The Diffusion of Artificial Intelligence: New evidence from German Online Job Vacancy data

Myrielle Gonschor[1,2] and Eduard Storm[1]

[1]RWI - Leibniz Institute for Economic Research

[2]Heinrich-Heine Universität Düsseldorf

February 2023

Abstract

This is the first paper documenting the diffusion of Artificial Intelligence (AI) across German firms and labor markets. Drawing on novel job vacancy data with access to original texts, we employ natural language process techniques to identify AI demand based on firm's skill requirements. We find the share of firms demanding AI skills has increased from 3% in 2017 to 4.8% in 2021 with substantial heterogeneity subject to firm characteristics, job location, and industry. Subsequently, we link our AI measures to administrative data to study the impact of rising AI demand on local employment and wages. We find no effects on employment. Yet, a 10 pp. increase in AI demand is associated with a wage increase of 1%. Contrary to previous automation technologies, our findings suggest AI has stronger wage than employment effects.

Keywords: Artificial Intelligence, Online Job Vacancies, Labor Demand, Local Labor Markets, Employment, Wages

JEL Codes: D22, J23, J24, J31, O33

1 Introduction

The growing diffusion of artificial intelligence (AI) technology has sparked a debate among economists regarding its potentials and limitations. AI stands out from previous automation technologies due to its increased level of autonomy, especially in tasks related to prediction and recommendation (Abrardi, Cambini & Rondi 2022). Because of this, some economists view AI as a general purpose technology (GPT) that has the potential of enhancing productivity across a wide range of occupations and industries (Brynjolfsson, Mitchell & Rock 2018, Agrawal, Gans & Goldfarb 2019, Cockburn, Henderson & Stern 2019). This perspective suggests that AI's superior prediction abilities can lead to better decisions for both firms and workers, thereby contributing to economic growth (Aghion, Jones & Jones 2019). Others, however, point to Solow's Paradox, i.e. a slowdown in productivity growth despite recent advancements in digital technologies (Gordon 2018). According to this view, AI's capabilities are limited to specific tasks, thereby limiting its overall impact on the economy compared to previous GPTs. In such a scenario, AI would merely be another "so-so" technology that automates some tasks previously performed by humans and thus displaces workers (Acemoglu & Restrepo 2019).

Despite AI being a nascent and specialized technology, analyzing its effect on labor markets is crucial for at least two reasons. First, a rapidly growing number of firms and workers are exposed to this technology. For example, ZEW (2022) report the share of AI-adopting firms in Germany has increased from 6% in 2019 to 10% in 2021 based on firm-level survey data. Moreover, a greater number of workers are potentially exposed to AI. According to individual-level German survey data from 2019, Giering, Fedorets, Adriaans & Kirchner (2021) find that up to 45% of workers already engage with AI technologies, though unbeknownst to more than half of them. Similarly, using firm-level survey data from the US, Acemoglu, Anderson, Beede, Buffington, Childress, Dinlersoz, Foster, Goldschlag, Haltiwanger, Kroff, Restrepo & Zolas (2022) show only 3% of US firms had adopted AI by 2019, though almost 13% of US workers have been exposed to this technology at work. This discrepancy is largely attributed to the fact that AI is primarily concentrated among large firms (Babina, Fedyk, He & Hodson 2021, Rammer, Fernández & Czarnitzki 2022). Second, because of its enhanced ability of autonomous learning, early evidence suggests AI may have different labor market effects than previous automation technologies, e.g., robots and software (Webb 2020, Gathmann & Grimm 2022). In contrast to previous technologies, which were viewed as substitutes to routine tasks typically performed by workers in the middle of the income spectrum, AI technology has the potential to replace workers in non-routine tasks, which could impact workers higher up the income distribution.

A young, yet growing, literature has analyzed the labor market effects of AI. Some studies use online job vacancy (OJV) data to study demand for AI skills and find the impact of AI to be concentrated at the establishment-level with negligible effects at more aggregated levels.[1] Others combine OJV with firm-level data (Alekseeva, Azar, Giné, Samila & Taska 2021, Babina, Fedyk, He & Hodson 2021, Bloom, Hassan, Kalyani, Lerner & Tahoun 2021) or use survey data (ZEW 2022) and usually find AI adoption to be concentrated in large, productive firms and specific industries, especially IT. More closely related to our paper, Gathmann & Grimm (2022) study the exposure and diffusion of AI using patent data and find that past exposure to AI is associated with positive employment and wage effects at the regional level. Most of the related literature focus on the US and analyze potential outcomes derived from firm's innovation and job search behavior. However, the existing literature provides limited insight about *realized* labor market outcomes. This paper aims to fill this gap linking OJV data to administrative data, allowing us to study how the rising demand for AI skills impacts local employment and wages.

In the first part of the paper, we use natural language processing (NLP) methods to identify AI skills from the near-universe of German OJV between 2017 - 2021. Our rich data gives us access to original texts and thus allows us to develop our own taxonomies. Taking advantage of this data feature, we provide a more comprehensive picture by differentiating

[1] See, for example, Alekseeva, Azar, Giné, Samila & Taska (2021), Bessen, Cockburn & Hunt (2021), Acemoglu, Autor, Hazell & Restrepo (2022), Goldfarb, Taska & Teodoridis (2023).

among distinct categories of AI skills. Specifically, we distinguish between AI tools, applications, and methods. Subsequently, we present novel stylized facts on demand for AI skills at both, the firm- and vacancy-level. Notably, the share of vacancies requiring AI skills has increased from 1.1% in 2017 to 1.7% in 2021. This increase has primarily been driven by business related services and manufacturing, where the share of AI vacancies has reached 3% by 2021. Similarly, the share of firms requiring AI skills in job postings has increased from 3% in 2017 to 4.8% in 2021. We inspect these trends more closely in various firm-level regressions and find that the increase in "AI firms" is driven by (i) younger firms, (ii) larger firms, (iii) firms demanding AI skills associated with Machine Learning (ML) methods and Robotics, (iv) firms in large metropolitan areas, and (v) firms in business related services and manufacturing.

While AI demand has broadly diffused across regions and industries, strong discrepancies still exist. We exploit this variation in the second part of the paper, where we aggregate differences in AI demand at the industry-region level. Defining local labor markets (LLMs) at this level, we merge our AI measures onto administrative data from the Institute of employment research (IAB). This strategy allows us to account for industry-specific uses of AI and, via employment shares, exposure to AI.

Our results indicate that the relationship between local AI demand and local wages is stronger than its relationship with employment. Specifically, we find no significant association between local AI demand and employment levels. However, we observe a positive association between AI demand and higher wages. Accordingly, a 10 pp increase in AI demand leads to an increase in wages by about 1%, though this effect is limited to AI skills pertaining to methods and algorithms such as Machine Learning and Deep Learning. Our results broadly align with prior evidence from the US, suggsting no significant link between increased AI demand and local employment growth.

Our paper contributes to several strands of the literature. First, it contributes to the literature on the local labor market effects of new technologies. Several studies analyze the local impact of specific technologies, e.g. robots (Acemoglu & Restrepo 2019, Dauth, Findeisen, Suedekum & Woessner 2021), software (Autor & Dorn 2013, Atalay, Sotelo & Tannenbaum 2021), or AI (Gathmann & Grimm 2022). These studies typically find mixed results, which is to some extent country-specific. For example, the impact of robots has negative employment effects in the US (Acemoglu & Restrepo 2019), yet, more muted effects in Germany (Dauth et al. 2021). Aside from Gathmann & Grimm (2022), all of these studies focus on older automation technologies associated with the ICT revolution that began in the 1980s. Our research adds to this literature by studying regional differences in demand for AI and their impact on local employment and wages. Gathmann & Grimm (2022) is most closely related to our paper, yet, they measure AI exposure via patents. In our view, our measures have two advantages. First, using OJV data allows us to directly link AI demand to specific firms, regions and industries. In contrast, using patents as a proxy provides insights on potential AI adoption, yet, this measure stays silent on *realized* AI adoption. Second, our AI measure is broader and more representative since most firms post vacancies, but not all firms are innovators.

Second, we contribute to the literature on regional skill differences (Hershbein & Kahn 2018, Deming & Kahn 2018, Modestino, Shoag & Ballance 2019). These studies often use OJV data and find large regional skill differences, often accompanied by "upskilling" patterns, i.e., firms raise skill requirements in the aftermath of recessions. More closely related to our paper are a series of papers studying factors related to AI demand. This literature finds AI does not display aggregate labor market effects yet. Instead, any effects are typically concentrated at the firm-level (Alekseeva, Azar, Giné, Samila & Taska 2021, Acemoglu, Autor, Hazell & Restrepo 2022), in part because so far AI adoption is concentrated among large, productive firms and specific industries, especially IT (Babina, Fedyk, He & Hodson 2021, Rammer, Fernández & Czarnitzki 2022). While all these papers offer important insights on the labor market effects of AI, they study potential labor market outcomes. Merging OJV data to administrative data, we instead study realized labor market outcomes by exploring how employment and wage outcomes change in LLMs with high AI demand. Moreover, we intent to instrument regional differences for variation in AI demand via past adoption of robots and software. Thus, we can also contribute to the existing literature by adding important insights on underlying mechanisms of AI adoption.

2 Data

2.1 Online Job Vacancies

We use the near-universe of online job vacancies posted in Germany between January 2017 and December 2022 to measure demand for AI skills. The job postings are collected by our private partner, a firm that is offering custom-made firm-, person- and job posting-data and market analysis. Our partner scrapes more than 2,000 web-pages for vacancies from the following platforms: (i) job boards (fee paying), (ii) job boards (free of charge), (iii) company websites, (iv) temporary employment agencies, and (v) head-hunters. They consistently update their online sources and scrape all sources on a daily basis. Subsequently, our partner performs some basic cleaning procedures and removes duplicates from the same source (i.e. sources from the same url address).

Compared to conventional OJV data used in the literature, which is usually provided by commercial providers, our data has two unique features. First, we have access to the original job vacancies, including all text included in the posting. This unique access allows us to have complete control over the data-generating process and to develop our own, transparent taxonomies. In contrast, data packages purchased from commercial providers are typically already pre-processed, diminishing researchers influence on data quality. Second, our partner merges posting firms with the German company registry ("Handelsregister"). This merge is successful for about 60% of firms and allows us to supplement vacancy contents with rich information information on firms.[2] Notable firm-level information we receive is the industry

[2]The data set is based on information from the trade register and includes all firms that are listed in

affiliation (5-digit level, WZ08), location at the zip-code level, sales, employee size, founding year, and even some basic information on the founders (including sex, function, and date of birth).

Upon receiving the data from our provider, we perform further cleaning procedures. First, we link firm and vacancy information, especially in order to assign job descriptions to specific industries. Second, we use this linked data set to assign each vacancy to a specific location, preferably at the zip-code level.

For our analysis, we focus on vacancies for regular work, i.e. full- or part-time. Thus, we remove vacancies seeking apprenticeships, trainees, and other types of irregular work. In particular, we drop vacancies for temporary employment as they are likely not representative of regular labor market developments (Stops, Bächmann, Glassner, Janser, Matthes, Metzger & Müller, Christoph, Seitz, Joachim 2021).[3] After cleaning and selecting the relevant data, we are left with 22 million job vacancies, comprising 240,000 firms and 2.1 million firm-month observations. In a final step, we perform a few more standard preprocessing steps on the job description. Specifically, we follow Gentzkow, Kelly & Taddy (2019) and preprocess the text data for the empirical analysis by (i) converting job descriptions with tokenization, (ii) removing stop words, and (iii) stemming words.

We identify AI skills by combining a keyword-based approach with manual annotation. First, we create a comprehensive keyword list of AI skills, using keywords that have previously been used in the literature.[4] We restrict our keyword search to the *skills* and *profile* section

the German trade register since 1991. About half of the 3,4 Mio. firms in Germany are noncommercial and therefore not listed in the trade register. In addition, firms from the public administration sector are not included. The firm level data includes information about the firm name, the complete address, legal status, industry, original stock and business volume, the number of employees and the formation date. The data can be merged through a firm identifier, which is available for about 60% of the job postings. Reasons why the firm identifier is not available are, on the one hand, that firms are not listed in the German trade register, or, on the other hand, because group of companies cannot be assigned to one specific firm.

[3] Temporary employment agencies are special in the sense that their postings may be counter-cyclical: If labor demand is small, they may increase the number of persons in their applicant pool, and publish less postings if labor demand is high in the labor market. Therefore, job vacancies of temporary employment agencies distort demand for labour demand and show patterns that are incompatible with official statistics.

[4] Specifically, we combine keywords that have been used by Büchel, Demary, Goecke, Kohlisch, Koppel, Mertens, Rusche, Scheufen & Wendt (2021), Bessen, Cockburn & Hunt (2021), Taska, O'Kane & Nania (2022). Especially Büchel, Demary, Goecke, Kohlisch, Koppel, Mertens, Rusche, Scheufen & Wendt (2021)

of vacancies to avoid counting vacancies that merely mention relevant keywords, but where AI skills are not part of the job. Second, we identify vacancies that include at least one AI skill and draw a random sample of 500 vacancies. Third, we manually annotate this random sample to verify that the identified vacancies indeed include AI skills. Moreover, we adjust our keyword list by adding further keywords we identified in the vacancies. This third step is important to avoid counting "buzzwords" or to exclude words or abbreviations that are used in areas not related to AI. Fourth, we train a training sample based on our manual annotation and subsequently apply the algorithm to a test sample. For our baseline analysis, we differentiate between three categories of AI skills: AI application, AI methods and AI tools. For parts of the analysis we broaden this definition by also including skills pertaining to supplementary AI technologies and common programming languages.

In section C.1 in Appendix C we provide external validity on our data quality by comparing trends by time and industries to official job vacancy statistics. Overall, we demonstrate our OJV data depicts similar trends between 2017 - 2021 and covers all industries properly.

2.2 Robot and Software Data

To approximate the technological frontier of LLMs, we use two automation technologies predating AI. First, we employ industry-level data on software adoption from the EU KLEMS & INTANProd database, combining the widely used EU KLEMS productivity database with information on intangible investments from INTAN-Invest.[5] This data not only updates previous releases of EU KLEMS, but also places a greater emphasis on intangible assets, making it a suitable database for our study. Specifically, we use data from the *DE capital accounts* database, comprising the EUR value of software for 58 industries. Data on German industries is available from 1995 - 2019.

Second, we employ industry-data on robot adoption from the International Federation

is helpful for us as this is the only comprehensive keyword list with German and English keywords, to our knowledge.

[5]For details of the data see EUKLEMS & INTANProd (2021). The data can be downloaded free of charge from the EUKLEMS & INTANProd website: `https://euklems-intanprod-llee.luiss.it/download/`

of Robotics (IFR).[6] This data provides information on the stock of industrial robots at the 2-digit level for manufacturing industries and at 1-digit level for the other industries. The IFR consolidates information collected from robot suppliers worldwide and harmonizes the data for more reliable cross-country comparisons. Data on German industries is available from 1993 - 2020, though we only use data from 1995 onward to match availability with software data from EU KLEMS & INTANProd.

In order to combine both databases, we need to harmonize IFR and EU KLEMS & INTANProd as they have differential aggregation procedures and thus distinct industry definitions. In particular, software is general-purpose technology employed heavily in all industries. Its adoption is thus well-documented across relatively fine industries. In comparison, robots are primarily used in manufacturing and related sectors but rarely in other industries. Accordingly, robot adoption is broken down across specific sectors within the broad manufacturing industry, but with little to no disaggregation in other industries.

For these reasons we harmonize industry definitions to increase reliability of our technology measures. As a result, we end up with 41 industries. A complete list of all industries can be found in Table 1 in Appendix B. In our empirical analysis we employ a shift-share framework to apportion these industry-level technology measures to LLMs based on local industry composition. Section 4 provides details on this approach.

Figure 1 highlights the distinct adoption of robots and software, where we aggregated industries into 8 broad sector for visual clarity. Between 1995 and 2016, robot exposure strongly increased in the production of food products and consumer goods, industrial goods, and capital goods, with an average yearly growth rate of adopted units ranging between 3.5 and 7.1% in these sectors. In contrast, the value of of software has increased across all industries, though, the growth rate is most pronounced in business related services with an average yearly growth rate of about 6%.

