From Code to Cash: The Impact of AI on Wages

Myrielle Gonschor[1] and Eduard Storm[1]

[1]RWI - Leibniz Institute for Economic Research

April 2024

Abstract

update: Artificial Intelligence (AI) can perform cognitively demanding tasks with more autonomy than previous technologies and is thus expected to have disruptive effects on labor markets. But empirical evidence is limited. Does AI already affect workers' wages? And how exactly does AI diffuse through labor markets? To answer these questions we combine novel job vacancy data from Germany with high-quality administrative data and contribute three main findings. First, using an IV approach, we find that a 10% increase in demand for AI skills implies average AI-induced wage returns of 2%. Second, we identify three key drivers behind our results and find that 95% of AI-induced wage effects are attributed to: (1) Employer Quality, (2) Socioeconomic, and (3) Occupational characteristics. Third, we explore mechanisms, suggesting that the primary beneficiaries of AI demand are male workers with: (i) only modest AI exposure, (ii) vocational education, (iii) 50+ years of age, (iv) occupational mobility, and (v) employment at high-quality firms. Our paper provides valuable insights for policymakers by identifying early winners and losers of growing AI diffusion and offers promising avenues for future research.

Keywords: AI, Online Job Vacancies, Labor Demand, Worker-level Analysis, Wages
JEL Codes: D22, J23, J24, J31, O33

1 Introduction

Artificial Intelligence (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, Webb 2020). While AI is still a nascent and specialized technology, a rapidly growing number of firms and workers are exposed to this technology. For example, the share of AI-adopting firms in Germany has increased from 6% in 2019 to 13% in 2023 based on firm-level survey data (**?**Schaller, Wohlrabe & Wolf 2023). Similarly, firm-level surveys from the US show only 3% of US firms had adopted AI by 2019, though almost 13% of US workers had been exposed to this technology at work.[1] **not sure if you should stress the firm perspective at the beginning**

The rapid adoption of AI may automate some tasks, while creating new ones, with uncertain implications for labor demand (**?**). These adjustments will likely hurt some workers, but also allow others to become more productive (Brynjolfsson, Li & Raymond 2023, Noy & Zhang 2023). Since economic theory dictates that wages are determined by a workers' marginal product, we should thus also expect that rising AI demand affects wages. In contrast to previous technologies, AI technology has the ability to perform many cognitively demanding tasks, which could impact workers higher up the income distribution disproportionately.

While a growing number of studies provide important insights on the diffusion of AI, there is still little evidence on actual labor market outcomes. A comprehensive analysis requires detailed and up-to-date data on AI diffusion and access to high-quality labor market data. In this paper, we fill this gap by studying the wage impact of rising AI demand and identifying its key drivers. We use natural language processing (NLP) methods to identify AI skills from the near-universe of German online job vacancies (OJV) between 2017 - 2021.

[1] See **?**. This discrepancy in firm-level adoption and worker-level exposure is largely attributed to the fact that AI is primarily concentrated among large firms (**?**Rammer, Fernández & Czarnitzki 2022). 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.

Subsequently, we use this measure to define AI exposure and merge it to administrative data at the occupation-region-year level —our baseline definition of a local labor market. Doing so allows us to study the worker-level wage effects associated with changes in AI demand in their local labor market. We perform this analysis using OLS as a baseline. However, we recognize that endogeneity concerns likely bias our OLS results due to non-random adoption of AI technologies. Therefore, we support our OLS results with an IV approach. To this end, we construct a Bartik-like instrument that exploits national trends in AI demand, which are plausibly orthogonal to local conditions. This identification strategy assumes shift-exogeneity, as proposed by Borusyak, Hull & Jaravel (2022), where the shift component of our instrument captures occupation-specific shocks that occur outside of a workers' home region.

In this paper, we measure AI exposure as the share of online job vacancies that require AI skills at the local labour market level, which also accounts for AI diffusion. However, one might be concerned that our wage effects of AI capture other shocks and trends at the local labour market level. We alleviate concerns regarding this identification threat with rich specifications and flexible models accounting for occupation-specific and region-specific demand shocks. Another potential concern is regarding our definition of AI skills. We address this concern by also constructing various alternative technology measures. We find our results robust to alternative AI measures and unique compared to broader measures that incorporate other digital technologies, not directly related to AI. These tests suggest our definition of AI skills is indeed valid and meaningful.

Our paper makes several contributes to the literature. First and foremost we contribute to the sparse literature on worker-level effects of digital technologies. Some studies combine occupation-level AI measures with survey data (Fossen & Sorgner 2022), others use self-administered firm-level surveys to measure adoption of different technologies. [2] These studies find mostly positive, though heterogenous, effects on outcomes such as wages and

[2]See Genz, Janser & Lehmer (2019), Genz, Gregory, Janser, Lehmer & Matthes (2021), Gathmann, Kagerl, Pohlan & Roth (2023), Barth, Bryson & Dale-Olsen (2022).

employment stability, both in terms of occupational exposure and among workers whose employer has adopted digital technologies. Given the nature of data collection, these studies cannot focus on specific technologies and instead bundle different technologies for survey questions, or, in the case of Fossen & Sorgner (2022), cannot develop own AI taxonomies. Using a Big Data approach, we add to this literature detailed insights on the wage implications of AI, one of the key emerging technologies. We especially identify the key drivers of AI-induced wage changes.

More recently, researchers have began studying the implications of Large Language Models (LLMs), a subfield of AI, and the foundation for tools such as Chat GPT. Noy & Zhang (2023) show in an experiment with 453 college graduates that Chat GPT substantially raises productivity and decreases between-worker inequality in productivity. Similarly, Eloundou, Manning, Mishkin & Rock (2023) show that many worker tasks in the US could be completed significantly faster using LLMs, e.g., annotation tasks (Gilardi, Alizadeh & Kubli 2023). While we do not analyze LLMs specifically (though included in our AI taxonomy), their productivity-enhancing features should in principle apply to a broader set of AI technologies. Economic theory then dictates that changes in productivity translate into changes in wages. We shed light on this channel and add further insights on idiosyncratic effects of specific AI applications in our robustness analysis.

We also contribute to the growing literature exploring the broader labor market impact of AI. Several studies study aggregate outcomes to gauge the AI exposure of occupations and industries or regions.[3] In the context of our paper, in which wages are of primary interest, we instead prefer a measure that captures the diffusion of AI more directly: the share of vacancies that demand AI skills in a worker's relevant local labor market - defined at the intersection of LMR and occupation.[4] We consider this the more relevant local labor market

[3]See Webb (2020), Felten, Raj & Seamans (2021), **?**), Brynjolfsson, Mitchell & Rock (2018), Albanesi, Dias da Silva, Jimeno, Lamo & Wabitsch (2023), **?**), Acemoglu, Autor, Hazell & Restrepo (2022) for papers exploiting occupation and/ or industry-level AI exposure. While these proxies highlight *potential* AI exposure, they are less informative on *actual* AI exposure. Notable studies that exploit regional variation are Bessen, Cockburn & Hunt (2021), Gathmann & Grimm (2022).

[4]AAHR construct a similar measure. However, in their paper the share of AI-related vacancies is the

for the analysis of technological change (Azar, Marinescu, Steinbaum & Taska 2020). Only a few studies have more detailed information on AI exposure than we do, using firm-level data. These studies have different objectives, however, as they focus on the innovation performance of AI-adopting firms (Rammer, Fernández & Czarnitzki 2022) or employment outcomes at the establishment-level (Copestake, Marczinek, Pople & Stapleton 2023, Peede & Stops 2023). Instead, we add the worker-level dimension to this literature and focus on wage implications. Doing so, we provide a detailed analysis of mechanisms that identify vulnerable groups and derive important policy implications.

Our paper also relates to many studies that provide important descriptive insights on the diffusion of AI. Some studies also use OJV data, usually from the US, and find the impact of AI to be concentrated at the establishment-level with negligible effects at more aggregated levels.[5] Others combine OJV with firm-level data (**?**Bloom, Hassan, Kalyani, Lerner & Tahoun 2021) or use survey data (**?**) and usually find AI adoption to be concentrated in large, productive firms and specific industries. **add stoppsis paper here: @Ede: you heard their paper hundred times and knows what their current results are**. We add new insights on AI skill demand in Germany, between 2017-2021, and contribute a new AI taxonomy for economic research. We make this taxonomy publicly available upon release of this working paper. Moreover, this literature typically studies potential labor market outcomes, e.g. by analyzing posted wages, instead we study *realized* outcomes.

Our paper is also broadly related to the vast literature studying the labor market effects of new technologies. Several studies analyze the impact of specific technologies, e.g. robots, software, or, more recent digital technologies.[6] These studies typically find mixed results,

---

dependent variable to test the implications of greater AI exposure on labor and skill demand. In contrast, we use the share of AI vacancies as key regressor to proxy worker's exposure to AI technologies and subsequent wage implications.

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

[6]For evidence on the impact of robots see **?**Dauth, Findeisen, Suedekum & Woessner (2021), Koch, Manuylov & Smolka (2021). Evidence on the effects of software can be found in Autor & Dorn (2013), **?**), Dillender & Forsythe (2022) and studies on the impact of recent digital technologies in Genz, Gregory, Janser, Lehmer & Matthes (2021), Arntz, Genz, Gregory, Lehmer & Zierahn-Weilage (2023).

which is to some extent country-specific, but also depends on data availability (firm-level vs aggregated-level data). Our research adds to this literature the analysis of AI, a relatively recent digital technology, and the use of online job vacancies to measure technology diffusion. Moreover, we contribute to the literature on regional skill differences.[7] These studies often use OJV data and find large regional skill differences. We contribute important insights on regional skill differences in the context of AI and their implications for wages. **maybe we frame this more as a tech 3.0 vs. we look at tech 4.0., I wouldn't refer to robots and software as new technologies**

2 Conceptual Background

In this section we outline the theoretical framework guiding our empirical analysis. Specifically, we conceptualize the framework proposed in Acemoglu, Autor, Hazell & Restrepo (2022) (henceforth AAHR), who explore the impact of rising AI exposure on establishment's labor demand.[8] Subsequently, we discuss two extensions to their model, from which we derive testable hypotheses for our empirical analysis.

In the AAHR model, establishments e produce output $y_e(x)$ by combining two kinds of inputs, labor and capital, in form of AI technologies. Each of these inputs performs tasks, which are necessary to produce output. The allocation of tasks across labor and AI technologies is subject to firms' profit-maximizing behavior and depends, among others, on input-specific productivity. For example, as AI technologies become more productive, firms will find it profitable to automate some production steps. Consequently, they allocate tasks, previously performed by workers, to AI technologies. Assuming perfect substitutability between labor and AI technologies, AAHR then show that the implications for labor demand are primarily characterized by two competing forces: the displacement effect and the productivity effect.

[7] See Hershbein & Kahn (2018), Deming & Kahn (2018), Modestino, Shoag & Ballance (2019).

[8] This conceptual framework is rooted in the pioneering models of **??**, which highlight the channels through which new technologies affect wages and employment.

On the one hand, as AI technologies become more productive, firms find it profitable to expand the set of tasks performed by these technologies. This reallocation comes at the expense of labor, however, as the displacement effect reduces labor demand. For example, the increasing prevalence of chatbots, such as ChatGPT, may reduce demand for customer service agents (Korinek 2023). On the other hand, adoption of AI technologies allows firms to generate cost savings, resulting from automation of certain production steps. Establishments can subsequently employ a more flexible allocation of tasks. These efficiency gains generate a positive *productivity effect*, which raises labor demand. Revisiting our chatbot example, these bots can increasingly handle routine queries, thereby allowing customer service agents to perform more complex tasks, such as consulting tasks.

The net effect of AI adoption on labor demand thus depends on the relative magnitude of the displacement and productivity effect. The magnitude of this labor demand shift is further magnified by an establishment's exposure to AI, i.e. the share of tasks that can be profitably performed by AI technologies. The AAHR framework emphasizes the impact of AI technologies on labor demand and offers important insights on the underlying forces operating through firms. In contrast, we are primarily interested in the implications for workers, as they are themselves increasingly exposed to AI. To this end, we focus on two extensions to the AAHR framework, from which we derive testable hypotheses. .

First, we explore the implications of rising AI exposure on wages. The same forces that operate on labor demand should also translate into changes in wages. And these wage implications should be exacerbated, the higher the exposure to AI technologies. Building upon the mechanisms laid out in this section, we present our first hypothesis:

Hypothesis 1: *Workers with higher exposure to AI technologies experience stronger wage changes. A positive change in wages is consistent with a relatively strong productivity effect, while a negative change is consistent with a relatively strong displacement effect.*

Second, AAHR assume worker homogeneity and perfect substitutability between labor and AI technologies in their framework. These assumptions are merely for simplicity, yet, unlikely to be true in reality —especially in the context of AI. What distinguishes AI technologies from previous technologies is a higher level of autonomy in performing tasks and their ability to perform more cognitively demanding tasks (e.g. higher-quality predictions and recommendations, assistance in high-stakes decision-making).[9] For these reasons, scholars have hypothesized that AI technologies may affect the allocation of tasks in a different manner than previous technologies, possibly affecting high-skilled workers disproportionately (Webb 2020, Felten et al. 2021, **?**).

Hence, we argue that the relative magnitude of productivity and displacement effects will differ, depending on worker's skill level. Consequently, we expect heterogeneous shifts to labor demand, as workers are differentially exposed to AI, and thus heterogeneous wage implications. This logic leads to our second hypothesis:

Hypothesis 2: *High-skilled workers face stronger wage changes resulting from higher AI exposure because AI technologies can perform cognitively more demanding tasks than previous automation technologies and thus affect these workers disproportionately.*

Through the lens of our second hypothesis, skill-specific implications on wages are consistent with heterogeneity in the relative size of displacement and productivity effects among different skill groups. We test both of our hypotheses by running wage regressions as a function of AI exposure, which we construct from online job vacancy data. In the next section, we proceed by describing our data and the construction of our AI exposure measure.

[9] See citet* for references

3 Data

In this section, we present our data and outline the steps necessary to identify AI skills. At the core of our study is the near-universe of German online job vacancies. We use data between January 2017 and June 2023 to identify AI skills, but restrict ourselves to 2017 - 2021 for our analysis (due to linkage of OJV with administrative data). We first present a description of our OJV data, including (i) a general overview, (ii) outline of NLP steps, (iii) sample selection, (iv) aggregation procedures, and (v) a discussion on external validity. In a second step, we outline the construction of our AI exposure measure, capturing demand for AI skills, and details on our identification strategy of AI skills from job postings. Lastly, we present our administrative data, comprising information on workers' wages.

3.1 Online Job Vacancies

3.1.1 General overview

Job postings are collected by our partner —Finbot AG, an IT-company from Meerbusch, Germany. Finbot is a subsidiary of Palturai GmbH, from Hofheim, Germany, and offers custom-made firm-, person- and job posting- data and market analysis. To this end, they scrape vacancies from job boards, company websites, temporary employment agencies, and head-hunters. Finbot consistently updates their online sources and scrapes all sources on a daily basis. Subsequently, Finbot performs basic cleaning procedures and removes duplicates from the same source (i.e. sources from the same url address).

