Labor Demand in the Age of Generative AI: Early Evidence from the U.S. Job Posting Data*

Yan Liu[†1], He Wang[‡1], and Shu Yu[§1]

[1]World Bank

October 30, 2025

Abstract: This paper examines the causal impact of generative artificial intelligence (GenAI) on U.S. labor demand using online job posting data. Exploiting ChatGPT's release in November 2022 as an exogenous shock, this paper applies difference-in-differences and event study designs to estimate GenAI's job displacement effects. Our identification strategy compares labor demand for occupations with high versus low AI-substitution vulnerability following ChatGPT's launch, conditioning on similar GenAI exposure levels to isolate substitution effects from complementary uses. Using 285 million job postings collected by Lightcast during 2018Q1-2025Q2, we find that between late 2022 and June 2025, the number of postings for occupations with above-median AI-substitution scores fell by an average of 12% relative to those with below-median scores. The effect intensified over time, rising from 6% in the first year after the launch to 18% by the third year. Losses were particularly acute for entry-level positions that require neither advanced degrees (18%) nor extensive experience (20%), as well as in administrative support (40%) and professional services (30%). While GenAI generates new occupations and enhances productivity, which may increase labor demand, early evidence suggests that some occupations may be less likely to be complemented by GenAI than others.

JEL codes: O33, J23, J21

Key words: Generative Artificial Intelligence, Technology Adoption, Labor Demand, Online Job Postings

*We would like to thank Aart Kraay, Daniel Lederman, Franziska Lieselotte Ohnsorge, Jonah Matthew Rexer, and participants at various seminars for helpful comments and suggestions. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

[†]Corresponding Author. Email: yanliu@worldbank.org

[‡]Email: hwang21@worldbank.org

[§]Email: syu2@worldbank.org

1 Introduction

Has Generative AI (GenAI) started displacing workers? And what kind of jobs face greater displacement risks? Since ChatGPT's debut in late 2022, the adoption of GenAI has spread rapidly across countries, sectors, and occupations (Liu and Wang 2024; Bick, Blandin, and Deming 2024; Bonney et al. 2024; Liu, Huang, and Wang 2025). The expanding capabilities of GenAI, coupled with its accelerating adoption, have reignited longstanding concerns about technology-driven job displacement, with warnings being raised by business leaders across industries in the U.S.–from IT to car manufacturing.[1] The fear of AI-driven job losses is not unfounded, with emerging evidence suggesting that freelancers are facing the sharpest initial impact.[2] Yet there is limited and inconclusive evidence on the job displacement effect of GenAI at the aggregate level.

This paper investigates the impact of GenAI on labor demand in the U.S. using the near-universe of online job posting data spanning 2018Q1 to 2025Q2.[3] We exploit the public release of ChatGPT in November 2022 as an exogenous shock and employ difference-in-differences (DiD) and event-study methodologies to identify GenAI's labor-displacement effects. Our analysis tracks how these effects evolve over time as GenAI capabilities improve and adoption deepens. It also examines the heterogeneous impacts of GenAI across education levels, experience requirements, and industries. The analysis draws on 285 million job postings from Lightcast, aggregated into a balanced panel of 6.8 million state–industry–occupation–quarter cells, providing a comprehensive view of how GenAI adoption is reshaping labor demand in the U.S. While this study focuses on the U.S. data, its findings on GenAI's impact on job displacement have broader implications for labor markets in other countries that are facing the same technology shock posed by GenAI.

[1] The warnings from industry leaders have been stark. In May 2025, Anthropic CEO Dario Amodei predicted that half of all entry-level jobs could disappear within one to five years, potentially driving U.S. unemployment 10-20 (link) A month later, Amazon CEO Andy Jassy declared that as the company increases adoption of GenAI and "AI agents," it "will need fewer people" in certain jobs and expects a shrinkage of its total corporate workforce in the next few years due to AI-driven efficiency gains (link). The CEO of Ford Motors suggested that AI will replace half of the white-collar jobs in the U.S. (link).

[2] See, for instance, Qiao, Rui, and Xiong (2023), Hui, Reshef, and Zhou (2024), Brynjolfsson, Chandar, and Chen (2025), Demirci, Hannane, and Zhu (2025), and Teutloff et al. (2025).

[3] The U.S. provides an especially strong empirical setting for this analysis. First, it leads global GenAI development, with major models (ChatGPT, Claude, Gemini) originating from U.S. firms. This proximity facilitates earlier, faster domestic labor market diffusion. Second, the U.S. labor market is large, dynamic, and well-documented, offering an ideal testing ground for new technology effects before broader diffusion. Third, U.S. Lightcast data has been demonstrated to provide superior coverage and consistency. It has been extensively used in prior research on automation and AI (e.g., Acemoglu and Restrepo (2020), Acemoglu et al. (2022), and Bonfiglioli et al. (2025)), making it particularly suitable for identifying emerging GenAI-driven labor demand shifts.

Our empirical framework combines two complementary dimensions that jointly determine how GenAI affects labor demand across occupations: GenAI exposure, which measures the theoretical technical applicability of GenAI, and AI-substitution vulnerability, which captures the practical likelihood that employers replace workers with AI. GenAI exposure reflects the extent to which AI is useful and applicable to specific tasks and occupations, but by itself it does not predict realized labor demand effects since it conflates automation and augmentation. To address this limitation, we incorporate an AI-substitution score adapted from Pizzinelli et al. (2023)–the inverse of their AI complementarity measure—constructed along six occupational dimensions: communication, responsibility, physical conditions, criticality, routine, and skill requirements. This measure accounts for contextual and non-technical factors such as societal preferences and regulatory constraints that shape actual displacement risk. Unlike prior studies that rely on exposure indices alone to define treatment groups, our strategy compares changes in job postings across occupations with above- versus below-median substitution vulnerability, conditional on similar GenAI exposure levels. This two-dimensional framework enables us to isolate the labor-displacement channel by holding constant differences in technical applicability and pre-existing automation potential. [4]

Our results demonstrate a large, statistically significant, and intensifying negative impact of GenAI on job postings for more substitutable occupations. In the U.S., we estimate that by mid-2025, postings in occupations with above-median AI substitution scores declined by an average of 12% relative to those with below-median scores, conditional on comparable levels of GenAI exposure. Event study analysis further confirms no significant difference in job posting trends between the two groups, consistent with the conditional parallel trends (CPT) assumption (Baker et al. 2025). Divergence appears only after ChatGPT's introduction, underscoring its causal impact. The displacement effect grew progressively stronger over time, increasing from 6% in the first year post-launch to 18% in the third year (by June 2025). The findings are robust to placebo tests where treatment is randomly assigned to occupations, alternative AI substitution measures, and different AI exposure measures as controls. Heterogeneity analyses reveal particularly pronounced negative effects for entry-level positions, which require neither advanced degrees (master's, professional, and doctoral degrees) nor extensive work experience (6 years and above), and for positions in professional services and administrative support industries.

