Calvino, Flavio; Fontanelli, Luca

Working Paper

AI Users Are Not All Alike: The Characteristics of French Firms Buying and Developing AI

CESifo Working Paper, No. 11466

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

CESifo WORKING PAPERS

11466
2024

November 2024

AI Users Are Not All Alike: The Characteristics of French Firms Buying and Developing AI

Flavio Calvino, Luca Fontanelli

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 11466*

AI Users Are Not All Alike: The Characteristics of French Firms Buying and Developing AI

Abstract

In this work we characterise French firms using artificial intelligence (AI) and explore the link between AI use and productivity. We distinguish AI users that source AI from external providers (AI buyers) from those developing their own AI systems (AI developers). AI buyers tend to be larger than other firms, but this relation is explained by ICT-related variables. Conversely, AI developers are larger and younger beyond ICT. Other digital technologies, digital skills and infrastructure play a key role for AI use, with AI developers leveraging more specialised ICT human capital than AI buyers. Overall, AI users tend to be more productive, however this is related to the self-selection of more productive and digital-intensive firms into AI use. This is not the case for AI developers, for which the positive link between AI use and productivity remains evident beyond selection.

JEL-Codes: D200, J240, O140, O330.

Keywords: technology diffusion, artificial intelligence, digitalisation, productivity.

Flavio Calvino
OECD Directorate for Science, Technology and Innovation
flavio.calvino@oecd.org

Luca Fontanelli
University of Brescia / Italy
luca.fontanelli@unibs.it

November 4, 2024
The views expressed here are those of the authors and cannot be attributed to the OECD or its member countries. We thank Simone Alfarano, Erik Brynjolfsson, Giovanni Dosi, Flora Bellone, Chiara Criscuolo, Timothy DeStefano, Marco Grazzi, Mattia Guerini, Francesco Manaresi, Luca Marcolin, Mauro Napoletano, Lionel Nesta, Ariell Reshef, Daniel Rock, Angelo Secchi, Federico Tamagni, Neil Thompson, Daniel Trefler, Elena Verdolini, Frederic Warzynski, the participants in the 2022 NBER Summer Institute on IT and digitization, the OECD Applied Economics seminar, the Pillars workshop on AI, Digitalization, and Labor Outcomes, the 2024 Royal Economic Society, the 2024 DRUID and the 2024 CESifo Economics of Digitization conference for insightful comments. Previous drafts of this work had the title “A portrait of AI adopters: evidence from France” and "Artificial intelligence, complementary assets and productivity: evidence from French firms" (LEMWorking Paper 2023/35). 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).

1 Introduction

Artificial intelligence (AI) is rapidly transforming economies and societies. It permeates a wide range of products and services used by consumers on a daily basis, it is changing the demand for skills, and may play an important role to tackle societal challenges such as climate change. Notably, AI has the potential to boost the productivity of its adopters (Brynjolfsson et al., 2021) and stimulate economic growth (Besiroglu et al., 2022), while spurring innovation (Cockburn et al., 2018; Bianchini et al., 2022) and enabling workers to operate more efficiently (Brynjolfsson et al., 2023; Noy and Zhang, 2023). However, empirical evidence characterising the diffusion of AI across the economy and assessing whether how AI adoption is related to firm productivity remains in its early stages (Acemoglu et al., 2022; Zolas et al., 2020; Alekseeva et al., 2021; Babina et al., 2024), particularly outside the United States. On the one hand, the diffusion of AI technologies among firms is still limited (see e.g. McElheran et al., 2023; Calvino and Fontanelli, 2023), also due to the high fixed costs of developing large datasets and acquiring the ICT skills necessary for deploying AI systems (Brynjolfsson et al., 2021; Fontanelli et al., 2024). On the other hand, productivity returns may fail to emerge because firms may not have yet learned how to effectively implement AI systems. The integration of AI systems within organisations is not merely an additive process to the firm's operations and activities, as discussed by Agrawal et al. (2022). Instead, it involves a comprehensive embedding of AI applications and functionalities across various dimensions of the firm. This integration process is complex and may need a careful alignment of AI systems with organisational workflows and objectives before generating significant productivity gains (Dell'Acqua et al., 2023).

In this context, make-or-buy decisions regarding AI may depend on the characteristics of AI users and may differently affect productivity returns from AI adoption. Purchasing AI solutions from external providers is likely less costly, as it limits the need for in-house IT assets necessary to develop AI in-house. Indeed, developing AI systems internally may require larger investments in IT resources and human capital.[1] Nevertheless, building AI systems in-house offers greater versatility and control over the implementation process, enabling firms to tailor AI solutions more closely to their specific needs, better integrate them into their organisational structures, and more readily adapt them based on feedbacks. In this sense, while in-house AI development may involve higher fixed costs and is more accessible to larger firms, it can also yield greater productivity gains, as the AI system can be more effectively and flexibly customised to meet firm-specific needs.

In this work we characterise French firms using AI and explore the link between AI use and productivity by distinguishing firms sourcing AI from external providers (AI buyers) and those developing their own AI systems (AI developers). In particular, we leverage detailed survey microdata – collected by the French statistical office – that include comprehensive information on the use of digital technologies in 2018, a period wherein the use of AI systems by firms predominantly captures *predictive* AI, aimed at generating out-of-sample predictions based on data (e.g., customer classification via machine learning algorithms), process automation (e.g., data collection via text mining, intelligent robots) and upgrading (e.g., anomalies detection and quality control). We match this survey with firms' balance sheets that contain additional information on their characteristics and financials. The level of detail of these data and their rep-

[1] For instance, highly-skilled occupations – such as data scientists, machine learning engineers and software developers – and up-to-date IT capital are necessary to build and maintain an AI infrastructure.

resentativeness of the French economy – differently from other commercial surveys – make them unique sources to analyse the patterns of AI use among firms in great detail.

Our contribution to the literature is twofold, and relates to the distinction between firms buying vis-á-vis developing in-house AI systems. First, we uncover a series of stylised facts concerning the diffusion of AI in France. In particular, we explore the characteristics of different types of AI users and the role played by IT assets (i.e. ICT skills, digital infrastructure and the use of other digital technologies) for AI use. On one hand, we focus on firm age, size and sectoral heterogeneity, first pooling together all AI users, and then distinguishing AI buyers from developers. We show that AI users tend to be overall younger and larger than other firms, in line with existing evidence (see for instance Acemoglu et al., 2022), and that sectoral shares are highest in the ICT and professional services sectors. However, we show that the AI-size relation appears largely explained by IT assets. Furthermore, when we separate out AI buyers and developers, several differences emerge. Differently from buyers, AI developers remain larger than other firms even after accounting for these confounding factors, and are also significantly younger than other firms. This result highlights that the development of AI systems may be linked to higher fixed costs and new managerial capabilities. Sectoral shares of firms buying AI are more homogeneous than the ones of AI developers, which exhibit a particularly high share in the ICT sector. On the other hand, once we focus on the role of various measures of firm digitalisation and human capital, we find that these are all positively and significantly related to AI use. AI use is indeed more likely among firms using larger bundles of business digital technologies, and it is thus likely fostered by the presence of a digital architecture within the firm, through which business data can be more easily stored and managed. ICT skills (i.e., the presence of ICT specialists and the provision of ICT training to non-ICT personnel) and digital infrastructure (i.e., the use of a fast broadband connection) also play a critical role for AI use. When distinguishing AI buyers and developers, results show that ICT training is associated with a higher probability of purchasing AI, while the presence of ICT specialists is positively linked with AI development. This suggests that different types of AI users leverage different types of ICT-related human capital, with AI developers hinging on more specialised ICT human capital than AI users.

Second, we analyse the relationship between AI use and labour productivity. On average, AI users tend to be more productive than other firms. However, this appears largely related to the selection of more productive and highly-digital firms into the use of AI (consistently with Acemoglu et al., 2023). Indeed, when considering all AI users together, the link disappears after controlling for IT assets and past productivity. Similar results characterise the productivity of AI buyers. However, this is not the case for AI developers, for which the AI-productivity link remains significant. This is robust to the inclusion of several confounding factors, to the use of alternative models (i.e., an endogenous treatment model, long-difference regressions) and measures of productivity (i.e., multi-factor productivity), to the exclusion of firms belonging to ICT business services industries, and to the estimation on comparable subsamples of AI users, buyers and developers.

Our findings provide valuable insights into how managers can effectively implement AI systems within their organisation and how policymakers can foster a wide diffusion of AI and its returns across firms and sectors, ensuring an inclusive digital transformation (see also Goos and Savona, 2024). On the one hand, it may not be enough to use a novel and disruptive technology such as AI to boost productivity. Indeed, our findings are coherent with the idea that firms better adapting AI systems to their productive and organisational structure by developing their own AI solutions already experiment productivity returns from AI use. On the other hand, our research indicates that AI buyers and developers are profoundly different, with AI development predominantly driven by large and young firms with specialised ICT skills. This underscores the importance of ICT human capital, the capacity to sustain high ICT-related expenses (e.g., the extensive cumulation of meaningful data), and new managerial capabilities in the development of AI systems. However, it also suggests that only a handful of firms have been already benefiting from predictive AI systems.

The remainder of the paper is organised as follows. In Section 2 we review recent contributions to the literature related to AI use by firms, focusing on those mostly related to our paper. In Section 3 we discuss in detail about the data sources used in this work, the French ICT survey and balance sheet data, and provide a series of basic summary statistics of the main variables used. In Section 4 we focus on the characteristics of firms using AI vis-à-vis those of other firms and explore the role of IT assets. In Section 5 we analyse the relation between the use of AI and productivity, using different empirical models. Finally, in Section 6 we draw some conclusions and point to next steps for future analysis.

2 Existing evidence on AI use by firms

Studies on AI use by firms primarily rely on US data and are mainly based on three data sources: firm-level ICT surveys, online job posting data that contain information on AI skills demand, and Intellectual Property (IP) records, in particular patents.[2] The literature more closely related to the current analysis on AI use by firms highlights four key facts.

First, notwithstanding the surge in the demand for AI jobs (Squicciarini and Nachtigall, 2021; Borgonovi et al., 2023; Alekseeva et al., 2021) and AI related innovations (Dibiaggio et al., 2022), the diffusion of AI technologies is still limited and heterogeneous across sectors. Evidence from ICT surveys carried out in the United States, Germany and Korea (Zolas et al., 2020; Rammer et al., 2022; Cho et al., 2022; McElheran et al., 2023), from matching different data sources in the United Kingdom (Calvino et al., 2022) and from recent cross-country analysis (Calvino and Fontanelli, 2023) shows that the use of AI technologies is still rare and concentrated in the ICT and professional services sectors. AI innovations occur more frequently in ICT-related services and manufacturing, and in the Wholesale & Retail sector (Igna and Venturini, 2023; Santarelli et al., 2022).

Second, existing evidence highlights a positive relation between AI adoption and firm size, which can be driven by both self-selection of larger firms into AI use and by a positive effect of AI on firm size. The probability to use AI is positively linked with size (McElheran et al., 2023; Rammer et al., 2022; Calvino et al., 2022; Calvino and Fontanelli, 2023; Dahlke et al., 2024). Ex-ante larger firms demand more intensively AI skills (Alekseeva et al., 2021; Babina et al., 2024). The self-selection of larger firms into AI use can be explained by scale advantages related to the availability of high computing power, massive amount of data and, more in general, of intangibles related to AI, which are necessary

[2]Recent contributions to the literature have also used other proxies of digital investments such as automation shocks or IT expenditures (see e.g. Jin and McElheran, 2018; Acemoglu and Restrepo, 2020; Aghion et al., 2020; Domini et al., 2021, 2022), which – although related – do not only proxy for AI use by firms but have instead a broader nature. Recently the literature has also used information from online websites to identify and characterise companies and organisations with an AI-related online presence (Dernis et al., 2023). Concerning ICT surveys, these are typically conducted by statistical offices. However, there is a part of literature that examines survey data collected by other institutions (see for instance Cette et al., 2022; Bessen et al., 2022; Czarnitzki et al., 2022; Cirillo et al., 2023).

to reap the full benefits of AI technologies (Brynjolfsson and McAfee, 2014; Brynjolfsson et al., 2021). However, part of the literature suggests that the relation between AI use and size is not exclusively limited to self-selection, as AI also has a positive impact on firm size in the US, mostly through innovations (see Babina et al., 2024; Alderucci et al., 2021; Damioli et al., 2023).[3] Relatedly, Conti et al. (2024) estimate a positive effect of the use of Big Data analytics on firms' value added and sales.

