Artificial Intelligence and Industrial Innovation: Evidence from Firm-Level Data

Christian Rammer, Gastón P. Fernández
and Dirk Czarnitzki

Artificial Intelligence and Industrial Innovation:
Evidence from Firm-Level Data

Christian Rammer[†], Gastón P. Fernández [‡] and Dirk Czarnitzki[‡,†,§]

This version: March 2021

Abstract

Artificial Intelligence (AI) represents a set of techniques that enable new ways of innovation and allows firms to offer new features of products and services, to improve production, marketing and administration processes, and to introduce new business models. This paper analyses the extent to which the use of AI contributes to the innovation performance of firms. Based on firm-level data from the German part of the Community Innovation Survey (CIS) 2018, we examine the contribution of different AI methods and applications to product and process innovation outcomes. The representative nature of the survey allows extrapolating the findings to the macroeconomic level. The results show that 5.8% of firms in Germany were actively using AI in their business operations or products and services in 2019. The use of AI generated additional sales with world-first product innovations in these firms of about €16 billion, which corresponds to 18% of total sales of world-first innovations in the German business sector. Firms that developed AI by combining in-house and external resources obtained significantly higher innovation results. The same is true for firms that apply AI in a broad way and have already several years of experience in using AI.

JEL-Classification: O14, O31, O32, O33, L25, M15

Keywords: Artificial Intelligence, Innovation, CIS data, Germany

[†] Department Economics of Innovation and Industrial Dynamics, ZEW - Leibniz Centre for European Economic Research, Germany

[‡] Department of Management, Strategy and Innovation, KU Leuven, Belgium

[§] Centre for R&D Monitoring (ECOOM) at KU Leuven, Belgium

[*] Corresponding author: Christian Rammer
Department Economics of Innovation and Industrial Dynamics
ZEW - Leibniz Centre for European Economic Research
L 7, 1 - 68161 Mannheim, Germany
Phone +49 621 1235 184; Email: rammer@zew.de

1 Introduction

Artificial Intelligence (AI) has gained great attention in innovation management and innovation policy as a new general purpose technology that may substantially change the way firms operate and innovate, with far-reaching consequences on markets, economies and societies (Agrawal et al. 2019). AI commonly describes information-technology (IT) methods that allow machines to perform human-like cognitive functions, such as understanding, learning, reasoning and interacting (Baruffaldi et al. 2020). While AI technologies have been developed and applied for several decades (see Haenlein and Kaplan 2019), recent years saw a huge surge in the use of AI as a consequence of the advancing process of digitalisation. The digital interconnection of product, services, machines and communication devices together with the ever increasing amount of data that is generated in digitalised systems offers entirely new opportunities of exploiting data for new applications and increasing the efficiency of operations. AI is a technology that allows an effective and comprehensive use of these data sources. The development of deep learning based on artificial neural networks and other automated machine learning processes offer a wide range of new applications in most industrial activities – from implementing data-based business models and optimising multi-machine systems to enhancing industrial research, potentially leading to a reorganisation of markets, supply chains and production systems (Nolan 2020).

At the same time, there are a number of challenges when it comes to fully utilising the innovative potential of AI (Brock and von Wangenheim 2019, Nolan 2020). Implementing AI methods often requires the adaptation of existing IT systems and raises compatibility issues. The availability and quality of data is another major challenge for effectively using AI methods, as are adequate skills of employees. As for other major new technologies in early diffusion stages, uncertainty on the technological feasibility and market acceptance of new AI applications is high. Potential users may question the credibility of decisions based on AI and may be reluctant to rely on AI-based processes. In addition, legal and regulatory issues (including data protection) as well as data security are potential hampering factors for successfully applying AI. As a consequence of these challenges, it is not guaranteed that using AI will result in more innovations or more successful innovation.

While there are high expectations about the potential of AI for disruptive innovation (OECD 2020a, Brynjolfsson et al. 2017), rather few data exist on the extent to which AI is currently

reshaping innovation in firms. Though some statistical offices and research institutes started to collect data on the use of AI within structural business statistics (see Montagnier et al. 2020 and Zhang et al. 2021), none of these data sources link AI use to innovation. Such a link is critical, however, to model and understand the role of AI for innovation, firm performance and wider economic impacts (Raj and Seamans 2019).

The aim of this paper is to fill this gap by analysing the diffusion of AI technologies in the business enterprise sector and identify the contribution of AI technologies to industrial innovation. We use data from the German part of the European Commission's Community Innovation Survey (CIS) which contained dedicated questions on the use of AI technologies, including items on the type of AI method used, the business areas where AI is applied, whether AI has been developed in-house or externally, and for how long a firm has been using AI technologies. We estimate innovation production functions to explore the contribution of AI to firms' innovation output and extrapolate the estimation results to arrive at total economy estimates.

Our results show that AI is used by only a small fraction of firms (5.8% of the target population of the German CIS), but these firms use AI successfully to increase their innovation output. For example, for one out of four AI using firms, the introduction of a world-first innovation could be assigned to the use of AI. The sales generated by these AI-based world-first innovations represent about 21% of total sales with world-first innovations in AI using firms. As AI using firms tend to be much larger as the average firm, the total economy contribution of AI is quite significant: 3.2% of all firms with world-first innovations in the German business enterprise sector, and 18.1% of total sales with world-first innovation in 2018, corresponding to €16.1 billion, could be attributed to AI.

The paper proceeds with a brief summary of the existing literature on the role of AI for innovation (section 2). Section 3 describes the way we measure AI use in firms and shows some descriptive results. Section 4 presents the model used to estimate the contribution of AI to innovation output and the estimation results. Section 5 discusses the total economy estimates and section 6 concludes.

2   Artificial Intelligence and Innovation

AI is a new type of general purpose technology that drives innovation in several ways (Trajtenberg 2019). The specific power of AI relates to the extensive and often real-time analysis of heterogeneous data on business processes and the use of products or services in order to identify regularities and patterns, to learn what drives the analysed phenomena, and to autonomously solve problems, including newly arising ones (Taddy 2019). Through the skills of perception, cognition and problem-solving, which characterise most types of human work (Brynjolfsson et al. 2017), AI can be employed to automate processes, improve the quality of operations and enhance the features of products and services, based on self-learning algorithms. The innovation impact of AI basically refers to three areas of innovation in firms:

- Products, services and business models: AI enables new ways of data-based business models that exploit, often in real time, information on customers, product use and product-relevant conditions to offering new types of products and services (see Reim et al. 2021, Lee et al. 2019, Garbuio and Lin 2019, Valter et al. 2018).

- Production, delivery and administrative processes: AI can be used to optimise operations (particularly by automating human activities) and helping humans to make the right diagnoses and decisions. For example, AI methods are used to identify patterns in production problems or defects in manufactured products and to implement predictive maintenance (Nolan 2020). There is also a huge rationalisation potential of AI in administrative operations (e.g. automated responses to telephone calls and e-mails). AI can support decision making, e.g. for interpreting x-rays by physicians.

- R&D and innovation processes: AI is reshaping the process of research and development (R&D) through the extensive use of large (often passively generated) datasets and enhanced prediction algorithms (Cockburn et al. 2019). AI can substantially fasten and broaden R&D processes, e.g. in pharmaceuticals and chemicals (compound identification and discovering new industrial materials through neural network approaches) or in the machinery and equipment industry (e.g. through virtual factories that allow to simulate and improve production processes, Nolan 2020).

In addition, the advance of AI applications drives complementary innovations often needed to leverage the full potential of AI, e.g. in digital communication (e.g. 5G), chip technology,

server infrastructure, new computing approaches (e.g. quantum technology) (see Brynjolfsson et al. 2017).

The literature also points to a number of challenges that can limit the innovation impact of AI (Nolan 2020, Reim et al. 2021, Haefner et al. 2020). First of all, data availability and data quality are often a main barrier to successfully implementing AI. High-value uses of AI typically combine diverse data types and require a constant data inflow of high quality (in terms of format, completeness, consistency and metadata information). The need for digitalising, cleaning, shaping, connecting and labelling data can easily eat up possible efficiency gains from using AI. Secondly, specific skills related to implementing AI methods are scarce and restrict firms in rolling out AI applications on a larger scale. In addition, AI projects often require a mix of skills, and setting up the necessary multidisciplinary teams can be challenging as well. Thirdly, for many firms AI is a rather new technology that is associated with uncertainty about its technological feasibility. A further challenge relates to a lack of transparency of how AI methods arrive at their results. The complex assembly of different functions, and its abstraction levels, impairs traceability ('black-box issue'). As a consequence, trust for AI may lack both among employees and among users when individuals do not understand how AI operates. In addition, AI applications may raise legal and regulatory issues, particularly if data from different owners are merged and the outcome of AI-based algorithms cannot be traced back to a responsible organisation or individual that can be made liable.

The trade-off between great innovation potentials and substantial challenges provides an interesting ground for studying the role of AI for industrial innovation. However, only few studies have analysed the contribution of AI to innovation in firms so far as representative data on the diffusion of AI and its role in innovation processes is largely lacking (see Raj and Seamans 2019). In the absence of survey data on AI and innovation, several authors attempted to identify the use of AI methods through patent data. Fujii and Managi (2017) used a code-based approach, focussing on international patent classification (IPC) code G06N ('computer systems based on specific computational models', corresponding to US patent classification code (USPC) 706 'data processing, artificial intelligence'). Cockburn et al. (2018) also used code 706, complemented by a keyword search on patent titles relating to AI. EPO (2017) used solely a code-based definition that should capture AI-related patents in the field of machine understanding. The OECD also developed a purely code-based approach that focuses on human interface, human cognition and meaning understanding (Inaba and Squicciarini 2017).

