# Augmenting or Automating Labor? The Effect of AI Exposure on New Work, Employment, and Wages*

David Marguerit

LISER

May 2024

**Abstract**

This study examines the impact of artificial intelligence (AI) exposure on the emergence of new work, employment, and wages in the United States from 2010 to 2022, mobilizing novel indices at the occupational and industry level. I differentiate between AI technologies that automate tasks and those that augment industry and occupational outputs. Using instrumental variable estimators, the analysis reveals that augmentation AI exposure promotes the creation of new work and increases employment, but it does not affect wages. Conversely, the findings indicate that automation AI exposure has a detrimental effect on hourly wages without significantly influencing new work emergence or employment levels. Further heterogeneity analysis demonstrates that the effects of both augmentation and automation AI exposures are contingent upon the educational requirements of occupations and the specific types of AI technologies involved.

**Keywords:** Artificial intelligence, Employment, New Work, Tasks, United States, Wages.

**JEL classification:** E24, J11, J21, J23, J24, J31, O33

*I thank Christina Gathmann, Konstantinos Tatsiramos, and Francis Kramarz for constructive and helpful comments. I am also grateful to the participants of the IAB 15th Interdisciplinary Ph.D. Workshop, 4th Junior Research Day, and 16th IAAEU Workshop on Labour Economics for their valuable feedback. This research has been funded by the Luxembourg National Research Fund (AFR-15666136). All errors are my own. Research results and conclusions expressed are those of the author.

# 1 Introduction

Artificial intelligence (AI) has seen substantial advancements in technology and performance over the past decade (Tegmark, 2018). Today, AI algorithms surpass human performance in certain tasks (Kiela et al., 2021; Maslej et al., 2024). This rapid progress has generated significant anxiety among workers about potential job displacement by AI (Gallup, 2023). Unlike previous wave of digital technologies, AI can perform non-routine tasks across most occupations and industries (Autor, 2022; Brynjolfsson et al., 2018). Despite significant advancements, assessing AI's labor market impact remains challenging due to insufficient data on AI adoption and commercial use (Seamans and Raj, 2018). Consequently, policymakers lack the necessary guidance to adapt public policies to AI-induced changes.

Building on the task-based framework (Acemoglu and Autor, 2011; Autor et al., 2003), recent frameworks have explored the effects of technological changes on the labor market (Acemoglu and Restrepo, 2018a, 2018b; Autor et al., 2024). These studies identify two main effects of technological changes. On one hand, Automation reallocates capital and labor towards capital, reducing labor demand and wages. However, it also increases productivity, which can have a positive effect on the labor demand and wages. On the other hand, the reinstatement effect creates new tasks for which human labor has a comparative advantage, thus increasing labor demand and wages. While these frameworks have explained the impacts of technologies like robotization, empirical evidence for AI remains limited, leaving a gap in understanding AI's role in these countervailing effects (Autor, 2022).

This study leverages novel measures of AI exposure to examine its effects on new work emergence, employment, and wages in the United States from 2010 to 2022. Guided by the task-based framework, I investigate whether AI exposure, which may substitute labor in task performance, reduces wages and labor demand. Furthermore, I explore whether AI exposure that augments occupational and industry outputs fosters the emergence of new work and increases labor demand and wages.

The analysis relies on three unique data sources. First, I develop a longitudinal measure of occupational AI exposure for task automation, based on Stack Overflow queries, a leading Q&A forum for developers. I identify AI-related questions and map them to occupational abilities necessary to perform tasks according to the Occupational Information Network (O*NET). The mapping is achieve with Semantic Textual Similarity, a natural language processing tool. I compute an occupational-level automation AI exposure score by tracking the yearly number of questions from 2010 to 2022.

The second data source measures AI exposure that augments occupational and industry outputs. Using a method similar to Autor et al. (2024), I compute an augmentation AI exposure score at the occupational-industry level by assessing the overlap between AI-related queries on Stack Overflow and micro-titles for occupations and industries from the Census Alphabetical Index (CAI). Micro-titles describe goods and services render by occupation and industry.

The third data source tracks the emergence of new work at the occupational level. Building on Lin (2011) and extended by Autor et al. (2024) and Kim (2022), I use the "alternate titles" rubric in O*NET to identify new work within occupations by comparing updates from 2015 to 2022. Alternate titles are micro-titles referring to the notion of jobs and describing someone's position in more detail than occupations. For instance, the micro-titles "Sprinkler Design Engineer" and "Remote Pilot" were added in 2018 and 2021, respectively.

To analyze AI exposure's impact on employment and wages, I match AI exposure measures with data from the Occupational Employment and Wage Statistics (OEWS) provided by the US Bureau of Labor Statistics (BLS). It offers detailed information on average hourly wages and employment size by industry and occupation.

Using these data sources, I provide new descriptive evidence on AI exposure and the emergence of new work. First, analytical non-routine tasks and decision-making skills are positively correlated to AI exposure, whereas routine manual task has a negative relationship. Second, occupations exposed to AI differ from those affected by robots and computers. Third, occupations exposed to automation AI differ augmentation AI are not similar. Fourth, high-paid occupations are more exposed to AI that substitutes human labor, while augmentation AI exposure is mainly concentrated among STEM occupations. Finally, STEM occupations exhibit the highest percentage of new work, with 24% of job titles in "Computer and Mathematical" occupations added between 2015 and 2022.

My main findings, based on IV estimators, are as follows. First, augmentation AI exposure increases the creation of new work and employment size, but has no discernible effect on wages. A one standard deviation increase in augmentation AI exposure raises the share of new work by 0.03 and employment size by 9.5%. Conversely, automation AI exposure negatively impacts hourly wages, with no significant effect on new work or employment size. Specifically, hourly wages decrease by 16.0% with a one standard deviation increase in automation AI exposure.

Second, heterogeneous analysis reveals underlying mechanisms by distinguishing occupations based on educational requirements. Automation AI exposure reduces hourly wages only for low-skilled occupations. Evidence suggests that augmentation AI exposure has a positive effect on the emergence of new work and wages only for high-skilled occupations. These results align with the predictions of skill-biased technical change hypothesis, where the labor supply for high-skilled workers is limited.

Finally, the effect of AI exposure varies according to AI technology types. Exposure to Computer Vision has a lower impact compared to Language Processing and Modeling, and Machine Learning and Deep Learning. This discrepancy can be explained by the varying degrees of technological advancement, with Computer Vision lagging behind in performance.

These results have several important economic implications. First, AI can perform tasks that were previously shielded from technological changes, necessitating a shift in focus from routine tasks to analytical non-routine tasks and decision-making skills when studying AI's effects. Second, as AI performs different tasks compared to robots and computers, occupations exposed to AI differ from those impacted by robotization and computerization. Third, the labor market effects of AI depend on its application: AI that augments output positively influences labor market outcomes, while AI that automates tasks has detrimental effects. Fourth, AI exposure can contribute to rising wage inequality, negatively affecting wages for low-skilled occupations while benefiting high-skilled occupations.

My paper contributes to three strands of economic literature. First, the literature proposing AI exposure indices. Brynjolfsson et al. (2018) develop a forward-looking index measuring the extent to which tasks in occupations could be suitable for machine learning algorithms, a subfield of AI, in the near future. Felten et al. (2018, 2021, 2023a, 2023b) leverage the AI Progress Measurement project by the Electronic Frontier Foundation (EFF) to measure exposure in 10 AI applications, and Engberg et al. (2024) extends this work to create a longitudinal version for 2010-2023. Webb (2020) uses patent data to quantify how much each occupation involves tasks that AI can perform. Tolan et al. (2021) develop an index linking AI research intensity to occupations using expert assessments. Lastly, Eloundou et al. (2023) create an index measuring exposure to large language models, such as ChatGPT. While these measures rely on the current and future capabilities of AI-algorithms, they fail to measure AI adoption.

I contribute to this strand of literature by proposing indices that are more closely aligned with the notion of AI adoption by leveraging developer questions. For this purpose, I suggest a new data source to track AI exposure: queries on Stack Overflow.

Second, this work speaks to the large and growing literature studying the effect of technological changes on the labor market (Acemoglu and Restrepo, 2020; Autor and Dorn, 2013; Autor et al., 2003; Goos and Manning, 2007; Goos et al., 2009; Graetz and Michaels, 2015; Krueger, 1993; Machin and Reenen, 1998; Michaels et al., 2014). Closer to my paper, Acemoglu et al. (2022) find no significant relationship between AI exposure and employment and wage growth at the occupational level. In contrast, Fossen and Sorgner (2022) use micro-data to find a positive relationship between AI exposure and wage growth. Babina et al. (2024) show that firms investing in AI experience increased employment growth, and Bonfiglioli et al. (2023) demonstrate that AI exposure negatively affects employment across U.S. commuting zones. Field experiments by Brynjolfsson et al. (2023), Noy and Zhang (2023) and Peng et al. (2023) find that AI exposure generates productivity gains. However, none of these papers distinguish between AI that automates tasks and AI complementing work. It has been shown that this distinction is crucial to fully understand the effect of technological changes (Acemoglu and Restrepo, 2018a, 2018b; Autor et al., 2024).

My paper contributes to this literature by disentangling the effects of AI that substitutes labor and AI that complements output, a distinction that has not been previously made.

Lin (2011) uses the Census Alphabetical Indexes of Industries and Occupations to show that new work emerges in areas dense with college graduates and industry variety. Atalay et al. (2020) analyze job ads from 1950 to 2000 to explore changes in job titles, reflecting real changes in occupational tasks. Kim (2022) examines the impact of trade on new work emergence, while Acemoglu and Restrepo (2019) link new task creation to labor demand. Autor et al. (2024) combine the Census Alphabetical Indexes with patent data to investigate innovation's impact on new work, finding that task-complementing innovations are associated with new work emergence. None of these papers focus on AI, a technology expected to be an important factor in creating new work.

Unlike these studies, my paper focuses specifically on the emergence of new work driven by artificial intelligence.

The rest of the study is organized as follows. Section 2 describes the data and explains how the different indices are created. In section 3, I validate the measures of AI exposure with previous indices. Section 4 explores which occupations are exposed to AI and where new work emerges. In section 5, I investigate the effect of AI exposure on new work, employment, and wages. I provide heterogeneity analysis and robustness checks in section 6. I conclude in section 7.

## 2 Data and measurement

In this section, I present the data sources and methodology used to construct the measures of AI exposure for automation and augmentation, as well as the measure of new work.

Figure 1 provides an overview of the construction of automation and augmentation AI exposures. To track AI development from 2010 to 2022, I exploit questions about AI asked on Stack Overflow, a Q&A website specializing in coding issues. To measure AI that substitutes tasks in occupations, these questions are matched to the abilities required for task performance as specified in the 2010 O*NET, a comprehensive dictionary describing occupational content in the United States. For augmentation AI (i.e., AI that complements output), I link AI-related questions from Stack Overflow to outputs produced at the micro-occupation and micro-industry levels, according to the Census Alphabetical Indexes of Industries and Occupations from the US Census Bureau. These indices are then merged with information on wages and employment size by occupation and industry from the U.S. Bureau of Labor Statistics.

The measure of new work is derived from updates to the "alternate titles" rubric in

O*NET from 2015 to 2022. After extensively cleaning the dataset to ensure that new work reflects new tasks and specializations rather than simple renaming or rewording, I perform a year-by-year comparison of alternate titles to identify new work.

## 2.1 Automation AI exposure

This section elaborates on the methodology used to construct the measure of automation AI exposure at the occupational level from 2010 to 2022. This measure is derived from two primary data sources: AI-related questions on Stack Overflow and occupational requirements from O*NET. Detailed descriptions of these data sources and the methodology for constructing the index are available in Appendix A.

To track the implementation of AI algorithms in the economy, I analyze questions asked on Stack Overflow between 2010 and 2022. Stack Overflow, established in 2008, is the leading Q&A platform for programming issues, with over 24 million questions asked and 20 million users as of early 2023. Approximately 80% of these users report that coding is part of their job responsibilities (Stack Overflow, 2022).

Each member of Stack Overflow can freely post questions, which must be tagged with 3-5 keywords related to technologies or tasks (e.g., scikit-learn, Python, text-to-speech, regex). These tags facilitate the classification of questions into categories. The community provides answers and comments to suggest solutions. Members also vote on the relevance of questions and answers.

By identifying and analyzing AI-related questions, I capture the types of AI algorithms developers implement for their organizations. The assumption here is that developers typically implement AI algorithms for their employers rather than for personal use.

For occupational content, I rely on the O*NET database, a comprehensive resource describing occupations across the US economy through various descriptors (Peterson et al., 2001). This database has been widely used in the literature to measure the task content of occupations (see, for instance, Acemoglu and Autor, 2011; Blinder, 2009; Brynjolfsson et al., 2018; Felten et al., 2021; Firpo et al., 2011; Peri and Sparber, 2009). This study uses O*NET 15.0, released in July 2010, ensuring that occupational descriptors are not influenced by the AI development measured in this research.

Following Felten et al. (2018, 2021), I measure occupational content using 52 abilities (e.g., oral comprehension, fluency of ideas, finger dexterity). Abilities represent fundamental capacities required to perform a wide range of tasks (Carroll, 1993; Fleishman, 1984). O*NET 15.0 provides detailed descriptions of 855 occupations in terms of these abilities. Each occupation-ability pair is assigned an importance and level score, indicating the significance and required proficiency of the ability for the occupation. These scores are based on surveys conducted with incumbent workers.

Abilities are chosen over other descriptors to measure AI exposure because AI algorithms are often described in broad terms, making them more analogous to general abilities rather than specific tasks or activities. This alignment allows for a more accurate and comprehensive assessment of how AI impacts various occupational requirements.[1]

Using abilities to measure the content of occupations for the index of automation AI exposure involves a conceptual assumption. Specifically, it is assumed that firms mobilize abilities derived from AI algorithms as inputs to perform tasks traditionally carried out by human labor. This assumption is consistent with the O*NET documentation, which states that each occupation requires a specific combination of abilities to effectively perform its tasks (Peterson et al., 2001). By aligning AI capabilities with occupational abilities, this approach provides a robust framework for assessing the impact of AI on different occupations.

The first step in constructing the index of automation AI exposure is to determine the location of Stack Overflow members and retain those living in the United States and its closest trading partners. The identification of member locations is conducted through a four-step process. First, I use the place of living mentioned in the members' profiles.Second, when the place of living is missing, I use ChatGPT to extract any geographical information from the members' personal descriptions.[2] Third, if both the place of living and personal descriptions are unavailable, I refer to the personal website domain. Fourth, the geographical information obtained from the previous steps is then passed to Google Maps to identify the country of living. Finally, I focus on members who live in the United States and its primary trading partners to ensure the accurate tracking of AI algorithm development relevant to the US market.[3] It is well-documented that international trade and the activities of multinational companies play a significant role in technological transfers (Bilir and Morales, 2020; Buera and Oberfield, 2020; Keller and Yeaple, 2013).

Next, I identify AI-related tags on Stack Overflow to pinpoint AI-related questions and subsequently match them to occupational requirements. Out of approximately 60,000 tags on Stack Overflow, the identification of those directly related to AI is performed in three steps. I start by searching for AI keywords in the tags' names and their technical descriptions provided by Stack Overflow. The list of keywords is derived from Alekseeva et al. (2021) and supplemented with additional keywords from computer science and technology literature. This step yields a conservative number of AI-related tags, limited by the comprehensiveness of the keywords list. Then, I identify additional tags that are used in conjunction with those identified in the first step to ensure the inclusion of the latest

[1]There is no guidance in the literature regarding the descriptors of the occupations that should be used to measure AI exposure. While Webb (2020) relies on the description of more than 18 000 tasks performed within occupations, Felten et al. (2018, 2021) use 52 abilities required to perform occupations. In contrast, Brynjolfsson et al. (2018) and Eloundou et al. (2023) take advantage of 2069 detailed work activities, which are merged to tasks performed in occupations.

[2]I use ChatGPT 3.5 Turbo.

[3]The closest trading partners include Bermuda, Canada, China, France, Germany, India, Ireland, Italy, Japan, Mexico, The Netherlands, Singapore, South Korea, Switzerland, Thailand, the United Kingdom, and Vietnam.

AI technologies. This broader search captures a more comprehensive set of AI-related tags, although some may not be directly related to AI (e.g., Python). Finally, I employ ChatGPT to verify which of the identified tags are directly related to AI. This step refines the list to ensure accuracy in identifying AI-specific tags. Through this methodology, I identify 1182 AI-related tags encompassing 181 233 questions.

Then, to integrate the information from Stack Overflow with the abilities required in occupations, I construct a transition matrix that links the descriptions of AI-related tags to the descriptions of abilities.[4] This transition matrix is populated using cosine similarity measures, which quantify the semantic similarity between the description of an AI-related tag and the description of an ability. The cosine similarity score ranges from -1 to 1, where 1 indicates identical meanings, -1 signifies opposite meanings, and 0 denotes orthogonality. Negative scores are replaced with 0, as an AI algorithm is unlikely to be used to perform an ability that is conceptually opposite. The cosine similarity measures are calculated for each pair of tag-ability descriptions using sentence embeddings generated by the Sentence-BERT model. Sentence-BERT creates embeddings in a 768-dimensional vector space (Reimers and Gurevych, 2019).[5]

To compute the index of automation AI exposure, I integrate the information from Stack Overflow and O*NET through a five-step process. In this initial step, I consider the votes attached to AI-related questions in 2022 and smooth these scores over time using a yearly decay factor.[6] A decay factor of 50% is applied, meaning that the impact of a question is halved for each additional year since its publication. This approach acknowledges that new technologies frequently emerge, rendering older questions less relevant. The formula for calculating the smoothed score of an AI-related question for a specific year is:

$$S_{qt} = V_{q2022} * \frac{0.5^{(t-k)}}{\sum_k^{2022} 0.5^{(2022-k)}} \tag{1}$$

Where $S_{qt}$ is the smoothed vote score for question $q$ in year $t$, $V_{q2022}$ represents the votes the question received in 2022, and $k$ is the year of publication. To illustrate the decay factor, a question published in 2020 with a score of 10 in 2022 has a smoothed score of 5.7, 2.9, and 1.4 in 2020, 2021, and 2022, respectively. This step give a table containing a smoothed yearly score for each AI-related question between its year of publication and 2022.

In the second step, I aim to calculate a yearly AI exposure score for each AI-related

[4]The AI-related tags' descriptions are sourced from ChatGPT, given that the descriptions provided by Stack Overflow are often highly technical and do not adequately convey the practical purposes that developers pursue when utilizing these technologies. Therefore, I leveraged ChatGPT to articulate the tasks that developers aim to accomplish with these tags.

[5]I utilize the "all-mpnet-base-v2" model.

[6]Questions with a negative score are excluded as they are deemed uninformative by the Stack Overflow community.

tag. To accomplish this, I divide the smoothed scores from equation 1 by the number of tags attached to each question, thereby avoiding double counting. These adjusted scores are then aggregated at the tag level as follows:

$$ST_{gt} = \sum_{q \in Q_g} \frac{S_{qt}}{n_q} \tag{2}$$

Here, $ST_{gt}$ represents the AI exposure score for the AI-related tag $g$ in year $t$. $Q_g$ denotes the set of questions tagged with $g$ and $n_q$ indicates the number of tags attached to question $q$. This method yields a yearly AI exposure score for each AI-related tag from 2010 to 2022.

Third, I utilize the yearly tag scores from equation (2) in conjunction with the transition matrix to compute the exposure to AI at the ability level:

$$A_{at} = \frac{\sum_{2010}^{t} \sum_{g=1}^{1182} ST_{gt} * C_{ag}}{1182} \tag{3}$$

In this equation, $A_{at}$ measures the exposure to AI for the ability $a$ in year $t$. The numerator represents the cumulative sum of yearly tag scores, weighted by the cosine similarity measures $C_{ag}$ between the descriptions of the tags and the ability $a$. This weighting gives greater significance to yearly tag scores where the tag's description closely aligns with the ability's description. The denominator, consisting of the total number of AI-related tags, normalizes this sum to produce an average exposure score at the ability level for each year.

