that the relationship between firm size and cloud adoption is not mediated by age. Our results challenge the traditional view that smaller firms inherently grow faster, a perspective that finds only mixed support in the empirical literature (Coad et al., 2014; Haltiwanger et al., 2013). Third, we examine whether growth differences between small and large firms adopting digital technologies reflect decreasing returns to scale rather than technology-specific effects. To do so, we compare the relation between growth performance and three distinct digital technologies (cloud, Big Data Analytics, and E-commerce). Similar growth patterns across these technologies (positive, but decreasing with size) would suggest that the cloud-growth link is driven by scale effects. We show that, differently from cloud, the association of Big Data Analytics and E-commerce to growth is positive and does not diminish with size. Overall, the analysis of the mechanisms confirms that the purchase of cloud services enables a reorganisation of production processes through digitalisation that is critical for scaling operations.

Finally, we examine whether the share of cloud-using firms is related to the concentration of industrial sales shares (see Brynjolfsson et al., 2023; Bessen, 2020) by estimating the relationship between the average cloud intensity and the market concentration of French industries aggregated at the 2-digit sectoral level. Our analysis uncovers a mild and negative

correlation between the two, suggesting that the greater impact of cloud on size growth rate of smaller firms may mitigate, or at least won't exacerbate, increasing concentration trends in industries by helping small firms to grow.

In this regard, we contribute to the literature on the use of cloud by identifying a specific channel through which cloud and ICTs differentially affect sales across firms of various sizes: the use of cloud software applications. This is consistent with the evidence that cloud technologies enable the use of other digital technologies (McElheran et al., 2023; Calvino and Fontanelli, 2023b). Our findings contribute to the growing body of research on the role of cloud technologies in firm growth and the mechanisms through which this process unfolds. The extant literature has privileged an analysis of cloud technologies varying based on the age, rather than size, of firms: for instance, DeStefano et al. (2023b) finds that cloud adoption boosts the growth of young firms, attributing this effect to a decline in IT investment per employee. Similarly, Jin and McElheran (2024) shows that young firms disproportionately benefit from IT outsourcing, pointing to mechanisms that mitigate the effects of uncertainty. More broadly, our findings relate to the literature on ICT and firm growth, which generally suggests that larger firms benefit more from ICT diffusion. Looking specifically at the heterogeneous effect of ICTs across smaller and larger firms, Brynjolfsson et al. (2023) finds that the impact of ICTs on firm size is more pronounced among larger firms, supporting the view that these technologies allow them to replicate business processes across additional production units, access new markets, and boost sales without a proportional increase in workforce. Likewise, Babina et al. (2024) demonstrates that larger firms have ramped up their investments in AI over the past decade, facilitating their expansion into additional markets, a rationale that aligns with the findings of Aghion et al. (2023). Finally, Bessen (2020) shows that in IT-intensive industries, the largest firms experienced faster sales growth, and links the diffusion of IT proprietary assets with the increases in industry concentration. Lashkari et al. (2024) documents that larger French firms invest a higher share of their sales in IT. In this respect, our evidence supports the idea that not all ICTs have the same effects on firms. The distinctive characteristics of cloud technologies – particularly their role in enabling software adoption crucial for organisational innovation necessary to scale up – suggest that cloud may be an exception to the broader trend in ICT-driven firm growth.

Our findings have significant policy implications in the current context of increasing trends in industry concentration at play in many countries (Gutiérrez and Philippon, 2017; Bajgar et al., 2023; Grullon et al., 2019; Autor et al., 2020). ICT diffusion policies that support the digital transition could worsen the current state of competition, if they fail to differentiate between small and large firms, as ICT adoption has been linked to increased industry concentration (Bessen, 2020; Bajgar et al., 2025; Brynjolfsson et al., 2023; Lashkari et al., 2024).

However, our evidence suggests that smaller firms, which are generally less digitalised, have substantial untapped growth potential that could be unlocked through the broader diffusion of cloud technologies, thereby contributing to the mitigation of rising concentration.

The remainder of the paper is organised as follows. Section 2 discusses the sources of data used for the analysis and reports key summary statistics. Section 3 describes the econometric framework and identification strategy applied in Section 4, where the main results of the analysis are reported. Section 5 estimates the cloud-concentration relation. Section 6 summarises the key findings and discusses possible avenues for future research.

2 Data

In this section, we discuss the data employed in the analysis and present key summary statistics. Our analysis is based on four sources of microdata.

First, we use the 2016, 2018, 2020 versions of the French ICT survey (*"Enquête sur les Technologies de l'Information et de la Communication (TIC)"*), which is managed by the INSEE (the French statistical office).[2] Each wave of the survey includes a rotating sample counting approximately 9000 firms from both manufacturing and non-financial market-services sectors. The sample is representative for firms with 10 or more employees and is exhaustive for those with over 500 employees.[3] We exclude firms belonging to sectors 62 ("Computer Programming, Consultancy And Related Activities") and 63 ("Information Service Activities") of the NACE classification. The possible positive effect of cloud adoption on the performance of these firms could indeed be driven by the sales of cloud and ICT services to other firms, rather than by increases in digital intensity. The survey questions focus on the use of advanced digital technologies in 2015, 2017 and 2019, respectively captured by the 2016, 2018, and 2020 survey waves.[4] This dataset is characterised by a greater level of detail and representativeness when compared to other commercial surveys. Additionally, it can be merged to other sources of French firms' data thanks to the *Siren* code, a unique identifier attributed to French companies at their birth.

Part of the ICT survey is dedicated to questions on cloud use by firms. Specifically, firms are asked whether they used cloud technologies in the previous year.[5] Firms are asked the

[2]Further information about each ICT survey can be found here for 2016 https://doi.org/10.34724/CASD.49.2273.V1, here for 2018 https://doi.org/10.34724/CASD.49.2876.V2 and here for 2020 https://doi.org/10.34724/CASD.49.4086.V1.

[3]It is therefore challenging to exploit the panel dimension of these datasets. Approximately six thousands firms are present in at least two of the waves of the survey, and are mostly large firms.

[4]The questions about advanced digital technologies are updated annually, although the ICT surveys run in different years may not include questions about the same technologies.

[5]Additional questions about cloud technologies are present in the 2021 wave, relative to the year 2020

following question:

"Did your enterprise buy cloud computing services? (Excluding cloud services provided for free.)".

Cloud services are defined as follows in the survey:

"Cloud computing (or cloud) refers to computing services used over the internet to access software, computing power, storage capacity, etc. These services must have the following characteristics:

- *They are delivered by servers from service providers.*
- *They are easily scalable up or down (for example, the number of users or changes in storage capacity).*
- *Once installed, they can be used "on-demand," without human interaction with the provider.*
- *They are paid either by the user or based on the capacity used or services provided.*

Cloud computing may include connections via a virtual private network (VPN)."

Furthermore, the survey provides information on the different types of cloud services purchased by firms, distinguishing them into six non-exclusive categories: mail, data storage, file storage, accounting software, office software, customer relationship management (CRM) software, and computing power. We define a cloud user as a firm that purchases cloud services in at least one of the latter five categories. We discard the first category of cloud usage (i.e., mail), as it is unlikely conducive to producing organisational changes in the firm's structure and, therefore, may not capture the effects of cloud on firm performance. Our main cloud use variable thus takes the form of a dummy, indicating whether firms use cloud technologies or not. Additionally, we provide results for different categories of cloud usage. We define the dummy variables 'Cloud - Storage', 'Cloud - Software', and 'Cloud - Computing Power.' A firm is considered to use cloud for storing data ('Cloud - Storage') when it purchases cloud services for storing data or files, to use cloud for software ('Cloud - Software') when using cloud services for accounting, office, or software for managing customers relationships, and to use cloud for computing power ('Cloud - Computing Power') when it purchases cloud services for borrowing external IT processing capacity.

respectively. However, in the context of our identification strategy (see Section 3), the use of this wave implies the inclusion of the COVID-19 pandemic years in the dataset. The pivotal role of digital technologies during the pandemic makes it challenging to precisely estimate the effect of cloud on performance in normal times. We chose to employ the 2016, 2018 and 2020 ICT surveys accordingly. We provide a robustness check confirming our results when using the 2021 version of the ICT survey in the Appendix (see Table A3).

Second, we match the ICT survey with the administrative data from French firms' balance sheets (FARE) covering the 2005–2022 period.[6] This dataset provides information on firm sales, age, employment, and exporter status, as well as physical and intangible capital.[7] Intangible capital is not available before 2009; this will limit our baseline estimation to the 2009–2018 period.[8] These variables allow us to provide a complete picture of firms buying cloud technologies, and to control for potential links between size growth and firm characteristics otherwise conflated with cloud usage, for instance the age of firms.

Third, we employ the information on the stocks of establishments by firm from the French business register. This data is used to build a binary variable indicating if the firm is multi establishment.

Finally, we match the ICT survey with French employer-employee data (DADS) in 2005–2020.[9] This data allow us to build the firm-level share of hours worked by ICT workers and by ones specialised in R&D hired by the firm (named ICT share and R&D share hereafter), and the average hourly wage of managers and engineers in a firm. [10] We consider ICT workers to be employees falling within the 4-digit classes 388a, 388b, 388c, 388d, 388e, 478a, 478b, 478c, 478d, and 544a of the 2003 French PCS classification. Instead R&D workers fall within the 4-digit classes 383a, 384a, 385a, 386a, 388a, 473a, 473b, 474a, 475a, and 478a, as suggested by the classification of occupations into functions provided by the French National Statistical Institute.[11] These classes specifically target occupations with a significant focus on ICTs and R&D.

2.1 Summary statistics

Before investigating the relationship between cloud adoption and firm growth rates, we sketch out the characteristics of the sample under consideration, highlighting some general differences between cloud users and non-users. To start with, the upper block in Table 1 shows

[6] Additional details about this dataset can be accessed here: https://doi.org/10.34724/CASD.42.3654.V1.

[7] Data on sales and capital are in real terms. Sales and physical capital have been deflated at the 2-digit sector level. Data on intangible capital have been deflated exploiting the deflators provided by INTANPRO-EUKLEMS (Bontadini et al., 2023).

[8] However, in Section 4.2 we discuss a series of robustness checks on the 2005–2019 period.

[9] Further information about DADS here https://www.casd.eu/en/source/all-employees-databases-business-data.

[10] It is worth noting that ICT workers are part of the techies definition used in Harrigan et al. (2021). The techies definition encompasses all occupations within the 2-digit classes 38 (executives and engineers) and 47 (Technicians) of the 2003 French PCS classification. The mentioned PCS codes cover roles such as R&D personnel in IT, computer engineers, developers, database administrators, and IT technicians. Further details and information on the PCS classification can be found here https://www.insee.fr/fr/statistiques/fichier/2401328/Brochure_PCS_ESE_2003.pdf.

[11] The classification can be found here https://www.insee.fr/fr/statistiques/1893116.

the share of cloud-adopting French firms grew of a remarkable 27.3 percent between 2015 and 2017 and of 55.8 percent between 2015 and 2019, indicating a rapid diffusion of cloud technologies.

Foreseeably, the share of cloud users increased in each sector considered (see Table 1). Notwithstanding the exclusion of IT services from the sample, cloud technologies are commonly adopted by firms in the ICT sector: their adoption rates span from 41.2 percent of total firms in 2015 to 58.7 percent in 2019. Large changes in the rate of adoption are registered across all sectors, suggesting that cloud technologies are rapidly diffusing everywhere. In 2019, professional and scientific, real estate and administrative firms display large rate of cloud use with respect to other sectors. Low-productivity sectors such as accommodation, transportation services, and wholesale & retail display lower levels of adoption, with limited growth over time. The most represented industry in the sample, manufacturing, also experienced sizeable growth in the use of cloud technologies, moving from 27.2 percent of adopters in 2015 to 41 percent in 2019 – a growth rate of 50.7 percent.

We now describe the general characteristics of cloud users versus non-users. As shown in Table 2, the most common use of cloud technologies in both years is to store data and files, although this particular type of cloud exhibits little growth between 2015 and 2019. This is likely due to the fact that cloud technologies for data storage were the first to be commercialised, and have reached earlier maturity. This is followed by software applications – which increased from 66,6 percent to 83.4 percent – and lastly by ICT applications, such as the acquisition of computing power available to the firm. Given also the lower share of adoption at the first available year observed, the latter type is the fastest growing cloud technology type among adopters, increasing of 44.8 percent in four years (from 21.2% in 2015 to 30.7% in 2019). Overall, Table 2 also shows that cloud-adopting firms tend to be older and remarkably larger in terms of sales. They also own a bigger stock of physical and intangible capital, employ a higher share of ICT workers (more than twice as bigger than non-adopters), and are more likely to export. Finally, cloud adopting firms are also more likely to deploy more than one productive plant or unit.

Table 1: Share of cloud users in 2015, 2017, and 2019: total for France and by industry.

France	2015	2017	2019
All Firms	25.64%	32.54%	40.00%
Count of Cloud Users	2047	2498	2490
Total Number of firms	7985	7676	6225
Industry	2015	2017	2019
Accomodation & Food	16.93%	20.05%	24.20%
Administrative	28.82%	34.41%	45.15%
ICT	40.88%	57.77%	57.04%
Manufacturing	27.21%	33.97%	41.01%
Professional & Scientific	38.45%	44.60%	50.00%
Real Estate	29.34%	42.86%	58.68%
Transportation & Storage	23.08%	30.68%	33.76%
Utilities & Construction	17.73%	24.22%	32.37%
Wholesale & Retail	22.71%	29.37%	37.75%

Table 2: Summary statistics by cloud user and year.