[4] A similar two-dimensional approach has also been applied in World Bank (2025a) and World Bank (2025b).

This paper makes several contributions to the emerging literature on GenAI and labor demand.[5] First, to our knowledge, this is the first study to identify the causal impact of GenAI on labor demand for long-term, stable jobs using large-scale online job posting data. Existing literature on this topic has primarily focused on the gig economy, with studies of freelancing platforms showing that GenAI reduces both labor demand and wages for highly automatable tasks, such as writing and translation (Demirci, Hannane, and Zhu 2025; Hui, Reshef, and Zhou 2024; Teutloff et al. 2025; Qiao, Rui, and Xiong 2023). Hui, Reshef, and Zhou (2024) find that top freelancers are disproportionately affected, likely because of their higher wage levels. In contrast, our research shifts the focus to formal, full-time positions that represent the vast majority of the labor market and are the primary source of income for most workers. A closely related study by Brynjolfsson, Chandar, and Chen (2025) uses employment data and finds a 13% relative decline in early-career worker employment within highly exposed occupations. We use online job posting data as they are more forward-looking and capture early signals of shifts in labor demand better than employment data.[6]

Second, our extended sample allows us to track the dynamic impact of GenAI over time. Previous studies using job posting data, such as Hampole et al. (2025) and Eisfeldt et al. (2023), were constrained by limited temporal coverage or data that ended in 2023, preventing the capture of evolving technological progress and diffusion effects. Our 2018-2025 dataset enables tracking how GenAI's negative effects intensified as capabilities advanced and adoption accelerated.

Third, our identification strategy isolates the displacement effect of GenAI by combining measures of occupations' GenAI applicability (GenAI exposure) with their practical vulnerability to AI substitution. Specifically, we compare how labor demand evolves for occupations with above- versus below-median substitution potential, conditional on the same level of GenAI exposure. This design allows us to distinguish GenAI's causal impact from differential hiring trends that arise solely from variation in technical applicability. While a few papers try to distinguish between automation

[5]The body of empirical literature on the labor market impacts of generative AI (GenAI) is expanding rapidly. While some experimental studies have documented substantial productivity gains from GenAI adoption (Brynjolfsson, Li, and Raymond 2025; Cui et al. 2024; Noy and Zhang 2023; Dell'Acqua et al. 2023), others have found mixed or minimal effects on wages and working hours (Jiang et al. 2025; Humlum and Vestergaard 2025; Chandar 2025; Hampole et al. 2025).

[6]Moreover, the ADP payroll data used in Brynjolfsson, Chandar, and Chen (2025) cover less than 15% of the U.S. workforce, while our online job posting data is more representative, consistently capturing 60% of the job vacancies in the U.S. since 2010.

and augmentation effects using task-level aggregation (Eisfeldt et al. 2023; Brynjolfsson, Chandar, and Chen 2025), our approach – building on Pizzinelli et al. (2023) – relies on work-context features and occupational skill requirements to capture displacement risk in reality. Existing methods that rely on task exposure scores, standard deviations of task-level scores, or subjective ratings derived from user interactions with GenAI tools do not fully capture the linkages between tasks and regulatory and societal factors. As a result, these methods may overestimate the automation potential of GenAI. Our substitution measure incorporates regulatory, societal, and contextual factors that influence GenAI adoption, providing a more realistic assessment of where GenAI is likely to substitute for labor in the short to medium term.

Fourth, our heterogeneity analysis offers novel insights into the distributional consequences of GenAI adoption. We show that the negative effects are disproportionately concentrated among entry-level and less-skilled workers, while more experienced workers are less adversely affected. The negative impacts are also more pronounced in administrative and support services, professional services, and the management of companies. These results are consistent with recent evidence (Brynjolfsson, Chandar, and Chen 2025; Teutloff et al. 2025), which shows that AI disproportionately reduces demand for entry-level and lower-skilled workers. Our findings underscore the necessity to consider targeted support to mitigate the potential effects of the technology shock on vulnerable groups.

The remainder of this paper is organized as follows. Section 2 describes the data used in the analysis, clarifies key concepts such as AI exposure, augmentation, automation, and presents descriptive evidence. Section 3 outlines the empirical strategy. Section 4 reports the baseline results. Section 5 conducts a series of robustness checks. Section 6 examines heterogeneous displacement effects across worker characteristics and industries. Section 7 concludes, discusses policy implications, and sheds light on potential avenues for future research.

2 Data and Descriptive Evidence

In this section, we begin by introducing the Lightcast job posting data used in our analysis. We then outline the AI exposure and substitution indices, clarifying key conceptual distinctions relevant to our framework. Lastly, we present descriptive evidence on how labor demand has diverged between

occupations with high and low AI-substitution vulnerability.

2.1 Job Posting Data from LightCast

In this paper, we use the real-time job posting data provided by Lightcast (previously Burning Glass) to capture labor demand in the U.S. Lightcast collects data from company websites and online job boards around the world.[7] It processes the raw data with a deduplication algorithm and further uses algorithms to clean and extract key information from detailed job descriptions. It reports information on posting dates, job location, company name, job title, hiring company's industry (at the NAICS 2-digit level), occupation (at the ISCO-08 4-digit level), and education and experience requirements. Existing studies suggest that job posting data offers more timely and more granular insights into employers' demand-side behavior than standard labor force surveys or job openings surveys do, making it more suitable for analyzing labor demand in response to external shocks.[8]

In the U.S., Lightcast data covers over 435 million job postings since 2010. As shown in Figure 1, Lightcast consistently captures around 60% of total vacancies recorded in Job Openings and Labor Turnover Survey (JOLTS) and closely tracks the evolution of overall vacancies.[9] Moreover, past studies show that the occupational and industry composition of Lightcast data is closely aligned with those found in Occupational Employment Statistics (OES), making the data representative enough for the purpose of our study. To capture labor demand dynamics before and after the release of ChatGPT in November 2022, we use high-frequency daily job-posting data from January 1, 2018 through June 30, 2025, which allows us to establish a long pre-treatment trend and observe post-release developments; the data are aggregated to the quarterly level for analysis.