Third, some analyses suggest that a wave of high-tech young firms has been driving – at least partly – the development of AI technologies (Calvino and Fontanelli, 2023; Acemoglu et al., 2022; Calvino et al., 2022), notwithstanding the role of high entry costs for AI startups (for instance, in terms of proprietary data, see Bessen et al., 2022). Relatedly, venture capital investments in AI startups has been significantly growing over time (Tricot, 2021), in line with the existence of a generation of AI start-ups. Furthermore, in the US, AI adoption by young firms is also related to indicators of high-growth entrepreneurship, with few cities and emerging hubs leading AI adoption by startups (McElheran et al., 2023).

Finally, the evidence on the firm-level relation between firm productivity and AI use is mixed.[4] Using measures of AI skills, Babina et al. (2024) and Alekseeva et al. (2020) do not find a robust effect of AI on the productivity of firms. A positive impact of AI-related innovations (i.e., patents) has been instead found on the productivity of SMEs and service firms and on the output and costs per worker in a dataset including worldwide patenting firms (Damioli et al., 2021), and on the output per workers of US patentees (Alderucci et al., 2021). This impact is found to be larger in more productive patentees from 15 European countries (Venturini et al., 2024, see also Ikeuchi et al., 2023 for related analysis on Japan) and driven by changes in workforce composition (Yang, 2022). When considering ICT surveys, Czarnitzki et al. (2022) find a positive impact of AI on German firms' productivity. However, Acemoglu et al. (2022) show that the relation between labour productivity and AI use is not significant for US firms. Finally, Calvino and Fontanelli (2023) show that the productivity premia of AI users tend to disappear – or to reduce – in several OECD countries when the role of IT assets and technologies is taken into account. This evidence clashes against studies finding a positive effect of ICT and digitalisation on productivity at the firm level (Brynjolfsson and Hitt, 2003; Commander et al., 2011; DeStefano et al., 2023), but it is consistent with the existence of a lag in the effect of AI on productivity after following its adoption (Brynjolfsson et al., 2021). Furthermore, an effective use of AI by firms may need a complex restructuring of their productive and organisational structure (Agrawal et al., 2022), suggesting that different patterns of adoption (e.g. developing in-house vs buying) may affect the returns to performance from AI use.

Our work is innovative with respect to the above literature in several aspects. First, we analyse representative

[3]Similarly, AI exposure is positively related to the change of employment of European regions (Guarascio et al., 2023) and employment share of sectors-occupations within European countries (Albanesi et al., 2023).

[4]A recent wave of works tends to find a positive impact of generative AI on the productivity of specific categories of workers (see also Brynjolfsson et al., 2023; Noy and Zhang, 2023; Peng et al., 2023; Eloundou et al., 2023; Kreitmeir and Raschky, 2023). However, the direction of such relation seems to depend on the extent to which tasks are within or outside the current capabilities of AI systems (Dell'Acqua et al., 2023). This evidence is complementary to our work. First, the use of generative AI by workers is just one of the types of AI systems that can be adopted by firms, and it was unlikely to be used by the firms considered in our analysis in 2018. Second, these analyses often refer to specific categories of workers. Finally, the extent to which AI-driven increases in workers' productivity translate into firms' productivity needs still to be explored. As a matter of fact, Dell'Acqua et al. (2023) provide experimental evidence suggesting that the introduction of AI may reduce team performance and increase coordination failures and Svanberg et al. (2024) highlights that only a limited share of worker compensations exposed to AI would be cost-effective to automate.

microdata for France, matching the ICT survey with balance sheet data. These data have not yet been explored for such purpose by the literature, which tends to focus on US firms. Furthermore, the ICT survey includes information on relevant confounding variables related to digital technologies and skills, whose role has been shown to be relevant by the above literature. Second, our data allow for a further characterisation of AI users, by providing information on the source of AI systems used by firms (also see Hoffreumon et al., 2024). Notably, we show that AI users are not all alike by distinguishing AI buyers and developers. As will be discussed in the next sections, the differences between these two groups of firms span over several dimensions, including size, age, sector, ICT skills of workers, and productivity.

3 Data

Our analysis is based on microdata from the 2019 French ICT survey.[5] The survey is administered by INSEE (the French statistical office). It consists of a rotating sample of about 9000 firms from manufacturing and market-services sectors with questions related to the use of advanced digital technologies in 2018. The sample is representative of the population of firms with employment greater than or equal to 10 and is exhaustive for firms with more than 500 employees. These data are characterised by a unique level of detail and representativeness compared to other commercial surveys, which allow an in-depth analysis of AI adoption patterns among firms. Furthermore, they can be easily merged with other sources of French firms' data thanks to the *Siren* code, which uniquely identifies French companies.

Part of the ICT survey is dedicated to questions on AI use by firms. In particular, firms are asked whether they used AI technologies in 2018.[6] Our main AI use variable takes thus the form of a dummy, which indicates whether firms use AI technologies or not. Furthermore, the ICT survey relevantly allows to separate out AI users into AI buyers vis-à-vis AI developers. In particular, AI buyers are firms using AI technologies developed by external providers (off-the-shelf or customised), while AI developers use AI systems developed in-house.

Our data refer to a period prior to the recent boom in generative AI, focusing therefore on predictive AI systems aimed at performing data-driven out-of-sample predictions (e.g., forecasts and classifications). On the one hand, predictive AI systems find application in several activities spanning multiple industries. These include applications relative to operations (e.g., pricing, sales and demand forecasting), supply chain (e.g., inventory churn prediction and energy management) and human resources (e.g., employees turnover forecasts, selection of candidates for hirings, employees classification based on performance analysis) management, customer behaviour classification for marketing purposes, predictive maintenance and anomalies detection in IT and industrial machineries and infrastructures (e.g., power plants), and cybersecurity (e.g., detection and prediction of IT fraud).

The ICT survey also includes questions on the use of other business digital technologies or tools, i.e. the use of Costumer Relation Management (CRM), Enterprise Resource Planning (ERP) software, and regarding e-commerce activities. We count the number of these business digital technologies to build the variable "Number of digital technologies", which takes values equal to the number of technologies used by the firm (from 0 to 3) and represents a proxy for its level of

[5] *"Enquête sur les Technologies de l'Information et de la Communication (TIC)"*, further information about the survey can be found here.

[6] Firms are asked the following question: "In 2018, did your company make use of software and/or equipment incorporating artificial intelligence technologies?".

digitalisation.[7] In the survey, firms are also asked about the presence of ICT specialists. We use this variable as a proxy for the presence of digital skills within the firm. In this respect, we also use the presence of ICT training for non-ICT specialists, as firms relying on new AI technologies may also need to build the IT skills of the existing workforce. Finally, we use the speed of the broadband connection as a measure of digital infrastructure. In particular, we build a dummy for the presence of a fast broadband connection, which takes value equal to 1 in presence of a speed greater or equal to 100 mbit/second.[8]

We match the ICT survey with French firms' balance sheet data between 2011 and 2019.[9] This data allow us to gather information on firm sales, age, employment and value added (i.e., sales net of intermediate costs) and to compute a proxy of labour productivity as the ratio of value added to the number of workers employed by the firm.[10]

All the regressions and summary statistics reported in this work have been weighted using probability weights available in the ICT survey.

		AI User		AI Buyer		AI Developer	
	All firms	User	Non-User	Buyer	Non-Buyer	Developer	Non-Developer
Age	24	24	24	24	24	21	24
Productivity	64000	70000	64000	68000	64000	92000	63000
Employment	63	154	51	137	55	268	56
ICT Specialists	0.17	0.30	0.16	0.23	0.17	0.59	0.16
ICT Training	0.19	0.28	0.18	0.25	0.18	0.40	0.18
Fast Broadband Connection	0.13	0.21	0.12	0.19	0.13	0.36	0.12
Number of digital tech.	0.90	1.16	0.86	1.11	0.88	1.54	0.88

Table 1: Averages for the whole sample and classifying firms by AI type. Age and employment statistics have been rounded to the closest unity, productivity to the closest thousand.

Based on the database described above, we report on Table 1 a series of summary statistics, which also allow for a first basic (unconditional) comparison of AI users, buyers and developers with other firms. AI users, buyers and developers are on average larger and more productive than their counterparts. They also more likely employ ICT specialists and provide ICT training to their workers. The presence of a fast broadband connection is more likely in AI firms, which are characterised by a higher average use of digital technologies other than AI. Finally, AI developers are younger than other firms, whereas AI buyers and AI users (when considered altogether) are very close in terms of age.

4 The characteristics of AI adopters

In this section we discuss the main characteristics of AI users in the database – disentangling between AI buyers and AI developers – and focus on the role of IT-related variables, such as digital infrastructure, skills, and other digital

[7]Business digital technologies are less likely to be related to sectoral specificities than other advanced technologies, whose use may be largely sector specific. For example, Robots and 3-D Printers are more likely to be used in manufacturing than in consulting services and would not be useful in identifying a more general proxy of firm digitalisation.

[8]In the 2019 wave of the ICT survey this is the highest speed included among the possible choices in the question about broadband connection speed.

[9]Further information about balance sheet data can be found here.

[10]Data on sales and value added are in real terms and have been deflated at the 2-digit sector level.

technologies for AI use. French AI users account for 11.26% of French firms with more than 10 employees. The use of AI technologies is thus still limited, but probably slightly higher than in some other countries (see the discussion in Section 2, although the comparison may be challenging given differences in timing and definitions). There is however a considerable difference between AI buyers and developers, that represent respectively 9.83% and 3.11% of French firms with more than 10 employees.[11]

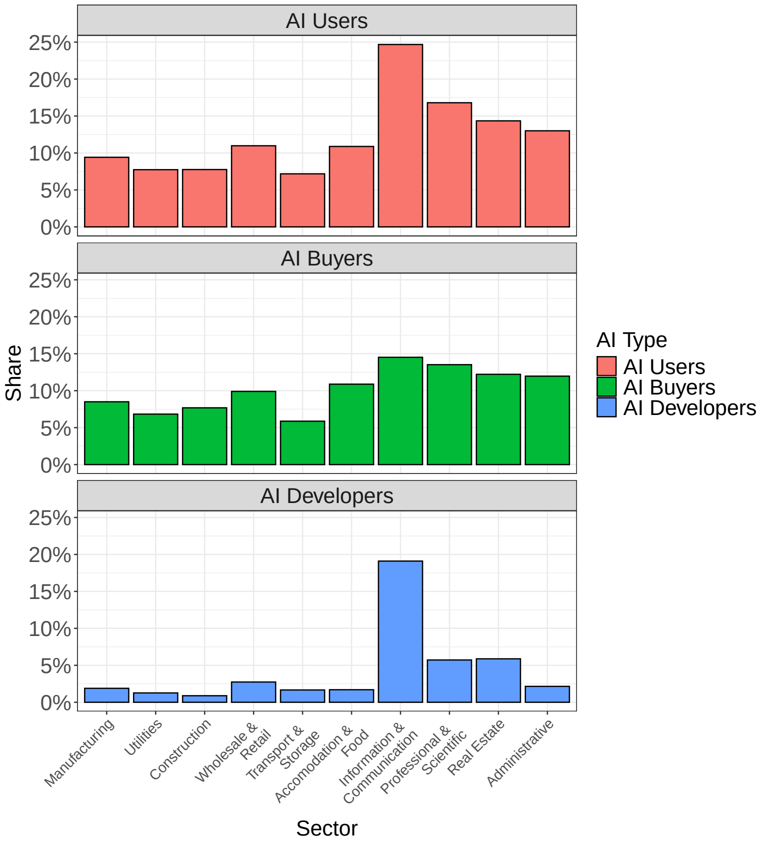

Figure 1: Shares of AI users, buyers and developers within sectors (see Table A.2).

