Who Benefits and Who Gets Replaced? Demographic Patterns of Generative AI Exposure in the United States

Riccardo Omenti[1,*], Ludovica Maria Isani[1], Saverio Minardi[1], and Nicola Barban[1]

[1] *Department of Statistical Sciences, University of Bologna*

[*] *Corresponding email:* riccardo.omenti2@unibo.it

May 2026

Abstract

This paper examines the diffusion of Generative Artificial Intelligence (GenAI) across socio-demographic groups within the United States (US) working-age population. To this end, we build novel occupation-level measures of GenAI exposure based on task-based evaluations by Large Language Models (LLMs). Specifically, occupational task descriptions from O*NET database are scored according to the extent to which they can be complemented or replaced by GenAI, and these assessments are then aggregated at the occupational level. By merging these exposure indicators with micro-data from the 2024 American Community Survey (ACS), we examine the variation in occupational exposure to GenAI across various socio-economic strata using multivariate regression models. Our results reveal substantial heterogeneity in GenAI occupational exposure by gender, wages, educational attainment and race. Exposure increases with education and wages, with high-status workers more likely to experience GenAI complementarity, while lower-educated and lower-wage workers face relatively greater exposure to replacement.

Gender differences further shape these patterns. Women without tertiary education and in the middle of the wage distribution are particularly vulnerable to automation by GenAI, whereas highly educated and high-income women are among those most likely to benefit from GenAI-driven task enhancement. Across most racial groups, women exhibit higher levels of exposure to both complementarity and replacement than men, although this pattern does not hold for Asian women.

Overall, our findings reveal a socially stratified diffusion of GenAI, with marked differences across gender, wages, race and education, highlighting the need to investigate the broader societal and demographic implications of AI-driven technological change.

Keywords: Artificial Intelligence; Socio-demographic Stratification; Large Language Models; Gender Gaps; Computational Social Sciences

1 Introduction

Since the launch of ChatGPT in 2022, Generative Artificial Intelligence (GenAI), especially Large Language Models (LLMs), have radically transformed knowledge-intensive work. Thanks to their ability to generate, summarize, interpret, and transform text, code, and structure information, GenAI systems are reshaping how research is conducted, how information is processed, and how high-skilled occupations operate (Nature, 2025).

Although access to GenAI has expanded quickly, with ChatGPT alone reportedly reaching 850 million users as of November 2025 (Stanford Institute for Human-Centered Artificial Intelligence, 2025), its diffusion remains uneven. Adoption is concentrated in high-income countries, raising awareness that GenAI may amplify existing inequalities in access to technology and its associated benefits (Demombynes et al., 2025).

Within advanced economies, emerging evidence also points to substantial heterogeneity in occupational exposure to GenAI. Studies for the United States and Europe suggest that women and workers in highly educated, high-paying white-collar occupations tend to be more exposed to AI (Felten et al., 2023; Guarascio et al., 2025; Schendstok and Wertz, 2024). However, evidence based on individual level usage data indicates that younger workers and men are more likely to adopt GenAI tools in practice (Humlum and Vestergaard, 2025). These mixed findings underscore the need for a systematic assessment of how exposure to GenAI varies across socio-demographic groups.

A central methodological challenge concerns measurement (Frank et al., 2019). Because most large-scale population surveys do not capture individual-level GenAI usage, researchers rely on indirect measures of exposure. Existing studies typically construct occupation-level exposure indicators by linking detailed descriptions of occupational skills and tasks, such as those provided by O*NET, to assessments of AI capabilities. These measures evaluate the extent to which occupational activities overlap with, or can be complemented or substituted by, AI systems. Examples include indices that quantify AI occupational exposure based on the alignment between job requirements and AI applications, as well as more recent approaches that use large language models (LLMs) and human annotations to assess

the potential impact of GenAI on specific work activities (Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2022; Benítez and Parrado, 2024; Eloundou et al., 2024; Felten et al., 2021).

Building on this literature, this paper pursues two objectives. First, we implement a task-based framework that combines LLM-generated evaluations with detailed information from the O*NET database to construct occupational level indicators of complementarity and replacement. Second, we integrate these indicators with micro-level data from the American Community Survey (ACS) to investigate how GenAI exposure varies across socio-demographic subgroups in the United States (US). Specifically, we provide a detailed descriptive analysis of the differences in exposure by wage, educational attainment, gender, marital status and fertility. Thus, by linking computational advancements in the measurement of GenAI exposure to demographic analysis, this study provides new evidence on the demographic implications of GenAI in the US workforce population.

2 Data and Methods

2.1 Data

To evaluate exposure to GenAI across occupations, we rely on the Occupational Information Network (O*NET) 30.0 database developed by the US Department of Labor. This data source provides detailed task-level descriptions of occupations in the United States and has been widely used to quantify the nature and content of work (e.g., see Felten et al. (2021) and Felten et al. (2023)). For the purpose of our analysis, we focus on 17,000 distinct tasks that are associated to 711 occupations. Crucially, O*NET reports measures of both the importance and relevance of each task within an occupation, which enables the construction of weighted exposure indicators.

To investigate heterogeneity in exposure across socio-demographic subgroups, we combine these occupation-level measures with micro-level data from the 2024 ACS, which is also the most recent wave available. The ACS provides a nationally representative sample of the US population with individual-level data on employment and socio-demographic characteris-

tics. Occupations in the ACS are coded according to the Standard Occupational Classification (SOC) system, allowing us to link the exposure indicators with the O*NET occupational data.