Year	2015		2017		2019	
Cloud – Storage		87.40%		89.19%		89.44%
Cloud – Software		66.59%		76.78%		83.45%
Cloud – Computing Power		21.20%		26.02%		30.76%
Age	27.93	31.83	28.49	33.02	29.15	32.82
Sales (Thousands €)	86057.38	343173.10	67374.23	357258.40	64121.75	370306.80
Physical Capital (Thousands €)	32366.84	168992.40	19946.08	182924.80	14536.88	183237.70
Intangible Capital (Thousands €)	375.41	2257.90	192.52	2184.91	150.06	1947.82
ICT Share	2.25%	5.16%	1.69%	4.67%	2.17%	5.38%
R&D Share	2.48%	5.31%	2.34%	5.04%	2.45%	5.57%
Exporter	42.62%	64.09%	41.31%	63.73%	39.54%	61.41%
Multi Establishment	42.05%	67.50%	40.42%	65.93%	39.68%	64.16%
Hourly Wage (Managers & Engineers, €)	24.20	31.77	23.94	32.75	24.88	33.10

3 Methods

In this work we aim at studying the effects of cloud purchases on firms' sales growth. Specifically, in our baseline estimation we choose as dependent variable the 5-year logarithmic difference in sales. Our baseline regression model reads as follows:

$$\begin{aligned}&\text{Sales Growth}_{i,t,t-5} = \\ &a + \beta_1 \text{Cloud}_{i,t} + \beta_2 \text{Cloud}_{i,t} \cdot \text{Log-Sales}_{i,t-5} + \beta_3 \text{Log-Sales}_{i,t-5} + \beta_X \text{X}_{i,t-5} + \\ &+ \text{2-digit Ind.}_j + \text{Region}_r + \text{Year}_t + \epsilon_{i,t}\end{aligned} \tag{1}$$

where $\text{Sales Growth}_{i,t,t-5}$ is the logarithmic difference between sales in t and $t-5$, and $\text{Cloud}_{i,t}$ is the cloud dummy which takes value 1 if the firm reported to adopt cloud technologies in or before the survey year t. Following the empirical strategy by Forman and McElheran (2025), we include a vector of controls $\text{X}_{i,t-5}$ measured at the beginning of the period in $t-5$. These include the logarithms of age, physical and intangible capital, the share of workers specialised in ICT and R&D roles, the average hourly wage of managers and engineers, and two dummies for exporter and multi establishment status. We include fixed effects for industries (2-digit Ind._j), regions (Region_r), and years (Year_t).

Long-run growth rates. We test the relationship between cloud adoption and growth using long-term growth rates. Doing otherwise (that is, employing short-term growth rates, such as annual growth rates, as the dependent variable) may fail to capture the relationship of interest for three key reasons. First, extensive literature shows that short-term firm growth rates often align with Gibrat's Law (Gibrat, 1931), particularly in the case of less young surviving firms (Santarelli et al., 2006; Lotti et al., 2003, 2009; Fontanelli, 2024).[12] This implies that short-term fluctuations are largely stochastic and may not capture growth drivers materialising in the long-run. Second, as highlighted in several studies (Brynjolfsson and Hitt, 2003; Tambe and Hitt, 2012; Brynjolfsson et al., 2018; Acemoglu and Restrepo, 2020; Babina et al., 2024), the effects of the diffusion of digital technologies often require time to materialise. This delay stems from from the uncertainties and implementation lags caused by the substantial and complex organisational changes associated with ICT adoption. This reasoning likely applies to cloud technologies as well (see the discussion in the Introduction).

Measuring cloud use in t. The coefficient of $\text{Cloud}_{i,t}$ captures the average long-term increase in sales due to cloud use, conditional on controls. The interaction term between sales and cloud accounts for potential differences in the relationship between performance and cloud usage between smaller and larger firms. While the ICT surveys do not provide

[12] According to Gibrat's model, firm size evolves as $\log s_{i,t} = \log s_{i,0} + \sum_{\tau=1}^{t} \epsilon_\tau$, where the growth rate ϵ_τ is an identically and independently distributed random variable.

information on the year of first cloud adoption by firms, they do indicate whether a cloud service was purchased by the firm in the survey year. The absence of information on the first year of cloud adoption by firms introduces a measurement challenge. Specifically, the adoption of cloud captured by $Cloud_{i,t}$ could have happened at any point before time t. Consistently with the empirical strategies adopted in the literature measuring the effects of ICT on firms (Forman et al., 2012; Forman and McElheran, 2025; Babina et al., 2024), we measure cloud adoption in t for three main reasons.

First, by measuring cloud closer to t, we ensure a more accurate comparison between firms using and not using cloud services. Indeed, the farther back we measure cloud adoption, the more likely the model described by Equation 1 will conflate cloud users and non-users. Many firms not using cloud in $t-5$ likely adopted it in subsequent years, given the rapid diffusion of cloud technologies (see Table 2). Indeed, the adoption of cloud services by firms before 2009 was highly unlikely in the US (Bloom and Pierri, 2018), due to the high prices associated with cloud service provision until the early 2010s (Byrne et al., 2018; Coyle and Nguyen, 2018). This suggests that French firms in our sample, observed in 2015, 2017, and 2019, likely began adopting cloud technologies in the early 2010s, that is before these technologies could produce the long term organisational changes that we aim to capture with our empirical strategy.

Second, measuring cloud adoption at $t-5$ would neglect the early diffusion period of cloud technologies (2010-2015), as well as the initial effects of cloud use on growth rates, which are particularly relevant in our case. Measuring cloud in t better captures these early adoption phases, and thus picks up the lion's share of performance gains owing to cloud use, even if its use since the start of the period cannot be ascertained.

Finally, measuring the controls before cloud adoption (at $t-5$ in our case) mitigates the risk of bad controls bias in estimating the cloud-growth relationship (Angrist and Pischke, 2008). Specifically, controls measured in 2010, 2012, or 2014 are less likely to have been influenced by cloud diffusion, which was either negligible or in its early stages at those times. We conduct a series of robustness checks by varying the year in which cloud use is measured, from t to $t-5$ (see Table A2 in the Appendix). These exercises demonstrate that our results are robust to different specifications of the timing of cloud use.

Control variables. The vector of variables $X_{i,t-5}$ includes a comprehensive set of time-varying firm characteristics. The logarithm of age, sales, physical capital, and intangible capital; the share of hours worked in ICT and R&D occupations; the logarithm of the average hourly wage of managers; dummies for export and multi establishment status; and fixed effects for 2-digit industries, regions, and years. The ICT and R&D shares serve as a proxy for the intensity of digitalisation and innovation within the firm and, in our regression setting,

clean the relationship between cloud and performance from the correlation between cloud and other performance-enhancing innovations and digital technologies. The average hourly wage of managers is an approximate measure of the quality of managers in the firm.

The inclusion of firms' sales among the control variables mitigates a key source of bias. Larger firms are more likely to innovate and adopt digital technologies, including cloud. Therefore, controlling for sales reduces the potential confounding effects of other digital technologies and innovations adopted by larger firms. Firm age is also included as a control to account for new managerial and ICT capabilities potentially affecting both cloud use and performance (Bloom and Pierri, 2018). Since younger firms may be more likely to adopt emerging technologies like AI (Calvino and Fontanelli, 2023a), controlling for age helps reduce bias from omitted variables while also cleaning the size-cloud interaction.

Controlling for firm capital further addresses endogeneity issues. Intangible capital, for instance, includes the firm-level value of complementary digital technologies, such as proprietary data and software, and factors which may impact firm growth rates, as it is also includes the value of assets such as patents, and trademarks (Corrado et al., 2021). Physical capital affects the feasibility of cloud adoption, as firms with lighter capital structures may find cloud technologies more advantageous. Export and multi establishment dummies are included to control for firms' access to multiple markets, addressing the self-selection of firms with higher growth potential and growth strategies prioritising growth into cloud usage. Finally, we include fixed effects for 2-digit NACE industries, regions, and years to capture average characteristics specific to industries, geographic locations, and time periods.

3.1 Identification strategy

We first estimate Equation 1 via Ordinary Least Squares (OLS). Notwithstanding the presence of several controls and fixed effects, the estimation could still exhibit biases owing to endogeneity in the relationship between cloud and size growth. This is because the adoption of the former may be associated with unobserved characteristics of firms, such as managerial and productive capabilities.

Endogenous Treatment model. We employ a measure of lightning strikes density at the postcode level to identify cloud adoption in Equation 1 in an Endogenous Treatment regression framework (referred to as ET hereafter, see Heckman, 1976, 1978; Maddala, 1983; Wooldridge, 2015), a latent variable approach widely used in research (Shaver, 1998; King and Tucci, 2002; Campa and Kedia, 2002) and closely related to conventional Instrumental Variable (IV) models such as Two-Stage Least Squares (TSLS, see Vella and Verbeek, 1999). This model allows to address endogeneity issues, such as self-selection into treatment, by simultaneously estimating a selection and an outcome model via Maximum Likelihood Es-

timation (MLE). The ET method offers two advantages. First, it employs a Probit model for the selection equation, which does not generate predicted values outside the unity range of the probability space, unlike Linear Probability Models (LPM) (Hamilton and Nickerson, 2003; Clougherty et al., 2016). This is particularly relevant in our case, where the endogenous variable – cloud adoption – is dichotomous. Second, the ET model belongs to the family of control function approaches. As such, it allows for the introduction of interaction terms between the endogenous treatment and other variables in a more parsimonious manner compared to the standard TSLS procedure (see the discussion in Wooldridge, 2015), which employs the interactions between the exogenous variable and the interacted variable as instruments. For robustness, we report TSLS results in Table A5 in the Appendix, which align closely with our main regression findings.

In the context of the ET framework, Equation 1 reads as follows:

$$\begin{aligned} &\text{Sales Growth Rate}_{i,t,t-5} = \alpha + \beta_1 \text{Cloud}_{i,t} + \beta_2 \text{Cloud}_{i,t} \cdot \text{Log-Sales}_{i,t-5} + \beta_{\mathbf{X}} \mathbf{X}_{i,t-5} \\ &+ \text{2-digit Ind.}_j + \text{Region}_r + \text{Year}_t + \epsilon_{i,t} \\ &\text{Cloud}_{i,t} = \begin{cases} 1, & \text{if } \beta_Z \text{Log-Lightning Density}_i + \beta_{\mathbf{X}} \mathbf{X}_{i,t-5} + \omega_{i,t} > 0 \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{2}$$

where $\text{Cloud}_{i,t}$ is the endogenous dummy variable for the use of cloud, $\mathbf{X}_{i,2011}$ is a vector of controls including the same variables of Equation 1 and the variable $\text{Log-Lightning Density}_i$, which is excluded from the outcome equation. The estimation of the ET model employs the Full Information Maximum Likelihood (FIML) method, concurrently estimating the selection equation (using the dummy variable for cloud use as the dependent variable) and the outcome equation (i.e., the sales growth rate regression).[13] The ET model produces estimates robust to the presence of specification errors when an additional variable is incorporated into the selection equation, but omitted from the outcome equation. This variable must adhere to two conditions (Puhani, 2000), similarly to the standard IV conditions in TSLS models: it must strongly predict the endogenous dummy variable (i.e., exhibit relevance) and must be exogenous (i.e., satisfy the exclusion restriction) in the presence of other controls.

Lightning strikes density. To estimate the ET model we exploit a source of spatial exogenous variation associated to the cost and quality of ICT investments. We build on evidence by Andersen et al. (2012) that ICT investments are slowed down by the incidence of a natural phenomenon: lightning strikes. The argument proposed by Andersen is based on the correlation between lightning strikes density and productivity across US states which emerged in the 1990s, years in which ICTs started to diffuse more widely. By causing energy spikes and

[13] This process assumes the joint normality of errors (ϵ_i, ω_i).

dips, lightning strikes increase the cost of digital infrastructures and technologies, slowing down their diffusion. In particular, it has been shown that the incidence of lightning strikes reduces the quality of broadband connections (Yu et al., 2023), with broadband network failures being four times more likely during thunderstorms (Schulman and Spring, 2011). However, in order to adopt cloud technologies, firms need to have access to reliable, fast and state-of-art internet connection (Nicoletti et al., 2020; Garrison et al., 2015; Ohnemus and Niebel, 2016), which is fostered by the presence of physical infrastructures providing high-quality broadband network.

Indeed, DeStefano et al. (2023b) identify cloud adoption by exploiting exogenous variation in access to broadband internet. Their instrument is based on a dummy variable measuring the enabled access to fibre broadband at a given time, and the distance between firms and the closest telephone exchange site. Absent this information in our data, we resort to an instrument that exogenously predicts the diffusion of broadband internet, which in turn is more likely to be available in areas with lower density of lightning density (also see Manacorda and Tesei, 2020; Guriev et al., 2021; Chiplunkar and Goldberg, 2022; Caldarola et al., 2023).

Instrument validity. To illustrate the relevance of lightning strike density in predicting the deployment of broadband internet, we report in Table 3 the estimated results of a linear probability model employing the presence of fast broadband as dependent variable and the log of lightning strike density as the independent variable. Fast broadband is a dummy variable taking value 1 when firms' broadband connection is faster than 100 Mbit per second. We source this information from the ICT surveys described at the beginning of Section 2. Information on lightning strike density is obtained from the World Wide Lightning Location Network (WWLLN) Global Lightning Climatology and Time Series (Kaplan, 2023). The raw WWLLN data is a grid of 5-arcminute cells, each containing information on the count of daily lightning strikes per square kilometre. Data collection occurs daily and spans from 2008 to 2020. Building on evidence that the incidence of lightning strikes is a stationary phenomenon (Andersen et al., 2012)[14], we are interested in constructing a measure that captures the geographical exposure of geographical areas to this phenomenon (in our case, the French municipality where firms are located).[15] We calculate the average lightning strike density for each French municipality, based on the density values contained in the cells that fall within each municipality over the period 2008 – 2017.[16] The resulting metric, representing average daily lightning strikes per square kilometre in each municipality, is then multiplied by the

[14]While it is well known that lightning strikes are a stationary phenomenon, we provide evidence on the stationarity of our time series in Table A1 and Figure A1 in the Appendix.

[15]French municipalities, or communes, are the smallest administrative subdivision in France, acting as local authorities. There are 34,826 communes in the country.

[16]In the case of cells that fall over a border between two or more communes, the cropped cells weight in each commune based on the percentage of the cell falling within each of them.

Table 3: Lightning strikes density and fast broadband internet adoption.