When focusing on the shares of AI use within sectors (see Figure 1), a highly uneven distribution of AI actors across sectors emerges. The share of AI users is highest in the Information & Communication (24.67%) and in the Professional & Scientific services sectors (16.7%), whereas other sectors tend to lag behind in terms of diffusion. Again, distinguishing between AI buyers and AI developers reveals significant heterogeneity. Even though the highest shares of AI users, buyers and developers are always in the ICT sector, the gap between ICT and other sectors is remarkably high for developers. Conversely, the sectoral shares of firms buying AI technologies from third parties are more homogeneous, which support conjectures about the general-purpose nature of AI technologies. The sectoral heterogeneity in adoption shares also suggests that users developing their own AI or buying from third parties are different. The fact that the sectoral gap between AI developers is the highest in the ICT sector is consistent with the necessity of advanced ICT and technical skills to develop AI systems, that workers in other sectors may not have.

[11]Some firms are both AI buyers and AI developers, this is why the sum of shares of AI buyers and developers is higher than the overall share of AI users. This way of grouping AI users allows for both a simple setting in the regression analyses and to focus on groups of firms that are large enough to assess relevant information.

Notwithstanding the high shares of firms using, buying and developing AI systems in the ICT sector, the sectoral composition of AI users reported on Table 2 highlights that AI users in our sample do not operate only in sectors where the use of AI is more frequent. Approximately half of firms using AI, independently from how the technology was sourced, have their main activity in the Wholesale & Retail and Manufacturing sectors, the largest French sectors in terms of number of firms.

Sector	AI Users	AI Buyers	AI Developers
Manufacturing	23.1%	24.6%	20.9%
Construction	7.1%	7.9%	5.7%
Utilities	1%	1%	1.1%
Wholesale & Retail	25.8%	26.5%	23.8%
Transportation & Storage	5%	4.8%	3.7%
Accommodation & Food	4.4%	10.4%	1.8%
Information & Communication	13.2%	9.9%	22%
Professional & Scient.	12.1%	10.9%	14.6%
Real Estate	1.7%	1.9%	1.1%
Administrative	6.6%	7%	5.2%
Total	100	100	100

Table 2: Sectoral composition of AI users, buyers and developers across sectors (see Table A.2).

Figure 2: Shares of AI users, buyers and developers by size class.

Figure 3: Shares of AI users, buyers and developers by age class.

Focusing on the characteristics of AI firms, we report the shares of AI users, buyers and developers by size and age classes in Figures 2 and 3 respectively. The relation between size and AI use is positive, with the largest firms (with more than 249 employees) exhibiting about two times the share of AI use than the smallest ones. Also considering the high fixed costs possibly characterising AI-related IT investments (Brynjolfsson et al., 2021), the AI use-size relation may be related to digital intensity, as IT assets are crucial for AI systems. Indeed, the association between AI use and firm size seems stronger for AI developers than for AI buyers, suggesting again that drivers of AI use by firms may differ by the source of AI technologies. Differently from size, the relation between AI use and age is negative. This is noteworthy, because in general size and age are positively correlated. Younger firms are more likely AI users, possibly suggesting a role of new managerial capabilities. Furthermore, young firms often introduce more radical innovations, especially when new technological paradigms – such as the one brought by AI – emerge. Again, the relation between AI use and age seems stronger in the case of AI developers, in line with the summary statistics in Table 1.

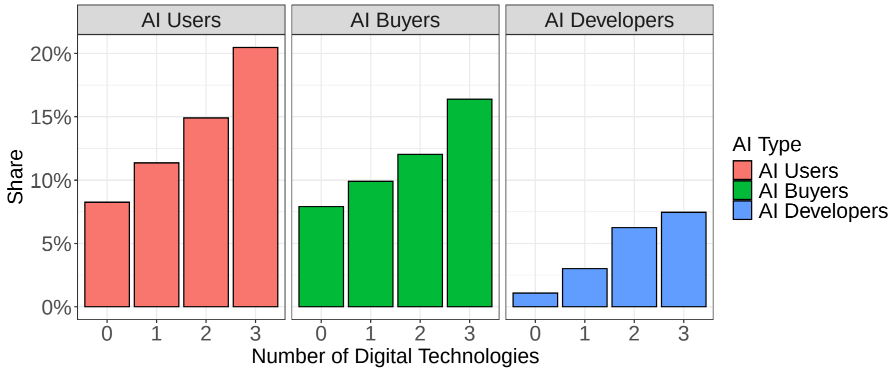

Figure 4: Shares of AI users, buyers and developers by number of digital technologies.

Finally, we focus on the shares of AI users by number of business digital technologies adopted in Figure 4. Firms using a higher number of digital technologies are also more likely to be users, buyers and developers of AI. This highlights that a proper digital architecture may be useful in order to exploit business data to use AI systems.

Although already informative, the descriptive unconditional patterns presented in Section 3 may be potentially influenced by a number of confounding factors that are not accounted for. This is why we turn to a regression analysis that further takes those into account. First, the relations discussed above may depend on the sectoral composition, such as the possible concentration of AI users in sectors in which average age is lower. Second, regressions may help further disentangle the role of other confounding factors affecting the use of AI, in particular considering intangibles, skills, digital infrastructure together and beyond the role of the firm characteristics explored above. We therefore estimate the following Probit regression:

$$P(\text{AI}_i^{\text{Type}} = 1|X) = \Phi(\alpha + \beta_1 \text{Log Size}_i + \beta_2 \text{Log Age}_i + \beta_2 \text{ICT Skills}_i + \beta_4 \text{Digitalisation}_i + \text{Ind.}_j + \text{Reg.}_r + \epsilon_i) \quad (1)$$

Where $\text{AI}_i^{\text{Type}}$ is a dummy variable that equals one if a firm uses, buys or develops AI (the first type focuses on all AI users, the second one on AI buyers, and the third one on AI developers), Log Size_i is the logarithm of the number of employees, Log Age_i is the logarithm of age, ICT Skills_i includes dummy variables indicating the presence of ICT specialists and the provision of ICT training for non-ICT employees, Digitalisation_i refers to a dummy variable that equals one if the firm uses a fast broadband connection and to a variable recording the number of other business digital technologies used (ERP, CRM, e-commerce), and Ind._j and Reg._r are vectors of controls including industry and geographic fixed effects respectively.[12]

Characteristics of AI Users, Buyers and Developers

	AI Users		AI Buyers		AI Developers	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Log Employment	0.0142***	-0.00165	0.00659*	-0.00359	0.0120***	0.00419**
	(0.00349)	(0.00415)	(0.00348)	(0.00409)	(0.00145)	(0.00175)
Log Age	-0.0121*	-0.0111*	-0.00652	-0.00622	-0.0124***	-0.0118***
	(0.00654)	(0.00646)	(0.00629)	(0.00624)	(0.00322)	(0.00324)
ICT Specialists		0.0356***		0.0118		0.0249***
		(0.0122)		(0.0122)		(0.00578)
ICT Training		0.0273**		0.0211**		0.00180
		(0.0110)		(0.0107)		(0.00492)
Fast Broadband		0.0266**		0.0225*		0.00820*
		(0.0119)		(0.0117)		(0.00480)
Number of Digital Technologies		0.0225***		0.0178***		0.0122***
		(0.00542)		(0.00525)		(0.00276)
Observations	8,268	8,268	8,268	8,268	8,268	8,268
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.0331	0.0480	0.0175	0.0260	0.153	0.191

Robust standard errors, clustered at the sampling strata level, in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Marginal effects from probit regressions employing the AI User (Models 1-2), AI Buyer (Models 3-4) and AI Developer (Models 5-6) dummy as dependent variable (see Equation 1).

We report the results of the estimation of Equation 1 in Table 3, which includes the estimated average marginal effects from the Probit regressions using dummies for AI users, buyers and developers as dependent variables in separate estimations. Reported numbers can therefore be interpreted as the average changes of AI use, buying or development probabilities due to a change in a covariate, ceteris paribus. The baseline regression model (Model 1 in Table 3) shows that in absence of other controls, larger firms have a higher probability of AI use, while older ones exhibit a lower likelihood of using AI, in line with the previously presented descriptive evidence. Distinguishing between AI buyers and developers confirms that, for both groups, being a large firm is associated with a higher probability of AI use. However, being a younger firm is related only to the probability of AI development, given the negative and significant coefficient of the logarithm of age. This may drive the overall result on the role of age evident for all AI users and suggest the relevance of new managerial capabilities for developing AI systems in house.

The link between firm size and AI use loses strength when accounting for other IT-related controls (fast broadband connection, other digital technologies, presence of ICT specialised workforce, training for non-ICT workers and the number of other digital technologies). This suggests that scale advantages in the adoption of AI technologies are at least partially driven by the joint presence of proper digital infrastructures, of a firm digital architecture and of ICT skills.

[12] Industry fixed effects are based on the OECD STAN A38 classification. More information can be found by opening this link. Geographic fixed effects correspond instead to French regions. Controlling for geographic fixed effects is relevant to capture differences in terms of regional development and other geographic factors possibly driving the diffusion of IT (e.g., see Andersen et al., 2012).

In particular, the coefficient of the logarithm of size becomes not significant when considering all AI users together and AI buyers, but its significance remains in the AI developer case – even though the coefficient's magnitude reduces to approximately one third of the coefficient reported. This hints at the presence of further drivers of AI development related to firm size – and therefore to fixed costs – with respect to the ones of AI buyers. These may concern for example the need of large amounts of data to develop and train large-scale AI systems or to costs related to acquiring computing power, as conjectured by Brynjolfsson and McAfee (2014).

Focusing on the coefficients of ICT-related controls suggests that these are crucial for AI use. Digital infrastructure is key to foster AI diffusion, as highlighted by the positive association between the presence of a fast broadband connection and the use of AI, which is significant throughout all regressions, in line with the evidence on the ICT-broadband nexus (DeStefano et al., 2018). The positive and significant coefficient of the number of digital technologies suggests that firms may leverage digitalised business information to adopt AI (see also McElheran et al., 2023).

In contrast, the positive relation between ICT specialists, ICT training and AI use depends on the type of AI users considered. In particular, the presence of ICT specialists is only linked to the use of AI when it is developed by the firm, whereas the training of non-ICT specialists is significant for AI buyers and not for AI developers. Albeit both suggest the importance of ICT skills for the diffusion of AI technologies (also see Babina et al., 2023), this result hints at the relevance of different skills for firms developing vis-à-vis buying AI systems. On the one hand, the development of AI requires a more in-depth knowledge of AI algorithms, as provided by ICT specialists. On the other hand, the use of ready-made AI systems may only require workers to know which inputs are needed to interact with AI systems in order to interpret their outputs, without necessarily a comprehensive understanding of the whole process behind them.

These results are robust to alternative modeling choices that further account explicitly for the relations between the decisions to make or buy AI systems. In this context, we estimate a Biprobit model, whose marginal effects can be found in Table A.1 of Appendix A. This model enables to control for the correlation between the decisions to buy and develop in-house AI systems by capturing omitted variables driving the joint decision to make and/or buy AI systems. Estimated marginal effects from the biprobit specification confirm the differences between AI buyers and AI developers, corroborating the Probit estimates and providing further support to the discussion above.

5 AI use and labour productivity

In this section we investigate the relation between AI use and labour productivity. We do so by estimating a series of regressions that employ the logarithm of productivity, measured by the ratio of the real value added to the number of employees, as dependent variable and AI use as main explanatory variable.

We organise the discussion in three sub-sections. In Section 5.1 we present the results of productivity regressions focusing on AI users, buyers and developers. We then explore in Section 5.2 the extent to which the AI-productivity nexus is heterogeneous across firms belonging to different sectors. Finally, we further address possible endogeneity issues in the AI-productivity link in Section 5.3 by estimating an endogenous treatment model. Further robustness checks are briefly discussed in Section 5.4 and placed in Appendix B.