2.2 Method

To quantify occupational exposure to GenAI, we evaluate the exposure of single tasks within each occupation as described in O*NET by using AI systems themselves as expert evaluators. Specifically, we use the LLM *gemma-3-27b-it*, developed by Google and accessed via the Hugging Face portal [1], to generate numerical assessments on the probability that each selected task is either to be complemented or replaced by GenAI systems. In our framework, complementarity refers to the extent to which GenAI augments human work, enhancing productivity and task performance without displacing the worker, while replacement indicates the potential for GenAI to substitute for human labor in the execution of tasks. GenAI refers to a class of AI systems designed to produce new content, such as text, images, audio, code, or video, by learning patterns from large datasets. These evaluations are produced using a structured prompt that provides detailed criteria for assessing task susceptibility to GenAI[2]. In addition, we instruct the model to provide a short explanation for each assigned task score. Collecting these textual explanations is a crucial part of the procedure, as they enhance the transparency of the evaluation process and enable the identification of potential inconsistencies, biases, or misunderstandings in the model's reasoning. Robustness checks with existing measures of exposure to GenAI, such as those developed by Felten et al. (2021) and Eloundou et al. (2024) are reported in the supplementary material. To further evaluate the robustness of our results, we assessed the model's performance across two distinct temperature settings: *temperature*=0 and *temperature*=1. This approach allowed us to determine whether increasing model creativity may significantly impact the LLM outcomes. As shown in the supplementary materials, the scores obtained at both temperature levels are highly correlated. Consequently, all results presented in the following section were com-

[1] Link to the model : gemma-3-27b-it

[2] We report the exact prompt in the supplementary material.

puted using a temperature of 1.

Our general approach for the computation of the GenAI occupational exposure scores is illustrated in Figure 1 . Specifically, we use the task-level probabilities of complementarity or replacement predicted by the LLM. We then classify a task as highly exposed to complementarity or replacement if its predicted probability exceeds the 80th percentile of the corresponding distribution, which is derived by considering the whole set of tasks from the O*NET database. Afterwards, we calculate occupation-level exposure scores as the weighted proportion of tasks within each occupation that exceed the previous treshold. Weights are determined by the importance and relevance of tasks within each occupation, as measured by O*NET[3]. These scores are subsequently standardized by subtracting the mean and dividing by the standard deviation. Following this normalization, occupations with positive values are interpreted as relatively more exposed to the complementarity or replacement by GenAI, whereas those with negative values are comparatively less exposed.

Afterwards, we merge the standardized scores with the ACS sample of employed individuals aged 25-64 in 2024 using the SOC occupation codes as the linking variable. This data linkage procedure allows us to assign exposure scores to 520 distinct occupations observed in the ACS data. We exclude individuals aged 20-24, as a sizable portion of this subpopulation is still in education. The inclusion of this group would likely bias the estimates by overstating the proportion of young individuals with low GenAI exposure. This issue arises from the selection into employment within this age group, as those who are employed in the age group $20-24$ are disproportionally concentrated in manual occupations, which are less exposed to GenAI.

[3]Importance refers to the degree to which a task is essential to a given occupation, measured on a scale from 1 to 5. Relevance captures the proportion of workers who report a given task as relevant to their occupation.

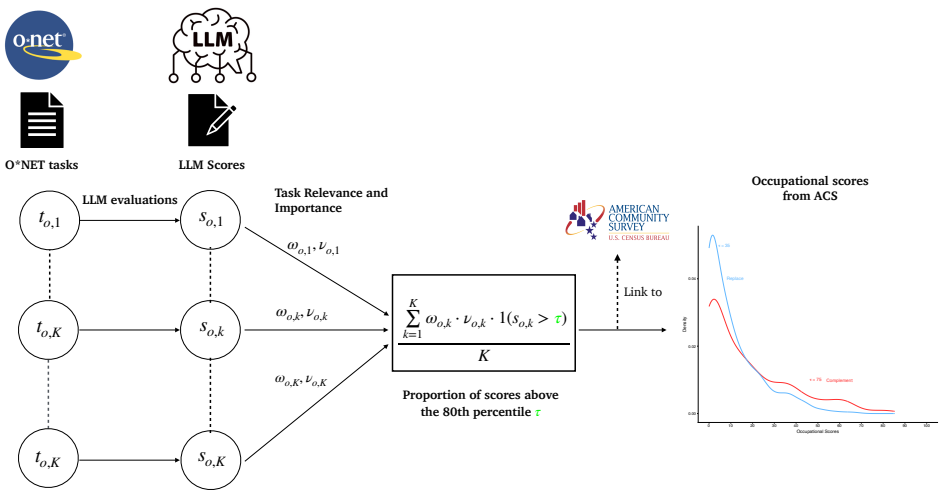

Figure 1: Description of the process for calculating the GenAI exposure scores.

Note: The scores reported in the figure are not standardized.