Fourth, I use the yearly ability AI exposure and the importance and level scores provided by O*NET to compute the automation AI exposure index at the occupational level. The computation is as follows:

$$AI\_auto_{ot} = \sum_{a=1}^{52} \frac{A_{at} * L_{ao} * I_{ao}}{\sum_{a=1}^{52} L_{ao} * I_{ao}} \tag{4}$$

Here, $AI\_auto_{ot}$ represents the automation AI exposure for occupation $o$ in year $t$. This index is the weighted average of the abilities' AI exposure scores from equation (3), where the weights are the importance $I_{oa}$ and level $L_{oa}$ scores, reflecting the varying requirements for abilities across occupations.[7] To account for the differing requirements across occupations, the weights are rescaled between 0 and 1, as commonly done in the literature (Brynjolfsson et al., 2018; Felten et al., 2021; Webb, 2020).[8] Since the importance and level scores are fixed at their 2010 values, all variations in $AI\_auto_{ot}$ are driven by the changes in questions asked on Stack Overflow over time.

[7] Higher importance and level scores indicate that an ability is crucial and frequently used within an occupation, while lower scores suggest that the ability is less relevant.

[8] Some occupations necessitate higher scores for many abilities, resulting in a larger total score. Without rescaling, this would disproportionately increase the AI exposure for these occupations compared to those with lower total scores.

Finally, I convert the index into the 2018 Standard Occupational Classification (SOC) system by using crosswalks from O*NET and computing the simple mean at the 6-digit level.[9] Additionally, I standardize the AI exposure index using the Z-score to facilitate interpretation of the results. A negative Z-score indicates that AI exposure is below the average, while a positive Z-score indicates above-average exposure. Ultimately, this methodology yields automation AI exposure scores for 758 occupations from 2010 to 2022.

Figure 2 presents the exposure to AI automation at the ability level for 2022 (see equation 3). The figure reveals a clear distinction between cognitive abilities and other types of abilities. Cognitive abilities are significantly more exposed to AI than physical, psychomotor, and sensory abilities, corroborating previous findings in the literature (Felten et al., 2021). Among the most exposed abilities are those related to combining information and extracting patterns, such as speed of closure and deductive reasoning. This is expected, given that AI algorithms excel at these tasks. Additionally, abilities related to understanding spoken words and sentences, such as oral comprehension, are also highly exposed.

Conversely, the least exposed abilities include sensory abilities (e.g., auditory attention and far vision), psychomotor abilities (e.g., multilimb coordination), and physical abilities (e.g., stamina and gross body equilibrium). These abilities are less susceptible to AI automation due to their reliance on physical and sensory inputs that AI currently finds challenging to replicate.

Table 1 presents the occupations most and least exposed to AI automation. The most exposed occupations are predominantly white-collar that require advanced educational degrees and high levels of cognitive abilities. Among these, "Purchasing Agents, Except Wholesale, Retail, and Farm Products," "Judges, Magistrate Judges, and Magistrates," and "Clinical and Counseling Psychologists" exhibit the highest scores. Conversely, the 20 least exposed occupations are exclusively blue-collar demanding substantial physical and psychomotor abilities. Occupations such as "Dancers," "Helpers–Painters, Paperhangers, Plasterers, and Stucco Masons," and "Structural Iron and Steel Workers" have the lowest scores. These findings align with those of Felten et al. (2018, 2021), who identified similar patterns in AI exposure across various occupations.

## 2.2 Augmentation AI exposure

The methodology used to create the index of augmentation AI exposure closely mirrors that employed for the index of automation AI exposure. The key difference lies in the use of micro-titles for occupations and industries instead of abilities. Autor et al. (2024) demonstrate that micro-titles effectively capture the services rendered in occupations and industries rather than the specific tasks required to render these services. They also show

[9] ONET classifies occupations using a more granular 8-digit system, whereas the SOC system used in official statistics and surveys employs a 6-digit classification.

that micro-titles can be employed to measure technologies that complement occupations and industries. By maintaining the rest of the methodology consistent, I ensure that any differences observed between automation AI exposure and augmentation AI exposure are attributable solely to the use of micro-titles instead of abilities, rather than methodological discrepancies.

Micro-titles are sourced from the Census Alphabetical Indexes of Industries and Occupations (CAI). The CAI provides a comprehensive list of micro-industry and micro-occupational titles reported by respondents of Census Bureau demographic surveys. Each micro-industry title is assigned a NAICS code, and each micro-occupational title is assigned a SOC code by the Census Bureau. For this study, I use the 2010 CAI, which includes 30 801 micro-occupational titles (e.g., Supervisor waterworks, Candle molder hand, Audit machine operator) and 22 001 micro-industry titles (e.g., Aluminum rolling and drawing, Bus washes, Yacht cleaning). Utilizing micro-titles from 2010 ensures that all variations in augmentation AI exposure are attributable to changes in the questions on Stack Overflow over time.

To construct the index of augmentation AI exposure, I follow a similar procedure to that used for automation AI exposure. First, I identify the locations of Stack Overflow members and retain questions using AI-related tags from members residing in the United States and its most significant trading partners. I then smooth the vote scores. Next, I integrate information from Stack Overflow and the CAI by creating two transition matrices: one linking AI-related tags with micro-industry titles and another linking AI-related tags with micro-occupational titles. Subsequently, I compute the augmentation AI exposure separately for industries and occupations and then average these exposures at the occupation-industry level. This process results in augmentation AI exposure measurements at the 6-digit 2018 SOC and 4-digit 2022 NAICS levels for the period 2010-2022. A more detailed explanation is provided in Appendix A.5.

Table 2 presents the occupations most and least exposed to augmentation AI. Among the most exposed occupations, those related to computers are highly represented, including "Computer and Information Research Scientists," "Computer Systems Analysts," and "Computer Programmers." In contrast, the least exposed occupations are more diverse and include roles requiring substantial physical effort, such as "Rock Splitters, Quarry," "Pourers and Casters, Metal," and "Stonemasons," as well as specialist physicians like "Oral and Maxillofacial Surgeons," "Orthodontists," and "Dental Hygienists."

Many variants were tested throughout the development of the automation AI exposure and augmentation AI exposure indices. However, none of these variants altered the results of this study. The creation of these indices involves several methodological choices that warrant discussion, such as the selection of AI-related tags and the locations of Stack Overflow members. Nevertheless, altering these parameters by adopting more or less restrictive definitions does not impact the observed effects of AI on new work, employment, and wages.

## 2.3 New work

The measure of new work is constructed by comparing the "alternate titles" rubric and its updates in O*NET between 2015 and 2022. The "alternate titles" rubric was introduced in version 20.1 of O*NET in 2015. Alternate titles are micro-titles developed to enhance keyword searches in O*NET (Gregory and Lewis, 2015). These micro-titles are closely related to the concept of jobs and provide a more detailed description of specific positions within occupations.

In O*NET 27.1, the latest version for 2022, there are 52 772 entries (42 889 unique alternate titles) covering 1016 occupations. For example, the occupation "Software Developer" includes 132 alternate titles, such as "Application Developer," "Artificial Intelligence Specialist," and "Computer Software Engineer." These alternate titles are distinct from one another and provide a more granular level of detail compared to the broader occupation categories. On average, there are 52 alternate titles per occupation, with most occupations having between 10 and 100 micro-titles.

O*NET and the alternate titles are regularly updated, with O*NET receiving two updates in 2015 and four updates per year subsequently. Analysis of the alternate titles reveals that there is typically one major update per year, along with several minor updates. Five sources are utilized to identify new alternate titles: incumbents and occupational experts, employer job postings, submitted job titles in the occupational code assignment process, analysis of search term data from customers, and requests from representative groups such as associations and professional organizations. When a new micro-title is identified, occupational analysts undergo a multi-step review process before adding it to the alternate titles. This review ensures that new alternate titles are not already present in the database, are sufficiently familiar to be included, and adhere to style and formatting guidelines.

A major challenge in measuring new work is distinguishing alternate titles that reflect task creation and specialization from those resulting from mere renaming or rewording. To address this, I extensively clean the alternate titles to ensure that the measurement captures only new work attributable to specialization and task creation.

The data cleaning process consists of several steps. First, I convert alternate titles to lowercase, replace acronyms with their full meanings, expand abbreviations, and remove stop words and punctuation. Then, I convert plural words to their singular forms and standardize gendered words to male forms, as male words are more commonly used in O*NET. Next, I retain unique words within alternate titles and order the remaining words alphabetically. Finally, I eliminate duplicate alternate titles within occupations. This cleaning process results in 49 749 unique entries for the period 2015-2022, compared to the initial 52 729 entries, highlighting the importance of thorough data cleaning.[10]

I identify new work by comparing the cleaned versions of alternate titles within occupations across two years. This process involves exact matching of alternate titles between years *t* and *t-1*, supplemented by fuzzy matching based on the meanings of the alternate titles. Fuzzy matching is necessary because the wording of some alternate titles might change over time, even if the task content remains the same.

For instance, in the occupation "Nuclear Technicians" (SOC 19-4051), the alternate titles "Nuclear Technician Worker" and "Nuclear Technician" appear in 2016, while "Nuclear Operating Technician" and "Nuclear Reactor Technician" were already present in 2015. The titles appearing in 2016 are broader and do not reflect changes in task content or specialization. Similarly, in 2019, the title "Medical Transport Driver" was added to the occupation "Ambulance Drivers and Attendants, Except Emergency Medical Technicians" (SOC 53-3011), even though "Transport Medic," "Medical Driver," and "Driver Medic" were already present. Approximate matching ensures that these new titles are considered variations of existing ones, rather than being incorrectly identified as new work.

Fuzzy matching relies on Semantic Textual Similarity. To achieve this, I use the sentence-BERT model to create sentence embeddings for each alternate title (Reimers and Gurevych, 2019). Then, I compute the cosine similarity measures for each pair of alternate titles within occupations. The cosine similarity scores range from -1 (indicating opposite meanings) to 1 (indicating identical meanings). Two alternate titles are considered identical if their similarity measure is 0.7 or higher. This threshold was determined by examining the matching results.[11] A threshold set too high would not sufficiently account for rewording, causing some alternate titles to erroneously appear as new work despite no changes in task content. Conversely, a low threshold would fail to capture genuine instances of new work.

For example, the alternate title "Medical Transport Driver" has a similarity score of 0.72 with "Transport Medic," 0.84 with "Medical Driver," and 0.79 with "Driver Medic." Therefore, "Medical Transport Driver" is not considered new work when it was added in 2019.

Between 2015 and 2022, 2159 instances of new work were added within occupations, averaging 308 new work additions per year. This figure is slightly below the estimates reported by Autor et al. (2024). Using updates from the Census Alphabetical Index of Occupations, Autor et al. (2024) identified 28 315 instances of new work during the period

[10]In 2020, O*NET updated its occupational classification to align with the 2018 Standard Occupational Classification provided by the US Bureau of Labor Statistics. This update introduced new occupations and involved splitting and merging some existing ones. To address these changes, I use the crosswalk provided by O*NET and perform a double matching between alternate titles in the previous classification with those in the updated version, as well as between the O*NET-SOC 2010 codes and the O*NET-SOC 2019 codes.

[11]The following thresholds were tested: 0.95, 0.90, 0.80, 0.70, 0.60, and 0.50. The choice of threshold does not affect the results concerning the effect of AI exposure on the creation of new work.

from 1940 to 2018, corresponding to an average of 363 new work additions per year.

Table 3 provides examples of new alternate titles added to O*NET between 2015 and 2022. Some new work can be directly attributed to technological advancements, such as "Autonomous Vehicle Design Engineer" introduced in 2018 and "Remote Pilot" in 2021. Other new work reflects emerging services unrelated to technological changes, such as "Culinary Artist" added in 2020 and "Cat Groomer" in 2022.

## 2.4 Wages and employment data

Data on wages and employment are sourced from the Occupational Employment and Wage Statistics (OEWS) database (BLS, 2023), maintained by the US Bureau of Labor Statistics (BLS). This database, built on establishment-based data series, provides more accurate estimates for wages and employment at the occupational level compared to household surveys (Acemoglu et al., 2022).

The OEWS provides annual employment and wage estimates for industry-occupation cells. In the May 2022 version, the database includes 831 occupations (at the 6-digit level) and 247 industries (at the 4-digit level). On average, each industry encompasses 147 occupations, and each occupation is represented in 44 distinct industries.

Wage estimates represent straight-time gross pay, excluding premium pay. I rely on the mean hourly wage, adjusted to 2022 dollars.[12] The data includes part-time and full-time employees who are paid a wage or salary. Data from establishments in farm industries and those in the Public Administration sector are excluded from the sample.[13]

I use wage and employment estimates from 2010 to 2022. However, the classifications for occupations and industries have changed over time. The occupation classification was updated in 2018, while the industry classification has been updated three times (2012, 2017, and 2022). To ensure consistency in occupation and industry groups over time, I use crosswalks provided by the BLS and the US Census Bureau.

Finally, I construct a balanced panel dataset by merging the AI exposure indices with information from the BLS. I retain only observations for occupations that are consistently present within industries throughout the entire period. The final dataset comprises 230 204 observations for analyzing the effect of AI exposure on wages and employment from 2010 to 2022, and 141 664 observations for examining the impact of AI on new work from 2015 to 2022. These observations cover 672 distinct occupations across 244 industries. For 2022, this dataset represents 72% of US employees.

Table 4 provides descriptive statistics on wages and employment size. The mean hourly wage is 34.6 with a standard deviation of 18.8. On average, an occupation-industry cell has 5993 individuals.

[12]Hourly wages are computed by dividing total wages by total worked hours.

[13]Excluded industries correspond to the following NAICS codes: 111, 112, 1131, 1132, 114, 1153, 814, and 92.

# 3 Comparison with previous proxies

In this section, I compare the AI exposure measures developed in this study with other indices proposed in the literature. First, I analyze the relationship with indices measuring AI exposure. Next, I examine indices that measure various tasks and skills impacted by previous waves of digital technologies. Finally, I explore the relationships with exposures to factors that have influenced labor demand in recent decades, including offshorability, software, and robots.

Table 5 presents the results of estimating the relationships between the AI exposure indices developed in this study and other AI indices proposed in the literature, using OLS estimators. To facilitate interpretation, the indices are converted into percentile ranks.

Columns 1 to 3 use the automation AI exposure index as the dependent variable. Automation AI exposure appears closely aligned with the index suggested by Felten et al. (2021), with a point estimate of 0.97 (column 1). This result is unsurprising since the methodologies for producing these two indices are similar, both relying on abilities to describe the content of occupations. Additionally, Felten et al. (2021) focus on AI applications that have experienced the fastest growth since 2010 and are believed to be more likely used in the medium term. There is significant overlap between their AI applications and the AI-related tags on Stack Overflow.

The correlation with the index from Brynjolfsson et al. (2018) is statistically significant but relatively weak (point estimate of 0.23) (column 2). This can be explained by the nature of their index, which focuses solely on machine learning, a subfield of AI, whereas this study employs a broader concept of AI.

The relationship between the measure of automation AI exposure and the index from Webb (2020) is not statistically significant (column 3). This discrepancy may be due to Webb (2020) using patents to measure AI exposure. AI systems are often protected as trade secrets, and protecting them under copyright and patent laws presents challenges (Foss-Solbrekk, 2021; Hattenbach and Snyder, 2018; Hu and Jiang, 2019). Furthermore, AI algorithms might be published as open source. Therefore, patents may reflect only a limited aspect of AI adoption due to these varied protection strategies.

Columns 4 to 6 present the relationships when the dependent variable is the augmentation AI exposure measure. The association with Felten et al. (2021) and Brynjolfsson et al. (2018) is much weaker than for automation AI exposure and is not significant for the latter (columns 4 and 5, respectively). These results indicate that augmentation AI exposure captures a different aspect of AI exposure compared to these previous indices. Specifically, Felten et al. (2021) and Brynjolfsson et al. (2018) focus on measuring AI that can perform tasks, rather than AI that complements output at the occupational and industry level.

Regarding Webb (2020), the coefficient is 0.31 and is significantly different from zero. This result suggests that firms might be more likely to patent AI algorithms that complement labor rather than replace it.

Table 6 explores the relationships between the AI exposure measures developed in this study and indices measuring different types of tasks and skills, all converted into percentile ranks. Previous research has shown that routine tasks are more likely to be automated by computers and robots because they are easier to codify (Acemoglu and Autor, 2011; Autor, 2015; Autor et al., 2003). In contrast, the demand for decision-making skills has increased in recent decades and has been relatively insulated from previous waves of automation (Deming, 2021). However, these decision-making skills are likely to be affected by AI algorithms (Agrawal et al., 2019; Choi et al., 2023; Shin et al., 2023).

Panel A presents the results when the response variable is the automation AI exposure.[14] In column 1, the routine task index shows a negative relationship with exposure to automation AI (point estimate of -0.47). This negative association indicates that AI can automate tasks distinct from those affected by previous waves of digital technologies.

In columns 2 to 6, the routine task index is decomposed into subcategories. Routine manual and routine cognitive tasks are negatively associated with AI automation exposure (point estimates of -0.79 and -0.16, respectively) (columns 2 and 3). While Eloundou et al. (2023) find a negative association between their measure of AI exposure and routine manual tasks, they detect a positive association with routine cognitive tasks. Column 4 shows a negative relationship between automation AI exposure and non-routine manual tasks (point estimate -0.89). Interestingly, the association is positive for non-routine interpersonal and non-routine analytical tasks (columns 5 and 6). Instead of being automated, these tasks have been complemented by previous waves of automation (Autor, 2022), and the skills required to perform them have been in demand (Deming, 2017; Deming and Kahn, 2018). Finally, automation AI exposure is also positively associated with decision-making skills (column 7), suggesting that these skills are at risk of automation due to AI.

In Panel B, augmentation AI exposure is used as the response variable. Overall, the associations are smaller than for automation AI exposure and are sometimes not statistically significant, highlighting the differences between automation AI exposure and augmentation AI exposure. Augmentation AI exposure is negatively associated with the routine task index and routine manual tasks (point estimates: -0.13 and -0.08, respectively) (columns 1 and 2). Conversely, it has a positive relationship with non-routine analytical tasks (point estimate: 0.29) (column 6). This result suggests that occupations relying on non-routine analytical tasks might benefit from AI, similar to the effects seen with robotization and computerization (Autor, 2022). Column 7 shows a positive association between the measure of augmentation AI exposure and decision-making skills (point estimate: 0.18).

In Table 7, I analyze the relationship between indices of AI exposure and measures

[14]The indices related to the routine task index (columns 1 to 6) are constructed following Acemoglu and Autor (2011). The measure of decision-making is based on Deming (2021).

of offshorability, software utilization, and robotics. Columns 1 to 3 present the results where automation AI exposure serves as the dependent variable. In Column 1, the results indicate a positive association with being exposed to offshorability (point estimate: 0.61). This finding implies that occupations vulnerable to offshorability are also susceptible to automation AI exposure. Conversely, occupations with high exposure to software are less likely to face automation AI exposure (Column 2). The significant negative coefficient for robot exposure (Column 3, point estimate: -0.73) further supports the notion that occupations impacted by automation AI exposure differ from those affected by earlier waves of automation.

Columns 4 to 6 shift the focus to augmentation AI exposure as the dependent variable. Here, the exposure to software emerges as the sole index with a positive and statistically significant association with augmentation AI exposure (Column 5). This result highlights that software-intensive occupations are more likely to benefit from augmentation AI technologies.

This section offers several critical implications. Firstly, it validates the use of Stack Overflow queries as a method for tracking AI development within the economy. The comparison of these measures with those found in existing literature demonstrates that the constructed indices for automation AI and augmentation AI exposure are consistent and reliable. Secondly, the findings provide evidence that AI impacts distinct tasks and skills compared to robots and traditional computers. Notably, there is a significant relationship between AI exposure and the prominence of analytical non-routine tasks and decision-making skills. Consequently, future research on the impact of AI should concentrate on these specific tasks and skills to gain a more nuanced understanding of AI's economic effects.

# 4 AI exposure and new work

In this section, I first provide descriptive statistics showing which occupations are more exposed to automation AI and augmentation AI. I look at the exposure by broad occupation groups, wages, and educational requirements. In a second subsection, I study where new work emerges. Similarly to AI exposure, I study the emergence of new work by broad occupation groups, wages, and educational requirements.

## 4.1 Which occupations are more exposed to AI?

Figure 3 Panel A shows the exposure to automation AI by broad occupations for 2022. Management, business, legal, and STEM occupations are the most exposed to AI automation in 2022. "Office and Administrative Support" and "Educational instruction and Library" occupations also have a high exposure to AI automation. On the contrary, occupations relying on physical activities appear less exposed (e.g., "Farming, Fishing, and Forestry", "Construction and Extraction", and "Transportation and Material Moving"). This result reflects the differences in ability requirements. Management, Business, Legal, and STEM occupations rely more on cognitive abilities, which are more exposed to automation AI (see Figure 2).