Our OJV data offers some key advantages compared to other vacancy data commonly used in economic research. Often, researchers purchase preprocessed data, leaving ambiguities about underlying data quality. While we also receive our data from a commercial provider, our data has two key features. First, we have access to the original job vacancies, including all text included in the posting. This unique access allows us to have more control over the data-generating process and to develop our own, transparent taxonomies. Second, Finbot merges job-posting firms with the German company registry ("Handelsregister"), which is possible for about 60% of the job postings. This linkage allows us to supplement firms' vacancy contents with a plethora of firm characteristics, e.g. their industry affiliation (5-digit level, WZ08) and location.[10]

3.1.2 Brief Outline: NLP steps

Upon receiving the data from Finbot, we link firm and vacancy information and perform necessary steps to preprocess the textual data, following conventions in the literature Gentzkow, Kelly & Taddy (2019), Ash & Hansen (2023) In particular, we tokenize the texts, lowercase tokens, and remove special characters. Beyond these basic steps, we enrich the data as follows. First, we assign each vacancy to a specific location, either at the zip code (39% of OJVs), municipality-level (48%), or county-level (10%).[11] Overall, we can thus assign 97% of job postings to a specific county. Second, we classify job titles according to the German Classification of Occupations 2010 (KldB2010). For this purpose, we use official, codified job titles at the 8-digit level, which are provided by the Federal Employment Agency (BA). Extracting job titles from our vacancies, and comparing their job description with the BA, we can immediately assign job titles to 3-digit occupations for about 60% of vacancies. In a follow-up step, we classify the remaining job titles by annotating a sample of not-yet-classified vacancies. **Here sentence for % of OJV for which we find KldB**. We will provide more details on NLP steps in Appendix A.1 in our upcoming draft.

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

[11] About 10% of job postings lack specific working place location information (typically smaller companies operating in one specific region). In such cases, we use the address provided in the imprint as the basis for regional allocation

3.1.3 Sample Selection

For our main analysis, we limit ourselves to the years 2017 - 2021 to match availability of our administrative data (see section 3.4 for details). Within this time horizon we only use vacancies advertising regular work, i.e. full- or part-time. To this end, we remove vacancies seeking apprenticeships, trainees, and other types of irregular work.

To focus on high-quality vacancies, we exclude job postings with fewer than 50 and more than 1,000 tokens. Our experience suggests that vacancies outside of this range do not represent standard job advertisements and instead add unnecessary noise to the data. **update and what %do we loose by applying our restrictions?** In this context we also drop postings for temporary employment and large recruitment agencies because these firms typically search for employees with more flexible work schedules and therefore advertise somewhat broader job descriptions and requirements.[12] Similarly, to focus on "regular firms", we restrict our data to vacancies from companies that can be linked to the business register. Lastly, we omit observations with missing information on either date, location or occupation.

After cleaning and selecting the relevant data, we are left with 8.3 million job vacancies from 242,000 firms. To put this data product in perspective, keep in mind there was a total of reported vacancies in Germany between 2017-2021 of 11 million vacancies.

3.1.4 Aggregation

We aggregate our OJV data at the LMR-occupation-year level and set the cutoff date to the 30th of June of each year. We do so in order to accommodate a merge with administrative data (details below), To this end, we aggregate 402 counties into 141 broader labor market regions (LMRs) that reflect commuting zones, following the classification of Kosfeld & Werner (2012).[13] To be consistent with previous studies, we drop postings for jobs in the armed forces

[12]See Stops, Bächmann, Glassner, Janser, Matthes, Metzger & Müller, Christoph, Seitz, Joachim (2021) for a detailed discussion on this issue.

[13]This classification has been used widely in research on LRMs in Germany. See, e.g., (Dauth, Findeisen, Suedekum & Woessner 2021, Dustmann, Lindner, Schönberg, Umkehrer & vom Berge 2021, Hirsch, Jahn, Manning & Oberfichtner 2022)

and in agriculture, fishing and forestry.

Subsequently, we combine these 141 LRMs with 3-digit occupations (KldB2010) and define the relevant local labor market at the occupation-LMR-Year level. To ensure that our results are not driven by outliers, we retain only those LMR-occupation combinations with at least 3 postings in a year.[14] Following these steps we are left 122 occupations and a total of (141 $\times$ 122) 17,202 local labor markets.[15] **Perhaps keeping only LMR-occups with ALWAYS > 3 OJVS per cell, but then maybe lose too many?**

3.1.5 External Validity

In Appendix A.2 we provide extensive external validity on our data quality. This analysis encompasses comparisons with two sources: (i) the German Job Vacancy Survey (JVS), a representative survey on reported job vacacncies, which is carried out by the Institute for Employment Research (IAB), and (ii) the "BA-Jobbörse", an employment website comprising job openings reported to the Federal Employment Agency (BA).[16] For brevity, we limit ourselves to three key takeaways in this section.

First, we demonstrate our OJV data depicts trends in the number of posted vacancies between 2017 - 2021 that mirror those from the JVS. Specifically, our data depicts an increasing trend of vacancies over time, but with a sharp decrease at the beginning of the COVID-19 pandemic in 2020 and a subsequent rebound of postings. Second, we show our data has representative regional coverage, when compared to the JVS. To this end, we illustrate similar trends in job postings between West and East Germany, along with state-level comparisons. Third, compared to the JVS and BA, our OJV data is tilted towards high-skilled jobs, requiring specialized expertise in jobs in professional services and similar white-collar occupations. While online job vacancies represent only of many search channels, they are by far the most

[14] **Fupnote mit Hinweis zu robustness check**.

[15] Other, related studies that use OJV data use similar definitions for local labor markets, see, e.g., Azar, Marinescu, Steinbaum & Taska (2020) and Schubert, Stansbury & Taska (2022).

[16] For details on the JVS, see Bossler, Gürtzgen, Kubis, Küfner & Popp (2021). Similarly, for details on the BA-Jobbörse, see Stops et al. (2021).

important channel through which firms recruit high-skilled workers (Carrillo-Tudela, Kaas & Lochner 2023)—including those required to possess AI-related skills. While the concentration on high-skilled jobs limits the representative nature of our data, it is also particularly suitable to identify AI skills.

3.2 Construction of AI exposure

Next, we describe the empirical counterpart to worker's AI exposure from section 2. Within the AAHR-framework, this measure reflects the share of tasks that can be performed by AI technologies. AAHR approximate this exposure, using distinct patent- and crowdsource-based indicators provided by existing literature (Brynjolfsson, Mitchell & Rock 2018, Webb 2020, Felten, Raj & Seamans 2021). We instead approximate AI exposure with the share of vacancies in a worker's relevant local labor market that demand AI skills. We refer to these vacancies throughout this paper as "AI vacancies".[17] While the theoretical counterpart to our exposure measure is task-based, our skill-based measure is closely related to this concept (see Acemoglu & Autor (2011) for a detailed discussion).[18]

We argue our vacancy-based exposure measures provides a more immediate measure to assess wage implications, compared to, say, patent-based data. It usually takes a lot of time for newly patented technologies to be used widely and unclear to what extent a technology is actually being adopted by firms. In comparison, OJV data offers near real-time insights on the diffusion of technologies in a worker's relevant labor market, and thus a more immediate measure for AI exposure.

[17] We follow conventions in this young literature and define a vacancy an "AI vacancy" if we find at least one AI skill (Alekseeva, Azar, Giné, Samila & Taska 2021, Acemoglu, Autor, Hazell & Restrepo 2022).

[18] Workers do perform tasks to produce output. However, in order to do so, they apply their skill endowments. Hence, AI vacancies are informative on the type of jobs in which workers may compete or collaborate with AI technologies. The more firms in their respective local labor market demand those skills, the stronger the diffusion of AI technologies. We view this slight deviation from our conceptual framework suitable for two reasons.

3.3 Identification of AI skills

Having established our vacancy-based measure on AI exposure, we now proceed with details on our strategy to identify AI skills. To this end, we combine a keyword-based approach with manual annotation and assistance by ChatGPT 4.0. First, we create an initial keyword list of 97 AI skills, using keywords that have previously been used in the literature.[19] However, this list lacks information on (i) more recent innovations in the AI Space, e.g., methods deployed with transformer-based models (Devlin, Chang, Lee & Toutanova 2019), (ii) specific tools commonly deployed to perform AI tasks, e.g. Python packages, and (iii) jargon commonly used in vacancies, e.g. abbreviations, German descriptions, etc. In a second step, we thus manually annotate a random sample to validate and extend existing AI skill keywords. After these adjustments, we end up with 140 relevant AI skills that we use for our analysis.

Figure 1 displays a word cloud, highlighting the most relevant AI skills. It shows the most important skills are applied to broad concepts in machine learning, data mining, or deep learning. These methods summarize algorithms, methods, and software libraries commonly deployed in AI. In addition, we also find more specific applications, which comprise specific domains in which AI skills are applied to, especially in the context of autonomous driving. See Table B1 in Appendix B for a full overview of our keywords.

We acknowledge our definition of AI skills is broader than most existing taxonomies and may thus be more susceptible to "false positive", i.e. erroneously classifying certain skills as AI skills. In our robustness section we present more narrow (and conservative) definitions and provide comparisons to alternative definitions from the literature. Importantly, neither the main message of our stylized facts nor our empirical results change fundamentally once we adopt alternative keywords, lending credence to our identification of AI skills. We provide more details on the validity of our AI skill measure in Appendix C.1.

[19] See, e.g., Büchel, Demary, Goecke, Kohlisch, Koppel, Mertens, Rusche, Scheufen & Wendt (2021), Bessen, Cockburn & Hunt (2021), Taska, O'Kane & Nania (2022), Acemoglu, Autor, Hazell & Restrepo (2022). Especially Büchel, Demary, Goecke, Kohlisch, Koppel, Mertens, Rusche, Scheufen & Wendt (2021) is helpful for this exercise because this is the only comprehensive keyword list with German and English keywords, to our knowledge.

3.4 Administrative Data and Summary Statistics

To facilitate our analysis of rising AI exposure on worker-level wages, 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 1975-2021.[20] 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 IAB and provides information on daily labor market spells, wages, and basic socio-economic characteristics (e.g., sex, nationality, education).[21] We use the June 30th of each year as a cutoff date.

As is common in administrative data, wage information is top-censored. Censoring affects about five percent of all spells, though, some skilled groups are more heavily affected (Dauth & Eppelsheimer 2020).[22] 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 terms of sample selection, we focus on full-time 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 of work, establishment and occupation. Currently, the data is available up until 2021. We thus restrict our OJV data to 2017-2021 to match data availability of the SIAB.

We further supplement the SIAB with data from the Establishment History Panel (BHP), comprising all establishments covered by the IAB employment history. We use information on employment, industry, and the location of work of establishments between 2017-2021.

[Table 1 here]

need to update summary stats Table::: ES will do that tonight

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

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

[22]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.

Table 1 provides summary statistics for workers in our sample. We provide several indicators that characterize workers in terms of skill: In terms of formal education, 71% of workers earned a vocational degree, while 23% are college graduates. Using detailed occupational codes, instead, suggests 57% of workers are skilled professionals. In comparison, 32% of workers are highly-skilled specialists or experts in their respective field.[23] Using the task structure of occupations, we observe very similar patterns. In particular, only 18% and 8% of (predominantly high-skilled) workers are employed in cognitively demanding occupations intensive in analytic and, respectively, interactive tasks.[24] We also observe that around two thirds of workers are men, primarily due to our sample restriction to full-time workers. Moreover, we see that 85% of workers are employed in medium-sized or large firms with at least 250 employees —especially in business organization jobs (26%), professional services (18%), and the public sector (18%) —and in large metropolitan areas (54%).

4 Stylized Facts: Diffusion of AI skills

In this section, we provide key stylized facts on the diffusion of AI skills in Germany from 2017 - 2021. To this end, we focus on the three dimension along which we exploit variation in demand for AI skills: (i) over time, (ii) across regions, and (iii) across occupations.

4.1 Demand for AI skills has increased by 12.6 % YoY between 2017-21

Figure 2 illustrates the evolution of AI demand from 2017/01 - 2021/12, illustrating the monthly share of AI vacancies, i.e. vacancies with at least one AI skill. The average share

[23] To facilitate this comparison, we use the fifth digit of the KldB 2010 classification, which assigns (broad) occupations into four skill groups: unskilled, skilled, specialist, expert.

[24] The data on the task structure of occupations is not originally included in the SIAB, but can be easily merged onto, using indicators from the German Occupational Panel (Dengler, Janser & Lehmer 2023). This data describes job characteristics and comprises, among others, the same (conceptual) task measures outlined in AAHR. The IAB Occupational Panel derives job descriptions from the BERUFENET database, comprising detailed job characteristics and maintained by the German Federal Employment Agency, and can be downloaded free of charge from the IAB webpage: `https://iab.de/en/daten/iab-occupational-panel/`.

of AI vacancies increased from 1.2% in 2017 to 2.1% in 2021, implying an annualized year-on-year (YoY) growth rate of 12.6%.

[Figure 2 here]

This upward trend aligns broadly with related literature, though we find higher shares of AI vacancies. Taska, O'Kane & Nania (2022) report the share of AI vacancies in Germany increased from 0.6% in 2017 to about 1% by 2021. Because we adopt a more extensive keyword list (outlined in section 3.3), our taxonomy implies that the share of AI vacancies has been about twice as large as in the existing literature. Once we adopt the original keyword list by Taska et al. (2022) (excluding German translations), our share of OJV vacancies merely increases from 0.5% to 1%, thus consistent with their findings and those from other studies (see Figure B1 in Appendix B.2).[25]

Overall, we find that alternative definitions of AI skills primarily lead to level differences. Importantly, however, they all display similar dynamics over time. This distinction is important because variation over time will be essential for our identification strategy (rather than the "exact" measurement of the level of AI vacancies). We show that our main results are robust to a variety of alternative definitions of AI skills in Appendix B.

4.2 Demand for AI skills has diffused broadly across regions

[Figure 3 here]

Having established intuitive trends over time in AI skill demand, we now illustrate regional diffusion of AI skills. Figure 3 displays the average share of AI vacancies for each of our 141 LRMs in 2017 and 2021. To this end we partition each LMR into one out of four equally sized groups. The first quartile comprises the 35 regions with highest share of AI

[25] For example, replicating Acemoglu, Autor, Hazell & Restrepo (2022), by using their exact keywords, the share of AI vacancies increases from 0.35% in 2017 to about 0.55% by 2021, implying a YoY growth rate of 8.8%. More recently, Borgonovi, Calvino, Criscuolo, Samek, Seitz, Nania, Nitschke & O'Kane (2023) document an increase in AI vacancies in Germany from 0.3% to 0.4% between 2019 - 2022. All these studies use similar taxonomies to identify AI skills, which are provided by Lightcast.

vacancies ("High-AI regions"). The three remaining quartils comprises regions with lower share of AI vacancies.

In 2017, demand for AI skills was concentrated in urban regions, especially in the Southwest, with large clusters around Berlin, Munich, and Stuttgart.[26] At least 1% of all vacancies in these regions in the first quartile required AI skills. In the remaining regions, AI demand was still mostly negligible.

To illustrate the regional diffusion of AI, we perform a descriptive counterfactual exercise. To this end, we reproduce the same map for the year 2021, but keep using the same boundaries (from 2017) to divide the four types of regions. Doing so, we find a broad diffusion across German regions. In fact, around 50% of LMRs would have been classified a "High-AI" region in 2017 in this counterfactual exercise. Overall, there are only a handful of regions, which would (in 2021) still be placed in the lowest quartile in 2017, in which AI has barely diffused, mostly in rural parts of East Germany.

4.3 Demand for AI skills is concentrated among "Pioneer Occupations"

Similar to the regional diffusion of AI skills, we finally assess the occupational diffusion. For this purpose, we first compute the share of AI vacancies for each of our 122 3-digit occupations for each year. The horizontal axis in Figure **??** displays these shares for our first period in 2017. We contrast this baseline level of occupational AI exposure the change in the share of AI vacancies between 2017 - 2021, depicted along the vertical axis.

This comparison shows a strong positive correlation ($\rho = 0.43$), suggesting that occupations with initially higher AI exposure are also those that experienced stronger exposure to AI demand in subsequent years. In particular, we identify 11 "AI Pioneer" occupations, characterized by high baseline demand for AI skills in 2017 (80th percentile) and high sub-

[26] Gathmann & Grimm (2022) likewise document a concentration of AI adoption in the Southwest of Germany, using patent data on AI innovations.

sequent change between 2017-21 (80th percentile). These occupations comprise: Teachers and researchers at universities and colleges (Overall share of AI vacancies: 7.2%), Computer science (7.1%), Mathematics and statistics (6.9%), Software development and programming (6.8%), IT-application-consulting (5.0%), Technical research and development (4.5%), Product and industrial design (4.4%), IT-network engineering, IT-coordination, administration (4.3%), IT-system-analysis, Business organisation and strategy (2.7%), Theatre, film, television productions (2.1%), and Teachers at educational institutions other than schools (1.6%).