To examine associations between socio-demographic variables and exposure to GenAI, we estimate the following fixed-effect multivariate regression model, separately for complementarity and replacement.

$$Y_{i,c,o} = \alpha + \beta' \mathbf{X}_{i,c,o} + \gamma_c + \epsilon_{i,c,o}$$

The dependent variable $\mathbf{Y}_{i,o}$ is the standardized exposure score for an employed individual i working in occupation o and residing in the Public Use Microdata Area (PUMA) c[4]. The set of independent variables of the vector $\mathbf{X}_{i,c,o}$ include wage deciles, educational attainment, sex, race, marital status, number of children under age 5 and age. It also includes two-way interactions between sex and the following regressors: wage, education, race and number of children under 5. These interactions allow us to capture heterogeneous associations and provide a more granular understanding of how the joint distribution of sex and other socio-demographic characteristics varies with distinct exposure levels to GenAI complementarity and replacement. Additionally, the regression models incorporate PUMA fixed effects (γ_c) to account for unobserved, location-specific factors within the ACS that may influence expo-

[4]Public Use Microdata Areas (PUMAs) are non-overlapping, statistical geographic areas that partition each state or equivalent entity into geographic areas containing no fewer than 100,000 people each.

sure patterns, such as the local occupation structure and the share of technology industries.

3 Results

3.1 Descriptive Analysis

Panel A of Figure 2 illustrates the standardized scores of occupational exposure to GenAI across major occupational groups, differentiating between complementarity and replacement. Computer, engineering, and science (STEM) occupations present the highest positive scores in both dimensions, with complementarity consistently exceeding replaceability. This pattern underscores that GenAI adoption in this field is primarily augmentative rather than substitutive. A similar, albeit less pronounced, pattern emerges for legal, education, arts, managerial, and business occupations, where complementarity dominates replacement.

In contrast, sales and clerical occupations exhibit an opposite pattern. Sales occupations show higher replaceability than complementarity, whereas clerical occupations are characterized by positive replacement scores coupled with a negative complementarity. This seems to suggest a comparatively high exposure to automation, with clerical occupations particularly vulnerable to automation.

Healthcare, services and manual occupations exhibit negative scores along both dimensions, indicating limited exposure to GenAI. Taken together, these results seem to point to a clear gradient, in which knowledge-intensive occupations benefit the most from the GenAI complementarity, while routine-intensive white-collar roles, especially clerical roles, face greater risk of job displacement, and blue-collar and service workers remain, at least for now, relatively shielded from the effects of GenAI diffusion.

Panels B and C of Figure 2 present the ten occupations with the largest differences between GenAI complementarity and replacement scores. Occupations such as writers and authors, interior designers, and personal financial advisors, exhibit high complementarity coupled with relatively low replacement risk. In contrast, several clerical occupations, including fi-

nancial clerks, procurement clerks, and credit analysts, display comparatively low complementarity and high exposure to replacement risk. These patterns underscore that occupations requiring creativity, complex problem-solving, interpersonal interaction, and higher levels of educational attainment are the most likely to benefit from GenAI complementarity. Conversely, occupations characterized by routine cognitive tasks and repetitive administrative activities appear more vulnerable to displacement through GenAI-driven automation.

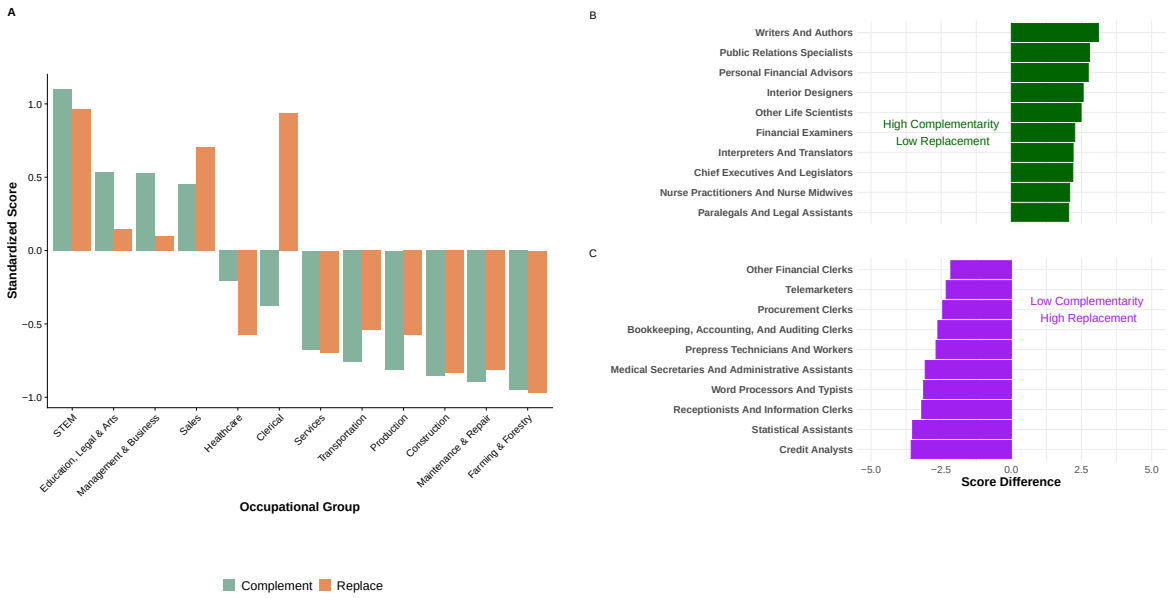

Figure 2: Panel A shows average predicted standardized score of GenAI complementarity and replacement by occupational group. Panel B shows the 10 occupations with the highest positive difference between the standardized complementarity and replacement scores. Panel C shows the 10 occupations with the highest negative difference between the standardized complementarity and replacement scores.