	Model 1	Model 2
Log Lightning Density	-0.0063***	-0.0038***
	(0.0014)	(0.0013)
Log Sales		0.0092***
		(0.0020)
Log Physical Capital		0.0007
		(0.0014)
Log Intangible Capital		-0.0020*
		(0.0011)
Log Age		-0.0031
		(0.0027)
ICT Share		0.0254**
		(0.0116)
R&D Share		-0.0189
		(0.0154)
Exporter		0.0125**
		(0.0050)
Multi Establishment		0.0066***
		(0.0022)
Log Average Hourly Wage Manag. & Ing.		0.0063***
		(0.0016)
Observations	21,886	21,886
Industry FE	Yes	Yes
Year FE	Yes	Yes
Region FE	Yes	Yes
Adjusted R^2	0.0153	0.0308

Note: The dependent variable measures the adoption of fast broadband at time t. Log lightning density is the population weighted density of daily lightning strikes per square kilometre in the INSEE area, averaged between 2008 and 2017. All the time-varying control variables are measured at $t-5$. Log Sales, Physical, and Intangible Capital are expressed in Log of Euros. Age is the firm age in years since the establishment. ICT and R&D shares express, respectively, the share of hours worked in ICT and R&D occupations. Exporter is a dummy that identifies whether the firms is engaged in exporting activities. Multi Establishment identify whether the firm has more than one economic establishment. Log Average Hourly Wage Manag. & Ing. measures the log of the hourly wage of management and engineers, expressed in Euros. All regressions are estimated using OLS. Standard errors are clustered at the NACE 2-digit level. *** p<0.01, ** p<0.05, * p<0.1.

surface area to retrieve the total amount of lightning strikes, and weighted by population. As shown in Table 3, lightning strike density is negatively associated to the use of fast broadband connections by firms, in line with our discussion above. The relation is strongly significant,

even when additional controls are included. Among these, we also find that the ICT share has a positive and significant association with fast broadband. This provides evidence that the lightning strike density is directly correlated with the presence of fast broadband, even after controlling for the share of hours worked by ICT specialists.

There are several reasons in support of the exogeneity of our instrument, conditional upon controlling for firms' observable characteristics. First of all, as documented by (Andersen et al., 2012), the economic effect of lightning strikes has only manifested after the beginning of the diffusion of ICTs and internet, as the incidence of this random natural phenomenon has discouraged investments in ICT equipments due to the additional costs imposed on infrastructure management. Moreover, given the stationary nature of this phenomenon (see Table A1), it is unlikely that other natural events such as climate shocks could account for the reduced form correlation between lightning and firm growth. Also, natural phenomena are not likely to align with administrative boundaries, where firm-specific economic policies are implemented. This rules out the concern that lightning-prone areas systematically differ in terms of regulatory environments that are directly related to firms' growth. It is also unlikely that firm-level growth determinants – such as managerial capabilities, industry composition, or business cycles – are directly affected.

It could be argued that the availability of fast broadband infrastructure would simultaneously encourage firms to adopt cloud technologies and drive broader digitalisation or innovation efforts, leading to a positive effect on firm performance. However, this is unlikely due to the inclusion of key controls in our regression model. Controlling for workers specialised in ICT and R&D occupations accounts for the firm's level of digitalisation and its propensity to innovate. Additionally, intangible capital captures the economic value of patents, trademarks, and digital assets owned by firms, further mitigating this concern.

4 Results

In this section we discuss our estimation results. In Section 4.1, we discuss the estimation of the econometric models described in Section 3. Next, in Section 4.2 we discuss a battery of robustness checks.

4.1 Cloud adoption and firm growth rates

OLS results. We discuss the association between adoption of cloud technologies and growth rates of French firms in 2015, 2017 and 2019, using the strategy set out in Section 3. It is worth recalling that, in our baseline model, we consider the growth rate as the log difference

Table 4: Baseline model.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Cloud	0.0405*	0.1064***	0.2738***	0.2676***	0.2780***	0.2649***
	(0.0209)	(0.0232)	(0.0850)	(0.0755)	(0.0860)	(0.0768)
Log Sales		-0.0476***	-0.0406***	-0.0267***	-0.0384***	-0.0265***
		(0.0041)	(0.0038)	(0.0040)	(0.0037)	(0.0042)
Cloud*Log Sales			-0.0170**	-0.0175***	-0.0177**	-0.0175***
			(0.0067)	(0.0060)	(0.0068)	(0.0061)
Log Physical Capital				0.0188***		0.0192***
				(0.0045)		(0.0044)
Log Intangible Capital				-0.0179***		-0.0183***
				(0.0046)		(0.0046)
Log Age				-0.0922***		-0.0907***
				(0.0091)		(0.0091)
Exporter				0.0212**		0.0186*
				(0.0104)		(0.0100)
Multi Establishment				-0.0029		-0.0022
				(0.0066)		(0.0069)
ICT Share					0.1350***	0.1184**
					(0.0498)	(0.0452)
R&D Share					0.2251***	0.1953***
					(0.0439)	(0.0445)
Log Average Hourly Wage Manag. & Ing.					-0.0075***	-0.0028
					(0.0028)	(0.0025)
Observations	21,886	21,886	21,886	21,886	21,886	21,886
Adjusted R^2	0.0334	0.0745	0.0760	0.114	0.0805	0.117
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table reports the results of the estimation of Equation 1. The dependent variable measures the sales rate of growth of the firm between t and $t-5$. Cloud is a dummy variable identifying the adoption of cloud technologies by time t. All the time-varying control variables are measured at $t-5$. Log Sales, Physical, and Intangible Capital are expressed in Log of Euros. Age is the firm age in years since the establishment. ICT and R&D shares express, respectively, the share of hours worked in ICT and R&D occupations. Exporter is a dummy that identifies whether the firms is engaged in exporting activities. Multi Establishment identify whether the firm has more than one economic establishment. Log Average Hourly Wage Manag. & Ing. measures the log of the hourly wage of management and engineers, expressed in Euros. All regressions are estimated using OLS. Standard errors are clustered at the NACE 2-digit level. *** p<0.01, ** p<0.05, * p<0.1.

between the sales of the firm five years before the survey period (2015, 2017 and 2019). The main explanatory variable identifies firms that use cloud technology on the survey year. All covariates are measured five years before the survey year. The results of the estimation are summarised in Table 4.

Model 1 of the table shows that there is a positive and statistically significant relationship

between the use of cloud technologies and the long run growth rates of sales. Adding the log of sales (Model 2) leads to a notable increase in the magnitude of the coefficient on cloud. This may be due to the higher frequency of cloud use among larger firms (see Section 2.1), coexisting with a smaller impact of cloud use on their performance. Also, this hints to the fact that possible effects of cloud on firm performance may be mediated by their size. Moreover, this proxy for firm size enters with a negative sign, consistently across specifications (Models 2 to 5). This suggests that large firms grow more slowly than smaller ones. This result is robust to the addition of relevant covariates (Model 3). Analogously, age has a negative relation with firm performance. The ICT and R&D shares are positively related with firm performance in the long run, in line with existing evidence (Acemoglu et al., 2022; Brynjolfsson et al., 2023). Finally, exporting firms are characterised by better long run performance, as they are able to leverage a larger number of markets.

We are particularly interested in understanding whether the adoption of cloud technologies affects differently smaller and larger firms. To test this hypothesis, we focus on the interaction term between cloud adoption and log sales (Models 3 to 6 in Table 4). In all specifications, the interaction term takes a negative sign, confirming that larger firms benefit less – in terms of sales growth – than smaller firms. Thanks to the diffusion of cloud technologies, smaller firms have access to affordable means to increase the scale of their operations by digitalising production processes. They can externalise digital storage space, software tools and computing power, thus avoiding irreversible investments into the physical facilities (such as servers or computing clusters) or specialised employees (engineers specialised in ICT, such as computer networks engineers) required to this aim.

Endogenous Treatment results. As shown in Section 2.1, cloud adopting firms are systematically different from non adopting ones, revealing the possible presence of self-selection into the use of this technology. In order to address this concern, and to offer a causal interpretation of the effect of cloud technologies on firm growth rates, we proceed by estimating the Endogenous Treatment model described by Equation 2. This empirical approach has been designed following the identification strategy illustrated in Section 3.1, based on the spatial exogenous variation in the density of lightning strikes.

The results of the ET estimation are reported in Table 5. We start by focusing on Model 1, which shows two columns: "Selection" for the selection equation of the ET (i.e. the Endogenous Treatment equation in the second row of Equation 2), and "Performance" which displays the results of the outcome equation (first row of Equation 2). We start from the selection equation in Model 1. As expected, the coefficient on our instrument based on the average lightning strike density shows a negative sign and strong statistical significance, indicating that firms located in areas with high incidence of lightnings per inhabitant are less

Table 5: Endogenous treatment model.

	Model 1		Model 2	
	Growth Rate	Selection	Growth Rate	Selection
Cloud	0.0611*		0.2989***	
	(0.0333)		(0.0715)	
Cloud*Log Sales			-0.0181***	
			(0.0055)	
Log Sales	-0.0314***	0.1797***	-0.0280***	0.1805***
	(0.0042)	(0.0212)	(0.0066)	(0.0216)
Log Physical Capital	0.0193***	0.0334**	0.0189***	0.0333**
	(0.0045)	(0.0131)	(0.0045)	(0.0131)
Log Intangible Capital	-0.0191***	-0.0086	-0.0183***	-0.0088
	(0.0048)	(0.0185)	(0.0045)	(0.0186)
Log Age	-0.0912***	-0.0904***	-0.0899***	-0.0897***
	(0.0091)	(0.0164)	(0.0093)	(0.0171)
Exporter	0.0215*	0.2064***	0.0168	0.2059***
	(0.0113)	(0.0796)	(0.0114)	(0.0798)
Multi Establishment	-0.0007	0.1382***	-0.0035	0.1377***
	(0.0065)	(0.0219)	(0.0069)	(0.0221)
ICT Share	0.1317***	0.5303***	0.1129**	0.5289***
	(0.0449)	(0.1689)	(0.0472)	(0.1697)
R&D Share	0.1932***	0.1030	0.1941***	0.1032
	(0.0446)	(0.1397)	(0.0446)	(0.1400)
Log Average Hourly Wage Manag. & Ing.	-0.0007	0.0354***	-0.0030	0.0349***
	(0.0024)	(0.0105)	(0.0025)	(0.0105)
Log Lightning Density		-0.0573***		-0.0573***
		(0.0108)		(0.0112)
Observations	21,886	21,886	21,886	21,886
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes

Note: The table reports the results of the estimation of Equation 2. In each model, the first column reports the results of the outcome equation, while the second column refers to the selection equation. The dependent variable measures the sales rate of growth of the firm between t and $t-5$. Cloud is a dummy variable identifying the adoption of cloud technologies by time t. All the time-varying control variables are measured at $t-5$. Log Sales, Physical, and Intangible Capital are expressed in Log of Euros. Age is the firm age in years since the establishment. ICT and R&D shares express, respectively, the share of hours worked in ICT and R&D occupations. Exporter is a dummy that identifies whether the firms is engaged in exporting activities. Multi Establishment identify whether the firm has more than one economic establishment. Log Average Hourly Wage Manag. & Ing. measures the log of the hourly wage of management and engineers, expressed in Euros. Log lightning density is the logarithms of the (population weighted) density of daily lightning strikes at the municipality level, averaged between 2008 and 2017. All regressions are estimated using a Full Information Maximum Likelihood (FIML) estimator. Standard errors are clustered at the NACE 2-digit level. *** p<0.01, ** p<0.05, * p<0.1.

likely to adopt cloud technologies.

The Performance column in Model 1 shows the results of the outcome equation in the ET model. The coefficient of cloud on firm growth rates turns out to be positive and statistically significant, and one third smaller than the one of the OLS coefficient in Table 4, Model 3. This indicates that the OLS coefficient is likely to suffer from upward bias, due to the sources of endogeneity discussed in the previous section. The results for Model 2 in Table 5 include the interaction term between firm size and cloud adoption, which indicate that smaller firms reap relatively more benefits in terms of sales growth after adopting cloud technologies, when compared to larger ones. Concerning the change in the coefficient of the interaction in the OLS, the stronger negative effect of the interaction term in the ET model (Table 5, Model 2) with respect to its OLS counterpart (Table 4, Model 6) indicates that OLS (slightly) underestimates the larger benefits reaped by smaller firms. This could be explained by the fact that OLS fails to account the positive selection of larger firms into cloud adoption, understating how less beneficial cloud is for them relative to smaller firms.

The evidence discussed in this section confirms that cloud has a positive impact on the performance of firms (see also Duso and Schiersch, 2025; DeStefano et al., 2023b; Jin and McElheran, 2024). This aspect is consistent with the existence of a scale without mass dynamic (Brynjolfsson et al., 2008), that is a process where ICT allows the expansion of firms operation, organisation and processes with less physical constraints. Nevertheless, concerning the diffusion of cloud technologies among French firms, we show that the returns to cloud become smaller as size increases. In light of the existing evidence on the higher returns gained by larger firms adopting digital technologies (Brynjolfsson et al., 2023; Babina et al., 2024), our evidence suggests that the intrinsic characteristics of cloud technologies make them different from other ICTs.

4.2 Robustness checks

We conduct a battery of robustness checks, in order to further support the results summarised in the previous subsection. For the sake of conciseness, the results of this final exercise are shown in the Appendix.

Shifting or enlarging the growth period. The first robustness check involves changing the time interval initially defined by Equations 1 and 2, and the timing of cloud (see Table A2 in the Appendix). First, we show that the results remain consistent when cloud usage is measured at the start of the growth period.[17] Next, we estimate the model using

[17]This approach uses balance sheet data up to 2022 and ICT survey data in 2015 and 2017 waves, as data for 2024 are not yet available to compute the forward 5-year growth rate of firms in the 2019. New waves of balance sheet data typically become 5-year with a lag of 2 to 3 years.

five-year sales growth rates, positioning cloud use within the period rather than at its extremes. Lastly, we calculate the growth rate from $t+1$ to $t-10$, measuring cloud usage in t. Table A2 shows that both the OLS and ET estimations exhibit substantial consistency with the baseline results.

The 2021 wave of the ICT Survey. As a second robustness check, we incorporate the 2021 wave of the ICT survey, which provides information on cloud adoption in 2020 (Table A3). Despite the significant disruptions caused by the global COVID-19 pandemic in early 2020, the estimation results remain consistent with the baseline findings. Specifically, cloud adoption continues to positively influence firm growth rates, with smaller firms benefiting disproportionately.