For around 70% of occupations, however, we find little to no demand for AI skills throughout our time horizon. Compared to regions, diffusion of AI demand across occupations has thus been much more concentrated. We provide a full overview of all occupations and their average share of AI vacancies from 2017-2021 in Table B2 in Appendix B.

[Figure 4 here]

5 Empirical Analysis

In this section, we empirically test the implications of our conceptual background, presented in section 2. In particular, we aim to test our first hypothesis, in which we argue that workers with higher AI exposure also experience stronger wage changes. To this end, we first run an OLS wage regressions accounting for worker FE and a rich set of control variables and fixed-effects. Second, we perform an IV estimation to address potential endogeneity concerns. Specifically, we use a leave-one-out instrument that exploits national trends in AI skill demand that are plausibly exogenous to local trends.

5.1 OLS Wage Regressions

5.1.1 Methodology

We begin our analysis by running OLS regressions, using the log daily wage w_{ilot}[27] for worker i, working in LMR l, and employed in occupation o in year t, as outcome variable:

$$lnw_{ilot} = \alpha_i + \beta_1 AI_{lot} + \beta_2 X_{it} + \beta_3 \psi_l + \beta_4 \omega_o + \beta_5 \theta_t + \epsilon_{ilt} \tag{1}$$

Our key covariate is A_{lot} which captures the share of vacancies requiring AI skills —our measure of AI demand —in each year t in regional labor market l and occupation o. We control for a rich set of covariates in X_{it} at the worker level, comprising socioeconomic characteristics (age, education, gender, nationality), work experience controls (firm tenure), and employer-related controls (firm size, industry 2-digit according to WZ08) [28], and employer quality —approximated by AKM effects [29].

To account for unobserved heterogeneity at the individual level we also include worker FE (α_i). By including LMR FE (ψ_l) at the commuting zone level (141 LMR), we also control for region-specific trends pertaining to productivity and technology adoption. For similar reasons, we include occupational FE (ω_o) at the 3-digit KldB (122 occupations). In addition, we include year FE (θ_t) to capture year-specific shocks such as the COVID-19 shock. Therefore, we exploit differential variation in demand for AI skills over time within a worker's LMR.

We are primarily interested in the sign and magnitude of β_1. Through the lens of our conceptual background (section 2), this coefficient is informative on the relative size of the displacement and productivity effect. We interpret $\beta_1 > 0$ as being consistent with a relatively strong productivity effect. Similarly, we consider $\beta_1 < 0$ to be consistent with a

[27] Since wages in the SIAB are censored above the contribution assessment ceiling, we use imputed wages.

[28] We combine smaller WZ08 2-digit industry groups to ensure a sufficient number of observations in each group. We end up with 40 different industry categories

[29] We use AKM effects for the time period 2007 to 2013 to avoid reverse causality of AI exposure on firms' productivity

comparably strong displacement effect.

In various robustness tests, we perform a battery of alternative specifications to check for model misspecification. In particular, we address three potential concerns. First, we use alternative definitions of our AI measure to alleviate concerns regarding mismeasurement of our key regressor. Second, we run more flexible specifications to explicitly account for region- and occupation-specific shocks (via LMR $\times$ Year FE and Occupation $\times$ Year FE). Third, we estimate the model in changes. One potential concern of our baseline model are rigidities in the level of wages, which may mask potential negative effects (consistent with a strong displacement effect). A specification in changes circumvents these issues. Overall, these robustness tests leave our key takeaways unchanged. We provide more details on these tests in section 7.

5.1.2 Results

We report our baseline results on demand of AI skills on worker-level wages in Table 2. All specifications include our worker-level controls, summarized in X_{it}, and worker FE. In the first column we add year FE to account year-specific shocks. Our point estimate of 0.09 shows positive wage implications in response to rising AI Diffusion. Adding LMR FE barely changes this estimate (column 2). However, once we additionally control for occupational FE, our point estimate drops to 0.073 (column 3). This observation suggests that occupational variation has a stronger impact on wages than regional variation, presumably because AI skill demand is more concentrated along the occupational dimension (see Figure 4). Evaluated at the mean value of AI Demand (0.009), this estimate implies that a 10% increase in AI Demand is associated with a wage increase of 0.66%.[30]

[30]For easier interpretation, we examine the impact of a 10% increase in AI demand, evaluated at its mean level. To normalize our coefficients accordingly, we first determine the percentage increase in AI demand that corresponds to a unit increase (i.e., in percentage points). Subsequently, we divide our point estimates by a normalizing factor that corresponds to a 10% increase in AI demand. For example, the mean level of AI demand in our baseline sample is 0.009. A one-unit increase corresponds to a move from 0.009 to 0.019, implying that our point estimate of 0.073 reflects an increase in AI demand by 111% (= 0.01/0.009). Dividing 0.073 by 11.1 then permits interpretation in response to a 10% increase in AI demand, yielding 0.0066.

Overall, our OLS results consistently show modest but positive effects of increasing AI diffusion on worker-level wages. These results support our first hypothesis from section 2, which argues that workers with higher AI exposure experience higher wage changes. Through the lens of our conceptual background, these estimates support the hypothesis that the demand for AI skills leads to a productivity effect. However, we emphasize that our baseline results should be interpreted with caution.

OLS results are helpful to learn about the relationship between wages and AI demand, but they can not be interpreted as capturing the causal impact of rising AI demand. Demand for AI skills is likely endogenous because skill demand is a function of underlying technology adoption. For example, we know from related literature that AI adoption is mainly concentrated among large, productive firms (Alekseeva, Azar, Giné, Samila & Taska 2021, Rammer, Fernández & Czarnitzki 2022). While we control for various differences in firm, LMR, and occupational characteristics in our OLS specifications to provide remedy against these identification threats, none of these measures resolves the inherent endogeneity problem. For this reason, we supplement our OLS analysis with an instrumental variable approach that permits causal interpretation.

[Table 2 here]

5.2 IV Regressions

5.2.1 Identification

To overcome the endogeneity concerns in our OLS specifications, we construct an instrument that exploits national trends in AI Demand that are plausibly exogenous to local conditions. To this end, we instrument for AI exposure in occupation o in LMR l by calculating the leave-one-out-mean (LOOM) of AI skill demand for each occupation, AI_{ot}, excluding its demand in workers' home LMR $l \neq l'$, i.e. $AI^{IV}_{lot} = \sum_{l \neq l'} AI_{lot}$.

Using our LOOM variable, we instrument for endogenous AI demand with a two-stage-least-squares (2SLS) approach:

$$lnw_{ilot} = \beta_1 AI^{IV}_{lot} + \beta_2 X_{it} + \beta_3 \psi_l + \beta_4 \omega_o + \beta_5 \theta_t + \epsilon_{ilt} \tag{2}$$

Our identifying variation comes from changes in national AI demand over time. Essentially, our LOOM instrument exploits the variation illustrated in Figure 2, showing that the (national) share of AI vacancies doubled between 2017 and 2021. To identify the causal effect of rising AI exposure on worker-level wages, our identifying assumption requires that firms' national skill requirements are orthogonal to local conditions. For this to be true, our LOOM instrument must satisfy the relevance condition and exclusion restriction, which we discuss in detail below.

I. Relevance Condition

First, to meet the relevance condition, national trends in AI demand must be positively correlated with wages of workers employed in occupation o in LMR l. LOOM instruments are, by construction, highly correlated with the endogenous variable, which is one of the reasons they have become increasingly popular in economic research (Schubert, Stansbury & Taska 2022, **?**). Intuitively, we expect high relevance of our instrument due to broad technology diffusion across regions (see Figure 3, implying high correlation between national and local trends in AI demand.

For our LOOM instrument to be relevant, we require firms to post in many LMRs. Otherwise, their national and local skill requirements collapse into one and suppresses diffusion of AI demand across regions. To provide supporting evidence for this claim, we inspect firms' posting behavior in more detail. To this end, we split firms into those that post AI vacancies in a given year ("AI firms") and those that do not post AI vacancies. This comparison underlines systematic differences in job posting behaviour between firms with differential (presumed) technology adoption. We report the distribution of the number of LMRs in which firms post in Figure 6. For the full sample, we see that the number of firms which post vacancies in only one or many LMRs (>= 10) is broadly balanced (Panel 6a). In contrast, 85% of AI firms post vacancies in ten or more LMRs.

[Figure 6 here]

These discrepancies in posting behavior between AI firms and Non-AI firms is primarily attributed to differences firm size. It is well-known that AI adoption is pronounced among large firms (Rammer, Fernández & Czarnitzki 2022) and we can support this stylized fact as well. Defining large firms as those with more than 275 employees, Figure B3 shows that 5-7% of all large firms post AI vacancies at any given time between 2017 - 2021. In contrast, only 2-3% of smaller firms post AI vacancies during this time.

Since large firms operate in many LMRs and are the primary adopters of AI technologies, they likely contribute substantially to national trends in AI demand —and thus the broad regional diffusion of AI skills. Our findings thus suggest that AI diffusion operates in large parts through those (few) large firms. Hence, we expect high relevance of our instrument, which we further validate by reporting the F-statistics of the first stage.

II. Exclusion Restriction

Second, to satisfy the exclusion restriction, national trends in AI demand must not affect local wages through any other channel than technology diffusion. The key concern with our identification strategy is that other confounding factors affect wages, but for unrelated reasons to AI demand. Especially concurrent demand shocks (e.g. globalization, local policies aimed at promoting AI adoption, ...) are hard to disentangle from technology diffusion, but are likely to have an impact on wages as well. Omission of these confounding factors shows up as unobserved residual in our IV specification, i.e. in ϵ_{ilt}, and thus biases our estimates. For example, positive demand shocks would also have a positive impact on wages. If present, these shocks may also be driving forces behind our baseline results.

To address this concern, we perform a placebo test in which we test the validity of shock orthogonality (see Borusyak, Hull & Jaravel (2022) for details). In our research design, we exploit national trends in AI Demand as "shocks". Running a regression with baseline characteristics that were realized *prior* to those shocks should thus result in null effects, conceptually similar to "pre-trend" test in Diff-in-Diff settings. In our context, we regress wages earned *prior* to 2017 on our AI measure, which we observe from 2017 onward. Formally, we run a modified version of our 2SLS specification:

$$lnw_{ilo,2012-16} = \alpha_i + \beta_1 AI_{lo,2017-2021} + \beta_2 X_{it} + \beta_3 \psi_l + \beta_4 \omega_o + \beta_5 \theta_t + \epsilon_{ilt} \tag{3}$$

Here, we regress wages from the period 2012 - 2016 on our AI demand between 2017 - 2021, such that wages in 2012 are regressed on AI demand in 2017 and so on. To support our claim that our research design satisfies the exclusion restriction, we require the null hypothesis $\beta_1 = 0$ to be true, i.e. no "pre-trends". Indeed, our results in Table 4 shows we cannot reject the null hypothesis in our preferred specification. Pre-trends would be a problem in specifications, which omit occupation FE (columns 1 and 2). However, once we control for occupation FE, we cannot reject the null hypothesis of no pre-trends (column 3). Running more flexible specifications by including Year × LMR FE (column 4), Year × Occupation FE (column 5), and LMR × Occupation FE (column 6) provide the same conclusions. Because our baseline IV specification accounts for LMR FE and Occupation FE, we find no evidence that rejects the validity of our LOOM instrument.

5.2.2 Results

We report our IV results on the impact of rising AI demand on worker-level wages in Table 3. For comparability, we also include the corresponding OLS results. Overall, OLS and IV specifications provide qualitatively similar results, reinforcing a positive link between rising AI demand and worker-level wages. In quantitative terms, IV results display coefficient sizes more than twice as large as those using OLS. For example, consider our preferred specification, using LMR FE and occupational FE (column 3).

The IV point estimate of 0.17 implies that a 10% increase in AI demand is associated with a wage increase of 1.5%. In contrast, OLS results imply only a wage increase by 0.7%. This comparison suggests that confounding factors tend to depress wages, causing attenuation bias. One might argue these discrepancies are driven by the COVID-19 pandemic. Yet, we control for Year FE in each of our specifications, making it unlikely that our results are circumstantial. We cannot rule out other confounding shocks are still present, for example productivity shocks concentrated in a broader region and encompassing many but not all regions. While these concerns warrant caution regarding interpretation, our identification tests in the section 5.2.1 make us confident that our IV results are informative on AI-induced implications on wages.

[Table 3 here]

6 Mechanisms

In this section, we build upon our baseline results and shed light on underlying mechanisms. Grounded in our conceptual framework in section 2, we hypothesize that the impact of AI on wages is heterogeneous across workers. In particular, in our second hypothesis, we argue that the complementarity between AI technology and higher-skilled workers leads to larger wage increases. Accordingly, we focus first on the role of skill heterogeneity.

In addition, we explore the role of monopsony power, using concentration of job postings as proxy (as in Azar, Marinescu, Steinbaum & Taska (2020)). Adverse wage implications of monopsony power are well-documented[31], yet, may even be exacerbated in the context of AI because these technologies are primarily adopted by large firms (see section 5.2.1). Workers who are exposed to concentrated AI demand, i.e. AI skills only demanded by few firms within their LMR, may thus not fully capitalize from the productivity effect. However, if workers are mobile, either in terms of regional or occupational mobility, they may circumvent

[31]See Manning (2021) for a recent overview.

constraints stemming from monopsony power. Hence, we will also investigate heterogeneity with respect to worker mobility.

To assess these heterogeneities, we run modified variations of our baseline specification, equation (3), as follows:

$$lnw_{ilot} = \alpha_i + \beta_1 AI_{lot} + \beta_2 GR_o + \beta_3 AI_{lot} \times GR_i + \beta_4 X_{it} + \beta_5 \psi_l + \beta_6 \omega_o + \beta_7 \theta_t + \epsilon_{ilt} \quad (4)$$

where we interact our AI measure with different group indicators GR_i. where we interact our AI measure with different group indicators GR_i. To do this, we assign workers to different groups and compare wage outcomes across these groups. To facilitate the interpretation of the interaction effects, we calculate marginal effects for each group. This allows us to identify heterogeneities in the relative size of productivity and displacement effects and how they differentially affect workers' wages.

6.1 Skill Heterogeneity

According to our second hypothesis in section 2 wage increases will be stronger for high-skilled workers due to complementarities with new technologies. To analyse skill heterogeneity, we explore several measures of skill. First, we use the fifth digit of the occupational code (KldB2010), which provides information on the underlying task complexity of an occupation and encompasses the range of tasks, problem-solving abilities, and relevant knowledge domains. Second, we consider education, which serves as a proxy for their formal training, knowledge acquisition and cognitive development. Third, we examine age because it embodies experience and accumulated human capital.

[Figure 7 here]

Starting with occupational task complexity, we distinguish between four skill groups: unskilled, professional, specialist and expert. Figure 7 shows the marginal effects on wages for each of these skill levels. Indeed, we find that the association between demand for AI skills and wages is stronger for higher-skilled workers, suggesting that our baseline results mask substantial heterogeneity across skill levels. In particular, we find significant marginal effects of 0.14 for specialists and 0.10 for experts. The mean AI exposure is higher among these groups as well, with a mean value of 0.015 and 0.02 for specialists and, respectively, experts. Evaluated at their respective mean AI exposure, these results imply that a 10% increase in AI exposure corresponds to a wage increase of 2.1% for both, professionals and specialists.[32] For skilled workers, the estimated coefficient is slightly positive while we find no effect for unskilled workers. Figure 8 illustrates the results by task group and underscores the complementarity between AI technology and cognitive tasks. It shows that the effect is largest for cognitive occupations.