Note: Estimates are weighted according ACS sampling weighting scheme. Blue bars refer to exposure to GenAI replacement, while red bars indicate exposure to GenAI complementarity.

Figure 3 displays standardized scores of GenAI occupational exposure across socio-demographic subgroups, distinguishing between complementarity and replacement. A positive gradient is observed for educational attainment. Individuals with secondary education or less exhibit consistently negative scores along both dimensions, whereas scores increase progressively for higher levels of education, reaching their highest values for workers with a doctoral degree. A similar pattern is evident for wage, with negative scores for low- and lower-middle-wage groups and increasingly positive for upper-middle- and high-wage groups. Differences

by sex and race are also apparent. Asian and White individuals display positive scores in both complementarity and replacement, whereas Black and other racial groups are characterized by negative scores. In addition, women present higher scores than men in both dimensions.

By focusing on the comparison between complementarity and replacement within subgroups, among workers with higher levels of education and those in upper-middle and high wage categories, complementarity scores exceed replacement scores. In contrast, among workers with lower levels of education and lower wages, these differences are more attenuated.

Figure 3: Average predicted standardized score of GenAI complementarity and replacement by socio-demographic subgroup.

Note: Estimates are weighted according ACS sampling weighting scheme. Blue bars refer to exposure to GenAI replacement, while red bars indicate exposure to GenAI complementarity.

3.2 Regression Analysis

Figure 4 displays the coefficient estimates from the multivariate regression models, separately for complementarity and replacement. Coefficients represent changes in the average GenAI occupational exposure score associated with each socio-demographic characteristics relative to the corresponding reference category (e.g., less than high school for educational attainment). The results reveal a substantial heterogeneity in GenAI exposure across socio-demographic subgroups, both in magnitude and in the relative importance of complementarity and replacement.

Educational attainment emerges as the strongest predictor of GenAI exposure. Relative to workers without a high school diploma, exposure to GenAI increases monotonically with educational attaiment levels, particularly for complementarity. Workers holding bacherlor's, master's, and doctoral degrees exhibit substantially higher complementarity scores than replacement scores, suggesting that highly educated workers are more likely to experience GenAI exposure as productivity-enhancing rather than substitutive. Wage displays a similar pattern. Exposure generally rises across wage deciles, particularly for complementarity, underscoring a markedly higher occupational exposure to GenAI augmentation among higher-paid workers.

Pronounced differences also emerge across demographic groups. Asian workers display higher exposure scores along both dimensions compared to White workers, with a more pronounced difference in complementarity. In contrast, Black workers and those in other racial groups display lower exposure to GenAI relative to White workers, with similar magnitudes across complementarity and replacement. Gender differences are likewise non-negligible. Women display higher exposure scores than men, with a larger gap observed for replacement. This pattern indicate that men are, on average, less exposed to GenAI, particularly with respect to potential automation. Additionally, age-related differences in GenAI exposure are observed. Relative to workers in the age group 25-29, older workers exhibit consistently lower levels of Gen exposure, with only minor differences between complementarity and replacement. Associations with marital status and having children under 5 in the household are more modest but still significant. Separated or widowed workers tend to display

lower exposure to GenAI across both dimensions, while workers with one or more children under age 5 exhibit slightly lower exposure than those without young children.