[Figure 1 here]

[6] For details of the data see Dauth, Findeisen, Suedekum & Woessner (2021), Bachmann, Gonschor, Lewandowski & Madon (2022)

2.3 Administrative Data

For our analysis on labor market effects of AI we use the Sample of Integrated Labor Market Biographies (SIAB), a 2 percent representative sample of administrative data on all workers who are subject to social security contributions (SSC) and all workers receiving unemployment benefits for the period 1975-2019.[7] The SSC requirement excludes certain individuals, notably the self-employed and civil servants. The SIAB is drawn from the Integrated Employment Biographies (IEB) of the Institute for Employment Research (IAB) and provides information about daily labor market spells, wages, and basic socio-economic characteristics (e.g., sex, nationality, education).[8] The cutoff date of firm characteristics is June 30th of each year. This information is important as we collapse our monthly OJV data to annual variables using the same cutoff date prior to merging them to the SIAB.

Besides the long sample period and information on daily labor market outcomes, the SIAB provides detailed information on establishment's industry affiliation at the 3-digit level. This granular data is essential for our study as we need reliable and detailed industry information in order to apportion industry-level technology measures to LLMs based on local industry shares. Moreover, the SIAB offers information on both, employee's place of residence and place of work, at the county level. This information allows us to account for spillover effects in case workers travel across LLMs to get to work.

As is common in administrative data, wage information is top-censored. Censoring affects about five percent of all spells between 1975-2017, though, some skilled groups are more heavily affected (Dauth & Eppelsheimer 2020).[9] To provide remedy and not disregard this data, we follow standard procedures and use the imputations for education and wages provided by the IAB-FDZ, which builds upon Fitzenberger, Osikominu & Völter (2006). In

[7] See vom Berge, Frodermann, Schmucker, Seth, Graf, Grießemer, Kaimer, Köhler, Lehnert, Oertel & Schneider (2019) for a detailed description of the data.

[8] If there are parallel employment spells for one individual, we only consider the employment spell with the highest pay.

[9] For example, Dauth & Eppelsheimer (2020) report that up to 44% of spells of regularly employed men with college degree are affected with an increasing trend over time.

terms of sample selection, we focus on workers aged 18-65 who are liable to social security and exclude workers with (i) zero wage and wages below the first percentile and (ii) missing information on place and industry. For our main analysis, we restrict the samples to the years 1995 to 2019 to match data availability of our technology measures on software and robots.

2.4 Regional Data and Local Labor Market Definition

We choose to perform our analysis at a broader definition than county-level. Counties have administrative boundaries that do not necessarily reflect LLMs in an economic context. For instance, counties do not account for common commuting zones. Disregarding these movements may introduce spillovers and thus bias our results. We therefore aggregate the 402 counties into 141 broader LLM, following the classification of Kosfeld & Werner (2012), which has been used widely in research on LLMs in Germany (Dauth, Findeisen, Suedekum & Woessner 2021, Dustmann, Lindner, Schönberg, Umkehrer & vom Berge 2021, Hirsch, Jahn, Manning & Oberfichtner 2022).

We moreover supplement our labor market and technology data with regional characteristics to account for systematic differences between LLMs that may confound our analysis on demand for AI skills. These variables are taken from a regional administrative database, Regionalstatistik.de (Regionalstatistik 2022), and comprise the unemployment rate and GDP at the county-level. Consistent with our definition of LLMs, we aggregate these variables for our 141 regions.

3 Descriptive Statistics: AI skills in Germany

In this section, we provide new stylized facts on the diffusion of AI skills in Germany from 2017 - 2021. First, we focus on the vacancy-level by describing our AI skill categories and their diffusion over time, industries, and LLMs. This exercise is especially insightful as we construct our proxies for the diffusion of AI skills via changes in the share of AI vacancies at the industry-region level over time. These proxies are subsequently merged to administrate data to perform our analysis on the LLM effects of AI in section 4.2. Second, we present stylized facts at the firm-level, describing trends over time, industries, LLMs, and by firm characteristics. This exercise gives a broad overview of the importance of AI skills across firms and serves as the foundation for our subsequent firm-level analysis in section 4.1.

3.1 Vacancy-level analysis

In this paper, we distinguish between three different types of AI skills regarding AI (i) tools, (ii) methods, and (iii) applications. Figure 2 displays a word cloud, illustrating the most relevant keywords in our search using all three skill categories. Figure 3 provides more detailed insights by breaking down the relative importance of each keyword for our three skill groups. AI tools comprise specific libraries and tools that specialists use in order to perform AI-related tasks. Among the most important AI tools are TensorFlow and PyTorch, two common software libraries used to perform AI-related tasks. AI methods summarizes algorithms and methods commonly deployed in AI, most notably machine learning techniques, predictive analytics and deep learning methods. Lastly, AI applications comprise specific domains in which AI skills are applied to. The most important domains are in robotics and autonomous driving.

We make this distinction as different AI skills likely apply to different people and industries. We assume AI tools apply primarily to IT-specialists, while AI applications and methods apply to generalists in various production processes and are also more common in industries outside of IT. In our baseline analysis we pool all types of AI skills together, though, to broaden the scope of our analysis. Overall, the most common type of AI skills in our data are skills regarding AI methods and algorithms. We discuss their prevalence in more detail in section 3.2 where we present our firm-level analysis.

For parts of the analysis we broaden these definitions and further include AI (iv) supplementary technologies and (v) languages. Supplementary technologies typically relate to IT-infrastructure necessary to perform some AI-related tasks while languages include the most common programming languages used by workers performing AI tasks. Figure 4 displays word clouds for these concepts.

[Figure 5 here]

Figure 5 illustrates the evolution of AI demand from 2017/01 - 2021/12. Panel 5a displays the monthly share of AI vacancies, i.e. vacancies with at least one AI skill. The share of AI vacancies increased from 1.1% in 2017 to 1.7% in 2021, implying a year-on-year (YoY) growth rate of 8.7%.[10] Panel 5b shows these trends are primarily driven by an increase in the importance of AI methods as the share of AI vacancies requiring these skills increased from 0.5% in 2017 to 1.2% in 2021. In comparison, the trend for AI tools and applications has remained flat since around 2018.

These findings broadly align with Taska, O'Kane & Nania (2022) who report the share of AI vacancies in Germany increased from 0.6% in 2017 to about 1% by 2021. While our evidence suggests the share of AI vacancies has been about twice as large, we also use a different and broader set of keywords, among others by including German translations of English concepts. Once we adopt their list of keywords (excluding German translations), the share of OJV vacancies merely increases from 0.5% to 1%, thus consistent with their findings (see Figure 24 in Appendix C). This comparison therefore highlights the importance of including keywords in the country-specific native language to account for different semantic variations of the same concept.

To better understand the key drivers of this development, we identify the most important keywords associated with AI skills and plot these over time. Figure 6 illustrates this exercise and shows the single most important AI skills are pertaining to *Artificial Intelligence*, *Machine Learning*, and *Robotics*. While the frequency of *Artificial Intelligence* mentions

[10] The YoY growth rate of vacancies requiring skills on AI tools is 10.5%, for AI applications -2.6%, and for AI methods 24.8%.

increased 10-fold between 2017 - 2021, the frequency of *Machine Learning* and *Robotics*, respectively, tripled and doubled over time.

[Figure 6 here]

A chief interest in our paper are differences regarding industry- and region-level demand for AI skills as we exploit variation along these dimensions for parts of our empirical analysis. Figure 7 affirms this rationale by underlining substantial differences in AI demand across eight broad industries. Overall, the share of AI vacancies has fluctuated between 0.5 - 1% for most industries. However, in business related services and capital goods & repair (sub-industry of manufacturing) the diffusion of AI vacancies has accelerated since 2017. In both of these industries, the share of AI vacancies has nearly doubled from about 1.5% to up to 3% over the past five years. These findings are intuitive considering the rising adoption of AI and ML techniques in both industries and the importance of AI skills pertaining to robotics.

[Figure 7 here]

Regarding spatial differences, Figure 8 displays the average share of AI vacancies for each of our 141 LLMs. As one would expect, AI vacancies tend to be clustered in large metropolitan areas, especially in Southern Germany around Stuttgart and Munich. Many of the regions with a relatively high share of AI vacancies in 2017 (Panel 8a) also experienced disproportionate growth between 2017 - 2021 compared to the national average (Panel 8b). For example, the average share of AI vacancies in a region at the 75th percentile (such as Heilbronn, near Stuttgart) was 1.6%. By 2021, this share increased to 2.2%. On the other hand, the average share of AI vacancies in a region at the 25th percentile (such as Altötting, Southeast Bavaria) was 0.5%. By 2021, this share decreased to 0.3%. This simple comparison suggests persistent gaps in demand for AI skills across LLMs.

[Figure 8 here]

3.2 Firm-level analysis

[Figure 14 here]

We define an AI firm as a firm that posts at least one AI vacancy a year, i.e. vacancies requiring AI skills. The most common type of AI skills demanded by firms are skills regarding methods and algorithms, see Table 2, displaying the demand for AI skills for AI and Non-AI firms. Among AI firms, 24% of all posted vacancies require these type of skills. In comparison, 17% (3%) of all posted vacancies among these firms require skills regarding AI applications (tools). The technical nature of these skills is also reflected in the job profile of AI firms. While AI and Non-AI firms have a broadly similar share of vacancies requiring interactive, routine, and manual tasks, AI firms have a substantially larger share of vacancies requiring analytical skills. More broadly speaking, AI and Non-AI firms differ along several firm characteristics. On average, AI firms are slightly younger, post about four times as many vacancies in a given month, employ four times as many workers, and generate revenues more than twice as large as Non-AI firms. These stylized facts corroborate with the existing literature, showing AI to be disproportionately adopted by large firms with higher sales (Babina, Fedyk, He & Hodson 2021, Rammer, Fernández & Czarnitzki 2022)

[Table 2 here]

Panel 9a displays the monthly share of AI firms. The share of AI firms increased from about 3.0% in 2017 to 4.8% in 2021, implying a YoY growth rate of 9.6%.[11] Panel 9b shows a breakdown of these aggregate trends and highlights the increase in the share of AI vacancies and firms is attributed to the rising importance of AI methods and, to a lesser extent, AI tools.

[Figure 9 here]

[11]The YoY growth rate for firm requiring skills on AI tools is 10.5%, for AI applications 0%, and for AI methods 23.1%.

How do this stylized fact corroborate with other firm- and OJV-level evidence? Regarding firm-level evidence, Rammer, Fernández & Czarnitzki (2022) and ZEW (2022) suggest the share of German firms adopting AI technologies increased from 5.8% in 2019 to up to 11% in 2021.[12] In contrast, Genz, Gregory, Janser, Lehmer & Matthes (2021) use a broader definition, reporting that 22% of German firms adopted 4.0 technologies by 2016. Therefore, our descriptive evidence is broadly consistent with Rammer, Fernández & Czarnitzki (2022) and ZEW (2022), suggesting we do indeed measure AI skills rather than some broad skill groups that are tangentially related to AI. Quantitative differences regarding the share of AI firms are unsurprising giving methodological differences (self-reported survey data versus posted OJV data) and different definitions. For example, once we broaden our definition of AI skills by also including more tangential skills pertaining to supplementary technologies and programming languages, our share of AI firms rises to around 13%, thus more in line with Rammer, Fernández & Czarnitzki (2022) and ZEW (2022).

A closer look at firm-level statistics reveals important patterns on AI demand subject to firm characteristics. For example, Figure 10 illustrates the evolution of AI demand by age groups. This comparison shows that the share of AI firms is higher among young firms. Among younger firms of up to 27 years, the share of AI firms increased from about 3% in 2017 to 5% in 2021. Importantly, however, the gap to older firms has shrunk over time as the share of AI firms among older firms doubled from 2.5% in 2017 to 5% in 2021. Consequently, old and young firms have about the same share of AI vacancies by the end of 2021.

[Figure 10 here]

With respect to firm size, large firms are more likely to be AI firms than small firms. Defining size based on employment shows that among large firms the share of AI firms increased from 4% in 2017 to 7% in 2021 (Panel 11a). In comparison, among smaller firms,

[12]Note that the share of AI firms drops to 1-3% after applying the taxonomy of Taska, O'Kane & Nania (2022), which excludes German translation of certain AI concepts. This comparison reaffirms the importance of including keywords in the country-specific native language to account for different semantic variations of the same concept. See Figure 24 in Appendix C.

the share of AI firms remained flat at around 3%. The gap in AI vacancies has thus increased over the past four years, suggesting rising firm inequality regarding AI demand in favor of large firms. Similar takeaways emerge from a comparison of firm size based on revenue (Panel 11b). Among firms with high revenues, the share of AI firms increased from from 3.5% in 2017 to 5.5% in 2021. In contrast, among firm with lower revenues, the share of AI firms increased slightly from from 2.5% in 2017 to 3% in 2021. Hence, our stylized facts for Germany are consistent with the US-centric OJV literature, documenting AI has so far been primarily adopted by large firms with high revenues (Alekseeva, Azar, Giné, Samila & Taska 2021, Babina, Fedyk, He & Hodson 2021), and the survey-based evidence on Germany (Rammer, Fernández & Czarnitzki 2022).

[Figure 11 here]

Two key characteristics for our study are (i) regional and (ii) industry-specific discrepancies in AI demand as we exploit variation along these dimension for our empirical analysis. Figure 13 displays differences in the share of AI firms in 2017 and subsequent changes between 2017 - 2021. Overall, this figure aligns broadly with the OJV-level analysis in section 3.1 in which we demonstrated persistent differences across LLMs regarding demand for AI skills. Compared to the OJV-level analysis, however, a firm-level analysis suggests somewhat less persistent effects.

[Figure 13 here]

To see this, consider Figure 12 in which we plot the average share of AI firms for different percentiles. The percentage point (pp) difference in the share of AI firms between regions at the 95th percentile and regions at the 5th percentile has remained constant at about 7.5 pp. between 2017 - 2021. More broadly speaking, however, there has been a convergence among structurally stronger and weaker regions, indicated, for example, by a declining gap in the interquartile range. That is, the difference in AI firms for regions at the 75th percentile and the 25th percentile has decresed from 7.5 pp. in 2017 to 5 pp. in 2021. Hence, while persistent differences remain, this comparison suggests a broad diffusion of AI across most LLMs.

[Figure 12 here]

Figure 14 likewise illustrates substantial gaps in the share of AI firms across industries. For reasons of visual clarity, we consider eight broad industries. Analogous to our stylized facts at the OJV-level, the share of AI firms is especially pronounced in business related services and capital goods & repair. In both industries, the share of AI firms increased from about 5% in 2017 to 7.5% by 2021. Similarly, the share of AI firms also increased for most of the other industries, though more muted, ranging from 1 - 3% over the past five years.

[Figure 14 here]

4 Methodology

In this section, we outline our empirical approach. First, we conduct a firm-level analysis to study sources of variation in AI across firms. Second, we perform analysis at the LLM-level to explore the labor market effects of AI. To this end, we merge proxies for AI diffusion at the industry-region level to administrative data, allowing us to study the impact of varying degrees of diffusion across LLMs on employment and wage outcomes. In the current manuscript, we employ an ordinary least squares regression for our empirical analysis. In a subsequent version of the paper, we intend to augment our analysis with an instrumental variables approach to address the issue of endogeneity inherent in the adoption of AI.

4.1 Firm-level analysis

We begin our empirical analysis at the firm-level. Our outcome variable that we want to explain is the change in AI postings at firm-level. To this purpose, we calculate the cumulative change in the number of postings, $\Delta AI_{ijl}^{(t+n)-t}$, of firm i between the first year of appearance t and the last year of appearance $t+n$. The subscripts j and l, respectively, denote industries and LLMs. A downside of this approach is that we measure firm's cumulative change in AI vacancies for different time horizons, limiting comparability. We choose to follow this approach, however, as restrictions on common time horizons cause a loss of many observations. Since many firms do not post vacancies regularly. In a robustness check, we thus restrict our analysis to firms that post vacancies in all years from 2017 to 2021, implying a loss of 58% of observation. We weigh the change in AI vacancies by the total number of firm-specific postings in their respective baseline year to account for differences in posting behaviour across firms.