Baruffaldi et al. (2020) used text mining techniques to search abstracts and patent documents that refer to AI-related papers in order to identify IPC codes that most frequently contain AI-related inventions (see Van Roy et al. 2020 for a summary of these methods in recent studies). All these studies are descriptive in nature and do not link AI use to innovation at the firm level. An exception is Behrens and Trunschke (2020) who used patent data on 'industry 4.0' technologies (a fraction of these patents relate to AI methods) to examine the impact on firms' sales, finding a stronger positive effect as compared to other patents, but which is diminishing with firm size.

Patent data, however, provide only an incomplete picture on the use and diffusion of AI as only a fraction of new AI methods are patented, and firms may implement and use AI methods based on technologies invented by others. The firm-level data used in this paper reveals that only 30% of firms that actively use AI in their products, services or operations are relying on patents to protect their intellectual property. With respect to IP related to AI, this share is most likely much smaller as many AI applications are based on existing AI technology and do not represent technological inventions in their own right.

Other studies looked at specific technologies that are closely linked to AI or rely on AI technologies for analysing the role of AI in innovation. One such technology are robots. They represent a specific area of AI application with respect to the automation of processes, though not all robots are based on AI. While there are a number of studies that examine the impact of robots on productivity and other firm performance measures (Stiebale et al. 2020, Acemoglu et al. 2020, Humlum 2019), only few works linked the use of robots to other area of innovation. Liu et al. (2020) used industrial robot data at the sector level to examine the relation of AI and technological innovation for Chinese manufacturing. They show that the use of robots fosters other technological innovation through accelerating knowledge creation and technology spillovers.

Another strand of literature examines the use of big data and firm innovation. Though big data is only one element of AI, and big data analysis can be carried out without employing AI methods, there is nevertheless a close connection between the two. Niebel et al. (2019) analysed the relationship between firms' use of big data and innovative performance in terms of product innovation and found higher likelihood for becoming a product innovator as well as higher market success of product innovations. Ghasemaghaei and Calic (2019) showed that the characteristics of big data are positively linked to the firms' innovation competency. Ferraris et

al. (2018) found a positive relation between big data analytics capabilities and firm performance which is stronger in case a firm has an effective knowledge management. Lozada et al. (2019) found a positive relation of big data capabilities and more agile processes of product and service co-creation.

To the best of our knowledge, no studies have been carried out yet that look at the entire field of AI application in firms and their role for innovation and that would allow to estimate the economy-wide relevance of AI for the innovation performance of the business sector. This paper fills this gap.

3   The Use of AI in Firms

3.1   Data source

This study employs firm-level data on the use of AI and innovation output in terms of new products and new processes. The database is the German part of the Community Innovation Survey (CIS). The CIS is a biennial exercise coordinated by the Statistical Office of the European Commission and constitutes the official innovation statistics for the EU. The CIS is a representative, large-scale survey designed to measure innovation inputs, innovation outputs and innovation-relevant characteristics of firms and their market environment. The survey is based on the definitions and measurement concepts for innovation data as laid down in the Oslo Manual (OECD and Eurostat 2018).

In the survey for the reporting year 2018 (CIS 2018), the CIS questionnaire used in Germany included a question on the use of AI (no other EU countries included this question).[1] The question identified the type of AI use based on a matrix design that correlates AI methods and application areas (see Figure 1). The phrasing was deliberately kept simple and short as the CIS questionnaire is not addressed to AI specialists, but to innovation and technology officers (in large corporations) or to general managers or firm owners (in small and medium firms). The aim was to briefly characterise main AI methods (allowing for a free text answer) and main application areas in order to determine whether firms were using AI at the time of the survey (February to July 2019). In addition to the matrix question, information on who mainly

---

[1] For more details on the German CIS, which is conducted as a panel survey ('Mannheim Innovation Panel'), see Peters and Rammer (2013).

developed the AI methods used by the firm (in-house and/or external) and the first year of AI use in the firm.

Figure 1: AI question in the German CIS 2018

Source: German CIS 2018.

Note that the chosen way to measure AI use in a firm is focussing on the *active* use of this technology. Passive forms of accessing AI methods, i.e. by placing own products on online sales platforms that are operated by others who use AI methods to run the platform, are not included. The AI measure does capture the state of AI use at the time of the survey irrespective of the time the AI methods have been introduced. It is hence a kind of stock variable, representing the accumulated investment into AI that is still in operation.

The German CIS 2018 targets firms with 5 or more employees in mining, manufacturing, utilities and a range of service sectors (wholesale, transportation, information & communication, banks & insurances, professional & technical services, business support services). The survey had a sample size of 43,672 firms.[2] Usable responses were recorded for 8,821 firms, resulting in a response rate of 20.2%. In order to evaluate a likely bias between responding and non-responding firms with respect to innovation activities, a comprehensive non-response survey was conducted, interviewing 10,250 non-responding firms (29.1% of all non-responding firms). The non-response survey revealed a higher share of innovation active firms among respondents (71.1%) than non-respondents (65.1%). For data extrapolation, this

---

[2] The sample size includes firms from the CIS core sectors and size classes as well as additional samples for some sectors and size classes beyond the CIS core as well as additional samples for four German regions. The figure excludes firms from the gross sample that could not be contacted because they were not existing anymore.

bias is corrected by using a correction factor for the firms' sampling weights (see Behrens et al. 2017: 27ff for the method used). In the following we report statistics that are extrapolated from the sample to the target population of the survey, i.e. the German manufacturing sector and business-related services.

3.2 Descriptive statistics on various dimensions of AI usage

The analysis of firm responses to the AI question reveals that in the first half of the year 2019, 5.8% of all firms in the target population actively used AI methods in their business operations (Table 1). This corresponds to about 17,500 firms. Only a small share of these firms developed these AI methods mainly in-house (0.9% of all firms, 16% of AI using firms) while most relied on externally developed AI methods (60% of AI using firms). 24% used AI methods that were developed both in-house and by others. One out of five AI using firms have first used AI prior to 2011 while one out of four only recently (2018 or 2019) started to apply AI in their firm.

Table 1: Use of AI in firms in Germany (first half of 2019)

	Share in all firms (%)
Firms with any active use of AI	5.8
thereof: mainly based on in-house developed AI methods	0.9
thereof: mainly based on externally developed AI methods	3.5
thereof: based both on in-house and externally developed AI methods	1.4
thereof: first use of AI before 2011	1.2
thereof: first use of AI before between 2011 and 2015	1.1
thereof: first use of AI before between 2016 and 2017	1.9
thereof: first use of AI in 2018 or 2019	1.6

All figures were extrapolated to the total population of firms in Germany with 5 or more employees in industries (Nace rev. 2) B to E, 46, H, J, K, 69, 70.2, 71 to 74, 78 to 82.
Source: German CIS 2018.

The share of 5.8% for AI using firms compares quite well to the results obtained from recent business surveys on the use of AI in other countries (see Montagnier et al. 2020). For Korea (firms with 10+ employees), a share of 1.5% has been reported for the year 2017. In the same year, the share of AI using firms in Canada (firms with 20+ employees) was 4.0%, and in Denmark it was 6.0% in 2019 for firms with 10+ employees. Higher shares than those found for the German business enterprise sector were found in France (11.0% in 2018 for firms with 10+ employees) and Japan (14.1% in 2017 for firms with 100+ employees).

The share of AI using firms varies greatly among industries and size classes (Table 2). The industry with the highest share of AI using firms is software and IT services (18.3%). 14.3% in

consulting and advertising, and 12.2% in financial services use AI in their business operations. In manufacturing industries, highest shares are found for the electronics and electrical equipment industry (11.0%) and the chemicals and pharmaceuticals industry (8.4%). The use of AI is very rare among firms from wholesale trade (1.0%) and transportation and logistics services (1.5%).

Table 2: Use of AI in firms in Germany by industry and size class (first half of 2019)

	Firms with AI use as a share in all firms (%)	Sales of firms with AI use as a share in total sales (%)
Sector (Nace rev. 2)		
Consumer goods, *manuf. of* (10-12, 14-15, 31-32)	2.2	7.6
Other materials, *manuf. of* (13, 16-18, 22-23)	2.6	10.1
Chemicals and pharmaceuticals, *manuf. of* (20-21)	8.4	30.7
Metals and metal products, *manuf. of* (24-25)	4.7	20.5
Electronics and electrical equipment, *manuf. of* (26-27)	11.0	32.8
Machinery and equipment, *manuf./repair of* (28, 33)	6.7	17.4
Vehicles, *manuf. of* (29-30)	5.1	38.0
Utilities, waste management, mining (5-9, 19, 35-39)	3.6	23.7
Wholesale trade (46)	1.0	7.4
Transport and logistics services (49-53)	1.5	16.5
Media services (58-60)	6.5	28.0
Software, IT services (61-63)	18.3	33.7
Financial services (64-66)	12.2	51.3
Consulting, advertising (69, 70.2, 73-74)	14.3	25.3
Engineering and R&D services (71-72)	6.5	15.7
Other producer services (78-82)	2.5	13.1
Size class (no. of employees)		
5 to 9	3.3	2.8
10 to 19	5.4	3.6
20 to 49	7.6	7.6
50 to 99	6.7	5.1
100 to 249	9.7	11.4
250 to 499	15.7	15.4
500 to 999	21.6	35.6
1,000 and more	30.8	65.5

All figures were extrapolated to the total population of firms in Germany with 5 or more employees in industries (Nace rev. 2) B to E, 46, H, J, K, 69, 70.2, 71 to 74, 78 to 82.
Source: German CIS 2018.

These figures are mainly driven by the AI adoption behaviour of small firms since small firms represent the largest number of all firms in any industry, and small firms show substantially lower AI adoption rates (3.3% for firms with 5 to 9 employees, 5.4% for firms with 10 to 19

employees) compared to large firms (30.8% for firms with 1,000 or more employees). For assessing the economic relevance of AI use, the share of sales in an industry that is represented by AI using provides a more accurate picture. In financial services, more than 50% of the industry's total sales were obtained by AI using firms. In manufacturing of vehicles, this share is 38%, and in the software and IT industry, it is 34%. Among the group of large firms with more 1,000 or more employees, AI using firms represent 66% of all sales of this size class. Among micro firms (5 to 9 employees), AI using firms contribute only 2.8% to the size class' total sales.