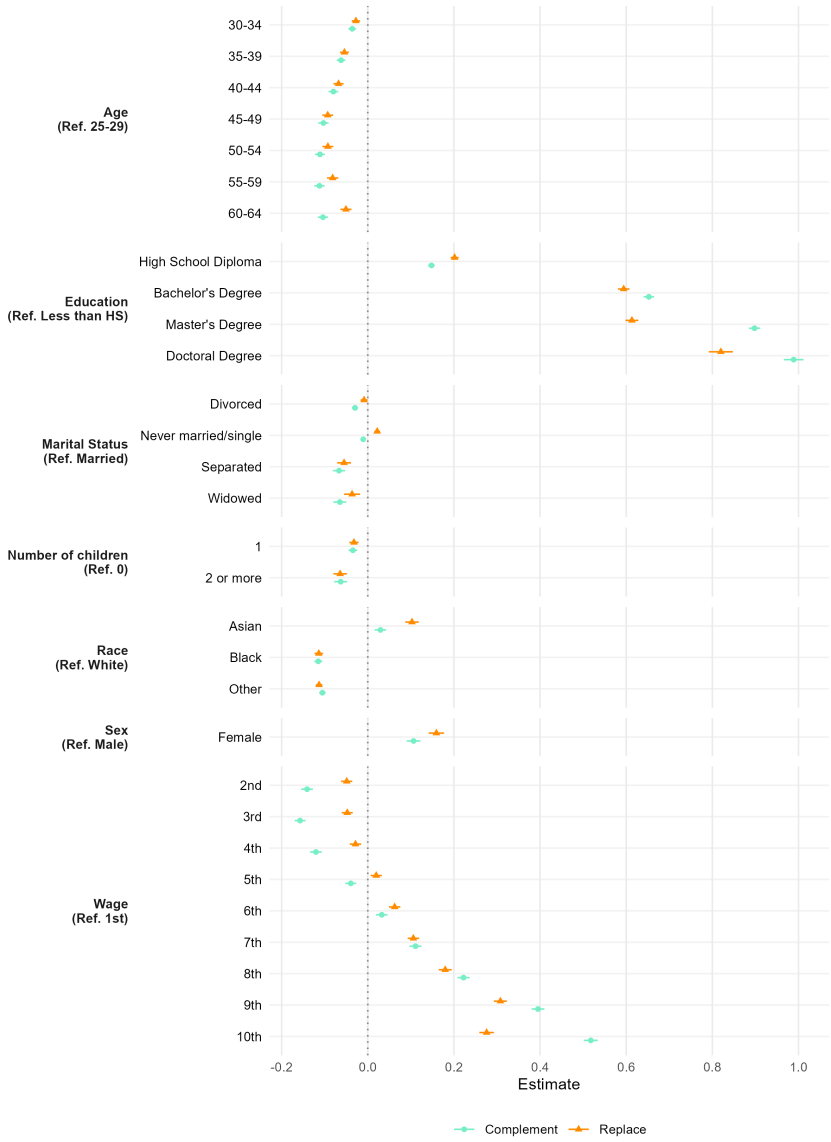

Figure 4: Coefficient plot displaying the association between the main socio-demographic predictors and GenAI occupational exposure scores differentiating between complementarity and replacement.

Note: The vertical axis lists predictors including age, educational attainment, sex, and wage, with their respective reference categories. The horizontal axis displays the regression coefficients, featuring a dashed vertical line at zero to indicate no effect. Each coefficient estimate includes horizontal error bars representing 95% confidence intervals. All models include PUMA fixed effects and interaction terms among sex, education, and wage. Standard errors are clustered at the PUMA level.

The previous results indicate that heterogeneity in GenAI exposure is primarily driven by

education and wages, with gender differences providing an additional layer of variation. Motivated by this pattern, we next examine how gender differences in GenAI exposure vary by education, wage and race.

Figure 5 presents gender differences in model-predicted GenAI exposure scores by wage, educational attainment and race, separately for complementarity and replacement. Positive values indicate higher predicted exposure among women relative to men, whereas negative values indicate higher exposure among men.

Substantial heterogeneity emerges across subgroups. Across nearly all wage levels, women exhibit higher predicted complementarity exposure than men. The complementarity gender gap increases from lower wage deciles, peaks among middle and upper-middle-income workers, and remains consistently positive among top earners. Gender differences in replacement exposure follow a pronounced inverted U-shaped pattern across the wage distribution. Women in lower- and middle-wage deciles are substantially more exposed to GenAI-related replacement risk than men, with the widest gaps occurring in the fourth and the fifth deciles. Nonetheless, the gap narrows at higher wage levels, reverses in the ninth decile, and becomes non-significant among workers in the top decile.

Educational attainment also displays a marked stratification. Among workers with low levels of education, women display higher exposure both across complementarity and replacement dimensions, with markedly higher gender gaps in GenAI replacement among workers with a high school degree. These gender gaps decreases with higher levels of educational attainment. Among workers holding a master's and doctoral degrees, the gaps become non-negligible or even negative, especially for replacement, underscoring that highly educated men may comparatively more exposed to GenAI-related automation risk than highly educated women.

Gender gaps in exposure to GenAI also differ by race category. Asian men are likely to be occupationally exposed to GenAI complementarity and replacement. In contrast, white and black women are more likely to complemented or replaced by GenAI.