In an alternative specification, we construct $\Delta AI_{ijl}^{(t+n)-t}$ via the change in the share of AI vacancies of firm i. This measure illustrates the intensity with which firms post AI vacancies more directly. We estimate the following model:

$$\begin{aligned}\Delta AI_{ijl}^{(t+n)-t} =& \beta_1 \Delta Size_{ijl}^{2017} + \beta_2 \Delta Age_{ijl}^{2017} + \beta_3 \Delta Qual_{ijl} \\ &+ \beta_4 \Delta Comp_{ijl}^{2017} + \beta_5 \Delta Conc_{ijl}^{2017} + \beta_6 \Delta Profile_{ijl}^{t} \\ &+ \gamma X_l^{2017} + \gamma Ind_j + \epsilon_{ijl}\end{aligned} \tag{1}$$

where we regress our outcome variable ΔAI_{ijl} on a set of firm-level characteristics. $Size_{ijl}^{2017}$ comprises information on a firm's employment and revenue, respectively, each in binned form as of 2017. Age_{ijl}^{2017} conditions on initial firm age, while $Qual_{ijl}$ controls for quality differences across firms, via a dummy, indicating whether the firm has turned insolvent during our time horizon. $Comp_{ij/l}^{2017}$ describes the baseline competition of a firm for workers with AI skills in the year 2017. We measure competition in two ways. On the one hand, we calculate the average number of AI vacancies within a firm's 5-digit industry in the baseline year ($Comp_{ij}^{2017}$). On the other hand, we calculate the average number of AI vacancies within a LLM ($Comp_{il}^{2017}$). A comparison of both competition measures thus allows us to draw conclusions whether competition for workers with AI skills at the industry-or region-level affects firm i's demand for AI skills.[13] $Conc_{ijl}^{2017}$ conditions on various concentration measures. We compute the Herfindahl–Hirschman Index (HHI) at the (2-digit) industry-region level to account for concentration of (i) employment, (ii) revenue, (iii) all job postings, and (iv) AI vacancies across all job-posting firms.

$Profile_{ijl}^{t}$ approximates differences in job profiles. To this end, we calculate the firm-specific task intensity by analyzing task requirements in a firm's first year of appearance. For our task classification, we follow literature and classify job activities into five broad tasks: (i) Non-Routine (NR) Analytic, (ii) NR Interactive, (iii) Routine Cognitive, (iv) Routine Manual, and (v) NR Manual.[14] We also condition on a rich set of local controls via X_l^{2017} to account for structural differences across regions.[15] Lastly, we include industry FE (at the 3-digit level) via Ind_j to account for industry-specific effects. As is common, ϵ_{ijl} represents an i.i.d. error term.

4.2 Local Labor Market analysis

This study also seeks to explore whether there is already a discernible influence of AI demand, as measured by AI job postings, on labor market outcomes at the local level. Given the data limitations of the SIAB database, the analysis is constrained to the period between 2017 and 2019. Specifically, we conduct our analysis on a set of 5,781 LLMs defined by the pairing of 41 different industries with 141 local regions. These LLMs are established in a manner similar to that found in existing literature, which characterizes labor market concentration using occupation × region pairs, as previously discussed by Azar, Marinescu, Steinbaum & Taska (2020) and Schubert, Stansbury & Taska (2020).

First, we use a level-level specification to analyse the relationship between our outcomes

[13] Acemoglu, Autor, Hazell & Restrepo (2022) construct a similar variable, though, in their case at the occupation-level. This way, they address the role of competition for AI workers at the occupation-level.

[14] This classification has been used by many studies in the literature, e.g., Autor, Levy & Murnane (2003), Spitz-Oener (2006), Storm (2022)

[15] Specifically, we account for the following local differences: (i) Skill composition (share of workers with college, high school, neither), (ii) Share of female workers, (iii) Share of foreign workers, (iv) Age composition (via seven age bins), (v) Employment share by (13) broad industries, and (vi) Unemployment rate.

and AI demand at LLM-level for the years 2017 to 2019. Our specification is as follows:

$$y_{jlt} = \beta AI_{jlt} + \gamma X_{jlt} + \delta_j + \delta_t + \delta_l + \epsilon_{jlt} \tag{2}$$

where y_{jlt} measures our labour market outcomes of interest, employment and log wages at LLM-level for the years $t = 2017, 2018, 2019$. The key regressor is AI_{jlt} measuring AI demand for each year t in each LLM. We approximate AI exposure in a LLM by calculating the share of all vacancies in a LLM that require AI skills (= AI vacancies). We run the specification for our broad AI measure and for each of its sub-categories: AI tools, methods and applications. X_{jlt} comprises a rich set of controls to account for confounding factors. Our controls are: LLM composition by age groups, gender, nationality, and education. In addition, we include year δ_t and industry group δ_j fixed-effects and include regional dummies δ_l to account for regional variation across Germany.

Next, for our baseline specification we follow Dauth, Findeisen, Suedekum & Woessner (2021) and conduct a First-Difference (FD) estimation at LLM-level. Our main specification is as follows:

$$\Delta y_{jl}^{19-17} = \beta \Delta AI_{jl}^{19-17} + \gamma X_{jl}^{2017} + +\delta_j + \delta_l + \epsilon_{jl} \tag{3}$$

where Δy_{jl}^{19-17} denotes our outcome variable: the growth in LLM employment and log wages between 2017-19. [16] We use controls X_{jl}^{2017} from the baseline year $t = 2017$ to avoid contamination due to endogenous adjustments. As for the level specification we control for the composition of the LLM by age, gender, nationality, and education. We also include industry group δ_j fixed-effects and regional indicators δ_l to control for systematic regional and industry differences in changes in AI demand.

One key concern with model (3) is endogeneity as AI-adopting firms likely settle in more urban regions with a relatively skilled workforce, as suggested by the directed technological

[16] Similar AI measures have been used in the existing literature, e.g. Alekseeva, Azar, Giné, Samila & Taska (2021).

change literature (Acemoglu 2002). Moreover, AI demand is somewhat industry-specific (see Figure 7). Therefore, demand for AI is likely predetermined by the local industry composition. In a subsequent version of the paper, we intend to account for these concerns by adopting a shift-share instrument, thus following Dauth, Findeisen, Suedekum & Woessner (2021). To this end, we will instrument contemporaneous AI demand with past robot and software adoption in LLMs, i.e. technologies predating AI:

$$\Delta \widehat{T}_{jl}^{2016-1995} = \underbrace{\frac{E_{j,l,t=1995}}{E_{l,t=1995}}}_{\text{Share}} \times \underbrace{\frac{\Delta T_{j}^{2016-1995}}{E_{j,t=1995}}}_{\text{Shift}} \tag{4}$$

where the term $\widehat{T}_{jl}^{2016-1995}$ on the left hand side measures the past change in previous technologies T (e.g., robots, software) for each LLM between the years 1995 and 2016. The first term on the right hand side is the *share* component, reflecting the initial employment share of industry j in local labor market l. The second term is the *shift* component, reflecting the nation-level increase in the value (€) of previous technologies T (robots, software) between 1995 - 2016 relative to the workforce size of industry j in 1995. Using past robot and software adoption as instrument for current AI demand permits causal interpretation of the LLM consequences of AI demand. Our methodology therefore solves an econometric problem and at the same time sheds light on channels giving rise to heterogeneous AI demand.

5 Results

5.1 Firm-level analysis

5.1.1 Baseline

Our main firm-level results are summarized in Figure 15, where we display the coefficients associated with the firm-level characteristics included in eq. (1). We also report 95% confidence intervals along with the number of firms included in our cross-sectional analysis and adjusted R-squared. The key determinants of AI demand are associated with initial firm size. Note we report coefficients with narrowly defined employment bins, allowing us to identify thresholds for firm size at which AI demand becomes more pronounced.

Regarding employment, we find a first threshold for firms employing about 275 workers. Firms of this size have had about two additional AI vacancies in subsequent years compared to smaller firms. A second threshold arises at a workforce of around 750. Firms of this size have posted an additional ten AI vacancies while firms surpassing a workforce of 3,000 even posted an additional 40 AI vacancies in the years following their first appearance in our data. Regarding revenue, we only find a meaningful threshold at about 550,000 annual sales (at the 10% level).

[Figure 15 here]

Combined, our insights are consistent with the repeated finding that AI adoption is pronounced among large firms wigh higher sales (Alekseeva, Azar, Giné, Samila & Taska 2021, Babina, Fedyk, He & Hodson 2021, Rammer, Fernández & Czarnitzki 2022). Overall, employment size is a more important determinant of AI demand than revenue, however. A novel aspect of our analysis is that we present specific thresholds at which firm size begins to matter pertaining to demand for AI.

Aside from the firm size measures, the remaining firm characteristics have limited explanatory power on subsequent demand for AI skills. Unsurprisingly, firm quality is positively correlated with demand for AI. Firms that have become insolvent at sometime in our sample have had about two AI postings less than surviving firms. Of course, this result is largely mechanical, as insolvent firms stop posting vacancies. Otherwise, competition matters as well, however, only with modest economic significance. The economic effects are stronger with respect to industry-specific competition, though imprecisely estimated and only significant at the 10% level. Accordingly, firms in industries with, on average, one more AI vacancy in 2017 have posted one more AI vacancy in subsequent years than firms in other industries. In comparison, firms in regions with, on average, one more AI vacancy in 2017 have posted 0.3 more AI vacancies in subsequent years than firms in other regions. The remaining initial firm characteristics (age, concentration, and job profile) are not correlated with a subsequent increase in AI demand.

We repeat above exercise, but replace the outcome variable with the change in the *share* of AI vacancies. Our previous outcome variable, i.e., change in AI vacancies, captures the absolute importance of AI skills at firm-level. This modified specification, however, explores the determinants of the relative importance of AI vacancies by explicitly accounting for trends in Non-AI vacancies as well. Figure 16 summarizes the results of this exercise. Overall, coefficients are estimated less precisely in this specification, rendering some coefficients insignificant. Notably, however, size (in terms of employment) remains a key determinant of the subsequent change in AI vacancies. The largest firms, with more than 3,002 employees, experienced a increase in the share of AI vacancies by 1 pp. after 2017 compared to other firms.

[Figure 16 here]

5.1.2 Heterogeneity

In this section we explore whether AI vacancies respond uniformly to initial firm-characteristics depending on the specific AI skill requirements. In our baseline analysis we define AI vacancies as those that contain skill requirements pertaining to AI tools, methods, and applications. Now, we distinguish between AI vacancies containing either of those skill groups specifically.

Figures 17 - 19 summarize our results from this exercise for AI methods, applications, and tools each. These graphs use the change in AI vacancies as outcome variable. Overall, our baseline results carry over to specifications defining AI vacancies as those requiring either AI methods, tools or applications. Using AI methods and employment as proxy for size, larger firms have had about 5 - 15 additional AI vacancies in subsequent years compared to smaller firms. Using revenue as proxy for size instead, larger firms have had about 3 - 10 additional AI vacancies in subsequent years compared to smaller firms. These numbers are broadly similar using AI applications , though estimated less precisely. We do not find any meaningful results using AI tools due too much noise in the estimates.

[Figures 17 - 19 here]

We repeat this exercise, but using the change in the *share* of AI vacancies as outcome variable. Compared to our baseline results in section 5.1.1, the main results carry over, especially pertaining to the importance of initial firm size.[17]

5.1.3 Robustness

A valid concern with our methodology is different time horizons for our outcome variables. In our baseline setting, we measure the change in AI vacancies and share of AI vacancies over time from the first to the last year of appearance in our data. Due to differential posting behavior, our baseline sample thus contains outcome variables whose change ranges from two to five years. In a robustness check, we impose common time horizons by calculating the change of AI vacancies and share of AI vacancies between 2017 - 2021. This requires firms to have posted vacancies both in 2017 and 2021. Though harmonizing time horizons for our outcome variables, this constraint removes 58% of firms in our sample.

[Figure 20 here]

The results of this exercise are summarized in Figure 20, where we use our baseline measure for AI vacancies, containing either AI tools, applications, or methods. We find that our main results are virtually unaffected by this robustness check. Similarly, our baseline results carry over when we define the outcome variable as the change in the *share* of AI vacancies (see Figure 21).

[Figure 21 here]

[17]The results are available from the authors upon request.

5.2 Local Labor Market analysis

Table 3 displays the findings from the level specification model, as specified in eq. (2), for employment. Notably, the coefficient of AI demand is statistically insignificant for both the overall measure and the three sub-categories. Thus, we find no evidence for a significant relation between the current AI demand and employment at the LLM level, indicating that the local demand for AI does not appear to have immediate effects on local employment outcomes. However, the low adjusted R-squared indicates the presence of factors affecting local employment outcomes that are not yet accounted for in the model.

[Table 3 here]

On the other hand, Table 4 depicts the outcomes from the level model specification model of log wages on AI demand. The results reveal that local AI demand is positively linked to higher wages at the LLM level. This relationship is statistically significant for the overall AI measure and the two subcategories of AI applications and tools. Nevertheless, the relationship is not significant for the AI methods. Moreover, the significantly higher adjusted R-squared value of 0.67 suggests that the proposed empirical model is adequately capturing wage variations among LLMs.

[Table 4 here]

Table 5 presents the outcomes of our baseline specification, described by eq. (3), for employment growth. Consistent with the results of the level specification, we do not find any significant effect of AI diffusion on employment growth during the period of 2017 to 2019.

[Table 5 here]

Regarding wages, Table 6 reveals a slightly positive and statistically significant effect of AI methods on wage growth. We find that a 10 percentage points increase in the share of vacancies demanding AI methods is related to an increase in wages by about 1%. However, the coefficient for AI tools and applications is negative, albeit statistically insignificant. As a result, the overall coefficient for the broad measure of AI is negative and insignificant. Consequently, we can infer that the change in AI demand has yet to impact the development of employment and wages at the local level.

[Table 6 here]

Due to data constraints, we can analyse changes solely over a short time span, i.e., 2017-2019, which might be too short-term for the labour market to adjust to the new technological environment. Furthermore, although the demand for AI has increased significantly, as outlined in Section 3, the proportion of AI-related vacancies may still be insufficient to have a discernible impact at the LLM level.

6 Conclusions

In this paper we study the diffusion of AI demand across Germany by using data from the near-universe of online job vacancies from 2017 - 2021. Our original data gives us access to the raw texts from job postings, allowing us to construct our own AI taxonomy in a transparent fashion. To this end, we employ natural language processing techniques to identify different types of AI skills by virtue of firms' job descriptions.

In the first part of the paper we employ firm-level analysis and present novel facts on the diffusion of AI in Germany. Overall, demand for AI skills has been on the rise. The share of vacancies requiring AI skills has increased from 1.1% in 2017 to 1.7% in 2021. Similarly, the share of firms demanding AI skills has increased from 3% in 2017 to 4.8% in 2021. These overarching trends mask substantial heterogeneity, however. We further document that demand for AI skills is especially pronounced among (i) younger firms, (ii) larger firms, (iii) firms demanding AI skills associated with machine learning (ML) methods and robotics, (iv) firms in large metropolitan areas, and (v) firms in business related services and manufacturing. A key takeaway from our firm-level analysis reveals persistent and substantial differences in AI demand across industries and regions.

In the second part of the paper we exploit this variation across industries and regions to study the impact of AI demand on local employment and wages between 2017 - 2019. To this end, we merge our AI measures onto administrative data at the industry-region level. Our results indicate that a 10 pp increase in demand for AI skills results in additional wage growth by up to 1 %. Regarding employment, we find no significant results yet. The latter result may be attributed to the fact that our analysis spans only three years. Hence, job transitions and aggregated employment effects as a result of varying degrees of AI skills may have not materialized yet. This interpretation is in line with evidence from the US (Acemoglu, Autor, Hazell & Restrepo 2022). Alternatively, our results may suffer bias from endogeneity concerns and other measurement errors.

In our upcoming revised draft we will tackle these challenges in more detail. To this end, we will address endogeneity directly by isolating the causal effects of AI exposure via instrumental variable (IV) estimation. Our instrument relies on a shift-share design in which we predict contemporaneous AI demand based on past adoption of technology, namely robots and software. This approach assumes that adoption of AI technologies builds upon previous automation technologies.

Moreover, we will elaborate on other mechanisms that may give rise to heterogeneous AI demand. Next to differences in labor market institutions, such as varying coverage of collective bargaining agreements, and public policies, such as local programs aimed at inducing adoption of AI technologies, we will pay special attention to concentration of AI vacancies. Inspired by our observation that AI vacancies are concentrated among large firms and metropolitan areas, these firms may use AI technologies to exert monopsony power. Related research from the US has shown higher concentration of job postings is negatively correlated with wages, suggesting such concentration measures are meaningful measures of employer power (Azar, Marinescu, Steinbaum & Taska 2020, Schubert, Stansbury & Taska

2020). Our new methodology will therefore solve an econometric problem and at the same time shed light on channels giving rise to heterogeneous AI demand.