Productivity regressions for AI Users, Buyers and Developers

	AI Users		AI Buyers & Developers	
	Model 1	Model 2	Model 3	Model 4
AI User	0.0464**	0.0310		
	(0.0210)	(0.0223)		
AI Buyer			0.0291	0.00572
			(0.0234)	(0.0201)
AI Developer			0.0936**	0.100***
			(0.0378)	(0.0381)
Log Employment	0.00633	-0.0242**	0.00542	-0.0247**
	(0.0145)	(0.0111)	(0.0145)	(0.0111)
Log Age	0.0661***	-0.0188**	0.0670***	-0.0180**
	(0.0115)	(0.00882)	(0.0116)	(0.00897)
Initial Log Productivity		0.522***		0.523***
		(0.0251)		(0.0250)
ICT Specialists		0.0656**		0.0628**
		(0.0291)		(0.0296)
ICT Training		0.0274*		0.0278*
		(0.0161)		(0.0160)
Number of Digital Technologies		0.0124		0.0118
		(0.00942)		(0.00941)
Fast Broadband		0.0905***		0.0898***
		(0.0190)		(0.0189)
Observations	8,268	8,268	8,268	8,268
Adj. R^2	0.277	0.497	0.277	0.498
Industry FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes

Robust standard errors, clustered at the sampling strata level, in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: OLS estimates of regressions using the logarithm of productivity in 2019 as dependent variable and AI users, buyers and developers as main explanatory variable.

5.1 Productivity regressions across different types of AI users

In this section we investigate the link between AI use and productivity by estimating the following regression models:

$$\begin{aligned} \text{Log-Productivity}_{i,2019} = &\alpha + \beta_{AI}\text{AI}_{i,2018} + \beta_2\text{Initial Log-Productivity}_i + \beta_3\text{Log Size}_{i,2018} + \\ &+ \beta_4\text{Log Age}_{i,2018} + \beta_5\text{ICT Skills}_{i,2018} + \beta_6\text{Digitalisation}_{i,2018} \\ &+ \text{Ind.}_j + \text{Reg.}_r + \epsilon_i \end{aligned} \tag{2}$$

where $\text{Log-Productivity}_{i,2019}$ is the logarithm of labour productivity (i.e., the ratio of real value added to the number of employees), $\text{AI}_{i,2018}$ is either the dummy variable indicating the use of AI by a firm or a vector of dummies including both AI buyers and AI developers, $\text{Initial Log-Productivity}_i$ is the productivity level either in 2011 or in the birth year of the firm (when later than 2011). As in equation 1, we include the logarithm of the number of employees ($\text{Log Size}_{i,2018}$) and of firm age ($\text{Log Age}_{i,2018}$), the dummy variables for the presence of ICT specialists and the provision of ICT training for non-ICT employees (ICT Skills$_{i,2018}$), the dummy for the use of fast broadband and the number of other digital technologies used by the firm (Digitalisation$_{i,2018}$), as well as industry (Ind.$_{j}$) and geographic (Reg.$_{r}$) fixed effects. We include initial productivity in the regression model because it allows to isolate the association between AI and productivity after accounting for firm productivity at a moment (2011) in which AI use was very unlikely.[13] The inclusion of the initial productivity term is also aimed at controlling for selection of most productive firms into AI use. This is relevant because the data do not contain information on the date of first use/adoption of AI.[14] Finally, the presence of ICT-related assets in the regression rules out potential under and over estimation biases related to productivity J-curves dynamics (see Brynjolfsson et al., 2021).[15]

We report the results of the estimation of Equation 2 in Table 4. We perform these regressions with and without initial productivity and IT-related controls other than industry, geographic fixed effects, size and age. On average, AI use is significantly linked to firm labour productivity, but the relation loses significance when the regression controls for IT-related variables and initial productivity, whose coefficients tend to be positive and significant. Next, we distinguish between users buying AI from external providers vis-à-vis those developing their own AI (see Models 3 and 4 of Table 4). Although the AI coefficients remain positive, AI use in the buyers' case is not significantly associated with productivity. Conversely, the relation between AI use and productivity is significant for developers, also when accounting for initial productivity and IT assets. The point estimate of the coefficient of AI developers is slightly higher and increases its significance in their presence, whereas the one of AI buyers is reduced to one fifth of the initial estimate.

In presence of other controls, the coefficients of the logarithm of size and age switch sign, becoming negative. Concerning size, our result hints at the relevance of intangibles, digital tools and investments for boosting firm productivity and at their greater availability in larger firms. The sign change of the size coefficient may thus be explained by lower productivity increases experienced by larger firms once the role of initial productivity and other IT assets are accounted for. Similarly, older firms tend to be more productive, but their productivity may thus grow less than younger firms when accounting for initial productivity and other IT assets.

These findings are novel to the literature and may contribute to explain why existing estimates of the AI-productivity relation are characterised by mixed results (see the discussion in Section 2). First, the absence of a significant AI-productivity link when accounting for IT controls and initial productivity supports the hypothesis of the self-selection of firms with higher productivity and digital capabilities into AI use. Second, the results highlight that the use of AI systems developed in house may already provide productivity returns. This is consistent with the idea that firms

[13]We chose to use 2011 as the year in which initial productivity is computed as this is a moment in which AI systems were likely not used by firms. In fact, several significant improvements in AI applications and technologies took place in 2012 (e.g., AlexNet neural network) and after that the use of deep learning and artificial neural networks started to outperform state-of-the-art non-AI related techniques in statistical analyses (see also Babina et al., 2024; Engberg et al., 2024). Accordingly, the boom in AI use by firms very likely started after 2011 in the US, and probably even later in other countries, such as France.

[14]When including the logarithm of initial productivity, Equation 2 captures the link between the use of AI and the change in the productivity of its users from the initial year considered (2011 or the birth year of firms). In that respect, the inclusion of the logarithm of age controls for changes across different periods of time considered (e.g., 2019-2011 for 2011 incumbents vis-á-vis 2019-2015 for a firm born in 2015).

[15]We did not include the logarithm of capital. This choice depends on the use of labour productivity as the dependent variable. Indeed, the inclusion of the logarithm of capital among controls would imply estimating a Cobb-Douglas production function with biased coefficients, and not the relation between AI and labour productivity. Instead, we provide in Section **??** of Appendix B a robustness check employing MFP as the dependent variable.

developing in-house their own AI were able to better adapt AI systems to their organisational structure, generating returns to its use.

5.2 Sectoral heterogeneity in the AI-productivity link

The findings presented in Section 5.1 have focused on the relation between AI and productivity in the overall sample, before and after accounting for initial productivity and the role of IT assets. Here we further explore the heterogeneity in the AI-productivity relation at the sector level. Indeed, the nexus between AI use and sector-specific organisational and productive structures may determine the applicability of AI technologies to firms and thus further explain the sources of the links between AI use and productivity. For instance, sectoral differences in the level of ICT intensity may determine the ability of firms to capture the productivity gains of AI.

Sectoral Productivity Regression for AI Users

	AI Users				
	Manufacturing	Wholesale & Retail	ICT	Other Services	Construction
AI User	0.0209	0.0713	0.142**	0.0207	0.0189
	(0.0510)	(0.0466)	(0.0660)	(0.0506)	(0.0345)
Adj. R^2	0.378	0.364	0.269	0.494	0.247
	AI Buyers & Developers				
	Manufacturing	Wholesale & Retail	ICT	Other Services	Construction
AI Buyer	0.000342	0.0257	-0.105	0.00139	0.0264
	(0.0589)	(0.0328)	(0.131)	(0.0354)	(0.0236)
AI Developer	0.125**	0.158**	0.240**	0.130	-0.0411
	(0.0620)	(0.0700)	(0.105)	(0.0850)	(0.0832)
Adj. R^2	0.378	0.365	0.277	0.495	0.247
Observations	2,145	2,092	603	2,413	892
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes

Robust standard errors, clustered at the sampling strata level, in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: OLS estimates of sector-specific regressions using the logarithm of productivity in 2019 as dependent variable and AI use as main explanatory variable. Estimated coefficients of size, age, ICT specialists, ICT training, Fast Broadband and Number of other technologies have not been reported, but can be found in Table A.3 in the appendix. Regressions for Utilities sector are not reported for the small sample size characterising this sector. The sectoral decomposition broadly corresponds to the one reported in Table A.2, with sectors beyond Wholesale & Retail and ICT aggregated into Other Services sector.

We estimate the main regression model (see Equation 2) on sectoral sub-samples broadly based on the decomposition presented in Table A.2.[16] The results are reported in Table 5. The relation between AI use and productivity is heterogeneous across sectors and type of users. A positive and significant coefficient for general AI users is only observed within the ICT sector. AI buyers do not exhibit significantly higher productivity, regardless of the sector considered. However, firms developing in-house AI show higher productivity not only in the ICT sector, but also in Wholesale & Retail and Manufacturing. First, firms in the Wholesale & Retail sector may have substantial amounts of data at their disposal, spanning across several domains (e.g., customers, inventories, prices). This abundant data availability represents an opportunity for these firms to effectively develop and deploy machine learning data analysis, possibly generating

[16]In several service sectors the number of firms and the one of AI users, buyers and developers is not large enough to produce meaningful results. We therefore aggregated service sectors different from ICT and Wholesale & Retail under the Other Service Sectors category.

productivity returns via data-driven predictions and decision making (in line with the discussion in Brynjolfsson et al., 2021). Second, our findings suggest that the higher productivity of firms developing AI is not confined to services and are in line with the idea that AI systems may enhance the productivity of manufacturers. AI applications in this sector can in fact lead to significant efficiency gains through both product and process innovations leveraging – for instance – AI-powered predictive maintenance, quality control, supply chain optimisation, and automation.

5.3 The AI-productivity relation beyond self-selection

In this section we develop a different specification aimed at further assessing whether the use of AI is linked with productivity beyond selection, tackling in a different way possible sources of estimation bias due to the selection of firms into use. We have indeed already showed that AI use is positively associated with productivity in the case of firms developing their own AI systems, even after controlling for a considerable set of factors that may affect such result. However, OLS estimates may still be subject to a degree of bias if there are other unobserved sources that are related to the higher productivity of AI developers. In this section, we use a different estimation technique that allows modeling explicitly the self-selection of firms into AI use, in order to further test the robustness of our results.

Following Hamilton and Nickerson (2003) and Clougherty et al. (2016), we rely on an endogenous treatment regression model (ET from now on), a latent variable approach that belongs to the family of selection models (see Heckman, 1976, 1978; Maddala, 1983; Wooldridge, 2015). This choice is consistent with two facts. First, we want to correct for self-selection of firms into AI use (see also Shaver, 1998; King and Tucci, 2002; Campa and Kedia, 2002; Dosi et al., 2022). Second, the relevant endogenous variable – the use of AI – is a dummy variable. The ET method corrects for the selection of firms into AI use by estimating an outcome (similar to Equation 2) and a selection (alike Equation 1) equation, and has the advantage of using a Probit model in the selection equation. Differently, the first stage of the 2SLS method is based on a linear probability model, whose coefficients can be interpreted as probabilities at the cost of generating predicted values outside of the unity range of the probability space and assuming a linear approximation to an underlying non-linear model (Clougherty et al., 2016). Furthermore, estimates include the coefficient ρ – the correlation between the errors from the outcome and selection equations – which tests the presence of selection into treatment and thus the endogeneity of the treatment variable considered.[17]

The ET model reads as follows:

$$\begin{aligned} \text{Log-Productivity}_{i,2019} =& \alpha + \beta_1 \text{AI}_{i,2018} + \beta_2 \text{Initial Log-Productivity}_i + \beta_3 \text{Log Size}_{i,2018} + \beta_4 \text{Log Age}_{i,2018} + \\ &+ \beta_5 \text{ICT Skills}_{i,2018} + \beta_6 \text{Digitalisation}_{i,2018} + \text{Ind.}_{\cdot j} + \text{Reg.}_{\cdot r} + \epsilon_i \\ \text{AI}^{Type}_{i,2018} =& \begin{cases} 1, & \text{if } \beta_Z \mathbf{Z}_i + \omega_i > 0 \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{3}$$

Where the two equations represent the outcome equation (i.e., the productivity regression described by Equation 2, using either the dummy for AI users or the two dummies for buyers and developers as main explanatory variables) and

[17]The sign of ρ accounts for the relation between unobservables affecting outcome variables with unobservables affecting selection.

the selection equation (i.e., using the dummy for one between AI users, buyers or developers – $\text{AI}_{i,2018}^{Type}$ – as dependent variable, see Equation 1), respectively. $\mathbf{Z}_i$ includes the logarithm of the initial productivity, size, age, ICT skills, proxies of digitalisation, industry, geographic fixed effects and an additional variable W_i, which is not included among covariates in the outcome equation. The selection equation employs $\text{AI}_{i,2018}^{Type}$, which is alternatively a dummy that equals one for AI users, buyers and developer. Therefore, we will account for the self-selection of firms into AI in three different specifications, respectively aimed at controlling for the self-selection of firms into the general use of AI, the purchase of AI solutions, or its development in-house.