The most frequently used AI method in German firms in 2019 was machine learning (55% of all AI using firms). AI-based image and pattern recognition methods were used by 49% of firms, and 46% had implemented knowledge and expert systems based on AI (Table 3). AI methods for language and text understanding were used by 30% of the firms. AI methods were most often applied to products and services (60% of AI using firms) and for the automation of processes (56%). 34% used AI for data analysis.

Table 3: AI methods and applications areas of AI in firms in Germany (2019)

AI method	Area of application					Total
	Products, Services	Automation of processes	Interaction with clients	Data analysis	Other areas	
Language/text understanding	15.1	9.5	7.9	7.0	5.5	30.3
Image/pattern recognition	24.1	30.8	4.9	11.3	3.3	48.9
Machine learning	32.3	30.4	9.1	16.7	4.3	54.6
Knowledge/ expert systems	24.9	19.1	9.8	16.4	4.4	46.2
Total	59.9	55.6	22.0	33.9	11.1	

All figures were extrapolated to the total population of firms in Germany with 5 or more employees in industries (Nace rev. 2) B to E, 46, H, J, K, 69, 70.2, 71 to 74, 78 to 82.
Note: The totals sum up to more than 100% as each firm could report multiple methods and areas of application.
Source: German CIS 2018.

# 4 Estimating the relationship between AI and innovation

The main aim of the paper is to assess the role of AI for the firms' innovation performance. For this purpose, we employ an innovation production function (Mairesse and Mohnen 2002) and regress a variety of variables of firms' innovation outcome on whether the firm uses AI (and in what way) while controlling for other variables that may affect innovation outcomes.

As the use of AI is surely embedded into a firm's broader innovation strategy, in particular also in its general digitalisation efforts, the goal of this paper is not to establish strong evidence of causality running from AI to innovation outcomes. This would require either an instrumental variable approach or the exploitation of some exogenous variation in the use of AI which we currently do not have at hand. We therefore see our study rather as an explorative study where we suggest potential causal relationships that could be studied in more rigorous econometric works in the future. We limit the ambition of our regression analysis to controlling for the most important variables driving innovation outcomes in general, and also firms' other (non-AI) digitalisation efforts within its innovation strategy in particular, in order to isolate the effect of AI on innovation. While we outline below that we are using a rich set of covariates which mitigates endogeneity concerns due to omitted variables, we cannot rule out that some remaining unobserved factors may drive innovation outcomes, the adoption and use of AI as well as other digitalisation efforts and key innovation input variables simultaneously. Remaining endogeneity concerns may only be ruled out further in the future when either panel data become available, or when other surely exogenous variation in the adoption and use of AI can be utilised.

4.1 Conceptual model

The conceptual model that guides the design of the empirical estimations consists of three main groups of variables (Figure 2). Innovation output is measured by a series of variables on new or improved products or processes that have been introduced to the market or implemented in the firm. These innovation outcome variables are related to the use of AI, other digitalisation efforts of the firm and other determinants including innovation input measures, general firm capabilities and market characteristics. The details are described in the following subsections.

Figure 2: Conceptual model on the role of AI for innovation output in firms
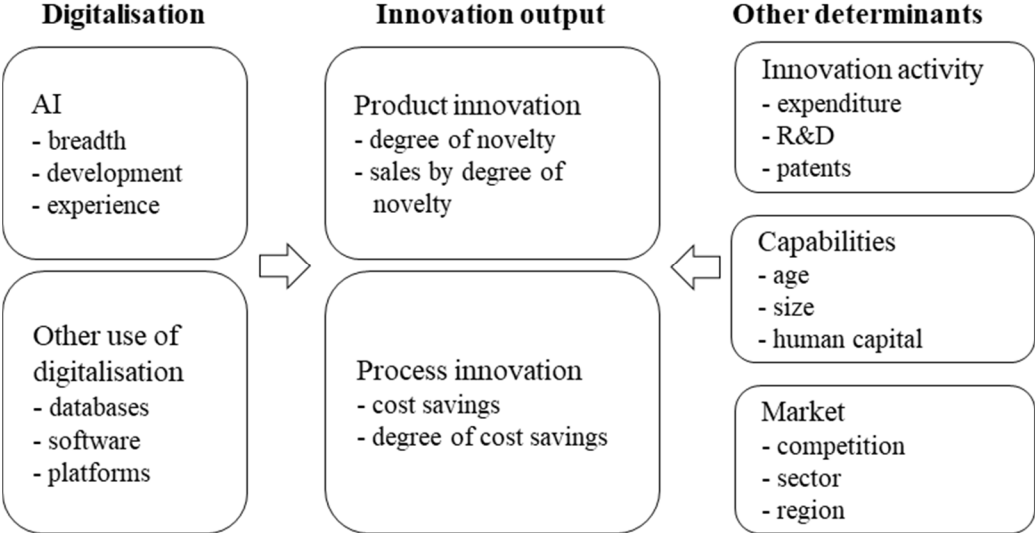
4.2 Dependent variables and descriptive statistics

As one of the main purposes of this paper is a first explorative analysis on where AI is intertwined with the innovation process in the business sector, we use a relatively large set of dependent variables that shed light on different dimensions of the innovation supply chain. We distinguish product from process innovation and use a number of subcategories of each innovation type. Product innovations are new or improved products that differ significantly from a firm's previous products and include both physical goods, services and digital products. We subsequently distinguish product innovations by their degree of novelty with respect to a firm's market. We separate new-to-market innovations from innovations that are only new to the firm. For new-to-market innovations, we further separate world-first innovations from those that are only new to a regional or sectoral sub-market. The importance of these innovations for the firm are measured by their respective shares in the firm's total sales.

Process innovations are considered separately from product innovations. A firm is a process innovator if it has implemented at least one new or improved process in the preceding three year period that differs significantly from the firm's previous processes. Process innovations can then be further separated into cost reducing process innovations and others (the latter may result in higher quality of the produced goods but not lower unit cost, for example). The economic returns of cost saving process innovations are approximated by the share of unit cost reduction.

All subsequent descriptive statistics are weighted results, i.e. the numbers are extrapolated from the sample to the population of firms using sampling weights.

Product innovation

For product innovation, we first use a dummy variable (*PDI*) indicating whether the firm introduced a new good or service (that differs significantly from the firm's previous goods or services) to its portfolio in the time period 2016-2018. On average, about 61% of firms with some AI activity report a product innovation. In the subsample of firms without AI activity this share amounts only to 35% (see Table 4).

We can separate product innovations by their degree of novelty in a firm's market. A product innovation can either be only new to the firm (*PDI new firm*), i.e. the product innovation is an imitation or adaption of other firms' products, or it can be new to the market that the firm is serving (either in geographical or product space) (*PDI new market total*). This market novelties can be further divided into world-first innovations or for new sub-markets (*PDI new sub-market* and *PDI world first*). Among the AI-using firms about 19.4% report to have had a market novelty of which 8.3% only had sub-market novelties and about 11.1% reported world-first innovations. These shares are higher than in the group of firms without AI activity, where they amount to about 8.8% that can be divided into 5% for sub-markets and 3.8% world-first, respectively.

In addition to the dummy variables of the product innovation types, we can also study the volume of new product sales. The survey also allows splitting the total sales of the firms into the following sub-categories:

1.	Sales with unchanged products (goods and services)
2.	Sales with product innovations
	2a. Sales with innovative products that were only new to the firm's portfolio
	2b. Sales with market novelties
	2bi. Sales with novelties only new for regional or sectoral sub-markets
	2bii. Sales with world-first innovations

Total sales: 100%

When looking at the sales shares rather than just at the event of product introductions, differences between AI-using firms and others are also striking. For instance, in the population of AI using firms, almost 21% of total sales are achieved with new products, while this figure is only 13% in the group of firms not using AI. The share of sales with world-first innovations is more than twice as high in the sub-population of AI users (2.9% vs. 1.3%).

Table 4: Descriptive statistics of product innovation variables (weighted results)

	Firms with AI activity	Firms without AI activity
	Mean	Mean
Dummy variables		
PDI total	0.607	0.346
PDI new firm only	0.413	0.258
PDI new market total	0.194	0.088
PDI new sub-market	0.083	0.050
PDI world first	0.111	0.038
Sales shares		
PDI share_all	0.207	0.125
PDI share_newfirm	0.152	0.098
PDI share_newmarket	0.055	0.027
PDI share_newsubmarket	0.026	0.014
PDI share_worldfirst	0.029	0.013

Note: Figures represent weighted results for 17,448 AI using firms and 282,190 not AI using firms within the firm population of the German CIS.

Process innovation

For process innovation, we use a dummy variable, *PCI total*, indicating whether the firm has introduced a process innovation during 2016-2018. On average, about 75% of firms with AI activity report to have introduced at least one new process. Among the firms without AI activity this share amounts only to 51% (see Table 5).

Table 5: Descriptive statistics of process innovation variables (weighted results)

	Firms with AI activity	Firms without AI activity
	Mean	Mean
Dummy variables		
PCI total	0.746	0.514
PCI no cost	0.481	0.393
PCI cost reduction	0.265	0.121
Share of unit cost reduction		
PDI cost share	0.052	0.034

Note: Figures represent weighted results for 17,448 AI using firms and 282,190 not AI using firms within the firm population of the German CIS.