References

Abrardi, L., Cambini, C. & Rondi, L. (2022), 'Artificial intelligence, firms and consumer behavior: A survey', *Journal of Economic Surveys* **36**(4), 969–991. 1

Acemoglu, D. (2002), 'Technical change, inequality, and the labor market', *Journal of Economic Literature* **40**(1), 7–72. 21

Acemoglu, D., Anderson, G., Beede, D., Buffington, C., Childress, E., Dinlersoz, E., Foster, L., Goldschlag, N., Haltiwanger, J. C., Kroff, Z., Restrepo, P. & Zolas, N. J. (2022), 'Automation and the workforce: A firm-level view from the 2019 annual business survey', *SSRN Electronic Journal* . 1

Acemoglu, D., Autor, D., Hazell, J. & Restrepo, P. (2022), 'Artificial intelligence and jobs: Evidence from online vacancies', *Journal of Labor Economics* **40**(S1), S293–S340. 2, 4, 19, 27

Acemoglu, D. & Restrepo, P. (2019), 'Automation and new tasks: How technology displaces and reinstates labor', *Journal of Economic Perspectives* **33**(2), 3–30. 1, 4

Aghion, P., Jones, B. F. & Jones, C. I. (2019), 'Artificial intelligence and economic growth', *The Economics of Artificial Intelligence* . 1

Agrawal, A., Gans, J. S. & Goldfarb, A. (2019), 'Artificial intelligence: The ambiguous labor market impact of automating prediction', *Journal of Economic Perspectives* **33**(2), 31–50. 1

Alekseeva, L., Azar, J., Giné, M., Samila, S. & Taska, B. (2021), 'The demand for ai skills in the labor market', *Labour Economics* **71**, 102002. 2, 4, 16, 20, 22

Arthur, R. (2021), 'Studying the uk job market during the covid-19 crisis with online job ads', *PloS one* **16**(5), e0251431. 60

Atalay, E., Sotelo, S. & Tannenbaum, D. (2021), 'The geography of job tasks', *SSRN Journal* . 4

Autor, D. H. & Dorn, D. (2013), 'The growth of low-skill service jobs and the polarization of the US labor market', *American Economic Review* **103**(5), 1553–1597. 4

Autor, D. H., Levy, F. & Murnane, R. J. (2003), 'The skill content of recent technological change: An empirical exploration', *The Quarterly Journal of Economics* **118**(4), 1279–1333. 19

Azar, J., Marinescu, I., Steinbaum, M. & Taska, B. (2020), 'Concentration in us labor markets: Evidence from online vacancy data', *Labour Economics* **66**, 101886. 19, 27

Babina, T., Fedyk, A., He, A. X. & Hodson, J. (2021), 'Artificial intelligence, firm growth, and product innovation', *mimeo* . 2, 4, 14, 16, 22

Bachmann, R., Gonschor, M., Lewandowski, P. & Madon, K. (2022), 'The impact of robots on labour market transitions in europe', *IZA DP No. 15303* . 8

Bamieh, O. & Ziegler, L. (2020), 'How does the covid-19 crisis affect labor demand? an analysis using job board data from austria', *IZA DP No. 13801* . 60

Bessen, J., Cockburn, I. & Hunt, J. (2021), 'Is distance from innovation a barrier to the adoption of artificial intelligence?', *mimeo* . 2, 6

Bloom, N., Hassan, T. A., Kalyani, A., Lerner, J. & Tahoun, A. (2021), 'The diffusion of disruptive technologies'. 2

Brynjolfsson, E., Mitchell, T. & Rock, D. (2018), 'What can machines learn and what does it mean for occupations and the economy?', *AEA Papers and Proceedings* **108**, 43–47. 1

Büchel, J., Demary, V., Goecke, H., Kohlisch, E., Koppel, O., Mertens, A., Rusche, C., Scheufen, M. & Wendt, J. (2021), 'Ki-monitor 2021: Status quo der künstlichen intelligenz in deutschland', *Report* . 6

Cockburn, I., Henderson, R. & Stern, S. (2019), 'The impact of artificial intelligence on innovation', *The Economics of Artificial Intelligence* . 1

Dauth, W. & Eppelsheimer, J. (2020), 'Preparing the sample of integrated labour market biographies (siab) for scientific analysis: a guide', *Journal for Labour Market Research* **54**(1). 9

Dauth, W., Findeisen, S., Suedekum, J. & Woessner, N. (2021), 'The adjustment of labor markets to robots', *Journal of the European Economic Association* **19**(6), 3104–3153. 4, 8, 10, 20, 21

Deming, D. & Kahn, L. B. (2018), 'Skill requirements across firms and labor markets: Evidence from job postings for professionals', *Journal of Labor Economics* **36**(S1), S337–S369. 4

Dustmann, C., Lindner, A., Schönberg, U., Umkehrer, M. & vom Berge, P. (2021), 'Reallocation effects of the minimum wage', *The Quarterly Journal of Economics* **137**(1), 267–328. 10

EUKLEMS & INTANProd (2021), 'Euklems industry accounts with intangibles: The euklems & intanprod productivity database: Methods and data description', *Data Manual* . 7

Fitzenberger, B., Osikominu, A. & Völter, R. (2006), 'Imputation rules to improve the education variable in the iab employment subsample', *Journal of Contextual Economics* **126**(3), 405–436. 9

Forsythe, E., Kahn, L. B., Lange, F. & Wiczer, D. (2020), 'Labor demand in the time of covid-19: Evidence from vacancy postings and ui claims', *Journal of public economics* **189**, 104238. 60

Gathmann, C. & Grimm, F. (2022), 'The diffusion of digital technologies and its consequences in the labor market', **mimeo**. 2, 4

Gentzkow, M., Kelly, B. & Taddy, M. (2019), 'Text as data', *Journal of Economic Literature* **57**(3), 535–574. 6

Genz, S., Gregory, T., Janser, M., Lehmer, F. & Matthes, B. (2021), 'How do workers adjust when firms adopt new technologies?', *IZA DP No. 14626* . 15

Giering, O., Fedorets, A., Adriaans, J. & Kirchner, S. (2021), 'Künstliche intelligenz in deutschland: Erwerbstätige wissen oft nicht, dass sie mit ki-basierten systemen arbeiten'. 1

Goldfarb, A., Taska, B. & Teodoridis, F. (2023), 'Could machine learning be a general purpose technology? a comparison of emerging technologies using data from online job postings', *Research Policy* **52**(1), 104653. 2

Gordon, R. J. (2018), 'Declining american economic growth despite ongoing innovation', *Explorations in Economic History* **69**, 1–12. 1

Gürtzgen, N., Lochner, B., Pohlan, L. & van den Berg, G. J. (2021), 'Does online search improve the match quality of new hires?', *Labour Economics* **70**, 101981. 60

Hensvik, L., Le Barbanchon, T. & Rathelot, R. (2021), 'Job search during the covid-19 crisis', *Journal of public economics* **194**, 104349. 60

Hershbein, B. & Kahn, L. B. (2018), 'Do recessions accelerate routine-biased technological change? evidence from vacancy postings', *American Economic Review* **108**(7), 1737–1772. 4, 59, 61

Hirsch, B., Jahn, E. J., Manning, A. & Oberfichtner, M. (2022), 'The urban wage premium in imperfect labor markets', *Journal of Human Resources* **57**(S), S111–S136. 10

Kosfeld, R. & Werner, A. (2012), 'Deutsche arbeitsmarktregionen – neuabgrenzung nach den kreisgebietsreformen 2007–2011', *Spatial Research and Planning* **70**(1). 10

Modestino, A. S., Shoag, D. & Ballance, J. (2019), 'Upskilling: Do employers demand greater skill when workers are plentiful?', *The Review of Economics and Statistics* **4**, 1–46. 4, 61

Rammer, C., Fernández, G. P. & Czarnitzki, D. (2022), 'Artificial intelligence and industrial innovation: Evidence from german firm-level data', *Research Policy* **51**(7), 104555. 2, 4, 14, 15, 16, 22

Regionalstatistik (2022), 'Regionaldatenbank deutschland'.
URL: *https://www.regionalstatistik.de/genesis/online/* 10

Rengers, M. (2018), 'Internetgestützte erfassung offener stellen: Machbarkeitsstudie im rahmen eines essnet-projekts zu big data', *WISTA - Wirtschaft und Statistik* **5**. 59

Schubert, G., Stansbury, A. & Taska, B. (2020), 'Monopsony and outside options', *SSRN Journal* . 19, 27

Shen, K. & Taska, B. (2020), 'Measuring the impacts of covid-19 on job postings in australia using a reweighting-estimation-transformation approach', *Australian Journal of Labour Economics* **23**(2), 153–171. 60

Spitz-Oener, A. (2006), 'Technical change, job tasks, and rising educational demands: Looking outside the wage structure', *Journal of Labor Economics* **24**(2), 235–270. 19

Stops, M., Bächmann, A.-C., Glassner, R., Janser, M., Matthes, B., Metzger, L.-J. & Müller, Christoph, Seitz, Joachim (2021), 'Extracting skill requirements from job ads - the machbarkeitsstudie kompetenz-kompass', *IAB-Forschungsbericht No. 7/2021* . 6

Storm, E. (2022), 'Task specialization and the native–foreign wage gap', *LABOUR* **36**(2), 167–195. 19

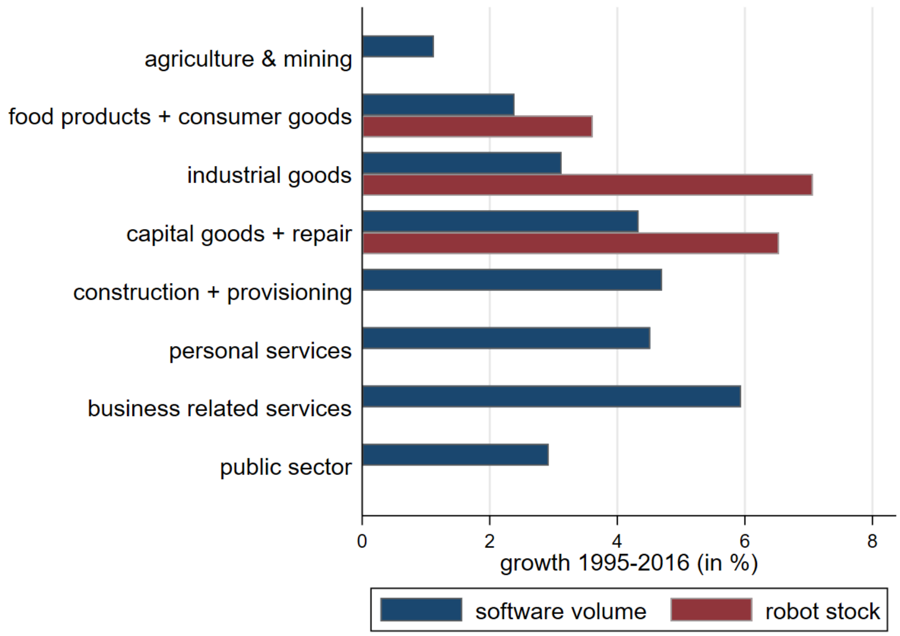
Taska, B., O'Kane, L. & Nania, J. (2022), 'Artificial intelligence in the uk: The relevance of ai in the digital transformation of the uk labour market', *Report* . 6, 12, 15, 63

vom Berge, P., Frodermann, C., Schmucker, A., Seth, S., Graf, T., Grießemer, S., Kaimer, S., Köhler, M., Lehnert, C., Oertel, M. & Schneider, A. (2019), 'Schwach anonymisierte version der stichprobe der integrierten arbeitsmarktbiografien (siab) – version 7519 v1'. 9

Webb, M. (2020), 'The impact of artificial intelligence on the labor market', *SSRN Electronic Journal* . 2

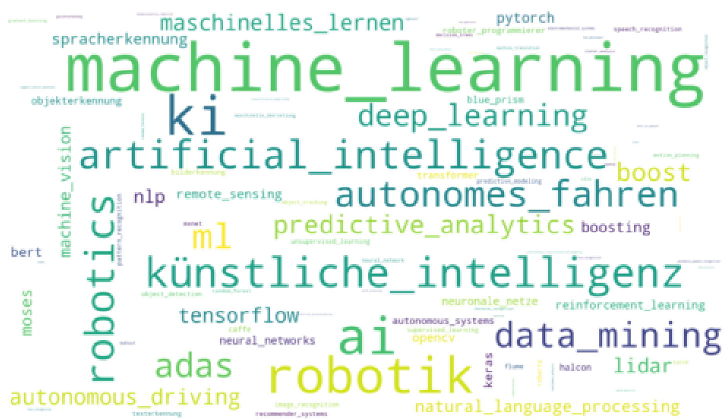
ZEW (2022), 'Kompetenzen und kooperationen zu künstlicher intelligenz: Ergebnisse einer befragung von ki-aktiven unternehmen in deutschland', *Report* . 1, 2, 15

A Figures



NOTE. —The chart displays the growth of software and robot adoption in Germany between 1995 and 2016 for eight broad industries. Growth of software is measured in terms of EUR, while growth of robots is measured in terms of units.

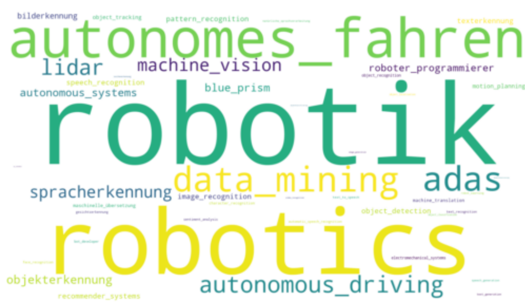
Figure 1: Expansion of software value and robots, 1995-2016



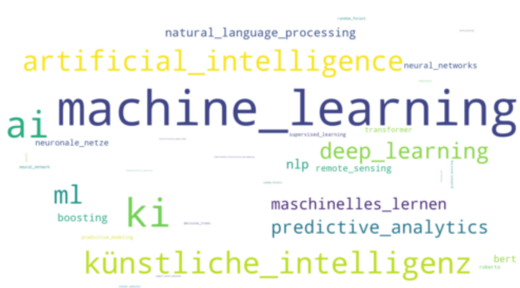
NOTE. —This word cloud comprises keywords that are associated with AI tools, applications, or domains.

Figure 2: Word clouds of AI Skills: Baseline definition

(a) AI Tools

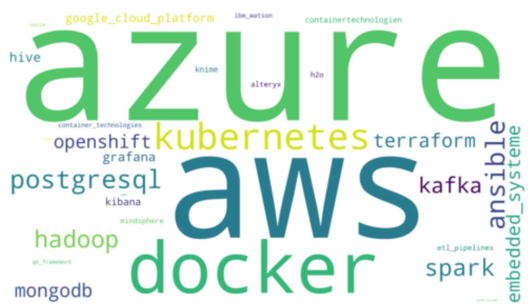


(b) AI Applications



(c) AI Methods

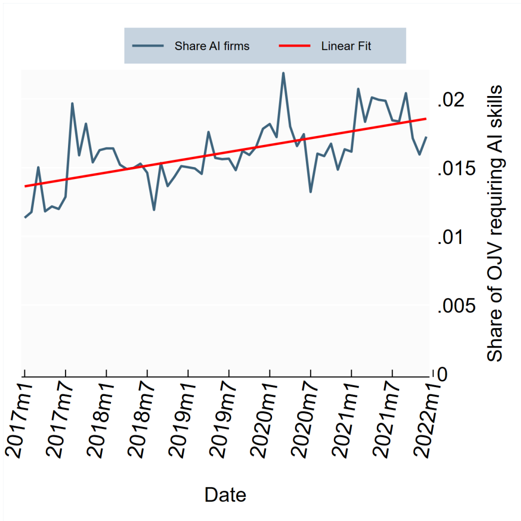
Figure 3: Word clouds of AI skills by category



