KU Leuven, Faculty of Economics and Business
Dept. of Management, Strategy and Innovation (MSI)

Artificial Intelligence and Firm-level Productivity

Dirk Czarnitzki, Gastón P. Fernández and Christian Rammer

MSI Discussion Paper No. 2203

Artificial Intelligence and Firm-level Productivity

Dirk Czarnitzki [a,b,c], Gastón P. Fernández [d], and Christian Rammer [c]

[a] *Dept. of Management, Strategy and Innovation, KU Leuven, Belgium*
[b] *Centre for R&D Monitoring (ECOOM) at KU Leuven, Belgium*
[c] *Department Economics of Innovation and Industrial Dynamics, ZEW - Leibniz Centre for European Economic Research, Germany*
[d] *Dept. of Economics, KU Leuven, Belgium*

February 2022

Abstract

Artificial Intelligence (AI) is often regarded as the next general-purpose technology with a rapid, penetrating, and far-reaching use over a broad number of industrial sectors. A main feature of new general-purpose technology is to enable new ways of production that may increase productivity. So far, however, only very few studies investigated likely productivity effects of AI at the firm-level; presumably because of lacking data. We exploit unique survey data on firms' adoption of AI technology and estimate its productivity effects with a sample of German firms. We employ both a cross-sectional dataset and a panel database. To address the potential endogeneity of AI adoption, we also implement IV estimators. We find positive and significant effects of the use of AI on firm productivity. This finding holds for different measures of AI usage, i.e., an indicator variable of AI adoption, and the intensity with which firms use AI methods in their business processes.

JEL-Classification: O14, O31, O33, L25, M15

Keywords: Artificial Intelligence, Productivity, CIS data

Corresponding author: Dirk Czarnitzki
KU Leuven
Dept. of Management, Strategy and Innovation
Naamsestraat 69
3000 Leuven, Belgium
Phone +32 478 405 255;
Email: dirk.czarnitzki@kuleuven.be

Acknowledgements: Czarnitzki gratefully acknowledges support by the Flemish Science Foundation (FWO, grant: G077218N)

1. Introduction

Artificial Intelligence (AI) is often regarded as the next general-purpose technology, with a rapid, penetrating, and far-reaching use over a broad number of industrial sectors (Brynjolfsson et al. 2017; Agrawal et al. 2019a; Nolan 2020). A main feature of general-purpose technology is to enable new and complementary production methods that may increase productivity over time (Bresnahan and Trajtenberg 1995; Bresnahan et al. 2002; Brynjolfsson and Hitt 2003; Cardona et al. 2013). As such, it could be expected that the adoption of AI technologies – especially its machine learning component – by firms enacts new business opportunities, spawns new innovative complementarities, and boosts productivity (Brynjolfsson and McAfee 2014).

In that sense, AI can be regarded as an intangible capital asset that firms may invest into and use to generate output through a production function. For instance, several AI's applications such as autonomous vehicles, voice-recognition and speech/text generating systems, or trained neural networks to optimize business energy consumption, would depict AI's potential to increase firm productivity (Brynjolfsson et al. 2017). Other authors have speculated about how AI-based machine and deep learning techniques may have the capacity to enhance firm productivity through its impact on R&D, innovation, and the generation of new ideas (Aghion et al. 2019; Cockburn et al. 2019). Notwithstanding these optimistic predictions about the transformative power of AI, there are other scholars that are more skeptical. Gordon (2014; 2018) claimed that U.S. productivity growth in the coming decades could be much slower, and that IT, and innovative progress in general, will not have any propeller role on the observed productivity slowdown. Gordon argues that, for the period 2004 to 2014, no concluding evidence on the link between fundamental inventions (e.g., smart phones) and U.S. productivity performance has been presented so far. More recently, the hypothesis that human labor could be automated by super-intelligent computers, leading to a rapid acceleration of a technology-driven growth and productivity increase has been rather rejected. Testing that hypothesis within the context of economic growth, Nordhaus (2021) suggests that AI would have to encompass all human tasks in order to reach such "economic singularity".

Besides this clash of predictions on the role of AI adoption, the debate lacks conclusive and rigorous empirical evidence mainly due to restricted availability of data on AI adoption in the business sector. Recent studies focused on analyzing the impact of AI using patent applications and scientific papers related to AI as a main measure of interest (Cockburn et al. 2019; Van Roy et al. 2020; Damioli et al. 2021). However, patent data might provide an incomplete and biased picture of AI's potential effect on productivity since not all AI methods are being patented, and many firms may adopt AI technologies invented by third parties. To analyze the effect of AI on innovative activities and productivity, other studies used data on specific components of AI technologies such as the use of robotics (Graetz and Michaels 2018; Acemoglu and Restrepo 2020) or the use of big data and data-driven managerial decisions (Brynjolfsson et al. 2011; Niebel et al. 2019; Ghasemaghaei and Calic 2019). While these measures represent important elements of AI, they do not account for the entire scope of AI that is used in firms. As stressed by Raj and Seamans (2018), more comprehensive data on the use of AI – especially at the firm level – would be required in order to truly understand the contribution of AI on productivity.

To the best of our knowledge, this is the first paper that addresses this issue in detail by studying the relationship between AI and firm productivity, using data from a representative, large-scale survey that contains rich information on firms' AI adoption. We analyze cross-section as well as panel data from the German part of the European Commission's Community Innovation Survey (CIS). Differently to the standard CIS, the German innovation survey for the reference year 2018 included specific questions on the adoption of AI which covered all types of AI methods that can be used in a firm as well as all kinds of business areas where AI may be applied (see Rammer et al. 2021 for more details). This data provides an ideal base for investigating productivity impacts of the entire diversity of AI applications in firms.

The main findings of this paper are in line with the notion that AI is a productivity-enhancing technology. We find positive and significant effects of the use of AI on firm productivity. This finding holds for different measures of AI usage, i.e., an indicator variable of AI adoption, and the intensity with which firms use AI methods in their business processes. The general positive effect of AI on firm productivity also holds both for sales-based and value-added based productivity measures. In addition, our instrumental variables results suggest a potential direct relationship between the use of AI and productivity that is significantly higher in magnitude than our reduced-form results.

The rest of the paper is organized as follows. Section 2 relates this work to previous studies linking artificial intelligence and productivity. Section 3 presents our theoretical framework and the empirical model. Section 4 describes the data and model variables while Section 5 contains the results of our empirical analysis, including several robustness checks. Finally, Section 6 discusses the findings and identifies further research questions.

2. Artificial Intelligence and Productivity

AI as a productivity-enhancing technology or faltering innovation?

There are several potential channels by which AI can propel firm productivity. For example, machine learning advances have encouraged cheaper and better predictive analyses allowing the full automation of tasks (e.g., self-driving vehicles), larger access to new relevant knowledge and data that can be combined to produce new ideas and know-how, and the generation of new innovations (Agrawal et al. 2019a; Agrawal et al. 2019b; Cockburn et al. 2019). As shown by Aghion et al. (2019) at a conceptual level, AI is an additional input in a firm's production process that can potentially change firm performance due to its effect on the generation of new ideas and technologies, and because it would become handy in solving difficult problems. According to Brynjolfsson et al. (2017), AI should be treated as an additional intangible capital in the production function of firms as the expansion of investment in AI technology may increase productivity similarly to other types of factor inputs. The effective use of AI technologies would result in additional intangible assets such as datasets, firm-specific human skills, and establishing new firm processes. As for other new technology, productivity impacts of AI technologies may not be observed right after its implementation but with some time lag only (Brynjolfsson et al. 2017) as firms may have to adopt other processes and invest in complementary assets in order to fully leverage the productivity enhancing potential of AI (Tambe et al. 2020).

Opposite to this strand of the literature, other economists have flagged the observed slowdown of productivity growth (Gordon 2014; 2018) and adopt a rather modest view on the transformative role of new (digital) technologies such as AI. Their arguments range from the notion that new ideas within firms are increasingly difficult to develop (Bloom et al. 2020) to the possible social, physical, and institutional constraints for accessing knowledge and data that are key for effectively exploiting AI techniques in business processes (see Agrawal et al. 2019b for a discussion).

Nonetheless, as claimed by Raj and Seamans (2018), until now there has not been a database available at the firm level that allows a rigorous study of the role of AI on productivity outcomes. Brynjolfsson et al. (2017) stressed the limitations of the currently available AI data and called for more comprehensive firm-level data on AI use. In that context, the literature reported above is rather speculative because it fundamentally lacks empirical evidence supporting either view. In the absence of measures that would cover the entire variety of AI use in firms, empirical research has so far mainly focused on three areas or approaches: (i) industrial robots and automation of tasks; (ii) patents or scientific publications related to AI; and (iii) data-driven managerial decisions or the use of big data.