In addition, the process innovation variable can be split into two categories, i.e. whether the process resulted in a reduction of unit costs of production (*PCI cost reduction)* or not (*PCI no cost)*. The latter could imply higher work safety, or higher quality of the production process and the like. We observe that the AI-using firms obtain higher shares in both dimensions. When looking at the share of reduction in the unit cost of production, we also find that AI-using firms achieve higher amounts with 5.2% versus 3.4% (see Table 5).

It will be interesting to see whether these descriptive findings regarding product and process innovation performance differences between AI users and other firms hold in multiple regression analyses when we control for the firms' general innovation strategy by accounting for several innovation input dimensions, other digitalisation efforts and additional structural firm characteristics.

4.3   Empirical measures of AI

The role of AI for a firm's innovation output is measured by an indicator on AI use. In addition to the AI indicator, we also examine whether heterogeneity in the use of AI is associated with different innovation outcomes. In addition to the AI indicator variable, we explore possible effects of AI measures that can be formed out of the survey data introduced in Section 3.2. One dimension is the *AI breadth* which is generated from the two sets of dummy variables representing the four AI methods and the five application areas of AI that have been distinguished in the AI question in the survey:

o (i) language/text understanding, (ii) image/pattern recognition, (iii) machine learning, (iv) knowledge/ expert systems;
o (i) products or services, (ii) automation of processes, (iii) interaction with clients (iv) data analysis, (v) other applications.

We define the breadth of AI usage as the count of items that a firm has reported affirmative in the survey. The breadth theoretically thus ranges from 0 to 20. On average, the firms employing AI technology use it in 2.77 of the described dimensions and the maximum reaches 15. This means no firm exhausts all possible combinations of application areas and AI methods.

Furthermore, we consider *AI experience*, measured as number of years that have elapsed since a firm's first adoption of AI technology. The average AI usage amounts to 4.4 years and the maximum to 29 years (i.e. AI technologies were first used in 1990).

Finally, a set of three dummy variables indicated whether the firm developed the AI technology mainly in-house, in-house and in collaboration with others, or whether the AI technology has been mainly developed by others and the focal firm is a mere adopter. Most firms, 51%, only adopt AI technology rather than customising it for their firm-specific production process. 20% develop the AI mainly internally, and 29% do that mainly in collaboration with external partners.

Table 6: Descriptive statistics of different AI dimensions conditional on AI use

Variable	Obs	Mean	Std. Dev.	Min	Max
AI breadth	442	2.77	2.28	1	15
AI experience	442	4.44	5.43	0.5	29
AI inhouse	413	0.20	0.40	0	1
AI others	413	0.51	0.50	0	1
AI inhouse+others	413	0.29	0.45	0	1

4.4 Supplemental digital capabilities and resources

AI methods are one technology to utilise the opportunities of digitalisation. Effectively leveraging the potential of AI requires additional digital capabilities and resources, which can both support the contribution of AI to innovation, and enable and advance innovation in their own right. We consider three such capabilities and resources: databases and data analytics, software programming capabilities, and digital platforms. It is possibly important to include these other digital capabilities and resources in the empirical model in order to avoid that their contributions to innovation is captured by the AI variables:

- The availability and quality of large data sets is one, if not the key, prerequisite for an effective use of AI (Agrawal et al. 2019, Obschonka and Audretsch 2019). The availability of big data, capacities to analyse these data, and data management capabilities that generate 'smart' data can create a number of innovation opportunities (George et al. 2014, Wamba et al. 2017), independent from using or not using AI technologies. We use an indicator on whether firm invested into setting-up, maintaining and analysing own databases (including the purchase of external data).
- Software programming capabilities are another digital competence that both can spur innovation in a variety of ways (see for example Arora et al. 2013) and supports the effective implementation of AI applications in existing IT systems and data structures. We consider firms that have own in-house programming capacity or purchased programming services externally as being equipped with software programming capabilities.

- Digital platforms are a tool for collecting data that are highly relevant for innovation-oriented AI applications (particularly with respect to social media), but they can also initiate new innovation approaches, particularly in re-organising marketing and interaction with business partners (Sedera et al. 2016) and developing new business models (Brousseau and Penard 2007, Täuscher and Laudien 2018). We use two items from a question on the use of different channels for acquiring knowledge (social web-based networks or crowd-sourcing, open business-to-business platforms or open-source software) to proxy a firm's use of digital platforms.

4.5 Further control variables

Aside from digitalisation, there are many firm and market characteristics that may influence innovation output. We consider three groups of variables:

- *Innovation input*: Following Crépon et al. (1998), we use the amount of innovation expenditure relative to a firm's sales as well as variables characterising the type of R&D activity of a firm (continuous or occasional) to capture a firm's input to innovation that will affect the type and scale of innovation output.

- *General firm capabilities*: Based on the extensive empirical literature on the determinants of innovation output (see Cohen 2010 for an overview), we include firm size, firm age and a human capital variable (share of graduated employees) for capturing heterogeneity among firms' capabilities to develop and successfully introduce market innovations. For the models on the economic returns from innovation, we also include the amount of marketing efforts, measured by marketing expenditure per employee. Higher marketing efforts are likely to increase sales independently from possibly superior characteristics of the innovative product or service.

- *Market characteristics*: There is ample evidence that the type and intensity of competition in a firm's market can be a major driver or barrier for innovation decisions and the outcome of innovation (Varian 2018, Aghion et al. 2005, Cohen and Levin 1989). We use an index on the intensity of competition that captures the relevance of various characteristics of the firm's market environment. In the survey the firms respond to the following eight characteristics of their competitive environment by rating each item into the categories "3: applies fully", "2: applies somewhat", "1: applies very little", "0: does not apply". The index is the sum of the scores on the statements (see Rexhäuser and Rammer 2014 for further details on the index):

- o products become outdated quickly;
- o the technological development is difficult to predict;
- o products / services from competitors are easily substituted for those of your enterprise;
- o major threat to market position because of entry of new competitors;
- o competitor's actions are difficult to predict;
- o demand development is difficult to predict;
- o strong competition from abroad;
- o price increases lead to immediate loss of clients.
- • In addition, industry and regional dummies control for further market characteristics possibly affecting innovation outcomes.

Table 14 in the Appendix shows descriptive statistics (and the definition) for all model variables. All variables are measured using data collected in the German CIS 2018 (except for age which is calculated using the firm's year of foundation as documented in Creditreform data, see Bersch et al. 2014). The total number of observations for model estimations varies between 6,738 and 6,283 for different dependent variables.

As we are interested in estimating the contribution of AI to innovation at the macroeconomic level, we run weighted estimations. Depending on the dependent variable, sampling weights either indicate the number of other firms that are represented by a firm in the sample (for innovation output variables that refer to the number of firms, e.g. firms that introduced a certain type of innovation), or the volume of sales that is represented by a firm in the sample (for economic returns from innovation such as sales with product innovation). The weights are calculated for the sample of the German CIS using 63 strata (21 industries, 3 firm size classes) from the firm population data obtained from the business register of the Federal Statistical Bureau of Germany.

5 Estimation Results

5.1 *Base models on AI use*

In the base models, we use a dummy variable for AI use. The results on the associations between AI use and innovation outcomes are shown in Table 7 (type of product innovation), Table 8 (process innovation) and Table 9 (economic returns from product innovation).

Table 7 shows the average marginal effects of AI use on the introduction of product innovations that were obtained from weighted Probit regressions. We find that firms employing AI technology are 8.5% more likely to introduce a product innovation than firms that do not use any AI. As the average probability to have a product innovation among AI using firms is about 60.7%, the economic magnitude of the AI contribution is sizeable. It amount to about 16% [= 8.5 / (60.7-8.5)]. When looking in the types of product innovation, we find that the firms employing AI are at the forefront of innovation, as the association of AI with product innovation mainly shows for market novelties and there especially for world-first innovations. The effect on world-first innovation amounts to 2.3% higher likelihood. The average value in the sample of AI-using firms amounts to 11.1%, and thus the marginal effect reflects an increase of about 26% [= 2.3 / (11.1-2.3)].

We also obtain interesting results for the control variables: the other digitalisation variables, i.e. software capabilities, data capabilities, and platform use, are positive and significant in the regressions. All are significant in the equation for any type of product innovation ('total'), and some variation occurs in the regressions on the different sub-types. Generally we can conclude, however, that the firms' digital affinity plays a role for product innovation and it seems useful to control for other IT-related variables as otherwise some general effects of IT affinity might be miss-assigned to our focal variable of AI. If the other IT usage would be omitted and their contribution would thus be in the error term of the econometric model specification, serious endogeneity concerns would arise.

Table 7: Marginal effects of AI use on the introduction of product innovation by degree of novelty (results of sampling-weighted Probit regressions)

	Product innovation				
	Total	Only new to firm	New to market	Only new to regional or sectoral market	World first
AI use	0.085***	0.024	0.025*	-0.005	0.023***
	(0.031)	(0.028)	(0.013)	(0.012)	(0.007)
Software capabilities	0.073***	0.071***	0.013	0.008	0.005
	(0.017)	(0.017)	(0.010)	(0.009)	(0.006)
Data capabilities	0.078***	0.031*	0.043***	0.032***	0.011**
	(0.018)	(0.018)	(0.010)	(0.009)	(0.005)
Platform use	0.096***	0.081***	0.018**	0.015*	0.004
	(0.015)	(0.015)	(0.009)	(0.008)	(0.006)
Continuous R&D	0.267***	0.038*	0.157***	0.101***	0.070***
	(0.021)	(0.022)	(0.011)	(0.011)	(0.008)
Occasional R&D	0.188***	0.098***	0.099***	0.062***	0.054***
	(0.021)	(0.022)	(0.012)	(0.010)	(0.009)
Innovation exp. / sales	0.256***	0.129**	0.067***	0.008	0.042***
	(0.063)	(0.052)	(0.022)	(0.022)	(0.011)
Share of graduates	0.014	-0.012	0.026	0.003	0.022**
	(0.033)	(0.033)	(0.019)	(0.017)	(0.011)
Age (ln # years)	-0.017**	-0.008	-0.008	-0.006	-0.002
	(0.009)	(0.009)	(0.005)	(0.004)	(0.002)
Size (ln # employees)	-0.002	0.001	-0.003	-0.005*	0.002
	(0.006)	(0.006)	(0.003)	(0.003)	(0.002)
# observations	6,475	6,475	6,475	6,475	6,475

Standard errors in parentheses. ***, **, *: significant at $p<0.01$, <0.05, <0.1.
All regressions include the competition index and its squared value as control variable, a set of 20 industry dummies, 14 regional dummies and an intercept.
Data source: German CIS 2018.