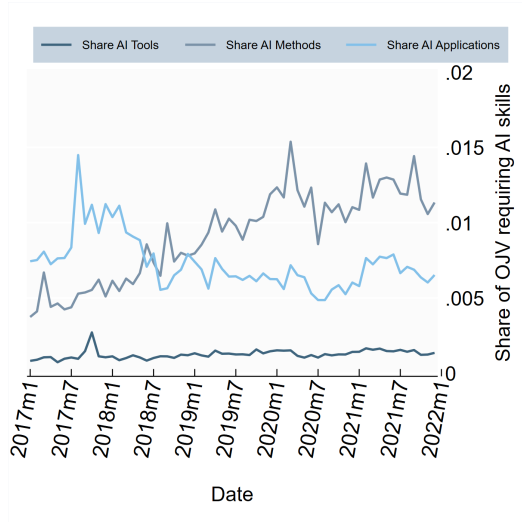
(a) AI Supplementary Skills

(b) AI Languages

Figure 4: Word clouds of skills related to AI



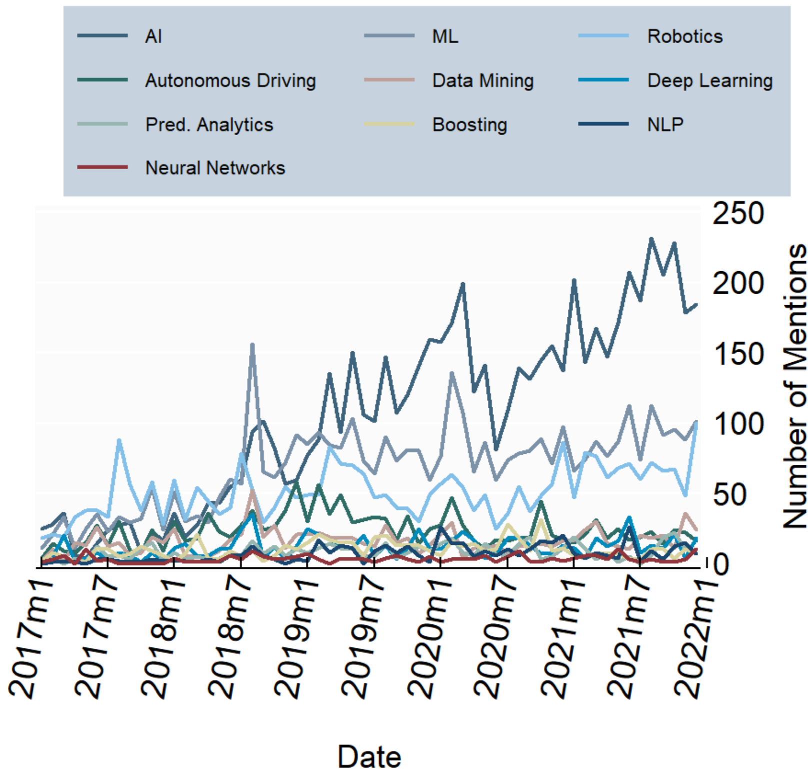
(a) Baseline definition



(b) Breakdown of baseline definition by AI categories

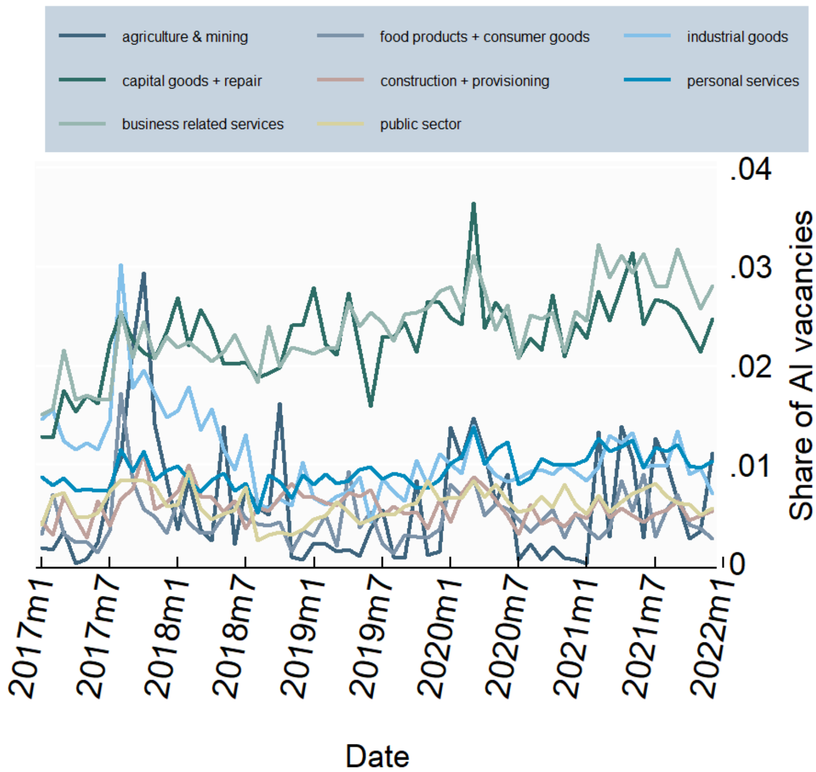
NOTE.—Vacancies are defined as an "AI vacancy" if a job posting contains at least one AI-related skill in a given month. Our baseline definitions comprises skills related to AI tools, methods, and applications.

Figure 5: Trends in AI demand at OJV-level, 2017/01 - 2021/12



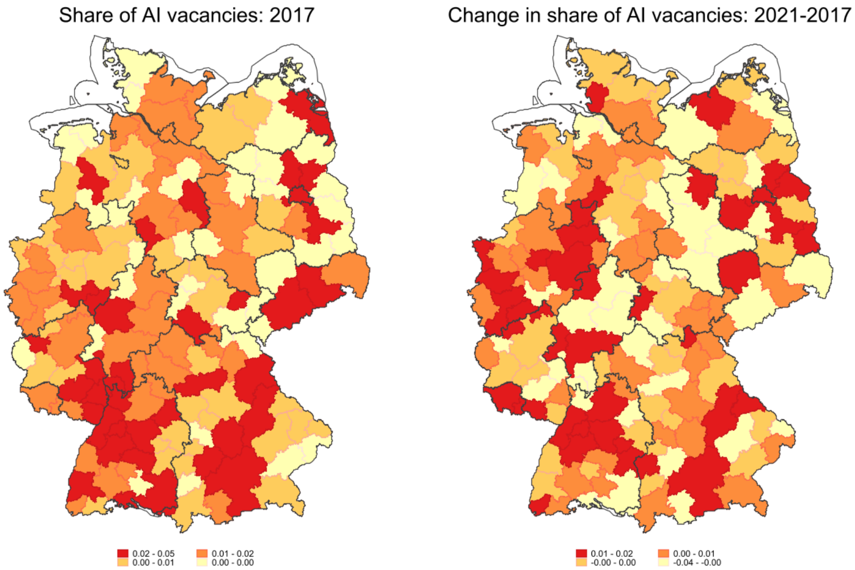
NOTE. —This graph displays the monthly mentions of the ten most important AI skills in German OJV over time. We summarize some keywords when they are closely related or ar merely German translations. The skill "AI" comprises the terms *AI*, *Artificial Intelligence*, *Kuenstliche Intelligenz*, and *KI*. The skill "ML" comprises the terms *ML*, *Machine Learning*, and *maschinelles Lernen*. The skill "Robotics" comprises the terms *Robotics* and *Robotik*. The skill "Autonomous Driving" comprises the terms *Autonomous Driving* and *autonomes Fahren*. The skill "Boosting" comprises the terms *Boost* and *Boosting*. The skill "NLP" comprises the terms *NLP* and *Natural Language Processing*. The skill "Neural Networks" comprises the terms *Neural Networks*, *Convolutional Neural Network*, *neural net*, *Neural Network*, *neurale Netze*, *neuronale Netze*, *neuronales Netz*, and *Recurrent Neural Network*.

Figure 6: Top 10 AI skills at OJV-level, 2017/01 - 2021/12



NOTE.—Vacancies are defined as an "AI vacancy" if a job posting contains at least one AI-related skill in a given month.

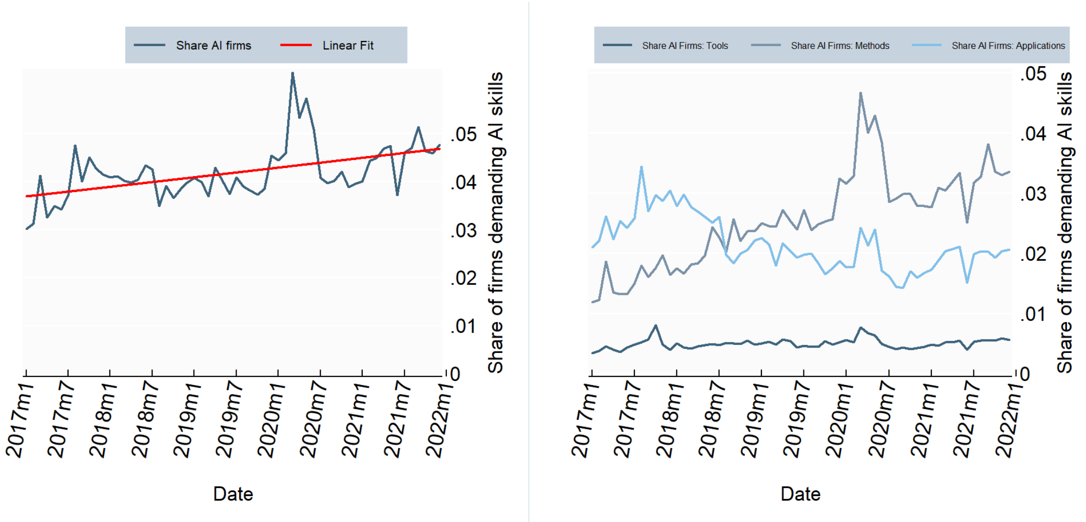
Figure 7: Share of AI vacancies by industry at OJV-level, 2017/01 - 2021/12



(a) 2017 (b) Change between 2017-2021

NOTE. —Local labor markets are assigned into four classes of task intensity. Each class corresponds to quartiles where lowest quartile implies lowest AI demand (yellow) and highest quartile implies hightest AI demand (red).

Figure 8: Demand for AI skills in Germany across local labor markets at OJV-level, 2017/01 - 2021/12

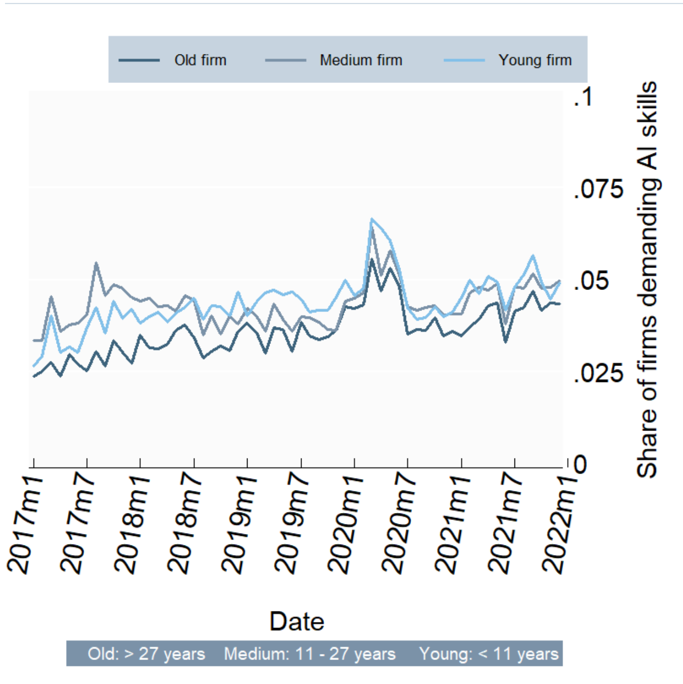


(a) AI firms: Baseline

(b) AI firms: Baseline by skill groups

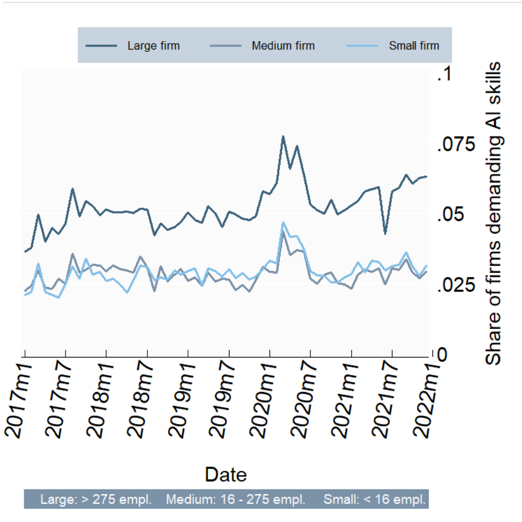
NOTE. —Firms are defined as an "AI firm" if they have at least one AI-related skill in a job posting in a given month. Vacancies are defined as an "AI vacancy" if a job posting contains at least one AI-related skill in a given month.

Figure 9: Trends in AI demand at firm-level, 2017/01 - 2021/12

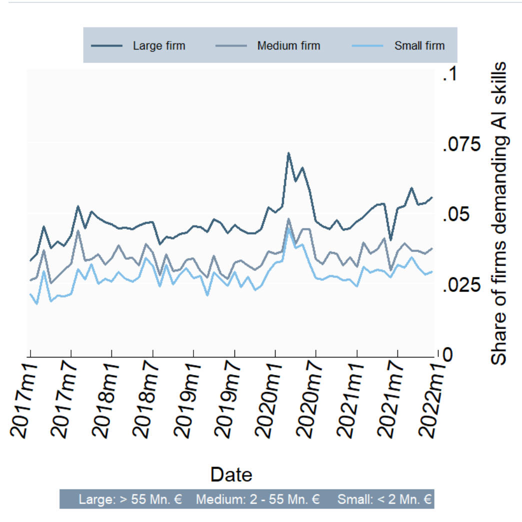


NOTE. —Young firms are defined as those at or below the 25th percentile of the firm age distribution. Old firms "AI firms" are defined as those at or above the 75th percentile of the firm age distribution. Medium firms are defined as those above the 25th and below the 75th percentile of the firm age distribution.

Figure 10: Share of firms posting AI skills in OJV: by firm age, 2017/01 - 2021/12



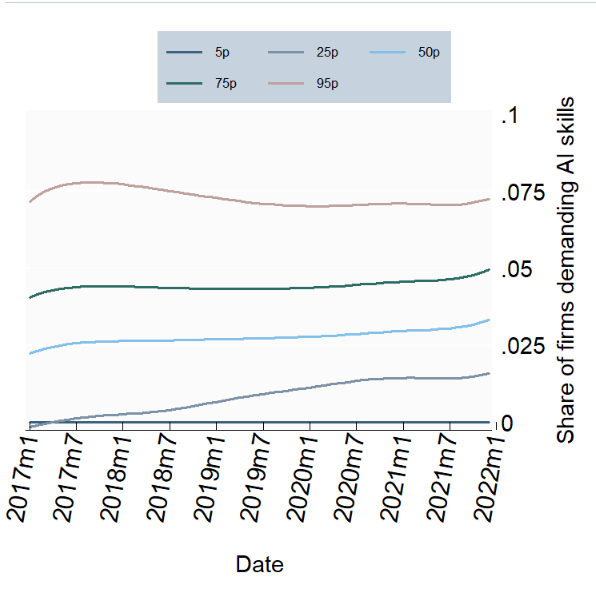
(a) By workforce



(b) By revenue

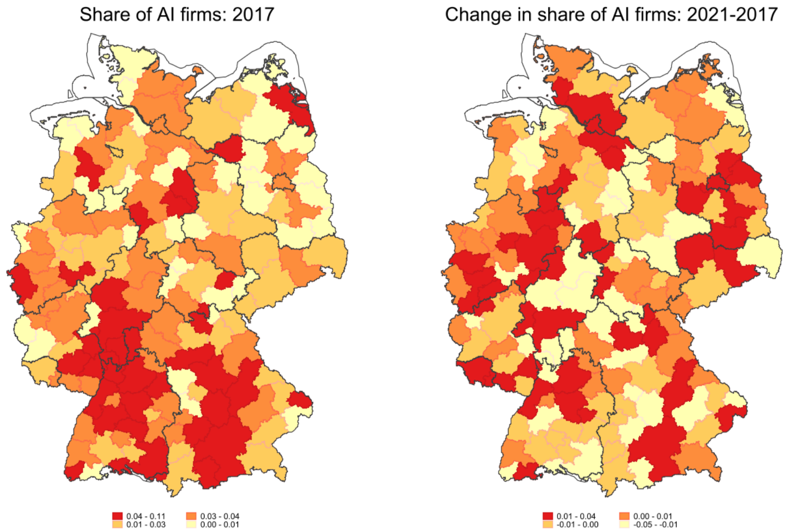
NOTE. —Small firms are defined as those at or below the 25th percentile of the firm size distribution. Large firms are defined as those at or above the 75th percentile of the firm size distribution. Medium firms are defined as those above the 25th and below the 75th percentile of the firm size distribution.

Figure 11: Share of firms posting AI skills in OJV: by firm size, 2017/01 - 2021/12



NOTE. —This graph is based on calculating the average share of AI vacancies for each of the 141 LLMs. We assign these shares to percentiles and plot the shares associated with five percentiles over time: the fifht, 25th, 50th, 75th, and 95th percentile.