[7] In total, the dataset covers over 2.5 billion job postings across more than 150 economies since 2010.

[8] See, for instance, Hershbein and Kahn (2018), Deming and Noray (2020), Acemoglu et al. (2022), Forsythe et al. (2022), and Braxton and Taska (2023).

[9] This is in line with findings of past studies, such as Hershbein and Kahn (2018) and Acemoglu et al. (2022).

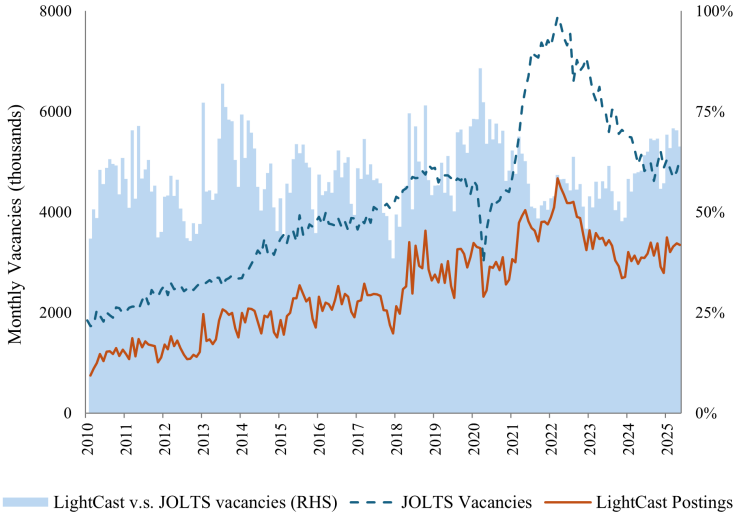


Figure 1: Comparing Lightcast Data with Job Openings and Labor Turnover Survey

Note: The figure shows the number of job openings recorded by Lightcast and the U.S. Job Openings and Labor Turnover Survey (JOLTS) from 2010 to June 2025. Lightcast closely tracks the evolution of overall U.S. vacancies and captures approximately 60% of total vacancies.

Table A1 shows the number of job postings by occupations, industries, as well as the education and experience requirements. During our sample period, half of the job postings are for professionals (31%) and technicians (19%), with managers (11%) and sales workers (16%) together accounting for more than a quarter. Within the 80% of postings with sectoral information, postings concentrated in healthcare and social assistance (12.4%), administrative and support and waste management (12.2%), retail trade (9.0%), professional, scientific and technical services (9.0%), and manufacturing (6.1%). Meanwhile, the four sectors with the fewest postings–agriculture, mining, company management services, and utilities–jointly account for only 1.1% of all postings.

The majority (59%) of job postings do not specify an educational requirement. Of those that do, one quarter of all postings require at least a high school diploma, while 10% and 6.4% require a college degree and advanced degrees (including master's, professional, and doctoral degrees), respectively. Similarly, most postings (55%) do not state an experience requirement. For those that do, 24% seek candidates with 0-2 years of experience, 15% look for 3-5 years, and 6% require at least 6 years. These detailed job requirements allow us to examine the heterogeneous effects of GenAI across sectors, education, and experience groups in Section 6.

2.2 GenAI Exposure and AI Substitution Measurement

As mentioned previously, two complementary dimensions jointly determine how GenAI affects labor demand across occupations: GenAI exposure reflects where GenAI may affect jobs, and AI-substitution vulnerability captures the probability of an occupation being displaced by AI in reality based on inherent occupational characteristics, societal preferences, and regulatory frameworks.

Researchers have developed various occupation-level AI exposure indices to estimate which tasks and occupations are most likely to be affected by AI (Frey and Osborne 2017; Felten, Raj, and Seamans 2018; Felten, Raj, and Seamans 2019; Webb 2019; Felten, Raj, and Seamans 2021; Pizzinelli et al. 2023), building on the task-based approach established in Acemoglu and Autor (2011), Autor (2015), and Acemoglu and Restrepo (2019).[10] More recent studies adapted these frameworks to focus on GenAI exposure more specifically.[11]

Researchers typically drew information from the O*NET database,[12] and asked human experts or GenAI systems to judge the extent to which each task can be performed or aided by AI. Then they aggregate the task-level score to generate an occupation-level exposure score (Felten, Raj, and Seamans 2023; Eloundou et al. 2024; Eisfeldt et al. 2023; Benitez-Rueda and Parrado 2024; Gmyrek, Berg, and Bescond 2023; Gmyrek et al. 2025). Some go further by using real-world user interaction data. For example, the Anthropic Economic Index (Handa et al. 2025) and Microsoft's AI applicability scores (Tomlinson et al. 2025) analyze millions of conversations with GenAI tools to identify where usage is most concentrated. Some indices made attempts to decompose automation and augmentation scores by task and occupation.

This paper mainly uses the GenAI exposure score developed by Gmyrek, Berg, and Bescond (2023) (GBB) for the following reasons. First, the GBB score was one of the first widely used GenAI exposure indices. Some other frequently used AI exposure measures, such as Felten, Raj, and Seamans (2021) and Pizzinelli et al. (2023), are not specifically about GenAI. Second, the GBB

[10] The task-based model is widely used to analyze how technology affects labor demand. This framework emphasizes that technologies such as AI affect labor demand not directly at the occupation level but through their impact on the tasks that make up occupations. Each occupation consists of a bundle of tasks. Technology can automate or augment existing tasks and create new tasks. The balance between these forces determines whether technology reduces or increases aggregate labor demand.

[11] See, for instance,Felten, Raj, and Seamans (2023), Eloundou et al. (2024), Benitez-Rueda and Parrado (2024), Gmyrek, Berg, and Bescond (2023), Gmyrek et al. (2025), Handa et al. (2025), and Tomlinson et al. (2025).