Long-differences equations. Third, we adopt a long-differences approach, which is frequently used in studies examining ICT impacts (e.g., Acemoglu and Restrepo, 2020; Babina et al., 2024). We report results in Table A4. This method enables us to estimate the effects of cloud adoption over its entire diffusion period, assumed to have begun in 2009 following Bloom and Pierri (2018). However, it implies the use of one ICT survey at a time. We implement the following cross-sectional regression:

$$\begin{aligned} &\text{Sales Growth}_{i,t,2009} = \\ &a + \beta_1 \Delta\text{Cloud}_i + \beta_2 \Delta\text{Cloud}_{i,t} \cdot \text{Log-Sales}_{i,2009} + \beta_3 \text{Log-Sales}_{i,2009} + \beta_X \text{X}_{i,2009} + \epsilon_i \end{aligned} \tag{3}$$

Here, ΔCloud_i is a dummy variable for cloud use in year t, indicating whether cloud technology was adopted between 2009 and t (2015, 2017, 2019 or 2020). Since cloud adoption was unlikely in 2009, ΔCloud_i effectively captures the transition to cloud use by firms (similarly to the identification strategy in Forman and McElheran, 2025). Table A4 shows that the results for all periods are consistent with the main findings discussed in Section 4.1.

Two-Stage Least Squares. Fourth, we perform an additional robustness check using a non-linear TSLS regression, with results presented in Table A5 in the Appendix. This model incorporates the interaction between cloud and sales as an additional endogenous variable, instrumented using the interaction between sales and lightning strike density, following Wooldridge (2015). The TSLS results corroborate the main analysis, producing coefficients slightly farther from zero than those from the endogenous treatment (ET) model, but characterised by much higher standard errors. This is in line with the discussion in Wooldridge (2015), showing that TSLS models are generally more consistent but less efficient than control function approaches, and that the latter should be preferred in presence of explanatory variables that are interacted with the instrumented endogenous variable.

Coarsened Exact Matching. Finally, we employ an alternative strategy to examine the association between cloud adoption and firm growth rates. Specifically, we implement a Coarsened Exact Matching (CEM) procedure using a subset of the covariates included as controls. Table A6 reports the average characteristics of firms using and not using cloud technologies, matched through CEM under different binning settings, and a t-test indicating whether differences between these averages are significant. The OLS estimation results of Equation 1 for these matched samples are presented in Table A7. Model 1 builds the matched sample based on the following binning criteria. Firm size is divided into bins based on 15 evenly spaced cut points in the logarithmic scale, while physical and intangible capital are each divided by 10 cut points, ensuring that the percentage distance between values remains constant across bins. Firm age is categorised as follows: start-ups (less than 5 years), young (5–10 years), mature (11–20 years), and old (more than 20 years). Firms are also grouped based on the presence of ICT (zero vs. positive) and R&D workers (zero vs. positive), as well as by year, the median wages of managers and engineers, and industry classes used in the ICT survey sampling structure.[18] Model 2 further includes binning based on exporting activity and whether a firm operates multiple establishments. Model 3 adds a bin distinguishing firms located within or outside Île-de-France. Models 4 and 5 modify the binning of firm size, reducing the number of cut points to 10 and increasing them to 20, respectively.

Despite a significant reduction in sample size, the results confirm a positive association between cloud adoption and firm growth rates, with the effect diminishing as firm size increases. The matched sample is well balanced across the considered variables, with only a few significant differences (see Table A6).

4.3 Mechanisms

The use of cloud has a positive effect on the growth rate of firms, diminishing for larger initial size. In this section we analyse possible mechanisms driving the findings discussed above. In particular, we provide evidence in support of the view that cloud affects firm growth rates via changes in firms' organisational processes, and that the larger effect for smaller firms is not driven by an age-size correlation, or by decreasing returns to scale. To support these hypotheses, we first disaggregate cloud into different categories (cloud for storage, software, or computing power) and we investigate how different cloud uses – each of which is linked to different operations and processes in firms – drive the relationship with

[18] Industries are classified based on the 3-digit NAF system: 100–129, 130–159, 160–189, 190–239, 240–259, 265–267, 261–264 + 268, 270–289, 290–309, 310–339, 350–399, 410–439, 450–459, 460–469 (excluding 465), 465, 47, 49–53, 55, 56, 582, 58–61 (excluding 582), 61, 62 and 631, 639, 68, 69–74, 77–78 combined with 80–82, 79, 951.

firm growth rates. Second, in order to rule out the hypothesis that the observed relationship is confounded by differential patterns between younger and older firms, we assess whether the size-cloud interaction coefficient is driven by age. Finally, we examine whether growth differences between small and large firms adopting cloud technologies stem from decreasing returns to scale – making the effect of cloud on growth a mechanical outcome – rather than the specific characteristics of the technology, and of the activities that it enables. Specifically, we estimate the firm-level relationship between growth and the adoption of other digital technologies, namely Big Data Analytics (BDA) tools and E-commerce activities. If firms adopting cloud, BDA, and E-commerce exhibit similar growth patterns, this would suggest that the cloud-growth relationship is not driven by the reorganisation of firms' operations but rather by the presence of decreasing returns to scale in the use of digital technologies.

Heterogeneity across cloud types. In this section we explore the relationship between different types of cloud technologies and the growth rates of firms. Each cloud category considered enables different complementary ICTs and operations in the firm, and is therefore likely to affect the growth rates of firms' sales differently. We test this conjecture by estimating Equation 1 and including three non-exclusive types of cloud uses as covariates (cloud for storage, software, and computing power). To recall, the first type includes data storage services; the second groups together various software services; and the third refers to services through which the firm can acquire computing power on demand.

Results of the OLS estimation for the three different classes of cloud services are shown in Table 6. Overall, the findings discussed in Section 4.1 are confirmed by the results in Table 6. However, the magnitude and significance of the cloud type variables' coefficients differ. The coefficients on cloud services for storage and software are positive and significant, unlike the one on cloud for computing power. Furthermore, when including the interaction terms, the only negative and significant coefficient relates to the use of cloud services for software applications, indicating a major role for cloud in enabling growth through the adoption of otherwise costly and demanding software tools, which can now be stored and managed remotely.

This finding suggests that the impact of cloud technologies on firm growth rates may depend on the specific functionalities enabled by different cloud services. The positive and significant coefficients on cloud for storage and software services indicate that the effect on growth may be mediated by the possibility to expand their internal IT architecture, business processes, and organisational structure in the cloud, without undertaking costly physical IT investments. Such finding is consistent with the scale without mass hypothesis (Brynjolfsson et al., 2008), suggesting that ICTs help firms scale their operations more easily.

In contrast, the lack of a significant coefficient on cloud for computing power (and on

Table 6: Mechanism analysis: cloud types.

	Model 1	Model 2
Cloud – Storage	0.0758***	0.1545**
	(0.0126)	(0.0606)
Cloud – Storage*Log Sales		-0.0084
		(0.0053)
Cloud – Computing Power	-0.0024	0.0655
	(0.0110)	(0.0600)
Cloud – Computing Power*Log Sales		-0.0055
		(0.0051)
Cloud – Software	0.0291*	0.1889***
	(0.0156)	(0.0576)
Cloud – Software*Log Sales		-0.0157***
		(0.0048)
Log Sales	-0.0334***	-0.0251***
	(0.0036)	(0.0041)
Log Physical Capital	0.0192***	0.0195***
	(0.0044)	(0.0044)
Log Intangible Capital	-0.0193***	-0.0183***
	(0.0048)	(0.0045)
Log Age	-0.0900***	-0.0903***
	(0.0089)	(0.0089)
ICT Share	0.1290***	0.1183**
	(0.0462)	(0.0465)
R&D Share	0.1894***	0.1909***
	(0.0447)	(0.0446)
Exporter	0.0195*	0.0182*
	(0.0098)	(0.0098)
Multi Establishment	-0.0020	-0.0023
	(0.0067)	(0.0068)
Log Average Hourly Wage Manag. & Ing.	-0.0007	-0.0032
	(0.0024)	(0.0025)
Observations	21,886	21,886
Adjusted R^2	0.116	0.119
Industry FE	Yes	Yes
Year FE	Yes	Yes
Region FE	Yes	Yes

Note: The table reports the results of the estimation of Equation 1 when including multiple types of cloud technologies. The dependent variable measures the sales rate of growth of the firm between t and $t-5$. Cloud technologies are captured by three different dummy variables, identifying the adoption of cloud services for data storage, software, or computing power, adopted by time t. All the time-varying control variables are measured at $t-5$. Log Sales, Physical, and Intangible Capital are expressed in Log of Euros. Age is the firm age in years since the establishment. ICT and R&D shares express, respectively, the share of hours worked in ICT and R&D occupations. Exporter is a dummy that identifies whether the firms is engaged in exporting activities. Multi Establishment identify whether the firm has more than one economic establishment. Log Average Hourly Wage Manag. & Ing. measures the log of the hourly wage of management and engineers, expressed in Euros. All regressions are estimated using OLS. Standard errors are clustered at the NACE 2-digit level. *** p<0.01, ** p<0.05, * p<0.1.

its interaction with size) suggests that firms may not immediately translate access to scalable computing resources into sales growth. This result may be influenced by the period covered in our empirical analysis, which ends in 2019 and coincides with the early diffusion of advanced digital technologies, such as AI, whose performance is linked to the availability of high computing power. However, these ICT may not generate immediate effects on firm growth due to the need of complementary investments and implementation lags (e.g. AI, Brynjolfsson et al., 2018, 2021), explaining why the relation between growth and cloud is not driven by services to acquire computing power.

Finally, unlike cloud storage, which primarily optimises internal data handling, the negative interaction between cloud for software and size indicate that its relationship with growth loses strength as the initial size of firms grows. Indeed, large firms may either already have an established physical IT infrastructure or may afford to build one in-house without purchasing external cloud services, reducing their dependency on externally provided cloud-based software. Conversely, small firms may face high barriers to accessing such software solutions due to cost constraints or gaps in ICT capabilities, which may limit the scalability of their operations. The results are consistent with the view that the adoption of cloud allows smaller firms to integrate software into their organisations while reducing costs.

Cloud, growth, and age. Our results indicate that the use of cloud technologies has a positive impact on firm's sales growth, particularly for smaller firms. However, it is well-established in the empirical literature that firm size and age are positively correlated. Recent studies on the impact of cloud on firm growth (DeStefano et al., 2023b) suggest that younger firms benefit more from cloud technology. Therefore, the negative interaction between cloud and sales may be moderated by the age of firms.

This distinction is crucial: depending on whether the effect is mediated by firm size or age, policies will target specific groups of firms and aggregate dynamics. Additionally, the distinction between the moderating effect of firm size and age on firm's growth helps clarify whether the nexus between growth and cloud is driven by factors related to size (easier access to digitalisation and the overcoming of fixed costs), or to age (for example, the lack of new ICT-related managerial capabilities).

We estimate Equations 1 and 2 with the inclusion of an interaction term between age and cloud, and report the results in Table 7. The coefficient on the cloud-age interaction is not statistically significant. Importantly, the interaction term between cloud and sales remains negative and significant, though it slightly decreases in magnitude when the product of cloud and age is added. We conclude that the mediating effect of firms' size in the relationship between cloud adoption and sales growth is correctly identified, and not confounded by firms' age.

Table 7: Mechanism analysis: the mediation role of firms' age.

	OLS	End. Treatment	
	Model 1	Model 2	
		Growth Rate	Selection
Cloud	0.2806***	0.3149***	
	(0.0954)	(0.0811)	
Cloud*Log Sales	-0.0158***	-0.0164***	
	(0.0048)	(0.0043)	
Cloud*Log Age	-0.0109	-0.0109	
	(0.0167)	(0.0166)	
Log Sales	-0.0271***	-0.0285***	0.1805***
	(0.0039)	(0.0063)	(0.0217)
Log Physical Capital	0.0192***	0.0189***	0.0333**
	(0.0044)	(0.0045)	(0.0131)
Log Intangible Capital	-0.0183***	-0.0182***	-0.0088
	(0.0046)	(0.0045)	(0.0186)
Log Age	-0.0874***	-0.0866***	-0.0896***
	(0.0078)	(0.0076)	(0.0172)
ICT Share	0.1178**	0.1123**	0.5289***
	(0.0455)	(0.0474)	(0.1697)
R&D Share	0.1949***	0.1937***	0.1032
	(0.0447)	(0.0447)	(0.1400)
Exporter	-0.0028	-0.0030	0.0349***
	(0.0025)	(0.0025)	(0.0105)
Multi Establishment	0.0183*	0.0166	0.2059***
	(0.0097)	(0.0112)	(0.0798)
Log Average Hourly Wage Manag. & Ing.	-0.0020	-0.0033	0.1377***
	(0.0070)	(0.0071)	(0.0222)
Log Lightning Density			-0.0573***
			(0.0112)
Observations	21,886	21,886	21,886
Adjusted R^2	0.117		
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes

Note: Estimation results of Equations 1 (Model 1) and 2 (Model 2), both including a cloud-age interaction. Model 1 is estimated using OLS, Model 2 uses a FIML estimator. In Model 2, the first column reports the results of the outcome equation, while the second column refers to the selection equation. The dependent variable measures the sales rate of growth of the firm between t and $t-5$. Cloud is a dummy variable identifying the adoption of cloud technologies by time t. All the time-varying control variables are measured at $t-5$. Log Sales, Physical, and Intangible Capital are expressed in Log of Euros. Age is the firm age in years since the establishment. ICT and R&D shares express, respectively, the share of hours worked in ICT and R&D occupations. Exporter is a dummy that identifies whether the firms is engaged in exporting activities. Multi Establishment identify whether the firm has more than one economic establishment. Log Average Hourly Wage Manag. & Ing. measures the log of the hourly wage of management and engineers, expressed in Euros. Log lightning density is the logarithm of the density of daily lightning strikes at the municipality level, averaged between 2008 and 2017. All regressions are estimated using a Full Information Maximum Likelihood (FIML) estimator. Standard errors are clustered at the NACE 2-digit level. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Mechanism analysis: decreasing returns to scale.