We estimate the model via full information maximum likelihood (FIML henceforth), which simultaneously estimates the selection and the outcome equations by assuming joint normality of errors (ϵ_i, ω_i).

Even though not necessary thanks to the joint normality assumption, the ET model provides estimates robust to specification errors when at least one additional variable W_i is included in the selection equation, but exluded from the outcome equation. This variable needs to comply with two conditions (Puhani, 2000): it must strongly predict the endogenous dummy variable (i.e., be relevant), and has to be exogenous (i.e., satisfy the exclusion restriction), given other controls.[18] We employ the Leave-One-Out (LOO henceforth) shares of ICT engineers at the aggregate industry (see Table A.2) and department level in 2011 computed using French worker-level data (i.e., DADS). Such variable captures the local supply of engineers with industry-specific and advanced IT capabilities. This may vary depending on the local supply of ICT engineers, the digital intensity of local industries and the nature of the production processes considered, which determines the type of AI applications best suited for a firm. We define as ICT engineers the employees classified in the 2003 French PCS classification under the 4-digit classes 388a, 388b, 388c, 388d and 388e.[19] The LOO shares are computed for each firm by excluding its employees from the computation of the variables.

Concerning the relevance of the exclusion restriction, the results in Section 4 show that AI developers are significantly and positively linked to specialised ICT human capital, differently from AI buyers (see Table 3). In this respect, the decision to buy or develop AI systems in-house may depend on the advanced ICT human capital pre-existing to the diffusion of AI, as measured by our exclusion restriction. Conversely, the absence of specialised ICT human capital may induce firms to buy, rather than develop, AI systems.

Furthermore, our specification corrects for three potential sources of endogeneity possibly affecting the share of ICT engineers in 2011. First, we compute the share of ICT engineers in 2011, when AI adoption was very unlikely in France. Therefore, the use of AI by a firm may unlikely have affected French firms and workers in 2011. Second, we rule out additional endogeneity sources by excluding the employees of a firm from the computation of its LOO shares. Third, we include the initial productivity, size and age of firms and the IT assets from previous regressions in the ET specification. The presence of these controls accounts for possible confounding factors in the relation between the LOO share of ICT engineers and productivity at the firm level. The initial productivity of firms accounts for possible relations

[18] As highlighted by Puhani (2000), in practice these are the same conditions that are required by instrumental variables in the 2SLS procedure.

[19] This classification focuses on ICT intensive occupations and is nested in the techies definition used in Harrigan et al. (2021), which includes all occupations in the 2-digit classes 38 (Technical managers and engineers) and 47 (Technicians) of the 2003 French PCS classification. The PCS codes listed above include for instance computer engineers, programmers, developers and database administrators. Further details and information on the PCS classification can be found here.

between the LOO shares of ICT engineers in 2011 and firm productivity in 2019. Furthermore, the joint presence of initial productivity, size and age of the firm also controls for the characteristics of firms self-selecting into AI use before adoption was possible or in the very first year of use. The presence of IT-related controls addresses the potential impact of the LOO share of ICT engineers in 2011 in terms of other ICTs adoption on productivity, and controls therefore for possible ICT-productivity relations beyond AI. Also, the presence of size in the ET specification helps further reduce possible sources of endogeneity to the extent to which ICT investments are more likely in larger firms (see e.g. McElheran et al., 2023; Zolas et al., 2020; Cirillo et al., 2023). In order to check whether the share of ICT engineers in 2011 is related to productivity, we use it as the main explanatory variable in a regression using the productivity as the dependent variable (see Table A.4). The share of ICT engineers in 2011 is not significant, suggesting that possible concerns regarding a significant relation between producitivity and the excluded variable – e.g., stemming from the higher productivity of firms in sectors-geographic regions with higher ICT intensity – are not supported by empirical estimates.

The role of AI for productivity beyond self-selection

	Model 1 - AI Users		Model 2 - AI Buyers		Model 3 - AI Developers	
	Outcome	Selection	Outcome	Selection	Outcome	Selection
AI User	0.066 (0.041)					
AI Buyer			0.007 (0.056)		-0.001 (0.020)	1.302*** (0.091)
AI Developer			0.101** (0.046)	1.475*** (0.123)	0.155** (0.066)	
2011 Share of ICT Engineers – Department & Aggr. Industry		0.006 (0.025)		-0.113*** (0.035)		0.112*** (0.023)
Initial Log Productivity	0.521*** (0.025)	-0.020 (0.047)	0.522*** (0.025)	0.050 (0.046)	0.522*** (0.025)	-0.177* (0.107)
Log Employment	-0.024** (0.011)	-0.008 (0.025)	-0.025** (0.011)	-0.037 (0.030)	-0.025** (0.011)	0.087*** (0.033)
Log Age	-0.018** (0.009)	-0.057 (0.039)	-0.018* (0.009)	-0.013 (0.053)	-0.017* (0.009)	-0.174** (0.077)
ICT Specialists	0.063** (0.028)	0.199*** (0.058)	0.061** (0.029)	-0.014 (0.081)	0.059** (0.029)	0.498*** (0.103)
ICT Training	0.027* (0.016)	0.152*** (0.057)	0.028* (0.016)	0.133** (0.054)	0.028* (0.016)	0.024 (0.081)
Fast Broadband	0.090*** (0.019)	0.153* (0.082)	0.090*** (0.019)	0.128 (0.079)	0.089*** (0.019)	0.069 (0.114)
Number of Digital Technologies	0.010 (0.009)	0.125*** (0.032)	0.011 (0.009)	0.077** (0.032)	0.010 (0.009)	0.200*** (0.047)
Observations	8,238	8,238	8,238	8,238	8,238	8,238
ρ	-0.0438	-0.0438	1.03e-05	1.03e-05	-0.0703	-0.0703
P-value$_\rho$	0.261	0.261	1	1	0.233	0.233
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors, clustered at the sampling strata level, in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Productivity regressions in endogenous treatment (ET) models. Rho is the correlation between errors from the first and the second stage, and p-value indicates the significance of Rho.

We report the estimation results of Equations 3 in Table 6. In Model 1, 2 and 3 the selection equation is estimated employing the dummies for AI users, buyers and developers respectively. Overall, the estimated results of the selection equations are in line with ones reported in Table 3 of Section 4. IT assets are significantly related to AI use, AI developers are larger and younger and significantly related to specialised ICT human capital, differently from AI buyers.

Furthermore, the share of ICT engineers is positively and significantly associated with AI development, suggesting that it is strongly associated with the endogenous treatment. Conversely, AI buyers are significantly but negatively related to the excluded variable. The local presence of ICT engineers specialised in the main sector of activity of the firm is necessary for firms to develop AI systems. Accordingly, in its absence, firms willing to adopt AI must purchase them from other firms. Finally, the LOO share is not significant for AI users in general. This is likely due to the opposite direction of the link with AI buyers and developers, that are both included in the general dummy for AI use.[20] Concerning the outcome equation, results are very similar to ones presented in previous sections, also in terms of magnitude (see Table 4 of Section 5.1), suggesting the existence of a positive relation between the use AI systems developed in house and firm productivity beyond selection. Conversely, such relation is not significant and very close to zero for AI buyers.

Finally, the error correlation term ρ is negative in the specification focused on AI developers. It is instead positive, but very close to 0 in the buyers case. In both cases, the term ρ is not significantly different from 0. This indicates that IT assets and initial productivity are able to rule out self-selection into AI, suggesting that the estimation reported in previous sections are likely robust to these issues.

5.4 Additional results and robustness checks

In previous sections we reported several econometric results supporting the existence of productivity returns from the development of AI technologies, which are absent for other users, also beyond the self-selection of firms into AI use and outside of the ICT sector. We further extend the analysis presented above by estimating three additional sets of models, whose results are reported in Appendix B.

First, we extend the analysis exploring the link between AI use and productivity employing specifications focused on labour productivity growth rates between 2019 and 2018, 2016, 2014 and 2011, corresponding to 2-periods firm fixed effects regressions. Second, we further check the robustness of our key findings by changing the productivity proxy used as dependent variable – moving from labour to multi-factor productivity –, the sample – excluding ICT business services (NACE 62-63) –, and the controls – using CRM, ERP and E-commerce separately in place of the number of digital technologies. Third, we employ a standard nearest-neighbor matching estimator to minimise the Mahalanobis distance between firms that use, purchase, or develop AI and those that do not. This approach allows us to estimate the impact of AI on productivity without relying on an IV strategy. However, this comes at the cost of information loss and limits the representativeness of the analysis for France.

These additional analyses and robustness checks confirm the absence of a significant AI-productivity link for AI users and buyers, and its existence for developers, once accounting for relevant IT assets and factors related to selection.

6 Concluding remarks

In this study we focus on the characteristics of French firms using AI and explore the association between AI use and firm productivity, distinguishing AI buyers and developers. We use detailed survey microdata collected by the French

[20] As a consequence, Model 1 will be weakly identified.

national statistical office, which provide detailed information on technology use in 2018 and are matched with firms' balance sheets containing additional firm-level information over a longer time period, between 2011 and 2019.

Our findings provide a novel outlook on the diffusion of AI among French firms and suggest that AI users are not all alike. First, we present several facts regarding the diffusion of AI technologies in France and the characteristics of different types of AI users, notably distinguishing AI buyers from developers. AI buyers tend to be larger than other firms, but ICT-related factors explain this relation. Conversely, AI developers remain larger even once these confounding factors are taken into account, and are also younger. Second, we uncover a critical role of intangibles and IT assets for AI adoption. Measures of firm digitalisation and human capital, such as the use of business digital technologies, digital infrastructure (i.e., fast-broadband connection), and the presence of ICT skills, are positively and significantly related to the use of AI. The distinction between AI buyers and developers also reveals that the latter group hinges on more specialised ICT human capital. Third, our work explores the link between AI use and firm productivity. On average, AI use tends to be positively linked with firm productivity. However, when considering all AI users together or firms buying AI, this link is driven by the self-selection into AI use of firms with higher productivity and that already leverage a higher intangible or IT intensity. In contrast, when focusing on AI developers, the link between AI use and productivity remains positive and significant beyond selection, hinting at a positive impact of AI on their productivity.

Firms developing in house their own AI systems may benefit from higher flexibility and control over the integration process of AI in their activities, and thus experiment positive and significant productivity returns, differently from other firms. However, the opportunities to leverage the potential of AI technologies may not be equally distributed among firms. Indeed, larger, younger and more digital firms turn out to be those more likely to develop AI, highlighting the relevant role of advanced IT human capital, of the capacity to incur high fixed costs, and of new managerial capabilities for developing AI systems in-house. This suggests that in the future productivity gains from the diffusion of AI technologies may be captured by a handful of firms, possibly widening productivity gaps between leaders and other firms (also see Corrado et al., 2021; Andrews et al., 2016). Policy makers can play a key role to foster an inclusive digital transformation, enabling AI use and its returns to be more widespread across firms and sectors (see also Calvino and Criscuolo, 2022; Calvino and Fontanelli, 2023).

Further research may extend the scope of the current analysis in different directions. First, additional work may focus on further exploring the role of IT assets for different groups of AI adopters. In particular, linking French matched employer-employee (DADS) data with ICT surveys and balance sheet data at the firm level could allow exploring in further detail the role of human capital for AI adoption and its productivity returns. Second, future analysis may also explore more directly the role of management and organisational layers for AI adoption and its returns, to further explore the links between AI use and organisational capabilities. Third, future work could focus on other outcomes beyond firm productivity, such as employment and wages, to further assess whether AI have an impact on firms' employment. Finally, future work may focus on the role of AI for innovation, trying to disentangle whether and which types of AI activities are linked to product and process innovations.

References

Acemoglu, D., G. Anderson, D. Beede, C. Buffington, E. Childress, E. Dinlersoz, L. Foster, N. Goldschlag, J. Haltiwanger, Z. Kroff, P. Restrepo, and N. Zolas (2023). Advanced technology adoption: Selection or causal effects? *AEA Papers and Proceedings 113*, 210–14.