We also find that all innovation input measures are positive and statistically significant in all models as one could have expected. The share of graduates is only positively associated with world-first product innovations. Interestingly, after controlling for AI, digitalisation and innovation inputs, we do not find strong size or age effects in the regressions. The effect of the competition index is inversely U-shaped (not shown in table). However, the inflexion point of the curve is at the very right of the data distribution and that implies that the product innovation propensity basically increases with competition. This is only found in the regressions on any kind of product innovation and the "only new to the firm" regression, though. Competition does not affect the introduction of market novelties which may create a temporary quasi-monopolistic position of the firm.

The results on process innovation are shown in Table 8. AI use is associated with an 8% higher likelihood to have any type of process innovation. When looking as cost-reducing process innovations versus others, it turns out that the effect of AI technology shows in the cost-reducing process innovations but not in the others. Firm with AI technology are 4.2% more likely to introduce cost-reducing processes. Of course, in the context of process innovations, AI might be part of the innovation itself.

Table 8: Marginal effects of AI use on the introduction of process innovation by type of impact (results of sampling-weighted Probit regressions)

| | Process innovation | | |
	Total	Not leading to cost reduction	Leading to cost reduction
AI use	0.080**	-0.017	0.042**
	(0.034)	(0.030)	(0.017)
Software capabilities	0.139***	0.107***	0.056***
	(0.017)	(0.019)	(0.013)
Data capabilities	0.122***	0.067***	0.037***
	(0.019)	(0.020)	(0.012)
Platform use	0.109***	0.073***	0.034***
	(0.016)	(0.017)	(0.011)
Continuous R&D	0.206***	0.071***	0.095***
	(0.022)	(0.024)	(0.014)
Occasional R&D	0.221***	0.086***	0.099***
	(0.024)	(0.025)	(0.015)
Innovation exp. / sales	0.130*	0.144**	0.023
	(0.067)	(0.060)	(0.034)
Share of graduates	-0.011	-0.009	0.002
	(0.035)	(0.036)	(0.025)
Age (ln # years)	-0.038***	-0.024***	-0.014**
	(0.009)	(0.009)	(0.006)
Size (ln # employees)	0.025***	0.004	0.017***
	(0.006)	(0.006)	(0.004)
# observations	6,738	6,738	6,738

Standard errors in parentheses. ***, **, *: significant at p<0.01, <0.05, <0.1.
All regressions include the competition index and its squared value as control variable, a set of 20 industry dummies, 14 regional dummies and an intercept.
Data source: German CIS 2018.

With respect to control variables, we find comparable effects as in the product innovation regressions. The controls on other IT usage, i.e. software and data capabilities and platform use, are positively related to process innovations. Also the two R&D dummies have positive, and statistically significant marginal effects. Interestingly, we find that older firms are less likely to

introduce new processes. This might either imply that their production processes are well calibrated or that they become inflexible over time.

For identifying the relationship between AI use and economic returns from innovations (sales with product innovation, cost reduction from process innovation), we run weighted OLS regressions accounting for sampling weights (see Table 9).

We find a relatively strong association between the use of AI and the sales shares of innovations with higher degrees of novelty. While there is only a weakly significant effect on product innovations in general and no effect on sales of products that a just new to a firm's portfolio (i.e. adoption or imitation), the sales of market novelties increase 1.7 percentage points with AI, and the share of world-first innovation sales is associated with a 1.3 percentage points increase in case firms employ methods of AI. As the sample averages of these variables are 5.5% and 2.9%, respectively, the use of AI is associated with an increase of 45% [= 1.7 / (5.5-1.7)] and about 81% [= 1.3 / (2.9-1.3)] respectively.

The unit cost-reduction in firms with AI use is 0.8 percentage points higher. As the average value of cost reductions is 5.2%, the marginal effect of AI is also not negligible in relative terms. The on first sight rather small coefficient of 0.008 accounts for a relative change of about 18% [= 0.8 / (5.2-0.8)].

With respect to the control variables we find quite similar results as in the Probit regressions on the corresponding dummy variables of the different innovation categories. Therefore we do not discuss those in detail. As we here accounted for sales volumes we had added controls on marketing expenditure and patent use but the results remain somewhat inconclusive. Marketing is never statistically significant. Patent use, however, shows a negative sign in the imitation regression which is not surprising as patenting firms might rather be innovation leaders and not imitators. This is consistent with the fact that we find strong effects of patent use on market novelties and especially world-first innovations that might be successfully protected by the firms' intellectual property rights.

Table 9: Coefficient estimates of AI use on sales of product innovation (by degree of novelty) and cost reduction from process innovation (results of sampling-weighted OLS regressions.)

	Sales share from product innovation					Share of unit cost reduction owing from process innovation
	Total	Only new to firm	New to market		World first	
			Total	Only new to regional or sectoral market		
AI use	0.027*	0.011	0.017**	0.004	0.013**	0.008**
	(0.015)	(0.013)	(0.008)	(0.006)	(0.006)	(0.004)
Software capabilities	0.007	0.009	-0.002	-0.001	0.000	0.004**
	(0.007)	(0.006)	(0.003)	(0.002)	(0.002)	(0.001)
Data capabilities	0.041***	0.030***	0.010***	0.005*	0.005**	0.006***
	(0.008)	(0.007)	(0.004)	(0.003)	(0.003)	(0.002)
Platform use	0.023***	0.024***	-0.001	0.002	-0.003	0.004**
	(0.007)	(0.006)	(0.003)	(0.002)	(0.003)	(0.002)
Continuous R&D	0.097***	0.060***	0.037***	0.028***	0.009***	0.012***
	(0.011)	(0.010)	(0.006)	(0.005)	(0.004)	(0.003)
Occasional R&D	0.055***	0.044***	0.011**	0.011***	0.000	0.011***
	(0.012)	(0.011)	(0.005)	(0.004)	(0.003)	(0.003)
Innov. exp. / sales	0.202***	0.116***	0.087***	0.033**	0.054***	0.021**
	(0.035)	(0.029)	(0.022)	(0.015)	(0.018)	(0.008)
Share of graduates	0.044**	0.024*	0.020	0.001	0.018	0.006
	(0.018)	(0.014)	(0.013)	(0.006)	(0.011)	(0.004)
Age (ln # years)	-0.015***	-0.011***	-0.003**	-0.002	-0.001	-0.003***
	(0.003)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)
Size (ln # employ.)	-0.010***	-0.006***	-0.004***	-0.003***	-0.001	0.000
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Competition index	0.002	0.003**	-0.001*	-0.001	-0.001	0.000
	(0.002)	(0.002)	(0.001)	(0.001)	(0.000)	(0.000)
Competition index^2	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Marketing expend.	0.001	0.001	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Patent use	-0.006	-0.019**	0.014**	0.002	0.012***	-0.004*
	(0.010)	(0.008)	(0.006)	(0.005)	(0.004)	(0.002)
# observations	6,283	6,283	6,283	6,283	6,283	6,626

Robust standard errors in parentheses. ***, **, *: significant at p<0.01, <0.05, <0.1.
All regressions include a set of 20 industry dummies, 14 regional dummies and an intercept.
Data source: German CIS 2018.

5.2   Robustness check: dependent variables measures in t+1

As a robustness check, we re-run the model estimations for dependent variables measured in t+1, using data from the German CIS 2019.[3] This allows to analyse lagged relations, particularly for firms that only recently started to use AI. However, the number of observations decreases for this analysis as not all firms participate in the innovation survey every year, making the estimation results more volatile. We were able to use three dependent variables from period t+1. The total innovation sales, and the split of those into sales of product innovations that were only new to the firm (imitation or adoption) and market novelties (see Table 10).

The estimation results for the lagged models largely confirm our base model results. We find positive effect of AI on all three variables. The effect for market novelties is only weakly significant though. Using the lead in t+1 of the dependent variables gives us some more confidence that the results identified before do not only arise because of simultaneity between innovation performance and the adoption of AI technology, but that potential causal effects might be running from the use of AI technology to innovation performance.

---

[3] The innovation survey in Germany, differently to most other countries, is conducted annually, based on a panel sample. See Peters and Rammer (2013) for more details on the panel nature of the German CIS.

Table 10: Coefficient estimates of AI use on sales of product innovation in period t+1 (results of sampling-weighted OLS regressions.)

| | Sales share from product innovation | | |
	Total	Only new to firm	New to market
AI use	0.054***	0.041***	0.013*
	(0.016)	(0.015)	(0.007)
Software capabilities	0.004	0.007	-0.004
	(0.009)	(0.008)	(0.003)
Data capabilities	0.026***	0.022***	0.003
	(0.009)	(0.008)	(0.004)
Platform use	0.005	0.004	0.001
	(0.008)	(0.007)	(0.003)
Continuous R&D	0.100***	0.064***	0.035***
	(0.013)	(0.011)	(0.007)
Occasional R&D	0.049***	0.048***	0.001
	(0.013)	(0.013)	(0.003)
Innovation exp. / sales	0.049***	0.032**	0.017**
	(0.018)	(0.016)	(0.008)
Share of graduates	0.231***	0.126***	0.106***
	(0.044)	(0.034)	(0.030)
Age (ln # years)	-0.016***	-0.011***	-0.005**
	(0.004)	(0.004)	(0.002)
Size (ln # employ.)	-0.009***	-0.006**	-0.002
	(0.003)	(0.003)	(0.001)
Competition index	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.001)
Competition index^2	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Marketing expend.	0.002	0.002	0.000
	(0.002)	(0.002)	(0.001)
Patent use	0.012	-0.003	0.015**
	(0.012)	(0.010)	(0.008)
# observations	4,143	4,143	4,143

Robust standard errors in parentheses. ***, **, *: significant at p<0.01, <0.05, <0.1.
All regressions include a set of 20 industry dummies, 14 regional dummies and an intercept.
Data source: German CIS 2018.

