WORKING PAPER Nº IDB-WP-1624

"Mirror, Mirror on the Wall: Which Jobs Will AI Replace After All?":

A New Index of Occupational Exposure

Miguel Benítez-Rueda
Eric Parrado

Inter-American Development Bank
Department of Research and Chief Economist

August 2024

“Mirror, Mirror on the Wall: Which Jobs Will AI Replace After All?”:

A New Index of Occupational Exposure

Miguel Benítez-Rueda*
Eric Parrado**

* University of California, Davis
** Inter-American Development Bank

Inter-American Development Bank
Department of Research and Chief Economist

August 2024

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Benitez-Rueda, Miguel.
"Mirror, mirror on the wall: which jobs will AI replace after all?": a new index of occupational exposure / Miguel Benitez-Rueda, Eric Parrado.
p. cm. — (IDB Working Paper Series ; 1624)
Includes bibliographical references.
1. Artificial intelligence-Social aspects. 2. Artificial intelligence-Economic aspects. 3. Job creation. 4. Occupational mobility. 5. Labor supply. 6. Labor market. 7. Training. 8. Employees-Effect of technological innovations on. I. Duenas, Juliana. II. Parrado, Eric. III. Inter-American Development Bank. Department of Research and Chief Economist. IV. Title. V. Series.
IDB-WP-1624

http://www.iadb.org

ABSTRACT[1]

This paper introduces the AI Generated Index of Occupational Exposure (GENOE), a novel measure quantifying the potential impact of artificial intelligence on occupations and their associated tasks. Our methodology employs synthetic AI surveys, leveraging large language models to conduct expert-like assessments. This approach allows for a more holistic evaluation of job replacement likelihood, reducing assumptions about the mechanisms through which AI innovations could replace job tasks and skills. Our findings reveal that the average occupational exposure to AI is 0.28 within one year, increasing to 0.38 and 0.44 over the five- and 10-year horizons, respectively. We also show evidence that the index not only considers task automation, but also contextual factors such as social and ethical considerations and regulatory constraints that may affect the likelihood of replacement. After calibrating the index with labor market microdata from the United States and Mexico, we show that approximately 43 million and 16 million jobs, respectively, are highly exposed to AI within a one-year time horizon. The GENOE index provides valuable insights for policymakers, employers, and workers, offering a data-driven foundation for strategic workforce planning and adaptation in the face of rapid technological change.

JEL classifications: C53, C81, J23, J24, O33
Keywords: Artificial intelligence (AI), Labor markets, Job displacement, Automation, Occupational exposure, Tasks, Synthetic AI surveys, AI Generated Index (GENOE)

[1]We thank Carlos Scartascini for his valuable comments, and Carlo Pizzinelli and his co-authors for sharing with us their complementarity index and C-AIOE estimates.

The concern that more autonomous instruments and machines would render work obsolete has recurred throughout history. From Aristotle speculating on the obsolescence of human labor if instruments could work on their own to Keynes' famous prediction of technological unemployment,[2] this concern arises whenever new technologies emerge and has particularly grown with recent advances in artificial intelligence (AI). Growing progress in natural language processing and generative models like OpenAI's GPT have significantly improved machines' abilities to mimic many human cognitive capabilities, such as understanding and generating human-like text, making predictions, and performing image and video recognition (Dwivedi *et al.*, 2023). Furthermore, the convergence of AI with robotics signals the emergence of highly autonomous systems capable of executing more complex physical tasks, potentially reshaping the landscape of both cognitive and manual labor.

This paper aims to identify which occupations are most exposed to the replacement due to AI adoption, developing a new Generated Index of Occupational Exposure (GENOE). Our approach uses a novel methodological approach that we call "synthetic AI surveys." This methodology leverages the extensive knowledge and advanced analytical capabilities of large language models (LLMs) to conduct expert-like assessments, provide estimations, and offer insights on a wide range of areas. We specifically applied this approach to estimate the exposure of various tasks and occupations to recent and anticipated advancements in AI-related technologies. The GENOE index measures the likelihood of job replacement by AI at the occupational level, based on the tasks typically performed in each role. We defined three time horizons for replacement: one year, five years, and 10 years. This accounts for short-term scenarios based on current AI capabilities as well as medium-term and long-term scenarios anticipating future technological advancements.

Our findings reveal that the average occupational exposure to AI is 0.28 within one year, increasing to 0.38 and 0.44 over the five- and 10-year horizons, respectively.[3] An exposure score of 0.28 suggests that, on average, occupations have a 28% likelihood of potentially being impacted by AI within the next year. To contextualize these results, we calibrated the GENOE index with labor market data from the United States and Mexico. Results show that 43 million and 16 million jobs are exposed to AI in the United States and Mexico, respectively, over the one-year horizon, increasing to 60 million and 22 million over five years, and 70 million and 26 million over ten years. Extrapolating these findings to Latin America and the Caribbean, we estimate that approximately 84 million jobs are exposed to AI within one year, rising to 114 million in 5 years, and 132 million in 10 years. Assuming this average exposure is applicable to worldwide labor markets, it implies that approximately 980 million, 1.33 billion, and 1.54 billion jobs are exposed to AI over the one-, five-, and 10-year horizons, respectively. It is crucial to note that this does not directly translate

[2] See Aristotle (350 B.C.E./1998), Politics and Keynes (1930).

[3] GENOE estimates are available online for use by academics, policymakers, and businesses to inform future research and strategic decision-making GENOE. Data can be found in this online repository.

to job losses or complete occupational obsolescence. Rather, it indicates the proportion of occupations that may undergo substantial changes due to AI integration. The impact on employment will likely be more complex, involving occupational transformations, the emergence of new roles, and potential displacement in some sectors.

Our analysis shows that in the United States and Mexico, women are more vulnerable to AI displacement, partly due to their significant presence in office and administrative jobs. In the United States, low-educated and low-wage workers are particularly exposed, while in Mexico, both low- and medium-skilled workers, especially those in middle-income formal jobs, are more vulnerable, whereas informal jobs remain less exposed. This suggests that AI adoption could exacerbate income inequalities by disproportionately impacting lower- and middle-income workers, echoing, to some extent, findings from previous technological advancements (Autor *et al.*, 2006; Autor and Dorn, 2013; Autor, 2019; Acemoglu and Restrepo, 2022).

To better understand the rationale behind the synthetic AI survey assessments, we conducted follow-up surveys, asking the model to explain the considerations behind each occupation's estimation. These responses revealed that the model's rationale focuses on several key factors: current and expected technological capabilities and limitations, the routine and non-routine nature of tasks, ethical considerations, the regulatory environment, and cost-benefit analyses. The models evaluated AI's current abilities and future potential, emphasized the likelihood of replacing routine tasks, and frequently mentioned ethical considerations and safety concerns, particularly for roles with direct human interaction. Regulatory factors were noted for industries with strict standards, and cost-benefit analyses influenced the economic feasibility of automation.

We conducted several alternative exercises to examine how the GENOE index changes with different methodologies and prompts. These variations included omitting occupation names, excluding ethical considerations, assessing tasks individually, and concentrating on specific AI technologies. Most notably, when we instructed the model to disregard ethical, regulatory, and social considerations and focus purely on AI's technical abilities, we observed a significant increase in the estimated exposure across occupations. This alternative GENOE index, which we term GENOE-Tech, provides an estimation of AI's potential impact based solely on its growing technical capabilities.

This paper makes several contributions to the literature on AI's impact on employment. First, it overcomes some limitations of the existing indexes that measure occupational exposure to AI (Brynjolfsson *et al.*, 2018; Webb, 2019; Felten *et al.*, 2021; Tolan *et al.*, 2021; Pizzinelli *et al.*, 2023; Eloundou *et al.*, 2024). Most existing indexes disaggregate occupations into collections of separate tasks, skills, or abilities. The first restriction of this approach is that it overlooks more complex interrelations between tasks or abilities. For instance, journalists and technical writers share several automatable routine tasks related to writing, editing, and generating content. Yet, routine roles in journalism often interact with non-routine roles like creating engaging narratives and storytelling, evaluating the significance of news, and engaging in investigative reporting, all of which require human judgment, critical thinking, and persuasion. This interrelation will result in a lower AI exposure of journalism routine tasks than those of technical writers.

Another limitation of this approach is that it does not account for the ethical and social context of different occupations performing similar types of tasks. Consider, for example, judges and credit analysts. Both make decisions based on evaluating extensive information, and both roles can significantly impact individuals' lives. However, the nature and scope of their impact differ in important ways. Judges deliver verdicts that can dramatically affect individuals' well-being, freedom, rights, and long-term societal standing. Their role demands a profound understanding of ethical principles and the ability to balance complex cases. The ethical weight and social impact of their decisions, in addition to requiring human judgment and empathy, make it unlikely that societies would accept these roles being legitimately replaced by AI. Credit analysts, while also impacting individual well-being through their decisions on loan approvals or credit ratings, typically operate within a more narrowly defined financial context. Their decisions can certainly affect a person's economic opportunities and quality of life, but generally do not carry the same weight of personal liberty or broader societal impact as judicial decisions. This comparison highlights the importance of considering not just the technical similarities in task types, but also the broader ethical, social, and institutional contexts in which different occupations operate when assessing the potential for AI integration or replacement.

Our approach corrects the limitations of previous indexes by integrating a more holistic task-based analysis that captures the interrelations and complementarities between tasks within occupations. Unlike previous methods that disaggregate occupations into isolated tasks, skills, or abilities, our AI-generated index considers the entire vector of tasks, reflecting a complete view of occupational contributions in production. To address the second limitation, we provide the model in the surveys with both the name of the occupation and its vector of tasks, allowing it to estimate the overall likelihood of AI replacement without imposing restrictive assumptions. This method allows the model to consider the ethical and social context of each occupation organically rather than through predefined criteria, offering a more holistic evaluation of exposure and replacement. Alternative prompt specifications and follow-up surveys suggest that our method effectively accounts for these considerations.