Acemoglu, D., G. W. Anderson, D. N. Beede, C. Buffington, E. E. Childress, E. Dinlersoz, L. S. Foster, N. Goldschlag, J. C. Haltiwanger, Z. Kroff, P. Restrepo, and N. Zolas (2022). Automation and the workforce: A firm-level view from the 2019 annual business survey. In *Technology, Productivity, and Economic Growth*, NBER Chapters.

Acemoglu, D., D. Autor, J. Hazell, and P. Restrepo (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics 40*(S1), 293–340.

Acemoglu, D. and P. Restrepo (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy 128*(6), 2188 – 2244.

Aghion, P., C. Antonin, S. Bunel, and X. Jaravel (2020). What are the labor and product market effects of automation?: New evidence from France. Sciences po publications, Sciences Po.

Agrawal, A., J. Gans, and A. Goldfarb (2022). Power and prediction: the disruptive economics of artificial intelligence. *Harvard Business Review Press.*

Albanesi, S., A. Dias da Silva, J. F. Jimeno, A. Lamo, and A. Wabitsch (2023, June). New technologies and jobs in europe. Working Paper 31357, National Bureau of Economic Research.

Alderucci, D., L. Branstetter, E. Hovy, A. Runge, and N. Zolas (2021). Quantifying the impact of AI on productivity and labor demand: Evidence from U.S. census microdata. *Mimeo.*

Alekseeva, L., J. Azar, M. Giné, S. Samila, and B. Taska (2021). The demand for AI skills in the labor market. *Labour Economics 71*(C).

Alekseeva, L., M. Gine, S. Samila, and B. Taska (2020). Ai adoption and firm performance: Management versus it. *Mimeo.*

Andersen, T. B., J. Bentzen, C.-J. Dalgaard, and P. Selaya (2012). Lightning, IT Diffusion, and Economic Growth Across U.S. States. *The Review of Economics and Statistics 94*(4), 903–924.

Andrews, D., C. Criscuolo, and P. N. Gal (2016). The Best versus the Rest: The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy. OECD Productivity Working Papers No. 2016/05, OECD Publishing.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Number 8769 in Economics Books. Princeton University Press.

Babina, T., A. Fedyk, A. He, and J. Hodson (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics 151*, 103745.

Babina, T., A. Fedyk, A. X. He, and J. Hodson (2023). Firm investments in artificial intelligence technologies and changes in workforce composition. Working Paper 31325, National Bureau of Economic Research.

Besiroglu, T., N. Emery-Xu, and N. Thompson (2022). Economic impacts of AI-augmented R&D. arXiv, 2212.08198.

Bessen, J., S. M. Impink, L. Reichensperger, and R. Seamans (2022). The role of data for AI startup growth. *Research Policy 51*(5).

Bianchini, S., M. Müller, and P. Pelletier (2022). Artificial intelligence in science: An emerging general method of invention. *Research Policy 51*(10), 104604.

Borgonovi, F., F. Calvino, C. Criscuolo, L. Samek, H. Seitz, J. Nania, J. Nitschke, and L. O'Kane (2023). Emerging trends in ai skill demand across 14 oecd countries. OECD Artificial Intelligence Papers No. 2, OECD Publishing, Paris.

Brynjolfsson, E. and L. M. Hitt (2003). Computing Productivity: Firm-Level Evidence. *The Review of Economics and Statistics 85*(4), 793–808.

Brynjolfsson, E., W. Jin, and K. McElheran (2021). The power of prediction: predictive analytics, workplace complements, and business performance. *Business Economics 56*(4), 217–239.

Brynjolfsson, E., D. Li, and L. Raymond (2023). Generative AI at work. Working Paper 31161, National Bureau of Economic Research.

Brynjolfsson, E. and A. McAfee (2014). *The second machine age: work, progress, and prosperity in a time of brilliant technologies.* New York : W.W. Norton & Company.

Brynjolfsson, E., D. Rock, and C. Syverson (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics 13*(1), 333–372.

Calvino, F. and C. Criscuolo (2022). Gone digital: Technology diffusion in the digital era. In Z. Qureshi and C. Woo (Eds.), *Shifting Paradigms: Growth, Finance, Jobs, and Inequality in the Digital Economy.* Brookings Institution Press.

Calvino, F. and L. Fontanelli (2023). 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.

Calvino, F., C. Morris, L. Samek, and M. Squicciarini (2022). Identifying and characterising AI adopters: a novel approach based on big data. OECD Science, Technology and Industry Working Papers No. 2022/06, OECD Publishing, Paris.

Campa, J. M. and S. Kedia (2002). Explaining the diversification discount. *The Journal of Finance 57*(4), 1731–1762.

Cette, G., S. Nevoux, and L. Py (2022). The impact of ICTs and digitalization on productivity and labor share: evidence from French firms. *Economics of Innovation and New Technology 31*(8), 669–692.

Cho, J., T. DeStefano, H. Kim, I. Kim, and J. H. Paik (2022). What's driving the diffusion of next-generation digital technologies? *Technovation* (119).

Cirillo, V., L. Fanti, A. Mina, and A. Ricci (2023). The adoption of digital technologies: Investment, skills, work organisation. *Structural Change and Economic Dynamics 66*, 89–105.

Clougherty, J. A., T. Duso, and J. Muck (2016). Correcting for self-selection based endogeneity in management research: Review, recommendations and simulations. *Organizational Research Methods 19*(2), 286–347.

Cockburn, I. M., R. Henderson, and S. Stern (2018). The impact of artificial intelligence on innovation: An exploratory analysis. In *The Economics of Artificial Intelligence: An Agenda*, NBER Chapters, pp. 115–146.

Commander, S., R. Harrison, and N. Menezes-Filho (2011). ICT and Productivity in Developing Countries: New Firm-Level Evidence from Brazil and India. *The Review of Economics and Statistics 93*(2), 528–541.

Conti, R., M. G. de Matos, and G. Valentini (2024). Big data analytics, firm size, and performance. *Strategy Science 9*(2), 135–151.

Corrado, C., C. Criscuolo, J. Haskel, A. Himbert, and C. Jona-Lasinio (2021). New evidence on intangibles, diffusion and productivity. OECD Science, Technology and Industry Working Papers No. 2021/10, OECD Publishing, Paris.

Czarnitzki, D., G. P. Fernández, and C. Rammer (2022). Artificial intelligence and firm-level productivity. ZEW Discussion Papers 22-005, ZEW - Leibniz Centre for European Economic Research.

Dahlke, J., M. Beck, J. Kinne, D. Lenz, R. Dehghan, M. Wörter, and B. Ebersberger (2024). Epidemic effects in the diffusion of emerging digital technologies: evidence from artificial intelligence adoption. *Research Policy 53*(2), 104917.

Damioli, G., V. V. Roy, and D. Vertesy (2021). The impact of artificial intelligence on labor productivity. *Eurasian Business Review 11*(1), 1–25.

Damioli, G., V. V. Roy, D. Vertesy, and M. Vivarelli (2023). AI technologies and employment: micro evidence from the supply side. *Applied Economics Letters 30*(6), 816–821.

Dell'Acqua, F., B. Kogut, and P. Perkowski (2023). Super Mario Meets AI: Experimental Effects of Automation and Skills on Team Performance and Coordination. *The Review of Economics and Statistics*, 1–47.

Dell'Acqua, F., E. McFowland, E. R. Mollick, H. Lifshitz-Assaf, K. Kellogg, S. Rajendran, L. Krayer, F. Candelon, and K. R. Lakhani (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. Harvard Business School Technology & Operations Mgt. Unit Working Paper 24-013.

Dernis, H., F. Calvino, L. Moussiegt, D. Nawa, L. Samek, and M. Squicciarini (2023). Identifying artificial intelligence actors using online data. OECD Science, Technology and Industry Working Papers No. 2023/01, OECD Publishing, Paris.

DeStefano, T., R. Kneller, and J. Timmis (2018). Broadband infrastructure, ICT use and firm performance: Evidence for UK firms. *Journal of Economic Behavior & Organization 155*(C), 110–139.

DeStefano, T., R. Kneller, and J. Timmis (2023). Cloud Computing and Firm Growth. *The Review of Economics and Statistics*, 1–47.

Dibiaggio, L., L. Nesta, and M. Keita (2022). *Artificial Intelligence - Technologies and Key Players.* SKEMA Business School.

Domini, G., M. Grazzi, D. Moschella, and T. Treibich (2021). Threats and opportunities in the digital era: Automation spikes and employment dynamics. *Research Policy 50*(7).

Domini, G., M. Grazzi, D. Moschella, and T. Treibich (2022). For whom the bell tolls: The firm-level effects of automation on wage and gender inequality. *Research Policy 51*(7).

Dosi, G., N. Mathew, and E. Pugliese (2022). What a firm produces matters: Processes of diversification, coherence and performances of indian manufacturing firms. *Research Policy 51*(8), 104152. Special Issue on Economic Complexity.

Eloundou, T., S. Manning, P. Mishkin, and D. Rock (2023). GPTs are GPTs: An early look at the labor market impact potential of large language models. *arXiv.org* (2303.10130).

Engberg, E., H. Görg, M. Lodefalk, F. Javed, M. Längkvist, N. P. Monteiro, H. Kyvik Nordås, S. Schroeder, and A. Tang (2024, January). AI Unboxed and Jobs: A Novel Measure and Firm-Level Evidence from Three Countries. IZA Discussion Papers 16717, Institute of Labor Economics (IZA).

Fontanelli, L., F. Calvino, C. Criscuolo, L. Nesta, and E. Verdolini (2024). The role of human capital for AI adoption: evidence from French firms. *Mimeo.*

Goos, M. and M. Savona (2024). The governance of artificial intelligence: Harnessing opportunities and mitigating challenges. *Research Policy 53*(3), 104928.

Guarascio, D., J. Reljic, and R. Stollinger (2023, October). Artificial Intelligence and Employment: A Look into the Crystal Ball. LEM Papers Series 2023/34, Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy.

Hamilton, B. H. and J. A. Nickerson (2003). Correcting for endogeneity in strategic management research. *Strategic Organization 1*(1), 51–78.

Harrigan, J., A. Reshef, and F. Toubal (2021). The march of the techies: Job polarization within and between firms. *Research Policy 50*(7).

Heckman, 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*, pp. 475–492. National Bureau of Economic Research.

Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica 46*(4), 931–959.

Hoffreumon, C., C. Forman, and N. van Zeebroeck (2024, March). Make or buy your artificial intelligence? Complementarities in technology sourcing. *Journal of Economics & Management Strategy 33*(2), 452–479.

Igna, I. and F. Venturini (2023). The determinants of AI innovation across European firms. *Research Policy 52*(2).

Ikeuchi, K., T. Inui, and G. Kim Young (2023). AI Adoption and Productivity of Japanese Firms: Spillover and innovation effects (Japanese). Discussion Papers (Japanese) 23034, Research Institute of Economy, Trade and Industry (RIETI).

Jin, W. and K. McElheran (2018). Economies before scale: Survival and performance of young plants in the age of cloud computing. Rotman School of Management Working Paper 3112901.

King, A. A. and C. 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).

Kreitmeir, D. and P. A. Raschky (2023). The unintended consequences of censoring digital technology - evidence from Italy's ChatGPT ban. *Center for Open Science SocArXiv* (v3cgs).

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

McElheran, K., J. F. Li, E. Brynjolfsson, Z. Kroff, E. Dinlersoz, L. S. Foster, and N. Zolas (2023). AI adoption in America: Who, what, and where. *Forthcoming in Journal of Economics & Management*.

Noy, S. and W. Zhang (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science 381*(6654), 187–192.

Peng, S., E. Kalliamvakou, P. Cihon, and M. Demirer (2023). The impact of AI on developer productivity: Evidence from GitHub copilot. *arXiv* (2302.06590).

Puhani, P. (2000). The Heckman Correction for sample selection and its critique. *Journal of Economic Surveys 14*(1), 53–68.

Rammer, C., G. P. Fernández, and D. Czarnitzki (2022). Artificial intelligence and industrial innovation: Evidence from German firm-level data. *Research Policy 51*(7).

Santarelli, E., J. Staccioli, and M. Vivarelli (2022). Automation and related technologies: a mapping of the new knowledge base. *Journal of Technology Transfer 71*(C).