5.3   Models on AI characteristics

We run a series of additional estimations for different characteristics of AI use to analyse heterogeneity among AI users: (i) the origin of the development of AI technology (in-house, others, inhouse + others) (ii) breadth of AI use, and (iii) firms' experience with AI.

For reasons of brevity, Table 11 only summarises the average marginal effect of the main variables of interest in the regressions, i.e. the AI variables. The results on the controls are omitted. These results are comparable to the results presented above.

Each column in the table is based on a separate regression, either Probit or OLS. Although the results are somewhat mixed, interesting nuances can be identified. World-first product innovations are mainly associated with in-house development of AI technologies. This seems intuitive. If the adoption of AI would be supplier-induced, for instance, these suppliers will almost surely also deliver AI technology to others. If AI is an integral part of the innovation, it will be unlikely that world-first innovation can be made with technology that is available to many other users. Interestingly, AI mainly developed by others is very positively associated with product and process innovation in general, though. This might be an indication that firms seeking some technological advancement of their products and processes in the dimension of AI may well rely on business partners, possibly such as suppliers or IT consultants, to upgrade their products and processes.

Table 11: Marginal effects / Coefficients of AI development on the introduction of product and process innovation (results of sampling-weighted Probit and OLS regressions)

	Product innovation (Probit)			Process innovation (Probit)		Sales share from product innovation (OLS)			Process innovation (OLS)
	Total			Total		Total			
		New to market			Leading to cost reduction		New to market		Share of unit cost reduction due to process innovation
		Only new to regional or sectoral market	World first				Only new to regional or sectoral market	World first	
AI in-house	-0.015	0.036	0.035***	-0.104*	0.021	0.010	0.030	0.045*	0.009
	(0.066)	(0.025)	(0.013)	(0.061)	(0.031)	(0.043)	(0.028)	(0.027)	(0.009)
AI others	0.119***	0.019	0.008	0.155***	0.064***	0.022	0.005	0.004	0.006
	(0.041)	(0.019)	(0.010)	(0.049)	(0.024)	(0.017)	(0.008)	(0.005)	(0.004)
AI in-house + others	0.026	0.024	0.028**	0.083	0.026	0.056	0.041	0.015	0.019
	(0.055)	(0.021)	(0.011)	(0.059)	(0.027)	(0.039)	(0.025)	(0.014)	(0.012)

All regressions include a full set of controls (software capabilities, data capabilities, platform use, continuous R&D, occasional R&D, innovation expenditure, share of graduates, age, size, competition, marketing, patent use) as well as 20 industry dummies, 14 regional dummies and an intercept.
Standard errors in parentheses. ***, **, *: significant at $p<0.01$, <0.05, <0.1.
Data source: German CIS 2018.

When exploring the effects of AI breadth and AI experience, we largely confirm the earlier results of the simpler regressions where an AI dummy was used, but also find some interesting nuanced results. We allow for non-linear effects by including also the squared values of breadth and experience in the regressions (Table 12).

We generally find positive effects of both AI breadth and AI experience on the innovation output variables. However, we find decreasing marginal returns in several of the regressions, i.e. the squared value of the coefficient is negative, and the curve thus describes an inverse U-shape. As average marginal effects might be somewhat misleading in such situation, because marginal effects may change signs in the data range, we report the coefficient estimates along with the inflexion point of the estimated curve when the squared term is significant in the regression. Otherwise, the relationship is basically linear.

Table 12: Effects of breadth and experience of AI use on the introduction of product and process innovation: estimated coefficients (results of sampling-weighted Probit and OLS regressions)

	Product innovation (Probit)			Process innovation (Probit)		Sales share from product innovation (OLS)			Process innovation (OLS)
	Total			Total		Total			Share of unit cost reduction due to process innovation
		Only new to regional or sectoral market	World first		Leading to cost reduction		Only new to regional or sectoral market	World first	
AI breadth	0.145***	0.046	0.117**	0.091	0.137***	0.016*	0.005	0.004	0.006**
	(0.054)	(0.053)	(0.059)	(0.056)	(0.045)	(0.009)	(0.006)	(0.004)	(0.003)
AI breadth^2	-0.013**	0.001	-0.001	-0.004	-0.011**	-0.001	0.000	0.000	-0.000**
	(0.006)	(0.005)	(0.005)	(0.006)	(0.004)	(0.001)	(0.001)	(0.001)	(0.000)
Inflexion$^\#$	6				6				5
AI experience	0.057*	0.067**	0.100***	0.052*	0.021	0.007*	0.006**	0.004*	0.001
	(0.032)	(0.031)	(0.033)	(0.031)	(0.028)	(0.004)	(0.003)	(0.002)	(0.001)
AI experience^2	-0.002	-0.002	-0.003*	-0.002	-0.000	-0.000*	-0.000*	0.000	0.000
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Inflexion$^\#$			18			12	13	13	

All regressions include a full set of controls (software capabilities, data capabilities, platform use, continuous R&D, occasional R&D, innovation expenditure, share of graduates, age, size, competition, marketing, patent use) as well as 20 industry dummies, 14 regional dummies and an intercept.
Note: $^\#$ the 90% quantile of AI breadth is at value 6, and the 90% (95%) quantile of AI experience is at 11 (17).
Standard errors in parentheses. ***, **, *: significant at p<0.01, <0.05, <0.1.
Data source: German CIS 2018.

For instance, the effect of AI breadth on product innovation is basically an upward-sloping curve until AI breadth reaches the value 6. The value 6 corresponds to the 90% quantile in the data. We therefore conclude that the relationship between AI breath and product innovation is basically positive for the majority of data points. The flat part and especially the negative part of the curve are induced by some extreme values of AI breadth. We refrain from interpreting that too much AI adoption might have negative consequences. The other results on both

breadth and experience are remarkably similar. First, the relationship between AI and innovation performance is statistically significant in the same regressions as in the base table where we used the AI dummy. Generally we find that both more AI adoption in terms of breadth and also the time elapsed since its adoption are positively associated with performance. The inflexion points of the U-shaped curves are remarkably similar in all cases. The only exception is the regression on the likelihood of world-first innovations and experience. There the curve peaks at 18 years which is beyond the 95% quantile of the experience distribution. In all other cases the estimated curves peak at the 90% quantile.

5.4   Macroeconomic extrapolations

The nature of the German CIS as a representative survey based on a random sample allows to calculate total economy estimates based on the estimated contributions of AI to innovation output. Each firm in the sample has been assigned weights that represent the firm's weight in the total number of firms and in total sales of the business enterprise sector in Germany for the firm's strata (combination of sector and size class). These weights are used to estimate the total number of AI using firms in Germany (which is about 17,500) and their total sales (about €1,235bn) as well as the number of AI using firms that have introduced different types of product and process innovation (about 10,623), and the volume of sales for different types of product innovation (about €256bn) as well as the amount of cost reduction from process innovation (about €64bn). The marginal effects estimated for the AI variable are used to then calculate the number of firms that have introduced a certain type of innovation as a result of using AI, as well as the volume of innovative sales and cost reductions that can be assigned to the use of AI. We use marginal effects from the base model, and only consider marginal effects that are statistically significant.

Table 13 reports the results of these calculations. The number of firms that introduced certain types of innovations with the use of AI are non-negligible, but not large. We estimate that almost 1,500 firms introduced new products because of their AI use, all else constant. Out of those, 436 achieved market novelties which are to a large extent also world-first innovators (401 firms). Similarly, almost 1,400 firms could implement new processes that were associated with the use of AI, and 733 firms among those achieved also reductions in unit cost of production.

When looking at the total sales with product innovations that are associated with the use of AI, we calculate that €33.3bn would not have realised in the absence of AI. Compared to the total sales of the AI-using firms (€1,235bn) and their total innovation sales (€256bn) this number is not high though (13% for the latter comparison). These rather low shares reflect that many AI using firms would have innovated also in the absence of applying AI, and that only a small share of innovating firms in Germany are actually using AI.

Table 13: Estimated contribution of AI to innovation output of the German business enterprise sector in 2018 (only statistically significant contributions)

	AI contribution (weighted results[a])	Unit of measure	Share in total innovators (and innovation output) of AI using firms (%)	Share in innovators (and innovation output) of all firms (%)
	(1)		(2)	(3)
Product Innovation (PDI) - total	1,483	k# firms	14.0	1.4
PDI - new-to-market	436	k# firms	12.9	1.5
PDI - world-first innovations	401	k# firms	20.7	3.2
PCI - total	1,396	k# firms	10.7	0.9
PCI - unit cost reduction	733	k# firms	15.9	1.9
Sales with product innovations (SPI)	33.3	bn€	13.0	4.4
SPI - new to the market	21.0	bn€	30.7	11.9
SPI - world-first innovations	16.1	bn€	44.8	18.1
Cost reduction	11.4	bn€	17.7	5.7

a) Firms with 5 or more employees in industries (Nace rev. 2) B to E, 46, H, J, K, 69, 70.2, 71 to 74, 78 to 82.
Source: German CIS 2018.