Similarly, Figure 3 Panel B documents the exposure to augmentation AI by broad occupations. In 2022, STEM occupations have the highest exposure to augmentation AI. On the other hand, occupations in "Sales and Related", "Food Preparation and Serving Related", and "Personal and Care Service" are the least exposed.

I dig deeper into AI exposure at the occupational level by looking at the overlap between automation AI and augmentation AI exposure. Figure 4 displays the scatterplot using automation AI exposure for the x-axis and augmentation AI exposure for the y-axis, both in percentile rank. A weak positive relationship emerges from this figure. This result has two important implications. First, it shows that occupations exposed to augmentation AI exposure and automation AI exposure can be different. Some occupations could mitigate the adverse effect of automation due to AI by using AI to complement their output (top right corner of the figure). Other occupations could gain from augmentation AI and not be hurt by automation AI (top left corner). Occupations in the bottom right might be negatively affected by automation AI and not gain from augmentation AI. Occupation on the bottom left should not be directly affected by AI exposure. The second implication is that AI seems different from previous historical innovations. Autor et al. (2024) show a similar figure for the relationship between exposure to automation and augmentation innovations at the occupation level between 1980 and 2018. They find a much stronger positive relationship.

Now, I explore the exposure to AI by the characteristics of the occupations. Figure 5 presents the average exposure to automation AI (Panel A) and augmentation AI (Panel B) by typical entry-level education.[15] I distinguish between occupations having a typical-level education equal to or below an associate's degree and those having higher requirements. This distinction reflects that decision-making skills are often necessary for occupations requiring a Bachelor's degree or higher (Deming, 2021). Those skills are more likely to be affected by AI exposure (Agrawal et al., 2019; Choi et al., 2023; Shin et al., 2023). Automation AI exposure increases with the level of educational requirements. Occupations requiring a lower level of education rely on abilities for which AI is not well suited, such as near-vision, arm-hand steadiness, finger dexterity, and manual dexterity. On the contrary, high-skilled occupations depend more on abilities more exposed to automation AI (e.g., written comprehension, fluency of ideas, or flexibility of closure). These findings echo results from Eloundou et al. (2023) and Webb (2020), who find a positive relationship between AI exposure and educational attainment.

Panel B reproduces a similar exercise but for augmentation AI exposure. On average,

[15]The typical entry-level education comes from the US Bureau of Labor Statistics. It represents the typical education level most workers need to enter an occupation.

occupations with a typical entry-level education equal to an associate's degree or lower are less exposed to augmentation AI.

Finally, I document the relationships between AI exposure and hourly wages (Figure 6). In both panels, the x-axis represents the average hourly wages for 2022, measured in percentile rank. Panel A shows a positive relationship between automation AI exposure and the level of hourly wages. This positive relationship appears much more pronounced for occupations with an average hourly wage above the median. These findings corroborate previous results in the literature (Acemoglu et al., 2022; Webb, 2020). In panel B, the y-axis measures the augmentation AI exposure. The regression line shows a positive association until the median hourly wages become flat.

This subsection highlights two main findings regarding AI exposure. First, occupations exposed to automation AI differ from those exposed to augmentation AI, suggesting that the effect of AI exposure can be heterogeneous. Second, high-paid occupations appear more exposed to automation AI. This finding points out the difference between automation induced by AI and recent waves of automation, such as robots, that affect mainly middle-paid workers (Autor and Dorn, 2013; Autor et al., 2003; Goos and Manning, 2007; Goos et al., 2009; Michaels et al., 2014).

## 4.2 Where does new work emerge?

In this subsection, I start by examining where new work emerges. Then, I investigate the correlation between new work and the measures of AI exposure.

Figure 7 documents the percentage of new work by broad occupation for 2015-2022. It shows that new work emerges mainly in STEM, business, and managerial occupations. This result confirms previous findings showing a shift in the emergence of new work towards high-paid occupations (Autor et al., 2024). The occupation "Computer and Mathematical" has the highest share of new work: 24% of the alternate titles in 2022 appeared during the last 7 years. This result reflects the recent progress in computer science and the creation of new tasks in this field. Within this broad occupation, and not surprisingly, the percentage of new work for the occupations "Data Scientists" and "Computer and Information Research Scientists" reach 41% and 21%, respectively.

After studying in which occupations new work emerges more frequently, I focus on the relationships between new work and exposure to AI (Figure 8). Panel A presents the automation AI exposure (x-axis) against the creation of new work (y-axis), both converted into percentile rank. The association between both measures is negative and then positive. Occupations with a middle exposure to automation AI are associated with less creation of new work. In contrast, Panel B displays a slightly positive relationship between the emergence of new work and the exposure to augmentation AI. The creation of new work appears slightly more important as occupational AI exposure increases.

This subsection provides evidence of where new work emerges. New work emerges more often in high-skilled occupations. This result supports the findings in Autor et al. (2024), which show that the locus of new task creation has shifted toward high-paid occupations in recent years. Interestingly, the emergence of new work appears to be more important in occupations facing higher augmentation AI exposure.

# 5 Labor market effect of AI exposure

In this section, I examine the impact of AI exposure on three key labor market outcomes: the emergence of new work, employment levels, and wages. The analysis begins with a detailed description of the empirical strategy employed to assess these effects. Subsequently, I present the main findings, elucidating the relationship between AI exposure and each of these labor market outcomes.

## 5.1 Empirical strategy

To investigate the effect of AI exposure on the emergence of new work, I employ fixed effect estimators. This methodological choice allows for the full exploitation of the dataset's information by examining yearly variations in AI exposure. The fixed effects regression model I use is specified as follows:

$$\begin{aligned}\frac{\Sigma_{2015}^{t} NewWork_{ot}}{Work_{o2015}} = \beta_1 auto\_AI_{ot} + \beta_2 augm\_AI_{oit} + \beta_3 log(Emp_{oit}) + \\ \alpha_{oi} + \gamma_{it} + \delta_t + \varepsilon_{oit}\end{aligned} \tag{5}$$

In this equation, $o$ indexes occupation, $i$ industry, and $t$ year. The dependent variable is the cumulative share of new work. By considering the cumulative share, I assume that a job is tagged as new from the year it first appears in the O*NET database onward. Given the short window of the dataset, this implies that a job is considered new for a maximum of six years. This duration is shorter than that used by Autor et al. (2024), who define a job as new for a maximum of ten years.

$AI\_auto_{ot}$ represents the measure of automation AI exposure, which varies at the occupation-year level (6-digit). This measure is derived from equation (4). $AI\_augm_{it}$ denotes the score for augmentation AI exposure at the occupation-industry-year level (6-digit for occupations and 4-digit for industries), as computed in equation (**??**). To facilitate the interpretation of the results, both AI indices are standardized to have a mean of zero and a standard deviation of one.

I introduce a set of fixed effects to control for unobserved characteristics. The match between occupations (6-digit) and industries (4-digit) is controlled by $\alpha$, while $\delta$ accounts for temporary shocks. Additionally, I capture temporary shocks at the industry level (3-digit) with $\gamma$. This fixed effect absorbs factors that have influenced labor market outcomes over recent decades, such as robotization and trade (Acemoglu and Restrepo, 2020; Acemoglu et al., 2016; Autor et al., 2013; Graetz and Michaels, 2015; Pierce and Schott, 2016). By incorporating fixed effects at the 3-digit industry-year level (n=84), I achieve a more granular control over the impact of robotization compared to most studies analyzing robots using data from the International Federation of Robotics (IFR). The IFR data typically measures robot adoption across approximately 20 broad industries.

I also control for the logarithm of employment size, and $\varepsilon$ represents the idiosyncratic error term. The estimations are weighted by employment size, and standard errors are clustered at both the occupational and industry levels.

The coefficients of interest are $\beta_1$ and $\beta_2$. It is hypothesized that new work will emerge in occupations where AI augments output, while automation AI exposure is not expected to significantly influence the creation of new work (Autor et al., 2024).

To examine the impact of AI exposure on employment and wages, I estimate the following fixed effects regression:

$$\begin{gathered} log(y)_{oit} = \beta_1 auto\_AI_{ot} + \beta_2 augm\_AI_{oit} + \\ \alpha_{oi} + \gamma_{it} + \delta_t + \varepsilon_{oit} \end{gathered} \tag{6}$$

In this context, $y$ represents either employment size or hourly wages. When estimating the effect of AI exposure on wages, I include an additional control for the logarithm of employment size.

The coefficients of interest are $\beta_1$ and $\beta_2$. The sign of $\beta_1$ is *a priori* unknown, since it depends on which effects is the strongest between the displacement effect and the productivity effect (Acemoglu and Restrepo, 2018a, 2018b). Theoretically, the displacement effect tends to reduce labor demand and wages due to capital's comparative advantage, whereas the productivity effect increases the labor demand and wages.

$\beta_2$ is expected to positively influence wages and employment. Technological changes that complement output have been shown to increase labor demand by creating new tasks and work where labor holds a comparative advantage (Acemoglu and Restrepo, 2019; Autor et al., 2024).

The estimated effects of AI exposures on the outcomes might be biased due to endogeneity issues. The identification of the effect of AI exposures on labor market outcomes relies on the assumption that automation AI and augmentation AI exposure measures are orthogonal to the error term. However, this assumption could be questionable for two main reasons. First, the estimates might be affected by reverse causality: higher wages could attract more AI investments, rather than AI exposure affecting wages. Second, the results could be biased due to omitted variables; factors such as education policies, regulatory changes, or industry shifts could simultaneously influence both AI exposure and the outcome variables.

To address these identification threats, I estimate the effect of AI exposures on labor market outcomes using instrumental variables (IV) regressions. The instrumental variables are constructed by measuring exposure to automation AI and augmentation AI using Stack Overflow questions from members living outside the United States and its closest trading partners. This approach is inspired by Acemoglu and Restrepo (2020) and Acemoglu et al. (2023), who leverage variation from other countries to assess the effect of robot adoption within a specific country.

The exposure to AI in the rest of the world serves as an exogenous instrument relative to wages in the United States, as US labor market conditions do not directly influence it. This mitigates concerns regarding reverse causality and simultaneity. Furthermore, this instrument captures broad, global patterns and unexpected changes in AI adoption that impact labor markets across multiple countries simultaneously. These global trends and shocks in AI adoption provide a source of variation not influenced by factors specific to the United States.

By excluding the most significant trading partners of the United States, I ensure that the instrumental variables do not affect US outcomes through trade channels. Moreover, AI exposure in the rest of the world is correlated with AI exposure in the United States due to global technological diffusion and interconnectedness in AI advancements. This correlation satisfies the relevance condition, ensuring that the instrument has predictive power over the endogenous regressors (see Table 8).

## 5.2 Results

In this subsection, I present the main results concerning the effect of automation AI and augmentation AI exposure on new work, employment, and wages.

### 5.2.1 New work

I report in Table 9 the estimates in equation (5) to test whether automation AI and augmentation AI exposures affect the emergence of new. Panel A shows the output when I use fixed effect estimators. In column 1, automation AI exposure is the only variable of interest included in the estimation, along with controls for the employment size (log) and fixed effects for years and occupation-industry matching. Automation AI exposure does not significantly affect the emergence of new work. In column 2, I add further fixed effects to control for temporary shocks at the industry level. The coefficient for automation AI exposure decreases and remains insignificant. I test for the effect of augmentation AI exposure in column 3. I find a positive association between the exposure of augmentation AI and the share of new work with a point estimate of 0.028 (SE=0.008). This positive effect is robust and even reinforced by adding further fixed effects (column 4). In column 5, I add the exposures for automation AI and augmentation AI. The point estimates for automation AI and augmentation AI exposures reduce compared to when introduced separately. However, the effect of augmentation AI exposure remains significant (point estimate = 0.027; SE = 0.008). In column 6, additional fixed effects are introduced in the estimation. Adding these controls reduces the coefficient of automation AI but increases the effect of augmentation AI exposure. According to this estimation, an increase of one standard deviation in augmentation AI exposure raises the cumulative share of new work by 0.03.

In Panel B, I re-do the same exercise when using IV estimators to mitigate the identification threats. Using IV estimators does not change the coefficient and the standard error compared to the fixed effect estimations. Automation AI exposure does seem to affect significantly the emergence of new work (columns 1 and 2). In contrast, augmentation AI exposure shows a positive and significant effect on the emergence of new work with a point estimate of around 0.03 (SE=0.008) (columns 3 and 4). In columns 5 and 6, the two measures of AI exposures are included together. Only the augmentation exposure predicts the emergence of new work. The Montiel-Pflueger F-test confirms that the IV estimations are not concerned by the threat of weak instruments (Olea and Pflueger, 2013). First-stage estimations are available in Appendix B.

Results reported in Table 9 echo the findings in Autor et al. (2024), though they focus on breakthrough innovations and not exclusively on AI. They find that new occupational tasks emerge in response to augmenting innovations. They do not find evidence that automation innovation affects the creation of new work.

### 5.2.2 Employment

In this subsection, I test whether employment size responds to the exposure to automation AI and augmentation AI.

Table 10 presents the estimates of equation (6) when employment size (in log) is used. Fixed effect estimators are used in Panel A. In column 1, only automation AI exposure is included in the regression with time and industry-occupation fixed effects. The point estimate is statistically insignificant, suggesting an absence of effect. In column 2, I add further fixed effects at the industry-year level to control for possible confounders. The point estimate increases to reach 0.184 (SE = 0.116) but remains insignificant. I test the effect of augmentation AI exposure in column 3. Augmentation AI exposure positively and significantly affects employment size (point estimate = 0.089; SE = 0.023). Including further controls in column 4 slightly increases the point estimate and the standard error, though the coefficient remains statistically significant at 1%. In Column 5, both measures of AI exposure are included in the estimation. While the point estimate for automation AI exposure is divided by more than two, the effect of augmentation AI exposure is only slightly reduced. In column 6, the complete set of fixed effects is added. The point estimate for automation AI exposure increases but remains statistically insignificant. In contrast, the effect of augmentation AI exposure slightly increases, though it is still significant at 1%. An increase of one standard deviation in the exposure to augmentation AI increases the employment size by 9.5%.

Panel B displays the results for the IV estimators. Using IV estimators instead of fixed effects estimators does not significantly affect the point estimates and standard errors. When automation AI exposure is the only variable of interest included in the regression (columns 1 and 2), its point estimate is statistically insignificant. In contrast, augmentation AI exposure positively and significantly affects employment size, regardless of the variety of fixed effects introduced (columns 3 and 4). The positive effect of augmentation AI exposure remains positive and statistically significant when included in the estimation with the exposure of automation AI (columns 5 and 6). The estimations pass the Montiel-Pflueger F-test (Olea and Pflueger, 2013), and First-stage estimations are available in Appendix B.

I provide different results than Acemoglu et al. (2022), who do not find discernible relationships between AI exposure and employment. In contrast, Bonfiglioli et al. (2023) estimate robust adverse effects of AI exposure on employment across commuting zones and time, and Babina et al. (2024) shows that AI-investing firms experience higher growth employment. Using data for three European countries, Engberg et al. (2024) find mixed results: an adverse effect of AI exposure on employment in Portugal, a positive one in Sweden, and no significant effect in Denmark. However, none of these studies distinguish between AI that augments output and AI used to automate tasks.

Not specifically focusing on AI, but still related to this study, Autor et al. (2024) and Kogan et al. (2023) find a positive and statistically significant effect of labor-augmenting technologies and innovations on employment. However, in contrast with this study, Autor et al. (2024) show a negative effect of automation innovation on employment.

### 5.2.3 Wages

Now, I turn on the effect of AI exposure on wages. To do so, I estimate the equation (6) using the hourly wages as the response variable.

Table 11 Panel A shows the estimates when fixed effect estimators are used. Automation AI exposure shows a negative and statistically significant effect on hourly wages when only time and industry-occupation fixed effects are included (point estimate = -0.193; SE = 0.042) (column 1). Including further industry-year fixed effects, reduce the point estimate and the standard deviation to -0.173 and 0.028, respectively (column 2). In column 3, automation AI exposure is replaced by augmentation AI exposure, and I control only for time and industry-occupation fixed effects. Augmentation AI exposure has a statistically significant negative point estimate. However, When adding the complete set of fixed effects, the point estimate of augmentation AI exposure turns statistically insignificant (column 4). In column 5, both measures of AI exposure are included simultaneously.

The effect of automation AI exposure appears negative and statistically different from zero (point estimate = -0.183; SE = 0.038). Augmentation AI exposure has a negative estimate (-0.012) and is statistically insignificant. In column 6, the whole model is estimated. The point estimate and standard error for automation AI slightly decreases, though it remains statistically significant: an increase of one standard deviation in automation AI exposure decreases by 16.0% the hourly wages. In contrast, while the point estimate of augmentation AI exposure increases and its standard error decreases, it remains statistically insignificant.

Panel B presents the results when using IV estimators. IV estimators do not affect the point estimates and standard errors. In columns 1 and 2, I include only automation AI exposure. The point estimate is negative and statistically significant, though both reduce when the complete set of fixed effects is included (column 2). I explore the effect of augmentation AI exposure in column 3. I find a negative effect of augmentation AI exposure on hourly wages (point estimate = -0.021; SE = 0.009). However, once all the fixed effects are included (column 4), the coefficient becomes statistically insignificant. Both measures of AI exposure are included in the estimation in columns 5 and 6. Automation AI exposure negatively affects hourly wages, while augmentation AI exposure is statistically insignificant but has a positive point estimate.

These results echo the findings in Hui et al. (2023), who find that freelancers offering tasks in an online platform and competing with ChatGPT experienced a decrease in their earnings after the release of the large language model. Conversely, Acemoglu et al. (2022) and Albanesi et al. (2023) do not find a statistically significant effect of AI exposure on wages. Fossen and Sorgner (2022) shows evidence that an increase in AI exposure is associated with wage growth. However, these studies use a broad measure of AI and do not disentangle between automation AI and augmentation AI. Studying the effect of important innovations, Autor et al. (2024) do not find statistically significant effects of exposure augmentation innovations and automation innovations on wages.

In summary, this section confirms some predictions of the task-based framework applied to AI (Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2018b; Autor et al., 2003, 2024). First, when AI complements output generated by occupations and industries, it creates new work and tasks. Second, by creating new work and tasks for which labor has a comparative advantage, augmentation AI exposure increases the demand for labor. However, the emergence of new work and increased labor demand do not result in a significant wage increase. This absence of effect on wages might come from the type of occupations that are complemented and the availability of labor supply. If augmenting AI creates new work and tasks in occupations for which there is a labor surplus, new work can be performed without creating a shortage in the labor market and leaving wages unaffected. This question is explored in the next section.

Regarding automation AI exposure, it does not affect employment size but reduces hourly wages. An explanation could be that AI is adopted to perform some tasks, reducing the return on labor. However, the performance of AI is not good enough to replace fully human labor in some tasks, so it does not affect employment size.

# 6 Extensions and robustness checks

This section starts by exploring heterogeneity analysis in the effect of AI exposure on new work, employment, and wages. Specifically, I look at whether AI exposure affects occupations differently according to their educational requirements, helping to understand the underlying mechanisms in play. Then, I study the effect of AI exposure by distinguishing the types of AI technologies. Finally, I run several robustness checks.

## 6.1 Effects of AI by educational requirements

Table 12 shows the effect of augmentation AI and automation AI exposures on the three labor market outcomes, breaking down by typical entry-level education in occupations. Results in Panel A rely on fixed effects estimators. Columns 1 and 2 focus on the share of new work. It shows that automation AI and augmentation AI exposures do not significantly affect the share of new work for occupations requiring an Associate's degree and a lower level.[16]. Conversely, augmentation AI exposure appears to positively affect the share of new work for occupations having a typical entry-level education equal to a Bachelor's degree or higher (point estimate = 0.048; SE = 0.009) (column 2).

In columns 3 and 4, I explore the effect of AI exposure on employment size (in log). When occupations with an educational requirement equal to an Associate's degree and lower are retained (column 3), I find a positive effect of augmentation AI exposure with a point estimate of 0.091 and a standard error of 0.045. In contrast, neither automation AI exposure nor augmentation AI exposure appears to affect the employment size for occupations with a Bachelor's degree and higher (column 4).