[12] The database includes a rich set of variables that describe work and worker characteristics (including skill requirements) associated with roughly 1,000 U.S. occupations

score was released soon after ChatGPT's public launch and remains uncontaminated by expanding model capabilities and rising adoption of GenAI, unlike its recently updated version (i.e., Gmyrek et al. 2025). Third, most existing AI exposure measures are highly positively correlated with each other. We use these alternative AI exposure indices as controls in Section 5.2 to test the robustness of our baseline results.

We identify treatment occupations using an AI-substitution score derived by reversing the AI-complementarity index of Pizzinelli et al. (2023). Their measure is constructed from five work-context features—communication, responsibility, physical conditions, criticality, and routine—and a composite skills dimension based on O*NET. Occupations that score high on these dimensions (except for "routine") are more likely to be complemented, rather than displaced, by AI. Reversing this scale allows us to capture the likelihood that employers actually substitute workers with AI, taking into account not only technical feasibility but also occupational characteristics, workplace context, and societal preferences. While the original measure was not designed specifically for GenAI, its focus on non-task-related job attributes makes it well-suited for our analysis. In Section 5, we further decompose this score to examine which occupational features protect against displacement and compare it with alternative substitution measures such as interpersonal-skill intensity.

We prefer this holistic substitution measure over task-based automation or augmentation indices for several reasons. First, occupations are not fixed bundles of tasks but dynamic, evolving sets of activities that differ across firms, industries, and countries. Task taxonomies such as O*NET often omit critical but intangible aspects of work—for example, while O*NET details the physical activities of truck drivers, it does not fully capture tasks such as coordinating with dispatchers or navigating disruptions. Moreover, where tasks are interdependent, automating or augmenting one component does not necessarily imply that the entire occupation can be automated or augmented. It remains highly uncertain how new tasks will emerge, how they will be distributed across existing occupations, or whether they will give rise to entirely new occupations.

Second, beyond these definitional and measurement challenges, aggregating task-level effects to labor demand at the occupational level requires tracing through broader economic mechanisms. Labor demand is derived from the demand for goods and services that an occupation produces. Product demand itself is influenced by multiple factors, including the emergence of AI. The effects of technological shocks at the task level are influenced by complex general equilibrium effects, which

in turn shape the ripple effect through various occupations and industries.

Third, the speed and extent of displacement depend on non-technical factors. Occupations with high tolerance for error (e.g., data-entry clerks) are more easily substituted than those with high responsibility or minimal room for error (e.g., surgeons or drivers). Jobs that require social interaction, empathy, and problem-solving—such as teaching—are also more resistant. Moreover, regulatory constraints, societal trust, and consumer expectations strongly influence adoption, none of which are reflected in task-level exposure scores. These limitations explain why task-based automation and augmentation scores often fail to predict occupation-level outcomes. In contrast, our substitution score provides a more accurate measure of where GenAI is likely to replace labor.

Figure 2 illustrates the complementarity of the two dimensions by plotting occupations' GenAI exposure scores (technical applicability) against AI substitution scores (vulnerability based on occupational characteristics). Occupations with identical exposure exhibit widely varying substitution potential. In the top-right quadrant, occupations with both high exposure and high substitution scores—including general secretaries, application programmers, and statistical, finance, and insurance clerks—face substantial automation risk. Conversely, the bottom-right quadrant contains occupations with high exposure but low substitution potential, such as life science technicians and university and higher education teachers, which are likely to be augmented rather than directly replaced by AI. The top-left quadrant reveals occupations like building structure cleaners and weaving and knitting machine operators that have high substitution potential despite low GenAI exposure. Finally, the bottom-left quadrant includes occupations such as police officers and dentists that exhibit both low GenAI exposure and minimal substitution risk.

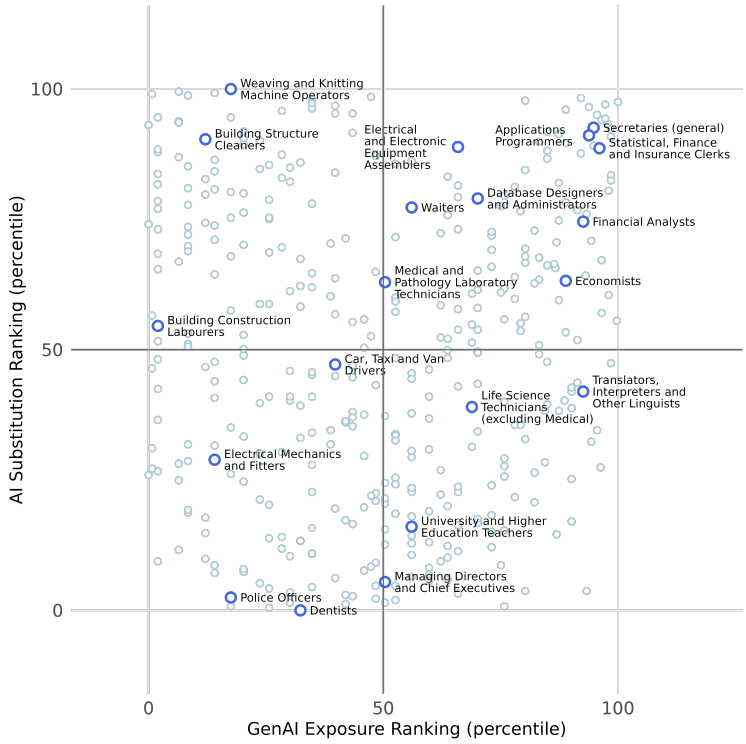


Figure 2: AI Substitution and GenAI Exposure Across Occupations

Note: The x-axis presents the percentile ranking of the GenAI exposure index from Gmyrek, Berg, and Bescond (2023) (GBB). The y-axis presents the percentile ranking of the AI-substitution score, which is defined as the inverse of the AI-complementarity measure in Pizzinelli et al. (2023). The AI-complementarity measure is constructed along six occupational dimensions: communication, responsibility, physical conditions, criticality, routine, and skill requirements. Typical occupations are highlighted.

A key concern is that occupations with high and low AI-substitution vulnerability may have followed different hiring trajectories even before the launch of ChatGPT, which could undermine the parallel-trends assumption required for our DiD design. This risk is especially pronounced if occupations with high AI exposure are compared against those with low exposure, since their underlying adoption potential differs. To address this, we condition our analysis on occupations with similar GenAI exposure scores. By holding exposure constant, we ensure that treatment and control groups are equally likely to adopt GenAI, thereby isolating the displacement effect from broader differences in technology exposure (see Section 3 for details).