This second concern is also raised by Pizzinelli *et al.* (2023) and is somewhat addressed by their complementarity-adjusted occupational exposure index (C-AIOE). This measure incorporates factors such as work contexts and job zones to account for the likelihood that AI will complement rather than replace human labor.[4] Our approach goes further by imposing fewer assumptions

[4]Notwithstanding its advancements over previous indexes, the work contexts and job zones retrieved from the O*NET contain several limitations and assumptions. Firstly, it assumes that higher levels of education and training automatically correlate with better integration of AI, which may not account for the specific skill sets and practical contexts required in each occupation. Additionally, the chosen contexts, such as face-to-face communication, responsibility for others' health, and physical proximity, might not fully capture the ethical reasoning required in certain roles, nor the degree of public resistance to AI replacing certain decision-making roles. Furthermore, as the index is built upon the AIOE, it relies solely on the isolated similarity of abilities to AI developments and thus does not account for the exposure of tasks and their interrelations in the workplace.

regarding the specific skills and contexts required for AI integration. Additionally, the GENOE index considers the complex interplay of tasks within occupations and integrating the ethical and social contexts organically.

Another contribution of our paper is the novel use of large language models to generate synthetic assessments that aim to replicate and extend expert assessments. Many studies rely on expert surveys or interviews to build analytical tools like indexes, including in the labor automation area (Frey and Osborne, 2017; Felten *et al.*, 2021; Tolan *et al.*, 2021). Synthetic AI surveys offer several advantages. They can efficiently process and synthesize vast amounts of information from diverse sources very rapidly, enabling a larger volume of questions to be addressed in a shorter time frame. Additionally, synthetic AI surveys can be easily scaled to include extensive sample sizes or repeated iterations, allowing greater granularity of the collected data.

While recent research has begun to explore the use of AI in workforce analysis, our methodology offers a novel approach in this evolving field. For example, Eloundou *et al.* (2024) combined human expert insights with AI-generated assessments to evaluate AI's potential for enhancing task efficiency across occupations. Our work is a step forward in this approach by exclusively and systematically using ChatGPT to build the entire index through multiple iterations and varying model parameters.

To validate the accuracy of this methodology, we replicated the well-known exposure index of Felten *et al.* (2021), replacing mTurk surveys with synthetic AI surveys. Our approach successfully replicated Felten's AI Occupational Exposure Index (AIOE) with a correlation of 0.99, suggesting our methodology is robust and reliable for these types of tasks and providing evidence of another area where AI-driven automation is highly feasible.

The remainder of this paper is organized as follows. After this introduction, Section 2 details our data sources and methodology used to build the GENOE index. Section 3 presents the GENOE index, including its key findings and implications. This section also includes several robustness checks to validate our results and replicates other human-like expert assessment using our novel synthetic survey methodology. Section 4 explores how the index can be used in practical applications, analyzing occupational exposure in the U.S. and Mexico labor market. Finally, Section 5 discusses some considerations and limitations of our research, while Section 6 concludes with implications and directions for future research.

2 DATA AND METHODOLOGY

We draw data from the O*NET database, which contains detailed descriptions and attributes for 1,016 occupations in the U.S. labor market under the O*NET-SOC 2019 classification. Occupational characterizations are built and continuously updated from surveys and input from multiple sources, including job analysts, incumbent workers, and industry experts. O*NET attributes—such as abilities, skills, knowledge, and tasks—are divided into six major domains: Worker Characteristics, Worker Requirements, Experience Requirements, Occupational Requirements, Workforce Characteristics, and Occupation-Specific Information.

Using O*NET's task module, which contains 18,156 specific tasks, we describe each occupation as a vector of tasks typically performed by that role. We focus on tasks rather than skills or abilities, following the extensive "task approach" literature (Autor *et al.*, 2003, 2006; Acemoglu and Autor, 2011; Autor, 2013). This framework suggests that technological advancements displace or reinstate specific tasks within occupations rather than entire jobs or skills, as neither skills nor entire jobs directly translate into output (Autor, 2013).

To measure occupational exposure to AI, we employed a novel methodology using large language models (LLMs), specifically the GPT-4o API from OpenAI. Our approach, which we term "synthetic AI surveys," involves asking these advanced models to estimate the likelihood of AI and its developments replacing each occupation based on its characteristic tasks.

Our synthetic AI survey methodology offers several key advantages in assessing occupational exposure to AI. First, the scalability of AI models allows for rapid processing of vast amounts of information, enabling a comprehensive analysis of all occupations in the database. This scalability ensures that no occupation is overlooked, providing a complete picture of potential AI impact across the labor market. Second, the use of a standardized AI model guarantees consistency in evaluation criteria across all occupations, eliminating potential human biases and ensuring comparability of results. Third, large language models like GPT-4o are trained on recent data, potentially incorporating the latest understanding of AI capabilities and trends, which keeps our analysis current and forward-looking. Lastly, these advanced models offer a more holistic view of tasks, occupations and technological innovations, capable of capturing interactions between tasks and considering broader contextual factors in their assessments.

Key features of our survey methodology include the following:

Independent Sessions: Each survey was conducted in a new session for every occupation, preventing carryover effects from previous responses. This ensures that each occupation is evaluated independently, based solely on its own characteristics.

Multiple Iterations: The process was repeated 100 times to test the accuracy and consistency of the estimates. This multiple-iteration approach allows us to account for potential variability in the AI model's responses and provides a measure of confidence in our results.

Varied Creativity: The temperature parameter, which controls the randomness and creativity in the model's output, was randomly set between 0 and 2 for each iteration. This introduces controlled variability in the model's responses, potentially capturing a wider range of perspectives on each occupation's AI exposure.

Time Horizons: We considered three exposure horizons: one year, five years, and 10 years. This allows us to capture both immediate and longer-term potential impacts of AI on occupations, recognizing that AI capabilities and their integration into various fields may evolve over time.

The GENOE index for each horizon was calculated as the average value across the 100 surveys. To ensure compatibility with standard occupational classifications, we translated these estimations to the SOC-18 classification using O*NET's crosswalks for a total of 759 occupations.

To enhance our understanding of the GENOE's effectiveness and to validate our methodology, we conducted several additional exercises. Firstly, we measured task-level exposure by assessing the proportion of tasks at risk within each occupation, allowing us to identify roles where AI might significantly impact certain tasks without necessarily replacing the entire job. We also evaluated tasks in isolation and compared the results with our holistic occupation-level approach to understand the value of considering tasks within their broader occupational context. Additionally, we tested the effect of withholding the occupation name from the model to gauge the importance of occupational context in the AI's assessment. To incorporate broader societal considerations, we explicitly prompted the model to consider ethical and social factors in its assessments.

Furthermore, we conducted follow-up surveys, asking the AI to explain its reasoning behind the exposure estimates, effectively "opening the black box" of its decision-making process. These additional analyses serve multiple purposes: they validate our methodology by providing alternative perspectives on AI exposure, offer insights into the factors influencing AI's potential impact on various occupations, and help us understand the decision-making process of the AI model. This approach not only strengthens the robustness of our GENOE index but also potentially reveals any biases or limitations in the AI's assessments, providing a more reliable picture of occupational exposure to AI.

To further validate our approach, we applied our synthetic survey methodology to replicate the Felten *et al.* (2021) index of occupational exposure. In this exercise, we substituted the original human expert surveys with synthetic surveys conducted using ChatGPT4o. The primary objective was to assess the capability of our AI-driven methodology to reproduce results comparable to those obtained through traditional human expert assessments. This replication serves as a crucial test of our approach's reliability and accuracy, offering insights into the potential of large language models to emulate human expertise in evaluating occupational exposure to AI.

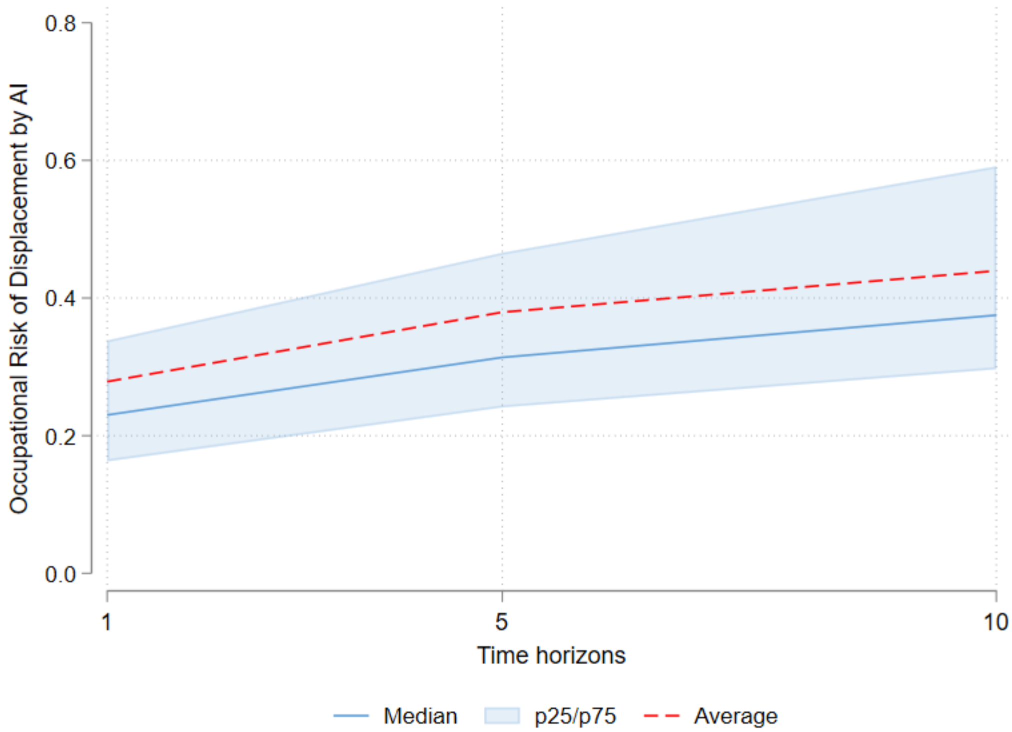
3 RESULTS

3.1 Baseline Estimations

Figure 1 presents a comprehensive view of AI exposure across 759 occupations, illustrating the average, median, 25th percentile, and 75th percentile for the three defined time horizons. The analysis reveals a progressive increase in average AI exposure: 0.28 within one year, rising to 0.38 over five years, and reaching 0.44 over a decade. This pattern suggests that as AI technologies advance and mature, their potential to reshape various professions grows significantly.[5]

[5] Complete descriptive statistics can be found in Appendix Table A.1.