Finally, I look at the effect of AI exposure on hourly wages in columns 5 and 6. I find that automation AI exposure has a negative and statistically significant effect on hourly wages for occupations requiring an Associate's degree and lower (column 5): one standard deviation in the exposure of automation AI decreases hourly wages by 13.6%. Regarding augmentation AI exposure, it affects significantly high-skilled occupations with a point estimate of 0.019 and a standard error of 0.005 (column 6).

In Panel B, I re-do the same exercise but using IV estimators. The coefficients and the standard errors are not significantly affected, and the results remain similar.

These results highlight four economic implications. First, augmentation AI exposure creates new work for high-skilled occupations, leading to increased wages but not a rise

[16]Automation AI exposure is statistically significant at 10% only, so I do not consider it.

in employment size. The inelastic labor supply for these occupations explains the discrepancy between wages and employment. Due to the scarcity and inelasticity of workers with the necessary skills, firms compete by raising wages to attract talent. Second, evidence suggests that augmentation AI exposure increases employment size for low-skilled occupations without creating new work or affecting wages. This may reflect that AI, when used to complement output, generates productivity gains that reduce the prices of goods and services. The price reduction boosts demand for these goods and services, increasing labor demand. With an elastic labor supply, this increased demand does not impact wages. Third, automation AI exposure negatively affects wages for low-skilled occupations without impacting employment. This can be attributed to the substitution of labor with capital, which decreases the marginal product of labor, leading to lower wages. Finally, AI can contribute to rising wage inequality by adversely affecting wages for low-skilled occupations while increasing wages for high-skilled occupations.

## 6.2 Effects of AI by type of AI technologies

Table 13 presents the estimates when I distinguish by the type of AI technologies. I characterize the type of AI technologies by grouping the AI-related tags. I rely on their descriptions and K-Means clustering to group AI-related tags together. I distinguish between three types of AI technologies: Computer Vision, Language Processing and Modeling, and Machine Learning and Deep Learning.

In Panel A, columns 1 to 3, I present the results when the share of new work is the response variable and fixed-effects estimators are used. All three types of AI technologies significantly and positively affect the emergence of new work when the technology is used to augment the output produced. Conversely, I do not find any effect of automation AI exposure.

In columns 4 to 6, I look at the effect of AI exposure on employment size (in log). Evidence suggests that Computer Vision technologies do not affect employment size (column 4). Technologies related to Language Processing and Modeling that are used to augment output positively impact employment size (point estimate = 0.073; SE = 0.033) (column 5). It is also the case of Machine Learning and Deep Learning technologies, with a point estimate of 0.101 and a standard error of 0.029 (column 6).[17]

The effect of AI technologies on hourly wages (in log) is explored in columns 7 to 9. Computer Vision shows no discernible effect whether it is used to automate tasks or augment output (column 7). In contrast, column 8 shows that automation Language Processing and Modeling has a negative and statistically significant effect on the hourly wages (point estimate = -0.072; SE = 0.015). However, augmentation Language Processing and Modeling does not impact hourly wages. Similar results are found for technologies related

[17] Automation Machine and Deep Learning exposure is statistically significant at 10% only, so I do not consider it.

to Machine Learning and Deep Learning. Those technologies have a negative relationship with hourly wages when used to automate tasks (point estimate = -0.173; SE = 0.025) (column 9). No discernible effects are found for augmentation Machine Learning and Deep Learning exposure.

In Panel B, I rely on IV estimators instead of fixed-effects estimators. These methodological changes do not significantly affect the results.

The level of technological advancement may explain these results. Although Computer Vision has seen rapid improvement over the past decade, its performance remains lower compared to other AI technologies (Maslej et al., 2024). Machine Learning and Deep Learning are fundamental to most AI algorithms, making them crucial for the development of other technologies. This study concludes in 2022, just as ChatGPT was released, which significantly enhanced language processing and modeling technologies.

### 6.3 Robustness checks

I perform several robustness checks to ensure that the effects of AI exposures on new work, employment, and wages are not spurious.

First, I consider placebo treatments where I randomize the relatedness scores in the transition matrix between abilities and AI-related tags for the construction of automation AI exposure.[18] Similarly, I randomize the relatedness scores between AI-related tags and occupational and industry micro-titles for augmentation AI exposure.[19] The results of the placebo treatment are presented in Appendix C. I do not find any statistically significant effects of AI exposure on new work and employment size. I find a positive and statistically significant effect of automation AI exposure on hourly wages, but the coefficient is too large to be realistic.

Second, I use the complete database and keep observations even if they are not present during the whole period and are considered outliers (Appendix D). Using the complete database slightly increases the positive effect of augmentation AI on the emergence of new work. The effect of augmentation AI on employment size is not affected. The negative effect of automation AI on wages is slightly reinforced when I use the complete database.

Third, I use constant weights over time by taking the employment size of the first year of analysis (i.e., 2015 for the share of new work and 2010 for the employment size and hourly wages). Results are presented in Appendix E. The point estimate concerning the positive effect of augmentation AI exposure on the emergence of new work slightly decreases. Regarding employment size, the point estimates are reduced. The positive effect of augmentation AI exposure on employment size becomes statistically significant at only 10%. Automation AI exposure negatively affects hourly wages but with a smaller

[18] I randomly shuffle the values for $C_{ag}$ in equation (3).

[19] The values in the transition matrices are randomly shuffled affecting $C_{ig}$ and $C_{og}$in equations (7) and (8), respectively.

magnitude than baseline estimates.

Finally, I verify that the choice of the wage measure does not affect the results. In Appendix F, I use the median hourly wages instead of the average. Automation AI exposure negatively affects hourly wages more than the baseline estimates.

# 7 Conclusions

In conclusion, this study provides robust evidence on the nuanced effects of AI exposure on the labor market, emphasizing the importance of distinguishing between automation and augmentation AI. Utilizing novel measures derived from Stack Overflow data, the analysis reveals that augmentation AI fosters the emergence of new work and increases employment size, particularly benefiting high-skilled occupations. Conversely, automation AI exposure is associated with a decline in wages, disproportionately affecting low-skilled occupations. These findings align with the skill-biased technological change hypothesis, underscoring AI's potential to exacerbate wage inequality.

The implications of these results are manifold. Policymakers must consider tailored strategies to mitigate the adverse effects of automation AI, particularly for low-skilled workers, while promoting the beneficial aspects of augmentation AI that can spur job creation and economic growth. Additionally, this research highlights the need for continuous monitoring of AI advancements and their diverse impacts across occupations, ensuring that labor market policies remain adaptive and responsive to technological progress.

This paper contributes significantly to the existing literature by introducing innovative methods to measure AI exposure and by disentangling the distinct effects of different types of AI on employment and wages. Future research should build on these findings to explore the long-term impacts of AI on various economic sectors and to develop comprehensive policy frameworks that address the dynamic nature of technological change. The ongoing evolution of AI technologies necessitates a proactive approach to understanding and managing their implications for the workforce, ensuring that the benefits of AI are broadly shared and that potential disruptions are effectively mitigated.

# References


Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics* (pp. 1043–1171, Vol. 4). Elsevier.

Acemoglu, D., Autor, D., Dorn, D., Hanson, G. H., & Price, B. (2016). Import competition and the great us employment sag of the 2000s. *Journal of Labor Economics*, *34*(S1), S141–S198.

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

Acemoglu, D., Koster, H. R. A., & Ozgen, C. (2023). *Robots and workers: Evidence from the netherlands* (Working Paper No. 31009). National Bureau of Economic Research.

Acemoglu, D., & Restrepo, P. (2018a). *Artificial intelligence, automation and work* (Working Paper No. 24196). National Bureau of Economic Research.

Acemoglu, D., & Restrepo, P. (2018b). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, *108*(6), 1488–1542.

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

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

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

Albanesi, S., Dias da Silva, A., Jimeno, J. F., Lamo, A., & Wabitsch, A. (2023). *New technologies and jobs in europe* (Working Paper No. 31357). National Bureau of Economic Research.

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

Atalay, E., Phongthiengtham, P., Sotelo, S., & Tannenbaum, D. (2020). The evolution of work in the united states. *American Economic Journal: Applied Economics*, *12*(2), 1–34.

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

Autor, D. (2022). *The labor market impacts of technological change: From unbridled enthusiasm to qualified optimism to vast uncertainty* (Working Paper No. 30074). National Bureau of Economic Research.

Autor, D., Chin, C., Salomons, A., & Seegmiller, B. (2024). New frontiers: The origins and content of new work, 1940–2018*. *The Quarterly Journal of Economics*, qjae008.

Autor, D., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, *103*(5), 1553–97.

Autor, D., Dorn, D., & Hanson, G. H. (2013). The china syndrome: Local labor market effects of import competition in the united states. *American economic review*, *103*(6), 2121–2168.

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

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

Barua, A., Thomas, S. W., & Hassan, A. E. (2014). What are developers talking about? An analysis of topics and trends in Stack Overflow. *Empirical Software Engineering*, *19*(3), 619–654.

Bilir, L. K., & Morales, E. (2020). Innovation in the global firm. *Journal of Political Economy*, *128*(4), 1566–1625.

Blinder, A. (2009). How many us jobs might be offshorable? *World Economics*, *10*(2), 41–78.

BLS. (2023). *Technical notes for may 2022 oews estimates* (tech. rep.).

Bonfiglioli, A., Crinò, R., Gancia, G., & Papadakis, I. (2023). Artificial Intelligence and Jobs: Evidence from US Commuting Zones. *CEPR*, (DP18495).

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). *Generative ai at work* (Working Paper No. 31161). National Bureau of Economic Research.

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.

Buera, F. J., & Oberfield, E. (2020). The global diffusion of ideas. *Econometrica*, *88*(1), 83–114.

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies* (1st ed.). Cambridge University Press.

Choi, S., Kang, H., Kim, N., & Kim, J. (2023). How does artificial intelligence improve human decision-making? evidence from the ai-powered go program.

del Rio-Chanona, M., Laurentsyeva, N., & Wachs, J. (2023). Are large language models a threat to digital public goods? evidence from activity on stack overflow.

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market*. *The Quarterly Journal of Economics*, *132*(4), 1593–1640.

Deming, D. J. (2021). *The growing importance of decision-making on the job* (Working Paper No. 28733). National Bureau of Economic Research.

Deming, D. J., & 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.

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.

Engberg, E., Görg, H., Lodefalk, M., Javed, F., Längkvist, M., Monteiro, N. P., Kyvik Nordås, H., Schroeder, S., & Tang, A. (2024). *Ai unboxed and jobs: A novel measure and firm-level evidence from three countries* (IZA Discussion Papers No. 16717). Institute of Labor Economics (IZA).

Felten, E., Raj, M., & Seamans, R. (2018). A method to link advances in artificial intelligence to occupational abilities. *AEA Papers and Proceedings*, *108*, 54–57.

Felten, E., Raj, M., & Seamans, R. (2019). The occupational impact of artificial intelligence: Labor, skills, and polarization. *Ewing Marion Kauffman Foundation Research Paper Series*.

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.

Felten, E., Raj, M., & Seamans, R. (2023a). How will Language Modelers like ChatGPT Affect Occupations and Industries? [Publisher: arXiv Version Number: 1].

Felten, E., Raj, M., & Seamans, R. (2023b). Occupational Heterogeneity in Exposure to Generative AI. *SSRN Electronic Journal*.

Firpo, S., Fortin, N. M., & Lemieux, T. (2011). *Occupational Tasks and Changes in the Wage Structure* (IZA Discussion Papers No. 5542). Institute of Labor Economics (IZA).

Fleishman, E. A. (1984). Systems for linking job tasks to personnel requirements. *Public Personnel Management*, *13*(4), 395–408.

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.

Foss-Solbrekk, K. (2021). Three routes to protecting AI systems and their algorithms under IP law: The good, the bad and the ugly. *Journal of Intellectual Property Law & Practice*, *16*(3), 247–258.

Gallea, Q. (2023). From mundane to meaningful: Ai's influence on work dynamics – evidence from chatgpt and stack overflow.

Gallup. (2023). *Bentley-gallup: Business in society report* (tech. rep.).

Goos, M., & Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in britain. *The Review of Economics and Statistics*, *89*(1), 118–133.

Goos, M., Manning, A., & Salomons, A. (2009). Job polarization in europe. *The American Economic Review*, *99*(2), 58–63.

Graetz, G., & Michaels, G. (2015). *Robots at work* (CEP Discussion Papers). Centre for Economic Performance, LSE.

Gregory, C., & Lewis, P. (2015). *O*net alternate titles procedures* (tech. rep. No. 24235). O*NET.

Handel, M. J. (2016). The o*net content model: Strengths and limitations. *Journal for Labour Market Research*, *49*(2), 157–176.

Hattenbach, B., & Snyder, G. (2018). Rethinking the mental steps doctrine and other barriers to patentability of artificial intelligence. *Science and Technology Law Review*, *19*(2).

Hu, S., & Jiang, T. (2019). Artificial intelligence technology challenges patent laws. *2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, 241–244.

Hui, X., Reshef, O., & Zhou, L. (2023). The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market. *Available at SSRN 4527336*.

Keller, W., & Yeaple, S. R. (2013). The gravity of knowledge. *American Economic Review*, *103*(4), 1414–1444.

Kiela, M., Douwe and Bartolo, Nie, Y., Kaushik, D., Geiger, A., Wu, Z., Vidgen, B., Prasad, G., Singh, A., Ringshia, P., Ma, Z., Thrush, T., Riedel, S., Waseem, Z., Stenetorp, P., Jia, R., Bansal, M., Potts, C., & Williams, A. (2021). Dynabench: Rethinking benchmarking in NLP. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4110–4124.

Kim, G. (2022). Trade-induced adoption of new work. *IZA Discussion Paper*, *15165*, 73.

Kogan, L., Papanikolaou, D., Schmidt, L. D., & Seegmiller, B. (2023). *Technology and labor displacement: Evidence from linking patents with worker-level data* (Working Paper No. 31846). National Bureau of Economic Research.

Krueger, A. B. (1993). How computers have changed the wage structure: Evidence from microdata, 1984-1989. *The Quarterly Journal of Economics*, *108*(1), 33–60.

Lin, J. (2011). Technological adaptation, cities, and new work. *The Review of Economics and Statistics*, *93*(2), 554–574.

Machin, S., & Reenen, J. V. (1998). Technology and changes in skill structure: Evidence from seven oecd countries. *The Quarterly Journal of Economics*, *113*(4), 1215–1244.

Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., & Clark, J. (2024). *Artificial intelligence index report 2024* (tech. rep.). Institute for Human-Centered AI. Stanford University, Stanford, CA.

Michaels, G., Natraj, A., & van Reenen, J. (2014). Has ict polarized skill demand? evidence from eleven countries over twenty-five years. *The Review of Economics and Statistics*, *96*(1), 60–77.

Montandon, J. E., Politowski, C., Silva, L. L., Valente, M. T., Petrillo, F., & Guéhéneuc, Y.-G. (2021). What skills do IT companies look for in new developers? A study with Stack Overflow jobs. *Information and Software Technology*, *129*, 106429.

Moutidis, I., & Williams, H. T. P. (2021). Community evolution on Stack Overflow (J. Soldani, Ed.). *PLOS ONE*, *16*(6), e0253010.

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

OECD.AI. (2023). Ai knowledge flow (stack overflow).

Olea, J. L. M., & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, *31*(3), 358–369.

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The impact of ai on developer productivity: Evidence from github copilot.

Peri, G., & Sparber, C. (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, *1*(3), 135–69.

Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P. R., Fleishman, E. A., Levin, K. Y., Campion, M. A., Mayfield, M. S., Morgeson, F. P., Pearlman, K., Gowing, M. K., Lancaster, A. R., Siler, M. B., & Dye, D. M. (2001). Understanding work using the occupational information network (o*net): Implications for practice and research. *Personnel Psychology*, *54*(2), 451–492.

Pierce, J. R., & Schott, P. K. (2016). The surprisingly swift decline of us manufacturing employment. *American Economic Review*, *106*(7), 1632–62.

Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *CoRR*, *abs/1908.10084*.

Rosen, C., & Shihab, E. (2016). What are mobile developers asking about? A large scale study using stack overflow. *Empirical Software Engineering*, *21*(3), 1192–1223.

Seamans, R., & Raj, M. (2018). *Ai, labor, productivity and the need for firm-level data* (Working Paper No. 24239). National Bureau of Economic Research.

Shin, M., Kim, J., van Opheusden, B., & Griffiths, T. L. (2023). Superhuman artificial intelligence can improve human decision-making by increasing novelty. *Proceedings of the National Academy of Sciences*, *120*(12), e2214840120.

Stack Overflow. (2022). 2022 developer survey.

Tegmark, M. (2018). *Life 3.0: Being human in the age of artificial intelligence*. Vintage.

Tolan, S., Pesole, A., Martínez-Plumed, F., Fernández-Macías, E., Hernández-Orallo, J., & Gómez, E. (2021). Measuring the occupational impact of AI: Tasks, cognitive abilities and AI benchmarks. *Journal of Artificial Intelligence Research*, *71*, 191–236.

Webb, M. (2020). The impact of artificial intelligence on the labor market. (3482150).

Yang, X.-L., Lo, D., Xia, X., Wan, Z.-Y., & Sun, J.-L. (2016). What Security Questions Do Developers Ask? A Large-Scale Study of Stack Overflow Posts. *Journal of Computer Science and Technology*, *31*(5), 910–924.