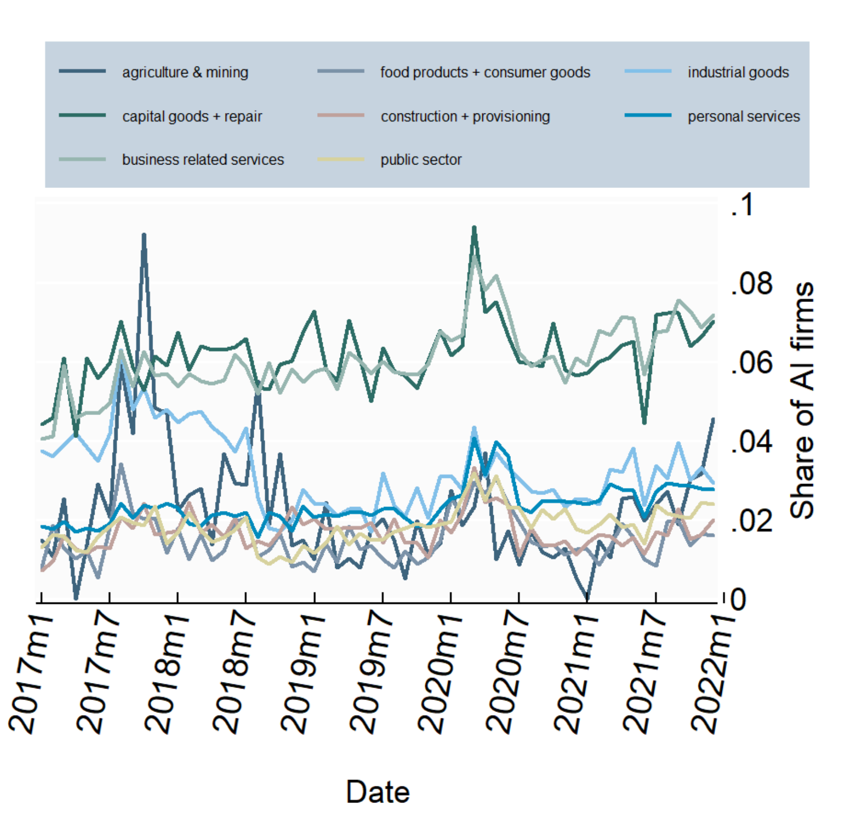
Figure 12: Share of firms posting AI skills in OJV: by regions, 2017/01 - 2021/12



(a) 2017 (b) Change between 2017-2021

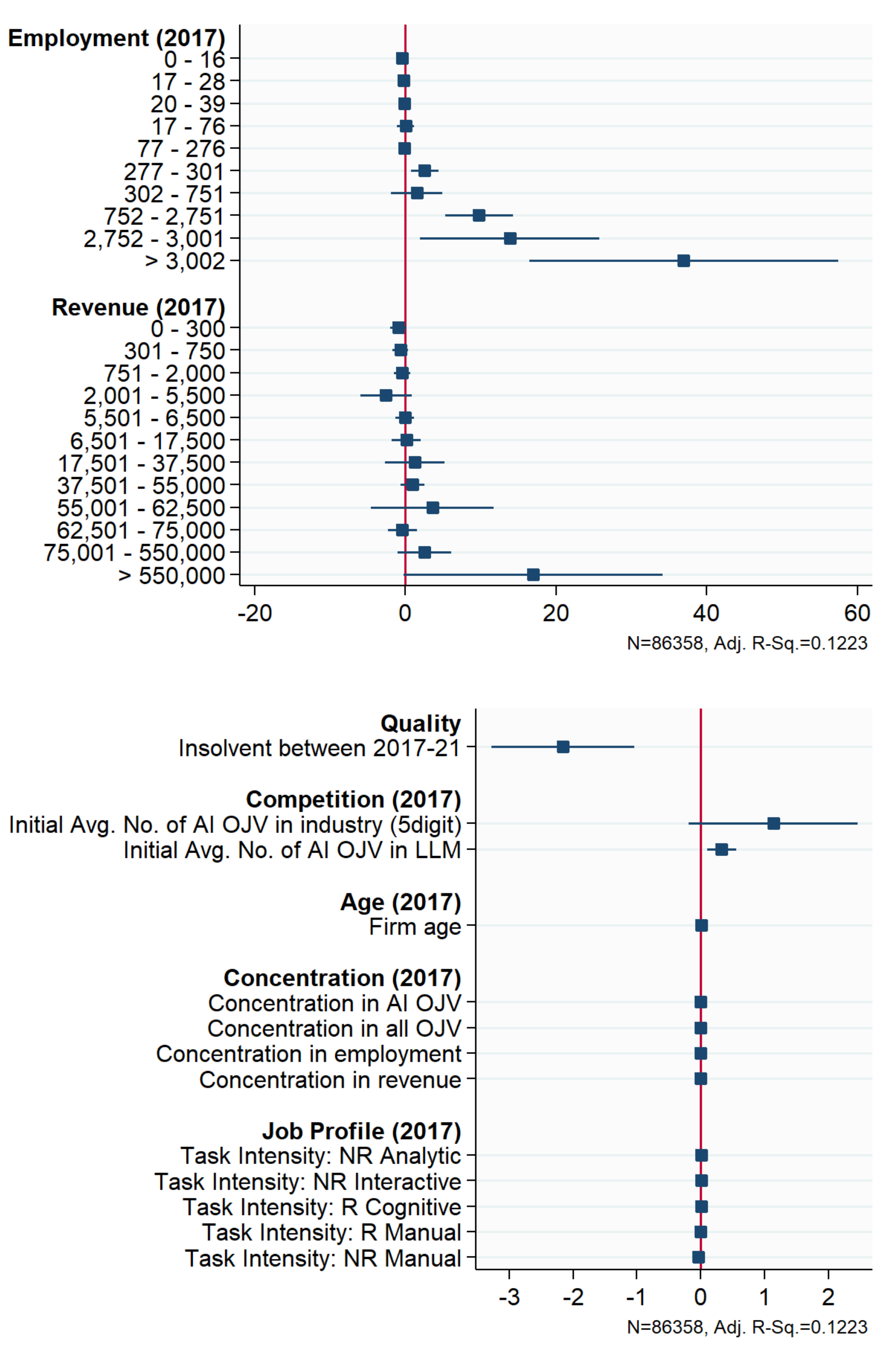
NOTE. —Local labor markets are assigned into four classes of task intensity. Each class corresponds to quartiles where lowest quartile implies lowest AI demand (yellow) and highest quartile implies hightest AI demand (red).

Figure 13: Share of AI firms across local labor markets in Germany, 2017/01 - 2021-12



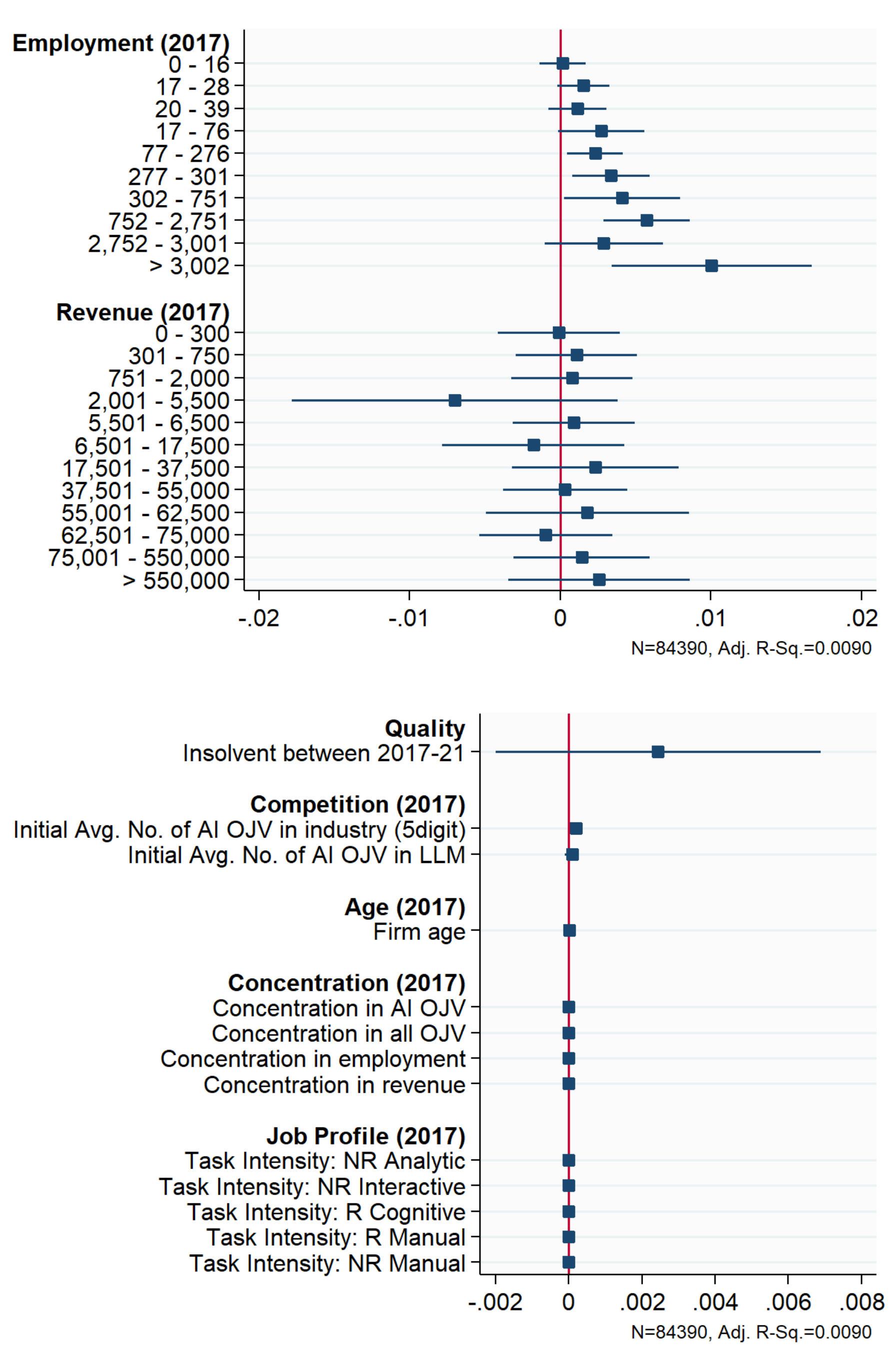
NOTE. —For each firm we first count the monthly relative share of vacancies requiring AI skills. Subsequently, we aggregate these shares for each of the eight displayed industries over time.

Figure 14: Share of AI firms by industry, 2017/01 - 2021/12



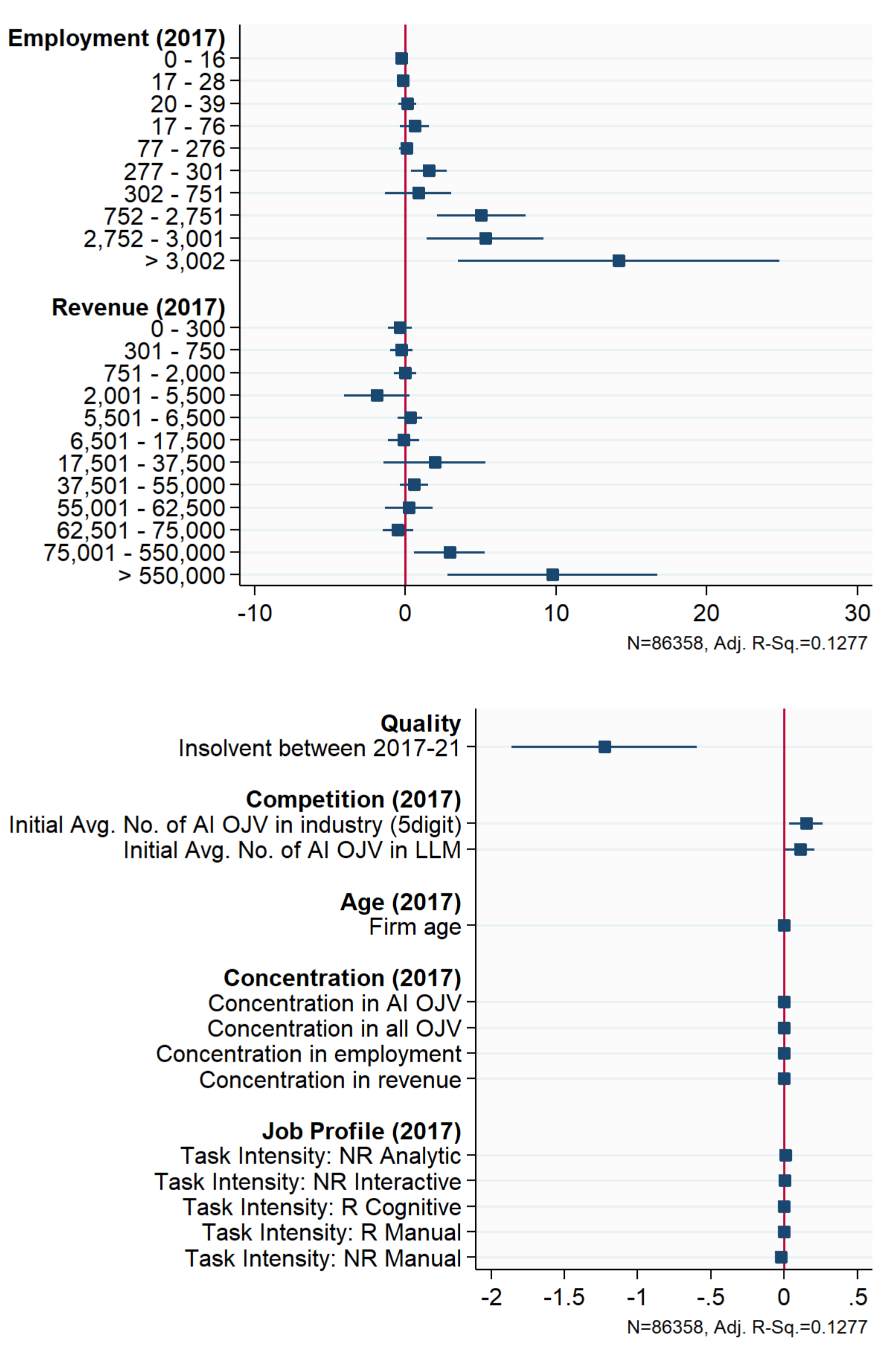
NOTE. —Point estimates based on a 95% CI and robust standard errors. All coefficients are based on (1). Next to the firm-level characteristics displayed in this graph, we include local differences regarding: (i) Skill composition (share of workers with college, high school, neither), (ii) Share of female workers, (iii) Share of foreign workers, (iv) Age composition (via seven age bins), (v) Employment share by (13) broad industries, and (vi) Unemployment rate.

Figure 15: Regression of the change in AI vacancies at firm-level: baseline (tools, applications, methods), 2017/01 - 2021/12



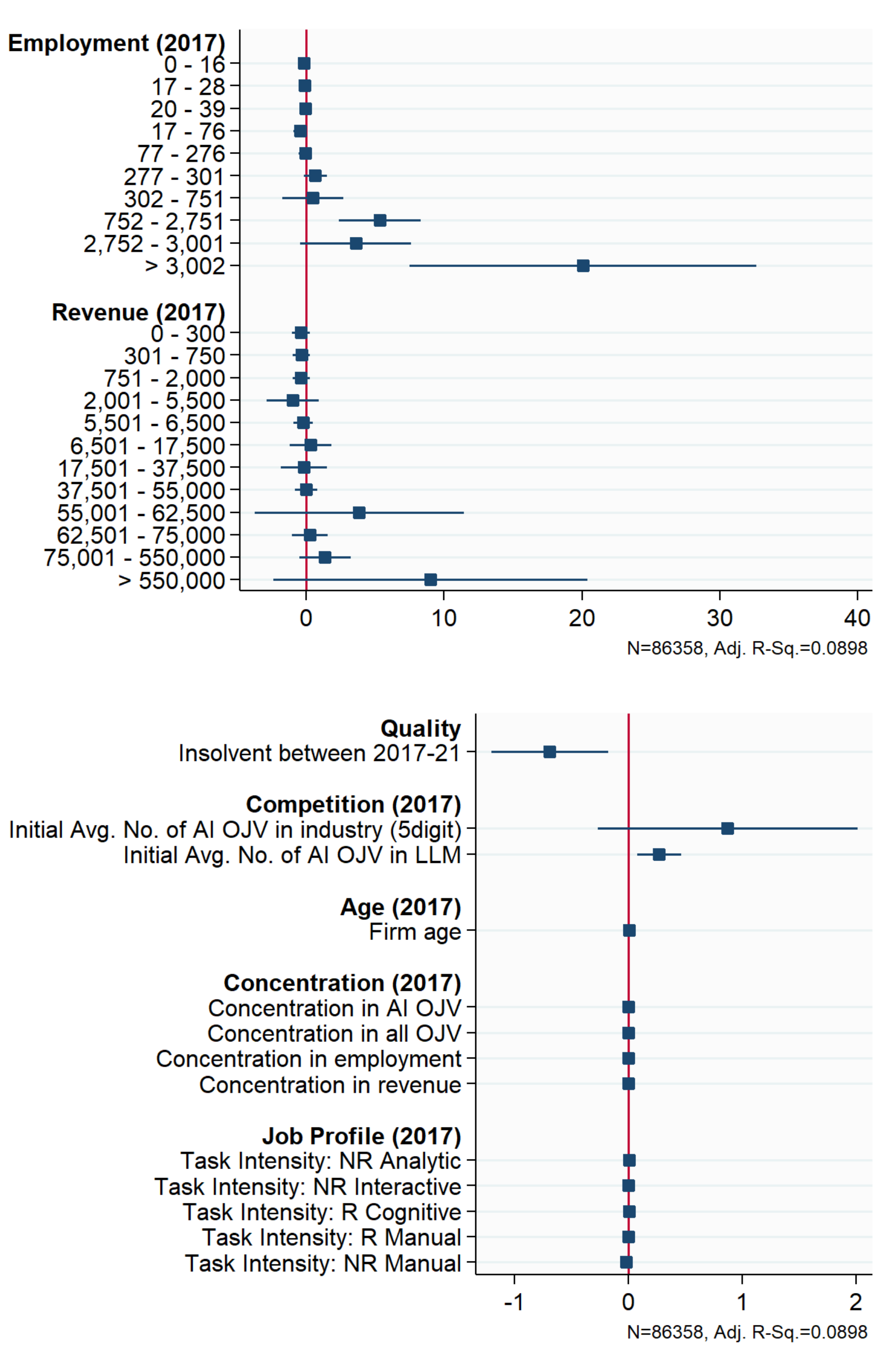
NOTE. —Point estimates based on a 95% CI and robust standard errors. All coefficients are based on (1). Next to the firm-level characteristics displayed in this graph, we include local differences regarding: (i) Skill composition (share of workers with college, high school, neither), (ii) Share of female workers, (iii) Share of foreign workers, (iv) Age composition (via seven age bins), (v) Employment share by (13) broad industries, and (vi) Unemployment rate.