Figure 1: AI-Generated Index of Occupational Exposure to AI (GENOE) over Time



Source: Authors' calculations using data from O*NET and surveys conducted with GPT-4o through OpenAI's API.
Note: This figure displays the GENOE for the next one, five, and 10 years. The index measures the likelihood that occupations will be displaced by AI within each time horizon.

Additionally, we analyze occupational exposure among the 23 major occupational groups. As depicted in Table 1, the most exposed groups are office and administrative support, production, and sales and related occupations. This result aligns with current trends in AI and machine learning, which have significantly improved the ability to automate tasks involving data entry, inventory management, and transaction processing. These roles often involve structured, predictable environments that are ideal for AI applications.

Table 1: **Average GENOE Score in 23 Major Occupational Groups**

Main occupations	1-year horizon	5-year horizon	10-year horizon
Management	0.21	0.28	0.33
Business and Financial Operations	0.35	0.45	0.52
Computer and Mathematical	0.29	0.38	0.45
Architecture and Engineering	0.24	0.33	0.38
Life, Physical, and Social Science	0.22	0.30	0.35
Community and Social Service	0.15	0.21	0.25
Legal	0.26	0.36	0.41
Educational Instruction and Library	0.16	0.22	0.26
Arts, Design, Entertainment, Sports, and Media	0.28	0.35	0.41
Healthcare Practitioners and Technical	0.15	0.24	0.30
Healthcare Support	0.21	0.31	0.36
Protective Service	0.21	0.29	0.36
Food Preparation and Serving Related	0.28	0.41	0.50
Building and Grounds Cleaning and Maintenance	0.22	0.33	0.40
Personal Care and Service	0.23	0.32	0.38
Sales and Related	0.42	0.51	0.57
Office and Administrative Support	0.60	0.70	0.75
Farming, Fishing, and Forestry	0.29	0.41	0.48
Construction and Extraction	0.18	0.28	0.35
Installation, Maintenance, and Repair	0.17	0.27	0.33
Production	0.44	0.59	0.65
Transportation and Material Moving	0.30	0.46	0.55

Source: Authors' calculations using data from O*NET and surveys conducted with GPT-4 through OpenAI's API. *Note:* This table displays the simple average GENOE score aggregated for the 23 major occupational groups according to the SOC 2018 classification.

However, there is significant heterogeneity in AI exposure across different occupations. To better illustrate this variation, Tables 2 and 3 presents the top 30 occupations with the highest AI exposure scores and the top 30 occupations with the lowest exposure scores, respectively. Roles such as firefighters, surgeons, priests, and athletes consistently show the lowest levels of AI exposure across all time horizons. These roles typically involve complex physical skills, high-stakes decision-making, or deeply human elements that current AI technologies struggle to replicate. Conversely, occupations like telemarketers, credit authorizers, word processors, and data entry keyers exhibit the highest levels of exposure, likely due to the routine and data-centric nature of their tasks.

Table 2: Top Exposed Occupations According to GENOE

	1-year horizon	5-year horizon	10-year horizon
Telephone Operators	0.92	0.93	0.95
Data Entry Keyers	0.90	0.90	0.91
Telemarketers	0.87	0.88	0.89
Word Processors and Typists	0.86	0.87	0.90
Switchboard Operators, Including Answering Service	0.84	0.85	0.86
Credit Authorizers, Checkers, and Clerks	0.83	0.84	0.85
Meter Readers, Utilities	0.81	0.85	0.86
Motion Picture Projectionists	0.81	0.83	0.86
Order Clerks	0.80	0.82	0.84
Shoe Machine Operators and Tenders	0.77	0.84	0.84
Cutting and Slicing Machine Setters, Operators, and Tenders	0.80	0.82	0.83
Textile Winding, Twisting, and Drawing Out Machine Setters, Operators, and Tenders	0.78	0.82	0.84
Machine Feeders and Offbearers	0.75	0.81	0.84
Billing and Posting Clerks	0.76	0.80	0.84
File Clerks	0.72	0.82	0.85
Statistical Assistants	0.76	0.80	0.83
Travel Agents	0.76	0.80	0.82
Graders and Sorters, Agricultural Products	0.76	0.79	0.83
Office Machine Operators, Except Computer	0.72	0.81	0.83
Correspondence Clerks	0.76	0.78	0.82
Coil Winders, Tapers, and Finishers	0.74	0.78	0.83
Medical Transcriptionists	0.73	0.79	0.82
Bookkeeping, Accounting, and Auditing Clerks	0.72	0.79	0.83
Payroll and Timekeeping Clerks	0.74	0.77	0.82
Mail Clerks and Mail Machine Operators, Except Postal Service	0.73	0.78	0.82
Gambling Change Persons and Booth Cashiers	0.72	0.77	0.84
Tellers	0.70	0.79	0.83
Packaging and Filling Machine Operators and Tenders	0.70	0.79	0.83
Desktop Publishers	0.73	0.79	0.80
Grinding, Lapping, Polishing, and Buffing Machine Tool Setters, Operators, and Tenders, Metal and Plastic	0.74	0.76	0.81

Source: Authors' calculations using data from O*NET and surveys conducted with GPT-4 through OpenAI's API. *Note:* This table shows the occupations most exposed to AI, along with the GENOE estimates for the next one, five, and 10 years. Occupations were ranked based on the average exposure value across the three horizons.

Table 3: Less Exposed Occupations According to GENOE

	1-year horizon	5-year horizon	10-year horizon
Athletes and Sports Competitors	0.05	0.09	0.11
Oral and Maxillofacial Surgeons	0.04	0.11	0.16
Midwives	0.05	0.11	0.16
Firefighters	0.06	0.11	0.16
Kindergarten Teachers, Except Special Education	0.07	0.12	0.15
Special Education Teachers, Secondary School	0.09	0.11	0.18
Dancers	0.09	0.12	0.18
Clergy	0.10	0.12	0.17
Nurse Midwives	0.07	0.13	0.19
Elementary School Teachers, Except Special Education	0.08	0.13	0.18
Special Education Teachers, Middle School	0.09	0.14	0.18
First-Line Supervisors of Police and Detectives	0.09	0.14	0.18
First-Line Supervisors of Firefighting and Prevention Workers	0.09	0.13	0.19
Massage Therapists	0.09	0.13	0.18
Special Education Teachers, Preschool	0.09	0.14	0.19
Dentists, General	0.08	0.14	0.21
Pediatricians, General	0.09	0.15	0.20
Helpers–Roofers	0.09	0.16	0.20
Commercial Divers	0.09	0.17	0.19
Clinical Nurse Specialists & Advanced Practice Psychiatric Nurses & Acute Care Nurses & Critical Care Nurses & Registered Nurses	0.09	0.16	0.20
Child, Family, and School Social Workers	0.10	0.16	0.19
Education Administrators, Kindergarten through Secondary	0.10	0.16	0.19
Nursing Instructors and Teachers, Postsecondary	0.09	0.17	0.20
Roofers	0.10	0.17	0.20
Middle School Teachers, Except Special and Career/Technical Education	0.10	0.16	0.21
Judges, Magistrate Judges, and Magistrates	0.08	0.17	0.21
Preschool Teachers, Except Special Education	0.10	0.17	0.20
Structural Iron and Steel Workers	0.09	0.17	0.21
Hairdressers, Hairstylists, and Cosmetologists	0.09	0.17	0.22
Music Therapists & Art Therapists	0.11	0.17	0.20

Source: Authors' calculations using data from O*NET and surveys conducted with GPT-4 through OpenAI's API. *Note:* This table shows the occupations less exposed to AI, along with the GENOE estimates for the next one, five, and 10 years. Occupations were ranked based on the average exposure value across the three horizons.

Occupations in the 25th percentile, with relatively low AI exposure, include medical scientists, stonemasons, and occupational therapy aides, suggesting that roles requiring specialized knowledge or physical dexterity remain relatively insulated from AI replacement in the near term. The 75th percentile encompasses occupations with moderately high exposure, such as purchasing agents and tutors, indicating that even some knowledge-based professions face significant AI impact. Interestingly, farmworkers and criminal investigators represent the median level of AI exposure, highlighting the broad reach of AI across both blue-collar and white-collar sectors.

Interestingly, our research reveals considerable heterogeneity in exposure within specific sectors, such as the healthcare industry. Pediatricians, neurologists, and general internal medicine physicians show low susceptibility to AI replacement, likely due to the high degree of human interaction, non-routine tasks, and complex decision-making required in these roles. In contrast, specialties like radiology, pharmacy, and dermatology face a higher potential for AI integration, which aligns with AI's growing capabilities in image analysis, data processing, and pattern recognition.

These findings underscore the importance of a granular, occupation-specific approach when assessing AI's potential impact on the workforce. They suggest that AI might reshape rather than wholly replace professions, with human expertise remaining crucial for complex cases and integrating specialized knowledge with broader contexts. This understanding is essential for effective workforce planning, educational curriculum development, and policy formulation, highlighting the need for targeted approaches to reskilling and upskilling initiatives that focus on cultivating distinctly human capabilities while leveraging AI's strengths.

Our analysis demonstrates high consistency and reliability in the GENOE estimates. As shown in Figure A.1 from the Appendix, the confidence intervals for the GENOE across the 100 iterations are remarkably narrow, indicating a high level of precision in the LLM's assessment of occupational exposure. Additionally, there is ordinal alignment across time horizons, with estimates for higher horizons being equal to or greater than those for lower horizons in almost all occupations. This pattern suggests a cumulative effect of AI exposure over time, with occupations generally becoming more susceptible to AI replacement as technology advances.

Importantly, our results show robustness to variations in the model's temperature setting. As observed in Table A.2 in the Appendix, the GENOE index remains unchanged whether it is calculated using iterations assigned higher temperatures or lower temperatures. This consistency across different levels of model creativity reinforces the reliability of our findings.

To address concerns about the "black box" nature of ChatGPT's assessments, we conducted follow-up surveys asking the model to explain its reasoning for each occupation's assessment in the five-year horizon. These responses allowed us to uncover patterns, ideas, and core topics influencing the models' decisions.