These results provide strong evidence in favor of our second hypothesis (see section 2), in which we argue higher-skilled workers face stronger wage changes resulting from higher AI exposure. In fact, our findings suggest that our baseline results are entirely driven by specialist and expert workers —a group that comprises only 33% of all workers in our sample. While the productivity effect appears to be concentrated among highly skilled workers, we find no evidence for a displacement effect.

[Figure 9 and 10 here]

Next, we look at the heterogeneity between educational groups as they proxy requirements for formal job training. Figure 9 shows the marginal effects by education group. As expected, we find that positive wage implications associated with rising AI exposure are concentrated among workers with a university degree. The marginal effect for these workers is 0.16. Evaluated at the mean value of AI demand for this group (= 0.018), this estimate implies that a 10% increase in AI skill demand is associated with higher wages on the order

[32]Note that AI exposure is mainly concentrated among specialists and expert workers. For expert workers, the average AI exposure is 0.02, which means that 2.0% of vacancies in their relevant local labour market require AI skills. Similarly, the average AI exposure for professionals is 0.015. In contrast, skilled (unskilled) workers face an average AI exposure of 0.006 (0.003). It is not surprising that highly skilled workers are more exposed to AI, given that these technologies can perform more cognitively demanding tasks than previous automation technologies (Felten, Raj & Seamans 2021, **?**, Webb 2020)

of 2.8%. For workers without university degree, we find no significant relationship between AI exposure and wages.

These findings broadly mirror our results on the skill levels. Similarly, college graduates are substantially more exposed to AI demand than other workers (0.018 vs 0.006) and also represent a minority —23% of the workforce. These results are all consistent with our key hypotheses, in which we posit that workers with more AI exposure experience higher wage implications and that high-skilled workers are more exposed. Therefore, our results are consistent with relatively stronger productivity effects of college graduates. However, we find no evidence for a displacement effect.

Next, we study age profiles, thus shedding light on the role of experience and accumulated knowledge. We distinguish three age groups: (i) young workers: 18-29, (ii) prime-Age workers: 30-49, and (iii) older workers: 50-65. We then interact these age groups with our AI share, using young workers as our reference group. Figure 10 displays substantial heterogeneities, despite all age groups facing similar levels of AI exposure (0.008 - 0.010). In particular, we find a sizeable point estimate of 0.28 for young workers (reference group) and net effect of 0.18 for prime-age workers. Evaluated at their respective mean AI exposures, a 10% increase in AI demand is associated with a 2.8% wage increase for young workers and, respectively, a 2.2% increase for prime-age workers. In contrast, we find a negative estimate for old workers. Accordingly, a 10% increase in AI demand corresponds to a 1.2% wage decrease for older workers.

These differences between younger and older workers suggest that AI skills are not complementary to work experience. We are able to confirm the pattern of the results when we repeat the analysis by interacting our variable for the demand for skills with an indicator for the experience groups (see Figure 11 in Appendix XC).[33]. Deming & Noray (2020) show that skill obsolescence lowers returns to experience, implying flat age-earnings profiles. This deterioration is especially pronounced in occupations associated with STEM degrees, such

[33] We distinguish between three experience groups: (i) workers with less than 5 years of experience, with experience between 5 and 15 years, and workers with more than 15 years of experience

as computer science. Importantly, these are exactly the type of jobs in which AI demand has diffused more intensely. Hence, young workers may be the primary beneficiaries of rising AI demand because their (newly acquired) skills are valued more than the (devalued) skills of older workers.

Alternatively, young workers may also benefit from long-lasting structural change. Dauth, Findeisen, Suedekum & Woessner (2021) document that robot exposure in Germany has been associated with a reallocation away from manufacturing jobs towards services. Moreover, this transformation has contributed to a substitution away from vocational training towards college education. Combined, these shifts towards higher rates of college and jobs in the service industry may have pushed young workers into higher-quality jobs, which benefit from AI diffusion disproportionately.

6.2 Market power of firms

In this section, we examine the role of local monopsony power. One channel through which firms can exercise monopsony power is labor market competition. We follow the literature and proxy the concentration of the labor market with a standard Herfindal-Hirschmann index (HHI), which we construct using the OJV data (Azar et al. 2020, Schubert et al. 2022, ?). The construction of the index is described in detail in Appendix X.[34]. By interacting the AI share with the HHI, we can examine whether the demand for AI skills affects wages differently in unconcentrated versus concentrated markets.

[Figure 12 here]

Indeed, our results in Figure 12 lend credence to this hypothesis. For ease of interpretation, we plot the marginal effects at specific points in the HHI values to show the

[34]Overall, our HHI likely overstates concentration to some extent because we do not have data on all firms, but only those that post vacancies. Accounting for non-posting firms as well would thus reduce concentration levels. Our HHI measure must thus be interpreted as an upper bound. Nonetheless, our measure is broadly in line with estimates from the literature. We thus view it informative to study the impact of AI skill on demand on wages in LMRs with varying degrees of concentration.

effects for less concentrated markets ($.01 < HHI < .15$), moderately concentrated markets ($.15 < HHI < .25$), and highly concentrated markets ($HHI > .25$). Figure 12 shows that the positive relationship between AI demand and wages is indeed muted in more concentrated markets.

Assessing the impact on wages at the mean AI share of 0.009 for the different labor markets, our results suggest that a 10% increase in AI demand is associated with a wage increase of 1.1% percent in unconcentrated markets, 0.9% in moderate concentrated markets, and only 0.7% in high concentration markets. These findings are consistent with monopsony power and corroborate with Azar, Marinescu, Steinbaum & Taska (2020).

Our findings have important policy implications. AI technologies are predominantly adopted by large firms (Rammer, Fernández & Czarnitzki 2022), hence these firms are also those that demand AI skills disproportionately (see Figure **XXX**). Because larger firm usually have higher monopsony power (**?**), there is reason to believe that rising AI adoption may reinforce labor market imperfections. We leave this open question for future research.

6.3 Mobility

It is well known that mobility can have positive effects on wages as workers take advantage of (better) external opportunities (Schubert, Stansbury & Taska 2022, **?**) as workers sort into LMR and occupations with higher wage prospects. In this section, we examine the role of regional and occupational mobility, as well as the role of labor market entrants. In our baseline specification, we allow for mobility across LMRs and occupations and do not restrict the analysis to incumbents. To examine the role of mobility, we impose different restrictions on mobility to test the extent to which the results are driven by labor market entrants, workers moving between LMRs and switching occupations.

For the analysis, we impose the following stepwise restrictions: First, we restrict the analysis to incumbents; second, we restrict the analysis to incumbents who stay within the same LMR, while still permitting occupational mobility within that LMR; third, we restrict the analysis to incumbents who remain in the same occupations while allowing for regional mobility; fourth, we restrict the analysis to incumbents who remain in the same LMR and occupation from 2017 to 2023.

[Figure 13 here]

Figure 13 illustrates the results of this exercise in comparison to our baseline results. Restricting the sample to incumbent workers has no effect on our results, implying that our results are not driven by labor market entrants. In addition, restricting regional mobility slightly reduces the estimated coefficient to 0.064, but restricting occupational mobility further reduces the estimated coefficient to 0.061. Imposing all restrictions reduces the estimate to 0.057.

These results suggest that occupational and regional mobility plays a role and that the ability to seek a better, higher paying job is partly driving our results. However, the size of the 95% confidence intervals of the estimates overlap. Therefore, we cannot reject that the estimates are not significantly different from each other.

Our findings on the importance of occupational mobility are related to those of Schubert, Stansbury & Taska (2022), who document that workers in occupations with a high degree of mobility experience less downward pressure on their wages due to labor market concentration. This points to the role of better outside options. Mobility can further help improve the matching quality between worker and firms. More broadly, the results also point to evidence that mobility can further help to improve the quality of matching between workers and firms. (**?**) show that high-quality workers are more attracted to larger cities, leading to greater geographic wage inequality (**?**). Importantly, this relationship is even stronger when local labor markets are defined at the city-occupation level (rather than simply cities).

7 Robustness

We perform a series of robustness checks to ensure that our results are not due to mismeasurement of our key variables or misspecification, such as omitted variables. Specifically, we address mismeasurement concerns using two strategies: First, we use alternative definitions of our measure of AI. Second, we validate our measure of AI exposure by comparing it to alternative measures of Technology 4.0. In addition, we conduct several robustness checks to address misspecification concerns: First, we run more flexible specifications to explicitly account for region- and occupation-specific shocks. Second, we estimate the model in terms of changes rather than levels to account for potential rigidity in the wage level that would underestimate negative effects. Third, we repeat the analysis using a more restrictive sample to account for the potential confounding effect of sample selection. The details of the robustness checks are discussed in Appendix C.

Our robustness checks help mitigate concerns about mismeasurement of our key AI measure. First, we test different specifications of our measure to rule out that our results are driven by specific industries (services or manufacturing), the inclusion of additional keywords, and broad AI keywords that do not specifically capture AI capabilities in a narrower definition. In addition, as an alternative approach, we also use an AI intensity measure that considers the number of AI skills within a job posting, rather than simply categorizing postings as either AI or non-AI. However, running the baseline specification using the alternative measure provides comparable estimates, we can show that our results are not industry-specific, nor are they driven by the inclusion of specific keywords or the use of an intensity-based measure. In addition, we show that our results remain robust even after accounting for the demand for skills related to Technology 4.0. This suggests that our measure of AI captures aspects distinct from the demand for skills related to 4.0 technologies.

In addition, we are able to mitigate concerns about misspecification and omitted variable bias. First, using a more flexible specification that accounts for occupation- and region-specific time trends reduces the estimates but maintains the positive estimate. Second, the model in changes yields larger estimates and rejects the concern that we are overestimating our effect by estimating in levels. Finally, we show that our estimates are not qualitatively affected when we impose further sample restrictions on the SIAB and OJV data. In particular, we find that our point estimates increase, especially when we restrict our analysis to LMR occupation cells with a limited number of postings across all years.

8 Conclusions

In this paper we study the diffusion of AI demand and perform detailed analysis on its worker-level wage implications. Our analysis combines the near-universe of German online job vacancies from 2017 - 2021 with administrative data. 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. We use NLP methods to measure AI skill demand from the vacancy data and merge this measure to administrative data at a detailed occupation-region level. Subsequently, we use regression-based methods to asses the impact of rising AI demand on wages. To address endogeneity on AI demand, we propose an instrument that exploits national trends in AI skill demand, which are plausibly orthogonal to local conditions.

Our key finding is that AI demand mostly has positive implications on wages. In our most restrictive model, a panel FE estimation, our IV estimates suggest that a a 10% increase in AI skill demand is associated with a 2% wage increase. OLS results, in turn, display more mixed evidence. We provide suggestive evidence, indicating this is due to demand shocks that attenuate the causal impact of AI. To identify the key drivers of our results, we perform a decomposition, revealing that 95% of AI-induced wage results are attributed to three overarching characteristics: (i) Employer quality, (ii) Occupational, and (iii) Socioeconomic characteristics. Using the results of our decomposition, we subsequently inspect the three key drivers in more detail to shed light on key mechanisms. This analysis reveals that workers who, at least so far, have benefited from rising AI demand are those that: (i) have had only moderate AI exposure, (ii) have a vocational degree, (iii) are older aged (50+ years), (iv) display occupational mobility, and (v) are employed at a high-quality firm. In contrast, workers negatively affected are those in highly exposed occupations, with highly specialized skills ("experts"), and those in labor markets with high concentration in labor demand. Potentially vulnerable groups include young workers, who experience negative AI-induced wage results, but also women, who benefit less from rising AI skill demand than men.

Our findings provide many important insights for policymakers. We highlight two in particular. First, we document widespread diffusion of AI skill demand. And while there is substantial heterogeneity in AI-induced wage results, we also show that occupations with high AI demand are generally of high quality (e.g., in terms of average wage, outside options). This feature makes these occupations desirable for workers and, combined with expected increase in AI adoption in coming years (Schaller, Wohlrabe & Wolf 2023), requires skill investments. Especially in light of detrimental effects on young workers, educational institutions, such as universities and vocational schools, should support these investments with targeted skill development in their curricula. Second, Germany, like many other countries, faces severe labor shortages. We have shown that occupational mobility is associated with wage gains. Therefore, policies aimed at supporting job mobility can help workers to escape negative AI-induced wage implications and move into occupations with acute shortages.

Future research can help guide these policy measures. Our paper suggests many important avenues for such research. To help guide policies on job mobility, future research may explore skill transferability in more detail. While we do provide a detailed analysis on AI skills, we are agnostic on their skill transferability. Combining our focus on AI skills with a conceptual framework that allows to study the portability of skills, as in Gathmann & Schönberg (2010), is a fruitful avenue for future research. We have also shown that concentration of AI demand has negative wage implications. Since AI technologies are disproportionately adopted by large firms with more labor market power, these firms may use this technology to exercise more market power. We believe, studying AI adoption and imperfect competition will be a key research area in coming years. Lastly, our results demonstrate the importance of employer quality. Workers employed at high-quality firms benefit much more from rising AI skill demand than other workers. Unfortunately, our data does not allow us to study worker-level effects in response to firm-level adoption of AI technologies. Future research should take a closer look at the role of firms, e.g. by combining our approach with survey-based approaches on firm-level adoption of digital technologies.[35].

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. & Autor, D. (2011), Skills, tasks and technologies: Implications for employment and earnings, Vol. 4 of *Handbook of Labor Economics*, Elsevier, pp. 1043–1171. 12

Acemoglu, D., Autor, D., Hazell, J. & Restrepo, P. (2022), 'Artificial intelligence and jobs: Evidence from online vacancies', *Journal of Labor Economics* **40**(S1). 3, 4, 5, 12, 13, 16, 70

Albanesi, S., Dias da Silva, A., Jimeno, J. F., Lamo, A. & Wabitsch, A. (2023), New technologies and jobs in europe, Working Paper 31357, National Bureau of Economic Research. Revision Date: July 2023. 3

Alekseeva, L., Azar, J., Giné, M., Samila, S. & Taska, B. (2021), 'The demand for ai skills in the labor market', *Labour Economics* **71**, 102002. 4, 12, 21

Antoni, M., Graf, T., Grießemer, S., Kaimer, S., Köhler, M., Lehnert, C., Oertel, M., Schmucker, A., Seth, S., Seysen, C. & vom Berge, P. (2019), 'Schwach anonymisierte version der stichprobe der integrierten arbeitsmarktbiografien (siab) – version 7519 v1', *Forschungsdatenzentrum der Bundesagentur für Arbeit (BA) im Institut für Arbeitsmarkt- und Berufsforschung (IAB)* . 14

[35]See Arntz, Genz, Gregory, Lehmer & Zierahn-Weilage (2023), Genz, Gregory, Janser, Lehmer & Matthes (2021) for notable examples of this young literature

Arntz, M., Genz, S., Gregory, T., Lehmer, F. & Zierahn-Weilage, U. (2023), Task shifts in the fourth industrial revolution – firm-level evidence. Mimeo. 4, 35

Arthur, R. (2021), 'Studying the uk job market during the covid-19 crisis with online job ads', *PloS one* **16**(5), e0251431. 54

Ash, E. & Hansen, S. (2023), 'Text algorithms in economics', *Annual Review of Economics* **15**, 659–688. First published as a Review in Advance on July 5, 2023. 9

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

Azar, J., Marinescu, I., Steinbaum, M. & Taska, B. (2020), 'Concentration in us labor markets: Evidence from online vacancy data', *Labour Economics* **66**, 101886. 4, 11, 25, 29, 30

Bamieh, O. & Ziegler, L. (2020), How does the covid-19 crisis affect labor demand? an analysis using job board data from austria, IZA Discussion Paper 13801, IZA - Institute of Labor Economics. 54

Barth, E., Bryson, A. & Dale-Olsen, H. (2022), Creative disruption: Technology innovation, labour demand and the pandemic, IZA Discussion Paper 15762, IZA - Institute of Labor Economics. 2

Bessen, J., Cockburn, I. & Hunt, J. (2021), Is distance from innovation a barrier to the adoption of artificial intelligence, Technical report, Boston University Working Paper. 3, 4, 13

Bloom, N., Hassan, T. A., Kalyani, A., Lerner, J. & Tahoun, A. (2021), The diffusion of disruptive technologies, Working Paper 28999, National Bureau of Economic Research. 4, 71

Borgonovi, F., Calvino, F., Criscuolo, C., Samek, L., Seitz, H., Nania, J., Nitschke, J. & O'Kane, L. (2023), Emerging trends in ai skill demand across 14 oecd countries, OECD Artificial Intelligence Papers 2, OECD. 16

Borusyak, K., Hull, P. & Jaravel, X. (2022), 'Quasi-experimental shift-share research designs', *Review of Economic Studies* **89**(1), 181–213. 2, 23

Bossler, M., Gürtzgen, N., Kubis, A., Küfner, B. & Popp, M. (2021), 'Iab-stellenerhebung – version 0018 v1', Forschungsdatenzentrum der Bundesagentur für Arbeit (BA) im Institut für Arbeitsmarkt- und Berufsforschung (IAB). 11

Brynjolfsson, E., Li, D. & Raymond, L. R. (2023), Generative ai at work, Working Paper 31161, National Bureau of Economic Research. 1

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. 3, 12

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'. Gutachten im Auftrag des Bundesverbandes Digitale Wirtschaft (BVDW) e.V. 13

Carrillo-Tudela, C., Kaas, L. & Lochner, B. (2023), Matching through search channels, Discussion Paper 16583, IZA. 12

Copestake, A., Marczinek, M., Pople, A. & Stapleton, K. (2023), Ai, firms and wages: Evidence from india. Mimeo. 4

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). 14

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, 10, 29

Deming, D. J. & Noray, K. (2020), 'Earnings dynamics, changing job skills, and stem careers*', *The Quarterly Journal of Economics* **135**(4), 1965–2005. 28

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.