2.3 Descriptive Evidence

Figure 3 plots online job posting trends for occupations above and below the median AI-substitution score, with values normalized to 1 in the quarter ChatGPT was launched (2022Q4). Prior to that

point, postings for the two groups moved in parallel, but their trajectories diverged sharply afterward.[13] Since ChatGPT's launch, postings for high-substitution occupations have fallen more steeply than those for low-substitution ones. Although normalization dampens the measured magnitude, the divergence is clearly visible.[14]

The timing of this divergence alone does not establish a causal link to GenAI. Other factors, such as the rising interest rate in 2022 and the post-pandemic normalization of white-collar hiring, may also contribute to this divergence. To distinguish GenAI's effect from these confounding trends, the subsequent section presents our empirical strategy, outlining the DiD and event study specifications.

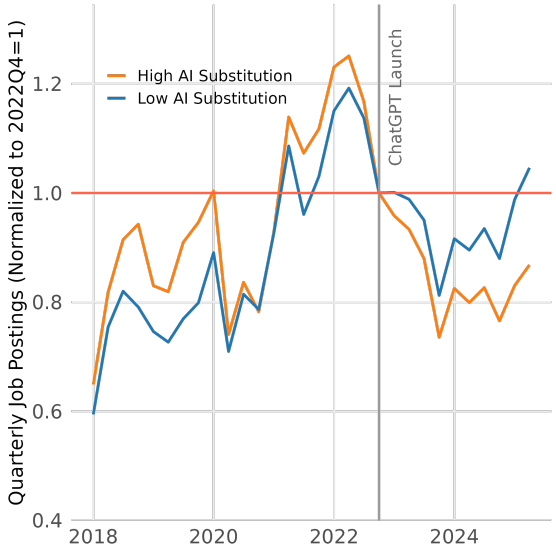


Figure 3: Normalized job postings trends for occupations with high vs. low AI-substitution score.

Note: This figure shows online job posting trends in LightCast for occupations that fall above and below the median AI-substitution score. The posting values have been normalized to 1 in the quarter of ChatGPT's launch (2022Q4). The AI-substitution score is defined as the inverse of the AI-complementarity measure in Pizzinelli et al. (2023), constructed along six occupational dimensions: communication, responsibility, physical conditions, criticality, routine, and skill requirements. The raw trends of job postings are reported in Appendix Table A1.

[13] Figure A1 shows the raw quarterly posting counts and confirms that both groups were at similar levels before 2022Q4.

[14] Figure A2 highlights this contrast for two occupations: general secretaries (high exposure, high substitution) and dentists (low exposure, low substitution). Both saw steady demand growth from 2018 to 2022, but their trajectories split after ChatGPT's launch: demand for dentists rose by nearly 50%, while postings for secretaries dropped by roughly 50%.

3 Empirical Specification

Our analysis centers on occupation-level treatment variation, evaluating GenAI's impact before and after ChatGPT's public release in November 2022. We aggregate the number of job postings by occupation and quarter across local labor markets, defined as the cross-section of 51 state-level units, 20 industries at the 2-digit NAICS level, and nearly 400 occupations at the 4-digit ISCO level. Our unit of analysis is chosen to align with the established theoretical models (Topel 1986; Moretti 2010), which suggest that labor demand is shaped by local market conditions and is closely related to empirical work examining local labor markets (e.g., Hershbein and Stuart (2024)). Compared to studies that examine the impact of AI at the firm level (Babina et al. 2023; Acemoglu et al. 2022), our approach has two key advantages. First, it avoids potential selection bias caused by firms entering or exiting the market during our sample period. Second, it addresses the issue of excessive zeros, since firms may not hire for every occupation every year. In general, aggregating job postings to the local market level smooths this firm-specific volatility and provides more reliable estimates of systematic labor demand shifts.

Baseline Specification. Our baseline identification strategy employs a difference-in-differences (DiD) framework that exploits the variation in occupations' vulnerability to AI substitution to define the treatment group and treats the release of ChatGPT in November 2022 as an exogenous shock. While the timing of the launch was largely unexpected, the launch of ChatGPT was followed by rapid adoption, making it necessary to control for the GenAI exposure levels in our model specification. Therefore, our baseline model specification takes the following form:

$$arsinh(Y_{siot}) = \alpha D_o \times \mathbf{1}_t^{Post} + \chi_{j(o)t} + \theta_{sit} + \gamma_{sio} + u_{siot}, \tag{1}$$

where Y_{siot} represents the number of job postings in state s, industry i, occupation o, and quarter t. Here, we include zero counts for state-industry-occupation-quarter cells with no observed job postings.[15] We apply the inverse hyperbolic sine ($arsinh$) transformation to address the prevalence of zero values while preserving interpretability. The transformation approximates the natural logarithm for large values but remains well-defined at zero, allowing coefficients to be interpreted as

[15] We retain state-industry-occupation cells with at least one posting and fill zeros for the rest of the quarters, thereby dropping cells with no postings in the sample period.

percent changes in job postings, similar to a logarithmic transformation.[16]

The key explanatory variable is the interaction term $D_o \times \mathbf{1}_t^{Post}$. The treatment indicator, D_o, is a binary variable that equals one for occupations with above-median AI substitution scores (i.e., the treatment group) and zero otherwise. The post-treatment period indicator, $\mathbf{1}_t^{Post}$, is also a dummy variable that equals one for quarters after ChatGPT's launch (including 2022Q4, the quarter when ChatGPT launched) and zero for earlier periods. The coefficient of interest, α, identifies the differential impact of ChatGPT's release on job postings for occupations with high versus low AI substitution vulnerability.

We include three sets of fixed effects as controls. First, state-industry-quarter fixed effects (θ_{sit}) are added to absorb all time-varying factors that can affect labor demand within a local labor market at the state-industry level.[17] Second, we control for state-industry-occupation fixed effects (γ_{sio}), which capture all time-invariant characteristics of specific occupation-location-industry triplets. These fixed effects capture persistent differences in baseline job posting levels, local comparative advantages, and other unobserved factors that influence the geographic and sectoral distribution of jobs.