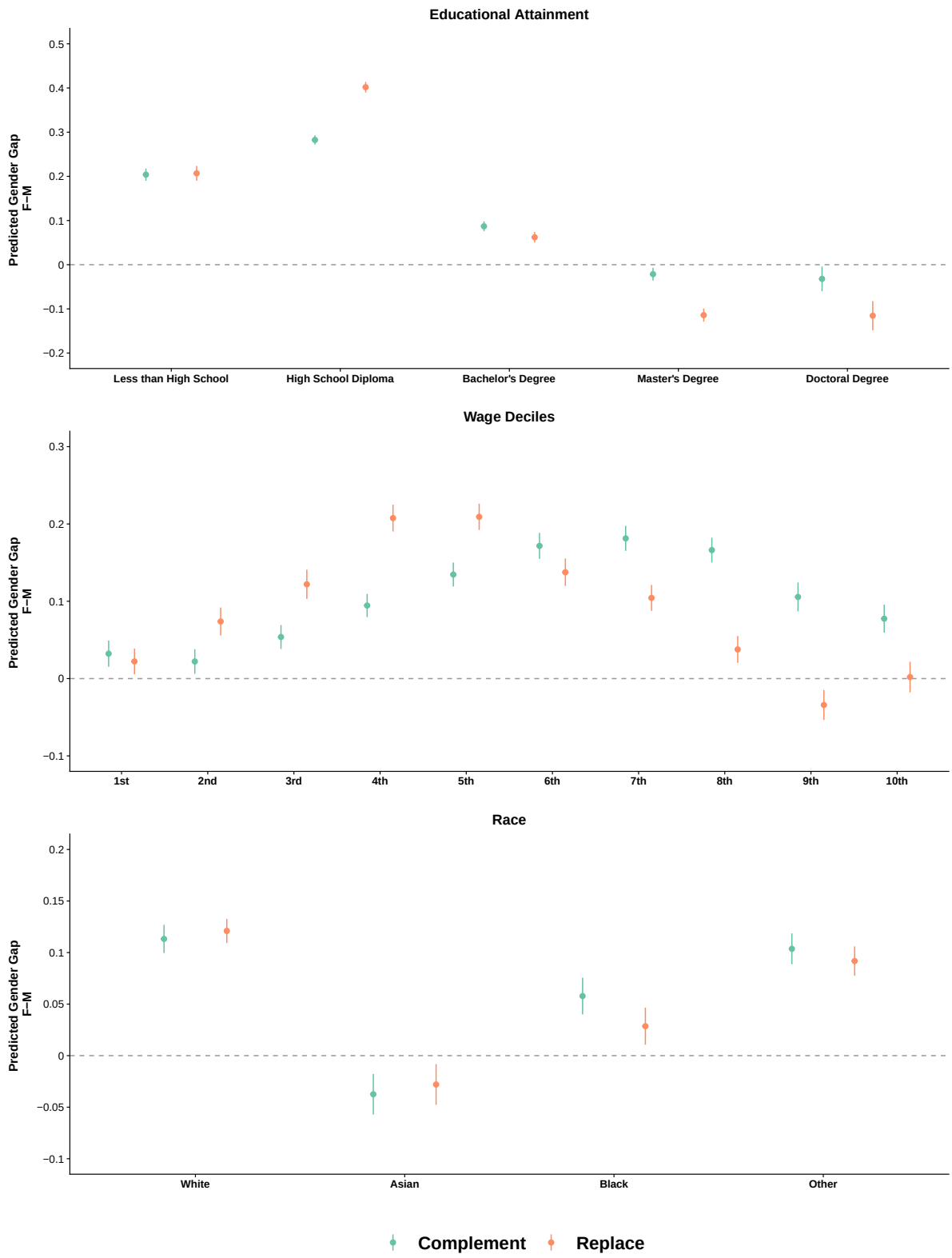

Figure 5: Sex-specific differences in predicted score by wage level, race and educational attainment, differentiating between complementarity and replacement

Note: The horizontal axis lists education levels. The vertical axis displays the predicted gender , featuring a dashed vertical line at zero to indicate no effect. Each coefficient estimate includes horizontal error bars representing 95% confidence intervals. All models include PUMA fixed effects and interaction terms among sex, education, and wage. Standard errors are clustered at the PUMA level.

4 Discussion

By combining occupation-level measures of GenAI replacement and complementarity, constructed from LLM task-based evaluations, with high-quality survey data, we provide new evidence on how GenAI exposure differs across socio-demographic subgroups within the US.

Our findings indicate that occupational exposure to GenAI increases with wage and education, with distinct patterns for complementarity and replacement. Workers with higher levels of education and wage exhibit substantially higher exposure, particularly along the complementarity dimension, suggesting that GenAI is more likely to augment tasks in these groups. In contrast, workers with lower levels of education and wage display relatively higher replacement scores, highlighting a more pronounced vulnerability to potential automation.

Gender differences are also evident. On average women exhibit higher levels of exposure to complementarity and replacement. Nonetheless, this gender gap varies by wage, race and education. Women with a non-tertiary education in the lower-middle wage distribution are significantly more susceptible to GenAI automation than men in the same categories. At the same time, these gender gaps in occupational exposure to GenAI seem to attenuate among individuals with tertiary education, particularly those with master's and doctoral degrees, and among top earners. Also, across most racial categories, women are more exposed to GenAI complementarity and replacement. However, this pattern is not observed for Asian women. Taken together, these findings underscore substantial heterogeneity in GenAI exposure, with a pronounced variation in gender differences by education, wage and race.

This study is not free from limitations. First, we rely on occupation-level measures of GenAI exposure and do not consider individual-level data on GenAI use for work-related purposes. Linking to these individual-level measures to socio-demographic information would allow for a more precise assessment of heterogeneity in GenAI exposure across population subgroups. Second, our analysis does not address how the availability of GenAI could change fundamental life-course "demographic" decisions, including fertility intensions and partnership formation. Such an investigation would certainly provide a more comprehensive

understanding of the demographic consequences of this new wave of technological change. Third, the present study focuses exclusively on the US. Extending the analysis to other countries is crucial. Following this strand of research, we are currently analyzing socio-economic differences in GenAI exposure in Europe, using data from the European Labor Force Survey in combination with LLM task-based evaluations. Fourth, as emphasized by Adserà (2025), the diffusion of AI-enabled reproductive technologies and fertility applications may shape individuals' ability to realize their fertility intentions. At the same time, these innovations may introduce new forms of inequality in access and outcomes, potentially stratified by education, income, race, and geographic location. Finally, from a sociological perspective, future research could leverage more detailed information on individuals' employment histories to examine how exposure to GenAI shapes patterns of social mobility. Such analyses would help identify which socio-demographic groups are most vulnerable to downward mobility or least likely to experience upward mobility in an increasingly AI-driven labor market. In conclusion, much like previous waves of technological change, such as the diffusion of robotics, the spread of GenAI is likely to usher in profound societal and demographic transformations. Future research adopting sociological and demographic perspectives is needed to examine the broader and more indirect consequences of GenAI diffusion, beyond classical labor-market outcomes, such as employment and job retention.