We identified at least five main patterns in LLM's rationale:

Technological Capabilities and Limitations. Many responses highlight what AI can currently achieve. For example, in the assessment of "File Clerks," it notes AI's current proficiency in data entry and document management, leading to a high likelihood of replacement (0.9). On the other hand, future advancements are often considered, especially regarding the merging of robotics and AI, a topic mentioned in the surveys 236 times—not only to see the potential, but also to cast doubts about the feasibility of AI replacement. For instance, in the case of "Barbers," ChatGPT notes that while advances in AI and robotics have made strides in automating various tasks, the significant manual dexterity, precision, and personal interaction involved in barbering are challenging to replicate with current or medium-term advances in technology.

Tasks Nature (Routine vs Non-Routine). The model consistently distinguishes between routine and non-routine tasks. Occupations with routine, repetitive tasks are considered more likely to be replaced by AI. The term ”routine” appears 72 times in the responses, indicating a strong emphasis on the predictability and repetitiveness of tasks. For example, the response for ”Anesthesiologists” (likelihood 0.2) notes that although AI might assist with some routine tasks such as diagnostic procedures, monitoring patients, and scheduling, the holistic and critical (non-routine) nature of anesthesiology, including patient care, emergency response, coordination with medical teams, and individualized decision-making, presents substantial challenges to full automation.

Ethical Considerations. The model frequently raises ethical concerns, particularly for occupations involving direct human interaction or safety-critical roles. The terms ”ethical” and “safety” appear 28 and 81 times, respectively, in the responses. For example, in the case of ”Judges,” ChatGPT notes that the complex and highly ethical nature of judicial tasks, such as interpreting laws and making nuanced judgments, poses significant challenges for AI, making it unlikely that judges will be replaced by AI soon (likelihood of 0.2). Similarly, in occupations like Pilots and Amusement and Recreation Attendants, safety concerns are raised to lower AI prospects of replacement.

Regulatory Environment. Legal and regulatory factors are also considered. Regulatory concerns were raised 63 times in the follow-up responses. This occurs in occupations like bakers, drivers, and civil engineers. Heavily regulated industries might be less likely due to strict regulations and standards that AI must meet.

Cost-Benefit Analysis. Economic feasibility also plays a role. For instance, jobs that require high (low) levels of human expertise and are less (more) economically viable for automation may have lower (higher) likelihoods. This factor is considered in occupations such as ”Farm Equipment Mechanics and Service Technicians” and “Hand Packers and Packagers.”

These findings not only provide insight into the model's decision-making process but also align with established theories in labor economics and technological change. The consideration of task routineness, for instance, echoes the task-based approach to technological change (Autor *et al.*, 2003, 2006; Acemoglu and Autor, 2011; Autor, 2013), while the emphasis on ethical and regulatory factors reflects the growing awareness of AI's societal implications.

3.2 Comparison with Existing Indexes

Figure 2 compares the GENOE index with other indexes of computerization and occupational exposure to technological displacement.[6] Interestingly, our index shows little correlation with

[6] Since most indexes were calculated using the SOC-2010 classification, we transformed the GENOE from SOC-2018 to SOC-2010 using the crosswalks from the Bureau of Labor Statistics for comparison purposes. Henceforth, in this document, whenever the generated indexes are compared to existing indexes, they are transformed to SOC-2010.

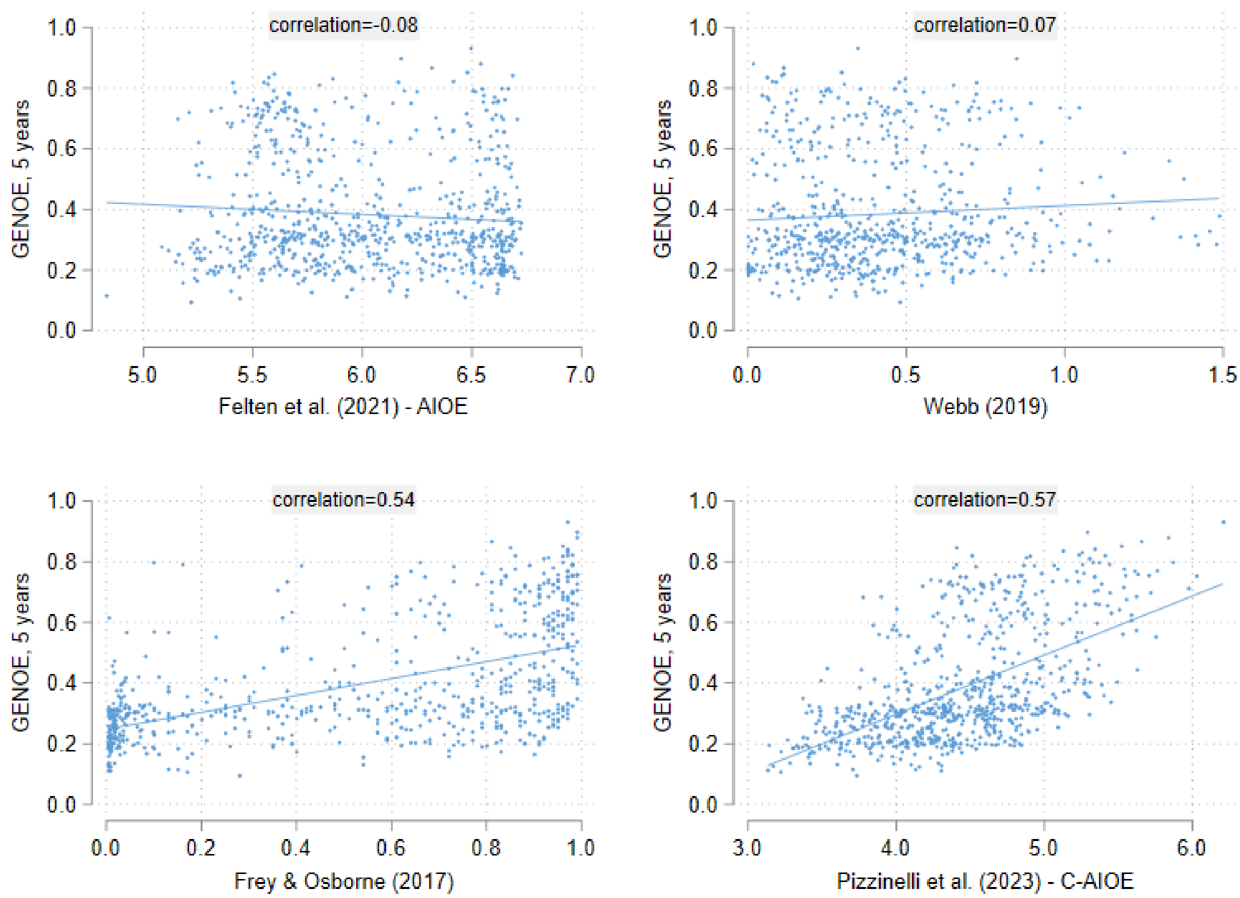
the indexes developed by Felten *et al.* (2021) and Webb (2019). However, GENOE demonstrates moderate correlations with the indexes created by Frey and Osborne (2017) and Pizzinelli *et al.* (2023). The similarities and differences can be attributed to several key factors in methodological approach and focus.

Firstly, Frey and Osborne's index places a higher emphasis on routinization, which aligns with our follow-up surveys, indicating a similar rationale behind our LLM models' assessment. As shown in Appendix Figure A.2, the GENOE index is higher for occupations intensive in routine manual and routine cognitive tasks. Both Pizzinelli et al.'s C-AIOE and Frey and Osborne's indexes also rank high for routine-intensive occupations, with Frey and Osborne's index higher for manual tasks and Pizzinelli et al.'s index higher for cognitive tasks.

Additionally, both Frey and Osborne (2017) and Pizzinelli *et al.* (2023) incorporate contextual elements beyond merely overlapping tasks or abilities with AI advancements. Pizzinelli et al.'s core adjustment to Felten et al.'s index involves considering potential complementarities between technology and labor arising from the physical and social factors of occupations. These include ethical considerations, training, and technology integration into the workplace. Similarly, Frey and Osborne's index adopts a holistic approach akin to ours. They evaluate not just isolated tasks or abilities but the overall occupation within a specific context and task composition to determine the extent of computerization.

The divergence from Felten et al.'s and Webb's indexes may be attributed to their more narrow focus on specific technological advancements or task-technology matches, without fully accounting for the broader occupational context or potential complementarities between AI and human labor.

Figure 2: Correlation of GENOE and Other Indexes of Occupational Exposure



Source: Authors' formulation based on data from O*NET Felten *et al.* (2021); Webb (2019); Pizzinelli *et al.* (2023). *Note:* This figure displays the GENOE for the next one, five, and 10 years. The index measures the likelihood that occupations will be displaced by AI within each time horizon.

3.3 Alternative Methodologies

We explore how GENOE estimates vary under different methodologies and prompt choices. Our baseline specification involves providing the model with the occupation name and a vector of tasks and asking for an estimate of the likelihood of job replacement. We then introduce six alternative indexes, each employing a distinct approach:

Blind Index: This approach removes the occupation name, relying solely on the task vector for estimations. This allows us to isolate the impact of task-specific factors from occupation-specific context factors in the model's predictions and analyze if the absence of the occupation name significantly alters the estimates.

Task-by-Task Index: Instead of aggregating tasks under an occupation, we estimate the likelihood of replacement for each task individually. These estimates are then averaged and weighted according to the relevance and importance of each task within the occupation, providing a granular view of task vulnerability.

Ethical Emphasis Index: Here, we explicitly ask the model to consider the ethical repercussions and potential social resistance to replacing certain tasks with AI. This approach highlights whether explicitly including ethical considerations causes a significant deviation from the baseline, which may already implicitly factor in such considerations.

Ethical Exclusion Index: This index explicitly asks the model not to consider ethical, regulatory, or social impediments to replacing tasks with AI. The objective of this index, compared to the previous one, is to measure the increase in the index when these considerations are excluded.

LLM-Specific Index: This variant focuses specifically on large language model (LLM) technologies. By narrowing the scope to LLM innovations, we assess the unique impact of advancements in natural language processing and related applications on occupational exposure, contrasting it with the general AI exposure measured in the baseline.

Non-AI Technologies Index: We estimate the probability of task replacement by non-AI technologies such as digital tools, robotics, and computer software. This provides a comparative baseline, distinguishing the impact of traditional technologies from AI-driven automation.

AI-Only Index: This index is derived by subtracting the Non-AI Technologies Index from the baseline GENOE, capturing the incremental exposure attributable solely to AI advancements. It allows us to identify occupations where AI-specific risks have significantly increased.