AI, Robotics, and productivity

In a pioneering paper, Graetz and Michaels (2018) analyzed industry-level data on industrial robots from the International Federation of Robotics (IRF) for six different countries from 1993 to 2007. These authors showed that country-industry pairs with a larger expansion in robot density were associated to larger benefits in labor productivity. Focusing on the German economy, Dauth et al. (2017) found that at the aggregate level the industrial use of robots enhances labor productivity. More recently, relying on the same IRF data, Acemoglu and Restrepo (2020) showed that the penetration of robots for different time periods had a positive effect on industry value added measures. Other studies analyzing the impact of industrial robotics at the aggregate level on productivity measures include, e.g., Humlum (2019) who estimated a structural model of firm's robotics adoption and found that firms expand the produced output when they embrace industrial robots (see also Stiebale et al. 2020 and Acemoglu et al. 2020 for further evidence).

Closely related, there is a strand of literature that focuses on the idea that technologies associated with AI (e.g., automatic guided vehicles or industrial robots) have the potential to automate tasks that are currently done by humans, and thus, possibly affect the labor market and productivity outcomes. For instance, Frey and Osborne (2017) used detailed information about tasks and occupations and estimated that around 47% of the U.S. jobs were in high risk of automation given recent advances in computerization and machine learning methods (e.g., data mining or machine vision). According to the authors, for example, "telemarketers", "cargo and freight agents", or "watch repairers" would be at a dramatically high risk of automation whereas "recreational therapists" or "nutritionists" would be on the opposite extreme (see also

Arntz et al. 2016). In a similar approach, Felten et al. (2021) classified industries with respect to their AI exposure (AIIE) based on expert assessments. They find highest AIIE scorings for financial services, legal, accounting and consulting services, and IT services, while the AIIE scores are rather low for most manufacturing industries except electronics.

In accordance with the framework developed by Acemoglu and Restrepo (2019a), AI-related or automation technologies could generate a strong displacement effect that may reduce the demand for labor, wages, and employment, and hence, contracting the share of human labor in national income. Notwithstanding, labor demand could either be expanded in those industrial sectors which are being automated or, alternatively, change the task content reinstating labor in a different and new way and generate a countervailing effect from automation that may boost productivity (see, e.g., Acemoglu and Restrepo 2019b for an empirical decomposition of these effects). However, not all robots or automation technologies are directly based on AI, and thus, this strand of the literature might not be strictly identifying the direct effect of AI adoption on productivity. In fact, as noted by Raj and Seamans (2018), the physical nature of robots, which make them a tangible capital asset easy to measure and track, on top of the availability of public data (e.g., IRF data on industrial robotics), have retained the attention on this area of most of the existing empirical work.

AI-related innovation and productivity

There is another body of work attempting to identify the impact of AI technologies through patent data. In general, patents are a relevant driver for productivity growth and performance of firms. For example, Van Roy et al. (2020) analyzed the economic performance of European firms patenting on AI (i.e., "AI inventors") for the period 2000-2016. Using a keyword-based method for identifying AI patents[1], the authors found a significant growth of annual sales in AI inventors with at least one granted patent – especially SMEs – compared to firms with only non-granted AI patent applications. Another recent study investigates the impact of patents associated to the so-called "Industry 4.0" technologies, which would include AI-methods, on the economic performance of firms. Behrens and Trunschke (2020) employed a panel dataset of German firms and found that the marginal effect of an additional "4.0 patent" would increase

[1] See Table 1 in Van Roy et al. (2020) for multiple methods usually used by the literature to identify AI patents.

firms' sales by 8.3%, which diminishes by firm size. Further studies analyzing the contribution of AI adoption on firm performance using patent information are De Prato et al. (2018), Cockburn et al. (2019) and Damioli et al. (2021).

AI, Big Data in the business, and productivity

A further group of papers study the impact of data-driven decisions or the use of big data on firm performance. AI-related methods such as machine learning algorithms and deep neural networks are usually used to analyze the ever-increasing amount of data that firms use and produce as part of their business activities (Taddy 2019). Given the essential role of data for AI technologies, many scholars looked on the impact that the use of big data might have on firms' decisions and performance. Brynjolfsson et al. (2011) analyze the impact of managerial decision-making based on big data on U.S. firm productivity for the period 2005 to 2009. They show that firms adopting data-driven decisions are more productive than competitors that do not use big data methods. More recently, Niebel et al. (2019) employed representative data of German manufacturing and services firms to analyze the effect of big data use on innovation performance such as the sales shares of new products. They show that the use of big data is related to a higher propensity and intensity to innovate (see Ghasemaghaei and Calic 2019 and Lozada et al. 2019 for related works). Notwithstanding the close relationship between big data and AI, not all big data analyses involve the development and adoption of AI, nor do the different varieties of AI applications in firms necessarily involve big data analysis.

In this paper, we aim to extend the empirical evidence about firm-level productivity impacts of AI in three ways. First, we cover all types of AI methods and technologies used for any kind of production process or output, overcoming the limitations of existing studies which focused on specific AI-related technologies such as robotics or big data use. Secondly, we consider all types of active use of AI in a firm, regardless of whether the AI technology was developed by the AI using firm or by others. By covering also the adoption of AI, we extend existing studies that focused on the development of novel AI technologies as indicated by patents. Thirdly, we analyze the role of how intensively a firm is using AI, by developing a measure of the breadth of AI use in a firm, i.e., the variety of different AI methods employed in different AI application areas.

3. Empirical Framework

We follow the standard approach to analyze firm productivity by linking inputs and outputs within a production function approach (Berndt 1991).[2] The production function (*f*) of firms describes the association between a firm's output (Y), measured by annual sales, and total factor productivity (A) as well as a set of inputs, such as capital (K), labor (L), and intermediate inputs such as materials, energy and purchased services (M). We will accommodate this framework and add an additional input to the production function that represents AI adoption (AI). This approach is similar to previous studies analyzing the role of IT or innovation technologies on firm's productivity. For instance, Brynjolfsson and Hitt (2003) estimated a production function with a firm's computer capital stock as an additional production input.[3] Assuming that AI is a sort of intangible asset that is accumulative and depreciable, and that firms can employ to generate output (Brynjolfsson et al. 2017), the production function for firm i in period t is defined as

$$Y_{it} = f(A_{it}, K_{it}, L_{it}, M_{it}, AI_{it}) \text{ with } i = 1, \ldots, N.$$

For simplicity, we assume that the functional form of the production function follows a four-input Cobb-Douglas form as

$$Y_{it} = A_{it} K_{it}^{\alpha_k} L_{it}^{\alpha_l} M_{it}^{\alpha_M} AI_{it}^{\alpha_{ai}},$$

or, equivalently,

$$lnY_{it} = lnA_{it} + \alpha_k lnK_{it} + \alpha_l lnL_{it} + \alpha_m lnM_{it} + \alpha_{ai} lnAI_{it},$$

---

[2] See Bartelsman and Doms (2000) and Syverson (2011) for literature reviews on studies analyzing the determinants of firm's productivity.

[3] Stiroh (2005) and Draca et al. (2006) provide surveys of studies considering IT technologies in firm's production functions. In a recent survey, Abrardi et al. (2021) review studies that consider AI as a new input of production.

where α_k, α_l, α_m, and α_{ai} are unknown parameters to be estimated. The term A, the total factor productivity (TFP), accounts for variations in productivity that are not due to observed inputs but that operate through the production function (Syverson 2011).

In order to get an empirical equation that can be estimated, we introduce stochastic disturbances, which represent random, nonsystematic shocks when firms seek to modify the amount of inputs employed to reach the necessary requirements for profit maximization (Zellner et al. 1966). Thus, for firm *i* in period *t*, we have the following estimable equation

$$lnY_{it} = \lambda_i + \alpha_k lnK_{it} + \alpha_l lnL_{it} + \alpha_m lnM_{it} + \alpha_{ai} lnAI_{it} + X_{it}\beta + \varepsilon_{it},$$

where $\lambda_i + X_i\beta + \varepsilon_i = lnA_{it}$, X_{it} is a matrix of firm's characteristics that are described below, λ_i is a firm-specific time-invariant productivity term, and ε_{it} is a random, unobserved error with mean zero. The coefficient of interest is α_{ai}, the ceteris paribus impact of AI adoption. Of course, we could conceptually also consider that AI is not a factor input but that it rather affects to total factor productivity, and would thus be included in the term *lnA*. We would arrive at the same equation to be estimated.

In the literature on production function estimation it is commonly hypothesized that the firm-specific term λ_i is known to the firm but unobserved to the researcher, and that the firm chooses its factor input optimally based on its knowledge of λ_i which gives rise to an endogeneity concern in econometric estimations. Several approaches have been suggested to overcome that problem. Among others, the most prominent examples are the works of Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg et al. (2015), and most recently Gandhi et al. (2020).