Figure 1: Linking AI-questions on Stack Overflow to occupations and industries

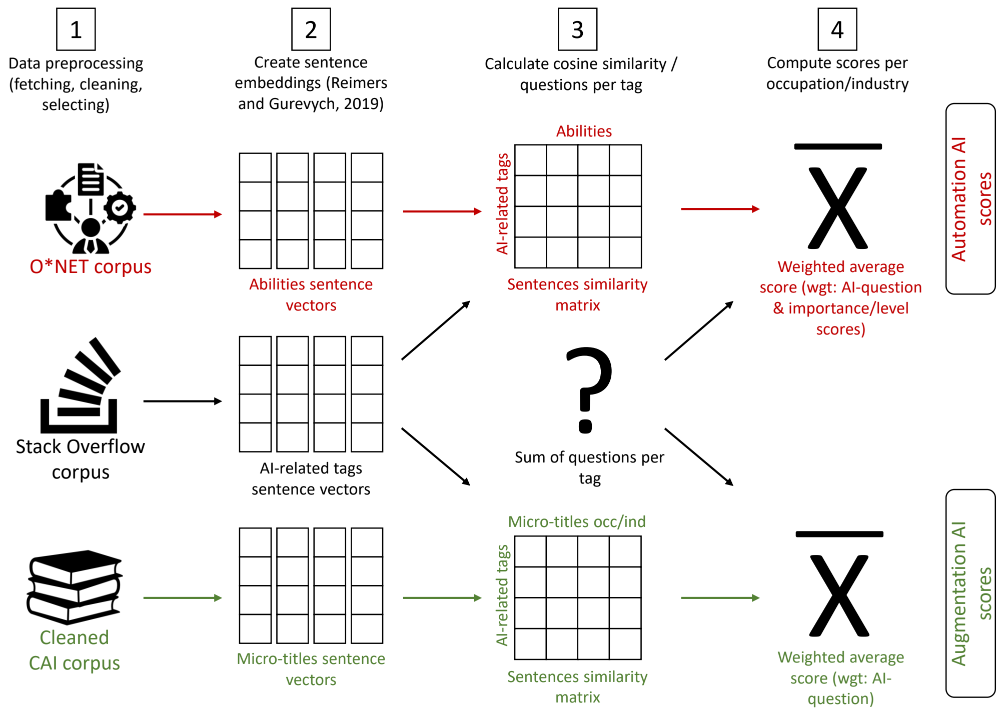
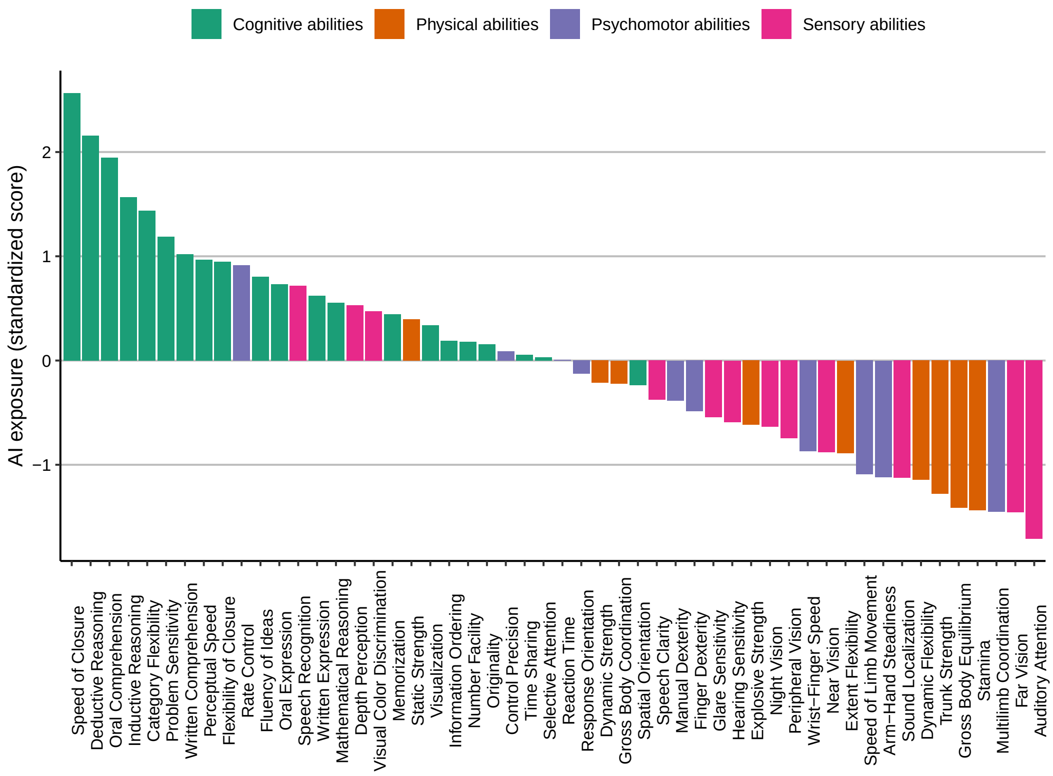
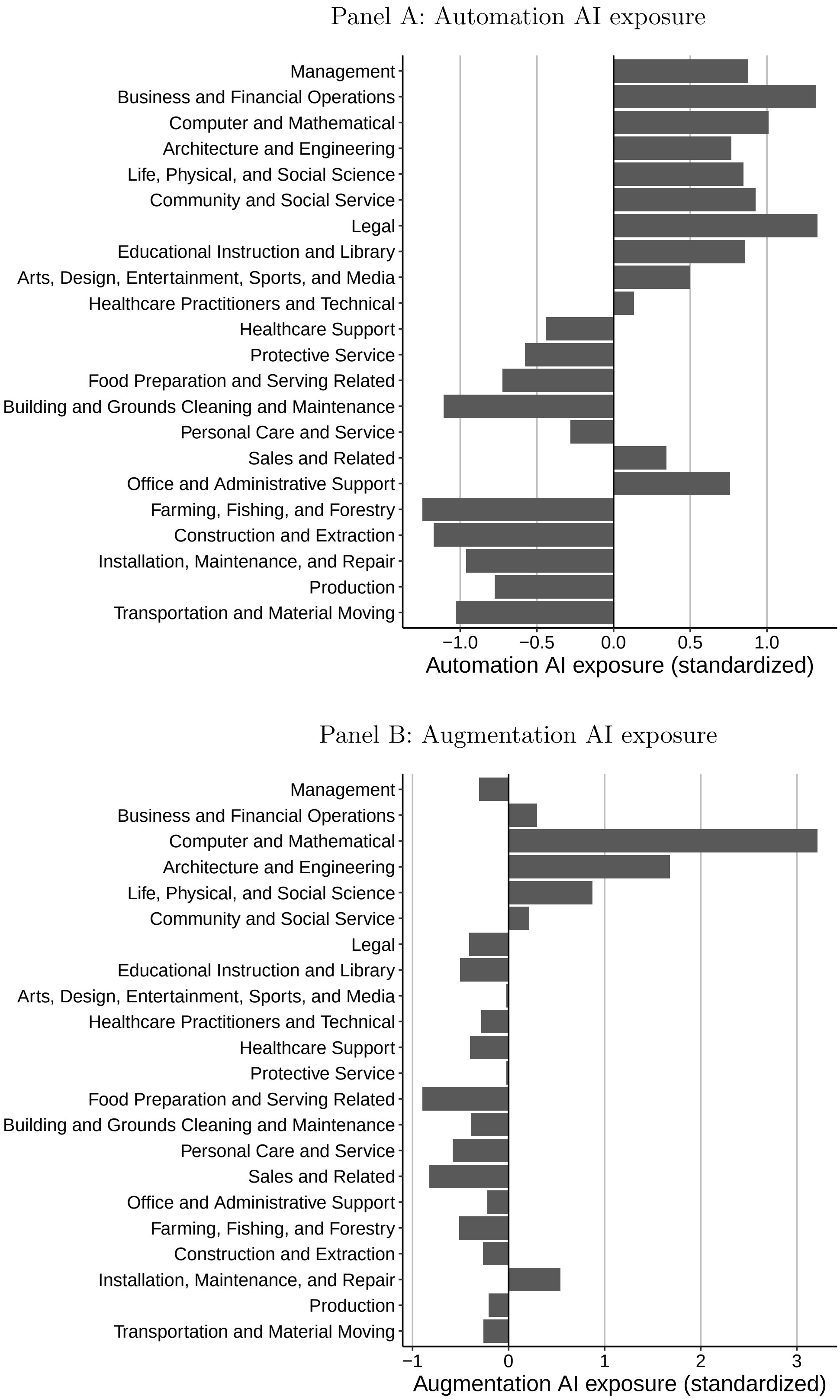



Figure 2: AI exposure for abilities



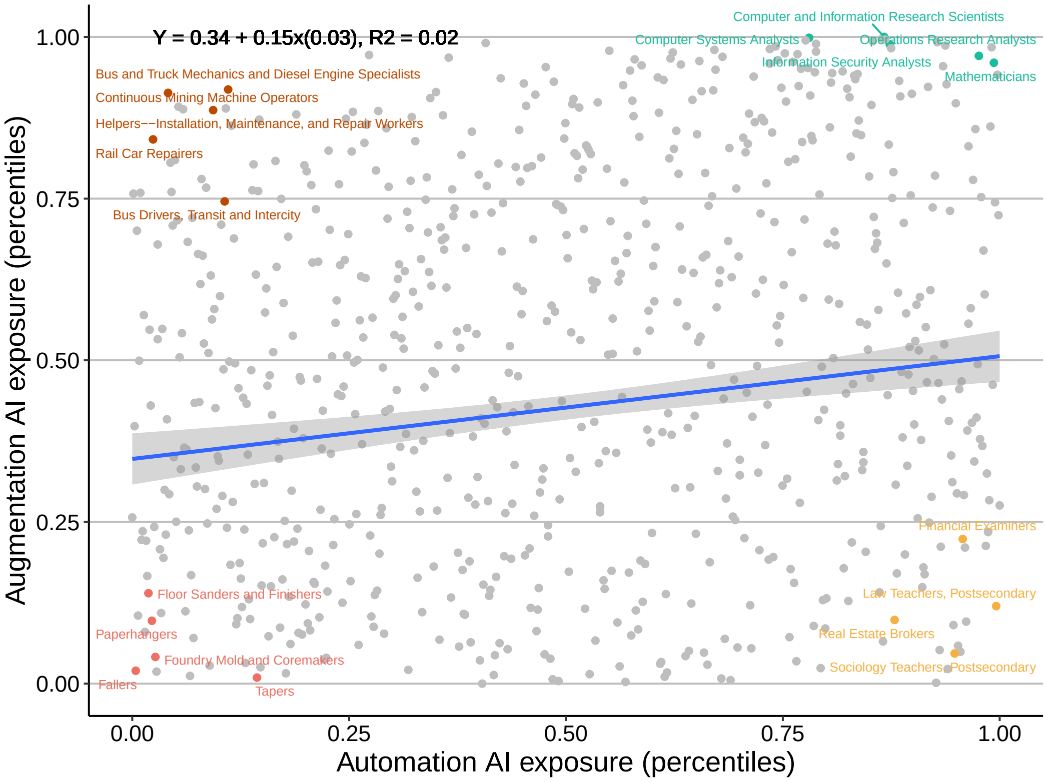

Note: The y-axis measures the standardized scores of AI exposure for abilities obtained from equation (3). Abilities are ranked in decreasing order of the score.

Figure 3: AI exposure by broad occupation in 2022




Note: The figure shows the exposure to automation AI (Panel A) and augmentation AI (Panel B) for broad occupations in 2022. Augmentation AI exposure is the weighted average exposure at the occupational level, using employment size in industries as weight. AI scores are the average of the exposure for occupations weighted by the employment size. Occupations are ranked in decreasing order of AI exposure. The data includes part-time and full-time employees who are paid a wage or salary. Data from establishments in farm industries are excluded from the sample (NAICS: 111, 112, 1131, 1132, 114, 1153, and 814), as well as industries within sector 92 Public Administration.

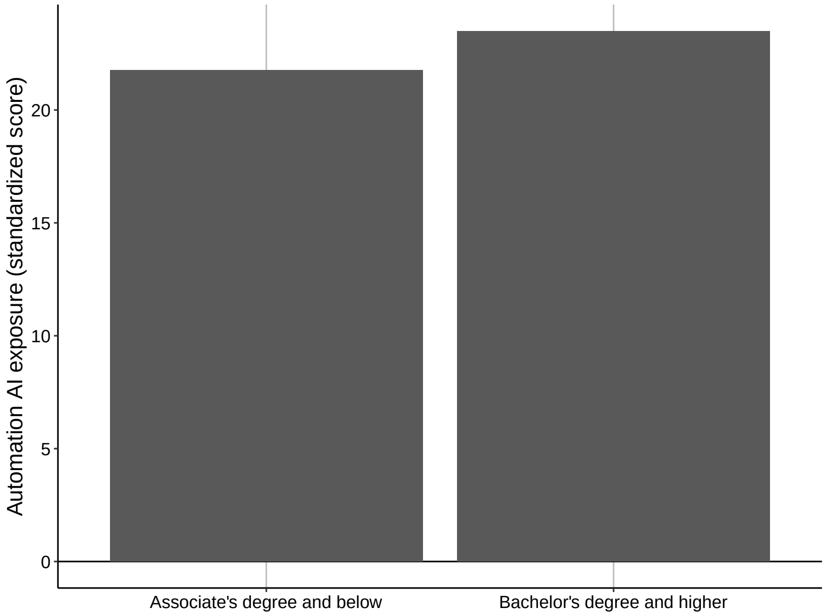
Figure 4: Correlation between automation AI and augmentation AI exposure, pooled 2010-2022



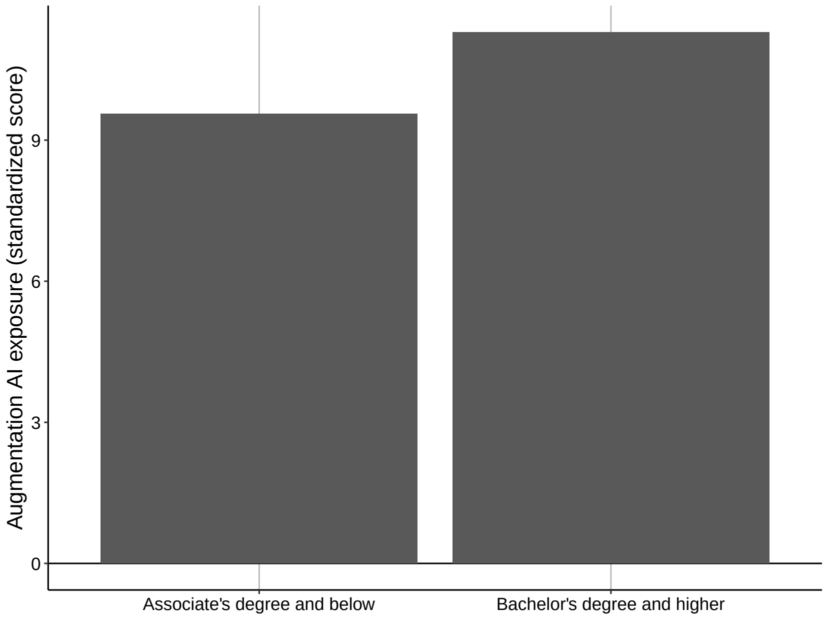

Note: The figure shows the relationship between automation AI exposure (x-axis) and augmentation AI exposure (y-axis) at the occupational level (6-digit). Augmentation AI exposure is the weighted average exposure at the occupational level, using employment size in industries as weight. The blue line represents the regression line weighted by the employment size. The regression line has a slope of 0.12 (SE = 0.03) and an intercept of 0.49 with R2 = 0.01. The data includes part-time and full-time employees who are paid a wage or salary. Data from establishments in farm industries are excluded from the sample (NAICS: 111, 112, 1131, 1132, 114, 1153, and 814), as well as industries within sector 92 Public Administration.

Figure 5: AI exposure by typical entry-level education, pooled 2010-2022

Panel A: Automation AI exposure



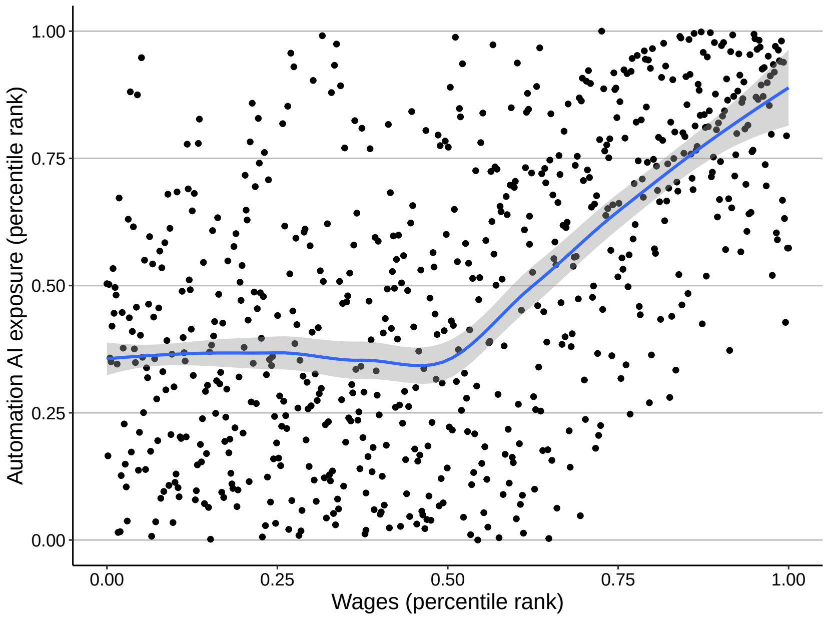

Panel B: Augmentation AI exposure



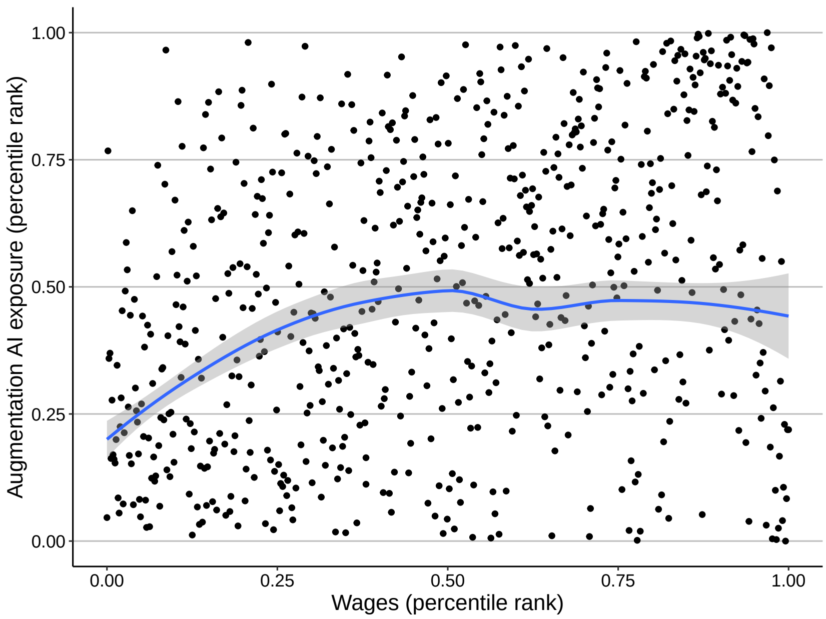

Note: The figure shows the average exposure to automation AI (Panel A) and augmentation AI (Panel B) by typical entry-level education. Augmentation AI exposure is the weighted average exposure at the occupational level, using employment size in industries as weight. AI scores are the average of the exposure for occupations weighted by the employment size. The data includes part-time and full-time employees who are paid a wage or salary. Data from establishments in farm industries are excluded from the sample (NAICS: 111, 112, 1131, 1132, 114, 1153, and 814), as well as industries within sector 92 Public Administration.

Figure 6: AI exposure and wages, pooled 2010-2022

Panel A: Automation AI exposure

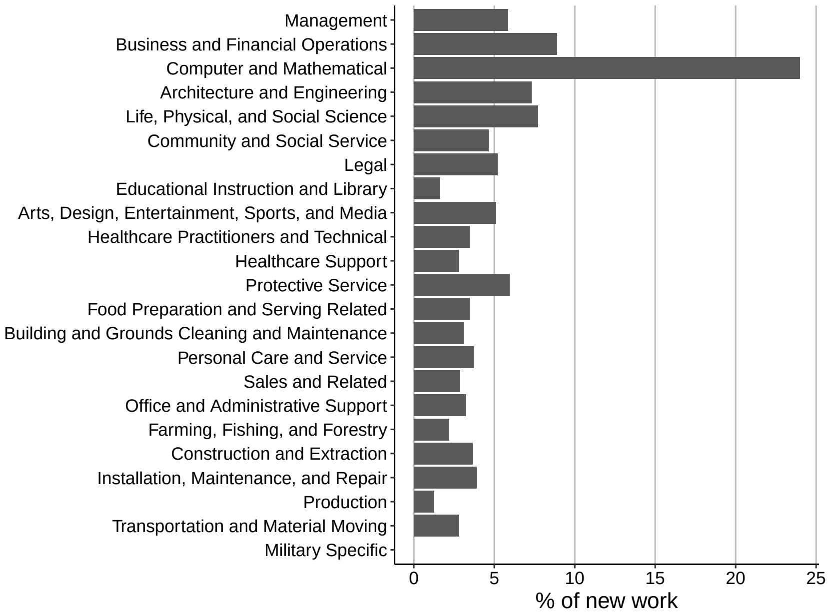



Panel B: Augmentation occupational AI exposure




Note: The figure shows the exposure to AI against the average hourly wages, both converted into percentile rank. The dots represent the occupations. The blue line is the smoothed conditional means modeling with the loess function using 70 points. The grey shadow displays the confidence interval at 0.95. The data includes part-time and full-time employees who are paid a wage or salary. Data from establishments in farm industries are excluded from the sample (NAICS: 111, 112, 1131, 1132, 114, 1153, and 814), as well as industries within sector 92 Public Administration.