Figure 16: Regression of the change in the AI vacancy share at firm-level: baseline (tools, applications, methods), 2017/01 - 2021/12



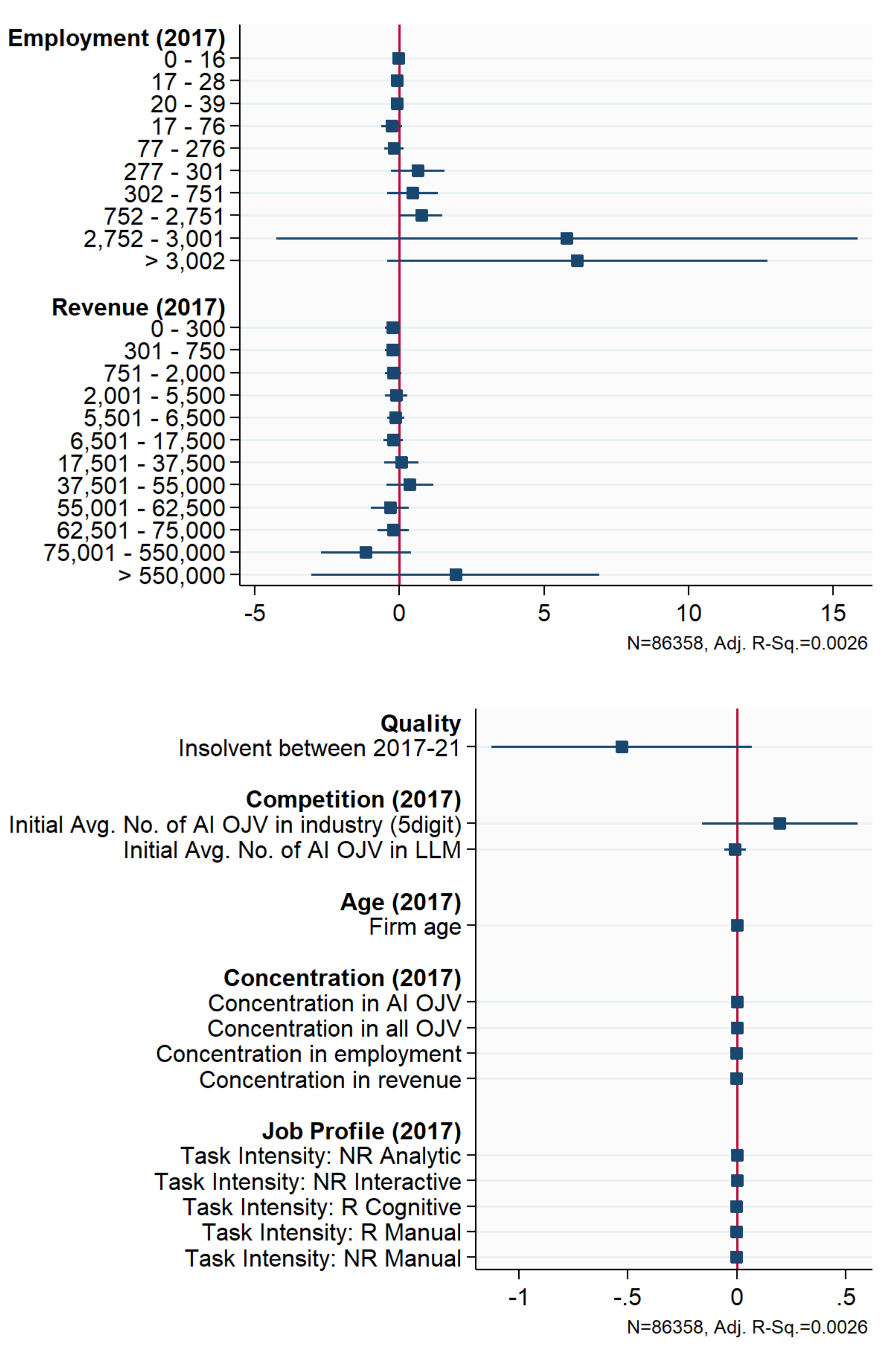
NOTE. —Point estimates based on a 95% CI and robust standard errors. All coefficients are based on (1). Next to the firm-level characteristics displayed in this graph, we include local differences regarding: (i) Skill composition (share of workers with college, high school, neither), (ii) Share of female workers, (iii) Share of foreign workers, (iv) Age composition (via seven age bins), (v) Employment share by (13) broad industries, and (vi) Unemployment rate.

Figure 17: Regression of the change in AI vacancies at firm-level: AI methods, 2017/01 - 2021/12



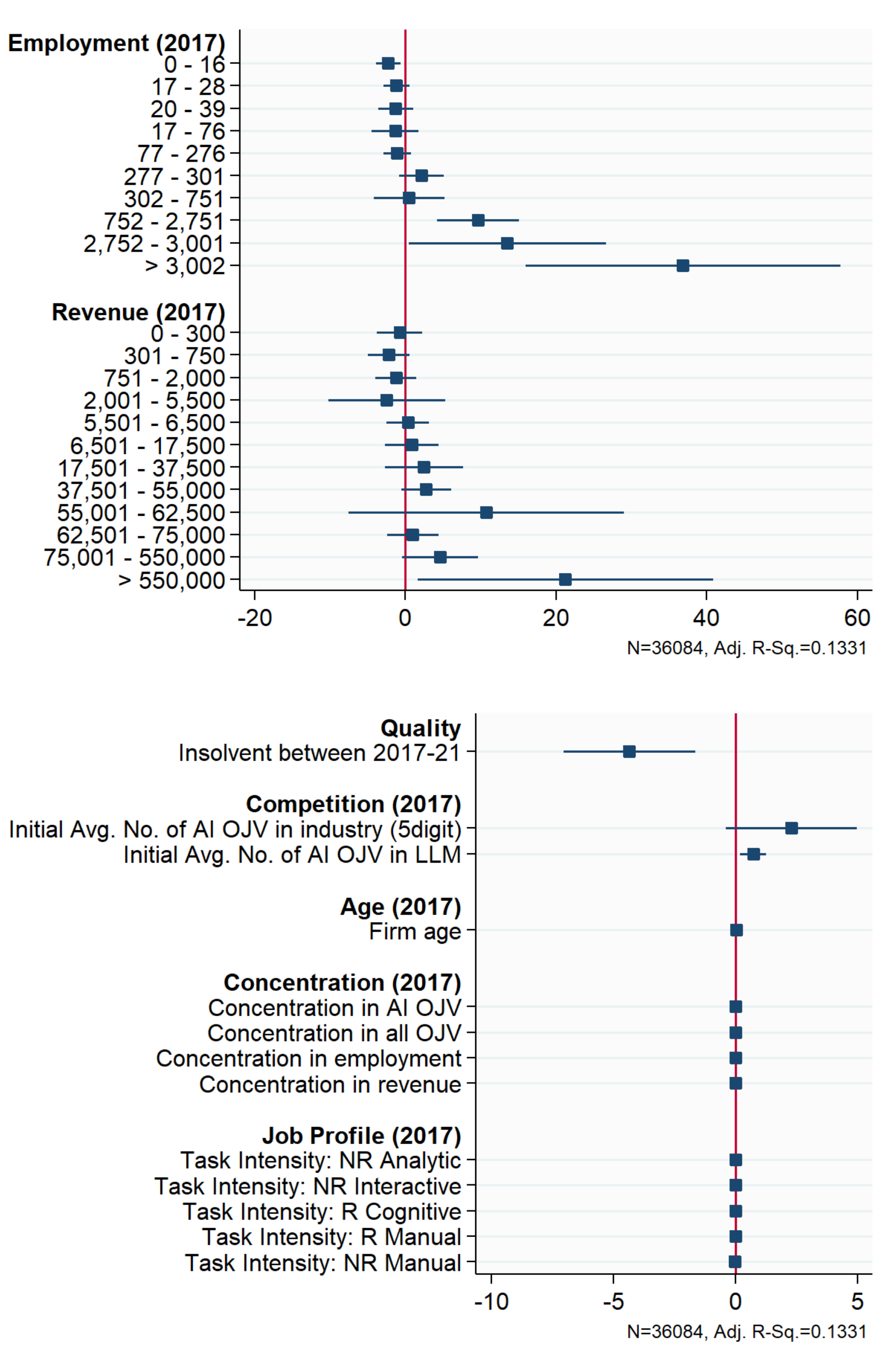
NOTE. —Point estimates based on a 95% CI and robust standard errors. All coefficients are based on (1). Next to the firm-level characteristics displayed in this graph, we include local differences regarding: (i) Skill composition (share of workers with college, high school, neither), (ii) Share of female workers, (iii) Share of foreign workers, (iv) Age composition (via seven age bins), (v) Employment share by (13) broad industries, and (vi) Unemployment rate.

Figure 18: Regression of the change in AI vacancies at firm-level: AI applications, 2017/01 - 2021/12



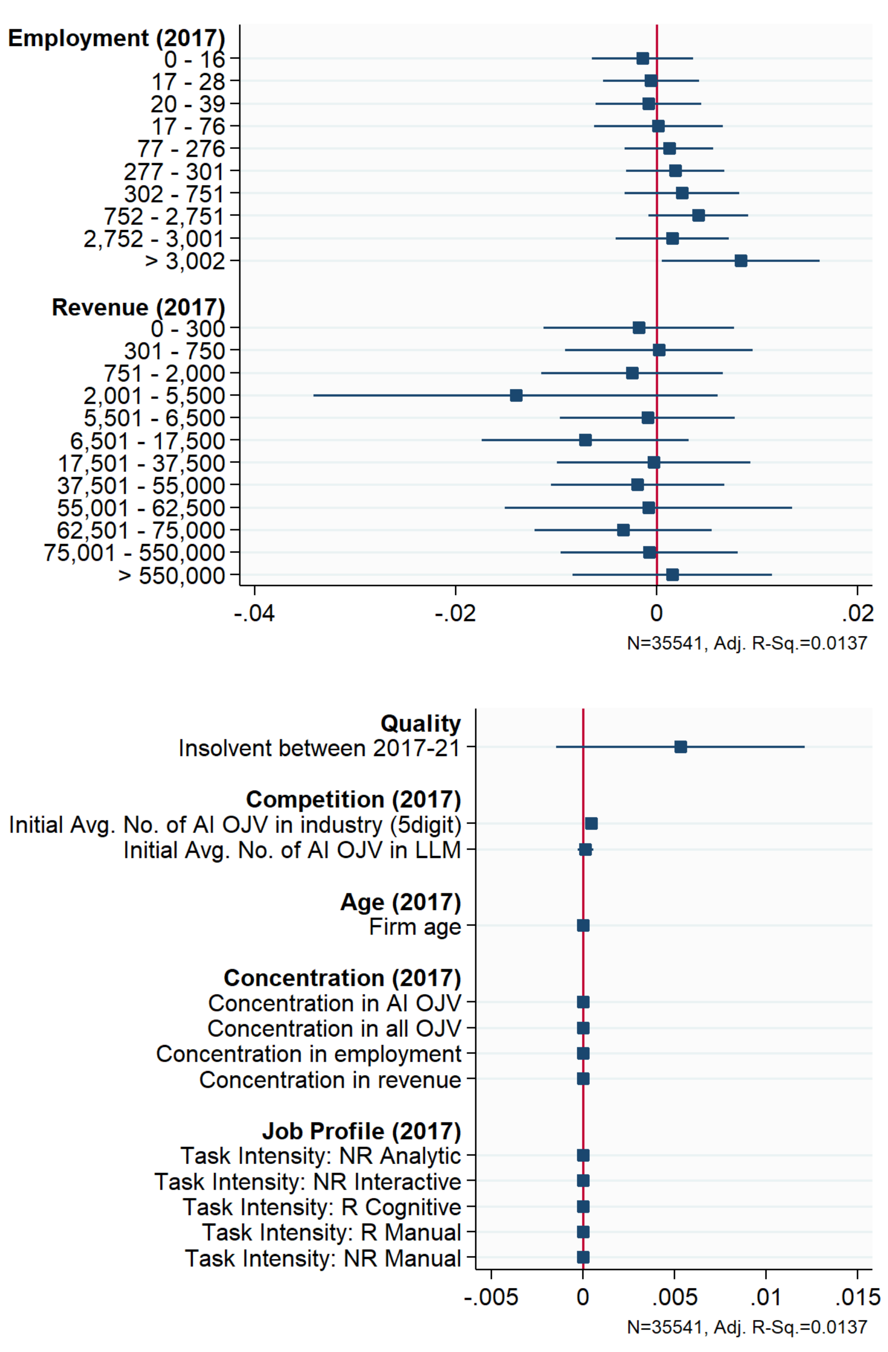
NOTE. —Point estimates based on a 95% CI and robust standard errors. All coefficients are based on (1). Next to the firm-level characteristics displayed in this graph, we include local differences regarding: (i) Skill composition (share of workers with college, high school, neither), (ii) Share of female workers, (iii) Share of foreign workers, (iv) Age composition (via seven age bins), (v) Employment share by (13) broad industries, and (vi) Unemployment rate.

Figure 19: Regression of the change in AI vacancies at firm-level: AI tools, 2017/01 - 2021/12



NOTE. —Point estimates based on a 95% CI and robust standard errors. All coefficients are based on (1). Next to the firm-level characteristics displayed in this graph, we include local differences regarding: (i) Skill composition (share of workers with college, high school, neither), (ii) Share of female workers, (iii) Share of foreign workers, (iv) Age composition (via seven age bins), (v) Employment share by (13) broad industries, and (vi) Unemployment rate.

Figure 20: Regression of the change in AI vacancies at firm-level: Robustness imposing common time horizon, 2017/01 - 2021/12



NOTE. —Point estimates based on a 95% CI and robust standard errors. All coefficients are based on (1). Next to the firm-level characteristics displayed in this graph, we include local differences regarding: (i) Skill composition (share of workers with college, high school, neither), (ii) Share of female workers, (iii) Share of foreign workers, (iv) Age composition (via seven age bins), (v) Employment share by (13) broad industries, and (vi) Unemployment rate.