Unfortunately, our data do not allow the application of the most commonly used estimator, as our main database is a cross-section of surveyed firms. We therefore have to limit the current analysis to simpler estimation techniques but offer a number of variations of the specification and conduct robustness test to the extent possible with the current database. In particular, we estimate the following models:

1. Our base models are cross-sectional OLS regressions where we use a dummy variable indicating whether a firm uses any AI technology in its production along with the factor inputs *K*, *L,* and *M*. We have to introduce the restriction that $\lambda_i = \lambda_0 \; \forall \, N$, but use a number of controls, *X*, to mitigate possible endogeneity concerns arising from omitted

variable bias and unobserved heterogeneity, such as industry dummies, firms' age, and variables on general innovation activity. The variables are described in detail in the subsequent data section.

2. We run IV regressions (2SLS) where we instrument the AI variable to address the concern that the more productive firms are those investing in AI. Unfortunately, our database is not rich enough to also instrument the common factor inputs *L*, *K,* and *M*.

3. We implement an entropy balancing procedure as a further approach to address bias due to unobserved heterogeneity. This method divides the sample into treated firms (i.e., AI adopters) and control firms (i.e., non-AI adopters) and attaches a unit weight to each firm based on an entropy balancing on the covariates' sample moments (see Hainmueller 2012). This procedure helps to maintain relevant information in the preprocessed firms by allowing the weights to change smoothly across observations.

4. In order to further address the concern of unobserved heterogeneity, we constructed a very small panel with $T = 2$, such that we can estimated the production function in first differences, i.e., we regress ΔlnY on $\Delta lnL, \Delta lnK, \Delta lnM$, and the AI dummy variable. The firm-fixed effect λ_i is thus accounted for by first-differencing. We run both OLS and IV regressions, where we instrument the AI dummy in the latter. As the AI dummy is just available from one cross-section of the survey, we cannot take the first difference though. We have to modify the interpretation slightly: in the first-difference fixed effects regression, the AI dummy is assumed to approximate that the firm invested into AI in the recent period rather than interpreting the dummy as a stock. The investment assumption is plausible as most AI using firms have only recently started to develop or adopt this technology. If they thus indicated in the survey that they have been using AI in the recent three-year period, it is also likely that they invested.

5. We offer a number of robustness tests of the specifications described above.
    a. Instead of the AI dummy variable, we will also use an AI index that measures the intensity of AI use in the firm.
    b. Instead of using sales as output variable, we also use value-added (and then omit *M* from the right-hand side of the production function).

c. For a sub-sample, we can control for accumulated, past investments that firms have made into their general IT architecture.

4. Data

Data source

We use cross-section as well as a panel data of firms taken from the German contribution to the Community Innovation Survey (CIS) of the European Commission. Differently to other national innovation surveys, the German survey is designed as a panel survey and conducted every year, called 'Mannheim Innovation Panel' (MIP, see Peters and Rammer 2013 for more details). The information collected is representative for all firms in Germany with at least 5 employees in manufacturing, mining, utilities, and business-oriented service sectors (wholesale trade, transportation, financing and insurance, information and communication, professional, scientific, technical, administrative and support services). The MIP follows the methodological guidelines of the CIS as laid down by the Statistical Office of the European Commission (Eurostat) in terms of sampling, data processing, and quality control. The survey is based on a stratified random sample. Data is collected through a standardized questionnaire that can be answered both on paper and online. The response rate of the MIP is between 25 and 35%. A likely bias among responding firms is analyzed through an extensive non-response survey (see Peters and Rammer 2013).

AI variables

In the survey for the reference year 2018, the questionnaire included questions on the use of AI which were not included in any other national CIS. One question asked in a matrix format whether a firm employs AI methods (distinguishing methods related to language understanding, image recognition, machine learning, knowledge-based systems) and in which application areas the method is used (distinguishing products/services, process automation, customer interaction, data analytics, and any other area). AI use includes both in-house developed AI technology and AI technology developed by others and adopted by the firm. Any firm that has developed or adopted at least one AI method by 2018 is considered as an AI user (*AI*). In addition, we use the matrix information to construct a measure of AI intensity (*AIint)* that corresponds to the sum of different AI methods and AI application areas used divided by the maximum number of 20 (as there are 4 methods and 5 application areas). As shown in Table 1, our cross-sectional sample contains 5,849 firms out of which 409 can be classified as AI users, i.e., about 7%. Among the AI using firms, the average value of our AI intensity variable amounts to 12.9%, i.e., the average firm used 2.5 out of the 20 possible combinations of AI technology and areas of application.

Other variables in the production function

Inputs to the production function are measured by the natural logarithms of employment (*lnEMP*), intermediate inputs such as material, energy and purchased services (*lnMAT*), and tangible assets (*lnCAP*). Output is measured by annual sales (*lnSALES*) and alternatively by value added, i.e., sales minus intermediate inputs (*lnVA*). As mentioned, we also estimate a first-differentiated fixed-effects production function, denoting the first difference as $\Delta lnSALES$, $\Delta lnVA$, $\Delta lnEMP$, $\Delta lnMAT$, and $\Delta lnCAP$.

In order to separate a likely productivity effect of AI from possible effects of other innovative and technological activities of the firm, we allow the TFP term, *A*, to vary with some further firm-specific variables *X*. First, firms that perform R&D on a permanent basis may obtain higher TFP levels than non-R&D-performers, or firms that conduct R&D only occasionally, reflecting their higher capacity to absorb and use relevant external knowledge (Cohen and Levinthal 1989). We thus include a dummy, *RDCON,* which equals 1 if the firm is performing R&D on a continuous basis. Besides inventions that may result from own R&D, firms may also benefit from supplier-induced innovation, i.e., they employ new or improved technology embedded in newly acquired machinery or equipment in their production process. The variable *TECHIMP* equals 1 if the firm has adopted, during 2016 to 2018, new or improved production technology relative to their machinery and equipment that has been used prior to the survey period (i.e., before 2016). As further explanatory variable, we use the natural logarithm of a firm's age (*lnAGE*). It could be expected that more mature firms had more opportunities to optimize their production more than younger firms and therefore achieve higher TFP, all else constant. Industry dummies control for unobserved TFP variation across sectors that are not yet captured by any of the structural variables described so far. In the panel fixed effects regressions, unobserved TFP heterogeneity is captured by firm-specific effects.

See Table 8 in the Appendix for further information regarding industry categories considered and the corresponding number of AI using firms and non-AI users per industry.

Endogeneity of AI

When extending the methodological application from cross-sectional OLS regressions to IV regressions in order to address the potential endogeneity of AI, we obviously need instrumental variables. There are several ways in which the impact of AI on productivity might be biased due to an endogenous (non-random) nature of the decision to employ AI methods. First, firms could decide to implement AI technologies as a consequence of higher profits or larger available economic resources. In this case, a firm's productivity level might drive the decision to use AI. Secondly, given the data at hand, omitted covariates not included in our estimated specifications might be correlated with the use of AI, leading to estimates that could be biased. For example, AI investment decisions might be associated to broader digitalization efforts of a firm or to general expansion of a firm's technological infrastructure.

We therefore need a set of instruments that must be correlated with AI usage, but not with unobserved productivity shocks. The following instruments were considered in the estimation process. First, as the MIP questionnaire also collected information on the year of first use of AI by the firms, we construct a measure of investment into AI at the industry level during the time period between the years 2011 to 2018 (*AI_IND*); that is, we consider the number of firms using AI methods by sector for the years 2011 to 2018. The frequency of AI use at the sector level may induce the focal firm to also employ AI, but the sector-level usage should not depend on a single firm's choice.

Second, making use of panel data, we compute the firms' average annual innovation expenses per employee for the period 2011 to 2017 (*PASTINNO*). Innovation expenditure cover internal and external R&D as well as other innovation-related expenses (e.g., acquisition of new equipment and external knowledge, training, marketing, design, and engineering work for innovations). We expect that the more a firm invested into innovation in the past, the more likely it is to use AI at some point. We measure the past innovation expenditure per employee in the regressions to avoid multicollinearity with firm size.

Lastly, as firms may be reluctant to employ AI methods in case they face organizational rigidities and reluctance to new technologies among the workforce, we construct a dichotomous variable which equals 1 if "internal resistance" was stated as an obstacle for the firm's innovation activities (*RESIST*). This instrument is much in the spirit of some of the instruments used by Brynjolfsson et al. (2011). They instrumented the impact of data-driven managerial decision on firm productivity by barriers to IT adoption within the firm.

Descriptive statistics

After removing observations with missing values, erroneous responses and outliers, we end up with a sample size of 5,849 firms. Table 1 provides summary statistics for all model variables.

We generally observe that firms which use AI are, on average, larger in all dimensions, i.e., sales and value added as output and employment, capital and materials as inputs. For instance, the average AI user realizes sales of about 114 million EUR and has 245 employees. For non-users, these numbers are about 24 million EUR sales and 85 employees. We also find that AI users' sales grow faster. The growth rates amount to about 5.4% versus 3.7% (see $\Delta lnSALES$). The adopters also engage more in R&D and innovation as can be seen from the share of firms engaging in R&D on a permanent basis (the mean of *RDCON* is 50% versus 19%) and the share of firms acquiring improved machinery and equipment (*TECHIMP* is 78% versus 59%). The means of past R&D and innovation expenses per employee are also higher for AI users than for other firms.