Lastly, we control for "GenAI exposure decile-by-quarter fixed effects" at the occupation level (i.e., $\chi_{j(o)t}$, where $j(o)$ denotes the decile ranking of occupation o based on our preferred GenAI exposure index). GenAI's impact on labor demand may differ across occupations based on their GenAI exposure scores. If not properly controlled for, this variation can lead to violations of the parallel trends assumption that is essential to the DiD approach. Therefore, we rely on the conditional parallel trend (CPT) assumption proposed by Baker et al. (2025), where we compare occupations with high- and low-AI-substitution-scores within the same decile ranking of GenAI exposure index (i.e., the GBB score). These controls absorb common trends affecting all occupations with similar levels of overall GenAI exposure, such as the AI adoption wave and changes in demand for AI-exposed work. With this specification, the parameter α identifies the AI-substitution effect

[16] Specifically, $arsinh(X) = \ln\left(X + \sqrt{X^2+1}\right)$ $arsinh$ is zero when X is zero, and it is close to $\ln(X)$, when X is large. This approach is widely used for count data, including for job postings as in Acemoglu et al. (2022). The $arsinh$ function, similar to ln(1+X), is a well-defined function, which is differentiable at zero with derivatives approximating ln(X) for large X. Nonetheless, recent work (e.g., Chen and Roth (2024)) shows that log-type transformations, including $arsinh$, can lead to unit-dependent and hard-to-interpret effects when handling zeros. Despite this, with large sample size and high-dimensional fixed effects in our specification, $arsinh$ remains a feasible way to approximate treatment effects, serving as a suitable transformation for capturing proportional changes in the dependent variable.

[17] Such factors could include local industry-specific business cycles, state policy changes, and seasonal employment patterns

of GenAI and separates it from the overall exposure effect.

Dynamic Effects. As GenAI model capabilities and adoption evolve over time, we modify the baseline specification and use the following event study design to estimate the dynamic treatment effects of GenAI:

$$arsinh(Y_{siot}) = \sum_{k \neq -1} \alpha_k D_o \times \mathbf{1}_t^k + \chi_{j(o)t} + \theta_{sit} + \gamma_{sio} + u_{siot}, \tag{2}$$

where $\mathbf{1}_t^k$ is a dummy indicating the k-th quarter after ChatGPT's launch (negative values indicating the pre-ChatGPT periods), with the quarter before the launch ($k = -1$) serving as the omitted period as the reference. The coefficients $\{\alpha_k\}$ trace out the dynamic treatment effect. To evaluate the annual effects, we also estimate a model that combines every four quarters after the launch of ChatGPT into one unit, while using all pre-ChatGPT quarters as the reference.[18]

Throughout our analysis, we weight observations by the number of job postings in each state-industry-occupation cell in 2017 to ensure our estimates reflect meaningful labor market segments. Standard errors are clustered at the occupation-quarter level to account for potential serial correlation in treatment effects and common shocks within occupational categories over time.[19]

4 Empirical Results

This section presents our empirical findings in two parts. First, we report baseline DiD results on GenAI's labor displacement effects. Second, we present event study results that both support the validity of our conditional parallel trends (CPT) assumption and reveal the quarterly dynamics of GenAI's impact.

4.1 Baseline Results

Table 1 reports our main DiD estimates of GenAI's impact on labor demand. We report results across three progressively restrictive specifications that add fixed effects to address potential con-

[18] This approach takes the following form:

$$arsinh(Y_{siot}) = \alpha_1 D_o \times \mathbf{1}_t^{Post\,1yr} + \alpha_2 D_o \times \mathbf{1}_t^{Post\,2yr} + \alpha_3 D_o \times \mathbf{1}_t^{Post\,3yr} + \chi_{j(o)t} + \theta_{sit} + \gamma_{sio} + u_{siot} \tag{3}$$

[19] As robustness checks, Appendix Figures A3 and A4 present results without weighting and with clustering at the occupation level, respectively. They are largely consistent with the baseline results.

founding factors. The first three columns follow Equation (1) and show the average difference in job postings between the treatment group and the control group after the launch of ChatGPT (i.e., from the last quarter in 2022 till mid-2025). The treatment group captures occupations with above-median AI substitution score, while the control group consists of occupations with below-median AI-substitution score. The last three columns display the dynamic impacts in the first, second, and third years after ChatGPT's launch, as specified in Equation (3).

Columns (1) and (4) present estimates only controlling for state-industry-occupation fixed effects and quarter fixed effects. The former absorb all time-invariant characteristics of specific occupation-location-industry combinations, while the latter control for time-varying factors (such as macroeconomic shocks, pandemic effects, and interest rate hikes). According to this specification, ChatGPT's launch resulted in a 23% average reduction in job postings for occupations with relatively higher substitution vulnerability from late 2022 through June 2025 (Column (1)). The dynamic estimates in Column (4) reveal that this displacement effect intensified over time, growing from 16% in the first year post-launch to 25% in the second year, and 31% in the third year.

Columns (2) and (5) replace aggregate time fixed effects with state-industry-year-quarter fixed effects. This stricter specification controls for all time-varying factors that affect labor demand uniformly across occupations within a specific state-industry. While these additional controls attenuate our estimates somewhat, the negative impact on high-substitution occupations remains economically and statistically significant at 18% on average. The dynamic pattern persists, with effects growing from 12% in year one to 20% in year two and 25% in year three (Column (5)).

Our preferred specification, shown in Columns (3) and (6), adds "GenAI exposure decile-by-quarter fixed effects". This approach allows us to compare occupations with above-median AI substitution scores to those with below-median scores, all within the same level of overall GenAI exposure. This method, using the GenAI exposure score from Gmyrek, Berg, and Bescond (2023), provides a more precise estimate of the displacement effect on labor demand.

Under this preferred specification, ChatGPT's launch reduced job postings for occupations with above-median AI substitution scores significantly by 12% on average. Similarly, this effect intensified over time, rising from 6% in the first year to 14% in the second year and 18% in the third year. We use the specifications from Columns (3) and (6) as the baseline for all subsequent analyses.