	Model 1	Model 2	Model 3		Model 4	
			Growth Rate	Selection	Growth Rate	Selection
Cloud	0.0841***	0.2559***	0.0556*		0.2882***	
	(0.0167)	(0.0614)	(0.0311)		(0.0654)	
Cloud*Log Sales		-0.0175***			-0.0180***	
		(0.0050)			(0.0047)	
E-commerce	0.0469**	0.0823	0.0499**	0.3401***	0.0803	0.3400***
	(0.0198)	(0.0924)	(0.0204)	(0.0523)	(0.0934)	(0.0523)
E-commerce*Log Sales		-0.0034			-0.0035	
		(0.0078)			(0.0078)	
Big Data Analysis	0.0425***	0.0798**	0.0459***	0.3594***	0.0772**	0.3597***
	(0.0079)	(0.0339)	(0.0091)	(0.0402)	(0.0357)	(0.0404)
Big Data Analysis*Log Sales		-0.0035			-0.0036	
		(0.0031)			(0.0031)	
Log Sales	-0.0357***	-0.0272***	-0.0342***	0.1621***	-0.0284***	0.1628***
	(0.0040)	(0.0051)	(0.0044)	(0.0203)	(0.0064)	(0.0207)
Log Physical Capital	0.0184***	0.0186***	0.0186***	0.0284**	0.0184***	0.0284**
	(0.0043)	(0.0044)	(0.0044)	(0.0133)	(0.0044)	(0.0133)
Log Intangible Capital	-0.0193***	-0.0182***	-0.0193***	-0.0101	-0.0181***	-0.0103
	(0.0047)	(0.0044)	(0.0047)	(0.0167)	(0.0044)	(0.0169)
Log Age	-0.0898***	-0.0903***	-0.0906***	-0.0863***	-0.0896***	-0.0858***
	(0.0088)	(0.0089)	(0.0088)	(0.0151)	(0.0088)	(0.0155)
ICT Share	0.1269**	0.1193**	0.1322***	0.5348***	0.1141**	0.5339***
	(0.0480)	(0.0475)	(0.0480)	(0.1401)	(0.0493)	(0.1409)
R&D Share	0.2012***	0.2045***	0.2032***	0.1594	0.2027***	0.1594
	(0.0453)	(0.0445)	(0.0453)	(0.1359)	(0.0444)	(0.1364)
Exporter	0.0175**	0.0158*	0.0191**	0.1874***	0.0143*	0.1867***
	(0.0085)	(0.0080)	(0.0094)	(0.0643)	(0.0085)	(0.0644)
Multi Establishment	-0.0036	-0.0039	-0.0025	0.1253***	-0.0050	0.1250***
	(0.0070)	(0.0071)	(0.0068)	(0.0226)	(0.0068)	(0.0228)
Log Average Hourly Wage Manag. & Ing.	-0.0006	-0.0028	-0.0004	0.0379***	-0.0031	0.0375***
	(0.0024)	(0.0026)	(0.0024)	(0.0099)	(0.0026)	(0.0099)
Log Lightning Density				-0.0525***		-0.0524***
				(0.0104)		(0.0106)
Observations	21,886	21,886	21,886	21,886	21,886	21,886
Adjusted R^2	0.119	0.121				
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Estimation results of Equations 1 (Models 1-2) and 2 (Models 3-4), both including dummies measuring the adoption of E-commerce and Big Data Analysis, and their interaction with firms' size. Models 1-2 are estimated using OLS, Models 3-4 use a FIML estimator. In Models 3-4, the first column reports the results of the outcome equation, while the second column refers to the selection equation. The dependent variable measures the sales rate of growth of the firm between t and $t-5$. Cloud is a dummy variable identifying the adoption of cloud technologies by time t. All the time-varying control variables are measured at $t-5$. Log Sales, Physical, and Intangible Capital are expressed in Log of Euros. Age is the firm age in years since the establishment. ICT and R&D shares express, respectively, the share of hours worked in ICT and R&D occupations. Exporter is a dummy that identifies whether the firms is engaged in exporting activities. Multi Establishment identify whether the firm has more than one economic establishment. Log Average Hourly Wage Manag. & Ing. measures the log of the hourly wage of management and engineers, expressed in Euros. Log lightning density is the logarithms of the (population weighted) density of daily lightning strikes at the municipality level, averaged between 2008 and 2017. All regressions are estimated using a Full Information Maximum Likelihood (FIML) estimator. Standard errors are clustered at the NACE 2-digit level. *** p<0.01, ** p<0.05, * p<0.1.

Decreasing returns to scale. The smaller effects of cloud use on firm performance for larger firms may be driven by decreasing returns to scale in the use of digital technologies. For instance, larger firms may operate in a less elastic part of the demand curve, limiting their ability to translate digital adoption into sales growth. However, the existence of such dynamics is not supported by empirical studies on the effects of ICT and AI investments on firms' size growth (Brynjolfsson et al., 2023; Babina et al., 2024). Furthermore, the view that smaller firms grow faster than larger firms, conditional on survival, has been challenged by various studies (see discussion in Coad et al., 2014), particularly when controlling for age (Haltiwanger et al., 2013). However, should a decreasing returns to scale dynamic hold in our sample, it would mitigate the role of cloud in allowing larger sales growth for smaller firms.

To test this hypothesis, we extend Equations 1 and 2 by incorporating two additional digital technologies: E-commerce and Big Data Analytics (BDA). E-commerce may improve firm performance by enabling access to larger markets, allowing firms to reach a broader customer base (Couture et al., 2021) and reducing search costs (Goldmanis et al., 2010). BDA, on the other hand, has been found to enhance productivity by enabling data-driven decision-making and fostering process innovation (see for instance, Andres et al., 2024; Conti et al., 2024).

We report the results in Table 8. Model 1 (OLS) and Model 3 (ET) show that both E-commerce and BDA are positively and significantly associated with long-run firm growth rates, suggesting that improvements in market access gained trough E-commerce activities and increases in productivity due to the use of BDA may contribute to better firm performance. However, when we include interaction terms with firm size (Models 2 and 4), the coefficients are not significant. In contrast, the coefficient for cloud adoption remains positive and significant when interacted with firm size.

This evidence indicates that, despite being positive and significant and differently from cloud, the effects of BDA and E-commerce on sales growth do not vary systematically with firm size. Such results are at odds with the existence of decreasing returns to scale in the use of digital technologies. It contrasts also with the claim that the growth rates of larger firms, facilitated by digital technologies, is limited by its position in less elastic parts of the demand curve. Furthermore, this evidence supports that the heterogeneous effect of cloud on firms' growth depends on the specific characteristics of this technology.

5 Cloud and industry concentration

In Section 4, we demonstrated that the adoption of cloud services positively impacts firms' long-term sales growth rate, with smaller firms growing more by leveraging cloud technologies to narrow their digital gap. This finding talks directly to a growing literature examining the link between ICT and industry concentration (Bajgar et al., 2025; Brynjolfsson et al., 2023; Babina et al., 2024; De Ridder, 2024). It is therefore legitimate to hypothesise that the widespread adoption of cloud technologies could counteract the increasing concentration trends observed across high-income markets (Bajgar et al., 2023).

To test this conjecture, we aggregate our firm-level data by 2-digit industrial sectors and estimate the following Equation (Bessen, 2020; Brynjolfsson et al., 2023):

$$\begin{aligned} &\text{Log Concentration}_{s,t} = \\ &\alpha + \beta_1 \text{Log Cloud Share}_{s,t} + \beta_x \text{Controls}_{s,t} + \text{2-digit Ind.}_s + \text{Year}_t \end{aligned} \tag{4}$$

Where s identifies 2-digit sectors at time t (either 2015, 2017, or 2019), $\text{Concentration}_{s,t}$ is the Herfindahl-Hirschman Index (HHI) of 2-digit industries at each time period, $\text{Cloud Share}_{s,t}$ is the sectoral share of firms using cloud at time t, $\text{Controls}_{i,t}$ is a vector including the the stock ICT and R&D Full Time Equivalent (FTE) occupations, the aggregate sales and employment of active firms, aggregate physical and intangible capital stocks, and the number of firms in sector s in logarithmic scale. We include also industry (2-digit Ind._s) and year (Year_t) fixed effects.[19]

Table 9 presents the estimation results of Equation 4. The baseline model (Model 1) reveals a negative relationship between industry concentration and the share of firms adopting cloud services. Consistent with findings from other studies (Bajgar et al., 2025; Babina et al., 2024; Brynjolfsson et al., 2023), the sectoral ICT intensity – in our case measured by the stock of ICT employees – is positively associated with concentration. In contrast, the stock of R&D personnel is negatively associated, suggesting that R&D intensive sectors tend to be characterised by lower concentration. As expected, the number of firms in a sector is negatively related to concentration.

Models 2, 3, and 4 provide the estimation results of Equation 4, distinguishing between different types of cloud services. These results indicate that the negative relationship estimated in Model 1 is primarily driven by the two most mature categories of cloud technologies: data and file storage, and, to a lesser extent, software applications. Conversely, the most re-

[19] We observe industries at the 2-digit level as the low number of firms in the samples provided by the ICT surveys does not allow us to measure the share of firms purchasing cloud at lower levels of aggregation. Our choice of aggregation allows to keep the sample representative.

Table 9: Different types of cloud technologies and industry concentration.

	Model 1	Model 2	Model 3	Model 4
Log Cloud Share	-0.0480**			
	(0.0237)			
Log Cloud Share (Storage)		-0.0501*		
		(0.0251)		
Log Cloud Share (Software)			-0.0464*	
			(0.0251)	
Log Cloud Share (Comp. Power)				-0.0175
				(0.0164)
Log ICT Employees	0.2575***	0.2536***	0.2719***	0.2730***
	(0.0788)	(0.0791)	(0.0796)	(0.0846)
Log R&D Employees	-0.3123***	-0.3106***	-0.3238***	-0.3135***
	(0.0794)	(0.0794)	(0.0786)	(0.0804)
Log Total Revenues	1.7249***	1.7398***	1.7116***	1.6993***
	(0.2243)	(0.2279)	(0.2305)	(0.2501)
Log Total Employment	-0.3907	-0.4188	-0.3938	-0.4240
	(0.4563)	(0.4579)	(0.4711)	(0.5091)
Log Number of Firms	-0.5834***	-0.5722***	-0.6073***	-0.5662***
	(0.1950)	(0.1955)	(0.1895)	(0.1951)
Log Intangible Capital	-0.2204	-0.2204	-0.2254	-0.2542
	(0.1427)	(0.1434)	(0.1454)	(0.1622)
Log Physical Capital	0.1545	0.1548	0.1579	0.2011
	(0.2163)	(0.2155)	(0.2191)	(0.2350)
Observations	174	174	174	174
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.985	0.985	0.985	0.985

Note: The table reports the results of the estimation of Equation 4. The dependent variable measures the Herfindahl-Hirschman Index (HHI) of 2-digit NACE industries, obtained by aggregating the sales of firms by industry. The Log Cloud Share variables measure the logarithm of the share of firms in each industry that have adopted respectively any, storage, software, or computing power cloud technologies by time t. Controls are measured at time t and include: the log of the number of employees in ICT and R&D occupations; the log of total revenues of the industry (in Euros); the log number of total employees and number of firms; the log of the total stock of tangible and intangible capital (in Euros). All regressions are estimated using OLS. Standard errors are clustered at the NACE 2-digit level. *** p<0.01, ** p<0.05, * p<0.1.

cent cloud functionality, associated with applications requiring strong computing power, is negatively but not significantly related to concentration.

Overall, the observed negative relationship between the share of cloud users in an industry and its concentration aligns with the idea that cloud services bolster higher growth for smaller firms than larger ones, and that this effect on individual firms is associated to a (mild) reduction in industrial concentration. Given the effect size of cloud adoption on industry concentration, it is unlikely that these technologies alone will systematically allow smaller firms to overturn concentration trends across industries, and to displace competing larger firms. This indicates that changes in industry concentration are more likely to manifest at the margin, with some catching up of smaller firms towards larger ones. These results also suggest that the impact of cloud adoption on concentration may take time to realise. In particular, the diffusion of newer digital technologies such as AI and Big Data Analytics, which heavily depend on computing power accessed via cloud services, remains limited with respect to former types of cloud services. These technologies, being less widely adopted at the time of writing, may however have a significant mitigating effect on concentration as their adoption increases.

6 Conclusions

In accordance with the extant literature on the impact of cloud adoption on firms' growth, our paper has shown that cloud technologies have a heterogeneous effect on the sales growth rates of firms. Notably, our results reveal that smaller firms experience greater benefits in terms of growth compared to their larger counterparts, when adopting cloud services. This finding underscores the importance of promoting their diffusion and adoption, particularly when considering the differential impacts on small relative to large firms. To rationalise our findings, we have shown that the disproportionate effect of cloud use on the growth rates of small firms passes through an internal reorganisation of firms' processes. This hypothesis is backed up by the evidence that cloud technologies that allow firms to use administrative, office, or CRM software applications are more positively associated to sales growth in small firms, unlike cloud services used for data storage or computing power. To prop up our findings, we have ruled out that the mediating effect of size is confounded by firms' age, and that decreasing returns could account for the smaller benefits on the growth rates of larger firms experienced from the use of cloud technologies.

Investigating the firm-level relation between cloud and performance as mediated by size holds key policy relevance in the current economic context, characterised by increasing trends in industry concentration (see, among others, Gutiérrez and Philippon, 2017; Bajgar et al.,

2023; Grullon et al., 2019; Autor et al., 2020; De Ridder, 2024). Previous evidence has suggested that digitalisation may positively affect market concentration by favouring the growth rates of larger firms, and that intangibles assets and ICT adoption are associated to higher concentration within industries (Bessen, 2020; Bajgar et al., 2025; Brynjolfsson et al., 2023; Babina et al., 2024; Lashkari et al., 2024). The relationship between ICT adoption and performance, as mediated by firm size, can be pointed out as one of the causes of the increase in concentration observed in several countries. This is in line with the idea that firms do not have equal access to new technologies (Terranova and Turco, 2022). However, the use of cloud by firms may be the exception to the rule. Contrary to the evidence on other ICTs, and as suggested by recent contributions Lu et al. (2024) – including our own – the diffusion of cloud technologies is not associated to increasing market concentration; conversely, we offer evidence on a mild reduction of concentration within industries associated to the diffusion of cloud technologies.