In a robustness check where we follow Brynjolfsson and Hitt (2003) we also consider firms' past software expenses to account for the general IT architecture of the firm. We use two alternative versions. First a lag of software expenses per employee, and secondly, the average of the firm's software expenses between 2011 and 2017. We cannot really calculate a stock variable as there are many gaps in the time series. We therefore average the expenses across the years that we observe between 2011 and 2017.

Interestingly, there is no difference in age among AI users and other firms, and AI users face higher internal resistance against innovation than other firms; 25% versus 15%.

Table 1 Summary statistics of model variables

Variable	Acronym	Non-AI users (5,440 obs.)				AI users (409 obs.)			
		Mean	**Std. dev.**	**Min**	**Max**	**Mean**	**Std. dev.**	**Min**	**Max**
AI use (D)	*AI*	0	0	0	0	1	0	1	1
AI intensity	*AIint*	0	0	0	0	0.129	0.097	0.050	0.750
Sales	*SALES*	33.689	115.170	0.097	1950.563	114.470	289.907	0.136	2151.413
Employment	*EMP*	84.504	219.719	1	2897	244.547	517.564	2	3150
Materials	*MAT*	15.040	53.116	0.005	791.133	32.651	77.198	0.007	617.418
Capital	*CAP*	17.180	131.436	0.001	6625.510	28.304	99.763	0.003	1098.981
First difference log sales*	$\Delta lnSALES$	0.037	0.140	-0.579	0.722	0.054	0.147	-0.493	0.559
First difference log employment*	$\Delta lnEMP$	0.019	0.105	-0.470	0.510	0.046	0.116	-0.405	0.510
First difference log materials*	$\Delta lnMAT$	0.049	0.228	-1.276	1.330	0.021	0.260	-1.321	1.098
First difference log capital*	$\Delta lnCAP$	0.053	0.214	-0.571	0.980	0.069	0.210	-0.559	0.911
Permanent R&D activities (D)	*RDCON*	0.191	0.393	0	1	0.498	0.500	0	1
New/improved technology (D)	*TECHIMP*	0.585	0.492	0	1	0.777	0.416	0	1
Log age	*lnAGE*	3.279	0.672	1.098	6.811	3.205	0.700	1.098	6.926
Industrial investment into AI	*AI_IND*	10.311	12.572	0	58	20.599	18.042	0	58
Past innovation expenses per empl.	*PASTINNO*	0.003	0.006	0	0.066	0.007	0.010	0	0.055
Internal resistance against inno. (D)	*RESIST*	0.147	0.354	0	1	0.254	0.435	0	1
Log value added**	*lnVA*	1.224	1.686	-4.291	7.490	2.088	2.129	-3.241	7.646
Log past software expenses***	*lnPASTSOFT*	- 2.353	3.105	-8.804	2.639	-1.253	2.242	-8.804	3.526
Log mean software exp. 2011-2017****	*lnAVGSOFT*	-3.214	3.265	-9.986	1.098	- 1.975	2.448	-9.986	1.098

Notes: D: dummy variable; N = 5,849 (*N = 5,567; **N = 5,691; ***N = 2,572, ****N = 3,449). Monetary units are in million EUR. Source: German CIS 2018.

5. Estimation Results

Main results: cross-sectional regressions

Table 2 presents the results for the cross-sectional regressions using the AI dummy as main variable of interest. First, we can see that the ceteris paribus effect of AI use on productivity based on sales as output measure (*lnSALES*) is positive and significant in almost all specifications. Looking at column (1) for the OLS results of the most parsimonious specification of the production function without additional covariates, AI use is associated with higher productivity: AI users annually sell, on average, 13.7% more than non-AI users. We also find a positive and significant relationship between AI use and productivity also after controlling for age (*lnAGE*), innovation engagement (*RDCON* and *TECHIMP*), and sectoral heterogeneity by industry dummies (column (2)). The marginal effect then implies about 5.9% higher sales.

In terms of the control variables, we find expected results. The coefficients of labor, capital, and intermediate inputs do roughly add up to 1. An F-test does not reject constant returns to scale. We also find that productivity is positively associated with firm age, permanent R&D and to a weaker extent to the investment in innovative machinery and equipment.

The positive result also holds, when we address the potential endogeneity of AI by instrumenting the variable with the sectoral adoption level, past innovation expenses and the internal resistance against innovation (see columns 3 and 4). A first-stage F-test of the excluded instruments clearly rejects the Null hypothesis, and the absolute F-value is also higher than 10 which does not lead to a concern about weak instruments (Staiger and Stock 1997). The Hansen J-test on instrument's validity does not reject the null in column (4).[4]

However, the estimated AI coefficients in the IV regressions are somewhat high, i.e. above unity. While possible, such high coefficients do not seem intuitive. The results may partly stem from the presence of unobserved heterogeneity that the IV regression cannot take into account. This motivates our further econometric applications. First, we employ an entropy balancing in

[4] The first-stage results of the IV regressions using *AI* are presented in Table 9 and Table 10 in the Appendix.

column (5) of Table 2. Subsequently we also estimate panel regressions to account for possible unobserved heterogeneity.

Table 2 Productivity effects of AI use (based on sales as output measure): results of OLS and 2SLS regressions (N = 5,849)

Dependent variable:	OLS		IV (2SLS)		IV (2SLS) with entropy balancing
lnSALES	(1)	(2) incl. additional covariates	(3)	(4) incl. additional covariates	(5) incl. additional covariates
AI	0.137***	0.059**	1.361***	1.358***	0.396**
	(0.028)	(0.029)	(0.171)	(0.310)	(0.163)
lnEMP	0.605***	0.595***	0.549***	0.563***	0.702***
	(0.011)	(0.012)	(0.014)	(0.015)	(0.025)
lnCAP	0.056***	0.063***	0.069***	0.063***	0.056***
	(0.005)	(0.006)	(0.006)	(0.007)	(0.014)
lnMAT	0.366***	0.369***	0.378***	0.374***	0.294***
	(0.007)	(0.008)	(0.008)	(0.009)	(0.017)
lnAGE		0.035***		0.050***	0.016
		(0.011)		(0.013)	(0.025)
RDCON		0.051**		-0.063*	-0.004
		(0.019)		(0.036)	(0.034)
TECHIMP		0.030*		-0.007	-0.009
		(0.015)		(0.020)	(0.040)
R-squared	0.903	0.909	0.874	0.878	0.936
F-stat. on joint sig. of IVs in 1st stage	-	-	51.474***	16.424***	16.585***
Hansen's J, p-value	-	-	0.023	0.689	
Industry dummies	No	Yes	No	Yes	Yes

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include an intercept. In the IV regressions, we use the following instruments: investment into AI per sector (*AI_IND*), past innovation expenses per employee (*PASTINNO*), and a dummy indicating internal resistance to innovative activities (*RESIST*). The following statistics were computed to test the joint significance of the instruments: F(3,5842) in column (3); F(3,5823) in column (4); F(3,5823) in column (5).

In column 5 of Table 2, we combine the IV regression of column 4 with entropy balancing, i.e., we perform weighted regressions such that the non-AI users that are similar in their production inputs to the AI users receive more weight in the regressions than other, less similar non-users. For the balancing of the sample with respect to AI, we use all covariates from the respective regression. The results also hold with propensity score matching methods but we prefer the entropy balancing as a more efficient method (see Hainmueller 2012). The results of the entropy-balanced IV regression in column (5) (the other weighted regressions are omitted for reasons of brevity) largely confirm the earlier results of a positive and significant coefficient of AI use. However, the marginal effect amounts to about 40%. This corresponds, on average, to roughly 16 million EUR higher sales (the unconditional average of sales is about 40 million EUR).

Cross-sectional regressions using AI intensity

Table 3 reports the regression results considering AI intensity as a variable of interest (*AIint*), i.e., the breadth with which a firm applies AI methods across application areas. Similarly to our baseline estimates, increasing the fraction of AI methods or areas would have, on average, a significant and positive ceteris paribus impact on productivity.