5

Dengler, K., Janser, M. & Lehmer, F. (2023), 'The occupational panel for germany', *Journal of Economics and Statistics* **243**(6), 711–724. 15

Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. (2019), Bert: Pre-training of deep bidirectional transformers for language understanding, *in* 'Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)', Vol. 1, pp. 4171–4186. 13

Dillender, M. & Forsythe, E. (2022), Computerization of white collar jobs, Working Paper 29866, National Bureau of Economic Research. 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

Eloundou, T., Manning, S., Mishkin, P. & Rock, D. (2023), 'Gpts are gpts: An early look at the labor market impact potential of large language models', *arXiv preprint arXiv:2303.10130* . arXiv:2303.10130v5. 3

Felten, E., Raj, M. & Seamans, R. (2021), 'Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses', *Strategic Management Journal* **42**(12), 2195–2217. 3, 7, 12, 27

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. 14

Fossen, F. M. & Sorgner, A. (2022), 'New digital technologies and heterogeneous wage and employment dynamics in the united states: Evidence from individual-level data', *Technological Forecasting and Social Change* **175**, 121381. 2, 3

Gathmann, C. & Grimm, F. (2022), The diffusion of digital technologies and its consequences in the labor market.beiträge zur jahrestagung des vereins für socialpolitik 2022: Big data in economics, Conference paper, ZBW - Leibniz Information Centre for Economicss. 3, 17

Gathmann, C., Kagerl, C., Pohlan, L. & Roth, D. (2023), The pandemic push: Digital technologies and workforce adjustments, IZA Discussion Paper 16062, IZA - Institute of Labor Economics. 2

Gathmann, C. & Schönberg, U. (2010), 'How general is human capital? a task–based approach', *Journal of Labor Economics* **28**(1), 1–49. 34

Gentzkow, M., Kelly, B. & Taddy, M. (2019), 'Text as data', *Journal of Economic Literature* **57**(3), 535–574. 9

Genz, S., Gregory, T., Janser, M., Lehmer, F. & Matthes, B. (2021), How do workers adjust when firms adopt new technologies?, IZA Discussion Paper 14626, IZA - Institute of Labor Economics. 2, 4, 35, 71

Genz, S., Janser, M. & Lehmer, F. (2019), 'The impact of investments in new digital technologies on wages – worker-level evidence from germany', *Jahrbücher für Nationalökonomie und Statistik* **239**(3), 483–521. 2

Giering, O., Fedorets, A., Adriaans, J. & Kirchner, S. (2021), 'Künstliche intelligenz in deutschland: Erwerbstätige wissen oft nicht, dass sie mit ki-basierten systemen arbeiten', *DIW Wochenbericht* **48**, 783–789. 1

Gilardi, F., Alizadeh, M. & Kubli, M. (2023), 'Chatgpt outperforms crowd workers for text-annotation tasks', *Proceedings of the National Academy of Sciences* **120**(30), e2305016120. 3

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. 4

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. 54

Hensvik, L., Le Barbanchon, T. & Rathelot, R. (2021), 'Job search during the covid-19 crisis', *Journal of public economics* **194**, 104349. 54

Hershbein, B. & Kahn, L. B. (2018), 'Do recessions accelerate routine-biased technological change? evidence from vacancy postings', *American Economic Review* **108**(7), 1737–1772. 5, 53, 55

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

Koch, M., Manuylov, I. & Smolka, M. (2021), 'Robots and firms', *The Economic Journal* **131**(638), 2553–2584. 4

Korinek, A. (2023), 'Generative ai for economic research: Use cases and implications for economists', *Journal of Economic Literature* **61**(4), 1281–1317. 6

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

Manning, A. (2021), 'Monopsony in labor markets: A review', *ILR Review* **74**(1), 3–26. 25

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. 5, 55

Noy, S. & Zhang, W. (2023), 'Experimental evidence on the productivity effects of generative artificial intelligence', *Science* **381**(6654), 187–192. 1, 3

Peede, L. & Stops, M. (2023), Ai skills on the labour market. Mimeo. 4

Rammer, C., Fernández, G. P. & Czarnitzki, D. (2022), 'Artificial intelligence and industrial innovation: Evidence from german firm-level data', *Research Policy* **51**(7), 104555. 1, 4, 21, 23, 30

Rengers, M. (2018), 'Internetgestützte erfassung offener stellen: Machbarkeitsstudie im rahmen eines essnet-projekts zu big data', *WISTA - Wirtschaft und Statistik* **5**. 53, 54

Schaller, D., Wohlrabe, K. & Wolf, A. (2023), 'Ki, cloud computing und blockchain – wo steht die deutsche wirtschaft?', *ifo Schnelldienst* **76**(8), 3–9. 1, 34

Schubert, G., Stansbury, A. & Taska, B. (2022), 'Monopsony and outside options', *SSRN Journal* . 11, 22, 29, 30, 31

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. 54

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* . 10, 11

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* . 13, 16

Webb, M. (2020), The impact of artificial intelligence on the labor market. 1, 3, 7, 12, 27

Figures

Figure 1: Word clouds of AI Skills: Baseline definition

NOTE. —This word cloud comprises keywords that are associated with AI skills.

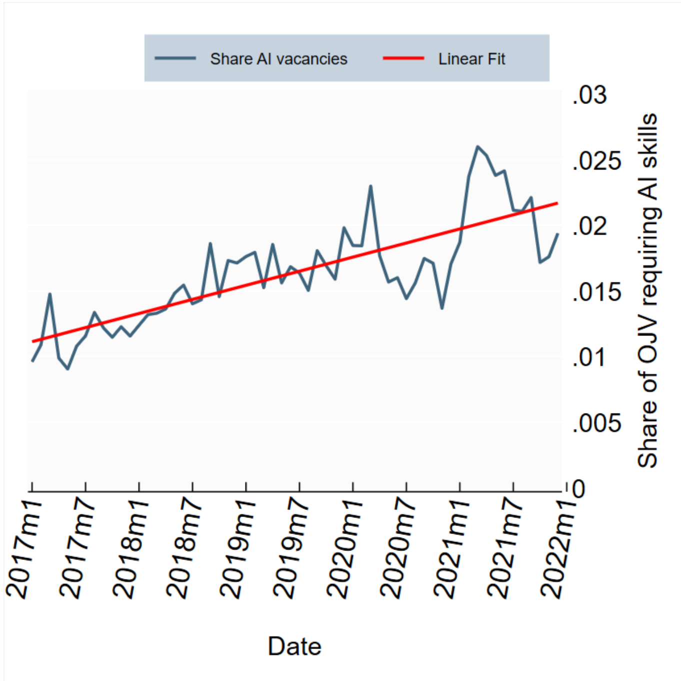
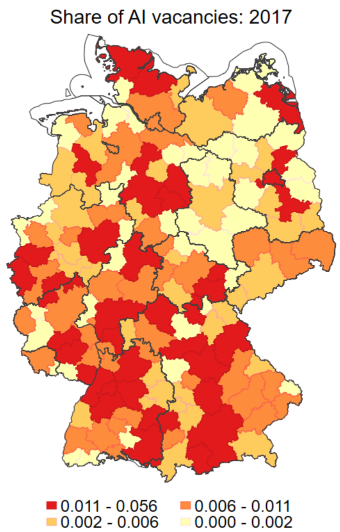
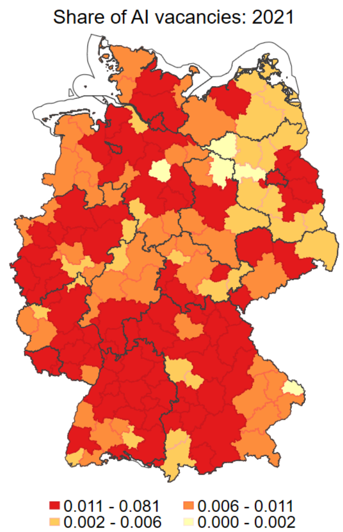


Figure 2: Trends in AI Demand, 2017/01 - 2021/12

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



(a) AI demand: 2017



(b) AI demand: 2021

Figure 3: Demand for AI skills in Germany across local labor markets, 2017-01 - 2021-12

NOTE. —Local labor markets are assigned into four classes of task intensity. Each class corresponds to quartiles as of 2017 where lowest quartile implies lowest AI demand (yellow) and highest quartile implies hightest AI demand (red).

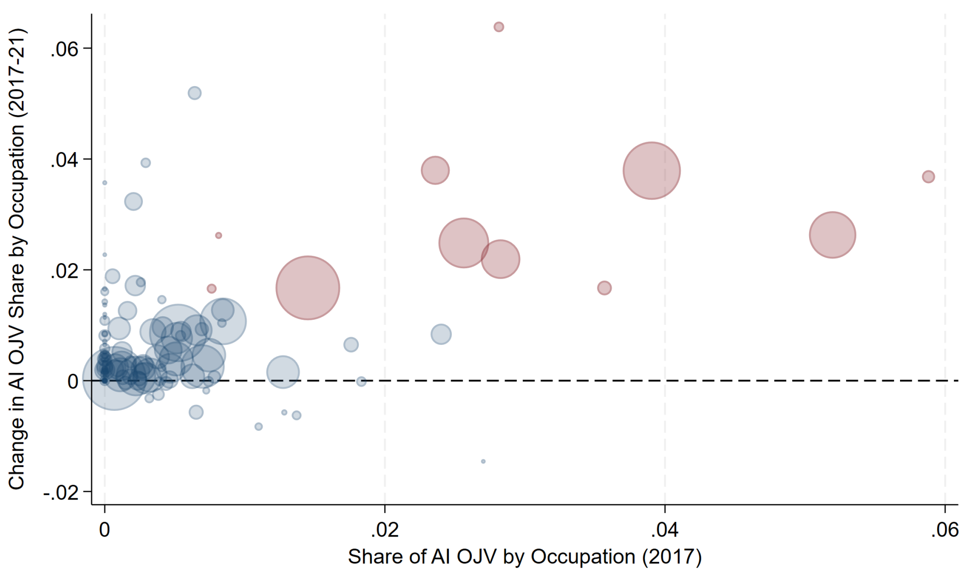
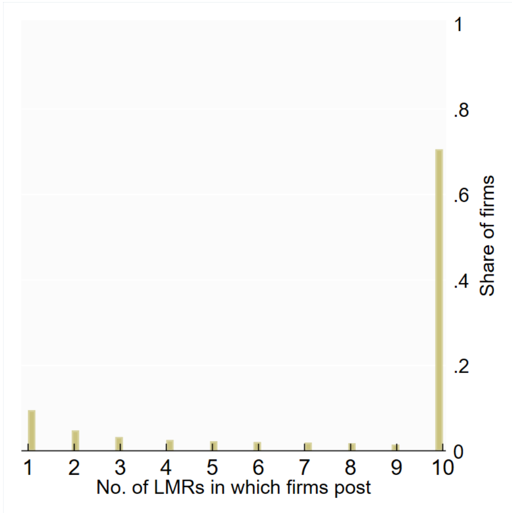
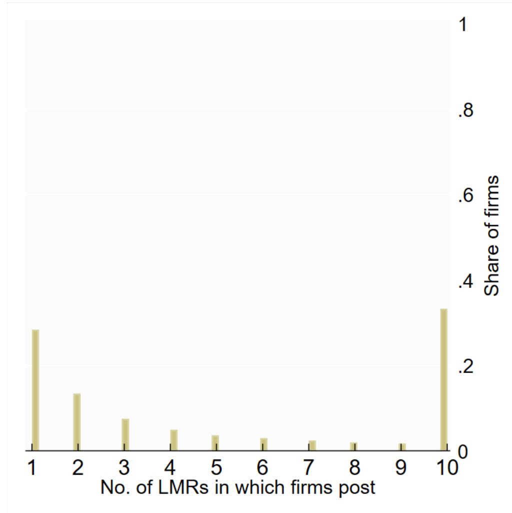


Figure 4: Dynamics in occupational demand for AI skills

NOTE. —The X-axis displays the share of OJV with AI demand ("AI Vacancies") for 140 3-d occupations as of 2017. The Y-axis displays the change in AI Vacancies between 2017 - 2021 for each occupation.



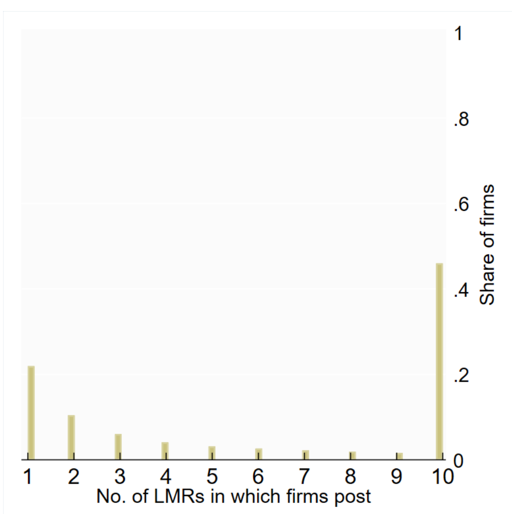
(a) Large firms



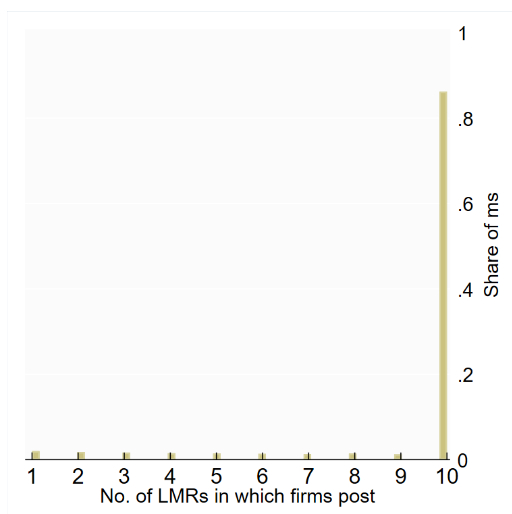
(b) Small firms

Figure 5: Distribution of number of local labor markets in which firms post vacancies: Large vs Small firms

NOTE. —Large firms are defined as those with more than 275 employees. Small firms have 275 or less employees. For ease of exposition, we cut off the number of LMRs at 10. Hence, firms included in these bars contain firms which post vacancies in at least ten LMRs.