Table 1: ChatGPT effect on labor demand

Dependent variable	arsinh(# job postings)					
	(1)	(2)	(3)	(4)	(5)	(6)
High-substitution occ. × Post ChatGPT	-0.232***	-0.183***	-0.122***			
	(0.066)	(0.05)	(0.036)			
High-substitution occ. × Post 1st year				-0.155*	-0.121*	-0.064*
				(0.064)	(0.047)	(0.031)
High-substitution occ. × Post 2nd year				-0.252***	-0.197***	-0.135**
				(0.072)	(0.055)	(0.041)
High-substitution occ. × Post 3rd year				-0.310***	-0.247***	-0.182***
				(0.064)	(0.051)	(0.04)
Quarter FE	Yes			Yes		
State × Industry × Occupation FE	Yes	Yes	Yes	Yes	Yes	Yes
State × Industry × Quarter FE		Yes	Yes		Yes	Yes
GenAI exp. Decile × Quarter FE			Yes			Yes
Num.Obs.	6,753,930	6,753,930	6,753,930	6,753,930	6,753,930	6,753,930
R squared	0.959	0.970	0.971	0.959	0.970	0.971

Note: Robust standard errors clustered at the occupation-quarter level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The regression unit is at the state-industry-occupation-quarter level, including cells with zero counts. The dependent variable is the inverse hyperbolic sine (arsinh) transformation of job postings, which approximates the logarithmic transformation and equals zero when job postings are zero, allowing coefficients to be interpreted as approximate log changes in job postings. "High-substitution occ." is a dummy indicating the treatment group, defined as occupations with above-median AI substitution score (or the inverse of the AI-complementary score in Pizzinelli et al. (2023)). "Post ChatGPT" is a dummy variable equal to 1 for quarters after ChatGPT's launch, including 2022Q4, the launch quarter. "Post 1st year", "Post 2nd year", and "Post 3rd year" refers to 2022Q4-2023Q3, 2023Q4-2024Q3, and 2024Q4-2025Q2, respectively. "GenAI exp. Decile" is a series of dummies based on the decile ranking of GenAI exposure index from Gmyrek, Berg, and Bescond (2023).

4.2 Event Study Results

Figure 4 presents event study estimates based on the specification from Equation (2). The figure plots quarterly gaps in job postings of occupations with above-median AI substitution scores relative to those with below-median scores, after controlling for the three sets of fixed effects (as in Columns (3) and (6)). The blue dots are the estimated gaps over our sample period, with the spikes capturing the corresponding 95% confidence intervals. The results confirm that the impacts shown in Columns (3) and (6) of Table 1 are causal and demonstrate the dynamic evolution of GenAI's labor market impact in detail.

Pre-Treatment Validation. Figure 4 shows that prior to the launch of ChatGPT, the estimates fluctuate around zero with no systematic trend, indicating that high- and low-substitution occupations exhibited similar labor demand trajectories. This visual evidence is corroborated by a formal pre-trend test (Wald statistic = 0.83, p-value = 0.66), which fails to reject the null hypothesis of common pre-trends. This validates our empirical identification strategy, which involves choosing occupations with above-median and below-median AI substitution scores as the treatment

and control groups, respectively, within each decile of AI exposure levels.

Post-Treatment Dynamics. It is also notable from Figure 4 that the treatment effect turns negative immediately in the first quarter after the launch of ChatGPT. In 2023Q1, the number of job postings for occupations with an above-median AI-substitution score declined by approximately 5% relative to the control group. This initial impact intensifies systematically over time, deepening to around 10% after mid-2023 (Q=3), 15% by the end of 2024 (Q=8), and reaching approximately 20% by the second quarter of 2025.[20]

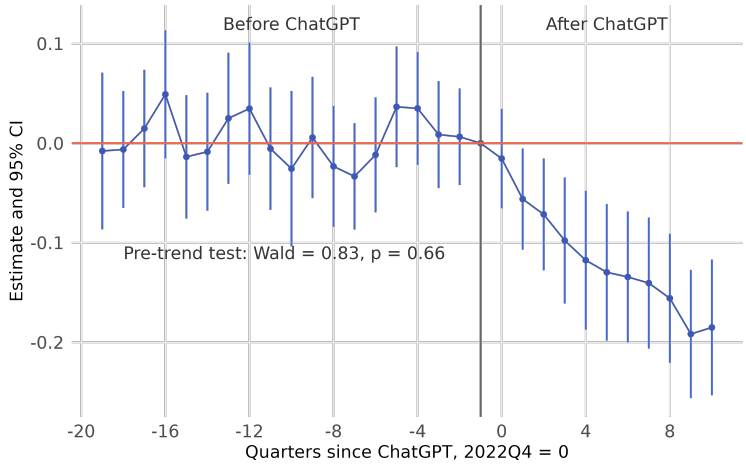


Figure 4: Event study: the "substitution" effect, conditional on GenAI exposure

Note: The figure plots the point estimates and 95% confidence intervals based on the event study specification in Equation 2. The reference period is the quarter preceding ChatGPT's launch (Q=-1). Dots represent quarterly gaps in job postings between the treatment group and the control group. The treatment group is defined as occupations with an above-median AI substitution score (or the AI complementary score, as in Pizzinelli et al. (2023)). The regression unit is at the state-industry-occupation-quarter level, including cells with zero counts. The specification controls for state-industry-occupation FE, state-industry-quarter FE, as well as a series of dummies, based on the decile ranking of the GenAI exposure index from Gmyrek, Berg, and Bescond (2023), interacting with quarterly dummies.

The dynamic effects show three distinctive features of GenAI's impact on labor demand. First, the immediate onset of negative effects suggests that labor markets responded quickly to the new technology's availability, contradicting narratives of delayed adjustment (Brynjolfsson, Rock, and Syverson 2017; Narayanan and Kapoor 2025). The rapid adoption of ChatGPT could have con-

[20]To examine treatment intensity effects, we further subdivide high-substitution occupations into two groups using the 50th and 90th percentiles as thresholds. Figure A5 compares the impact on medium-high substitution occupations (50th-90th percentile) and highest substitution occupations (90th-100th percentile) relative to below-median occupations. As a robustness check, Figure A6 uses the above-median AI exposure score to identify the treatment group.

tributed to the immediate effect.[21] Second, the quick and intensifying response can also be attributed to the more flexible and forward-looking nature of hiring trends, as new hiring is usually easier to adjust than existing employment, which requires layoffs or organizational restructuring. Third, the sustained downward trajectory likely reflects both the gradual organizational adoption of AI tools and continued improvements in model capabilities over time.