Table 3 Productivity effects of AI intensity (based on sales as output measure): results of OLS and 2SLS regressions (N = 5,849)

Dependent variable:	OLS		IV (2SLS)	
lnSALES	(1)	(2)	(3)	(4)
AIint	0.763***	0.293*	9.673***	9.474***
	(0.163)	(0.160)	(1.311)	(2.314)
lnEMP	0.606***	0.595***	0.547***	0.564***
	(0.011)	(0.012)	(0.014)	(0.016)
lnCAP	0.056***	0.063***	0.070***	0.063***
	(0.005)	(0.006)	(0.007)	(0.007)
lnMAT	0.366***	0.369***	0.377***	0.371***
	(0.007)	(0.008)	(0.008)	(0.009)
lnAGE		0.034***		0.051***
		(0.011)		(0.013)
RDCON		0.053**		-0.043
		(0.019)		(0.035)
TECHIMP		0.031*		-0.010
		(0.015)		(0.021)
R-squared	0.903	0.909	0.861	0.867
F-stat. on joint sig. of IVs in 1st stage	-	-	37.223***	12.106***
Hansen's J, p-value	-	-	0.170	0.804
Industry dummies	No	Yes	No	Yes

Robust standard errors are in parentheses. All regressions include an intercept. In the IV regressions, we use the following instruments: industrial investment on AI (*AI_IND*), past innovation expenses per employee (*PASTINNO*), and the dummy indicating internal resistance to innovative activities (*RESIST*). The following statistics were computed to test the joint significance of the instruments: F(3,5842) in column (3); F(3,5823) in column (4). *** p<0.01, ** p<0.05, * p<0.1

In terms of the magnitude of the marginal effect, we find similar results to the specification using the dummy variable. When looking at the IV results in column (4), the estimated coefficient is 9.474. A firm that uses one AI technology in one application area (which corresponds to the majority of AI users) would have a value of *AIint* = 0.05 (1 out of 20 combinations of AI technology and area). It would thus realize about 47% higher sales than non-users (0.05 x 9.474), all else constant.

Fixed effects panel regressions

When inspecting our IV regressions carefully, one might be concerned that the estimated marginal effects are much higher than in the OLS regressions. This is a phenomenon that often arises in IV regressions and in our case, the benefits of AI seem possibly unintuitively high. Even though the regression diagnostics do not suggest a weak instrument bias or endogeneity of instruments, remaining unobserved heterogeneity among firms might still result in biased estimates. Therefore, we constructed a panel database to check the robustness of the results once we account for unobserved heterogeneity by including firm-fixed effects. Unfortunately, the AI data is so recent in the survey that we can only build a panel with two time periods, 2017 and 2018. We implement the fixed effect regressions as first-differences approach (1D-FE) where we regress the first difference of *lnSALES* on the first differences of the factor inputs and the AI dummy, i.e., we consider the growth rates of sales and factor inputs between 2017 and 2018. Compared to the cross-sectional regressions, the interpretation of the estimated AI effect has to be adapted accordingly: an AI dummy of 1 is now assumed to indicate that the firm invested into AI in the survey period (flow interpretation) rather than employing AI technology (stock interpretation). Consequently, the estimated coefficient now refers to the productivity growth associated with recent AI investments.

In order to account for any potential correlation between the error term and our covariates, i.e., a violation of the strict exogeneity assumption, we also estimate IV FE regressions.

Table 4 and Table 5 report the results of FE regressions using AI use (*AI*) and AI intensity (*AIint*), respectively. The estimates of interest show the percentage change or growth over time in firm productivity when employing AI methods. The 1D-FE regression for AI use shows a significant impact of AI use on productivity. The marginal effect amounts to about 6% in the IV regression and to 4% when entropy balancing is used additionally.

Table 4 Productivity effects of AI use (based on sales as output measure): results of fixed effect panel regressions (N = 5,567)

Dependent variable: $\Delta lnSALES$	OLS (1)	IV (2SLS) (2)	IV (2SLS) with entropy balancing (3)
AI	0.012*	0.061**	0.044**
	(0.006)	(0.030)	(0.020)
$\Delta lnEMP$	0.363***	0.359***	0.398***
	(0.021)	(0.022)	(0.042)
$\Delta lnCAP$	0.013	0.012	0.056***
	(0.008)	(0.008)	(0.018)
$\Delta lnMAT$	0.198***	0.200***	0.150***
	(0.012)	(0.012)	(0.020)
R-squared	0.214	0.207	0.191
F-stat. on joint sig. of IVs in 1st stage	-	51.739***	127.625***
Hansen's J, p-value	-	0.554	

Robust std. err. in parentheses. All regressions include an intercept. In the IV regressions, we use as instruments: industrial investment on AI (*AI_IND*), past innovation per employee (*PASTINNO*), and the dummy indicating internal resistance to innovative activities (*RESIST*). The following statistics were computed to test the joint significance of the instruments: F(3,5560) in column (2); F(3,5560) in column (3). *** p<0.01, ** p<0.05, * p<0.1

When considering AI intensity, we also find a positive coefficient. On average, the AI firms realize about 5% higher sales based on the IV regression results (0.433 x 0.12, where the latter is the mean of *AIint* among AI users).

Table 5 Productivity effects of AI intensity (based on sales as output measure): results of first difference fixed effect panel regressions, 2017-2018 (N = 5,567)

Dependent variable: $\Delta lnSALES$	OLS (1)	IV (2SLS) (2)
AIint	0.040	0.433**
	(0.041)	(0.211)
$\Delta lnEMP$	0.364***	0.358***
	(0.021)	(0.022)
$\Delta lnCAP$	0.013	0.012
	(0.008)	(0.008)
$\Delta lnMAT$	0.198***	0.200***
	(0.012)	(0.012)
R squared	0.214	0.201
F-stat. on joint sig. of IVs in 1st stage	-	36.706***
Hansen's p-value	-	0.687

Robust std. err. in parentheses. All regressions include an intercept. In the IV regressions, we use as instruments: number of firms using AI per sector (*AI_IND*), past innovation expenses per employee (*PASTINNO*), and the dummy indicating internal resistance to innovative activities (*RESIST*). The following statistics were computed to test the joint significance of the instruments: F(3,5560) in column (2). *** p<0.01, ** p<0.05, * p<0.1

Further robustness checks

We conduct two further robustness checks. First, instead of annual sales as output measure, we use value-added defined as the logarithm of annual sales net of intermediate inputs (*lnVA*). Secondly, instead of controlling only for *TECHIMP* and *R&DCON* as proxies of the firm's innovation affinity, we also considered the firm's general efforts towards digitalization by using expenses for software and databases, measured as the logarithm of lagged software expenses (*lnPASTSOFT*) and, alternatively, as the logarithm of the average annual software expenses in the period 2011 to 2017 (*lnAVGSOFT*). These variables are in the spirit of Brynjolfsson and Hitt (2003) and intended to account for the general IT architecture of the firm. Not all firms in the sample had information on software expenses: 44% of the firms had data in the case of *lnPASTSOFT,* whereas 59% of the firms had information on *lnAVGSOFT*. See Table 1 for descriptive statistics on these variables. In the following, we report the results for the AI indicator (*AI*). Tables 11 and 12 in the Appendix shows the results for AI intensity (*AIint*).

As shown in Table 6, using value-added as output indicator produces results that are consistent with the ones presented above based on sales. Note that our sample is slightly smaller because of some missing values in the value-added measure. The ceteris paribus effect of AI use is positive and significant in almost all specifications. The estimation results in column (1) suggest that AI users achieve 14.5% higher value-added. In the IV regressions, the magnitude of the estimated coefficients increases again as in the case where sales were used as dependent variable. This increase is reduced, however, if the IV regression is run with weights obtained from entropy balancing.

Table 6 Productivity effects of AI use (based on value added as output measure) results of OLS and 2SLS regressions (N = 5,691)

Dependent variable: *lnVA*	OLS (1)	OLS (2)	IV (2SLS) (3)	IV (2SLS) (4)	IV with entropy balancing (5)
AI	0.145***	0.053	1.772***	1.621***	0.528**
	(0.036)	(0.037)	(0.216)	(0.382)	(0.216)
lnLEMP	0.905***	0.903***	0.842***	0.868***	0.995***
	(0.012)	(0.012)	(0.015)	(0.016)	(0.026)
lnCAP	0.142***	0.136***	0.164***	0.138***	0.096***
	(0.007)	(0.008)	(0.008)	(0.009)	(0.017)
lnAGE		0.039**		0.061***	0.012
		(0.015)		(0.018)	(0.037)
RDCON		0.100***		-0.042	0.001
		(0.026)		(0.046)	(0.046)
TECHIMP		0.025		-0.020	-0.005
		(0.020)		(0.025)	(0.051)
R-squared	0.827	0.836	0.770	0.787	0.884
F-stat. on joint sig. of IVs in 1st stage	-	-	52.282***	16.651***	17.218***
Hansen's J, *p*-value	-	-	0.904	0.809	
Industry dummy	No	Yes	No	Yes	Yes

Robust standard errors are in parentheses. All regressions include an intercept. In the IV regressions, we use the following instruments: number of firms using AI per sector (*AI_IND*), past innovation expenses per employee (*PASTINNO*), and the dummy indicating internal resistance to innovative activities (*RESIST*). The following statistics were computed to test the joint significance of the instruments: F(3,5685) in column (3); F(3,5666) in column (4); F(3,5666) in column (5). *** p<0.01, ** p<0.05, * p<0.1

Table 7 shows the estimations results when including lagged software expenses or the average of software expenses for the period 2011 to 2017. This exercise may isolate the productivity impact of AI use more convincingly from the impact of other digitalization efforts on the firm's productivity. At the same time, however, the software expenses might also include specific AI investments and therefore we might also underestimate the AI effect. As we cannot fully separate the AI effect with the given data, we only present this as a robustness test.