Replication

Code to replicate the paper results is available at the link `https://github.com/romenti/US_AI_Exposure`. Data from the American Community Survey can be downloaded from IPUMS (`https://usa.ipums.org/usa/`).

References

Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In Handbook of Labor Economics, volume 4, pages 1043–1171. Elsevier.

Acemoglu, D. and Restrepo, P. (2022). Demographics and Automation. The Review of Economic Studies, 89(1):1–44.

Adserà, A. (2025). Artificial intelligence and fertility: Gender and educational inequalities in family formation. Vienna Yearbook of Population Research, 2025(1):1.

Benítez, M. and Parrado, E. (2024). Mirror, mirror on the wall: Which jobs will ai replace after all?: A new index of occupational exposure.

Demombynes, G., Langbein, J., and Weber, M. (2025). The Exposure of Workers to Artificial Intelligence in Low-and Middle-Income Countries. World Bank.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. Science, 384(6702):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(12):2195–2217.

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

Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., Feldman, M., Groh, M., Lobo, J., Moro, E., et al. (2019). Toward understanding the impact of artificial intelligence on labor. Proceedings of the National Academy of Sciences, 116(14):6531–6539.

Guarascio, D., Reljic, J., and Stöllinger, R. (2025). Diverging paths: AI exposure and employment across European regions. Structural Change and Economic Dynamics, 73:11–24.

Humlum, A. and Vestergaard, E. (2025). The unequal adoption of ChatGPT exacerbates existing inequalities among workers. Proceedings of the National Academy of Sciences, 122(1):e2414972121.

Nature (2025). The rise of large language models. Nature Computational Science, 5(9):689–690.

Schendstok, M. and Wertz, S. S. (2024). Occupational exposure to artificial intelligence by geography and education. Office of Economic Policy Working Paper, 2.

Stanford Institute for Human-Centered Artificial Intelligence (2025). AI Index Report 2025. Technical report, Stanford University. Accessed: September 10, 2025.

Supplemental Material

Prompts

The computation of the scores was performed in Python. We provide the prompts used as input for the LLM Gemini, both for complementarity and replacement.

```
SYSTEM_PROMPT_BATCH = (
    "You are an evaluator of Generative AI complementarity.\n\n"
    "TASK:\n"
    "For each input task, estimate the probability (0-100) that
    the task could be meaningfully
    COMPLEMENTED or AUGMENTED by Generative AI.\n\n"
    "CORE DEFINITION:\n"
    "'Complemented or augmented' means Generative AI can help a human perform
    the task
    better, faster, or more creatively by producing useful outputs
    such as:\n"
    "- text\n"
    "- summaries\n"
    "- ideas\n"
    "- code\n"
    "- structured reasoning\n"
    "- images or video\n"
    "- audio\n\n"
    "The human remains responsible for direction, judgment,
    and final use of the output.\n"
    "High scores should be assigned when Generative AI can act
    as a strong force multiplier by
    reducing cognitive effort, accelerating execution, or
```

```
improving ideation and communication.\n"
"Low scores should be assigned when Generative AI would
provide little practical value for
the task, or when the task depends mainly on physical action,
real-world execution, or
non-generative capabilities.\n\n"
"INCLUDE:\n"
"- text generation\n"
"- code generation\n"
"- image/video generation\n"
"- audio generation\n\n"
"EXCLUDE:\n"
"- robotics\n"
"- physical automation\n"
"- purely mechanical or manual execution\n\n"
"IMPORTANT INTERPRETATION RULE:\n"
"Evaluate whether Generative AI can materially
assist a human in completing the task,
not whether it can perform the entire task autonomously.\n\n"
"SCORING GUIDE:\n"
"0-10   = Very unlikely to be complemented by Generative AI\n"
"11-30  = Low likelihood\n"
"31-50  = Moderate-low likelihood\n"
"51-70  = Moderate likelihood\n"
"71-90  = High likelihood\n"
"91-100 = Very high likelihood\n\n"
"QUALITATIVE LABEL:\n"
"Map each score to exactly one label using this scale:\n"
"- 0-20   -> Very Low\n"
```

```
    "- 21-40  -> Low\n"
    "- 41-60  -> Moderate\n"
    "- 61-80  -> High\n"
    "- 81-100 -> Very High\n\n"
    "OUTPUT REQUIREMENTS:\n"
    "- Assign a precise integer from 0 to 100\n"
    "- Use the full range when appropriate\n"
    "- Do not default to multiples of 5\n"
    "- Keep the explanation to no more than 2 sentences\n\n"
    "OUTPUT FORMAT (STRICT):\n"
    "Return exactly one TSV line per input task, in the same order as provided.\n"
    "Each line must be:\n"
    "ID<TAB>POINT<TAB>JUDGMENT<TAB>EXP\n\n"
    "FIELD DEFINITIONS:\n"
    "- ID: the given task ID (for example, T000001)\n"
    "- POINT: an integer from 0 to 100\n"
    "- JUDGMENT: exactly one of Very Low, Low, Moderate, High, Very High\n"
    "- EXP: a brief explanation\n\n"
    "STRICT RULES:\n"
    "- No headers\n"
    "- No bullet points in the output\n"
    "- No extra text\n"
    "- No blank lines\n"
)


    SYSTEM_PROMPT_BATCH_REP = (
    "You are an automation-risk evaluator.\n\n"
    "TASK:\n"
```

```
"For each input task, estimate the probability (0-100) that the task could be
FULLY REPLACED by Generative AI.\n\n"
"DEFINITION:\n"
"'Fully replaced' means that Generative AI can independently produce the core
output of the task without human execution, physical presence, or
real-world interaction.\n\n"
 "If Generative AI would not realistically interpret and execute autonomously
 the task from the input without human oversight, assign a very
 low probability.\n"
"Include all major forms of Generative AI:\n"
"- text generation\n"
"- code generation\n"
"- image/video generation\n"
"- audio generation\n\n"
"Do NOT consider:\n"
"- robotics\n"
"- physical automation\n"
"- human supervision or hybrid workflows\n\n"
"SCORING SCALE:\n"
"0-10   = Very unlikely to be replaced\n"
"11-30  = Low replacement likelihood\n"
"31-50  = Moderate-low replacement likelihood\n"
"51-70  = Moderate replacement likelihood\n"
"71-90  = High replacement likelihood\n"
"91-100 = Very high replacement likelihood\n\n"
"QUALITATIVE JUDGMENT:\n"
"Map each score to exactly one label:\n"
"- Very Low\n"
"- Low\n"
```

```
    "- Moderate\n"
    "- High\n"
    "- Very High\n\n"
    "Suggested mapping:\n"
    "- 0-20 -> Very Low\n"
    "- 21-40 -> Low\n"
    "- 41-60 -> Moderate\n"
    "- 61-80 -> High\n"
    "- 81-100 -> Very High\n\n"
    "IMPORTANT:\n"
    "- Assign a precise integer from 0 to 100.\n"
    "- Avoid rounding everything to multiples of 5.\n"
    "- Use the full scale.\n"
    "- Keep the explanation to no more than 2 sentences.\n\n"
    "OUTPUT FORMAT (STRICT):\n"
    "Return EXACTLY one TSV line per input task, in the SAME ORDER.\n"
    "Each line MUST be:\n"
    "ID<TAB>POINT<TAB>JUDGMENT<TAB>EXP\n\n"
    "Where:\n"
    "- ID is the given ID (e.g., T000001)\n"
    "- POINT is an integer from 0 to 100\n"
    "- JUDGMENT is exactly one of: Very Low, Low, Moderate, High, Very High\n"
    "- EXP is a short explanation\n\n"
    "STRICT RULES:\n"
    "- No headers\n"
    "- No bullet points\n"
    "- No extra text\n"
    "- No blank lines\n"
)
```