Figure 21: Regression of the change in the AI vacancy share at firm-level: Robustness imposing common time horizon, 2017/01 - 2021/12

B Tables

Table 1: Overview of industries from IFR and EU KLEMS

1-digit industry	2-digit industry	industry name
A	01-03	Agriculture, forestry and fishing
B	05-09	Mining and quarrying
C	10-12	Manufacture of food products; beverages and tobacco products
C	13-15	Manufacture of textiles, wearing apparel, leather and related products
C	16	Manufacture of wood, paper, printing and reproduction
C	17-18	Manufacture of wood, paper, printing and reproduction
C	19	Coke and mineral oil processing
C	20-21	Chemicals; basic pharmaceutical products
C	22-23	Manufacture of rubber and plastic products and other non-metallic mineral products
C	24-25	Manufacture of basic metals and fabricated metal products, except machinery and equipment
C	26-27	Computer, electronic, optical products; electrical equipment
C	28	Manufacturing of machines
C	29-30	Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment
C	31-33	Manufacture of furniture; jewellery, musical instruments, toys; machinery and equipment
D	35	Energy supply
E	36-39	Water supply; sewerage, waste management and remediation services
F	41	Structural engineering
F	42	Civil engineering
F	43	Preliminary site work
G	45	Sale, maintenance and repair of motor vehicles and motorcycles
G	46	Wholesale trade
G	47	Retail trade
H	49	Land transport and transport in pipelines
H	50-51	Air and sea transport
H	52-53	Warehousing, postal, courier and express services
I	55-56	Hospitality industry
J	58-61	Publishing, motion picture, video, television programme production; and similar
J	62-63	Computer programming, consultancy, and information service activities
K	64-66	Financial and insurance activities
L	68	Real estate and housing
M	69-75	Scientific, technical and freelance services
N	77-82	Other economic service activities
O	84	Public administration, defence; social security
P	85	Education and teaching
Q	86	Healthcare
Q	87	Homes
Q	88	Social services (without homes)
R	90-93	Art, entertainment and recreation
S	94-96	Other services
U	99	Extra-territorial organisations and bodies

Table 2: Summary Statistics: AI Firms vs Non-AI Firms

	AI Firms	Non-AI Firms	Difference
Age	19.81	20.91	1.10***
Avg. No. Job Postings	19.84	5.60	-14.24***
Workforce size	3196.57	791.69	-2404.89***
Revenue (in Thousands)	247.90	104.30	-142.60***
Share of OJV w/ AI skills: TOOLS	0.03	0.00	-0.03***
Share of OJV w/ AI skills: METHODS	0.24	0.00	-0.24***
Share of OJV w/ AI skills: APPLICATIONS	0.17	0.00	-0.17***
Share of OJV requiring NRA tasks	0.79	0.60	-0.18***
Share of OJV requiring NRI tasks	0.83	0.78	-0.05***
Share of OJV requiring RC tasks	0.45	0.43	-0.03***
Share of OJV requiring RM tasks	0.38	0.37	-0.01***
Share of OJV requiring NRM tasks	0.36	0.35	-0.00**
Observations	90160	2022643	2112803

NOTE. —Firms are defined as an "AI firm" if they have at least one AI-related skill in a job posting in a given month. Vacancies are defined as an "AI vacancy" if a job posting contains at least one AI-related skill in a given month.

Table 3: LLM-level regression of employment on AI demand, 2017 - 2019

	AI (all)	AI tool	AI application	AI method
AI demand	102.945	964.792	73.809	119.832
	(64.321)	(960.558)	(59.690)	(107.330)
East/West dummy	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
Adj. R^2	0.092	0.093	0.092	0.092
N	12072	12072	12072	12072

NOTE. —Standard errors corrected for heteroskedasticity clustered at labour market regions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is region west and the industry group agriculture & mining. The following variables are included as controls: share of women, university graduates, workers with vocational training, foreign workers.

Table 4: LLM-level regression of log wages on AI demand, 2017 - 2019

	AI (all)	AI tool	AI application	AI method
AI demand	0.162*	1.078***	0.234***	0.009
	(0.083)	(0.324)	(0.085)	(0.096)
East/West dummy	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
Adj. R^2	0.671	0.671	0.671	0.670
N	12072	12072	12072	12072

NOTE. —Standard errors corrected for heteroskedasticity clustered at labour market regions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is region west and the industry group agriculture & mining. The following variables are included as controls: share of women, university graduates, workers with vocational training, foreign workers.

Table 5: LLM-level regression of employment growth on the change in AI demand, 2017 - 2019

	AI (all)	AI tool	AI application	AI method
AI demand	0.027	-0.286	-0.019	0.074
	(0.064)	(0.293)	(0.085)	(0.108)
East/West dummy	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
Adj. R^2	0.020	0.020	0.020	0.020
N	3536	3536	3536	3536

NOTE. —Standard errors corrected for heteroskedasticity clustered at labour market regions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is region west and the industry group agriculture & mining. The following variables are included as controls: share of women, university graduates, workers with vocational training, foreign workers.

Table 6: LLM-level regression of wage growth on the change in AI demand, 2017 - 2019

	AI (all)	AI tool	AI application	AI method
AI demand	-0.006	-0.075	-0.035	0.096*
	(0.037)	(0.261)	(0.050)	(0.054)
East/West dummy	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
Adj. R^2	0.040	0.040	0.040	0.041
N	3532	3532	3532	3532

NOTE. —Standard errors corrected for heteroskedasticity clustered at labour market regions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is region west and the industry group agriculture & mining. The following variables are included as controls: share of women, university graduates, workers with vocational training, foreign workers.

C Appendix

C.1 External Validity of OJV Data

Figure 22a shows the number of OJV over time by source platforms. Overall, we see an increasing trend of the number of postings over time. In principle, this pattern can be explained by two factors. First, an increasing trend over time, i.e., firms may use their websites and job boards more often to post jobs online. Second, methodological changes, e.g., our private partner updates its scraping method and thus adds more sources. Rising levels of digitalization and the growing popularity of online job search by job seekers likely contribute to the increasing trend in OJV. We further find evidence that methodological changes matter as well since the composition of source platforms has changed over time. While (fee paying) job boards represented about 50% of all postings in 2017, their share increased to 70% by the end of 2021. This increase has come primarily at the expense of headhunters whose share decreased from 17% to less than 2% during the same time. These compositional changes demonstrate the need to validate the representativeness of OJV data.

[Figure 22 here]

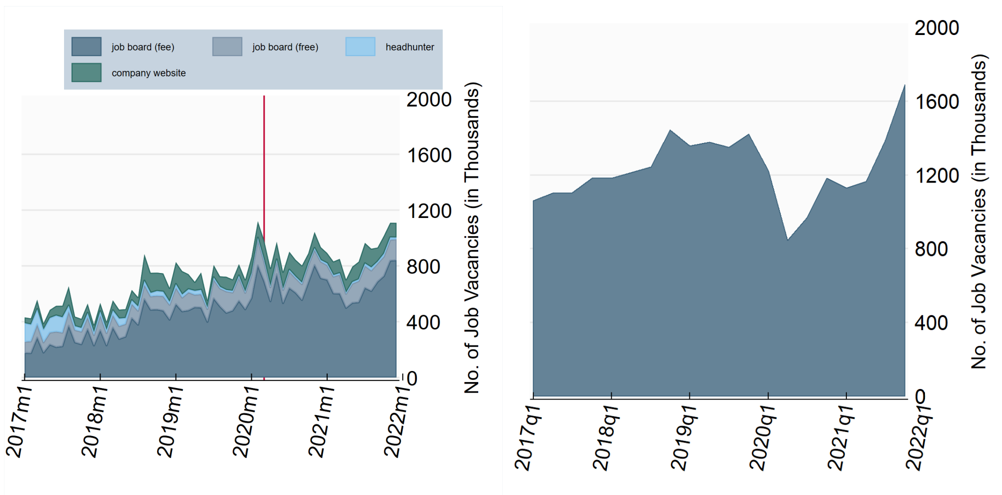
We follow common practice in the literature by comparing our OJV data with representative information on vacancies from official sources (Hershbein & Kahn 2018, Rengers 2018). Hershbein & Kahn (2018) compare characteristics of the job postings from Lightcast (formerly Burning Glass Technologies) with the Bureau of Labor Statistics' Job Openings and Labor Market Turnover (JOLTS) survey and other data sources for the US at the aggregate level and by industries. Likewise, Rengers (2018) makes similar comparisons for Germany with data from the Federal Employment Agency (BA) and the IAB Job Vacancy Survey. Especially relevant for our purposes, the IAB Job Vacancy Survey is a representative survey and measures the aggregate labor demand and the recruiting behavior of firms in Germany since 1989, making it a well-suited survey for the analysis of recruitment processes

(Gürtzgen, Lochner, Pohlan & van den Berg 2021). Below, we address these concerns by first studying aggregate trends and subsequently breaking down our OJV data by industries.

First, Figure 22 compares the (aggregate) evolution of vacancies taken from the IAB Job Vacancy Survey from 2017Q1 - 2021Q4 (2021 values are estimates) with our OJV data. Note that the IAB data reflects stock information, while our data is a measure for inflows of job postings. Despite these methodological differenes, the two graphs display similar trends. Both display a steady increase in postings from 2017 until early 2020 with a sharp decrease at the onset of the pandemic in March 2020. While the stock of vacancies decreased by 40% between 2019Q4 and 2020Q2 based on the IAB Vacancy Panel, the inflows of vacancies in our OJV data decreased by 30% from December 2019 until June 2020. Both time series display a sharp subsequent rebound, leading to a catch-up to pre-COVID vacancy levels by the end of 2020. Moreover, the magnitude of the drop and rebound in job vacancies during the pandemic is consistent with previous findings in the literature from comparable countries, such as Australia (Shen & Taska 2020), Austria (Bamieh & Ziegler 2020), Sweden (Hensvik, Le Barbanchon & Rathelot 2021), the UK (Arthur 2021), and the US (Forsythe, Kahn, Lange & Wiczer 2020). Hence, both, the cyclicality of job postings and the magnitude in collapse and recovery of postings, lend credence to the validity of our data.

[Figure 23 here]

Second, we divide our vacancies into six broad industries for ease of exposition: (i) manufacturing, (ii) retail & hospitality, (iii) information & communication, (iv) professional services, (v) personal services, and (vi) other industries. Figure 23 summarizes this industrial breakdown and provides three key takeaways. First, all industries are covered and well-represented in our data. Second, service industries, comprising professional and personal services, are the most important industry groups. On average, these broad industries comprise around half of all vacancies. Third, the industry composition in our data has become more balanced over time. While the share of services decreased from 60% to 45% from 2017 until 2021, manufacturing and retail & hospitality have experienced rising coverage (in each industry from 15% to 20%). We interpret these developments favorably as the descriptive statistics support the quality of our data and its broadly representative nature. Part of this takeaway is attributed to the fact that our data begins in 2017. Internet access and especially online job search have already been common at this point, a distinguishing feature from the earliest possible OJV data in the US in the mid 2000s, a time during which online job posting was concentrated among professionals (Hershbein & Kahn 2018, Modestino, Shoag & Ballance 2019).



(a) OJV data, by source (Inflow) (b) IAB Vacancy Panel (Stock)

NOTE. —Panel 22a displays the number of online job vacancies that are posted each month in our data, i.e., monthly inflows, broken down by the type of source platform. Panel 22b displays the stock of vacancies firms report to the IAB for each quarter. The values for 2021Q1 onward are estimates as final numbers are not available yet.

Figure 22: Number of online job vacancies over time, 2017/01 - 2021/12

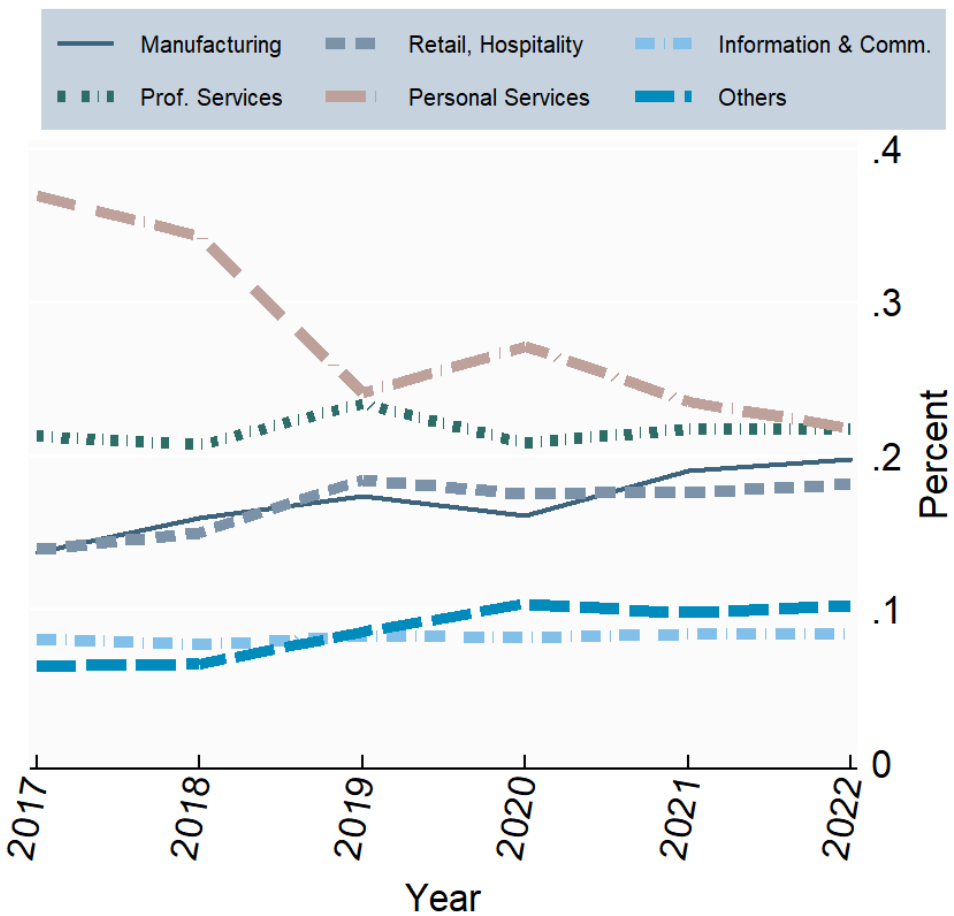
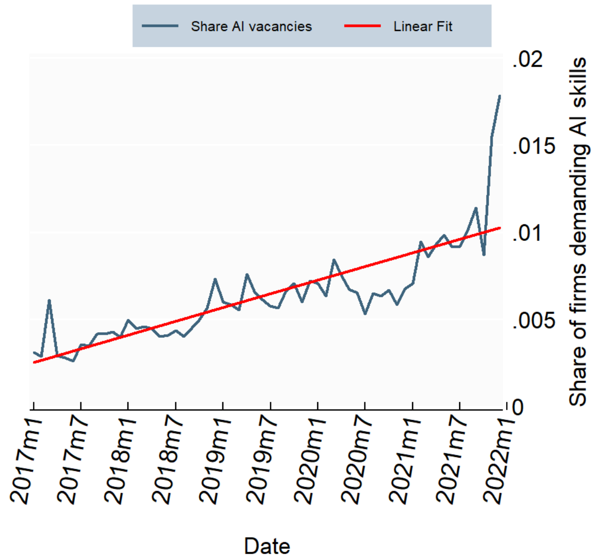
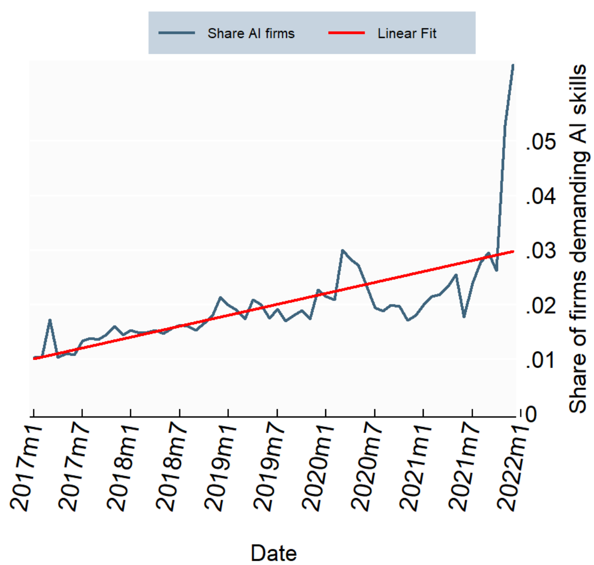


Figure 23: Industry composition of online job vacancies, 2017/01 - 2021/12



(a) AI vacancies



(b) AI firms

NOTE. —Firms are defined as an "AI firm" if they have at least one AI-related skill in a job posting in a given month. Vacancies are defined as an "AI vacancy" if a job posting contains at least one AI-related skill in a given month. In this graph we used keywords provided by Taska et al. (2022) as a benchmark to our own taxonomy.

Figure 24: Trends in AI demand: Taska et al. (2022) taxonomy, 2017/01 - 2021/12