As we have the lagged software expenses only for 44% of our observations, we generate a dummy variable indicating when the information is a missing value, *D*(*MISSING*), and we impute zeros in the software variable in order to not lose more than half of the sample for the regression. The dummy captures the effect of imputing zeros and the estimated slope coefficient of *lnPASTSOFT* is obtained from the non-missing values. As an alternative approach we have used the average software expenses as we can then use information from more than one previous survey. When scanning the last six surveys, we have information on software expenses for 59% of the current sample. As the data may now stem from different years, and the available frequencies per firm vary, we average the software expenses across the different years and create the variable *lnAVGSOFT*. We again combine this with a missing value indicator to keep the observations where no software information is available.

Table 7 Productivity effects of AI use (based on sales as output measure) including past software expenses variables: results of OLS and 2SLS regressions (N = 5,849)

Dependent variable:	OLS		IV (2SLS)	
lnSALES	(1)	(2)	(3)	(4)
AI	0.056*	0.054*	1.230***	1.183***
	(0.029)	(0.029)	(0.295)	(0.293)
lnEMP	0.589***	0.586***	0.561***	0.560***
	(0.012)	(0.012)	(0.014)	(0.014)
lnCAP	0.061***	0.060***	0.062***	0.061***
	(0.006)	(0.006)	(0.007)	(0.007)
lnMAT	0.367***	0.366***	0.372***	0.371***
	(0.008)	(0.008)	(0.009)	(0.009)
lnAGE	0.033***	0.030**	0.047***	0.044***
	(0.011)	(0.011)	(0.013)	(0.013)
RDCON	0.046**	0.045**	-0.055	-0.052
	(0.019)	(0.019)	(0.034)	(0.034)
TECHIMP	0.024	0.022	-0.008	-0.008
	(0.015)	(0.015)	(0.019)	(0.019)
lnPASTSOFT	0.027***		0.022***	
	(0.004)		(0.005)	
lnAVGSOFT		0.030***		0.025***
		(0.004)		(0.005)
D(MISSING)	-0.163***	-0.185***	-0.158***	-0.174***
	(0.027)	(0.026)	(0.032)	(0.030)
R-squared	0.910	0.910	0.885	0.887
F-stat. on joint sig. of IVs in 1st stage	-	-	16.691***	16.462***
Hansen's *p*-value	-	-	0.530	0.492
Industry dummy	Yes	Yes	Yes	Yes

Robust standard errors are in parentheses. All regressions include an intercept. In the IV regressions, we use the following instruments: number of firms using AI per sector (*AI_IND*), past innovation expenses per employee (*PASTINNO*), and the dummy indicating internal resistance to innovative activities (*RESIST*). *D(MISSING)* corresponds to a dummy that is equal to 1 if a missing value was imputed by a 0 in the corresponding software expenses variable. The following statistics were computed to test the joint significance of the instruments: F(3,5821) in column (3); F(3,5821) in column (4). *** p<0.01, ** p<0.05, * p<0.1

Similarly as in our reference model (see Table 2), the OLS estimates for AI use are positive and significant. Once we instrument the dummy for AI use (columns (3) and (4)), we observe larger coefficients for AI use than in the OLS regressions. Interestingly, the coefficient of lagged software expenses remains significant even after addressing the endogeneity issues of AI (column (3)). This is virtually the same when we include the annual average of software expenses for the period 2011 to 2017 (column (4)). The coefficients for AI use become slightly smaller in the IV regressions compared to the estimations excluding past software expenses (compare to Table 2). This may partly reflect that some software expenses are related to implementing AI technology, implying that a part of the AI productivity effect is captured by software expenses, or that the AI variable in Table 2 is picking up some effect of the general digitalization efforts measured by software investment. The fact that the AI coefficient remains positive and significant, however, supports our confidence in the presence of a genuine contribution of AI to firm-level productivity.

6. Conclusion

This paper studied the extent to which the use of Artificial Intelligence (AI) technologies by firms contributes to the firms' productivity. We used firm-level panel data from the German innovation survey which provides rich information on firms' performance and technological activities. Most importantly, the data contain information on the use of different AI methods and the business areas in which AI methods have been applied. This database allows to derive measures of AI use that overcome shortcomings in the existing literature. So far, most studies on AI and productivity either relied on specific technologies related to AI (e.g., robotics, big data analysis, see Graetz and Michaels 2018; Acemoglu and Restrepo 2020; Brynjolfsson et al. 2011; Niebel et al. 2019; Ghasemaghaei and Calic 2019) or employed patent data. Though these studies obtained important insights into how AI may drive productivity, they failed to capture the entire variety of how AI is used by firms, including firms that are mere adopters of AI technology, and the intensity of AI use. Our study tried to close this gap.

Using both a dichotomous and a continuous measure of AI usage, we examined the impact of AI on productivity using sales and valued-added as alternative output variables. To overcome potential endogeneity issues of AI use, we employed instrumental variable regressions using AI diffusion at industry level, the firm's past investment in R&D and innovation, and organizational rigidities as instruments.

We found that employing AI technologies has a positive and significant impact on firm productivity. In particular, we showed that both the use of AI and the intensity with which firms exploit the potentials of AI significantly increases both sales and valued-added. This effect remains robust after controlling for several technological features of the firm. The evidence presented in this paper therefore confirms what has previously been hypothesized in the literature, i.e., AI use contributes positively to firms' productivity (Abrardi et al. 2021).

Our results provide evidence that AI is a productivity-enhancing technology; at least in the short-run. This finding has important implications for policy. Fostering the adoption of AI in firms could lead to substantial productivity gains. However, a common preoccupation about the diffusion of AI-related technologies is their possible impact on jobs and inequality, as AI may displace workers and affect low-skilled jobs more than high-skilled ones (see Lane and Saint-Martin 2021; Arntz et al. 2019). As pointed out by Agrawal et al. (2019b), policy measures in education or taxation may be required to counter-balance such developments. At the same time, policy measures should encourage firms to use AI on a broader scale, by tackling barriers related to AI use such as a lack of specialized skills, insufficient IT infrastructure (e.g., scarcity of access to secure cloud computing or low digit rates), and privacy regulation on data usage (see Agrawal et al. 2019b; Reim et al. 2020; Nolan 2020). Our results also suggest that managers need to be better aware of the potential of AI for increasing productivity, as only a small fraction of firms are currently using AI (see Montagnier et al. 2020).

Our empirical findings are subject to several limitations. First, we rely on data from one country and we can currently construct only a very short panel database. While currently unique, our data has several shortcomings. For instance, we cannot employ state-of-the-art techniques to estimate production functions where one can appropriately account for the endogeneity of (all) factor inputs. We can only offer to estimate IV regressions in which we account for the possible endogeneity of AI use. Even though our IV regressions pass the common specification tests, the estimated coefficients of AI use seem somewhat high in cross-sectional regressions. We try to address this issue by supplementing the analysis with entropy balancing methods and fixed effect panel regressions to account for unobserved heterogeneity. In addition, we control for additional covariates, past software expenses, to mitigate further risk of omitted variable bias. However, in the future it would be desirable to compile a database that allows the application of the latest production function estimation methodologies.

A broader country coverage and time-series data on AI use would be highly useful to better identify causal contributions of AI to productivity under a quasi-experimental setting. For example, one could capitalize on policy changes in the regulation of AI or utilize data on technological shocks (e.g., the emergence of new features in AI methods). In order to pin down the channels by which the adoption of AI is boosting firm productivity, we would have to complement our data with more detailed information, particularly on a firm's labor force as well as on its innovative capacity and strategy. For policy, it would be highly important to disentangle the productivity impact of AI into a labor saving (e.g., from automation) and a business expansion one. For example, AI could increase labor productivity by complementing human labor and automating specific tasks (Acemoglu and Restrepo 2019a). AI could also be contributing to the creation of new types of innovation or business models that will lead to new sales and increase productivity through output growth (Rammer et al. 2021).

References

Abrardi, L., Cambini, C., Rondi, L. (2021). Artificial intelligence, firms and consumer behavior: a survey. *Journal of Economic Surveys,* forthcoming.

Acemoglu, D., Restrepo, P. (2019a). Artificial intelligence, automation and work. In Agrawal, A., Gans, J., Goldfarb, A. (eds.). *The Economics of Artificial Intelligence: An Agenda.* University of Chicago Press, 197–236.

Acemoglu, D., Restrepo, P. (2019b). Automation and new tasks: how technology displaces and reinstates labor. *Journal of Economic Perspectives* 33(2), 3–30.