Shaver, J. M. (1998). Accounting for endogeneity when assessing strategy performance: Does entry mode choice affect FDI survival? *Management Science 44*(4), 571–585.

Squicciarini, M. and H. Nachtigall (2021). Demand for AI skills in jobs: Evidence from online job postings. OECD Science, Technology and Industry Working Papers No. 2021/03, OECD Publishing, Paris.

Svanberg, M., W. Li, M. Fleming, B. Goehring, and N. Thompson (2024). Beyond AI Exposure: Which Tasks are Cost-Effective to Automate with Computer Vision? SSRN 4700751.

Tricot, R. (2021). Venture capital investments in artificial intelligence. OECD Digital Economy Working Paper No. 2021/319.

Venturini, F., L. d. S. Marioni, and A. Rincon-Aznar (2024). Productivity performance, distance to frontier and AI innovation: Firm-level evidence from Europe. *Mimeo.*

Wooldridge, J. (2015). Control function methods in applied econometrics. *Journal of Human Resources 50*(2), 420–445.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters 104*(3), 112–114.

Yang, C.-H. (2022). How artificial intelligence technology affects productivity and employment: Firm-level evidence from taiwan. *Research Policy 51*(6), 104536.

Zolas, N., Z. Kroff, E. Brynjolfsson, K. McElheran, D. N. Beede, C. Buffington, N. Goldschlag, L. Foster, and E. Dinlersoz (2020). Advanced technologies adoption and use by U.S. firms: Evidence from the annual business survey. Working Paper 28290, National Bureau of Economic Research.

Appendix

A Other tables

Estimated marginal effects from Biprobit regressions

	Model 1		Model 2	
	AI buyers	AI developers	AI buyers	AI developers
Log Employment	0.00684**	0.0116***	-0.00363	0.00346**
	(0.00344)	(0.00138)	(0.00402)	(0.00173)
Log Age	-0.00633	-0.0123***	-0.00612	-0.0109***
	(0.00629)	(0.00299)	(0.00621)	(0.00299)
ICT Specialists			0.0132	0.0259***
			(0.0121)	(0.00555)
ICT Training			0.0223**	0.00387
			(0.0106)	(0.00486)
Fast Broadband			0.0236**	0.00715
			(0.0116)	(0.00478)
Number of Digital Technologies			0.0177***	0.0118***
			(0.00523)	(0.00258)
Observations	8,268	8,268	8,268	8,268
Industry FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes

Table A.1: Estimated marginal effects from Biprobit regressions jointly estimating the probability to use AI by buying it from third parties or developing it in-house.

Sectoral disaggregation

Group of sectors	Sector 2–digit code (ISIC rev.4)
Manufacturing	10-33
Utilities	35-39
Construction	41-43
Wholesale & Retail	45-47
Transport & Storage	49-53
Accommodation & Food	55-56
Information & Communication	58-63, 951
Real Estate	68
Professional & Scientific Activities	69-75
Administrative	77-82

Table A.2: The sectoral disaggregation used for computing shares in Figure 1. This aggregation is aimed at capturing common features between sectors and encompasses ten different macro-sectors.

Complete results of sectoral regressions from Table 5

	AI User					AI Buyer & Developer				
	Manuf.	Whol. & Ret.	ICT	Other Serv.	Constr.	Manuf.	Whol. & Ret.	ICT	Other Serv.	Constr.
AI User	0.0209	0.0713	0.142**	0.0207	0.0189					
	(0.0510)	(0.0466)	(0.0660)	(0.0506)	(0.0345)					
AI Buyer						0.000342	0.0257	-0.105	0.00139	0.0264
						(0.0589)	(0.0328)	(0.131)	(0.0354)	(0.0236)
AI Developer						0.125**	0.158**	0.240**	0.130	-0.0411
						(0.0620)	(0.0700)	(0.105)	(0.0850)	(0.0832)
Log Employment	-0.0115	-0.0152	-0.0218	-0.0637***	0.0175	-0.0129	-0.0154	-0.0220	-0.0645***	0.0180
	(0.0234)	(0.0267)	(0.0265)	(0.0181)	(0.0103)	(0.0233)	(0.0265)	(0.0270)	(0.0183)	(0.0106)
Log Age	-0.00840	-0.0272	-0.105***	-0.00646	-0.0291***	-0.00787	-0.0274	-0.109***	-0.00398	-0.0293***
	(0.0198)	(0.0197)	(0.0318)	(0.0213)	(0.00364)	(0.0198)	(0.0195)	(0.0317)	(0.0214)	(0.00353)
ICT Specialists	0.0751*	0.0424	0.0731	0.160***	-0.127	0.0731*	0.0399	0.0713*	0.155***	-0.124
	(0.0435)	(0.0305)	(0.0448)	(0.0557)	(0.0704)	(0.0433)	(0.0313)	(0.0412)	(0.0565)	(0.0744)
ICT Training	0.0665**	-0.0405*	0.0786	0.0680**	0.0726**	0.0664**	-0.0391*	0.0776*	0.0671***	0.0728**
	(0.0288)	(0.0213)	(0.0499)	(0.0259)	(0.0270)	(0.0288)	(0.0211)	(0.0446)	(0.0248)	(0.0262)
Fast Broadband	0.0750*	0.134***	-0.0556	0.162***	0.00130	0.0725*	0.132***	-0.0545	0.162***	0.000899
	(0.0404)	(0.0480)	(0.0540)	(0.0327)	(0.0138)	(0.0402)	(0.0472)	(0.0450)	(0.0326)	(0.0133)
Number of Digital Technologies	0.0564***	-0.0137	0.0114	0.00231	0.0173**	0.0562***	-0.0147	0.0115	0.00112	0.0174**
	(0.0163)	(0.0160)	(0.0384)	(0.0144)	(0.00557)	(0.0164)	(0.0155)	(0.0362)	(0.0146)	(0.00587)
Initial Log Productivity	0.598***	0.525***	0.425***	0.624***	0.513***	0.597***	0.526***	0.424***	0.624***	0.513***
	(0.0386)	(0.0366)	(0.0509)	(0.0492)	(0.0231)	(0.0383)	(0.0354)	(0.0475)	(0.0494)	(0.0223)
Observations	2,145	2,092	603	2,413	892	2,145	2,092	603	2,413	892
Adj. R^2	0.378	0.364	0.269	0.494	0.247	0.378	0.365	0.277	0.495	0.247
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table A.3: OLS estimates of sector-specific regressions using the logarithm of productivity in 2019 as dependent variable, AI use as main explanatory variable, and complementary assets and initial productivity as controls.

Relation between the productivity and the 2011 Share of ICT Engineers	
2011 Share of ICT Engineers – Department & Aggr. Industry	0.0157
	(0.0124)
Initial Log Productivity	0.521***
	(0.0250)
Log Employment	-0.0239**
	(0.0112)
Log Age	-0.0187**
	(0.00918)
ICT Specialists	0.0655**
	(0.0282)
ICT Training	0.0287*
	(0.0159)
Fast Broadband	0.0904***
	(0.0195)
Number of Digital Technologies	0.0122
	(0.00936)
Observations	8,238
R-squared	0.500
Adj. R^2	0.498
Region FE	Yes
Industry FE	Yes

Table A.4: Testing the exogeneity of the share of ICT engineers. The dependent variable is the logarithm of labour productivity in 2019.

B Robustness checks and additional analysis

The relation between productivity growth rates and AI use

	2018-2019		2016-2019		2014-2019		2011-2019	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
AI User	0.0112		0.00251		0.0144		0.00340	
	(0.0115)		(0.0183)		(0.0166)		(0.0197)	
AI Buyer		0.00647		-0.00739		0.00129		-0.0124
		(0.0127)		(0.0210)		(0.0200)		(0.0206)
AI Developer		0.0295*		0.0529**		0.0691*		0.0658*
		(0.0178)		(0.0215)		(0.0374)		(0.0384)
Log Employment	0.0121	0.0119	-0.00179	-0.00225	0.0246***	0.0241***	0.0289***	0.0283***
	(0.0106)	(0.0106)	(0.00602)	(0.00602)	(0.00867)	(0.00875)	(0.00845)	(0.00855)
Log Age	-0.0922***	-0.0922***	-0.00812	-0.00757	-0.0156**	-0.0151**	-0.0181***	-0.0175***
	(0.0104)	(0.0104)	(0.00639)	(0.00649)	(0.00633)	(0.00631)	(0.00600)	(0.00610)
Initial Log Productivity	0.00610	0.00593	-0.118***	-0.118***	-0.235***	-0.234***	-0.320***	-0.319***
	(0.00424)	(0.00426)	(0.0156)	(0.0157)	(0.0223)	(0.0224)	(0.0220)	(0.0221)
ICT Specialists	-8.24e-05	0.000221	0.0128	0.0110	0.0150	0.0129	0.0254	0.0234
	(0.00432)	(0.00433)	(0.0158)	(0.0154)	(0.0223)	(0.0217)	(0.0223)	(0.0221)
ICT Training	-0.00287	-0.00367	-0.000470	-0.000503	0.00102	0.00105	-0.00908	-0.00906
	(0.0108)	(0.0107)	(0.0137)	(0.0137)	(0.0116)	(0.0115)	(0.0160)	(0.0159)
Fast Broadband	-0.00110	-0.00133	0.0365***	0.0358***	0.0284**	0.0276**	0.0476**	0.0465**
	(0.00490)	(0.00484)	(0.00997)	(0.00990)	(0.0116)	(0.0116)	(0.0196)	(0.0193)
Number of Digital Technologies	0.00228	0.00236	-0.000434	-0.000881	-0.00588	-0.00640	0.00178	0.00136
	(0.00897)	(0.00898)	(0.00704)	(0.00702)	(0.00886)	(0.00875)	(0.00825)	(0.00822)
Observations	8,227	8,227	8,033	8,033	7,814	7,814	7,223	7,223
Adj. R^2	0.0551	0.0554	0.0492	0.0498	0.119	0.119	0.179	0.180
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table B.1: Regression results estimating growth rates (logarithmic differences) computed over different time horizons as a function of AI Users, AI Buyers and AI Developers and other controls. The time horizons span between 2019 and the year reported on top of each model (e.g., 2015 refers to the growth rate computed between 2019 and 2015).

In this section we further test the robustness of the key findings presented and provide additional analysis, extending the previous results in several respects. First, we focus on productivity growth rates computed over different time horizons. We compute labour productivity growth as the logarithmic difference between productivity in a base year $t0$ (from 2011 to 2018) and 2019. This is relevant because AI may have diffused very recently in France, fostered by a series of high-tech startups. Accordingly, such regressions may help understand whether and when the use of AI started to generate productivity returns. Furthermore, this model is equivalent to a 2-period firm-year fixed effect estimation. However, this is not our baseline model because, differently from the specifications estimated in Section 5, firms born after $t0$ will be excluded from the sample. Furthermore, the presence of the logarithm of productivity computed in $t0$ may capture not only self-selection into use, but also part of the effect of AI on productivity from the moment of adoption until $t0$. Accordingly, bad control issues may arise from the inclusion of initial productivity among controls (Angrist and Pischke, 2009). Notwithstanding this issue, results from Section 5 are broadly confirmed.

In order to carry out such analysis, we estimate the following productivity growth regression:

$$\begin{aligned}\text{Productivity Growth}_{i,t0,2019} =& \alpha + \beta_1 \text{AI}^{\text{Type}}_{i,2018} + \beta_2 \text{Log-Productivity}_{i,t0} + \beta_3 \text{Log Size}_{i,t0} + \\ &+ \beta_4 \text{Log Age}_{i,t0} + \beta_5 \text{ICT Skills}_{i,2018} + \beta_6 \text{Digitalisation}_{i,2018} \\ &+ \text{Ind.}_{j} + \text{Reg.}_{r} + \epsilon_i\end{aligned} \tag{B.1}$$

Where $\text{Productivity Growth}_{i,t0,2019}$ is the productivity growth rate computed over the period between $t0$ and 2019, where $t0$ takes values from 2011 to 2018 (either 2018, 2016, 2014, or 2011).