Table 4 provides the descriptive statistics of these alternative indexes, their correlation with the baseline GENOE, and other existing exposure indexes, and the extent to which they differ from the baseline. Both the Blind Index and the Ethical Emphasis Index do not deviate significantly from the baseline, although the mean differences are statistically significant. This indicates that there are slight changes in the reasoning of the model when the occupation name is omitted or when ethical considerations are explicitly prompted, but most of these factors are already incorporated in the baseline index. The overall correlations of these indexes with the baseline GENOE are 0.95 and 0.99, respectively. This consistency is supported by follow-up surveys, where ethical considerations and occupation-specific criteria were crucial elements for many of the model's assessments.

GENOE's ability to reflect ethical, regulatory, and social considerations in its assessment of AI's potential to replace human labor, even when not explicitly directed to do so, is a critical aspect of our argument. This aspect was initially confirmed by our follow-up survey analysis and is also supported by the comparisons with alternative indexes. Notably, when we explicitly instructed the model to include these factors, the resulting index showed only a minor reduction of one percentage point over the 5-year horizon. However, when we asked the model to exclude these considerations, the index increased by four percentage points, from 0.38 to 0.42. A similar pattern is observed for the 10-year horizon, as illustrated in Appendix Figure A.3, though the effect is less marked over the one-year horizon.

For the Task-by-Task Index, the average exposure increases substantially to 0.56, compared to 0.38 in the baseline. This highlights the fact that exposure to individual tasks does not linearly translate into the same exposure to job replacement. In other words, some occupations have high task-exposure but are embedded in occupation-specific contexts that mitigate overall exposure. Additionally, as previously mentioned, the disaggregated approach does not account for how the interrelation of tasks may result in different levels of exposure to replacement. However, it is worth noting that this index, despite having higher values, exhibits a high correlation with the baseline index (0.89).

The specific LLM Index estimations suggest that although a significant source of exposure emerges from the enhanced capabilities of language model technologies (with an index value of 0.20), it accounts for only 54% of the general AI index. Thus, the model considers broader potential applications of AI in the baseline scenario. The overall correlation with the GENOE is modest, at about 0.47. It is also noteworthy that the LLM Index is more correlated with some of the existing indexes of AI exposure, such as Felten et al. (2021)'s AIOE and Pizzinelli et al. (2023)'s C-AIOE. This may reflect the fact that these indexes also capture the specific impacts of language model technologies and their applications.

Exposure to non-AI technologies is also considerable, with an average index value of about 0.34. Interestingly, most occupations exposed to AI are also exposed to other types of technologies, as suggested by the high correlation of the Non-AI Index with the baseline (0.95). Additionally, this index is most correlated with the Frey and Osborne (2017) measure of exposure, probably because that index relates more closely to the broader exposure to computerization, including non-AI innovations.

The final "Only AI" Index allows us to determine which occupations will experience a higher increase in exposure to technology specifically due to AI. The average value of this index is low, at 0.04, but it reaches up to 0.39 for the highest values. Occupations with the highest values include clerical workers and data-intensive and writing-intensive roles such as translators, credit analysts, insurance underwriters, customer service representatives, paralegals, loan officers, and proofreaders.

Table 4: Descriptive Statistics of Alternative Methodologies for GENOE

	Baseline GENOE	Blind to occup. name	Task-by-task	Ethical emphasis	Ethical exclusion	LLMs	Non-AI technologies	Only AI
mean	0.382	0.391	0.560	0.370	0.423	0.204	0.343	0.039
p25	0.245	0.283	0.446	0.238	0.277	0.100	0.216	0.005
p50	0.317	0.325	0.546	0.312	0.355	0.175	0.295	0.031
p75	0.475	0.469	0.680	0.438	0.594	0.261	0.399	0.068
min	0.095	0.117	0.250	0.098	0.079	0.009	0.089	-0.108
max	0.932	0.862	0.891	0.885	0.945	0.857	0.919	0.391
Corr with baseline	1.000	0.947	0.888	0.992	0.982	0.475	0.946	0.346
Corr with FN	-0.081	-0.043	0.065	-0.102	-0.082	0.661	-0.261	0.507
Corr with PZ	0.569	0.559	0.616	0.560	0.564	0.666	0.430	0.511
Corr with WB	0.072	0.061	0.066	0.061	0.062	-0.036	0.048	0.085
Corr with FO	0.545	0.501	0.470	0.546	0.553	0.029	0.613	-0.082
P-value of diff with baseline	.	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Source: Authors' calculations using data from O*NET and surveys conducted with GPT-4o through OpenAI's API. *Note:* This table shows the descriptive statistics of the alternative indexes at the SOC-2010 occupational level for the five-year horizon, the correlation with the baseline GENOE and other existing indexes, and the p-value of the mean difference with the baseline GENOE. FN refers to Felten *et al.* (2021), PZ to Pizzinelli *et al.* (2023), WB to Webb (2019) and FO to Frey and Osborne (2017).

These alternative methodologies serve a dual purpose: they validate the robustness of our baseline GENOE index while offering deeper insights into the complex landscape of AI's potential impact on the workforce. By examining occupational exposure through various lenses—from task-level granularity to ethical considerations and AI-specific advancements—we gain a more comprehensive understanding of the challenges and opportunities that lie ahead. Our findings underscore the necessity of a multifaceted approach in assessing the future of work, one that considers not only the technical capabilities of AI but also the contextual, ethical, and broader technological factors that shape occupational vulnerability. This perspective is crucial for policymakers, employers, and workers as they navigate the evolving dynamics of human-AI interaction in the workplace and prepare for a future where the boundaries between human and artificial intelligence continue to shift.

3.4 Replication of Felten *et al.* (2021)'s AIOE Using Synthetic Surveys

To validate the effectiveness of synthetic surveys in producing assessments comparable to those of human experts, we replicated the well-established AI Occupational Exposure Index (AIOE) developed by Felten *et al.* (2021), which has been widely used in various empirical applications

(Acemoglu *et al.*, 2022; Fossen and Sorgner, 2021).[7] In our replication, we replaced the original human expert surveys with synthetic surveys conducted using GPT-4o via OpenAI's API.

The AIOE measures the degree of overlap between advancements in AI and workplace abilities. They define 10 key AI applications from the Electronic Frontier Foundation, which have seen considerable development since 2010: abstract strategy games, real-time video games, image recognition, visual question answering, image generation, reading comprehension, language modeling, translation, and speech recognition. They then link these AI applications to the 52 labor abilities recorded in O*NET by conducting 2,000 surveys with gig workers from Amazon's Mechanical Turk (MTurk).[8] Each participant was asked whether they believe each application is related to each ability, with a yes or no answer. These responses were processed as binary variables and averaged across all respondents to construct a matrix of relatedness, with a score between 0 and 1 for each application and each ability. The ability-level exposure was calculated as the sum of the scores of all applications for Ability j (A_j). Finally, they calculated the occupation-level ($AIOE_k$) for every occupation k as follows:

$$\text{AIOE}_k = \frac{\sum_{j=1}^{52} A_j \times L_{jk} \times I_{jk}}{\sum_{j=1}^{52} L_{jk} \times I_{jk}} \tag{1}$$

Using the replication files for the AIOE made available online by the authors, we re-estimated the $AIOE_k$ by replacing the matrix built with mTurk surveys with a vector resulting from surveys conducted with GPT-4o through OpenAI's API. Instead of defining the 10 AI applications, we asked the language model the following: "Based on recent advancements in AI, please answer 'Yes' or 'No' depending on whether you believe that AI is related to or could be used for the following ability: [ability name] [ability description]." This was done for the 52 abilities defined in O*NET and repeated 100 times, with the temperature setting randomized. Although this approach differs from the original methodology by not specifying 10 applications, we made this change to minimize conditioning the model. This method likely results in a more conservative estimate of the index by reducing potential similarities between the original and replicated indexes.

Since we do not have specific AI applications, we calculated the A_j as 10 times the average score from the 100 iterations. This adjustment was made solely for comparability purposes and does not affect the index calculation. We then ran the same equation (1) and computed the new index.

The results from the estimation, and the comparison with the original index, can be observed in Figure 3. The correlation between both indexes is 0.99, indicating a very high degree of replicability and very small variations between the answers from the mTurk workers and the answers

[7]The AIOE was initially introduced in Felten *et al.* (2018) and subsequently modified and updated in Felten *et al.* (2021).

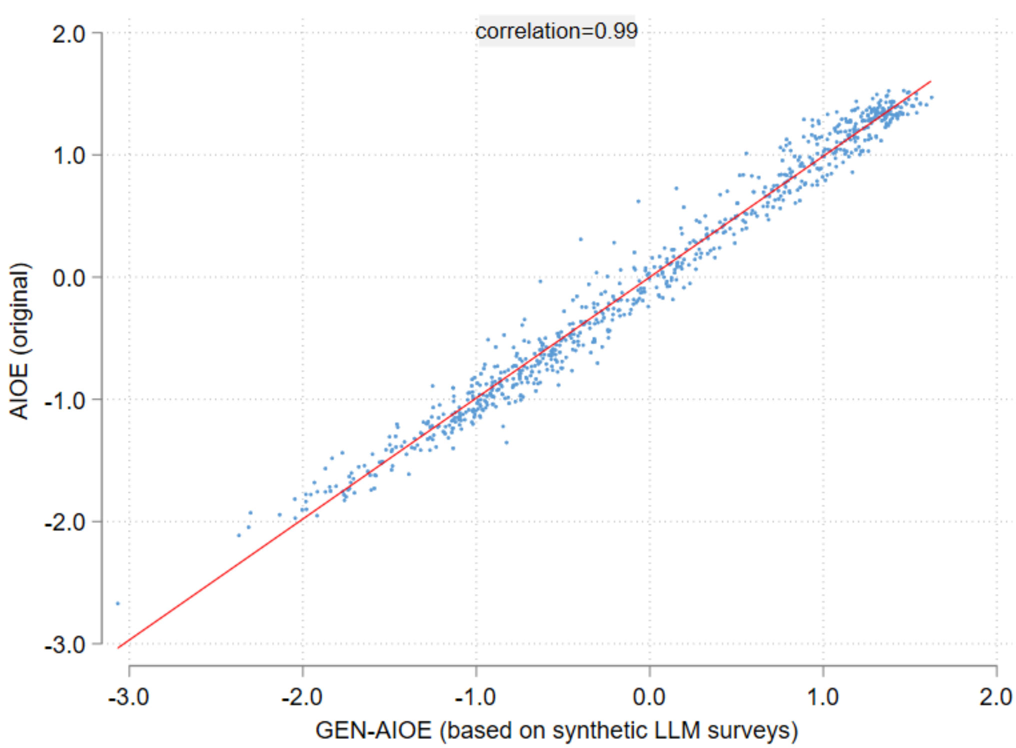
[8]Amazon Mechanical Turk (MTurk) is an online crowdsourcing platform that allows businesses to outsource a wide range of tasks to a global, distributed workforce. These tasks can vary from data validation and research to content moderation and survey participation.

provided by the synthetic surveys with GPT-4o.