Bringing together the firm- and industry-level findings, we argue that cloud technologies can enhance the digitalisation and growth rates of smaller firms more than larger ones, lending support to a more inclusive digital transition. Future research will be aimed at investigating the broader implications of the Industry 4.0 transition beyond cloud services. Emerging digital tools such as AI and 3D printing have the potential to reshape firm performance, particularly in terms of size, productivity, and survival. Such exploration would contribute to the understanding of the firm-level determinants of existing trends in productivity divergences and business dynamism, that have both been linked to digital intensity at the sector level (Calvino and Criscuolo, 2019; Calvino et al., 2020; Corrado et al., 2021). Moreover, these technologies may yield different impacts depending on firm characteristics, as observed with cloud services, potentially accelerating or mitigating the aforementioned trends. Understanding how these new digital technologies interact with firm-specific factors is crucial, as technological adoption increasingly becomes a key determinant of market competitiveness in the current digital transition.

References

Acemoglu, Daron, Gary W Anderson, David N Beede et al. 2022. "Automation and the Workforce: A Firm-Level View from the 2019 Annual Business Survey." Working Paper 30659, National Bureau of Economic Research. 10.3386/w30659.

Acemoglu, Daron, and Pascual Restrepo. 2020. "Robots and Jobs: Evidence from US Labor Markets." *Journal of Political Economy* 128 (6): 2188 – 2244. 10.1086/705716.

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li. 2023. "A Theory of Falling Growth and Rising Rents." *The Review of Economic Studies* 90 (6): 2675–2702. 10.1093/restud/rdad016.

Andersen, Thomas Barnebeck, Jeanet Bentzen, Carl-Johan Dalgaard, and Pablo Selaya. 2012. "Lightning, IT Diffusion, and Economic Growth Across U.S. States." *Review of Economics and Statistics* 94 (4): 903–924. 10.1162/rest_a_00316.

Andres, Raphaela, Thomas Niebel, and Robin Sack. 2024. "Big data and firm-level productivity: A cross-country comparison." ZEW Discussion Papers 24-053, ZEW - Leibniz-Zentrum für Europäische Wirtschaftsforschung, Mannheim, https://hdl.handle.net/10419/300678.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen. 2020. "The Fall of the Labor Share and the Rise of Superstar Firms*." *The Quarterly Journal of Economics* 135 (2): 645–709. 10.1093/qje/qjaa004.

Babina, Tania, Anastassia Fedyk, Alex He, and James Hodson. 2024. "Artificial intelligence, firm growth, and product innovation." *Journal of Financial Economics* 151 103745. 10.1016/j.jfineco.2023.103745.

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo, and Jonathan Timmis. 2023. "Industry Concentration in Europe and North America." *Industrial and Corporate Change* dtac059. 10.1093/icc/dtac059.

Bajgar, Matej, Chiara Criscuolo, and Jonathan Timmis. 2025. "Intangibles and Industry Concentration: A Cross-Country Analysis." *Oxford Bulletin of Economics and Statistics*. 10.1111/obes.12659.

Bessen, James. 2020. "Industry Concentration and Information Technology." *The Journal of Law and Economics* 63 (3): 531–555. 10.1086/708936.

Bloom, Nicholas, and Nicola Pierri. 2018. "Cloud Computing Is Helping Smaller, Newer firms Compete." *Harvard Business Review* 51 (7): .

Bontadini, F, C Corrado, J Haskel, M Iommi, and C Jona-Lasinio. 2023. "EUKLEMS & INTANProd: Industry productivity accounts with intangibles." *Sources of growth and productivity trends: methods and main measurement challenges, Luiss Lab of European Economics, Rome.*

Brynjolfsson, Erik, and Lorin M. Hitt. 2003. "Computing Productivity: Firm-Level Evidence." *The Review of Economics and Statistics* 85 (4): 793–808. 10.1162/003465303772815736.

Brynjolfsson, Erik, Wang Jin, and Xiupeng Wang. 2023. "Information Technology, Firm Size, and Industrial Concentration." NBER Working Papers 31065, National Bureau of Economic Research.

Brynjolfsson, Erik, Andrew McAfee, Michael Sorell, and Feng Zhu. 2008. "Scale without mass: business process replication and industry dynamics." Mimeo.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2018. "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics." In *The Economics of Artificial Intelligence: An Agenda*, 23–57, National Bureau of Economics Research.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2021. "The Productivity J-Curve: How Intangibles Complement General Purpose Technologies." *American Economic Journal: Macroeconomics* 13 (1): 333–372. 10.1257/mac.20180386.

Byrne, David, Carol Corrado, and Daniel E Sichel. 2018. "The Rise of Cloud Computing: Minding Your P's, Q's and K's." Working Paper 25188, National Bureau of Economic Research. 10.3386/w25188.

Caldarola, Bernardo, Marco Grazzi, Martina Occelli, and Marco Sanfilippo. 2023. "Mobile internet, skills and structural transformation in Rwanda." *Research Policy* 52 (10): 104871. 10.1016/j.respol.2023.104871.

Calvino, Flavio, and Chiara Criscuolo. 2019. "Business dynamics and digitalisation." OECD Science, Technology and Industry Policy Papers 62, OECD Publishing, https://ideas.repec.org/p/oec/stiaac/62-en.html.

Calvino, Flavio, Chiara Criscuolo, and Antonio Ughi. 2024. "Digital adoption during COVID-19: Cross-country evidence from microdata." OECD Science, Technology and Industry Working Papers 2024/03, OECD Publishing. 10.1787/f63ca261-en.

Calvino, Flavio, Chiara Criscuolo, and Rudy Verlhac. 2020. "Declining business dynamism: Structural and policy determinants." OECD Science, Technology and Industry Policy Papers 94, OECD Publishing, https://ideas.repec.org/p/oec/stiaac/94-en.html.

Calvino, Flavio, and Luca Fontanelli. 2023a. "Artificial intelligence, complementary assets and productivity: evidence from French firms." LEM Papers Series 2023/35, Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy.

Calvino, Flavio, and Luca Fontanelli. 2023b. "A portrait of AI adopters across countries: Firm characteristics, assets' complementarities and productivity." OECD Science, Technology and Industry Working Papers No. 2023/02, OECD Publishing, Paris.

Campa, Jose Manuel, and Simi Kedia. 2002. "Explaining the Diversification Discount." *The Journal of Finance* 57 (4): 1731–1762. 10.1111/1540-6261.00476.

Chiplunkar, Gaurav, and Pinelopi Koujianou Goldberg. 2022. "The Employment Effects of Mobile Internet in Developing Countries." Working Paper 30741, National Bureau of Economic Research. 10.3386/w30741.

Cirillo, Valeria, Lucrezia Fanti, Andrea Mina, and Andrea Ricci. 2023. "The adoption of digital technologies: Investment, skills, work organisation." *Structural Change and Economic Dynamics* 66 89–105. 10.1016/j.strueco.2023.04.011.

Clougherty, Joseph A., Tomaso Duso, and Johannes Muck. 2016. "Correcting for Self-selection Based Endogeneity in Management Research: Review, Recommendations and Simulations." *Organizational Research Methods* 19 (2): 286–347. 10.1177/1094428115619013.

Coad, Alex, Anders Bornhäll, Sven-Olov Daunfeldt, and Alexander McKelvie. 2024. *Scale-ups and High-Growth Firms: Theory, Definitions, and Measurement*. Springer Nature.

Coad, Alex, Sven-Olov Daunfeldt, Werner Hölzl, Dan Johansson, and Paul Nightingale. 2014. "High-growth firms: introduction to the special section." *Industrial and Corporate Change* 23 (1): 91–112. 10.1093/icc/dtt052.

Conti, Raffaele, Miguel Godinho de Matos, and Giovanni Valentini. 2024. "Big Data Analytics, Firm Size, and Performance." *Strategy Science* 9 (2): 135–151. 10.1287/stsc.2022.0007.

Corrado, Carol, Jonathan Haskel, and Cecilia Jona-Lasinio. 2021. "Artificial intelligence and productivity: an intangible assets approach." *Oxford Review of Economic Policy* 37 (3): 435–458. 10.1093/oxrep/grab018.

Couture, Victor, Benjamin Faber, Yizhen Gu, and Lizhi Liu. 2021. "Connecting the Countryside via E-Commerce: Evidence from China." *American Economic Review: Insights* 3 (1): 35–50. 10.1257/aeri.20190382.

Coyle, Diane, and David Nguyen. 2018. "Cloud Computing and National Accounting." Economic Statistics Centre of Excellence (ESCoE) Discussion Papers ESCoE DP-2018-19, Economic Statistics Centre of Excellence (ESCoE).

Crémer, Jacques, Gary Biglaiser, and Andrea Mantovani. 2024. "The Economics of the Cloud." TSE Working Papers 24-1520, Toulouse School of Economics (TSE), https://ideas.repec.org/p/tse/wpaper/129265.html.

De Ridder, Maarten. 2024. "Market Power and Innovation in the Intangible Economy." *American Economic Review* 114 (1): 199–251. 10.1257/aer.20201079.

DeStefano, Timothy, Richard Kneller, and Jonathan Timmis. 2023a. "Cloud Computing and Firm Growth." *The Review of Economics and Statistics* 1–47. 10.1162/rest_a_01393.

DeStefano, Timothy, Richard Kneller, and Jonathan Timmis. 2023b. "Cloud Computing and Firm Growth." *The Review of Economics and Statistics* 1–47. 10.1162/rest_a_01393.

Duso, Tomaso, and Alexander Schiersch. 2025. "Let's switch to the cloud: Cloud usage and its effect on labor productivity." *Information Economics and Policy* 70 101130. https://doi.org/10.1016/j.infoecopol.2025.101130.

Fontanelli, Luca. 2024. "Theories of market selection." *Journal of Economic Surveys* 1–35. 10.1111/joes.12625.

Forman, Chris, Avi Goldfarb, and Shane Greenstein. 2012. "The Internet and Local Wages: A Puzzle." *American Economic Review* 102 (1): 556–75. 10.1257/aer.102.1.556.

Forman, Chris, and Kristina McElheran. 2025. "Production Chain Organization in the Digital Age: Information Technology Use and Vertical Integration in U.S. Manufacturing." *Management Science* 71 1027–1049. 10.1287/mnsc.2019.01586.

Gal, Peter, Giuseppe Nicoletti, Christina von Rüden, Stéphane Sorbe, and Théodore Renault. 2019. "Digitalization and Productivity: In Search of the Holy Grail - Firm-level Empirical Evidence from European Countries." *International Productivity Monitor* 37 39–71.

Garrison, Gary, Robin L. Wakefield, and Sanghyun Kim. 2015. "The effects of IT capabilities and delivery model on cloud computing success and firm performance for cloud supported processes and operations." *International Journal of Information Management* 35 (4): 377–393. 10.1016/j.ijinfomgt.2015.03.001.

Gibrat, Robert. 1931. *Les Inégalités Économiques*. Paris: Librairie du Recueil Sirey.

Goldmanis, Maris, Ali Hortaçsu, Chad Syverson, and Onsel Emre. 2010. "E-Commerce and the Market Structure of Retail Industries." *The Economic Journal* 120 (545): 651–682. 10.1111/j.1468-0297.2009.02310.x.

Grullon, Gustavo, Yelena Larkin, and Roni Michaely. 2019. "Are US Industries Becoming More Concentrated?*." *Review of Finance* 23 (4): 697–743. 10.1093/rof/rfz007.

Guriev, Sergei, Nikita Melnikov, and Ekaterina Zhuravskaya. 2021. "3G Internet and Confidence in Government." *The Quarterly Journal of Economics* 136 (4): 2533–2613. 10.1093/qje/qjaa040.

Gutiérrez, Germán, and Thomas Philippon. 2017. "Declining Competition and Investment in the U.S." NBER Working Papers 23583, National Bureau of Economic Research, Inc, https://ideas.repec.org/p/nbr/nberwo/23583.html.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda. 2013. "Who Creates Jobs? Small versus Large versus Young." *The Review of Economics and Statistics* 95 (2): 347–361. 10.1162/REST_a_00288.

Hamilton, Barton H., and Jackson A. Nickerson. 2003. "Correcting for Endogeneity in Strategic Management Research." *Strategic Organization* 1 (1): 51–78. 10.1177/147612700300100121.

Harrigan, James, Ariell Reshef, and Farid Toubal. 2021. "The March of the Techies: Job Polarization Within and Between Firms." *Research Policy* 50 (7): . 10.1016/j.respol.2020.104008.

Heckman, James J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models." In *Annals of Economic and Social Measurement, Volume 5, number 4*, 475–492, National Bureau of Economic Research.

Heckman, James J. 1978. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica* 46 (4): 931–959. 10.2307/1909757.

Jin, Wang, and John Bai. 2022. "Cloud Adoption and Firm Performance: Evidence from Labor Demand."Technical Report Available at SSRN 4082436, SSRN.

Jin, Wang, and Kristina McElheran. 2024. "Economies Before Scale: IT Strategy and Performance Dynamics of Young US Businesses." *Management Science* Forthcoming. 10.1287/mnsc.2019.03116.

Kaplan, Jed O. 2023. "The World Wide Lightning Location Network (WWLLN) Global Lightning Climatology (WGLC) and time series." December. 10.5281/zenodo.10246857.

King, Andrew A., and Cristopher L. Tucci. 2002. "Incumbent Entry into New Market Niches: The Role of Experience and Managerial Choice in the Creation of Dynamic Capabilities." *Management Science* 48 (2): . 10.1287/mnsc.48.2.171.253.

Lashkari, Danial, Arthur Bauer, and Jocelyn Boussard. 2024. "Information Technology and Returns to Scale." *American Economic Review* 114 (6): 1769–1815. 10.1257/aer.20220522.

Lotti, Francesca, Enrico Santarelli, and Marco Vivarelli. 2003. "Does Gibrat's Law hold among young, small firms?." *Journal of Evolutionary Economics* 13 (3): 213–235. 10.1007/s00191-003-0153-0.

Lotti, Francesca, Enrico Santarelli, and Marco Vivarelli. 2009. "Defending Gibrat's Law as a long-run regularity." *Small Business Economics* 32 (1): 31–44. 10.1007/s11187-007-9071-0.

Lu, Yao, Gordon M Phillips, and Jia Yang. 2024. "The Impact of Cloud Computing and AI on Industry Dynamics and Concentration." Working Paper 32811, National Bureau of Economic Research. 10.3386/w32811.

Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press.

Manacorda, Marco, and Andrea Tesei. 2020. "Liberation Technology: Mobile Phones and Political Mobilization in Africa." *Econometrica* 88 (2): 533–567. 10.3982/ECTA14392.

McElheran, Kristina, J. Frank Li, Erik Brynjolfsson, Zachary Kroff, Emin Dinlersoz, Lucia S Foster, and Nikolas Zolas. 2023. "AI Adoption in America: Who, What, and Where." *Journal of Economics and Management Strategy* Forthcoming. 10.1111/jems.12576.

Nicoletti, Giuseppe, Christina von Rueden, and Dan Andrews. 2020. "Digital technology diffusion: A matter of capabilities, incentives or both?." *European Economic Review* 128 (C): . 10.1016/j.euroecorev.2020.

Ohnemus, Jörg, and Thomas Niebel. 2016. "On the Determinants of Cloud Computing Adoption." 27th European Regional ITS Conference, Cambridge (UK) 2016 148694, International Telecommunications Society (ITS).

Puhani, Patrick. 2000. "The Heckman Correction for Sample Selection and Its Critique." *Journal of Economic Surveys* 14 (1): 53–68. 10.1111/1467-6419.00104.

Santarelli, Enrico, Luuk Klomp, and A. Roy Thurik. 2006. "Gibrat's Law: An Overview of the Empirical Literature." In *Entrepreneurship, Growth, and Innovation: The Dynamics of Firms and Industries*, edited by Santarelli, Enrico 41–73, Boston, MA: Springer US, . 10.1007/0-387-32314-7_3.

Schulman, Aaron, and Neil Spring. 2011. "Pingin' in the Rain." In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, IMC '11 19–28, New York, NY, USA: Association for Computing Machinery, . 10.1145/2068816.2068819.

Shaver, J. Myles. 1998. "Accounting for Endogeneity When Assessing Strategy Performance: Does Entry Mode Choice Affect FDI Survival?" *Management Science* 44 (4): 571–585. 10.1287/mnsc.44.4.571.

Tambe, Prasanna, and Lorin M. Hitt. 2012. "The Productivity of Information Technology Investments: New Evidence from IT Labor Data." *Information Systems Research* 23 (3-part-1): 599–617. 10.1287/isre.1110.0398.

Terranova, Roberta, and Enrico M. Turco. 2022. "Concentration, stagnation and inequality: An agent-based approach." *Journal of Economic Behavior & Organization* 193 569–595. 10.1016/j.jebo.2021.11.002.

Vella, Francis, and Marno Verbeek. 1999. "Estimating and Interpreting Models with Endogenous Treatment Effects." *Journal of Business & Economic Statistics* 17 (4): 473–478.

Wooldridge, Jeffrey. 2015. "Control Function Methods in Applied Econometrics." *Journal of Human Resources* 50 (2): 420–445, https://EconPapers.repec.org/RePEc:uwp:jhriss:v:50:y:2015:i:2:p:420-445.

Yu, Li, Tiemeng Ma, Sirong Wu, and Zhuoyang Lyu. 2023. "How does broadband internet affect firm-level labor misallocation: The role of information frictions." *China Economic Review* 82 102067. 10.1016/j.chieco.2023.102067.

Zolas, Nikolas, Zachary Kroff, Erik Brynjolfsson et al. 2020. "Advanced Technologies Adoption and Use by U.S. Firms: Evidence from the Annual Business Survey." Working Paper 28290, National Bureau of Economic Research.

A Appendix

Table A1: Lightning strikes in density: stationarity test results.

Test Name	Test Statistic	Lag order	P-value
Augmented Dickey-Fuller (ADF)	-6.12	5	0.01
Kwiatkowski-Phillips-Schmidt-Shin (KPSS)	0.30	4	0.11
Phillips-Perron (PP)	-52.59	4	0.01

Note: The table reports the results of three different stationarity tests conducted on the time series of lightning strikes density (average number of daily lightning strikes per square kilometre) in France. Lightning strikes density is obtained from the WWLLN dataset (Kaplan, 2023) and is measured monthly between 2010 and 2020. The null hypotheses of the ADF and PP tests states that the time series is non-stationary, while for the KPSS the null hypothesis states that the series is stationary.

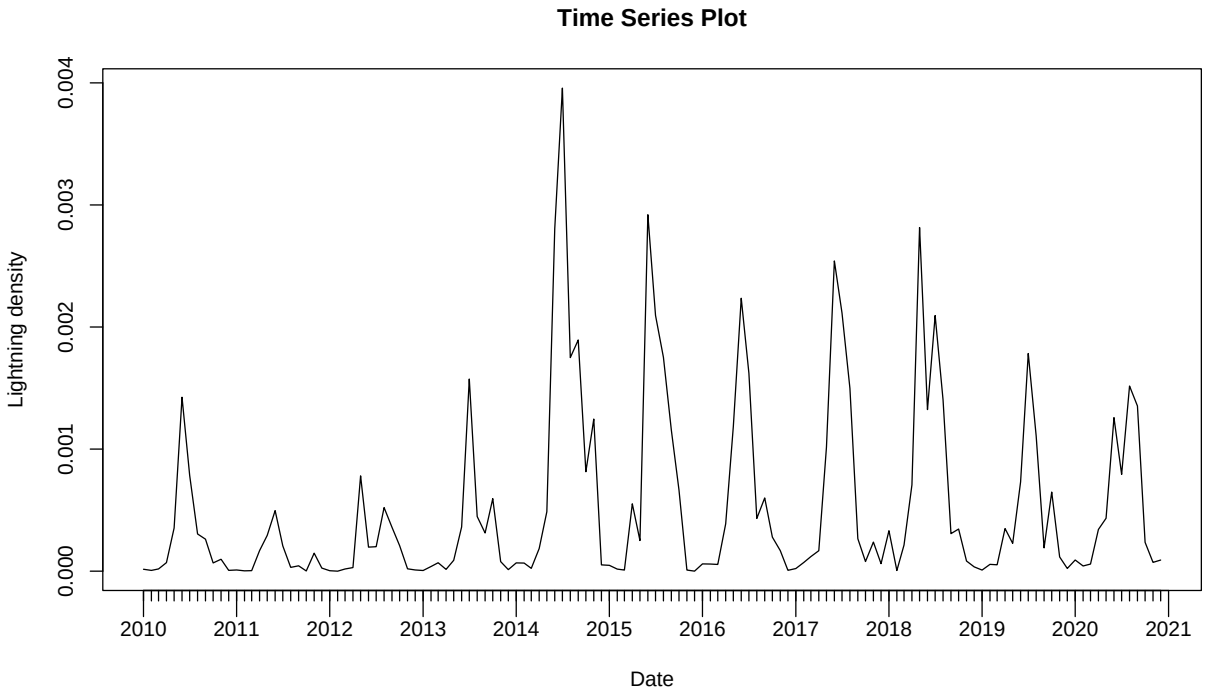


Figure A1: Lightning density monthly time series for France, 2010-2021

Table A2: Alternative definitions of sales growth rates.

	Cloud in t, growth between t and $t+5$			Cloud in t, growth between $t-4$ and $t+1$			Cloud in t, growth between $t-3$ and $t+2$			Cloud in t, growth between $t-10$ and t		
	OLS	End. Treatment		OLS	End. Treatment		OLS	End. Treatment		OLS	End. Treatment	
		Growth Rate	Selection		Growth Rate	Selection		Growth Rate	Selection		Growth Rate	Selection
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Cloud	0.1983***	0.2288***		0.2643***	0.2814***		0.2068***	0.1986***		0.5141***	0.7107***	
	(0.0342)	(0.0444)		(0.0771)	(0.0761)		(0.0498)	(0.0551)		(0.1269)	(0.1076)	
Log Sales	0.0165**	0.0180**	0.1777***	-0.0205***	-0.0212***	0.1757***	-0.0101	-0.0097	0.1773***	-0.0866***	-0.0951***	0.1761***
	(0.0075)	(0.0084)	(0.0230)	(0.0050)	(0.0055)	(0.0216)	(0.0065)	(0.0068)	(0.0208)	(0.0126)	(0.0131)	(0.0220)
Cloud*Log Sales	-0.0176***	-0.0198***		-0.0189***	-0.0192***		-0.0155***	-0.0154***		-0.0309***	-0.0344***	
	(0.0033)	(0.0034)		(0.0063)	(0.0061)		(0.0043)	(0.0042)		(0.0104)	(0.0091)	
Log Physical Capital	0.0217***	0.0216***	0.0159	0.0204***	0.0203***	0.0266**	0.0234***	0.0235***	0.0324***	0.0078	0.0067	0.0232*
	(0.0039)	(0.0040)	(0.0126)	(0.0049)	(0.0049)	(0.0129)	(0.0045)	(0.0045)	(0.0121)	(0.0074)	(0.0079)	(0.0137)
Log Intangible Capital	-0.0325***	-0.0322***	0.0233	-0.0198***	-0.0199***	0.0053	-0.0269***	-0.0268***	0.0028			
	(0.0048)	(0.0050)	(0.0177)	(0.0043)	(0.0043)	(0.0173)	(0.0046)	(0.0046)	(0.0165)			
Log Age	-0.0244***	-0.0207**	-0.0876***	-0.0778***	-0.0774***	-0.1055***	-0.0653***	-0.0655***	-0.1015***	-0.1204***	-0.1162***	-0.0795***
	(0.0057)	(0.0085)	(0.0315)	(0.0090)	(0.0091)	(0.0178)	(0.0086)	(0.0086)	(0.0151)	(0.0108)	(0.0113)	(0.0196)
ICT Share	0.2401***	0.2142***	0.7109***	0.2083***	0.2046***	0.7189***	0.2036***	0.2049***	0.5429***	0.2663**	0.2396**	0.4452***
	(0.0704)	(0.0806)	(0.1991)	(0.0378)	(0.0405)	(0.1747)	(0.0402)	(0.0415)	(0.1817)	(0.1017)	(0.1043)	(0.1333)
R&D Share	-0.0388	-0.0575	-0.0608	0.1694***	0.1687***	0.1173	0.1360**	0.1366**	0.2105	0.3366***	0.3228***	0.2312
	(0.0584)	(0.0602)	(0.1736)	(0.0505)	(0.0505)	(0.1474)	(0.0550)	(0.0549)	(0.1295)	(0.1028)	(0.0982)	(0.1757)
Exporter	-0.0093	-0.0102	0.1924**	0.0185*	0.0177	0.1951**	0.0093	0.0097	0.2037**	0.0029	-0.0062	0.1842**
	(0.0125)	(0.0133)	(0.0831)	(0.0107)	(0.0113)	(0.0816)	(0.0081)	(0.0086)	(0.0821)	(0.0187)	(0.0198)	(0.0797)
Multi Establishment	-0.0193	-0.0154	0.1407***	-0.0176***	-0.0183***	0.1455***	-0.0146**	-0.0143**	0.1396***	0.0180	0.0087	0.1653***
	(0.0138)	(0.0124)	(0.0261)	(0.0064)	(0.0064)	(0.0224)	(0.0062)	(0.0063)	(0.0240)	(0.0159)	(0.0170)	(0.0428)
Log Average Hourly Wage Manag. & Ing.	0.0055**	0.0034	0.0416***	-0.0011	-0.0012	0.0396***	0.0022	0.0023	0.0331**	0.0115**	0.0106*	0.0223***
	(0.0025)	(0.0026)	(0.0133)	(0.0026)	(0.0027)	(0.0123)	(0.0025)	(0.0025)	(0.0132)	(0.0055)	(0.0057)	(0.0083)
Log Lightning Density			-0.0521***			-0.0568***			-0.0547***			-0.0656***
			(0.0094)			(0.0109)			(0.0107)			(0.0107)
Observations	15,686	14,428	14,428	22,250	22,250	22,250	22,266	22,266	22,266	17,478	17,479	17,479
Adjusted R^2	0.0981			0.107			0.0858			0.159		
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Estimation results of Equations 1 and 2, for alternative definitions of the sales growth rate. Across models, the dependent variable is measured as: the growth rate between time t and $t+5$ (columns 1-3); the growth rate between time $t-4$ and $t+1$ (columns 4-6); the growth rate between time $t-3$ and $t+2$ (columns 7-9); the growth rate between time $t-10$ and t (columns 10-12). Cloud is a dummy variable identifying the adoption of cloud technologies by time t. All the time-varying control variables are measured at $t-5$. Log Sales, Physical, and Intangible Capital are expressed in Log of Euros. Age is the firm age in years since the establishment. ICT and R&D shares express, respectively, the share of hours worked in ICT and R&D occupations. Exporter is a dummy that identifies whether the firms is engaged in exporting activities. Multi Establishment identify whether the firm has more than one economic establishment. Log Average Hourly Wage Manag. & Ing. measures the log of the hourly wage of management and engineers, expressed in Euros. Log lightning density is the logarithms of the (population weighted) density of daily lightning strikes at the municipality level, averaged between 2008 and 2017. All regressions are estimated using a Full Information Maximum Likelihood (FIML) estimator. Standard errors are clustered at the NACE 2-digit level. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Inclusion of an additional wave of the ICT Survey (2021).