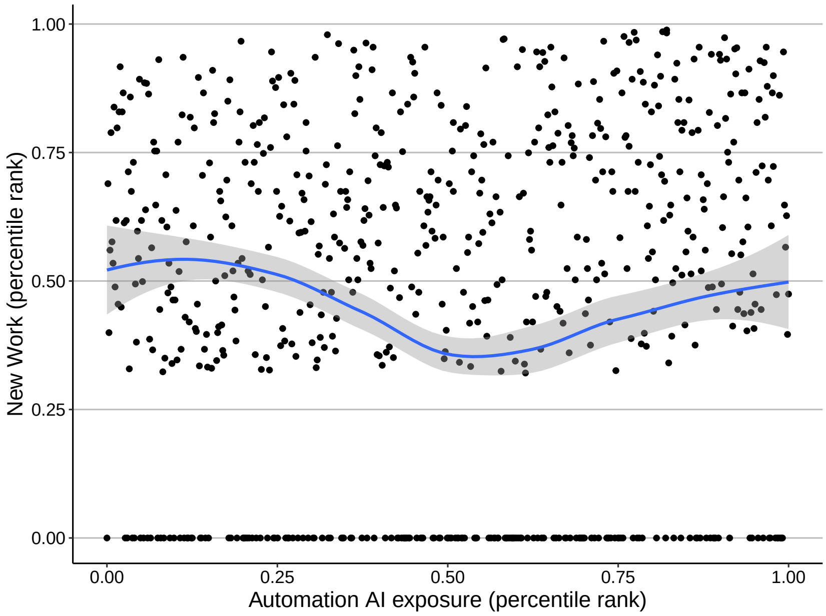
Figure 7: Share of new work by broad occupation, 2015-2022



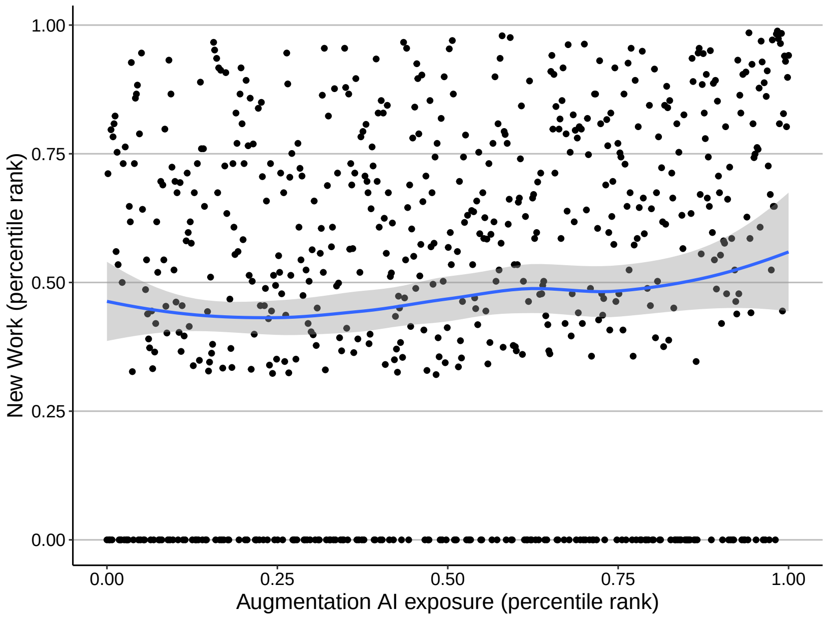

Note: The figure shows the percentage of new work by broad occupation between 2015 and 2022.

Figure 8: AI exposure and new work, pooled 2015-2022

Panel A: Automation AI exposure




Panel B: Augmentation occupational AI exposure




Note: The figure shows the percentage of new workk against the exposure to automation AI (Panel A) and augmentation AI (Panel B). The dots represent the occupations. The blue line is the smoothed conditional means modeling with the loess function using 70 points. The grey shadow displays the confidence interval at 0.95. The data includes part-time and full-time employees who are paid a wage or salary. Data from establishments in farm industries are excluded from the sample (NAICS: 111, 112, 1131, 1132, 114, 1153, and 814), as well as industries within sector 92 Public Administration.

Table 1: Occupations with the highest and lowest automation AI exposure, pooled 2010-2022

| Rank | Occupation | Automation AI score |
|---|---|---|
| 1 | Purchasing Agents, Except Wholesale, Retail, and Farm Products | 1.97 |
| 2 | Judges, Magistrate Judges, and Magistrates | 1.91 |
| 3 | Clinical and Counseling Psychologists | 1.85 |
| 4 | Management Analysts | 1.71 |
| 5 | Law Teachers, Postsecondary | 1.70 |
| 6 | Sociologists | 1.66 |
| 7 | Mathematicians | 1.63 |
| 8 | Environmental Science Teachers, Postsecondary | 1.59 |
| 9 | Mathematical Science Occupations, All Other | 1.58 |
| 10 | Political Scientists | 1.58 |
| 11 | Procurement Clerks | 1.57 |
| 12 | English Language and Literature Teachers, Postsecondary | 1.57 |
| 13 | Health Specialties Teachers, Postsecondary | 1.55 |
| 14 | Geography Teachers, Postsecondary | 1.55 |
| 15 | Psychology Teachers, Postsecondary | 1.54 |
| 16 | Library Science Teachers, Postsecondary | 1.53 |
| 17 | Epidemiologists | 1.52 |
| 18 | Mental Health Counselors | 1.52 |
| 19 | Business Teachers, Postsecondary | 1.51 |
| 20 | Operations Research Analysts | 1.51 |
| ... | ... | ... |
| 739 | Tire Builders | -1.63 |
| 740 | Rail Car Repairers | -1.64 |
| 741 | Paperhangers | -1.68 |
| 742 | Plasterers and Stucco Masons | -1.69 |
| 743 | Industrial Truck and Tractor Operators | -1.69 |
| 744 | Floor Sanders and Finishers | -1.71 |
| 745 | Fence Erectors | -1.72 |
| 746 | Dining Room and Cafeteria Attendants and Bartender Helpers | -1.74 |
| 747 | Dishwashers | -1.79 |
| 748 | Terrazzo Workers and Finishers | -1.79 |
| 749 | Insulation Workers, Floor, Ceiling, and Wall | -1.82 |
| 750 | Reinforcing Iron and Rebar Workers | -1.84 |
| 751 | Roof Bolters, Mining | -1.86 |
| 752 | Helpers–Brickmasons, Blockmasons, Stonemasons, and Tile and Marble Setters | -1.86 |
| 753 | Pressers, Textile, Garment, and Related Materials | -1.86 |
| 754 | Helpers–Roofers | -1.90 |
| 755 | Fallers | -1.92 |
| 756 | Structural Iron and Steel Workers | -1.97 |
| 757 | Helpers–Painters, Paperhangers, Plasterers, and Stucco Masons | -2.00 |
| 758 | Dancers | -2.17 |

Note: Occupations are ranked by their score of automation AI exposure at 6-digit Standard Occupational Classification (SOC) level and computed following equation (4). AI score is standardized to have a mean of 0 and a standard deviation of 1.

Table 2: Occupations with the highest and lowest augmentation AI exposure, pooled 2010-2022

| Rank | Occupation | Augmentation AI score |
|---|---|---|
| 1 | Computer and Information Research Scientists | 5.81 |
| 2 | Computer Systems Analysts | 4.30 |
| 3 | Computer Programmers | 3.93 |
| 4 | Software Developers | 3.85 |
| 5 | Computer Network Architects | 3.82 |
| 6 | Software Quality Assurance Analysts and Testers | 3.77 |
| 7 | Computer Hardware Engineers | 3.14 |
| 8 | Veterinary Technologists and Technicians | 2.88 |
| 9 | Web and Digital Interface Designers | 2.83 |
| 10 | Information Security Analysts | 2.83 |
| 11 | Industrial-Organizational Psychologists | 2.80 |
| 12 | Computer Occupations, All Other | 2.79 |
| 13 | Mathematical Science Occupations, All Other | 2.70 |
| 14 | Computer Science Teachers, Postsecondary | 2.67 |
| 15 | Database Architects | 2.63 |
| 16 | Civil Engineering Technologists and Technicians | 2.54 |
| 17 | Environmental Engineering Technologists and Technicians | 2.51 |
| 18 | Web Developers | 2.47 |
| 19 | Computer User Support Specialists | 2.39 |
| 20 | Agricultural Engineers | 2.36 |
| ... | ... | ... |
| 733 | Molders, Shapers, and Casters, Except Metal and Plastic | -1.51 |
| 734 | Residential Advisors | -1.53 |
| 735 | Family Medicine Physicians | -1.55 |
| 736 | Cooks, Institution and Cafeteria | -1.56 |
| 737 | Fallers | -1.57 |
| 738 | Rock Splitters, Quarry | -1.59 |
| 739 | First-Line Supervisors of Gambling Services Workers | -1.60 |
| 740 | Pourers and Casters, Metal | -1.61 |
| 741 | Chiropractors | -1.63 |
| 742 | Dental Hygienists | -1.63 |
| 743 | Stonemasons | -1.67 |
| 744 | Tellers | -1.75 |
| 745 | Tapers | -1.77 |
| 746 | Lodging Managers | -1.80 |
| 747 | Morticians, Undertakers, and Funeral Arrangers | -1.86 |
| 748 | Clergy | -1.93 |
| 749 | Dentists, General | -2.01 |
| 750 | Orthodontists | -2.10 |
| 751 | Loan Officers | -2.18 |
| 752 | Oral and Maxillofacial Surgeons | -2.18 |

Note: The table shows the weighted average exposure at the occupational level, using employment size in industries as weight. Occupations are ranked by their score of augmentation AI exposure at 6-digit Standard Occupational Classification (SOC) level and computed following equation (9). AI score is standardized to have a mean of 0 and a standard deviation of 1.

Table 3: Example of new alternate titles per year, 2016-2022

| Year | Alternate titles | Occupations |
|---|---|---|
| 2016 | Family Reunification Specialist | Social and Human Service Assistants (21-1093) |
| | Scrum Master | Computer Occupations, All Other (15-1299) |
| 2017 | Executive Cyber Leader | Chief Executives (11-1011) |
| | Online Health and Fitness Coach | Health Education Specialists (21-1091) |
| 2018 | Sprinkler Design Technician | Civil Engineering Technologists and Technicians (17-3022) |
| | Autonomous Vehicle Design Engineer | Engineers, All Other (17-2199) |
| 2019 | Safety Research Professional | Occupational Health and Safety Technicians (19-5012) |
| | Route Diver | Commercial Divers (49-9092) |
| 2020 | Blockchain Penetration Tester | Computer Occupations, All Other (15-1299) |
| | Culinary Artist | Cooks, Private Household (35-2013) |
| 2021 | Solar Site Surveyor | Surveyors (17-1022) |
| | Remote Pilot | Life, Physical, and Social Science Technicians, All Other (19-4099) |
| 2022 | Recruiter Sourcing | Human Resources Specialists (13-1071) |
| | Cat Groomer | Animal Trainers (39-2011) |

Note: Examples of new alternate titles added in O*NET per year from 2016 to 2022. Occupations correspond to the Standard Occupational Classification 2018, and codes are given in parentheses.

Table 4: Descriptive statistics: Wages and employment

| | Mean | SD | Min | Max |
|---|---|---|---|---|
| Hourly wages | 34.8 | 18.9 | 10.0 | 182.6 |
| Employment size | 6070 | 40 990 | 30 | 3 244 470 |

Table 5: Relationships with previous AI exposure indices

| | Automation AI exposure | | | Augmentation AI exposure | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Felten et al. (2021) | 0.966*** | | | 0.204*** | | |
| | (0.009) | | | (0.036) | | |
| Brynjolfsson et al. (2018) | | 0.232*** | | | 0.030 | |
| | | (0.036) | | | (0.037) | |
| Webb (2020) | | | −0.021 | | | 0.310*** |
| | | | (0.036) | | | (0.035) |
| Intercept | 0.017*** | 0.384*** | 0.510*** | 0.398*** | 0.485*** | 0.345*** |
| | (0.005) | (0.021) | (0.021) | (0.021) | (0.021) | (0.020) |
| $R^2$ | 0.933 | 0.053 | −0.001 | 0.040 | 0.000 | 0.095 |
| Observations | 749 | 749 | 749 | 749 | 749 | 749 |

Note: Augmentation AI exposure is the weighted average exposure at the occupational level, using employment size in industries as weight. Indices are converted into percentile ranks. The indices for automation AI exposure and augmentation AI exposure are for 2022. Columns 1 to 3 use automation AI occupation for the dependent variable, while columns 4 to 6 use augmentation AI exposure. The table shows the results using OLS estimators. $\star\star\star$ Significant at the 1% level; $\star\star$ significant at the 5% level; $\star$ significant at the 10% level.

Table 6: AI exposure and Routine tasks indices

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | *Panel A: Automation AI exposure* | | | | | | |
| Routine task index | −0.468*** | | | | | | |
| | (0.032) | | | | | | |
| Routine manual | | −0.792*** | | | | | |
| | | (0.022) | | | | | |
| Routine cognitive | | | −0.165*** | | | | |
| | | | (0.036) | | | | |
| Non-routine manual | | | | −0.892*** | | | |
| | | | | (0.016) | | | |
| Non-routine interpersonal | | | | | 0.460*** | | |
| | | | | | (0.032) | | |
| Non-routine analitycal | | | | | | 0.694*** | |
| | | | | | | (0.026) | |
| Decision-making | | | | | | | 0.592*** |
| | | | | | | | (0.029) |
| Intercept | 0.734*** | 0.896*** | 0.582*** | 0.946*** | 0.270*** | 0.153*** | 0.205*** |
| | (0.019) | (0.013) | (0.021) | (0.010) | (0.019) | (0.015) | (0.017) |
| $R^2$ | 0.218 | 0.628 | 0.026 | 0.796 | 0.211 | 0.482 | 0.349 |
| | *Panel B: Augmentation AI exposure* | | | | | | |
| Routine task index | −0.133*** | | | | | | |
| | (0.036) | | | | | | |
| Routine manual | | −0.086** | | | | | |
| | | (0.036) | | | | | |
| Routine cognitive | | | 0.057 | | | | |
| | | | (0.036) | | | | |
| Non-routine manual | | | | −0.028 | | | |
| | | | | (0.036) | | | |
| Non-routine interpersonal | | | | | −0.022 | | |
| | | | | | (0.036) | | |
| Non-routine analitycal | | | | | | 0.292*** | |
| | | | | | | (0.035) | |
| Decision-making | | | | | | | 0.178*** |
| | | | | | | | (0.036) |
| Intercept | 0.567*** | 0.543*** | 0.472*** | 0.514*** | 0.511*** | 0.354*** | 0.411*** |
| | (0.021) | (0.021) | (0.021) | (0.021) | (0.021) | (0.020) | (0.021) |
| $R^2$ | 0.016 | 0.006 | 0.002 | −0.001 | −0.001 | 0.084 | 0.030 |
| Observations | 752 | 752 | 752 | 752 | 752 | 752 | 752 |

Note: Augmentation AI exposure is the weighted average exposure at the occupational level, using employment size in industries as weight. The routine and non-routine indices in columns 1 to 6 are constructed following Acemoglu and Autor (2011). In column 7, the index of decision-making follows Deming (2021).Indices are converted into percentile ranks. The indices for automation AI exposure (Panel A) and augmentation AI exposure (Panel B) are for 2022. Panel A uses automation AI occupation for the dependent variable, while Panel B uses augmentation AI exposure. The table shows the results using OLS estimators. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

Table 7: Relationships with other factors affecting the labor demand

| | Automation AI exposure | | | Augmentation AI exposure | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Offshorability | 0.611*** | | | 0.049 | | |
| | (0.029) | | | (0.037) | | |
| Software exposure | | −0.340*** | | | 0.239*** | |
| | | (0.034) | | | (0.036) | |
| Robot exposure | | | −0.739*** | | | 0.036 |
| | | | (0.025) | | | (0.037) |
| Intercept | 0.195*** | 0.670*** | 0.869*** | 0.476*** | 0.380*** | 0.482*** |
| | (0.017) | (0.020) | (0.014) | (0.021) | (0.021) | (0.021) |
| $R^2$ | 0.372 | 0.115 | 0.546 | 0.001 | 0.056 | 0.000 |
| Observations | 749 | 749 | 749 | 749 | 749 | 749 |

Note: Augmentation AI exposure is the weighted average exposure at the occupational level, using employment size in industries as weight. Indices are converted into percentile ranks. The indices for automation AI exposure and augmentation AI exposure are for 2022. The indices measuring exposure to robots and software are sourced from Webb (2020). The index of offshorability is built following Autor et al. (2013). Columns 1 to 3 use automation AI occupation for the dependent variable, while columns 4 to 6 use augmentation AI exposure. The table shows the results using OLS estimators. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

Table 8: Relationships AI exposures and instrumental variables

| | Automation AI | Augmentation AI |
|---|---|---|
| Automation AI IV | 1.072*** | |
| | (0.001) | |
| Augmentation AI IV | | 1.019*** |
| | | (0.003) |
| $R^2$ | 0.999 | 0.999 |
| Observations | 230 204 | 230 204 |

Note: The table shows the relationships between the measures of AI exposures and their corresponding instrumental variables. The estimates are computed using OLS estimators with year fixed effects. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

Table 9: Effect of AI exposure on share of new work

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| | *Panel A: Fixed effect estimators* | | | | | |
| Automation AI | 0.029 | 0.006 | | | 0.009 | 0.000 |
| | (0.029) | (0.024) | | | (0.025) | (0.022) |
| Augmentation AI | | | 0.028*** | 0.037*** | 0.027*** | 0.037*** |
| | | | (0.008) | (0.009) | (0.008) | (0.009) |
| $R^2$ | 0.673 | 0.689 | 0.688 | 0.705 | 0.688 | 0.705 |
| | *Panel B: IV estimators* | | | | | |
| Automation AI | 0.030 | 0.006 | | | 0.008 | 0.000 |
| | (0.029) | (0.024) | | | (0.025) | (0.022) |
| Augmentation AI | | | 0.028*** | 0.037*** | 0.028*** | 0.037*** |
| | | | (0.008) | (0.009) | (0.008) | (0.009) |
| $R^2$ | 0.673 | 0.689 | 0.688 | 0.705 | 0.688 | 0.705 |
| *Montiel-Pflueger F-test* | $7 \times 10^5$ | $8 \times 10^5$ | $3 \times 10^4$ | $1 \times 10^5$ | | |
| *Covariates included:* | | | | | | |
| Employment (log) | X | X | X | X | X | X |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 141 664 | 141 664 | 141 664 | 141 664 | 141 664 | 141 664 |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| SD outcome | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |

Note: The table presents the outputs for the regressions (5) when the dependent variable is the cumulative share of new work. Panel A shows the output when fixed estimators are used, whereas Panel B is when IV estimators is applied. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. The estimations are weighted by employment size. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level

Table 10: Effect of AI exposure on employment (log)

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| | *Panel A: Fixed effect estimators* | | | | | |
| Automation AI | 0.113 | 0.184 | | | 0.040 | 0.160 |
| | (0.120) | (0.116) | | | (0.108) | (0.112) |
| Augmentation AI | | | 0.089*** | 0.094*** | 0.087*** | 0.091*** |
| | | | (0.023) | (0.031) | (0.023) | (0.030) |
| $R^2$ | 0.993 | 0.994 | 0.993 | 0.994 | 0.993 | 0.994 |
| | *Panel B: IV estimators* | | | | | |
| Automation AI | 0.120 | 0.189 | | | 0.044 | 0.164 |
| | (0.121) | (0.117) | | | (0.109) | (0.113) |
| Augmentation AI | | | 0.090*** | 0.093*** | 0.087*** | 0.090*** |
| | | | (0.024) | (0.031) | (0.024) | (0.030) |
| $R^2$ | 0.993 | 0.994 | 0.993 | 0.994 | 0.993 | 0.994 |
| *Montiel-Pflueger F-test* | $5 \times 10^5$ | $8 \times 10^5$ | $5 \times 10^4$ | $1 \times 10^5$ | | |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 10.9 | 10.9 | 10.9 | 10.9 | 10.9 | 10.9 |
| SD outcome | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 |

Note: The table presents the outputs for the regressions (6) when the dependent variable is the employment size in log. Panel A shows the output when fixed estimators are used, whereas Panel B is when IV estimators is applied. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. The estimations are weighted by employment size. Standard errors are reported in brackets and are clustered at the occupational and industry level. $\star\star\star$ Significant at the 1% level; $\star\star$ significant at the 5% level; $\star$ significant at the 10% level

Table 11: Effect of AI exposure on hourly wages (log)

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| | *Panel A: Fixed effect estimators* | | | | | |
| Automation AI | −0.193*** | −0.173*** | | | −0.183*** | −0.174*** |
| | (0.042) | (0.028) | | | (0.038) | (0.028) |
| Augmentation AI | | | −0.021** | 0.000 | −0.012 | 0.003 |
| | | | (0.009) | (0.006) | (0.008) | (0.006) |
| $R^2$ | 0.994 | 0.996 | 0.994 | 0.995 | 0.994 | 0.996 |
| | *Panel B: IV estimators* | | | | | |
| Automation AI | −0.197*** | −0.176*** | | | −0.187*** | −0.177*** |
| | (0.043) | (0.028) | | | (0.038) | (0.028) |
| Augmentation AI | | | −0.022** | 0.000 | −0.012 | 0.003 |
| | | | (0.009) | (0.006) | (0.007) | (0.006) |
| $R^2$ | 0.994 | 0.996 | 0.994 | 0.995 | 0.994 | 0.996 |
| *Montiel-Pflueger F-test* | $9 \times 10^6$ | $8 \times 10^6$ | $4 \times 10^5$ | $4 \times 10^5$ | | |
| *Covariates included:* | | | | | | |
| Employment (log) | X | X | X | X | X | X |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 3.2 | 3.2 | 3.2 | 3.2 | 3.2 | 3.2 |
| SD outcome | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |

Note: The table presents the outputs for the regressions (6) when the dependent variable is the employment size in log. Panel A shows the output when fixed estimators are used, whereas Panel B is when IV estimators is applied. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. The estimations are weighted by employment size. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level

Table 12: Effect of AI exposure by typical entry-level education in occupations

| | *Share New Work* | | *Employment (log)* | | *Hourly wages (log)* | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| | Associate's degree and lower | Bachelor's degree and higher | Associate's degree and lower | Bachelor's degree and higher | Associate's degree and lower | Bachelor's degree and higher |
| | *Panel A: Fixed effect estimators* | | | | | |
| Automation AI | −0.044* | 0.090 | −0.074 | 0.312 | −0.146*** | 0.120 |
| | (0.025) | (0.082) | (0.142) | (0.333) | (0.036) | (0.109) |
| Augmentation AI | 0.012 | 0.048*** | 0.091** | 0.046 | −0.012 | 0.019*** |
| | (0.014) | (0.009) | (0.045) | (0.037) | (0.010) | (0.005) |
| $R^2$ | 0.716 | 0.698 | 0.994 | 0.993 | 0.992 | 0.987 |
| | *Panel B: IV estimators* | | | | | |
| Automation AI | −0.044* | 0.089 | −0.070 | 0.295 | −0.148*** | 0.125 |
| | (0.025) | (0.082) | (0.143) | (0.337) | (0.036) | (0.112) |
| Augmentation AI | 0.011 | 0.049*** | 0.091** | 0.045 | −0.012 | 0.019*** |
| | (0.013) | (0.010) | (0.045) | (0.038) | (0.010) | (0.005) |
| $R^2$ | 0.716 | 0.698 | 0.994 | 0.993 | 0.992 | 0.987 |
| *Covariates included:* | | | | | | |
| Employment (log) | X | X | | | X | X |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | X | X | X | X | X | X |
| Observations | 89 056 | 52 608 | 144 716 | 85 488 | 144 716 | 85 488 |
| Unique SOC (6-digit) | 470 | 202 | 470 | 202 | 470 | 202 |
| Unique NAICS (4-digit) | 244 | 239 | 244 | 239 | 244 | 239 |
| Mean outcome | 0.02 | 0.03 | 11.1 | 10.1 | 3.0 | 3.9 |
| SD outcome | 0.03 | 0.06 | 2.0 | 2.0 | 0.3 | 0.3 |

Note: The table presents the outputs for the regressions (5) (columns 1 and 2) and (6) (columns 3-4 and 5-6). In columns 1 and 2, the response variable is the share of new work, in columns 3 and 4 the employment size (log), and in columns 5 and 6 the hourly wages (log). In column 1, 3, and 5, the sample size is composed of occupations having a typical entry-level education equal to Associate's degree or lower. In columns 2, 4, and 6, only occupations with a typical entry-level education equal to Bachelor's degree of higher are retained. Panel A present the results using fixed effect estimators, while Panel B relies on IV estimators. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level

Table 13: Effect of AI exposure by AI technologies

| | *Share New Work* | | | *Employment (log)* | | | *Hourly wages (log)* | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| | *Panel A: Fixed effect estimators* | | | | | | | | |
| Auto. Computer Vision | −0.124 | | | 0.013 | | | −0.015 | | |
| | (0.098) | | | (0.371) | | | (0.086) | | |
| Augm. Computer Vision | 0.037*** | | | 0.028 | | | 0.009 | | |
| | (0.013) | | | (0.031) | | | (0.006) | | |
| Auto. Lang. processing & modeling | | −0.014 | | | 0.019 | | | −0.072*** | |
| | | (0.013) | | | (0.056) | | | (0.015) | |
| Augm. Lang. processing & modeling | | 0.043*** | | | 0.073** | | | 0.010 | |
| | | (0.012) | | | (0.033) | | | (0.006) | |
| Auto. Machine & Deep Learning | | | 0.008 | | | 0.195* | | | −0.173*** |
| | | | (0.019) | | | (0.101) | | | (0.025) |
| Augm. Machine & Deep Learning | | | 0.032*** | | | 0.101*** | | | −0.001 |
| | | | (0.007) | | | (0.029) | | | (0.005) |
| $R^2$ | 0.697 | 0.703 | 0.705 | 0.994 | 0.994 | 0.994 | 0.995 | 0.996 | 0.996 |
| | *Panel B: IV estimators* | | | | | | | | |
| Auto. Computer Vision | −0.118 | | | −0.074 | | | 0.068 | | |
| | (0.097) | | | (0.363) | | | (0.092) | | |
| Augm. Computer Vision | 0.037*** | | | 0.034 | | | 0.006 | | |
| | (0.013) | | | (0.031) | | | (0.006) | | |
| Auto. Lang. processing & modeling | | −0.014 | | | 0.016 | | | −0.072*** | |
| | | (0.013) | | | (0.056) | | | (0.015) | |
| Augm. Lang. processing & modeling | | 0.044*** | | | 0.074** | | | 0.010 | |
| | | (0.012) | | | (0.033) | | | (0.006) | |
| Auto. Machine & Deep Learning | | | 0.008 | | | 0.198* | | | −0.174*** |
| | | | (0.019) | | | (0.101) | | | (0.025) |
| Augm. Machine & Deep Learning | | | 0.033*** | | | 0.102*** | | | 0.000 |
| | | | (0.007) | | | (0.029) | | | (0.005) |
| $R^2$ | 0.697 | 0.703 | 0.705 | 0.994 | 0.994 | 0.994 | 0.995 | 0.996 | 0.996 |
| *Covariates included:* | | | | | | | | | |
| Employment (log) | X | X | X | | | | X | X | X |
| *Fixed effects:* | | | | | | | | | |
| Year | X | X | X | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X | X | X | X |
| NAICS*year (3-digit) | X | X | X | X | X | X | X | X | X |
| Observations | 141 664 | 141 664 | 141 664 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 0.02 | 0.02 | 0.02 | 10.9 | 10.9 | 10.9 | 3.2 | 3.2 | 3.2 |
| SD outcome | 0.04 | 0.04 | 0.04 | 2.1 | 2.1 | 2.1 | 0.5 | 0.5 | 0.5 |

Note: The table presents the outputs for the regressions (5) (columns 1, 2, and 3) and (6) (columns 4-6 and 7-9). In columns 1 to 3, the response variable is the share of new work, in columns 4 to 6 the employment size (log), and in columns 7 to 9 the hourly wages (log). Panel A present the results using fixed effect estimators, while Panel B relies on IV estimators. Standard errors are reported in brackets and are clustered at the occupational and industry level. $\star\star\star$ Significant at the 1% level; $\star\star$ significant at the 5% level; $\star$ significant at the 10% level

# APPENDIX

# A Data sources and AI indices

## A.1 Stack Overflow

Created in 2008, Stack Overflow is a Q&A website for questions about programming problems, software algorithms, and software tools for developers. Access to Stack Overflow is free, and everyone can create an account and start asking, answering, and commenting on questions. Stack Overflow counts 24 million questions for approximately 35 million answers and 20 million users at the beginning of 2023.

Stack Overflow is the most famous website for debugging code and the primary source of help for programmers. In 2022, it received 250 million monthly visitors, more than three times the monthly visitors of its direct competitors, such as w3schools.com (70.4 million monthly visits) and geeksforgeeks.com (64.7 million monthly visits).[20]

A key feature of Stack Overflow for this project is the importance of developers among the website users. A high proportion of developers among the users ensures that the questions asked on SO are related to algorithms deployed in the labor market and do not concern leisure. This aspect is crucial for this study since it aims at measuring AI exposure in the labor market. According to the annual survey conducted by Stack Overflow, 73% of the users are developers by profession, and 80% write code as part of their work (Stack Overflow, 2022).

Stack Overflow presents several advantages compared to alternative sources of information previously used in the literature, such as patents and surveys filled by workers or experts in the field of AI (see, for instance, Brynjolfsson et al., 2018; Felten et al., 2021; Tolan et al., 2021; Webb, 2020). Patents are likely to provide an incomplete measure of AI exposure for two reasons. First, it has been shown that AI systems are primarily protected as trade secrets because protecting them under copyright and patent laws encounters difficulties (Foss-Solbrekk, 2021; Hattenbach and Snyder, 2018; Hu and Jiang, 2019). Second, there are strong incentives to release AI algorithms in open source. It is a convenient way to quickly test and prototype AI solutions, allowing developers to gain insights and iterate without needing extensive in-house development. The release of the large language model Llama2 by Meta in open-source and the success of the online platform HuggingFace are examples of these incentives. Regarding surveys filled by workers or experts, they usually have small sample sizes, and the questions asked concern more often the potential of AI rather than its actual development and implementation in the economy.

Stack Overflow also comes with some drawbacks. The first drawback concerns AI

[20]For more details, see Similarweb, an online company measuring online audiences: https://www.similarweb.com/website/stackoverflow.com/competitors.

algorithms developed internally by firms, which Stack Over might not capture. However, this drawback is negligible and should not affect the results of this study. Indeed, an AI algorithm is built on various sub-algorithms and software; some are likely to be referenced in Stack Overflow. For instance, a developer aiming to create an internal chatbot for her company will probably use Python libraries like Tensorflow, NLTK, or Chatterbot. All these packages are referenced in Stack Overflow, and developers may ask questions about them during the production of the chatbot. A second potential drawback is the use of commercial AI solutions by firms. This drawback appears limited since creators of commercial AI solutions might want to be referenced in Stack Overflow for two reasons: gaining in popularity and providing support to their clients to debug their code themselves. As an illustration, 650 tags refer to commercial Google solutions, including AI solutions such as Google Speech-to-Text, Google Translate, and Google Natural Language.

Data from Stack Overflow has already been used to explore related topics to this study in computer science and technology research. For instance, questions on Stack Overflow have been used to analyze topics discussed among developers (Barua et al., 2014; Rosen and Shihab, 2016; Yang et al., 2016). Moutidis and Williams (2021) rely on Stack Overflow to document technologies used by developers, and Montandon et al. (2021) use job adverts published on Stack Overflow to study skills demand from IT companies.

Despite a few recent examples, data from Stack Overflow has been neglected in the economics literature. Gallea (2023) uses questions asked on Stack Overflow to study the effect of AI on work dynamics, and del Rio-Chanona et al. (2023) looks at potential threats of large language models to digital public goods. Closer to my study, the OECD.AI Policy Observatory develops a set of indicators to measure AI knowledge flows using questions and answers on Stack Overflow (OECD.AI, 2023).

In this project, I consider only questions asked between 2010 and 2022. While it would be feasible to add further years, the release of ChatGPT in November 2022 drastically hurts the attractiveness of the website (del Rio-Chanona et al., 2023), making Stack Overflow less relevant for more recent years.

## A.2 Identifying AI-related questions

The first step in building the measures of AI exposure is to identify AI-related questions. To do so, I exploit the information the tags attached to the questions provide. I identify tags about AI and consider any questions attached to these tags as AI-related (see Appendix **??** for more details).

Tags are keywords that sort the questions into categories. A tag may refer to a technology, a programming language, or a task developers want to perform. For instance, the following tags are on Stack Overflow: Python, GitHub, web scraping, indexing, and Tensorflow. A question requires between 3 to 5 tags to be published. Stack Overflow counts more than 63 000 tags, which usually come with a short technical description.

The identification of tags related to AI proceeds in three steps. In the first step, I search for 164 AI keywords in the tags and their technical descriptions. The list of AI keywords is based on the keywords identified in Alekseeva et al. (2021), which I complement with keywords from the computer science and technology literature and AI-specialized websites. Table A1 presents the list of keywords. I identify934 tags in this step.

This approach has the advantage of being straightforward, but it also has drawbacks. The identification quality depends on the keywords, and some tags might be missing if the list of keywords is not comprehensive. This limit is probably more problematic for AI than any other topic because this field evolves quickly, and new algorithms and methods appear often.

To mitigate this limitation, I supplement the first identification step with a second one using tags co-appearance. Tags co-appearance consists of adding to the selection the tags that appear together in questions with tags identified in the first step. This second step ensures that all tags related to AI are identified. For instance, the following tags are added with this procedure: seaborn, csv, imbalanced-data, and scikit-multilearn. In this step, 21 837 tags are added, and the number of potential tags related to AI reaches 22 771. In the end of this step, I end up with a list of potential AI-related tags.

The second identification step significantly increases the number of potential tags related to AI and the false positives. False positives concern all tags wrongly classified as related to AI. For instance, taking the examples found with the tags co-appearance procedure, the tags seaborn and csv are misclassified and should not considered as related to AI.

In the last identification step, I drop the false positives from the list of potential tags related to AI with the help of ChatGPT 3.5 Turbo. For instance, the tag "Python" is identified during the co-appearance step. Python is widely used among developers creating AI algorithms, but not exclusively, and, therefore, should be dropped from the selection. For each tag identified in the first two steps, I ask ChatGPT 3.5 Turbo whether they are related to AI. I use the following prompt for this task:

> *I will provide a tag and its description from the website www.stackoverflow.com. Please return 1 if the tag is related to or used in artificial intelligence and 0 otherwise.*

For the final list of AI-related tags, I keep only tags that ChatGPT3.5 Turbo classifies as related to AI. I end up with 1182 AI-related tags, which encompasses 678 115 questions.[21].

Once the AI-related tags are identified, I need to describe the type of tasks that one wants to perform with them. This information is necessary to match the number of questions per tag with the abilities required in occupations. I perform this task with ChatGPT 3.5 Turbo and the following prompt:

[21] In a previous version of this paper, I made the checking manually and ended up with a lower number of AI-related tags and questions. Running the analysis with this previous selection of AI-related tags does not affect the results of this study.

*I will provide an AI-related tag and its description from the website www.stackoverflow.com. Please describe which tasks someone wants to perform with this tag. The answer must fit into 60 tokens maximum.*

I explicitly ask for the type of task one wants to perform with the tag used because this is the information I need to match AI-related questions with the abilities required for occupations. The length of the answer is limited to 60 tokens to be consistent with the length of the descriptors in O*NET.

Figure A1 shows the yearly number of questions asked on Stack Overflow. The number of questions increases sharply between 2010 and 2017. During this period, AI has seen considerable progress with the introduction of Generative Adversarial Networks (2014), Residual Networks (2015), Recurrent Neural Networks (2015), Long Short-Term Memory (2015), and Transformer Structure (2017). The number of yearly questions stabilizes around 80 000 from 2017 to 2020 and drops in 2021. In 2022, the number of newly AI-related questions reaches 73 430.

Figure A2 shows the distribution of questions per AI-related tags. The distribution is right-skewed, indicating that some tags are highly used. The top 5 AI-related tags are Tensorflow (79 729 questions), Apache Spark (78 435), OpenCV (69 694), Machine Learning (52 299), and Keras (40 940).

Table A1: AI keywords

| | | | |
|---|---|---|---|
| Activity-Recognition | Generative AI | Mahout | Sdscm |
| Ai Chatbot | Gesture-Recognition | Marf | Semi-Supervised Learning |
| Ai Kibit | Google Cloud Machine Learning Platform | Microsoft Cognitive Toolkit | Semantic Driven Subtractive Clustering Method |
| Antlr | Google-Cloud-Ml | Microsoft-Cognitive | Sentence-Transformers |
| Apache-Spark-Ml | Google-Colaboratory | Microsoft-Translator | Sentiment Analysis |
| Apache-Spark-Mllib | Google-Speech-Api | Midjourney | Sentiment Classification |
| Apertium | Gpt-2 | Mlpack | Sentiment-Analysis |
| Artificial-Intelligence | Gpt-3 | Mlpy | Sklearn-Pandas |
| ASR | Gpt-4 | Modular Audio Recognition Framework | Spacy-Transformers |
| Automatic Speech Recognition | Gradient Boosting | Moses | Speech Recognition |
| Automl | H2O | Multilabel-Classification | Speech-Recognition |
| Azure-Cognitive-Services | Handwriting-Recognition | Mxnet | Speech-Synthesis |
| Azure-Language-Understanding | Huggingface-Transformers | Named-Entity-Recognition | Speech-To-Text |
| Bert | Ibm Watson | Natural Language Processing | Stable Diffusion |
| Bert-Language-Model | Image Processing | Natural Language Toolkit | Stanford-Nlp |
| Caffe Deep Learning Framework | Image Recognition | Nd4J | Supervised Learning |
| Chatbot | Image-Processing | Nearest Neighbor Algorithm | Support Vector Machines |
| Chatgpt | Image-Segmentation | Neural-Network | SVM |
| Chatgpt-Api | Information-Extraction | NLP | Tensor |
| Classification | Information-Retrieval | NLTK | Tensorflow |
| Computational Linguistics | Ipsoft Amelia | Object Recognition | Tensorflow2.0 |
| Computer Vision | Iris-Recognition | Object Tracking | Text Mining |
| Computer-Vision | Ithink | Object-Detection | Text To Speech |
| Conv-Neural-Network | Keras | Object-Detection-Api | Text-Classification |
| Copilot | Language-Detection | Opencv | Text-Extraction |
| Dall-E | Languages Modeler | Opencv3.0 | Text-To-Speech |
| Decision Trees | Large-Language-Model | Opennlp | Tf.Keras |
| Deep Learning | Latent Dirichlet Allocation | Opinion Mining | Tokenization |
| Deep-Learning | Latent Semantic Analysis | Pattern-Recognition | Topic-Modeling |
| Deeplearning4J | Lexalytics | Progen | Torch |
| Dialogflow-Es | Lexical Acquisition | Pybrain | Transformer |
| Distinguo | Lexical Semantics | Python-Imaging-Library | TTS |
| Edge-Detection | Libsvm | Pytorch | Unsupervised Learning |
| Emgucv | Llama | Random-Forest | Virtual Agents |
| Face-Api | LSTM | R-Caret | Visual-Recognition |
| Face-Detection | Machine Learning | Recommender Systems | Voice-Recognition |
| Face-Recognition | Machine Translation | Recurrent-Neural-Network | Vowpal |
| Facial-Identification | Machine Vision | Reinforcement-Learning | Wabbit |
| Feature-Extraction | Machine-Learning | Roberta-Language-Model | Word2Vec |
| Feature-Selection | Machine-Translation | Scikit-Image | Word-Embedding |
| Form-Recognizer | Madlib | Scikit-Learn | Xgboost |

Note: This list has been built on the keywords from Alekseeva et al. (2021) and keywords found in the computer science and technology literature and AI-specialized websites.

Figure A1: Yearly number of AI-related questions asked on Stack overflow

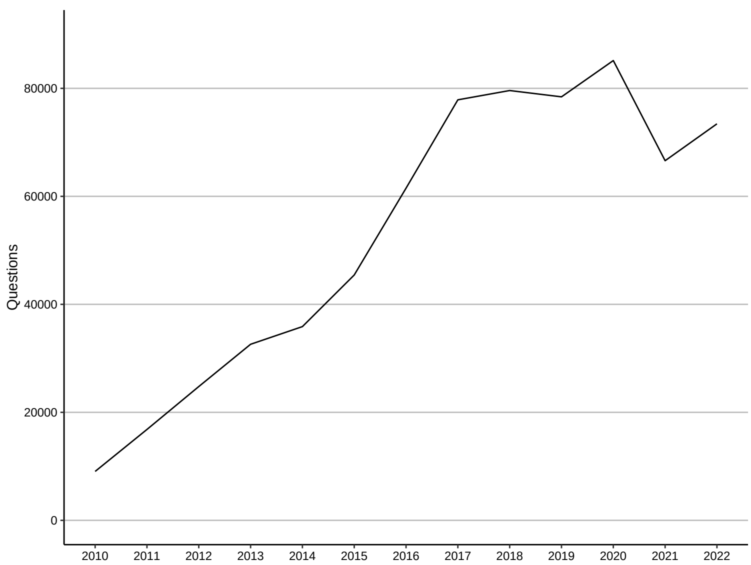
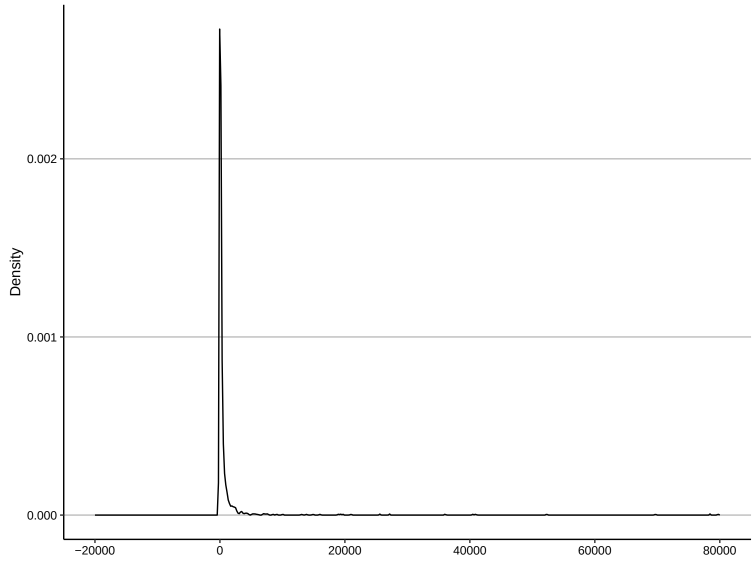


Figure A2: Distribution of questions per AI-related tag




## A.3 Identifying users location

Identifying the geolocation of the users is possible thanks to the information provided in the users' profiles. When one registers to Stack Overflow, she can fill in some information about herself, including the place of living, a personal description, and a website. The information in the user profile is filled in voluntarily. However, it is in the interest of the developers to provide these fields since Stack Overflow has become an important place for hiring developers.