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

Acemoglu, D., Lelarge, C., Restrepo, P. (2020). Competing with robots: firm-level evidence from France. *AEA Papers and Proceedings* 110, 383–388.

Ackerberg, D.A., Caves, K., Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica* 83(6), 2411–2451.

Aghion, P., Jones, B.F., Jones, C.I. (2019). Artificial intelligence and economic growth. In Agrawal, A., Gans, J., Goldfarb, A. (eds.). *The Economics of Artificial Intelligence: An Agenda.* University of Chicago Press, 237–282.

Agrawal, A., Gans, J., Goldfarb, A. (eds.) (2019a). *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press.

Agrawal, A., Gans, J., Goldfarb, A. (2019b). Economic policy for artificial intelligence. *Innovation Policy and the Economy* 19(1), 139–159.

Autor, D. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives* 29(3), 3–30.

Arntz, M., Gregory, T., Zierahn, U. (2017). Revisiting the risk of automation. *Economics Letters* 159, 157–160.

Bartelsman, E. J., Doms, M. (2000). Understanding productivity: lessons from longitudinal microdata. *Journal of Economic Literature* 38(3), 569–594.

Behrens, V., Trunschke, M. (2020). *Industry 4.0 Related Innovation and Firm Growth*. ZEW Discussion Paper 20-070, Mannheim.

Berndt, E. (1991). *The Practice of Econometrics: Classic and Contemporary Reading*. Addison-Wesley.

Bloom, N., Jones, C.I., Van Reenen, J., Webb, M. (2020). Are ideas getting harder to find? *American Economic Review* 110(4), 1104–1144.

Bresnahan, T.F., Trajtenberg, M. (1995). General purpose technologies "engines of growth"? *Journal of Econometrics* 65(1), 83–108.

Bresnahan, T.F., Brynjolfsson, E., Hitt, L.M. (2002). Information technology, workplace organization, and the demand for skilled labor: firm-level evidence. *Quarterly Journal of Economics* 117(1), 339–376.

Brynjolfsson, E., Hitt, L.M. (2003). Computing productivity: firm-level evidence. *Review of Economics and Statistics* 85(4), 793–808.

Brynjolfsson, E., McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. WW Norton & Company.

Brynjolfsson, E., Hitt, L.M., Kim, H.H. (2011). *Strength in numbers: How does data-driven decision making affect firm performance?* Available at SSRN: http://dx.doi.org/10.2139/ssrn.1819486.

Brynjolfsson, E., Rock, D., Syverson, C. (2017). Artificial intelligence and the modern productivity paradox: a clash of expectations and statistics. In Agrawal, A., Gans, J., Goldfarb, A. (eds.). *The Economics of Artificial Intelligence: An Agenda.* University of Chicago Press, 23–57.

Cardona, M., Kretschmer, T., Strobel, T. (2013). ICT and productivity: Conclusions from the empirical literature. *Information Economics and Policy* 25(3), 109–125.

Cockburn, I. M., Henderson, R., Stern, S. (2019). The impact of artificial intelligence on innovation: an exploratory analysis. In Agrawal, A., Gans, J., Goldfarb, A. (eds.). *The Economics of Artificial Intelligence: An Agenda.* University of Chicago Press, 115–146.

Cohen, W.M., Levinthal, D.A. (1989). Innovation and learning: the two faces of R & D. *The Economic Journal* 99, 569–596.

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

Dauth, W., Findeisen, S., Südekum, J., Woessner, N. (2017). *German Robots – The Impact of Industrial Robots on Workers*. CEPR Discussion Paper No. 12306.

De Prato, G., Cobo, M.L., Samoili, S., Righi, R., Baillet, M.V.P., Cardona, M. (2019). *The AI Techno-Economic Segment Analysis*. JRC Working Papers No. 118071, Joint Research Centre.

Draca, M., Sadun, R., Van Reenen, J. (2006). *Productivity and ICT: A Review of the Evidence*. CEP Discussion Paper No. 0749, London.

Felten, E., Raj, M. Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: a novel dataset and its potential uses. *Strategic Management Journal* 42(12), 2195–2217.

Frey, C.B., Osborne, M.A. (2017). The future of employment: how susceptible are jobs to computerisation? *Technological Forecasting and Social Change* 114, 254–280.

Gandhi, A., Navarro, S., Rivers, D.A. (2020). On the identification of gross output production functions. *Journal of Political Economy 128(8*), 2973–3016.

Ghasemaghaei, M., Calic, G. (2019). Does big data enhance firm innovation competency? The mediating role of data-driven insights. *Journal of Business Research* 104, 69–84.

Gordon, R.J. (2014). *The Demise of US Economic Growth: Restatement, Rebuttal, and Reflections.* NBER Working Paper No. 19895, Cambridge, MA.

Gordon, R.J. (2018). *Why has Economic Growth Slowed When Innovation Appears to be Accelerating*? NBER Working Paper No. 24554, Cambridge: MA.

Graetz, G., Michaels, G. (2018). Robots at work. *Review of Economics and Statistics* 100(5), 753–768.

Hainmueller, J. (2012). Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis* 20(1), 25–46.

Humlum, A. (2019). *Robot Adoption and Labor Market Dynamics*. Unpublished manuscript, Princeton University.

Lane, M., Saint-Martin, A. (2021). *The Impact of Artificial Intelligence on the Labour Market: What Do We Know So Far?* OECD Social, Employment and Migration Working Papers No. 256. OECD Publishing.

Levinsohn, J., Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies* 70(2), 317–342.

Lozada, N., Arias-Pérez, J., Perdomo-Charry, G.G. (2019). Big data analytics capability and co-innovation: an empirical study. *Heliyon* 5(10), e02541.

Montagnier, P., Ek, I., Perset, K. (2020). *AI Measurement in ICT Usage Surveys: A Review*. Document for the OECD Working Party on Measurement and Analysis of the Digital Economy (DSTI/CDEP/MADE(2020)3). Organization of Economic Cooperation and Development.

Niebel, T., Rasel, F., Viete, S. (2019). BIG data–BIG gains? Understanding the link between big data analytics and innovation. *Economics of Innovation and New Technology* 28(3), 296–316.

Nolan, A. (2020). Artificial intelligence, digital technology and advanced production. In OECD (ed.). *The Digitalisation of Science, Technology and Innovation: Key Developments and Policies*. OECD Publishing, 119–142.

Nordhaus, W.D. (2021). Are we approaching an economic singularity? Information technology and the future of economic growth. *American Economic Journal: Macroeconomics* 13(1), 299–332.

Olley, G.S., Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64(6), 1263–1297.

Peters, B., Rammer, C. (2013). Innovation panel surveys in Germany. In Gault, F. (ed.). *Handbook of Innovation Indicators and Measurement*. Edward Elgar, 135–177.

Raj, M., Seamans, R. (2018). Artificial intelligence, labor productivity, and the need for firm-level data. In Agrawal, A., Gans, J., Goldfarb, A. (eds.). *The Economics of Artificial Intelligence: An Agenda.* University of Chicago Press, 553–565.

Rammer, C., Czarnitzki, D., Fernández, G.P. (2021). *Artificial Intelligence and Industrial Innovation: Evidence from Firm-Level Data*. ZEW Discussion Paper No. 21-036, Mannheim.

Reim, W., Aström, J., Eriksson, O. (2020). Implementation of Artificial Intelligence (AI): a roadmap for business model innovation, *AI* 1(2), 180–191.

Roy, V., Vertesy, D., Damioli, G. (2020). AI and robotics innovation. In Zimmermann, K.F. (ed.). *Handbook of Labor, Human Resources and Population Economics*. Springer International Publishing, 1-35

Staiger, D., Stock, J.H. (1997). Instrumental variables regression with weak instruments. *Econometrica* 65, 557–586.

Stiebale, J., Südekum, J., Woessner, N. (2020). *Robots and the Rise of European Superstar Firms*. CEPR Discussion Paper No. 15080, London.

Stiroh, K.J. (2005). Reassessing the impact of IT in the production function: a meta-analysis and sensitivity tests. *Annales d'Economie et de Statistique* 79/80, 529–561.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature* 49(2), 326–365.

Taddy, M. (2019), The technological elements of Artificial Intelligence. In Agrawal, A., Gans, J., Goldfarb, A. (eds.). *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press, 61–87.

Tambe, P., Hitt, L., Rock, D., Brynjolfsson, E. (2020). *Digital Capital and Superstar Firms.* NBER Working Paper No. 28285, Cambridge, MA.

Zellner, A., Kmenta, J., Dreze, J. (1966). Specification and estimation of Cobb-Douglas production function models. *Econometrica* 34(4), 784-795.

.