When looking at world-first innovations -which may reflect the technological frontier in many sectors- the relative contribution of AI is high. For instance, almost 45% of the total world-first innovation sales of AI-using firms are based on AI, and these reflect 18.1% of all world-first sales of German firms which is a substantial share. The relationship between sales of market novelties and AI are also high.

From these extrapolations, we generally conclude that the use of AI seems still expandable but that the contribution of AI in frontier innovations such as market novelties and especially world-first innovations starts to be essential.

6   Conclusions

This paper analysed the extent to which the use of AI contributes to innovation results of firms in Germany. We employed data from the German part of the CIS 2018 which included a number of questions on how and where firms were using AI. We examined the contribution of various AI variables to different dimensions of product and process innovation outcomes. The estimated marginal effects of AI use were then extrapolated to total economy levels, utilising the representative nature of the survey.

We found that AI plays a significant role for introducing innovations and obtaining economic returns from these innovations. AI is particularly relevant for more ambitious product innovations like product innovations that were new to a market. The relatively strongest AI contribution was found for world-first innovations. AI methods are also a major driver for process innovation including cost savings (see Acemoglu and Restrepo 2018).

Firms that developed AI by combining in-house and external resources obtained significantly higher original innovation results, i.e. market and especially world first novelties, than firms that mainly used externally developed AI methods (though the latter is the largest group among AI using firms). Firms that apply AI in a broad way and the have already several years of experience in using AI tend to yield higher innovation outputs.

The results of this paper are a first step to quantify the role of AI for industrial innovation which need to be further developed, extended and broadened by future research. The findings of our paper are limited by the fact that we have to rely on a cross-sectional database. Even though we made an effort to establish that the causality runs from AI to innovation performance (by having a rich set of covariates and by a robustness check looking at future innovation outcomes), we cannot rule out some endogeneity issues. Unfortunately, instrumental variables approaches were not really feasible with the data at hand. Panel data could help in the future to shed more light on causality. Panel data would also enable investigations into the temporal nature of the link between AI and innovation and whether the effects found in this study also hold for other periods in the diffusion of AI. Finally, international comparisons would be useful to evaluate the role of economic framework conditions such as digital infrastructures or availability of specific skills for the role of AI for industrial innovation.

7 Acknowledgments

References

Acemoglu, D., P. Restrepo (2018), *Artificial Intelligence, Automation and Work*, NBER Working Paper 24196, Cambridge, MA.

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

Aghion P., N. Bloom, R. Blundell, R. Griffith, P. Howitt (2005), Competition and innovation: an inverted-U relationship, *Quarterly Journal of Economics* 120, 701–728.

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

Arora, A., L.G. Branstetter, M. Drev (2013), Going soft: how the rise of software-based innovation led to the decline of Japan's IT industry and the resurgence of Silicon Valley, *Review of Economics and Statistics* 95(3), 757–775.

Bag, S., J.H.C. Pretorius, S. Gupta, Y.K. Dwivedi (2020), Role of institutional pressures and resources in the adoption of big data analytics powered artificial intelligence, sustainable manufacturing practices and circular economy capabilities, *Technological Forecasting and Social Change*, 120420.

Baruffaldi, S., B. van Beuzekom, H. Dernis, D. Harhoff, N. Rao, D. Rosenfeld, M. Squicciarini (2020), *Identifying and Measuring Developments in Artificial Intelligence: Making the Impossible Possible, OECD Science*, Technology and Industry Working Papers 2020/05, Paris.

Behrens, V., M. Trunschke (2020), *Industry 4.0 Related Innovation and Firm Growth*, ZEW Discussion Paper No. 20-070, Mannheim.

Behrens, V., M. Berger, M. Hud, P. Hünermund, Y. Iferd, B. Peters, C. Rammer, T. Schubert (2017), *Innovation Activities of Firms in Germany – Results of the German CIS 2012 and 2014. Background Report on the Surveys of the Mannheim Innovation Panel Conducted in the Years 2013 to 2016*, ZEW Documentation No. 17-04, Mannheim.

Bersch, J., S. Gottschalk, B. Müller, M. Niefert (2014), *The Mannheim Enterprise Panel (MUP) and Firm Statistics for* Germany, ZEW Discussion Paper No. 14-104, Mannheim.

Brock, J.K.-U., F. von Wangenheim (2019), Demystifying AI: what digital transformation leaders can teach you about realistic Artificial Intelligence, *California Management Review* 61(4), 110–134.

Brousseau, E., T. Penard (2007), The economics of digital business models: a framework for analyzing the economics of platforms, *Review of Network Economics* 6(2), 81–114.

Brynjolfsson, E., D. Rock, C. Syverson (2017), *Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics,* NBER Working Paper 24001, Cambridge, MA.

Cockburn, I.M., R. Henderson, S. Stern (2018), *The Impact of Artificial Intelligence on Innovation,* NBER Working Paper No. 24449, Cambridge, MA.

Cohen, W.M. (2010), Fifty years of empirical studies of innovative activity and performance, in B.H. Hall, N. Rosenberg (eds.), *Handbook of the Economics of Innovation*, Volume 1, Amsterdam: North-Holland, 129–213.

Cohen, W.M., R. Levin (1989), Empirical studies of R&D and market structure, in R. Schmalensee, R. Willig (eds.), *Handbook of Industrial Organization*, Amsterdam: North-Holland, 1059–1107.

Crépon, B., E. Duguet, J. Mairesse (1998), Research, innovation and productivity: an econometric analysis at the firm level, *Economics of Innovation and New Technology* 7(2), 115–158.

EPO (2017), *Patents and the Fourth Industrial Revolution*, European Patent Office, Munich.

Fujii, H., S. Managi (2017), *Trends and Priority Shifts in Artificial Intelligence Technology Invention: A Global Patent Analysis*, RIETI Discussion Paper Series 17-E-066, Tokyo.

Garbuio, M., N. Lin (2019), Artificial intelligence as a growth engine for health care startups: emerging business models. *California Management Review* 61, 59–83.

George, G., M.R. Haas, A. Pentland (2014), Big data and management, *Academy of Management Journal* 57, 321–332.

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

Haefner, N., J. Wincent, V. Parida, O. Gassmann (2020), Artificial intelligence and innovation management: a review, framework, and research agenda, *Technological Forecast and Social Change* 162, 120392.

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

Haenlein, M., A. Kaplan (2019), A brief history of Artificial Intelligence: on the past, present, and future of Artificial Intelligence, *California Management Review* 61(4), 5–14.

Inaba, T., M. Squicciarini (2017), *ICT: A New Taxonomy Based on the International Patent Classification*, OECD Science, Technology and Industry Working Papers, No. 2017/01, OECD Publishing, Paris.

Lee, J., T. Suh, D. Roy, M. Baucus (2019), Emerging technology and business model, innovation: the case of Artificial Intelligence, *Journal of Open Innovation: Technology, Market, and Complexity* 5(3), 5030044.

Liu, J., H. Chang, J.Y.-L. Forrest, B. Yang (2020), Influence of artificial intelligence on technological innovation: Evidence from the panel data of china's manufacturing sectors, *Technological Forecasting and Social Change* 158, 120142.

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

Mairesse, J., P. Mohnen (2002), Accounting for innovation and measuring innovativeness: an illustrative framework and an application, *American Economic Review* 92(2), 226–230.

Montagnier, P., I. Ek, K. Perset (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), Paris.

Niebel, T., F. Rasel, S. Viete (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, Paris.

Obschonka, M., D.B. Audretsch (2019), Artificial intelligence and big data in entrepreneurship: a new era has begun, *Small Business Economics* 55, 529–539.

OECD (ed.) (2020a), *The Digitalisation of Science, Technology and Innovation: Key Developments and Policies*, OECD Publishing, Paris.

OECD and Eurostat (2018), *Oslo Manual 2018. Guidelines for Collecting, Reporting and Using Data on Innovation*, 4th Edition, OECD Publishing, Paris.

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

Raj, M., R. Seamans (2019), Artificial Intelligence, labor, productivity, and the need for firm-level data, in A.J. Agrawal, J. Gans, A. Goldfarb (eds.), *The Economics of Artificial Intelligence: An Agenda*, Chicago: University of Chicago Press, 553–565.

Rammer, C., D. Czarnitzki, A. Spielkamp (2009), Innovation success of non R&D performers: substituting technology by management in SMEs, *Small Business Economics* 33, 35–58.

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

Rexhäuser, S., C. Rammer (2014), Environmental innovations and firm profitability: unmasking the Porter Hypothesis, *Environmental and Resource Economics* 57, 145–167.

Sedera, D., S. Lokuge, V. Grover, S. Sarker, S. Sarker (2016), Innovating with enterprise systems and digital platforms: a contingent resource-based theory view, *Information & Management* 53(3), 366–379.

Stiebale, J., J. Suedekum, N. Woessner (2020), *Robots and the Rise of European Superstar Firms*, CEPR Discussion Paper No. DP15080, London.

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

Täuscher, K., S.M. Laudien (2018), Understanding platform business models: a mixed methods study of marketplaces, *European Management Journal* 36(3), 319–329.

Trajtenberg, M. (2019), Artificial Intelligence as the next GPT: a political-economy perspective, in A.J. Agrawal, J. Gans, A. Goldfarb (eds.), *The Economics of Artificial Intelligence: An Agenda*, Chicago: University of Chicago Press, 175–186.

Valter, P., P. Lindgren, R. Prasad (2018), Advanced business model innovation supported by artificial intelligence and deep learning, *Wireless Personal Communications* 100, 97–111.

Van Roy, V., Vertesy, D., Damioli, G. (2020), AI and robotics innovation, *Handbook of Labor, Human Resources and Population Economics*, 1-35.

Varian, H. (2018), Artificial intelligence, economics, and industrial organization, *National Bureau of Economic Research,* No. w24839, Cambridge, MA.