A potential concern might be that GPT-4o could generate similar responses to mTurk workers because it searches the web or because its training data includes information from the AIOE and the mTurk surveys. However, GPT-4o does not perform real-time web searches unless explicitly instructed to do so, and in our study, we did not prompt GPT-4o to search the web. Regarding the second concern of the training data, as shown in follow-up surveys from the GENOE, the reasoning of the model is more holistic, based on its understanding of AI advancements and specific labor and workplace tasks (in this case, abilities), and does not rely on specific unique sources. Therefore, while it is possible that the training data could include some pre-existing information on Felten *et al.* (2021)'s work, among millions of other sources, the responses are synthesized from a broader knowledge base and not directly copied from any particular source.

Therefore, the high correlation observed suggests that the model's responses are consistent with human judgment for this specific task. This consistency supports the robustness and reliability of using artificial intelligence tools to emulate human analytical capabilities in these types of assessments.

Figure 3: Correlation of Original AIOE and AIOE Built with Synthetic Surveys



Source: Authors' formulation based on data from O*NET and Felten *et al.* (2021) *Note:* Each dot represents an occupation at the SOC 10-digit code.

A key strength of the GENOE index is its capacity for a detailed analysis of occupational exposure to AI-driven automation across different labor markets. In this section, we use the index to examine two distinct contexts: the United States, the world's largest developed economy, and Mexico, one of the largest emerging market economies.

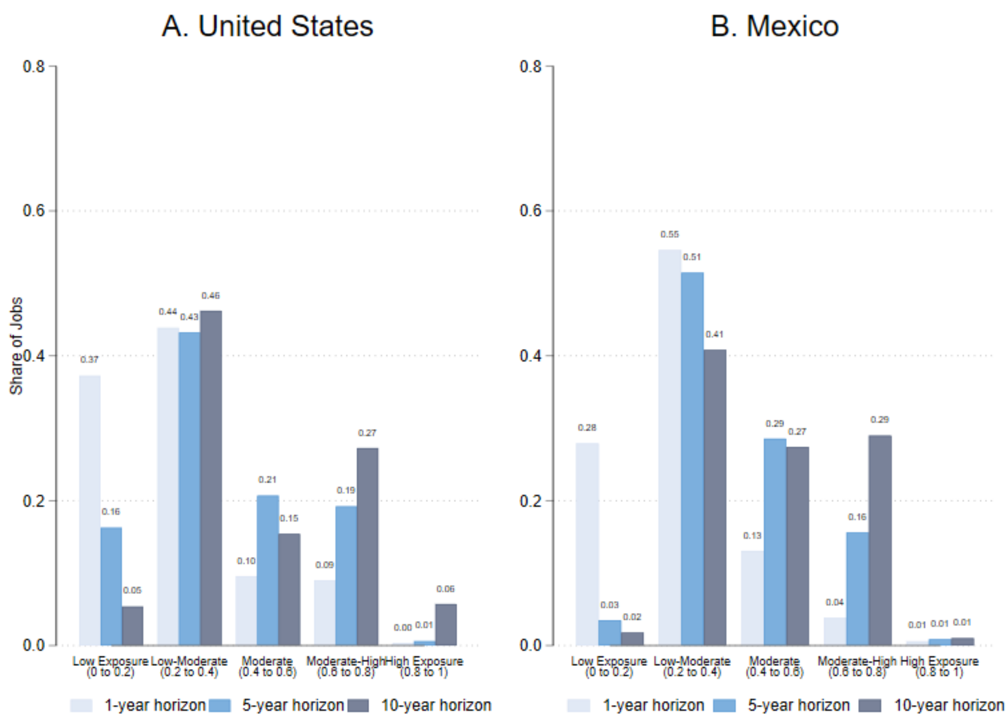
We retrieved data from the U.S. Bureau of Labor Statistics' Current Population Survey (CPS) and Mexico's INEGI Encuesta Nacional de Ocupación y Empleo (ENOE), focusing on surveys from January to December 2023. We mapped the GENOE estimates from the SOC 10 classification to each survey using the respective employment classifications provided by each survey, with crosswalks from the U.S. Bureau of Labor Statistics and the International Labour Organization (ILO). Our analysis concentrated on workers between 15 and 65 years old.

Figure 4 displays the distribution of jobs in the United States and Mexico according to the GENOE index across different time horizons. For the one-year horizon, 81% of jobs in the United States fall into the lower exposure categories (0 to 0.4), compared to 83% in Mexico. In the middle exposure category (0.4 to 0.6), Mexico has a slightly higher concentration of jobs (13%) compared to the United States (10%). In the high exposure categories (0.6 to 1), there is a low concentration of jobs in both countries, with 9% in the United States and only 5% in Mexico, and virtually no jobs at the highest risk level (0.8 to 1) in both countries.

As the time horizon extends, a shift towards higher exposure becomes evident for both countries. Looking at the five-year horizon, the share of jobs in the low exposure categories decreases to 59% in the United States and 54% in Mexico. The middle exposure category increases to 21% in the United States and 29% in Mexico. The high exposure categories remain relatively low, with 20% in the United States and 17% in Mexico, indicating a strong increase in vulnerability over this time frame.

The long-term outlook shows a significant increase in potential AI impact. In the 10-year horizon, the low exposure categories comprise 51% of jobs in the United States and 43% in Mexico, while the high exposure categories increase to 33% and 30% in respectively. This showcases a significant increase in potential AI impact over the long term, though the highest risk level (0.8 to 1) remains low at 6% in the United States and 1% in Mexico.

Figure 4: Distribution of Employment in the United States and Mexico by Occupational Exposure



Source: Authors' calculations using data from the Current Population Survey (CPS), Encuesta Nacional de Ocupación y Empleo (ENOE), O*NET, and surveys conducted with GPT-4 through OpenAI's API. *Note:* This figure shows the share of total jobs in each range of occupational exposure according to the GENOE index for each time horizon and country.

Table 5 presents the average exposure of the U.S. labor market to AI displacement, broken down by various socioeconomic factors. The average exposure is 0.28 for the one-year horizon, indicating that approximately 42.6 million jobs, or roughly one-quarter of total employment, are at risk of being displaced by AI adoption. Over longer time horizons, the average exposure increases to 0.39 (59.8 million jobs) and 0.46 (70.2 million jobs) for the five-year and 10-year projections, respectively.

Notably, women face higher exposure levels across all three time horizons, primarily because many work in office, administrative, and support occupations, which are among the most vulnerable to AI. Additionally, workers with lower levels of education, such as those who have not completed high school or only have a high school diploma, and individuals in the lowest earnings quintile, exhibit higher exposure scores compared to other groups. These findings suggest that AI adoption could exacerbate income inequality, echoing patterns observed in previous technological advancements (Autor et al., 2006; Autor and Dorn, 2013; Autor, 2019; Acemoglu and Restrepo, 2022).

Table 5: Occupational Exposure in the U.S. Labor Market According to GENOE

	1-year horizon	5-year horizon	10-year horizon
All workers	0.280	0.392	0.461
Sex			
Men	0.262	0.381	0.455
Women	0.298	0.404	0.466
Education level			
No High-School	0.293	0.423	0.503
High-School	0.309	0.434	0.509
Associate Degree	0.276	0.386	0.451
Bachelor's	0.261	0.357	0.417
Master's	0.199	0.281	0.333
Professional School	0.158	0.242	0.313
Doctorate	0.175	0.260	0.319
Quintiles of weekly earnings			
1st	0.339	0.469	0.549
2nd	0.335	0.464	0.538
3rd	0.320	0.446	0.518
4th	0.301	0.422	0.492
5th	0.215	0.318	0.378
Jobs exposed (in millions)	42.642	59.763	70.249

Source: Authors' calculations using data from the 2023 Current Population Survey (CPS), O*NET, and surveys conducted with GPT-4 through OpenAI's API. *Note:* This table shows the average GENOE score for each time horizon by gender, educational level, and quintiles of weekly earnings. The total jobs exposed are calculated as the average exposure multiplied by total U.S. employment in 2023.

Average occupational exposure in Mexico is quite similar to that in the United States, with 28%, 41%, and 48% of jobs exposed to AI in the one-year, five-year, and 10-year horizons, respectively (Table 6). In the immediate future, the analysis indicates that approximately 16 million jobs are at risk of being displaced by AI adoption. In contrast to the United States, the distribution of exposure by educational level and earnings is different. In Mexico, both low- and medium-skilled workers—those with post-secondary technical and professional education—face higher exposure levels. Additionally, exposure is higher for workers in the middle of the wage distribution.

This pattern probably reflects the occupational structure in Mexico (and probably of many developing economies), where office and administrative workers are more commonly found in more middle-income positions, unlike in the United States and the developed world, where these roles are typically at the lower end of the income distribution. In Mexico, these positions are often formal jobs, more vulnerable to AI-driven automation, as also shown in the table. Conversely, informal jobs, which are typically lower on the income spectrum, show less exposure to AI. Lower exposure of informal jobs is likely due to their less intensive use of technology and computer-based tasks, as these roles tend to be more manual and less reliant on digital processes.