7. Appendix

Table 8 Industries and number of AI using firms per industry

Industry group	**Non-AI users**	**AI users**	**All firms**
Crop and animal production; hunting; fishing; and manufacture of food products, beverages, tobacco products.	234 (4.30%)	5 (1.22%)	239 (4.09%)
Manufacture of textiles, wearing apparel, leather and related products, wood and products of wood and cork except furniture, and articles of straw and plaiting materials.	216 (3.97%)	3 (0.73%)	219 (3.74%)
Manufacture of chemicals and chemical products, and of basic pharmaceutical products and pharmaceutical preparations.	148 (2.72%)	15 (3.67%)	163 (2.79%)
Manufacture of rubber and plastic products, other non-metallic mineral products, basic metals, and fabricated metal products, except machinery and equipment.	727 (13.36%)	35 (8.56%)	762 (13.03%)
Manufacture of computer, electronic and optical products, and electrical equipment.	315 (5.79%)	40 (9.78%)	355 (6.07%)
Manufacture of machinery and equipment n.e.c., motor vehicles, trailers and semi-trailers, other transport equipment; and repair and installation of machinery and equipment.	486 (8.93%)	43 (10.51%)	529 (9.04%)
Manufacture of furniture, other manufacturing, paper and paper products; printing and reproduction of recorded media; and repair of computers and personal and household goods.	285 (5.24%)	13 (3.18%)	298 (5.09%)
Manufacture of coke and refined petroleum products; mining of coal and lignite; extraction of crude petroleum and natural gas; mining of metal ores; other mining and quarrying; mining support service activities; electricity, gas, steam and air conditioning supply; construction of buildings; civil engineering; and specialized construction activities.	314 (5.77%)	7 (1.71%)	321 (5.49%)
Water collection, treatment, supply, and material recovery; sewerage; and remediation activities and other waste management services.	260 (4.78%)	6 (1.47%)	266 (4.55%)
Wholesale and retail trade and repair of motor vehicles and motorcycles; wholesale trade, except of motor vehicles and motorcycles; and retail trade, except of motor vehicles and motorcycles.	302 (5.55%)	10 (2.44%)	312 (5.33%)
Land transport and transport via pipelines; water transport; air transport; warehousing and support activities for transportation; and postal and courier activities.	353 (6.49%)	13 (3.18%)	366 (6.26%)
Publishing activities; motion picture, video and television program production, sound recording and music publishing activities; programming and broadcasting activities; and printing and reproduction of recorded media.	212 (3.90%)	15 (3.67%)	227 (3.88%)

Table 8 Cont.

Industry group	**Non-AI users**	**AI users**	**All firms**
Telecommunications; computer programming, consultancy and related activities; and information service activities.	239 (4.39%)	68 (16.63%)	307 (5.25%)
Financial service activities, except insurance and pension funding; insurance, reinsurance and pension funding, except compulsory social security; activities auxiliary to financial services and insurance activities.	126 (2.32%)	20 (4.89%)	146 (2.50%)
Architectural and engineering activities; technical testing and analysis; scientific research and development; education; human health activities; residential care activities; creative, arts and entertainment activities; libraries, archives, museums, other cultural activities.	456 (8.38%)	44 (10.76%)	500 (8.55%)
Legal and accounting activities; activities of head offices; management consultancy activities; advertising and market research; and public administration and defense; compulsory social security.	309 (5.68%)	52 (12.71%)	361 (6.17%)
Accommodation; food and beverage service activities; real estate activities; other professional, scientific and technical activities; administrative and support service activities; other services.	458 (8.42%)	20 (4.89%)	478 (8.17%)

Source: NACE Rev. 2, Statistical classification of economic activities in the European Community.

Table 9 First stage regressions on AI use. See Table 2 for Second Stage results. (N = 5,849)

Dependent variable: *AI*	(1)	(2)
lnEMP	0.035***	0.025***
	(0.004)	(0.004)
lnCAP	-0.003	-0.001
	(0.002)	(0.002)
lnMAT	-0.008**	-0.003
	(0.002)	(0.002)
lnAGE		-0.012***
		(0.005)
RDCON		0.064***
		(0.011)
TECHIMP		0.028***
		(0.006)
PASTINNO	4.006***	3.132***
	(0.687)	(0.715)
AI_IND	0.041***	0.036***
	(0.010)	(0.010)
RESIST	0.003***	0.002***
	(0.0003)	(0.0006)
Industry dummy	No	Yes

Robust standard errors are in parentheses. All regressions include an intercept. *** p<0.01, ** p<0.05, * p<0.1

Table 10 First stage fixed effect panel regressions on AI use. See Table 4 for Second Stage results. (N = 5,567)

Dependent variable: *AI*	(1)
Δ*lnEMP*	0.116***
	(0.034)
Δ*lnCAP*	0.014
	(0.014)
Δ*lnMAT*	-0.048***
	(0.016)
PASTINNO	7.541***
	(1.360)
AI_IND	0.003***
	(0.0003)
RESIST	0.049***
	(0.011)

Robust standard errors are in parentheses. Regression includes an intercept. *** p<0.01, ** p<0.05, * p<0.1

Table 11 Productivity effects of AI intensity (based on value added as output measure): results of OLS and 2SLS regressions. (N = 5,691)

Dependent variable:	OLS		IV (2SLS)	
lnVA	(1)	(2)	(4)	(5)
AIint	0.956***	0.453**	12.341***	11.311***
	(0.209)	(0.207)	(1.633)	(2.834)
lnEMP	0.906***	0.903***	0.837***	0.866***
	(0.011)	(0.012)	(0.016)	(0.017)
lnCAP	0.142***	0.136***	0.166***	0.137***
	(0.007)	(0.008)	(0.008)	(0.009)
lnAGE		0.039**		0.060***
		(0.015)		(0.017)
RDCON		0.099***		-0.019
		(0.025)		(0.045)
TECHIMP		0.025		-0.022
		(0.020)		(0.026)
R-squared	0.827	0.836	0.752	0.771
F-stat. on joint sig. of IVs in 1st stage	-	-	37.704***	12.155***
Hansen's J, p-value	-	-	0.936	0.782
Industry dummy	No	Yes	No	Yes

Robust standard errors are in parentheses. All regressions include an intercept. In the IV regressions, we use the following instruments: number of firms using AI per sector (*AI_IND*), past innovation expenses per employee (*PASTINNO*), and the dummy indicating internal resistance to innovative activities (*RESIST*). The following statistics were computed to test the joint significance of the instruments: F(3,5685) in column (3); F(3,5666) in column (4). Note that our sample is slightly smaller than for the regressions using annual sales because to missing values in the value-added indicator. *** p<0.01, ** p<0.05, * p<0.1

Table 12 Productivity effects of AI intensity (based on sales as output measure) including past software expenses variables: results of OLS and 2SLS regressions. (N = 5,849)

Dependent variable:	OLS		IV (2SLS)	
lnSALES	(1)	(2)	(3)	(4)
AIint	0.267*	0.270*	8.732***	8.334***
	(0.159)	(0.163)	(2.211)	(2.177)
lnEMP	0.589***	0.587***	0.562***	0.560***
	(0.012)	(0.012)	(0.015)	(0.015)
lnCAP	0.061***	0.060***	0.061***	0.060***
	(0.006)	(0.006)	(0.007)	(0.007)
lnMAT	0.367***	0.365***	0.370***	0.368***
	(0.008)	(0.008)	(0.009)	(0.009)
lnAGE	0.033***	0.029**	0.047***	0.043***
	(0.011)	(0.011)	(0.013)	(0.013)
RDCON	0.048**	0.047**	-0.039	-0.036
	(0.019)	(0.019)	(0.034)	(0.033)
TECHIMP	0.025	0.023	-0.010	-0.011
	(0.015)	(0.015)	(0.020)	(0.020)
lnPASTSOFT	0.027***		0.020***	
	(0.004)		(0.006)	
lnAVGSOFT		0.030***		0.025***
		(0.004)		(0.005)
IMPUTEDSOFT	-0.163***	-0.185***	-0.140***	-0.172***
	(0.027)	(0.026)	(0.037)	(0.033)
R-squared	0.910	0.910	0.874	0.877
F-stat. on joint sig. of IVs in 1st stage	-	-	12.504***	12.284***
Hansen's *p*-value	-	-	0.689	0.641
Industry dummy	Yes	Yes	Yes	Yes

Robust standard errors are in parentheses. All regressions include an intercept. In the IV regressions, we use the following instruments: number of firms using AI per sector (*AI_IND*), past innovation expenses per employee (*PASTINNO*), and the dummy indicating internal resistance to innovative activities (*RESIST*). *IMPUTEDSOFT* corresponds to a dummy that is equal to 1 if a missing value was imputed by a 0 in the corresponding software expenses variable. The following statistics were computed to test the joint significance of the instruments: F(3,5821) in column (3); F(3,5821) in column (4). *** p<0.01, ** p<0.05, * p<0.1

KU LEUVEN
Faculty of Economics and Business
Management, Strategy and Innovation (MSI)
Naamsestraat 69 bus 3535
3000 LEUVEN, Belgium
tel. + 32 16 32 67 00
msi@econ.kuleuven.be
https://feb.kuleuven.be/research/MSI/