(a) All firms



(b) Firms demanding AI skills

Figure 6: Distribution of number of local labor markets in which firms post vacancies: AI vs Non-AI firms

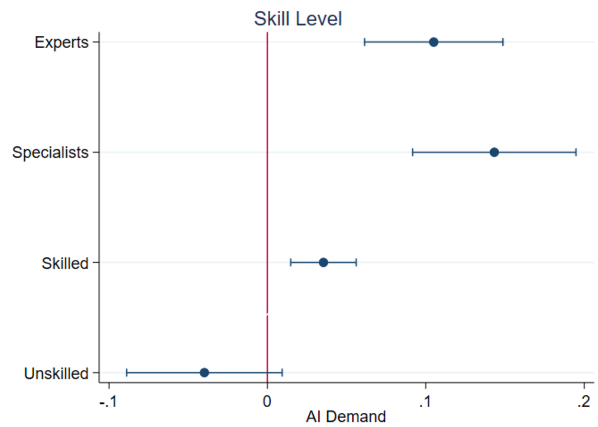


Figure 7: Marginal Effect of AI Demand by Occupational Skill Level

NOTE. —The figure shows the marginal effect of AI demand on wages by skill level. The horizontal lines present 95% confidence intervals. The model includes worker FE, year FE, LMR and occupation FE. The model also includes socioeconomic, work, and firm quality controls. Socioeconomic controls include age, citizenship, education, and gender. Work controls include LMR experience, establishment tenure, and job tenure. Firm controls include establishment size and industry (WZ08, 2-digit). Firm quality is the establishment AKM for 2007-2013.

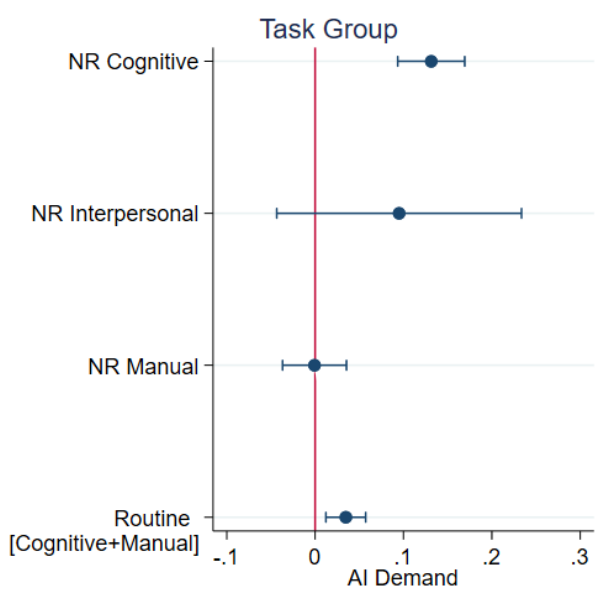


Figure 8: Marginal Effect of AI Demand by Task Group

NOTE. —The figures shows the marginal effect of AI demand on wages by task group. See notes Figure 7.

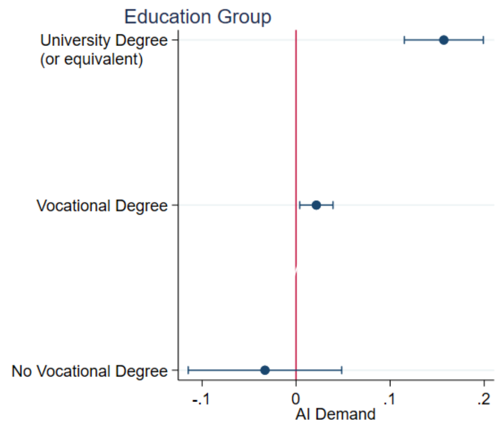


Figure 9: Marginal Effect of AI Demand by Education Group

NOTE. —The plot shows the marginal effect of AI demand on wages by education group. See notes to Figure 7.

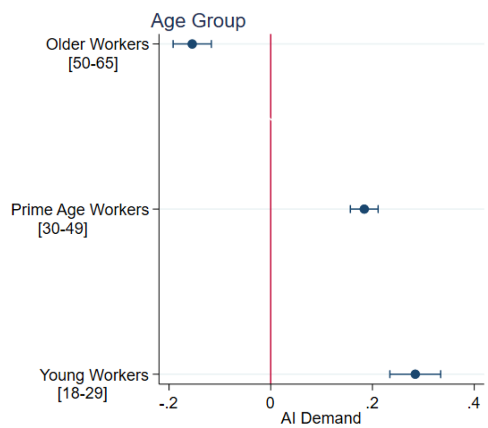


Figure 10: Marginal Effect of AI Demand by Age Group

NOTE. —The plot shows the marginal effect of AI demand on wages by age groups. See notes to Figure 7.

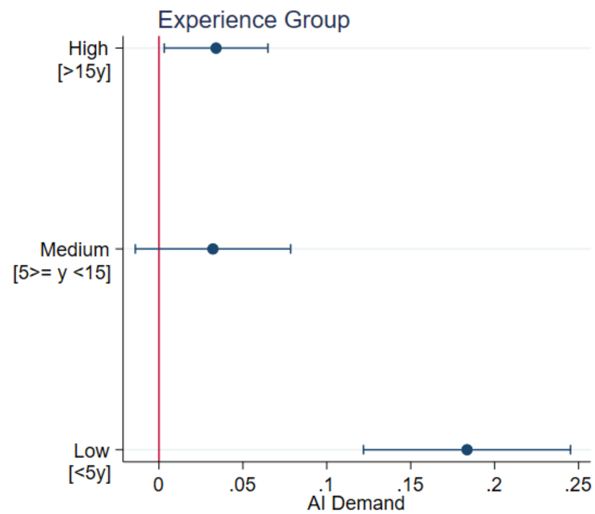


Figure 11: Marginal Effect of AI Demand by Experience Group

NOTE. —The plot shows the marginal effect of AI demand on wages by experience groups. See notes to Figure 7.

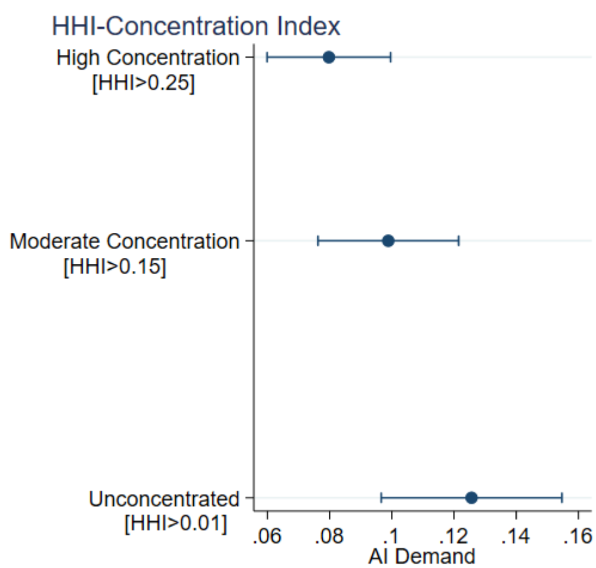


Figure 12: Marginal Effect at HHI Thresholds

NOTE. —The figures shows the marginal effect of AI demand on wages at specific HHI values. See notes Figure 7.

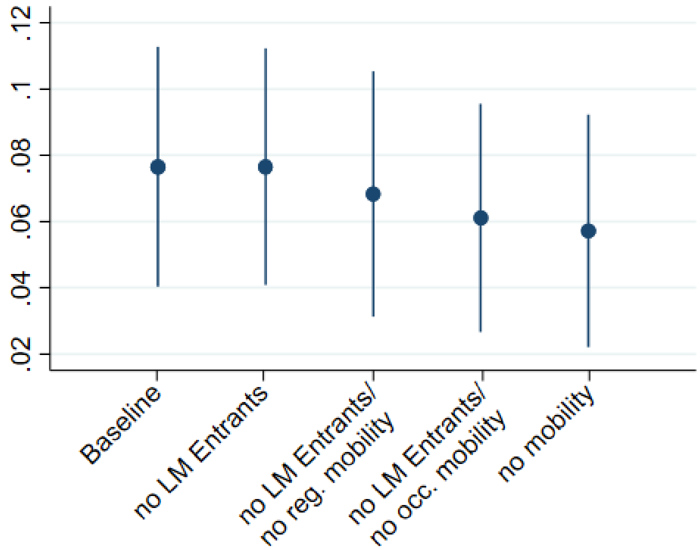


Figure 13: Coefficient Plot of AI Demand by Mobility Pattern

Tables

Table 1: AI exposure and average wages by socioeconomic characteristics

	(1) Share	(2) AI OJV Share	(3) log wage
Men	54%	1.3%	4.69
Women	46%	1.3%	4.53
College	22%	2.2%	4.96
Vocational	71%	1.1%	4.45
No degree	7%	1.0%	4.17
Young (18-29)	15%	1.3%	4.36
Mid-Old (30-49)	48%	1.4%	4.53
Old (50-65)	37%	1.2%	4.56
Observations	2,239,971	2,239,971	

NOTE. —The first column displays the share of workers in our sample with specific socioeconomic characteristics. The second column shows the group-specific AI exposure, defined as the share of vacancies with AI skills posted in their local labor market. The third column displays the average log daily wage for each group.

Table 2: Wage regressions AI Exposure: Occupation-LMR-level

	Dependent Variable: Log Wages		
	(1)	(2)	(3)
AI Share	0.089***	0.087***	0.073***
	(0.018)	(0.018)	(0.016)
Worker FE	✓	✓	✓
Year FE	✓	✓	✓
LMR FE		✓	✓
Occupation FE			✓
AI Share Mean		0.009	
Observations	1,500,688	1,500,688	1,500,688
R^2	0.89	0.89	0.89

Note: All specifications include the following controls: Socioeconomic controls include age, citizenship, education, and gender. Work controls include LM experience. Firm controls include establishment size and industry (WZ08, 2-digit). AKM effects for the time period 2007-2013.

Table 3: IV regressions AI Exposure: Occupation-LMR-level

	Dependent Variable: Log Wages		
	(1)	(2)	(3)
IV:AI Share (Occ-LMR)	0.270***	0.273***	0.172***
	(0.024)	(0.024)	(0.027)
OLS:AI Share	0.089***	0.087***	0.073***
	(0.018)	(0.018)	(0.016)
Worker FE	✓	✓	✓
Year FE	✓	✓	✓
LMR FE		✓	✓
Occupation FE			✓
AI Share Mean		0.009	
Observations	1,500,688	1,500,688	1,500,688

Note: All specifications include the following controls: Socioeconomic controls include age, citizenship, education, and gender. Work controls include LM experience. Firm controls include establishment size and industry (WZ08, 2-digit). AKM effects for the time period 2007-2013.

Table 4: Wage regressions AI Exposure: "Pre-Trends"

	Dependent Variable: Log Wages		
	(1)	(2)	(3)
AI Share	0.079***	0.074***	0.032
	(0.022)	(0.022)	(0.020)
Worker FE	✓	✓	✓
Year FE	✓	✓	✓
LMR FE		✓	✓
Occupation FE			✓
AI Share Mean		0.009	
Observations	1,275,961	1,275,961	1,275,958
R^2	0.93	0.93	0.93

Note: All specifications include the following controls: Socioeconomic controls include age, citizenship, education, and gender. Work controls include LM experience. Firm controls include establishment size and industry (WZ08, 2-digit). AKM effects for the time period 2007-2013.

Table 5: Wage regressions AI Exposure: Mobility

	Dependent Variable: Log Wages				
	(1)	(2)	(3)	(4)	(5)
AI Share	0.073***	0.073***	0.064***	0.061***	0.057***
	(0.016)	(0.016)	(0.017)	(0.017)	(0.018)
Worker FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
LMR FE	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
AI Share Mean			0.009		
Observations	1,500,699	1,458,002	1,305,227	1,223,314	1,156,076
R^2	0.90	0.89	0.90	0.90	0.90

Note: All specifications include the following controls: Socioeconomic controls include age, citizenship, education, and gender. Work controls include LM experience. Firm controls include establishment size and industry (WZ08, 2-digit). AKM effects for the time period 2007-2013.

Appendix

A Details on OJV Data

A.1 NLP Steps

TBD.

A.2 External Validity

Figure A1a 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.

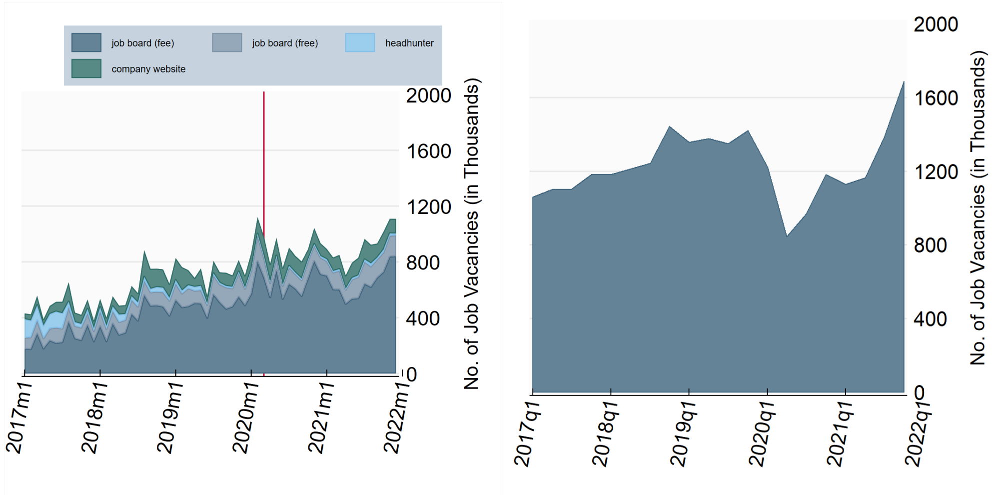
[Figure A1 here]

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 A1 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 (?). 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 A2 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 A2 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)

Figure A1: Number of online job vacancies over time, 2017/01 - 2021/12

NOTE. —Panel A1a 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 A1b 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.