I proceed in three steps to retrieve the location of the users. First, I use Google Maps to find the country corresponding to the place of living mentioned by the users. Second, when the place of living is missing, I ask ChatGPT 3.5 Turbo to extract any information in the personal description that could be used to geolocate the users.[22] Then, I pass this information to Google Maps and retrieve the country of living. Third, I use the personal website domain to geolocate users when the place of living and the personal description are missing.

With this method, 35.2% of the users are geolocated, which counts for 38.0% of the AI-related questions asked worldwide.

## A.4 Job content

The content of the occupations comes from O*NET 15.0, which was released in 2010. O*NET is a database containing standardized descriptors for the occupations composing the entire US economy (Peterson et al., 2001). O*NET has been widely used in the literature to measure the task content of occupations (see, for instance, Acemoglu and

[22] I use the following prompt: *I have some text describing people. I want to know where they live. Please return information found in the text about where they live. Return "none" if you do not find information.*

Autor, 2011; Blinder, 2009; Brynjolfsson et al., 2018; Felten et al., 2021; Firpo et al., 2011; Peri and Sparber, 2009).

In O*NET, every occupation comprises a different mix of knowledge, skills, and abilities and performs various activities and tasks. O*NET is continually updated to keep the occupations' descriptors aligned with the evolution of the labor market. The update system is based on ratings from responses by sampled workers who fill out O*NET questionnaires.[23]

There is no guidance in the literature regarding the descriptors of the occupations that should be used to measure AI exposure. While Webb (2020) relies on the description of more than 18 000 tasks performed within occupations, Felten et al. (2018, 2021) use 52 abilities required to perform occupations. In contrast, Brynjolfsson et al. (2018) and Eloundou et al. (2023) take advantage of 2069 detailed work activities, which are merged to tasks performed in occupations.

In this paper, I follow Felten et al. (2019, 2021) and use the abilities to describe the content of occupations. This choice is motivated by the terminology used to describe AI-related tags in Stack Overflow. They are more frequently defined by their abilities than by the precise tasks they can perform. For instance, AI-related tags in Stack Overflow include "Speech Recognition", "Pattern Recognition", and "Image Recognition", which appear closer to the notion of abilities used in O*NET than tasks or detailed work activities.

O*NET 15.0 describes 855 occupations through the lens of 52 abilities required to perform occupations. For each occupation, O*NET provides an *Importance*[24] and *Level*[25] scales to measure to what extent an ability is required for this occupation. The abilities encompass cognitive abilities (e.g., Oral Comprehension and Deductive Reasoning), psychomotor abilities (e.g., Finger Dexterity and Reaction Time), physical abilities (e.g., Static Strength and Dynamic Flexibility), and sensory abilities (e.g., Near Vision and Auditory Attention).

## A.5 Augmentation AI exposure

I start building the index that links each description of AI-related tags with the micro-industry and micro-occupational titles from CAI. I populated the transition matrices with cosine similarity measures following the same procedure as for automation AI exposure. I create sentences embenddings using

I derive the index of augmentation AI exposure for micro-industries by applying the following equation:

[23]See Handel (2016) for a detailed description.

[24]The scale importance indicates the degree of Importance a particular descriptor is to the occupation. The possible ratings range from "Not Important" (1) to "Extremely Important" (5).

[25]The scale Level indicates the degree, or point along a continuum, to which a particular descriptor is required or needed to perform the occupation.

$$I_{it} = \frac{\sum_{2010}^{t} \sum_{g=1}^{1182} ST_{gt} * C_{ig}}{1182} \tag{7}$$

Where $I_{it}$ gives the augmentation AI exposure for micro-industry $i$ in year $t$. $ST_{gt}$ is the yearly-tags scores and comes from equation (2). $C_{ig}$ is the cosine similarity measure between micro-industry $i$ and tag $g$.

Similarly, the index of augmentation AI exposure for micro-occupations is given by:

$$O_{ot} = \frac{\sum_{2010}^{t} \sum_{g=1}^{1182} ST_{gt} * C_{og}}{1182} \tag{8}$$

Here, $O_{ot}$ is the exposure to AI that complements micro-occupation $o$ and $C_{og}$ is the cosine similarity measure between micro-occupation $i$ and tag $g$.

Finally, I compute the simple mean at 6-digit 2018 SOC for the augmentation AI exposure for micro-occupations and at 4-digit 2022 NAICS for the augmentation AI exposure for micro-industries, and I take the average of both scores as follows:[26]

$$AI\_augm_{oit} = \frac{I_{it} + O_{ot}}{2} \tag{9}$$

The index of augmentation AI exposure at the occupational*industry level is given by $AI\_augm_{oit}$, where $I_{it}$ is given by equation (7) and $O_{ot}$ by equation (8).

# B First Stage IV estimators

Table C1: First Stage IV estimators for New Work

| | Auto AI | Auto AI | Augm AI | Augm AI | Auto AI | Auto AI | Augm AI | Augm AI |
|---|---|---|---|---|---|---|---|---|
| Automation AI IV | 1.025*** | 1.026*** | | | 1.026*** | 1.026*** | −0.105*** | −0.066*** |
| | (0.001) | (0.001) | | | (0.001) | (0.001) | (0.020) | (0.013) |
| Augmentation AI IV | | | 0.999*** | 0.995*** | −0.001*** | −0.001*** | 1.003*** | 0.996*** |
| | | | (0.005) | (0.003) | (0.000) | (0.000) | (0.005) | (0.003) |
| $R^2$ | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 |
| Observation | 141 664 | 141 664 | 141 664 | 141 664 | 141 664 | 141 664 | 141 664 | 141 664 |
| Employment (log) | X | X | X | X | X | X | X | X |
| FE Year | X | X | X | X | X | X | X | X |
| FE NAICS*SOC | X | X | X | X | X | X | X | X |
| FE NAICS*year (3-digit) | | X | | X | | x | | X |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 | 244 | 244 |

Note: Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. The estimations are weighted by employment size. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

[26] I use crosswalks provided by the US Census Bureau.

Table C2: First Stage IV estimators for employment

| | Auto AI | Auto AI | Augm AI | Augm AI | Auto AI | Auto AI | Augm AI | Augm AI |
|---|---|---|---|---|---|---|---|---|
| Automation AI IV | 1.060*** | 1.061*** | | | 1.061*** | 1.061*** | −0.096*** | −0.057*** |
| | (0.001) | (0.001) | | | (0.001) | (0.001) | (0.019) | (0.013) |
| Augmentation AI IV | | | 1.011*** | 1.008*** | −0.001*** | −0.001*** | 1.016*** | 1.010*** |
| | | | (0.004) | (0.002) | (0.000) | (0.000) | (0.004) | (0.002) |
| $R^2$ | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 |
| Observation | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 |
| *Fixed effects:* | | | | | | | | |
| FE Year | X | X | X | X | X | X | X | X |
| FE NAICS*SOC | X | X | X | X | X | X | X | X |
| FE NAICS*year (3-digit) | | X | | X | | x | | X |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 | 244 | 244 |

Note: Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. The estimations are weighted by employment size. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

Table C3: First Stage IV estimators for wages

| | Auto AI | Auto AI | Augm AI | Augm AI | Auto AI | Augm AI | Auto AI | Augm AI |
|---|---|---|---|---|---|---|---|---|
| Automation AI IV | 1.060*** | 1.061*** | | | 1.061*** | −0.096*** | 1.061*** | −0.057*** |
| | (0.001) | (0.001) | | | (0.001) | (0.019) | (0.001) | (0.013) |
| Augmentation AI IV | | | 1.011*** | 1.008*** | −0.001*** | 1.016*** | −0.001*** | 1.010*** |
| | | | (0.004) | (0.002) | (0.000) | (0.004) | (0.000) | (0.002) |
| $R^2$ | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 |
| Observation | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 |
| *Covariates included:* | | | | | | | | |
| Employment (log) | X | X | X | X | X | X | X | X |
| *Fixed effects:* | | | | | | | | |
| Year | X | X | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | | X | X |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 | 244 | 244 |

Note: Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. The estimations are weighted by employment size. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

# C Placebo AI exposures

Table D1: Effect of AI exposure on share of new work using placebo AI exposures

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Placebo Automation AI | −0.074 | 0.016 | | | −0.068 | 0.012 |
| | (0.106) | (0.087) | | | (0.107) | (0.086) |
| Placebo Augmentation AI | | | −0.040 | −0.030 | −0.038 | −0.030 |
| | | | (0.034) | (0.043) | (0.034) | (0.043) |
| $R^2$ | 0.672 | 0.689 | 0.673 | 0.690 | 0.673 | 0.690 |
| *Covariates included:* | | | | | | |
| Employment (log) | X | X | X | X | X | X |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 141 664 | 141 664 | 141 664 | 141 664 | 141 664 | 141 664 |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| SD outcome | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |

Note: The table presents the outputs for the regressions 5 using fixed effect estimators and when the dependent variable is the cumulative share of new work. Augmentation AI and automation AI exposures are placebo measures.The placebo measure for automation AI exposure is constructed by randomly shuffling the sentence similarity scores between AI-related tags and abilities. Similar procedure is followed to build placebo augmentation AI exposure for the sentence similarity scores between AI-related tags and micro-titles for occupations and industries. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. The estimations are weighted by employment size. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

Table D2: Effect of AI exposure on share of employment using placebo AI exposures

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Placebo Automation AI | −1.330 | −1.446 | | | −1.333 | −1.465 |
| | (0.828) | (0.956) | | | (0.835) | (0.937) |
| Placebo Augmentation AI | | | −0.005 | −0.077 | 0.019 | −0.104 |
| | | | (0.247) | (0.181) | (0.249) | (0.182) |
| $R^2$ | 0.993 | 0.994 | 0.993 | 0.994 | 0.993 | 0.994 |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 10.9 | 10.9 | 10.9 | 10.9 | 10.9 | 10.9 |
| SD outcome | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 |

Note: The table presents the outputs for the regressions 5 using fixed effect estimators and when the dependent variable is the employment size (in log). Augmentation AI and automation AI exposures are placebo measures.The placebo measure for automation AI exposure is constructed by randomly shuffling the sentence similarity scores between AI-related tags and abilities. Similar procedure is followed to build placebo augmentation AI exposure for the sentence similarity scores between AI-related tags and micro-titles for occupations and industries. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. The estimations are weighted by employment size. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

Table D3: Effect of AI exposure on wages using placebo AI exposures

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Placebo Automation AI | 1.331*** | 0.889*** | | | 1.318*** | 0.901*** |
| | (0.300) | (0.195) | | | (0.285) | (0.188) |
| Placebo Augmentation AI | | | 0.101 | 0.049 | 0.077 | 0.066 |
| | | | (0.083) | (0.049) | (0.069) | (0.047) |
| $R^2$ | 0.994 | 0.996 | 0.994 | 0.995 | 0.994 | 0.996 |
| *Covariates included:* | | | | | | |
| Employment (log) | X | X | X | X | X | X |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 3.2 | 3.2 | 3.2 | 3.2 | 3.2 | 3.2 |
| SD outcome | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |

Note: The table presents the outputs for the regressions 5 using fixed effect estimators and when the dependent variable is the hourly wages (in log). Augmentation AI and automation AI exposures are placebo measures.The placebo measure for automation AI exposure is constructed by randomly shuffling the sentence similarity scores between AI-related tags and abilities. Similar procedure is followed to build placebo augmentation AI exposure for the sentence similarity scores between AI-related tags and micro-titles for occupations and industries. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. The estimations are weighted by employment size. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

# D Full sample

Table E1: Effect of AI exposure on share of new work using the full sample

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Automation AI | 0.022 | 0.005 | | | 0.003 | 0.000 |
| | (0.029) | (0.025) | | | (0.025) | (0.022) |
| Augmentation AI | | | 0.031*** | 0.040*** | 0.031*** | 0.040*** |
| | | | (0.008) | (0.009) | (0.008) | (0.009) |
| $R^2$ | 0.655 | 0.672 | 0.672 | 0.689 | 0.672 | 0.689 |
| *Covariates included:* | | | | | | |
| Employment (log) | X | X | X | X | X | X |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 272 300 | 272 300 | 272 300 | 272 300 | 272 300 | 272 300 |
| Unique SOC (6-digit) | 751 | 751 | 751 | 751 | 751 | 751 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| SD outcome | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |

Note: The table presents the outputs for the regressions (5) using fixed effect estimators and when the dependent variable is the cumulative share of new work. The database is composed of all observations. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. The estimations are weighted by employment size. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

Table E2: Effect of AI exposure on employment in log using the full sample

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Automation AI | 0.079 | 0.219* | | | 0.016 | 0.199* |
| | (0.118) | (0.115) | | | (0.105) | (0.111) |
| Augmentation AI | | | 0.091*** | 0.093*** | 0.090*** | 0.090*** |
| | | | (0.024) | (0.031) | (0.024) | (0.031) |
| $R^2$ | 0.992 | 0.993 | 0.992 | 0.993 | 0.992 | 0.993 |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 442 935 | 442 935 | 442 935 | 442 935 | 442 935 | 442 935 |
| Unique SOC (6-digit) | 752 | 752 | 752 | 752 | 752 | 752 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 10.8 | 10.8 | 10.8 | 10.8 | 10.8 | 10.8 |
| SD outcome | 2.2 | 2.2 | 2.2 | 2.2 | 2.2 | 2.2 |

Note: The table presents the outputs for the regressions 6 using fixed effect estimators and when the dependent variable is the employment size. The database is composed of all observations. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. The estimations are weighted by employment size. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

Table E3: Effect of AI exposure on hourly wages in log using unbalanced database

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Automation AI | −0.197*** | −0.186*** | | | −0.189*** | −0.187*** |
| | (0.042) | (0.030) | | | (0.039) | (0.031) |
| Augmentation AI | | | −0.019** | 0.002 | −0.010 | 0.005 |
| | | | (0.009) | (0.007) | (0.008) | (0.007) |
| $R^2$ | 0.993 | 0.994 | 0.992 | 0.994 | 0.993 | 0.994 |
| *Covariates included:* | | | | | | |
| Employment (log) | X | X | X | X | X | X |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 431 467 | 431 467 | 431 467 | 431 467 | 431 467 | 431 467 |
| Unique SOC (6-digit) | 695 | 695 | 695 | 695 | 695 | 695 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 3.2 | 3.2 | 3.2 | 3.2 | 3.2 | 3.2 |
| SD outcome | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |

Note: The table presents the outputs for the regressions 6 using fixed effect estimators and when the dependent variable is the hourly wages. The database is composed of all observations. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. The estimations are weighted by employment size. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

# E Estimations with constant weights

Table F1: Effect of AI exposure on share of new work with constant weights (employment size in 2015)

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Automation AI | 0.023 | 0.004 | | | 0.006 | $-0.001$ |
| | (0.028) | (0.023) | | | (0.025) | (0.022) |
| Augmentation AI | | | 0.025*** | 0.034*** | 0.025*** | 0.034*** |
| | | | (0.007) | (0.008) | (0.007) | (0.008) |
| $R^2$ | 0.667 | 0.683 | 0.681 | 0.698 | 0.681 | 0.698 |
| *Covariates included:* | | | | | | |
| Employment (log) | X | X | X | X | X | X |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 141 664 | 141 664 | 141 664 | 141 664 | 141 664 | 141 664 |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| SD outcome | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |

Note: The table presents the outputs for the regressions 5 using fixed effect estimators and when the dependent variable is the cumulative share of new work. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.

Table F2: Effect of AI exposure on employment size in log with constant weights

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Automation AI | 0.055 | 0.170 | | | 0.002 | 0.153 |
| | (0.118) | (0.119) | | | (0.110) | (0.115) |
| Augmentation AI | | | 0.069** | 0.071** | 0.069** | 0.068* |
| | | | (0.027) | (0.036) | (0.027) | (0.035) |
| $R^2$ | 0.991 | 0.992 | 0.991 | 0.992 | 0.991 | 0.992 |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 10.8 | 10.8 | 10.8 | 10.8 | 10.8 | 10.8 |
| SD outcome | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 |

Note: The table presents the outputs for the regressions 6 using fixed effect estimators and when the dependent variable is the employment size. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. Standard errors are reported in brackets and are clustered at the occupational and industry level. $\star\star\star$ Significant at the 1% level; $\star\star$ significant at the 5% level; $\star$ significant at the 10% level.

Table F3: Effect of AI exposure on hourly wages in log with constant weights

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Automation AI | −0.181*** | −0.159*** | | | −0.171*** | −0.160*** |
| | (0.042) | (0.026) | | | (0.038) | (0.026) |
| Augmentation AI | | | −0.021** | 0.001 | −0.013* | 0.003 |
| | | | (0.009) | (0.005) | (0.008) | (0.005) |
| $R^2$ | 0.994 | 0.995 | 0.994 | 0.995 | 0.994 | 0.995 |
| *Covariates included:* | | | | | | |
| Employment (log) | X | X | X | X | X | X |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 | 230 204 |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 3.2 | 3.2 | 3.2 | 3.2 | 3.2 | 3.2 |
| SD outcome | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |

Note: The table presents the outputs for the regressions 6 using fixed effect estimators and when the dependent variable is the hourly wages. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. Standard errors are reported in brackets and are clustered at the occupational and industry level. $\star\star\star$ Significant at the 1% level; $\star\star$ significant at the 5% level; $\star$ significant at the 10% level.

# F Median hourly wages

Table G1: Effect of AI exposure on median hourly wages in log

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Automation AI | −0.229*** | −0.216*** | | | −0.213*** | −0.216*** |
| | (0.042) | (0.032) | | | (0.039) | (0.032) |
| Augmentation AI | | | −0.029*** | −0.003 | −0.019** | 0.001 |
| | | | (0.010) | (0.007) | (0.009) | (0.007) |
| $R^2$ | 0.991 | 0.994 | 0.991 | 0.993 | 0.992 | 0.994 |
| *Covariates included:* | | | | | | |
| Employment (log) | X | X | X | X | X | X |
| *Fixed effects:* | | | | | | |
| Year | X | X | X | X | X | X |
| NAICS*SOC | X | X | X | X | X | X |
| NAICS*year (3-digit) | | X | | X | | X |
| Observations | 229 296 | 229 296 | 229 296 | 229 296 | 229 296 | 229 296 |
| Unique SOC (6-digit) | 672 | 672 | 672 | 672 | 672 | 672 |
| Unique NAICS (4-digit) | 244 | 244 | 244 | 244 | 244 | 244 |
| Mean outcome | 3.1 | 3.1 | 3.1 | 3.1 | 3.1 | 3.1 |
| SD outcome | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |

Note: The table presents the outputs for the regressions (6) using fixed effect estimators and the median hourly wages as dependent variable. The estimations are weighted by employment size. Augmentation AI exposure and automation AI exposure are standardized to have a mean of 0 and a standard deviation equal to 1. Standard errors are reported in brackets and are clustered at the occupational and industry level. ⋆⋆⋆ Significant at the 1% level; ⋆⋆ significant at the 5% level; ⋆ significant at the 10% level.