Table 6: Occupational Exposure in Mexico's Labor Market According to GENOE

	1-year horizon	5-year horizon	10-year horizon
All workers	0.288	0.410	0.481
Sex			
Men	0.279	0.402	0.472
Women	0.303	0.423	0.494
Type of worker			
Formal	0.306	0.422	0.487
Informal	0.274	0.401	0.476
Education level			
No High-School	0.292	0.423	0.497
High-School	0.208	0.295	0.352
Technical	0.298	0.417	0.482
Bachelor's	0.284	0.385	0.447
Master's	0.213	0.296	0.349
Doctorate	0.182	0.253	0.302
Quintiles of monthly earnings			
1st	0.278	0.409	0.484
2nd	0.301	0.432	0.508
3rd	0.300	0.426	0.497
4th	0.285	0.402	0.470
5th	0.257	0.358	0.421
Jobs exposed (in millions)	15.679	22.307	26.141

Source: Authors' calculations using data from the 2023 Encuesta Nacional de Ocupación y Empleo (ENOE), O*NET, and surveys conducted with GPT-4 through OpenAI's API. *Note:* This table shows the average GENOE score for each time horizon by gender, type of worker, educational level, and quintiles of weekly earnings. The total jobs exposed are calculated as the average exposure multiplied by the total Mexican employment in 2023.

Extrapolating these findings to Latin America and the Caribbean, we estimate that approximately 84 million jobs are exposed to AI within one year, rising to 114 million in 5 years, and 132 million in 10 years. Extending this average exposure to global labor markets suggests that approximately 980 million, 1.33 billion, and 1.54 billion jobs may be exposed to AI over the one-, five-, and 10-year horizons, respectively. However, it is important to note that this projection assumes uniform AI adoption rates and labor market structures across countries, which may not accurately reflect the global economic landscape.

5 LIMITATIONS AND CONSIDERATIONS

While our methodology offers a novel and comprehensive approach to assessing occupational exposure to AI, it is important to acknowledge its limitations. Firstly, the assessments are based on the knowledge and potential biases inherent in the GPT-4o model. Although state-of-the-art, this model may have limitations in understanding certain specialized fields or future AI developments. Additionally, predictions about future AI capabilities, especially in the five and 10-year horizons, are inherently uncertain and may not account for unforeseen breakthroughs or obstacles in AI development. Our method attempts to capture broader contextual factors, but it may not fully account for all social, economic, and regulatory factors that could influence the adoption of AI in various occupations. Lastly, while the O*NET task descriptions are detailed, they may not capture all ways in which tasks are performed in real-world settings, potentially affecting the accuracy of AI exposure estimates.

Another limitation of our index—and indeed, of all indexes that aim to capture occupational exposure to technology—is that it does not account for general equilibrium effects that can influence social preferences and the demand for certain jobs. For instance, while firefighters currently exhibit low exposure estimates to technological displacement, advances in electrical systems, domestic safety features, and industrial processes could lead to a significant reduction in the occurrence of fires. This, in turn, could decrease the overall demand for firefighters, despite their low exposure to direct technological replacement. Thus, our index may not fully capture how broader technological progress can indirectly impact the demand for certain occupations through changes in societal needs and preferences.

Despite these limitations, the GENOE index offers a unique and holistic perspective on the future of work in the age of artificial intelligence. By combining a robust data source, a novel AI-driven survey methodology, and comprehensive validation exercises, we have created a valuable tool for researchers, policymakers, and industry leaders to anticipate and prepare for the potential impacts of AI on the labor market. The index's strengths lie in its comprehensive coverage, consistent evaluation criteria, and ability to capture more complex aspects like interactions between tasks and contextual factors. As such, while acknowledging its constraints, we believe the GENOE index provides a significant contribution to our understanding of occupational exposure to AI and can inform strategic decision-making in workforce planning and policy development.

6 CONCLUDING REMARKS

This paper introduces the AI-Generated Index of Occupational Exposure (GENOE), a novel measure of susceptibility to job replacement by artificial intelligence. Using synthetic AI surveys conducted with advanced language models, we estimate the likelihood of job replacement by AI over one, five, and 10-year horizons. Our findings reveal a significant increase in exposure over time, with the average likelihood of job replacement rising from 0.28 in the next year to 0.38 and 0.44 over the next five and 10 years, respectively, underscoring the accelerating impact of AI on the labor market.

The GENOE methodology of synthetic AI surveys leverages the extensive knowledge and analytical capabilities of AI models to provide expert-like assessments. This approach offers a comprehensive view of occupational exposure by considering not only the tasks performed within each occupation but also ethical and social contexts that influence the likelihood of AI-driven automation. By doing so, it overcomes the limitations of previous measures by incorporating a more holistic task-based analysis and integrating other occupation-specific, social, and ethical considerations more organically and with fewer assumptions. This contrasts with prior methods that isolated tasks or abilities separately or relied on more predefined criteria for measuring exposure.

We validated the consistency of the index by varying key elements of the methodology and the prompts, such as calculating the task-by-task exposure, occupation name omission, and explicitly nudging the model to consider ethical and social resistance to automation considerations. Additionally, follow-up surveys with AI models uncovered key patterns including understanding and projections of current and expected technological capabilities, the routine versus non-routine nature of tasks, ethical considerations, regulatory environments, and cost-benefit analyses.

The GENOE index offers significant and practical uses and implications for academia, policymakers, industries, and individual workers. By identifying which types of tasks, occupations, and industries are most susceptible to AI-driven automation, the GENOE index can inform strategic decisions and policy formulations. For academia, it offers a robust framework for future research on labor market dynamics and technology's impact, opening new avenues for interdisciplinary studies on AI and employment.

Policymakers can use the index to develop targeted interventions and support measures for workers in highly exposed occupations, ensuring a smoother transition in the face of technological change. This data-driven approach can inform education and training policies, unemployment insurance programs, and economic development strategies. Industries can leverage the index to guide strategic decision on workforce development, technology integration, and long-term business planning. It can help companies anticipate skill gaps, plan for reskilling initiatives, and optimize human-AI collaboration.

For individual workers, the index provides valuable insights for career planning and skill development, helping them prepare for potential shifts in their occupational landscape. In the broader societal context, the GENOE index contributes to the ongoing dialogue about the future of work, AI ethics, and the need for responsible AI development and deployment.

While the GENOE index represents a significant advancement in understanding AI's impact on employment, it is important to acknowledge its limitations. The index is based on current AI capabilities and projected advancements, and unforeseen technological breakthroughs or obstacles could alter these projections. Additionally, the index does not account for potential new job creation resulting from AI advancements.

Future research could expand on this work by incorporating more dynamic elements, such as the potential for job creation and transformation due to AI, and by exploring the interaction between AI exposure and other economic and social factors. Regular updates to the index will be crucial to reflect the rapidly evolving AI landscape.

In conclusion, the GENOE index provides a valuable tool for understanding and preparing for the impact of AI on the workforce. By offering a holistic, forward-looking perspective on occupational exposure to AI, it equips stakeholders across society with the insights needed to navigate the challenges and opportunities of the AI-driven future of work.

REFERENCES

Acemoglu, D. and Autor, D. (2011) Skills, tasks and technologies: Implications for employment and earnings, in *Handbook of Labor Economics*, vol. 4, pp. 1043–1171.

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

Acemoglu, D. and Restrepo, P. (2022) Tasks, automation, and the rise in us wage inequality, *Econometrica*, **90**, 1973–2016.

Aristotle (350 B.C.E./1998) *Politics*, Hackett Publishing. (Original work published around 350 B.C.E.).

Autor, D. (2013) The "task approach" to labor markets: an overview, *Journal for Labour Market Research*, **46**, 185–199.

Autor, D. H. (2019) Work of the past, work of the future, in *AEA Papers and Proceedings*, American Economic Association, vol. 109, pp. 1–32.

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

Autor, D. H., Katz, L. F. and Kearney, M. S. (2006) The polarization of the us labor market, *American Economic Review*, **96**, 189–194.

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

Brynjolfsson, E., Mitchell, T. and Rock, D. (2018) What can machines learn and what does it mean for occupations and the economy?, in *AEA papers and proceedings*, American Economic Association, vol. 108, pp. 43–47.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M. *et al.* (2023) Opinion paper:"so what if chatgpt wrote it?" multidisciplinary perspectives on opportunities, challenges and implications of generative conversational ai for research, practice and policy, *International Journal of Information Management*, **71**, 102642.

Eloundou, T., Manning, S., Mishkin, P. and Rock, D. (2024) Gpts are gpts: Labor market impact potential of llms, *Science*, **384**, 1306–1308.

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

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

Fossen, F. M. and Sorgner, A. (2021) Digitalization of work and entry into entrepreneurship, *Journal of Business Research*, **125**, 548–563.

Frey, C. B. and Osborne, M. A. (2017) The future of employment: How susceptible are jobs to computerisation?, *Technological forecasting and social change*, **114**, 254–280.

Keynes, J. M. (1930) Economic possibilities for our grandchildren, in *Essays in Persuasion*, Springer, pp. 321–332.

Pizzinelli, C., Panton, A. J., Tavares, M. M. M., Cazzaniga, M. and Li, L. (2023) *Labor market exposure to AI: Cross-country differences and distributional implications*, International Monetary Fund.

Tolan, S., Pesole, A., Martínez-Plumed, F., Fernández-Macías, E., Hernández-Orallo, J. and 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. (2019) The impact of artificial intelligence on the labor market, *Working Paper*.