Wamba, S.F., A. Gunasekaran, A. Akter, S.J. Ren, R. Dubey, S.J. Childe (2017), Big data analytics and firm performance: effects of dynamic capabilities, *Journal of Business Research* 70, 356–365.

Zhang, D., Mishra, S., Brynjolfsson, E., Etchemendy, J., Ganguli, D., Grosz, B., Lyons, T., Manyika, J., Niebles, J. C., Sellitto, M., Shoham, Y., Clark, J., Perrault, R., (2021), The AI Index 2021 Annual Report, AI Index Steering Committee, Human-Centered AI Institute, Stanford University, Stanford, CA.

8 Appendix

Table 14: Model variables: definition and descriptive statistics

Variable	Definition	Type	Mean	Std.d.	Min.	Max.
Innovation output						
PDI total	Firm has introduced product innovations (new or improved products that differ significantly from the firm's previous products) during 2016 and 2018	D	0.363	0.481	0	1
PDI newfirm	Product innovations 2016-2018 were not new to the market (only new-to-firm)	D	0.268	0.443	0	1
PDI newmarket	Product innovations 2016-2018 were new-to-market	D	0.095	0.293	0	1
PDI newsubm	Product innovations 2016-2018 were new to a regional or sectoral market, but no world first		0.057	0.232	0	1
PDI worldfirst	Product innovations 2016-2018 were world first innovations	D	0.038	0.191	0	1
PDI share_all	Share of sales in 2018 with product innovations introduced during 2016-2018	S	0.514	0.500	0	1
PDI share_ntf	Share of sales in 2018 with product innovations 2016-2018 that were only new-to-firm	S	0.380	0.485	0	1
PDI share_ntm	Share of sales in 2018 with product innovations 2016-2018 that were new-to-market	S	0.134	0.341	0	1
PDI share_nrsm	Share of sales in 2018 with product innovations 2016-2018 that were new to a regional/sectoral market only		0.083	0.184	0	1
PDI share_wof	Share of sales in 2018 with product innovations 2016-2018 that were world first innovations	S	0.069	0.162	0	1
PCI total	Firm has introduced process innovations (new or improved processes that differ significantly from the processes used by the firm before) during 2016 and 2018	D	0.014	0.079	0	1
PCI nocostred	Process innovations 2016-2018 not led to a reduction in unit costs	D	0.008	0.055	0	1
PCI costred	Process innovations 2016-2018 led to a reduction in unit costs	D	0.006	0.054	0	1
PCI sh_costred	Share of unit cost reduction in 2018 through process innovations 2016-2018	S	0.062	0.242	0	1
Artificial intelligence						
AI use	Firm uses AI methods in 2018	D	0.065	0.247	0	1
AI breadth	Number of combinations of AI methods and areas of applications	I	0.176	0.888	0	15
AI in-house	Firm uses AI methods in 2018 that were mainly developed in-house	D	0.298	1.763	0	29
AI others	Firm uses AI methods in 2018 that were mainly developed by others	D	3.196	33.80	0	841
AI both inh/oth	Firm uses AI methods in 2018 that were mainly developed both in-house and by others	D	0.011	0.105	0	1
AI years	Number of years (by the end of 2018) a firm has used AI methods	M	0.034	0.182	0	1
Other digitalisation variables						
Software	Firm conducted in-house programming of software or purchased software programming services or implemented new software packages during 2016 and 2018	D	0.472	0.499	0	1
Databases	Firm set up or maintained own data bases or purchased data bases from others or conducted systematic analysis of big data during 2016 and 2018	D	0.300	0.458	0	1
Platforms	Firm used social web-based networks or crowd sourcing or open source software or open business platforms for acquiring external knowledge during 2016 and 2018	D	0.333	0.471	0	1
Innovation input						
Innoexp_sales	Innovation expenditure in 2018 per sales	M	0.046	0.132	0	1
RD_continuous	Firm conducted R&D activities continuously during 2016 and 2018	D	0.145	0.352	0	1

Variable	Definition	Type	Mean	Std.d.	Min.	Max.
RD_occasional	Firm conducted R&D activities occasionally during 2016 and 2018	D	0.098	0.297	0	1
Patent use	Firm used patent to protect its intellectual property during 2016 and 2018	D	0.094	0.292	0	1
Capabilities						
Graduates	Share of employed persons in the firm that hold a university degree in 2018	S	0.237	0.278	0	1
Age	Age of the firm in midyear 2018 (years, log)	L	3.093	0.866	-0.693	6.50
Size	Number of employed persons in the firm in 2018 (annual average at full-time equivalents, log)	L	2.845	1.143	-0.693	10.64
Marketing	Marketing expenditure in 2018 per sales, relative to industry average	M	0.959	2.959	0	121
Market						
Competition	Sum of the evaluation of eight characteristics of the firm's competitive environment (3: applies fully, 2: applies somewhat, 1: applies very little, 0: does not apply): Products become outdated quickly; The technological development is difficult to predict; Products / services from competitors are easily substituted for those of your enterprise; Major threat to market position because of entry of new competitors; Competitor's actions are difficult to predict; Demand development is difficult to predict; Strong competition from abroad; Price increases lead to immediate loss of clients	I	10.86	4.43	0.000	24
Nace_10-12	Firm operates in Nace divisions 10, 11 or 12 (food, beverages, tobacco)	D	0.049	0.216	0	1
Nace_13-15	Firm operates in Nace divisions 13, 14 or 15 (textiles, clothing, leather)	D	0.007	0.084	0	1
Nace_16-17_31	Firm operates in Nace divisions 16, 17 or 31 (wood, paper, furniture)	D	0.015	0.122	0	1
Nace_20-21	Firm operates in Nace divisions 20 or 21 (chemicals, pharmaceuticals)	D	0.008	0.090	0	1
Nace_22	Firm operates in Nace division 22 (rubber, plastics)	D	0.017	0.128	0	1
Nace_23	Firm operates in Nace division 23 (non-metallic mineral products)	D	0.012	0.108	0	1
Nace_24-25	Firm operates in Nace divisions 24 or 25 (metals, metal products)	D	0.073	0.260	0	1
Nace_26-27	Firm operates in Nace divisions 26 or 27 (electronics, instruments, electrical equipment)	D	0.026	0.159	0	1
Nace_28	Firm operates in Nace division 28 (machinery and equipment)	D	0.035	0.183	0	1
Nace_29-30	Firm operates in Nace divisions 29 or 30 (automotive, other vehicles)	D	0.007	0.084	0	1
Nace_32-33	Firm operates in Nace divisions 32 or 33 (other consumer products, repair/installation)	D	0.051	0.219	0	1
Nace_5-9_19_35	Firm operates in Nace divisions 5-9, 19, 35 (mining, petroleum, energy supply)	D	0.008	0.092	0	1
Nace_36-39	Firm operates in Nace divisions 36, 37, 38 or 39 (water supply, sewerage, waste disposal)	D	0.016	0.126	0	1
Nace_46	Firm operates in Nace division 46 (wholesale trade)	D	0.135	0.341	0	1
Nace_49-53	Firm operates in Nace divisions 49, 50, 51, 52, or 53 (transportation and storage)	D	0.113	0.316	0	1
Nace_18_58-60	Firm operates in Nace divisions 18, 58, 59 or 60 (printing, publishing, film, broadcasting)	D	0.026	0.159	0	1
Nace_61-63	Firm operates in Nace divisions 61, 62 or 63 (telecommunications, IT/information services)	D	0.060	0.237	0	1
Nace_64-66	Firm operates in Nace divisions 64, 65 or 66 (financial and insurance services)	D	0.021	0.144	0	1
Nace_71-72	Firm operates in Nace divisions 71 or 72 (architecture, engineering, R&D services)	D	0.072	0.259	0	1
Nace_69-70_73	Firm operates in Nace divisions 69, 70 or 73 (legal, consulting, advertising services)	D	0.116	0.320	0	1
Nace_74_78-82	Firm operates in Nace divisions 74, 78, 79, 80, 81 or 82 (other business services)	D	0.132	0.338	0	1
Nace_other	Firm operates in any other Nace division[a]	D	0.001	0.038	0	1
State_SH	Firm located in the state of Schleswig-Holstein	D	0.018	0.132	0	1
State_HA	Firm located in the state of Hamburg	D	0.012	0.110	0	1
State_LS	Firm located in the state of Lower Saxony	D	0.048	0.213	0	1
State_BR	Firm located in the state of Bremen	D	0.012	0.110	0	1
State_NW	Firm located in the state of Northrhine-Westfalia	D	0.102	0.303	0	1
Sate_HE	Firm located in the state of Hesse	D	0.036	0.186	0	1

Variable	Definition	Type	Mean	Std.d.	Min.	Max.
State_RP	Firm located in the state of Rhineland-Palatine	D	0.023	0.149	0	1
State_BW	Firm located in the state of Baden-Wuerttemberg[b]	D	0.184	0.387	0	1
State_BV	Firm located in the state of Bavaria	D	0.082	0.274	0	1
State_SR	Firm located in the state of Saar	D	0.008	0.089	0	1
State_BE	Firm located in the state of Berlin[b]	D	0.087	0.282	0	1
State_BB	Firm located in the state of Brandenburg[b]	D	0.110	0.313	0	1
State_ME	Firm located in the state of Mecklenburg-West Pomerania	D	0.022	0.147	0	1
State_SX	Firm located in the state of Saxony[b]	D	0.178	0.382	0	1
State_ST	Firm located in the state of Saxony-Anhalt	D	0.034	0.182	0	1
State_TH	Firm located in the state of Thuringia	D	0.044	0.205	0	1

D: dummy, I: index, L: logarithmic value, M: metric value, S: share. Std.d.: standard deviation; Min.: minimum value; Max.: maximum value.
a) Firms from these industries are not part of the random sample of the German CIS. b) Firms from these states are oversampled to allow for data evaluation at the level of these states.
Source: German CIS 2018.