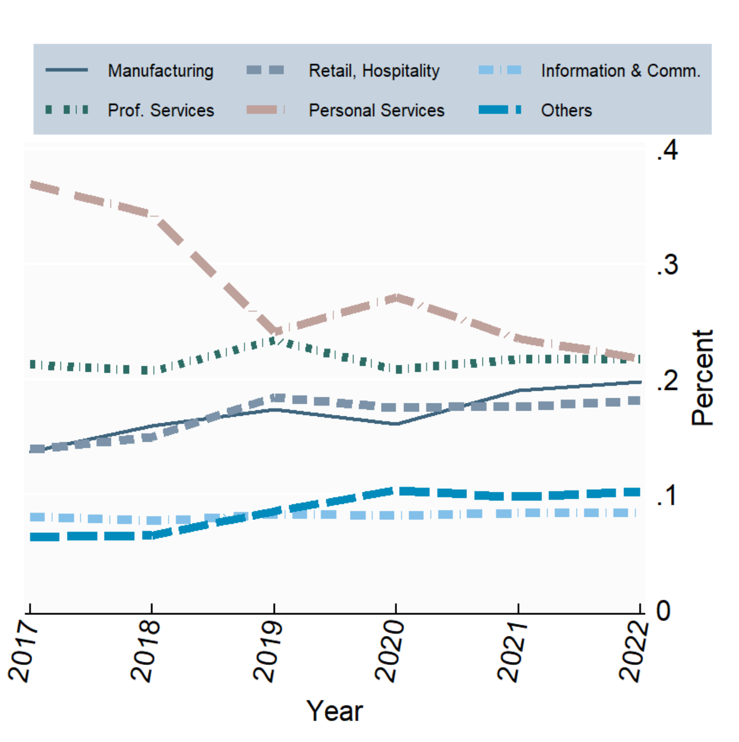


Figure A2: Industry composition of online job vacancies, 2017/01 - 2021/12

B Additional Descriptives

B.1 Overview: AI skill keywords and list of occupations

Table B1: Share of counts AI keywords

Rank	AI keywords	Share of counts
0	machine learning	15.02%
1	ki	11.75%
2	ai	9.39%
3	künstliche intelligenz	5.42%
4	chatbot	4.76%
5	artificial intelligence	4.37%
6	autonomes fahren	3.96%
7	spark	3.77%
8	data mining	3.69%
9	ml	3.60%
10	adas	3.40%
11	deep learning	2.63%
12	text mining	2.37%
13	predictive analytics	2.08%
14	computer vision	1.74%
15	maschinelles lernen	1.27%
16	tensorflow	1.24%
17	lidar	1.18%
18	autonomous driving	1.04%
19	machine vision	0.95%
20	ros	0.95%
21	nlp	0.92%
22	natural language processing	0.78%
23	spracherkennung	0.68%
24	new mobility	0.58%

Continued on next page

Table B1 – continued from previous page

Rank	AI keywords	Share of counts
25	ocr	0.55%
26	boosting	0.54%
27	pytorch	0.52%
28	sw design	0.50%
29	remote sensing	0.47%
30	bert	0.47%
31	asr	0.39%
32	neural networks	0.38%
33	keras	0.36%
34	neuronale netze	0.35%
35	adtf	0.34%
36	objekterkennung	0.32%
37	opencv	0.31%
38	reinforcement learning	0.29%
39	v2x	0.29%
40	gan	0.29%
41	structured data	0.29%
42	unstructured data	0.27%
43	transformer	0.24%
44	autonomous systems	0.24%
45	halcon	0.22%
46	cobots	0.19%
47	bilderkennung	0.14%
48	recommender systems	0.13%
49	caffe	0.13%
50	model validation	0.13%
51	flume	0.13%
52	predictive modeling	0.12%
53	abb roboter	0.11%

Continued on next page

Table B1 – continued from previous page

Rank	AI keywords	Share of counts
54	speech recognition	0.10%
55	supervised learning	0.10%
56	gpt	0.09%
57	image recognition	0.09%
58	nlu	0.09%
59	machine translation	0.09%
60	cobot	0.09%
61	sensorfusion	0.08%
62	vit	0.08%
63	motion planning	0.07%
64	data labeling	0.07%
65	neural network	0.07%
66	random forests	0.07%
67	v2v	0.07%
68	object tracking	0.07%
69	unsupervised learning	0.07%
70	electra	0.07%
71	bard	0.06%
72	object detection	0.06%
73	mxnet	0.06%
74	pattern recognition	0.06%
75	text to speech	0.06%
76	texterkennung	0.06%
77	model training	0.06%
78	sw implementation	0.06%
79	v2h	0.05%
80	feedback loop	0.05%
81	roberta	0.05%
82	decision trees	0.05%

Continued on next page

Table B1 – continued from previous page

Rank	AI keywords	Share of counts
83	random forest	0.05%
84	language models	0.05%
85	feature extraction	0.05%
86	elmo	0.05%
87	transfer learning	0.05%
88	dnn	0.05%
89	ros2	0.04%
90	gans	0.04%
91	humanoide roboter	0.04%
92	electromechanical systems	0.04%
93	maschinelle übersetzung	0.04%
94	autonome mobile roboter	0.04%
95	neuronale netzwerke	0.04%
96	federated learning	0.04%
97	gesichtserkennung	0.04%
98	chatgpt	0.04%
99	computervision	0.03%
100	adaptive learning	0.03%
101	text recognition	0.03%
102	torch	0.03%
103	path planning	0.03%
104	support vector machines	0.03%
105	dimensionality reduction	0.03%
106	image segmentation	0.03%
107	mahout	0.03%
108	fahrerlose transportfahrzeuge	0.03%
109	xgboost	0.03%
110	roboterarme	0.03%
111	automl	0.03%

Continued on next page

Table B1 – continued from previous page

Rank	AI keywords	Share of counts
112	automatic speech recognition	0.03%
113	entity recognition	0.03%
114	gradient boosting	0.03%
115	face recognition	0.02%
116	tokenization	0.02%
117	parking assistance	0.02%
118	nmt	0.02%
119	voice recognition	0.02%
120	natürliche sprachverarbeitung	0.02%
121	object recognition	0.02%
122	ai chatbot	0.02%
123	cluster analysis	0.02%
124	robot perception	0.02%
125	object classification	0.02%
126	synthetic data	0.02%
127	robot learning	0.02%
128	nltk	0.02%
129	simultaneous localization and mapping	0.02%
130	v2g	0.02%
131	collaborative robots	0.02%
132	meta learning	0.02%
133	adaptive cruise control	0.02%
134	opennlp	0.02%
135	cntk	0.02%
136	classification algorithms	0.01%
137	image generation	0.01%
138	sentiment analysis	0.01%
139	video generation	0.01%

Table B2: Adoption of AI in Various Occupations

Rank	Occupation (3-digit KLdB 2010)	AI vacancy share
1	Teachers and researchers at universities and colleges	7.21%
2	Computer science	7.07%
3	Mathematics and statistics	6.85%
4	Software development and programming	6.77%
5	IT-system-analysis, IT-application-consulting	4.97%
6	Technical research and development	4.48%
7	Product and industrial design	4.36%
8	IT-network engineering, IT-coordination/admin	4.28%
9	Technical media design	2.94%
10	Business organisation and strategy	2.69%
11	Physics	2.30%
12	Social sciences	2.27%
13	Theatre, film, television productions	2.10%
14	Laboratory medicine	1.92%
15	Editorial work and journalism	1.78%
16	Advertising and marketing	1.62%
17	Public relations	1.61%
18	Teachers at educational institutions other than schools	1.59%
19	Publishing and media management	1.58%
20	Construction scheduling, supervision, architecture	1.57%
21	Managing directors, executive board members	1.39%
22	Biology	1.36%
23	Event organisation and management	1.22%
24	Human resources management, personnel service	1.20%
25	Legal services, jurisdiction, court officers	1.18%
26	Insurance and financial services	1.15%
27	Media, documentation, information services	1.11%
28	Production planning and scheduling	1.09%
29	Purchasing and sales	1.07%

Continued on next page

Table B2 – continued from previous page

Rank	Occupation (3-digit KLdB 2010)	AI vacancy share
30	Economics	1.05%
31	Electrical engineering	1.00%
32	Photography, photographic technology	0.94%
33	Geology, geography, meteorology	0.93%
34	Draftspersons, technical designers, model makers	0.87%
35	Environmental protection management and consulting	0.87%
36	Accounting, controlling and auditing	0.85%
37	Plastic- and rubber-making and -processing	0.84%
38	Management assistants in transport, logistics	0.84%
39	Chemistry	0.81%
40	Beverage production	0.81%
41	Textile making	0.80%
42	Printing technology, print finishing, book binding	0.77%
43	Pharmacy	0.77%
44	Horsekeeping	0.77%
45	Mechatronics, automation, control technology	0.75%
46	Legislators, senior officials of interest organisations	0.74%
47	Automotive, aeronautic, aerospace, ship building	0.70%
48	Event technology, cinematography, sound engineering	0.66%
49	Machine-building and -operating	0.66%
50	Nutritional advice, health counselling, wellness	0.65%
51	Natural stone, minerals, building materials	0.64%
52	Metal-making	0.63%
53	Office clerks and secretaries	0.62%
54	Public administration	0.61%
55	Technical occupations in medicine, orthopaedic	0.57%
56	Precision mechanics and tool making	0.57%
57	Musicians, singers and conductors	0.51%
58	Surveying and cartography	0.51%

Continued on next page

Table B2 – continued from previous page

Rank	Occupation (3-digit KLdB 2010)	AI vacancy share
59	Metal constructing and welding	0.50%
60	Technical occupations in railway, aircraft, ships	0.49%
61	Doctors' receptionists and assistants	0.46%
62	Psychology, non-medical psychotherapy	0.45%
63	Production of clothing, textile products	0.44%
64	Interior design, visual marketing, interior decoration	0.44%
65	Hotels	0.44%
66	Building services, waste disposal	0.43%
67	Housekeeping, consumer counselling	0.42%
68	Sales selling drugstore products, pharmaceuticals	0.40%
69	Human medicine and dentistry	0.39%
70	Non-medical therapy, alternative medicine	0.38%
71	Energy technologies	0.36%
72	Driving, flying, sports instructors, educational inst.	0.35%
73	Artisans working with metal	0.35%
74	Tax consultancy	0.34%
75	Physical security, personal protection, fire safety	0.33%
76	Education, social work, pedagogic specialists	0.33%
77	Warehousing, logistics, postal, delivery services	0.33%
78	Driver of vehicles in road traffic	0.32%
79	Interior construction, dry walling, insulation	0.31%
80	Building services engineering	0.30%
81	Farming	0.29%
82	Industrial glass-making and -processing	0.28%
83	Real estate, facility management	0.27%
84	Gastronomy occupations	0.27%
85	Cleaning services	0.27%
86	Construction, transportation vehicles, equipment	0.26%
87	Painters, varnishers, plasterers, waterproofing	0.26%

Continued on next page

Table B2 – continued from previous page

Rank	Occupation (3-digit KLdB 2010)	AI vacancy share
88	Traffic surveillance and control	0.26%
89	Treatment of metal surfaces	0.26%
90	Colour coating and varnishing	0.25%
91	Underground, surface mining, blasting engineering	0.25%
92	Gardening	0.25%
93	Body care	0.23%
94	Wood-working and -processing	0.23%
95	Building construction	0.22%
96	Teachers in schools of general education	0.22%
97	Trading occupations	0.21%
98	Nursing, emergency medical services, obstetrics	0.21%
99	Foodstuffs, confectionery, tobacco production	0.21%
100	Geriatric care	0.21%
101	Metalworking	0.20%
102	Veterinary medicine, non-medical animal health	0.19%
103	Sales (retail) selling clothing, electronics, furniture	0.18%
104	Cooking occupations	0.18%
105	Actors, dancers, athletes, related occupations	0.17%
106	Leather- and fur-making and -processing	0.15%
107	Plumbing, sanitation, heating, air conditioning	0.15%
108	Service occupations in passenger traffic	0.15%
109	Animal husbandry	0.14%
110	Sales (retail) selling books, art, antiques	0.13%
111	Sales in retail trade (without product specialization)	0.12%
112	Civil engineering	0.12%
113	Tourism and the sports (and fitness) industry	0.12%
114	Floristry	0.11%
115	Floor layers	0.10%
116	Forestry, hunting, landscape preservation	0.08%

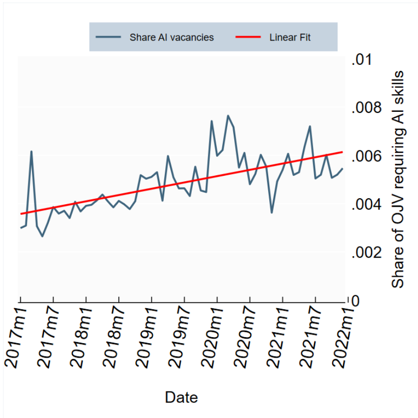
Continued on next page

Table B2 – continued from previous page

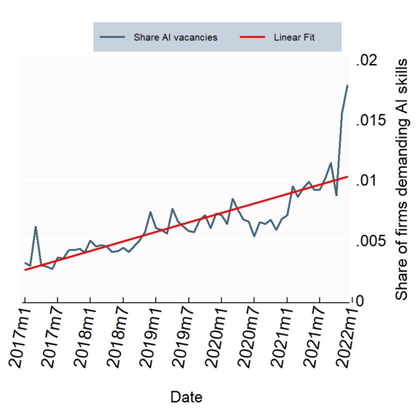
Rank	Occupation (3-digit KLdB 2010)	AI vacancy share
117	Teachers for specific subjects, vocational training	0.07%
118	Drivers of vehicles in railway traffic	0.06%
119	Paper-making and -processing, packaging	0.05%
120	Sales (retail) selling foodstuffs	0.05%
121	Occupational health, safety admin, public health	0.03%
122	Animal care	0.00%

B.2 Alternative AI measures

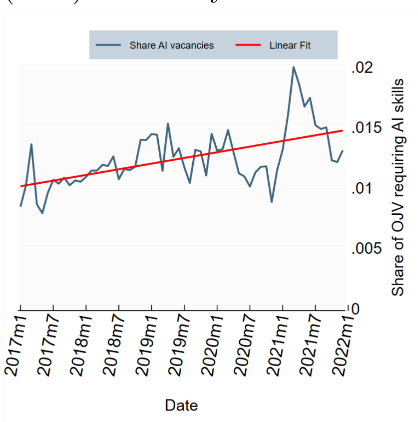
- word clouds, trends over time,



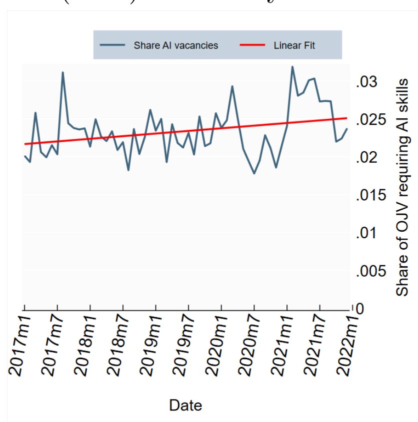
(a) Trends AI Demand: AAHR (2022) Taxonomy



(b) Trends in AI Demand: BGT (2022) Taxonomy



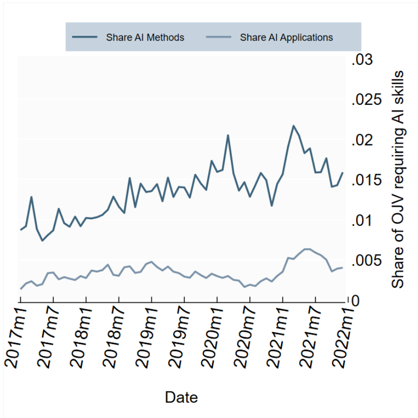
(c) Trends AI Demand: Excluding Top 4 Keys Taxonomy



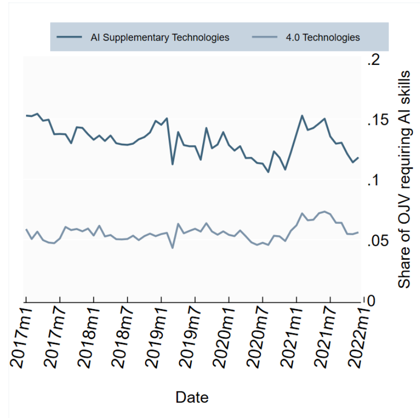
(d) Trends in AI Demand: IW (2021) Taxonomy