Appendix A. Additional Tables and Figures

Table A.1: GENOE Descriptive Statistics

	1-year horizon	5-year horizon	10-year horizon
mean	0.279	0.379	0.440
p25	0.163	0.241	0.297
p50	0.230	0.314	0.375
p75	0.338	0.466	0.591
min	0.043	0.095	0.112
max	0.919	0.932	0.946

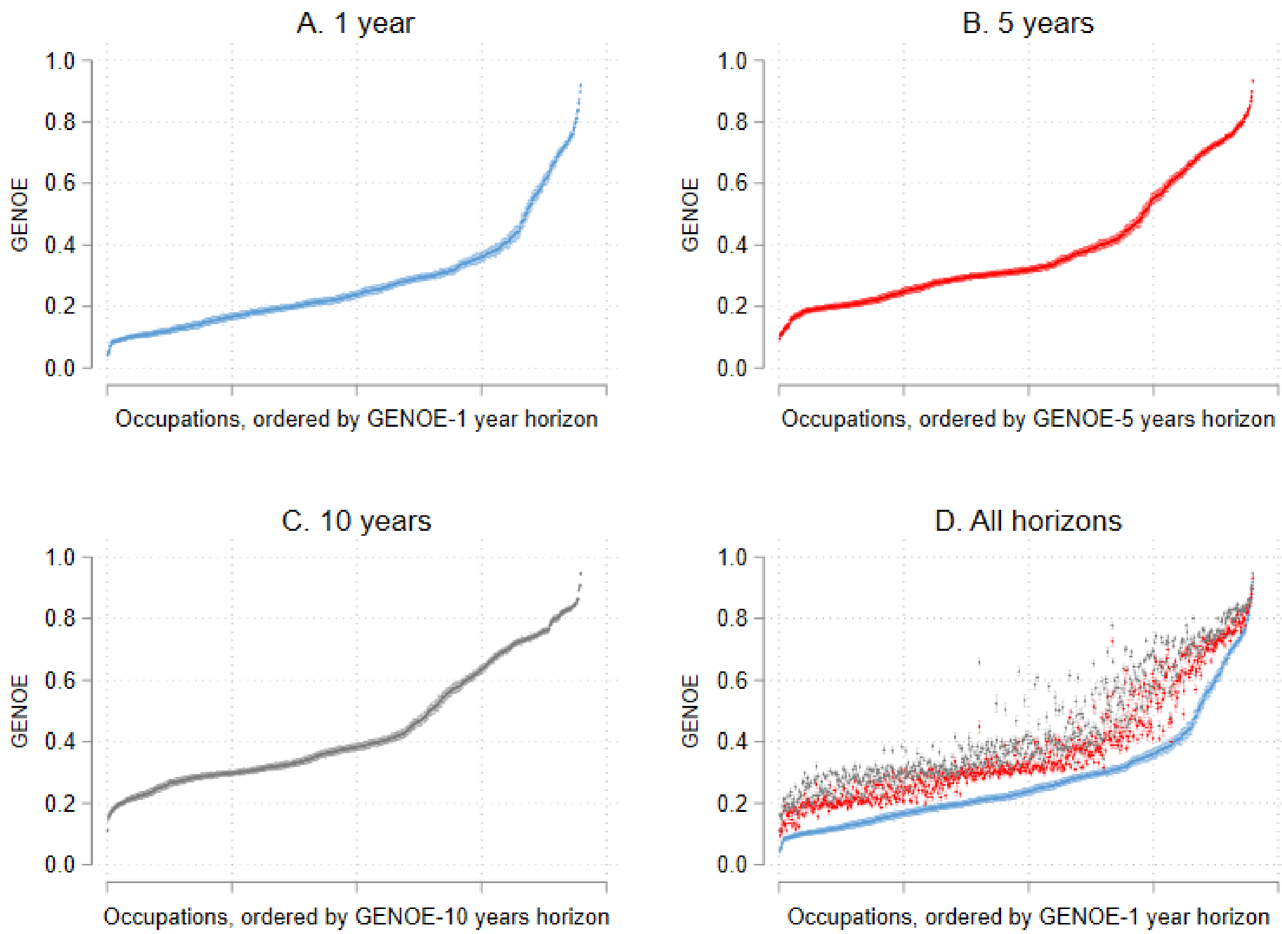
Source: Authors' calculations using data from O*NET and surveys conducted with GPT-4o through OpenAI's API. *Note:* This table shows the descriptive statistics of the GENOE at the SOC-18 occupational level for the one, five, and 10-year horizons.

Table A.2: Differences in the GENOE Index by Model Temperature

	High-temperature mean	Low-temperature mean	P-value of dif.
A. Above vs. below median temperature			
1 year horizon	0.278	0.279	0.873
5 year horizon	0.379	0.380	0.933
10 year horizon	0.440	0.440	0.996
B. Fifth vs. first temperature quintiles			
1 year horizon	0.278	0.279	0.897
5 year horizon	0.379	0.380	0.926
10 year horizon	0.440	0.439	0.959

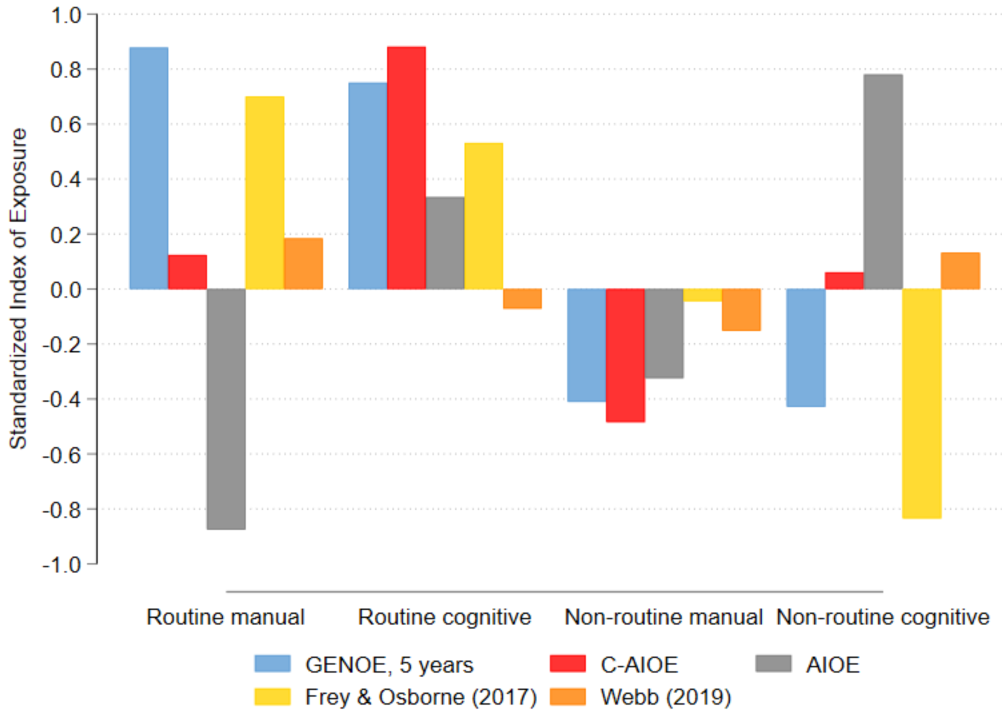
Source: Authors' calculations using data from O*NET and surveys conducted with GPT-4o through OpenAI's API. *Note:* This figure shows the mean GENOE index for iterations with different temperature settings across three time horizons. In Panel A, the low-temperature GENOE is the average of iterations below the median temperature, while the high-temperature GENOE is the average of iterations above the median. In Panel B, the low-temperature GENOE is the average of iterations in the first temperature quintile, and the high-temperature GENOE is the average of iterations in the fifth quintile.

Figure A.1: GENOE Distribution at the Occupational Level



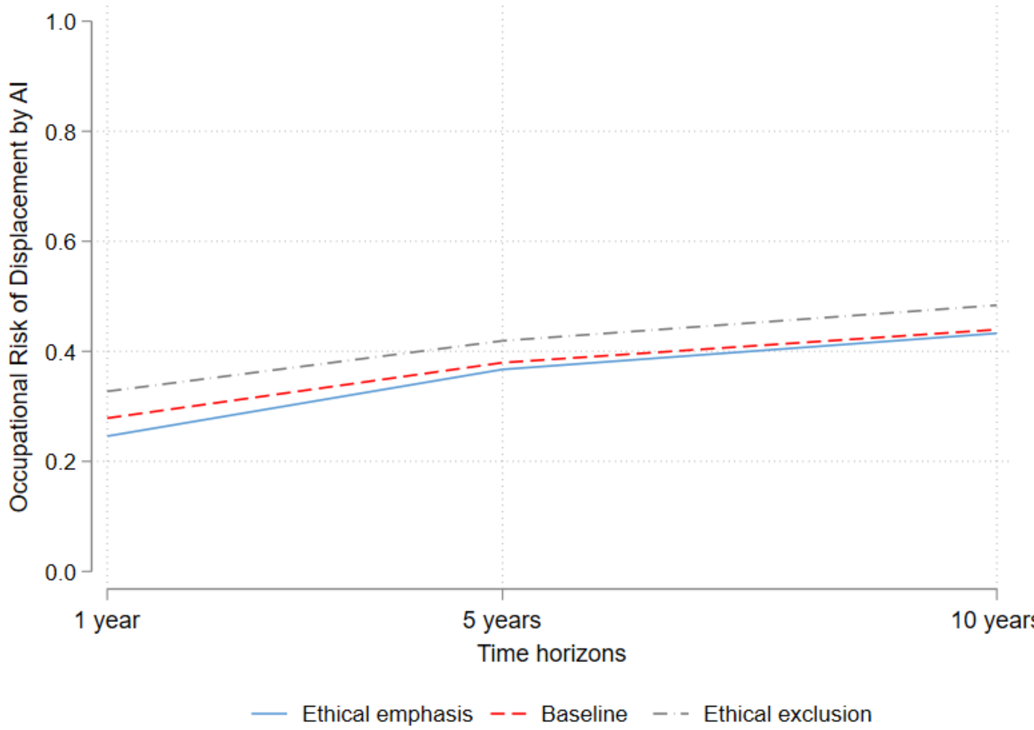
Source: Authors' calculations using data from O*NET and surveys conducted with GPT-4o through OpenAI's API.
Note: This figure plots the GENOE point estimate, averaged over 100 iterations, and the 95% confidence interval for each of the 759 occupations, ordered from lower to higher exposure. Each panel represents a different time horizon, and the bottom-right panel combines all horizons, ordered by the one-year horizon.

Figure A.2: Occupational Exposure by Task Nature Across Various Indexes



Source: Authors' formulation based on data from O*NET, Felten *et al.* (2021); Webb (2019); Frey and Osborne (2017); Pizzinelli *et al.* (2023). *Note:* This figure displays the average values of several indexes of occupational exposure, categorizing occupations into four groups: routine-manual, routine-cognitive, non-routine manual, and non-routine cognitive. These classifications were made following the methodology proposed by Acemoglu and Autor (2011). Indexes were standardized for comparability.

Figure A.3: AI-Generated Index of Occupational Exposure to AI (GENOE) Over Time Including and Excluding Ethical Considerations



Source: Authors' formulation based on data from O*NET, Felten *et al.* (2021); Webb (2019); Frey and Osborne (2017); Pizzinelli *et al.* (2023). *Note:* This figure displays the baseline GENOE for the next one, five, and 10 years, along with alternative indexes constructed both with and without ethical considerations. The index measures the likelihood that occupations will be displaced by AI within each time horizon and under each methodology.