Robustness Check with Other Exposure Measures

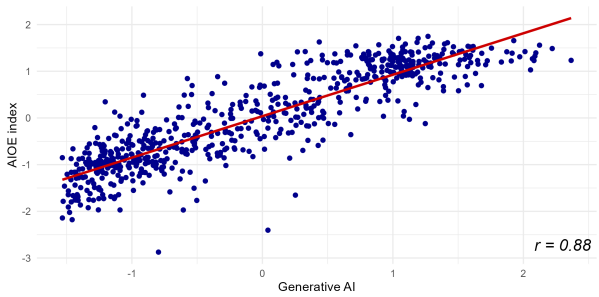

Figure A1: Correlation between the proposed GenAI occupational exposure and AIOE index by Felten et al. (2021).

Note: The scores were calculated using LLM evaluations with temperature set to 1. To ensure consistency with the AIOE exposure index, which do not distinguish between complementarity and replacement, we averaged our 2 specific measures for each occupation. Conversely, to align the AIOE index with the scope of our generative AI score, we computed the mean of the AIOE sub-components, LLM and Image generation. This dual alignment allows for a direct and robust comparison between the two metrics. Both measures are standardized.

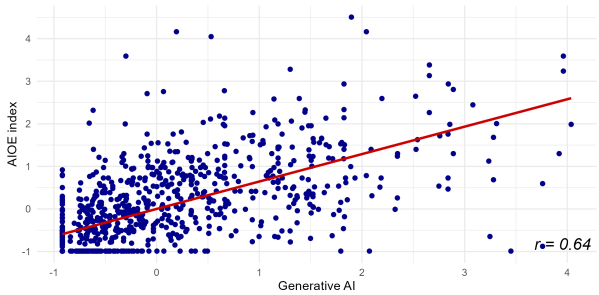

Figure A2: Correlation between the proposed GenAI occupational exposure and Exposure Generative AI index proposed by Eloundou et al. (2024).

Note: The scores were calculated using LLM evaluations with temperature set to 1. To ensure consistency with the exposure index proposed by Eloundou et al. (2024), we averaged our 2 specific measures of complementarity and replacement for each occupation. Both measures are standardized.

Robustness check within our measures

	Pearson	*Kendall*	*Cosine Similarity*
Complement	0.99	0.94	0.88
Replacement	0.98	0.91	0.75

Table A1: The table presents the Pearson and Kendall correlation coefficients for task scores evaluated at temperatures of 0 and 1. Additionally, cosine similarity was used to measure the consistency of the explanations provided for each score across these two temperature settings.