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

Figure B1: Trends in AI Demand: Alternative AI Measures



(a) Trends AI Demand: AI Applications vs Methods

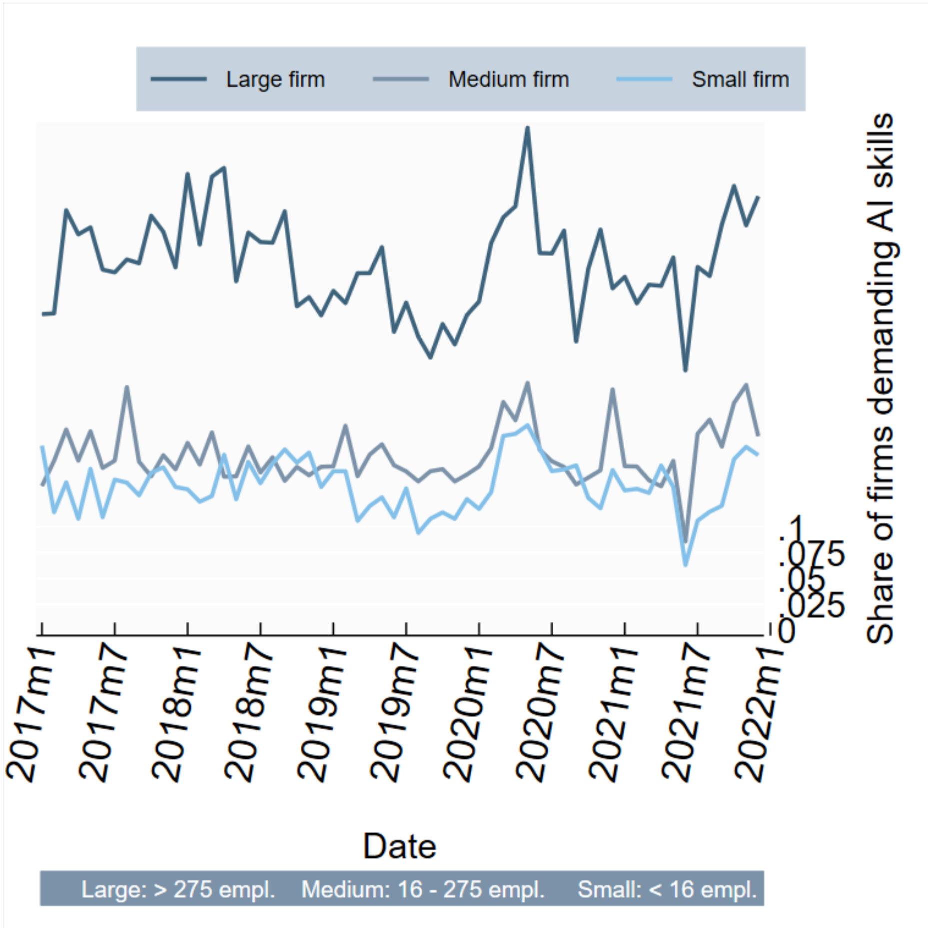


(b) Trends in Demand for 4.0 Technologies

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

Figure B2: Trends in AI Demand: Alternative Technology Measures

B.3 Supporting evidence for IV approach



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 B3: Share of German firms posting AI skills in online job vacancies, 2017 - 2021

C Robustness

C.1 Measurement of AI demand

I: Alternative Definitions of AI Demand

One concern with our AI measure might be that our results are driven by specific industries, e.g. car manufacturing. Therefore, we divide our baseline measure into two subcategories: (i) AI methods and (ii) AI applications. The former category comprises general methods, such as machine learning and deep learning, and thus captures AI skills that are widely applicable across many applications. In contrast, the latter category reflect specific domains in which AI skills are applied to (e.g., autonomous driving); these keywords are relatively more industry-specific.

For our baseline definition of AI skills we enrich keywords from the existing literature in order to create an up-to-date keyword list. Therefore, another potential concern is that our enrichment process inflates the concept the AI skills. If this enrichment is too generous, our AI skill measure will be too broad. To alleviate such concerns, we create alternative keyword lists: (i) using the same keywords as in Acemoglu et al. (2022), (ii) removing very generic keywords, such as ”Artificial Intelligence”. Both of these can be considered more conservative definitions of AI skill demand.

[Figure C1 here]

Figure C1 shows that the positive relationship between AI exposure and wages holds when alternative definitions of AI skill demand are used. Overall, the results from the alternative models are very similar to those from the baseline model.

II: Intensity of AI demand

Our baseline measure is share-based, i.e. defines the exposure to AI via the share of vacancies requiring AI skills. Following this definition, one AI-related skill suffices to declare it an AI vacancy. However, this share-based measure might mask underlying heterogeneity in AI exposure as firms differ in the number of distinct AI skills required in a vacancy (”intensity”). Therefore, the impact of AI exposure on wages might be different depending on, e.g., whether firms demand a variety of specialized AI skills compared to only one specific domain.

We address this concern by constructing an intensity-based AI measure that measures the average number of AI skills required within a posting. Otherwise, our aggregation procedure across local labor markets remains the same. The mean of the AI intensity measure is 0.017, which is about twice as high as our baseline AI share measure. This suggests that when firms demand AI skills, they often demand more than one AI-specific skill. In our sample, the maximum AI intensity is 4.6 AI skills at the LMR occupational year level. Using this alternative definition of AI demand does not qualitatively affect our baseline results either, see Table C2.

III: "4.0 Technologies" (excluding AI)

To empirically validate our AI measure, we construct a similar measure, however, using exposure to 4.0 technologies instead. This allows us to test if we actually capture the impact of AI skills —and not any tangent skills associated with related technologies. We collect a broad keyword list of recent digital technologies that are often used alongside AI technologies. We construct this keyword list, using two kind of sources. First, we extract a comprehensive list of technologies from the European Skills, Competences, Qualifications, and Occupations (ESCO) framework (**?**). ESCO provides, among others, a harmonized classification of ICT technologies. Second, because ESCO is incomplete, we add these types of technologies from state-of-the-art literature. Especially Bloom, Hassan, Kalyani, Lerner & Tahoun (2021) and **?**. In total, we end up with 300 distinct 4.0 technologies. Importantly, this list of 4.0 technologies is orthogonal to our AI keyword list, i.e. we exclude any AI-related technologies. These "4.0 technologies" (Genz, Gregory, Janser, Lehmer & Matthes 2021) encompass technologies that have been introduced to mass markets in the 2010s and comprise, among others, cloud technologies, virtual reality, and embedded systems (see **??** for an overview).

[Figure C3 here]

Subsequently, we add this "4.0 exposure" variable as additional regressor in our baseline model. Figure C3 reports the results of this exercise, showing that inclusion of other 4.0 technologies does affect not our main findings substantially. Explicitly accounting for related technologies, a 10% increase in AI exposure is still associated with a wage increase of 0.6% —in line with our baseline results.

We have perform a similar robustness exercise, using other technologies that are broadly related to AI. These "supplementary technologies" are those that we have often found accompanying our baseline AI skills in job vacancies. Using these supplementary technologies in a similar fashion as our 4.0 technologies likewise does not affect our main results. This evidence bolsters our baseline AI measure, showing it has unique explanatory power on wages, even when controlling for a broad set of related technologies.

C.2 Econometric robustness checks

I: Flexible fixed-effect models

To ensure our baseline results are not driven by misspecified model, we run more flexible specifications that account for LMR- and occupation-specific demand shocks. We modify our baseline wage regression, eq. 3, by interacting LMR FE and, respectively, occupational FE with year FE. This way, we can capture a broader range of idiosyncratic shocks that only affects workers in certain regions or occupations. We report the results of this exercise in Table C4.

First, we account for LMR-occupation-specific trends by including (LMR $\times$ occupation) FE (column 4). Conceptually, this allows us to follow the trend of a (LMR $\times$ occupation) cell over time. In fact, including the interaction has little impact on the baseline result, the estimate even slightly increases up to 0.077. Next, we account for flexible LMR and occupation-specific shocks by adding (Year $\times$ LMR) and (Year $\times$ occupation) FE to our model. The point estimate reduces to 0.057 when we include (Year $\times$ LMR) FE, and further reduces to 0.036 when we include (Year $\times$ Occupation). This underlines the importance of the occupational dimension in capturing important structural changes in the labor market over time. Evaluated at the mean AI diffusion, these last two models with flexible FE imply that a 10% increase in AI diffusion is associated with a wage increase between 0.5% and 0.3%. Overall, we find positive and significant estimates in all specifications, supporting the positive impact of AI diffusion on wages.

II: Specification in differences

Our baseline methodology estimates the impact of AI exposure on wages in level. A potential concern with this approach is that it limits downward movements in wages, especially for continuously employed workers. Hence, our baseline approach may not be able to capture any displacement effects properly. To address this concern, we re-run our model in differences. In Table C5 we report our results, comparing outcomes over (i) 1-year differences, (ii) 2-year differences, and (iii) the maximum of 4-your differences. Overall, this exercise confirms the main takeaway from our baseline results. An increase in the change of AI exposure is associated with a positive change in wages in all specifications. Interestingly, the size of the coefficients increases substantially for longer time-differences. This observation suggests that the cumulative impact of rising AI exposure increases over time. Given that demand for AI skills is still relatively small in our time horizon, wage implications might become more pronounced in future years.

III: Sample restrictions

To address concerns about potential sample bias, we refine our sample selection process. First, we restrict our sample to workers who are present in the SIAB data across all years to ensure consistency. This step aims to mitigate potential biases from positive selection into regions and occupations with higher wage growth and negative selection out of the labor force, both of which could impact our estimates. We then restrict the OJV data to include only LMR occupation cells with at least three postings across all years, focusing on meaningful labor markets. Finally, we apply both restrictions simultaneously.

The results are detailed in Table C6, with the baseline result reported in column (1). First, imposing restrictions on the SIAB data has minimal impact on the estimated effect. Second, imposing restrictions on the OJV data increases the estimated effect to 0.092, indicating a stronger relationship between AI demand and wages in more substantial labor markets. Moreover, imposing both restrictions further increases the estimated effect to 0.102, underscoring a more robust relationship between AI and wages in meaningful local labor markets, particularly for workers with stronger labor market attachment.

C.3 Figures

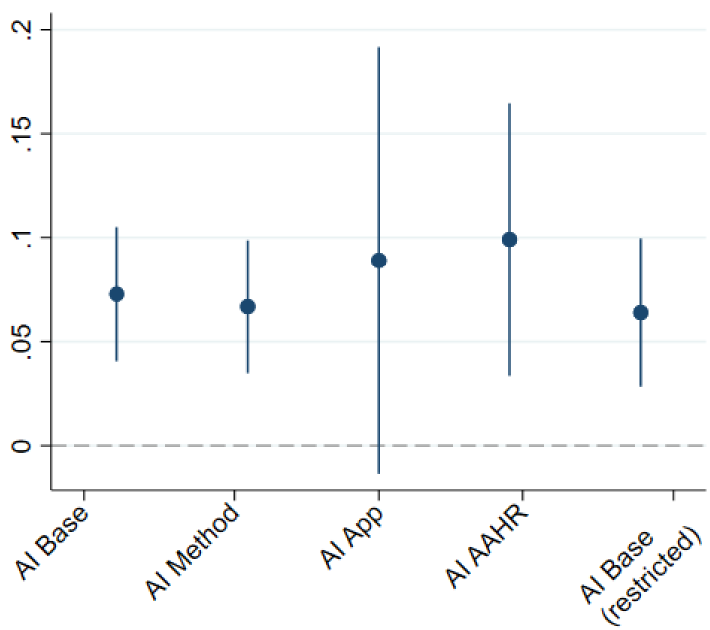


Figure C1: Coefficient Plot of AI Demand using Alternative Measures Pattern

C.4 Tables

Table C1: Wage regressions with alternative AI demand measures

	Dependent Variable: Log Wages				
	(1)	(2)	(3)	(4)	(5)
AI baseline	0.073***				
	(0.016)				
AI method		0.067***			
		(0.016)			
AI app			0.089*		
			(0.052)		
AI AAHR				0.099***	
				(0.033)	
AI Base (restricted)					0.064***
					(0.018)
Worker FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
LMR FE	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
Observations	1,500,688	1,500,688	1,500,688	1,500,688	1,500,682
R-squared	0.89	0.89	0.89	0.89	0.89

Note: All specifications include the following controls: Socioeconomic controls include age, citizenship, education, and gender. Work controls include LM experience. Firm controls include establishment size and industry (WZ08, 2-digit). AKM effects for the time period 2007-2013.

Table C2: Wage regressions AI Intensity: Occupation-LMR-level

	Dependent Variable: Log Wages		
	(1)	(2)	(3)
AI Intensity	0.027***	0.027***	0.023***
	(0.007)	(0.007)	(0.007)
Worker FE	✓	✓	✓
Year FE	✓	✓	✓
LMR FE		✓	✓
Occupation FE			✓
AI Share Mean		0.009	
Observations	1,500,688	1,500,688	1,500,688
R^2	0.89	0.89	0.89

Note: All specifications include the following controls: Socioeconomic controls include age, citizenship, education, and gender. Work controls include LM experience. Firm controls include establishment size and industry (WZ08, 2-digit). AKM effects for the time period 2007-2013.

Table C3: Wage regressions other 4.0 technologies

	Dependent Variable: Log Wages			
	(1)	(2)	(3)	(4)
AI Baseline		0.0676***		0.0724***
		(0.016)		(0.016)
4.0 Technologies	0.0264***	0.0217**		
	(0.009)	(0.009)		
AI Supplementary Technology			0.0044	0.0021
			(0.005)	(0.005)
Worker FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
LMR FE	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓
Observations	1,500,688	1,500,688	1,500,688	1,500,682
R^2	0.89	0.89	0.89	0.89

Note: All specifications include the following controls: Socioeconomic controls include age, citizenship, education, and gender. Work controls include LM experience. Firm controls include establishment size and industry (WZ08, 2-digit). AKM effects for the time period 2007-2013.

Table C4: Wage regressions with flexible interactions

	Dependent Variable: Log Wages					
	(1)	(2)	(3)	(4)	(5)	(6)
AI Share (Occ-LMR)	0.089***	0.087***	0.073***	0.077***	0.057***	0.036***
	(0.018)	(0.018)	(0.016)	(0.018)	(0.014)	(0.012)
Worker FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
LMR FE		✓	✓	✓	✓	✓
Occupation FE			✓	✓	✓	✓
LMR X Occupation FE				✓		
LMR X Year FE					✓	
Occupation X Year FE						✓
AI Share (Occ-LMR) Mean	0.009					
Observations	1,500,688	1,500,688	1,500,688	1,500,688	1,500,682	1,500,149
R-squared	0.89	0.89	0.89	0.90	0.89	0.90

Note: All specifications include the following controls: Socioeconomic controls include age, citizenship, education, and gender. Work controls include LM experience. Firm controls include establishment size and industry (WZ08, 2-digit). AKM effects for the time period 2007-2013.

Table C5: Wage regressions in differences over years

	Dependent Variable: Difference Log Wages		
	(1)	(2)	(3)
Δ AI Demand (1 year)	0.051***		
	(0.017)		
Δ AI Demand (2 years)		0.137***	
		(0.030)	
Δ AI Demand (4 years)			0.310***
			(0.059)
Year FE	✓	✓	✓
LMR FE	✓	✓	✓
Occupation FE	✓	✓	✓
Observations	1,126,404	795,358	230,464
R^2	0.019	0.049	0.103

Note: All specifications include the following controls: Socioeconomic controls include age, citizenship, education, and gender. Work controls include LM experience. Firm controls include establishment size and industry (WZ08, 2-digit). AKM effects for the time period 2007-2013.

Table C6: Wage regressions restricted dataset

	Dependent Variable: Log Wages			
	(1)	(2)	(3)	(4)
AI Demand	0.073***	0.078***	0.092***	0.102***
	(0.016)	(0.019)	(0.023)	(0.026)
Worker FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
LMR FE	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓
AI Share Mean	0.009			
Observations	1,500,688	1,500,688	1,500,688	1,500,688
R^2	0.89	0.90	0.89	0.89

Note: All specifications include the following controls: Socioeconomic controls include age, citizenship, education, and gender. Work controls include LM experience. Firm controls include establishment size and industry (WZ08, 2-digit). AKM effects for the time period 2007-2013. Column (1) Baseline result. Column (2) sample is restricted to workers observed in all years in administrative data. Column (3) sample is restricted to $LMR \times occupation$ cells with at least 3 postings in each year between 2017 and 2021. Column (4) sample is restricted to workers observed in all years and $LMR \times occupation$ with at least 3 postings in each year.