The estimated results of Equation B.1 are reported in Table B.1. These further suggest that the relation between AI adoption and productivity growth rates depends on the type of AI users considered. AI developers grow significantly more than other firms during the time period considered. Conversely, the AI use variable, as well as the one indicating the use of AI technologies bought from third parties are not significantly linked with productivity growth in the time periods considered. The initial productivity coefficient is negative and significant, suggesting that firms closer to the frontier grew less than other firms.

These findings confirm the evidence emerging from the productivity regressions discussed in Section 5.2. Furthermore, when less recent years (2011 and 2014) are included in the growth period, the coefficient of AI developers loses precision and remains similar to the one of year 2016. Accordingly, they show how the relation between AI use and productivity may have emerged in more recent years, because AI systems became available recently or due to the presence of high-tech startups driving the AI revolution. Also, firms may have realised how to effectively use AI systems only few years after its adoption.

Second, we discuss a series of additional robustness check by using CRM, ERP and E-commerce separately in place of the number of digital technologies, excluding ICT business services from the sample, or employing multi-factor productivity (MFP) as dependent variable when estimating Equation 2. Firm-level MFP has been computed at the 2-digit industry level via the procedure described by Wooldridge (2009) for the population of French firms between 2008 and 2019. Indeed, a production function cannot be consistently estimated by employing the 2018 data only – corresponding to our matched ICT-balance sheet data –, as the GMM procedure suggested by Wooldridge (2009) to tackle endogeneity in firms' characteristics requires a panel dataset.

We start by reporting the results when the number of digital technologies in disaggregated into CRM, ERP and E-commerce (Model 1 and 2). The estimation results confirm the findings of Section 5, suggesting that the AI-productivity link found above does not depend on the form of the digital controls employed for the analysis.

The results reported in Table B.2 already suggest that the AI-productivity relation for AI developers does not only originate from the ICT sector only. This evidence sheds light on the fact that AI developers may realise productivity returns which are not linked to ICT services only. To investigate this further, we re-estimate the productivity Equation 2 excluding NACE sectors 62-63, which encompass firms primarily engaged in IT and data services. The estimation results are presented in Models 3 and 4 of Table B.2. On average, the relation between general AI use and productivity is positive, but not statistically significant. Similarly, AI buyers outside ICT business services do not exhibit significantly higher productivity. The results for AI developers remain consistent with ones discussed in Section 5, although the

Additional robustness checks on Equation 2

	Changing ICT controls		Excluding 62-63		MFP	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
AI User	0.0314		0.0289		0.0350	
	(0.0224)		(0.0234)		(0.0266)	
AI Buyer		0.00583		0.0137		0.00947
		(0.0201)		(0.0187)		(0.0262)
AI Developer		0.102***		0.0904**		0.0855**
		(0.0385)		(0.0441)		(0.0388)
Log Employment	-0.0244**	-0.0250**	-0.0246**	-0.0253**	0.159***	0.159***
	(0.0111)	(0.0112)	(0.0115)	(0.0115)	(0.0168)	(0.0168)
Log Age	-0.0189**	-0.0180**	-0.0184**	-0.0177*	-0.0662***	-0.0655***
	(0.00880)	(0.00895)	(0.00891)	(0.00905)	(0.00962)	(0.00967)
Initial Log Productivity	0.522***	0.522***	0.527***	0.528***	0.563***	0.563***
	(0.0251)	(0.0250)	(0.0256)	(0.0255)	(0.0289)	(0.0289)
ICT Specialists	0.0662**	0.0635**	0.0674**	0.0648**	0.00636	0.00406
	(0.0291)	(0.0295)	(0.0300)	(0.0305)	(0.0347)	(0.0348)
ICT Training	0.0271*	0.0276*	0.0251	0.0252	0.0228	0.0232
	(0.0161)	(0.0160)	(0.0167)	(0.0166)	(0.0164)	(0.0162)
Fast Broadband	0.0909***	0.0902***	0.0991***	0.0988***	0.0815***	0.0812***
	(0.0190)	(0.0188)	(0.0189)	(0.0187)	(0.0203)	(0.0202)
CRM	0.00559	0.00291				
	(0.0142)	(0.0146)				
ERP	0.0207	0.0214				
	(0.0163)	(0.0163)				
E-commerce	0.00769	0.00796				
	(0.0213)	(0.0214)				
Number of Digital Technologies			0.0123	0.0118	0.00589	0.00546
			(0.00961)	(0.00959)	(0.0110)	(0.0110)
Observations	8,268	8,268	8,018	8,018	8,223	8,223
Adj. R^2	0.497	0.498	0.490	0.490	0.689	0.690
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes

Table B.2: Additional robustness checks employing the dummies for digital technologies instead of the number of other digital technologies as control, excluding ICT business services (NACE 62-63) and using MFP in place of labour productivity.

coefficient slightly decreases.

Second, we report the estimation results for Equation 2 employing MFP in 2019 as the dependent variable in Models 5 and 6 of Table B.2. Similarly to the results for labour productivity, the relation between MFP and AI use is positive and significant for AI developers only. AI users and buyers are positively linked to the MFP, but their coefficients are not significantly different from zero.

Finally, we discuss a series of OLS results estimated on balanced subsamples of treated (AI users, buyers or developers) and untreated units. In particular, we employ a standard nearest neighbour matching estimator that minimizes the Mahalanobis distance between treated and untreated units. This procedure allows us to estimate the causal effect of a treatment by creating a comparable control group that has similar covariate distributions as the treated group, but drops many observations during the matching process. Accordingly, despite confirming the results discussed in Section 5 in terms of sign, significance and magnitude of coefficients, our estimated results will not be representative for French firms.

We build three balanced samples, one for each treatment considered (AI users, buyers or developers). Treated and control units are matched along all controls included in Equation 2: size, age, initial productivity, fast broadband, ICT specialists, ICT training, number of other digital technologies, industry and regional dummies. The match is performed exactly on industries. Also, the sample for AI buyers includes exact matching on AI developers, and viceversa. In order to validate our matching procedure, we report the standardised mean differences on Table B.3 for all variables considered. These differences are high when computed on the full sample, but are always lower than 0.1 after balancing the sample, suggesting that the groups of treated and untreated firms considered are on average very close to each other in terms of characteristics.

We report on Table B.4 the OLS estimation of Equation 2 performed on balanced subsamples. Model 1 and 2 use the sample balanced on AI users. AI users result to be more productive thanks to developers, whereas AI buyers are not significantly more productive. Next we balance the sample based on being an AI buyer in Model 3 or an AI developer in Model 4. Results confirm that AI developers are significantly more productive than other firms, differently from AI buyers.

Balance in subsamples

	Model 1-2 Treatment: AI User		Model 3 Treatment: AI Buyer		Model 4 Treatment: AI Developer	
	Original	Balanced	Original	Balanced	Original	Balanced
AI Developer			0.873	0		
AI Buyer					0.558	0
Log Employment	0.439	0.026	0.697	0.091	0.352	0.018
Log Age	0.053	-0.014	-0.008	-0.016	0.079	0.065
Initial Log Productivity	0.161	0.058	0.187	0.016	0.171	0.095
ICT Specialists	0.514	0.035	1,008	0.057	0.347	0.045
ICT Training	0.45	0.006	0.695	0.019	0.37	0.054
Fast Broadband	0.427	0.003	0.653	0.061	0.369	-0.026
Number of Digital Technologies	0.495	0.047	0.828	0.056	0.395	0.028
CA Food products	-0.13	0	-0.338	0	-0.085	0
CB Textiles	-0.043	0	-0.275	0	-0.023	0
CC Wood and paper products	-0.042	0	-0.054	0	-0.033	0
CD Coke and refined petroleum products	0.036	0	0.065	0	0.028	0
CE Chemicals and chemical products	0.051	0	0.064	0	0.068	0
CF Pharmaceutical products	0.008	0	0.009	0	0.005	0
CG Rubber and plastics products, non-metallic mineral products	-0.013	0	-0.026	0	0.003	0
CH Basic metals and fabricated metal products	-0.07	0	-0.118	0	-0.054	0
CI Computer, electronic and optical products	-0.063	0	0.008	0	-0.09	0
CJ Electrical equipment	0.067	0	0.071	0	0.06	0
CK Machinery and equipment n.e.c.	-0.006	0	-0.011	0	0.009	0
CL Transport equipment	0.066	0	0.09	0	0.068	0
CM Furniture; other manufacturing; repair and installation of machinery equipment	-0.048	0	-0.111	0	-0.026	0
DD Electricity, gas, steam and air-conditioning supply	0.02	0	0.046	0	0.023	0
EE Water supply; sewerage, waste management and remediation activities	-0.111	0	-0.157	0	-0.097	0
FF Construction	-0.16	0	-0.245	0	-0.114	0
GG Wholesale and retail trade	0.005	0	-0.032	0	0.02	0
HH Transportation and storage	-0.088	0	-0.168	0	-0.093	0
II Accommodation and food service activities	-0.061	0	-0.268	0	0	0
JA Publishing, audiovisual and broadcasting activities	0.087	0	0.154	0	0.034	0
JB Telecommunications	0.052	0	0.097	0	0.016	0
JC ICT business services	0.18	0	0.312	0	0.098	0
LL Real estate activities	-0.024	0	-0.084	0	-0.003	0
MA Professional activities	0.132	0	0.157	0	0.1	0
MB Scientific research & development	0.053	0	0.111	0	0.026	0
MC Advertising and market research	0.014	0	0.027	0	0.021	0
NN Administrative	-0.035	0	-0.122	0	-0.012	0
SS Other service activities	0.047	0	0.068	0	0.021	0
Île-de-France	0.289	0.023	0.501	0.007	0.215	-0.027
Centre-Val de Loire	-0.051	0.048	-0.051	0.07	-0.029	0.018
Burgundy-Free-County	-0.071	0.013	-0.076	0.033	-0.04	0.034
Normandy	-0.053	-0.045	-0.146	-0.012	-0.024	-0.005
Heights-of-France	0.008	-0.028	-0.071	-0.031	0.013	0.008
Grand Est	-0.094	-0.016	-0.129	-0.057	-0.076	-0.025
Lands of the Loire	-0.11	0.004	-0.151	-0.05	-0.094	-0.025
Brittany	-0.079	0.004	-0.152	0.011	-0.046	0.005
Nouvelle-Aquitaine	-0.105	0.03	-0.38	-0.013	-0.058	0.03
Occitanie	-0.123	-0.007	-0.226	-0.02	-0.092	0.013
Auvergne-Rhône-Alps	-0.047	-0.016	-0.041	0.087	-0.054	-0.012
PACA	0.028	-0.003	-0.015	-0.06	0.019	0.028
Corsica	0.012	-0.029	-0.031	0	0.02	0

Table B.3: Measures of distance between original and matched subsamples across different regressions.

Productivity across balanced subsamples

Treatment →	AI Users Model 1	AI Users Model 2	AI Developers Model 3	AI Buyers Model 4
AI User	0.052***			
	(0.019)			
AI Buyer		0.009	-0.002	-0.005
		(0.020)	(0.032)	(0.022)
AI Developer		0.085***	0.092***	0.092***
		(0.026)	(0.030)	(0.030)
Log Employment	-0.027***	-0.028***	-0.028**	-0.035***
	(0.007)	(0.007)	(0.011)	(0.009)
Log Age	-0.030**	-0.028*	0.011	0.004
	(0.015)	(0.015)	(0.025)	(0.017)
Initial Log Productivity	0.326***	0.326***	0.335***	0.373***
	(0.032)	(0.032)	(0.042)	(0.038)
ICT Specialists	0.069**	0.065**	-0.052	0.098***
	(0.029)	(0.029)	(0.043)	(0.034)
ICT Training	0.041*	0.041*	0.042	0.013
	(0.024)	(0.024)	(0.037)	(0.026)
Fast Broadband	0.061***	0.059**	0.042	0.070**
	(0.023)	(0.023)	(0.034)	(0.028)
Number of Digital Technologies	0.036***	0.034***	0.039*	0.041***
	(0.013)	(0.013)	(0.021)	(0.015)
Num.Obs.	2524	2524	1045	1887
R2 Adj.	0.486	0.487	0.464	0.487
Region Fixed Effects	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes

Robust standard errors, clustered at the strata level, in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table B.4: OLS estimates of Equation 2 on matched subsamples.