	Model 1 OLS	Model 2 End. Treatment	
		Growth Rate	Selection
Cloud	0.2686***	0.2852***	
	(0.0720)	(0.0642)	
Log Sales	-0.0236***	-0.0243***	0.1832***
	(0.0055)	(0.0066)	(0.0242)
Cloud*Log Sales	-0.0180***	-0.0183***	
	(0.0060)	(0.0057)	
Log Physical Capital	0.0192***	0.0191***	0.0306**
	(0.0038)	(0.0038)	(0.0142)
Log Intangible Capital	-0.0209***	-0.0209***	-0.0042
	(0.0046)	(0.0046)	(0.0172)
Log Age	-0.0907***	-0.0903***	-0.0953***
	(0.0087)	(0.0089)	(0.0167)
ICT Share	0.1744***	0.1714***	0.5756***
	(0.0524)	(0.0525)	(0.1477)
R&D Share	0.1286***	0.1276***	0.1711
	(0.0482)	(0.0486)	(0.1304)
Exporter	0.0240**	0.0231*	0.1971**
	(0.0118)	(0.0123)	(0.0807)
Multi Establishment	-0.0017	-0.0024	0.1416***
	(0.0056)	(0.0051)	(0.0216)
Log Average Hourly Wage Manag. & Ing.	-0.0055**	-0.0056**	0.0348***
	(0.0026)	(0.0025)	(0.0087)
Log Lightning Density			-0.0554***
			(0.0106)
Observations	29,243	29,243	29,243
Adjusted R^2	0.130		
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes

Note: Estimation results of Equations 1 (Model 1) and 2 (Model 2), including an additional wave of the ICT survey (2021). Model 1 is estimated using OLS, Model 2 uses a FIML estimator. In Model 2, the first column reports the results of the outcome equation, while the second column refers to the selection equation. The dependent variable measures the sales rate of growth of the firm between t and $t-5$. Cloud is a dummy variable identifying the adoption of cloud technologies by time t. All the time-varying control variables are measured at $t-5$. Log Sales, Physical, and Intangible Capital are expressed in Log of Euros. Age is the firm age in years since the establishment. ICT and R&D shares express, respectively, the share of hours worked in ICT and R&D occupations. Exporter is a dummy that identifies whether the firms is engaged in exporting activities. Multi Establishment identify whether the firm has more than one economic establishment. Log Average Hourly Wage Manag. & Ing. measures the log of the hourly wage of management and engineers, expressed in Euros. Log lightning density is the logarithms of the (population weighted) density of daily lightning strikes at the municipality level, averaged between 2008 and 2017. All regressions are estimated using a Full Information Maximum Likelihood (FIML) estimator. Standard errors are clustered at the NACE 2-digit level. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Long differences across waves of the ICT survey.

	2009-2019			2009-2017			2009-2015		
	Model 1	Model 2		Model 3	Model 4		Model 5	Model 6	
	OLS	End. Treatment		OLS	End. Treatment		OLS	End. Treatment	
		Growth Rate	Selection		Growth Rate	Selection		Growth Rate	Selection
Cloud	0.4193***	0.5980***		0.4048***	0.6423***		0.3073***	0.7588***	
	(0.1200)	(0.1293)		(0.1212)	(0.1867)		(0.0922)	(0.1886)	
Log Sales	-0.0550***	-0.0646***	0.1776***	-0.0411***	-0.0540**	0.2323***	-0.0337***	-0.0469***	0.1612***
	(0.0157)	(0.0190)	(0.0270)	(0.0130)	(0.0227)	(0.0290)	(0.0079)	(0.0117)	(0.0230)
Cloud*Log Sales	-0.0270**	-0.0287***		-0.0256**	-0.0304***		-0.0194**	-0.0313***	
	(0.0111)	(0.0109)		(0.0101)	(0.0075)		(0.0078)	(0.0062)	
Log Physical Capital	0.0347**	0.0319**	0.0570**	0.0412***	0.0415***	-0.0052	0.0229***	0.0154*	0.0826***
	(0.0155)	(0.0155)	(0.0275)	(0.0138)	(0.0132)	(0.0255)	(0.0074)	(0.0091)	(0.0250)
Log Intangible Capital	-0.0373**	-0.0359**	-0.0239	-0.0414***	-0.0413***	-0.0013	-0.0211**	-0.0153	-0.0675*
	(0.0169)	(0.0169)	(0.0353)	(0.0143)	(0.0136)	(0.0399)	(0.0086)	(0.0096)	(0.0371)
Log Age	-0.1108***	-0.1037***	-0.1347***	-0.1071***	-0.1028***	-0.0615**	-0.0935***	-0.0874***	-0.0368
	(0.0126)	(0.0126)	(0.0219)	(0.0129)	(0.0127)	(0.0272)	(0.0116)	(0.0114)	(0.0401)
ICT Share	0.3275**	0.3005**	0.4974***	0.2977***	0.2617**	0.4972***	0.2195***	0.2003***	0.1201
	(0.1327)	(0.1316)	(0.1744)	(0.0998)	(0.1135)	(0.1831)	(0.0677)	(0.0661)	(0.2246)
R&D Share	0.1919	0.1539	0.7474***	0.2121	0.2218	-0.1508	0.1485*	0.1014	0.3726*
	(0.1643)	(0.1572)	(0.2467)	(0.1507)	(0.1569)	(0.2933)	(0.0839)	(0.0897)	(0.2135)
Exporter	0.0306	0.0213	0.1772*	0.0374*	0.0252	0.2047**	0.0418***	0.0277*	0.1554**
	(0.0196)	(0.0225)	(0.1009)	(0.0191)	(0.0209)	(0.0866)	(0.0133)	(0.0168)	(0.0721)
Multi Establishment	-0.0104	-0.0231	0.2138***	0.0022	-0.0045	0.0945**	-0.0198	-0.0356**	0.1389***
	(0.0208)	(0.0197)	(0.0542)	(0.0133)	(0.0120)	(0.0428)	(0.0123)	(0.0150)	(0.0404)
Log Average Hourly Wage Manag. & Ing.	0.0051	0.0032	0.0347**	-0.0020	-0.0035	0.0321**	-0.0007	-0.0036	0.0367**
	(0.0065)	(0.0061)	(0.0135)	(0.0065)	(0.0067)	(0.0128)	(0.0039)	(0.0040)	(0.0180)
Log Lightning Density			-0.0597***			-0.0589***			-0.0586***
			(0.0176)			(0.0170)			(0.0130)
Observations	5,264	5,264	5,264	6,829	6,829	6,829	7,594	7,594	7,594
Adjusted R^2	0.144			0.124			0.123		
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Estimation results of Equation 3 over different waves of the ICT surveys. Across models, the beginning of the period is set to 2009. The period ends respectively in 2019 (Model 1), 2017 (Model 2), and 2015 (Model 3). In each model, the first column reports the OLS results, the second and the third include the results of the ET model (respectively for the outcome and selection equations). Cloud is a dummy variable identifying the adoption of cloud technologies within the period under consideration. All the time-varying control variables are measured in 2009. Log Sales, Physical, and Intangible Capital are expressed in Log of Euros. Age is the firm age in years since the establishment. ICT and R&D shares express, respectively, the share of hours worked in ICT and R&D occupations. Exporter is a dummy that identifies whether the firms is engaged in exporting activities. Multi Establishment identify whether the firm has more than one economic establishment. Log Average Hourly Wage Manag. & Ing. measures the log of the hourly wage of management and engineers, expressed in Euros. Log lightning density is the logarithms of the (population weighted) density of daily lightning strikes at the municipality level, averaged between 2008 and 2017. All regressions are estimated using a Full Information Maximum Likelihood (FIML) estimator. Standard errors are clustered at the NACE 2-digit level. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Two stages least squares.

	2nd Stage	1st Stage – Cloud	1st Stage – Interaction
Cloud	0.497*		
	(0.274)		
Cloud*Log Sales	-0.042**		
	(0.020)		
Log Sales	-0.016**	0.034***	0.282***
	(0.008)	(0.008)	(0.089)
Log Physical Capital	0.020***	0.009***	0.102***
	(0.004)	(0.003)	(0.035)
Log Intangible Capital	-0.017***	0.000	0.047
	(0.004)	(0.004)	(0.035)
Log Age	-0.092***	-0.029***	-0.305***
	(0.005)	(0.004)	(0.035)
ICT Share	0.110**	0.184***	1.379***
	(0.050)	(0.043)	(0.399)
R&D Share	0.201***	0.046	0.659
	(0.049)	(0.050)	(0.485)
Exporter	0.018*	0.062***	0.556***
	(0.010)	(0.007)	(0.072)
Multi Establishment	-0.002	0.043***	0.427***
	(0.008)	(0.007)	(0.067)
Log Average Hourly Wage Manag. & Ing.	-0.005	0.006***	-0.036*
	(0.003)	(0.002)	(0.019)
Log Lightning Density		0.003	0.391***
		(0.006)	(0.065)
Log Lightning Density*Log Sales		-0.002***	-0.063***
		(0.001)	(0.008)
Observations	21,886	21,886	21,886
Adjusted R^2	0.0819		
Cragg-Donald test		33.81	
Kleibergen-Paap test		34.47	
F-statistic		40.39	55.88
Industry FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Note: The table reports the results of the estimation of Equation 1, using Two Stages Least Squares with lightning strike density as instrument. In each model, the first column reports the results of the second stage of the estimation, while the second and third columns refer to the first stage results of cloud and its interaction with size. The dependent variable measures the sales rate of growth of the firm between t and $t-5$. Cloud is a dummy variable identifying the adoption of cloud technologies by time t. All the time-varying control variables are measured at $t-5$. Log Sales, Physical, and Intangible Capital are expressed in Log of Euros. Age is the firm age in years since the establishment. ICT and R&D shares express, respectively, the share of hours worked in ICT and R&D occupations. Exporter is a dummy that identifies whether the firms is engaged in exporting activities. Multi Establishment identify whether the firm has more than one economic establishment. Log Average Hourly Wage Manag. & Ing. measures the log of the hourly wage of management and engineers, expressed in Euros. Log lightning density is the logarithms of the (population weighted) density of daily lightning strikes at the municipality level, averaged between 2008 and 2017. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic and Cragg-Donald test statistic, and the results of the F-statistic testing the equivalence between the first stages and the nested model restricting the coefficients of instruments to 0. Standard errors are clustered at the NACE 2-digit level. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Comparing cloud users and non users after CEM

	Baseline		Export + Multi Est.		Île-de-France		10 Size Cutpoints		20 Size Cutpoints	
Cloud Users	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Sales	132741	139502.90	160079.60	172275.70	117086.00	125834.30	137176.50	148352.00*	113881.40	116829.40
Physical Capital	57980.13	59237.67	76247.31	78469.16	57320.97	60127.99	57623.95	53110.34	62796.28	56185.67*
Intangible Capital	655.31	717.18	851.49	934.49	613.73	707.36	673.62	686.72	680.70	676.97
Average Hourly Wage Manag. & Ing.	25.88	25.92	25.84	25.80	24.76	24.76	26.79	26.49	25.29	25.31
Age	30.59	30.45	31.00	31.14	29.90	29.86	30.50	30.58	29.98	29.74
ICT Share	2.95%	2.67%	2.31%	1.92%	2.23%	2.20%	3.15%	3.17%	2.72%	2.72%
R&D Share	3.56%	3.21%	3.52%	3.12%	2.86%	2.54%	3.62%	3.54%	3.40%	3.16%
Exporter	52.11%	54.01%	53.42%	53.42%	48.85%	49.53%	54.06%	55.75%	50.68%	52.55%
Multi Establishment	54.99%	56.10%	56.15%	56.15%	51.92%	52.84%	54.67%	57.18%	53.52%	54.63%
Count	3000	1820	1759	1170	2305	1391	3850	2228	2532	1589

Note: The table reports the averages of controls computed for cloud users and non users across samples balanced via the CEM procedure described in 4.2 and used for estimating Equation 1 in Table A7. Stars refer to the p-value of t-tests on the difference between the averages of users and non users. The averages are computed using CEM weights as well as the corresponding t-tests. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: CEM-balanced OLS estimation.

	Baseline	Exp and Multi Est.	Île-de-France	10 Size Cutpoints	20 Size Cutpoints
	Model 1	Model 2	Model 3	Model 4	Model 5
Cloud	0.1798***	0.1847***	0.1795**	0.1724***	0.1856***
	(0.0620)	(0.0705)	(0.0711)	(0.0597)	(0.0689)
Log Sales	-0.0070	-0.0124	-0.0033	-0.0146	-0.0134
	(0.0108)	(0.0137)	(0.0122)	(0.0097)	(0.0113)
Cloud*Log Sales	-0.0126**	-0.0131**	-0.0133*	-0.0113**	-0.0116*
	(0.0059)	(0.0066)	(0.0068)	(0.0057)	(0.0065)
Log Physical Capital	0.0175	0.0000	0.0207	0.0196*	0.0019
	(0.0131)	(0.0159)	(0.0158)	(0.0114)	(0.0137)
Log Intangible Capital	-0.0255**	-0.0040	-0.0302**	-0.0224**	-0.0048
	(0.0127)	(0.0161)	(0.0148)	(0.0110)	(0.0148)
Log Average Wage Manag. & Ing.	-0.0053	-0.0023	-0.0043	0.0026	-0.0035
	(0.0064)	(0.0077)	(0.0072)	(0.0056)	(0.0073)
Log Age	-0.0846***	-0.0886***	-0.0950***	-0.0876***	-0.1001***
	(0.0133)	(0.0162)	(0.0146)	(0.0129)	(0.0160)
ICT Share	-0.0712	-0.0689	-0.0169	0.0137	0.0607
	(0.1000)	(0.1751)	(0.1241)	(0.0993)	(0.1202)
R&D Share	0.3823***	0.2701	0.3811**	0.3129**	0.3159*
	(0.1280)	(0.1684)	(0.1604)	(0.1315)	(0.1703)
Exporter	0.0467**	0.0640**	0.0572***	0.0243	0.0427**
	(0.0188)	(0.0291)	(0.0204)	(0.0159)	(0.0204)
Multi Establishment	-0.0057	-0.0247	-0.0113	-0.0083	-0.0250
	(0.0162)	(0.0247)	(0.0181)	(0.0147)	(0.0172)
Observations	4,820	2,929	3,696	6,078	4,121
Adjusted R^2	0.104	0.118	0.126	0.106	0.122
Industry FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Note: The table reports the results of the estimation of Equation 1 after balancing the sample using Coarse Exact Matching as described in Section 4.2. The dependent variable measures the sales rate of growth of the firm between t and $t-5$. Cloud is a dummy variable identifying the adoption of cloud technologies by time t. All the time-varying control variables are measured at $t-5$. Log Sales, Physical, and Intangible Capital are expressed in Log of Euros. Age is the firm age in years since the establishment. ICT and R&D shares express, respectively, the share of hours worked in ICT and R&D occupations. Exporter is a dummy that identifies whether the firms is engaged in exporting activities. Multi Establishment identify whether the firm has more than one economic establishment. Log Average Hourly Wage Manag. & Ing. measures the log of the hourly wage of management and engineers, expressed in Euros. All regressions are estimated using OLS. Standard errors are clustered at the NACE 2-digit level. *** p<0.01, ** p<0.05, * p<0.1.