The magnitude of the negative impact underscores the disruptive potential of GenAI. A nearly 20% decline in postings for occupations with above-median AI-substitution scores is economically significant, aligning with estimates in recent studies (Brynjolfsson, Chandar, and Chen 2025; Hui, Reshef, and Zhou 2024; Eisfeldt et al. 2023). Notably, this contraction occurred within just three years following the launch of ChatGPT, highlighting the rapid pace of technological disruption. For comparison, Acemoglu and Restrepo (2020) find that one additional robot per thousand workers reduced the employment-to-population ratio by only 0.18–0.34 percentage points in the U.S. between 1990 and 2007, while Acemoglu et al. (2022) report no discernible impact of earlier AI exposure on employment or wages during 2010–2018. Unlike robotics or offshoring, where employment declines accumulated gradually over decades, GenAI's effects are rapid, concentrated, and disproportionately hitting white-collar jobs. The speed and scale of this impact suggest that, over a short period, GenAI could profoundly reshape labor markets, underscoring the need for rapid policy responses and timely adaptation from education and training systems.

5 Robustness Checks

To ensure the reliability of our main findings, we conduct two sets of robustness checks. First, we validate our approach to identify the treatment group by running a series of placebo tests and using alternative AI-substitution scores (see Section 5.1). Second, we test the robustness of our results to the use of alternative AI exposure measures, as our identification strategy relies on the proper control of GenAI exposure levels (see Section 5.2).

[21] GenAI has spread faster than any other technology in history, with ChatGPT becoming the fastest application to reach 100 million users in just two months. By mid-2025, over a third of U.S. adults have used ChatGPT (link), and the tool has reached 700 million weekly active users in August 2025, representing nearly 20% of the global labor force(link).

5.1 Selection of the Treatment Group

5.1.1 Placebo tests

To verify that our treatment group is well identified, we implement placebo tests by randomly assigning occupations to the control and treatment groups. We conducted 400 placebo trials, randomly assigning half of all occupations to a "pseudo-treatment" group in each iteration, ignoring their actual AI substitution scores. For each random trial, we re-estimate our preferred specification in Column (6) of Table 1.

Figure 5 presents the results across the three post-treatment periods. Each panel displays the distribution of placebo coefficients for one post-treatment year (gray histogram) alongside our baseline estimate (red vertical line). The placebo distributions are approximately normally distributed and centered around zero, consistent with the null hypothesis of no systematic treatment effects under random assignments. In contrast, our actual estimates fall far into the left tails of these distributions. For Year 1, our coefficient of -0.06 represents the 6.5th percentile of the placebo distribution. For Year 2, the baseline coefficient of -0.14 falls at the 0.25th percentile, and for Year 3, our coefficient of -0.18 is visibly outside the range of coefficients produced under random assignments.

We then extend our robustness checks to our event study results. We ran 400 placebo trials using the same randomization procedure and re-estimated the model specification detailed in Equation (2). Figure 6 overlays our baseline event study results (thick blue line) with the full distribution of placebo estimates (light blue lines). Unlike our baseline results, the placebo estimates are largely clustered around zero in both the pre- and post-treatment periods, indicating no systematic pattern or trend. Meanwhile, our baseline estimates fluctuate around zero in the pre-treatment period, which is consistent with the parallel trends assumption. More importantly, they exhibit a notable downward trend falling both below zero and below the lower boundary of the estimates from the placebo tests after the launch of ChatGPT (marked by the gray vertical line).

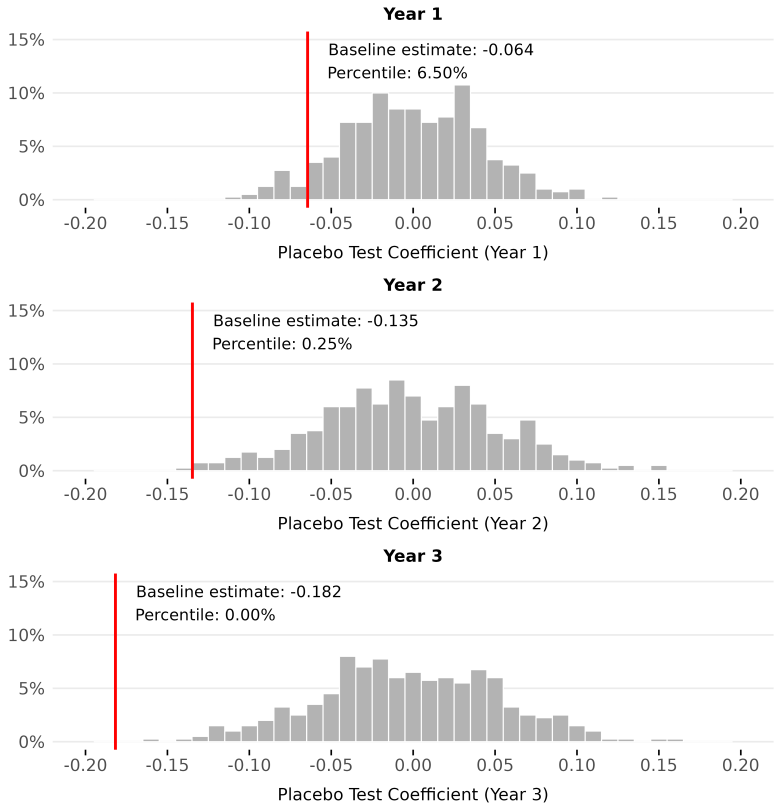


Figure 5: Placebo tests: DiD specification

Note: In each panel, the red vertical line indicates the baseline coefficients as in Column 6 of Table 1 for each post-treatment year, and the gray bars show the distribution of coefficients for each post-treatment year from 400 placebo tests with randomly assigning half of all occupations to a "pseudo-treatment" group and estimated with specification in Equation 3.

These results from the placebo tests largely validate our identification of the treatment group, as random treatment assignment yields null effects under the same model specification. Most importantly, the placebo tests confirm that ChatGPT's launch disproportionately affected the demand for occupations with above-median AI-substitution scores in the U.S. The intensification of the treatment effects over time, visible in our baseline results but absent from the placebo tests, provides compelling evidence that we are capturing the causal impact of GenAI adoption on labor demand.

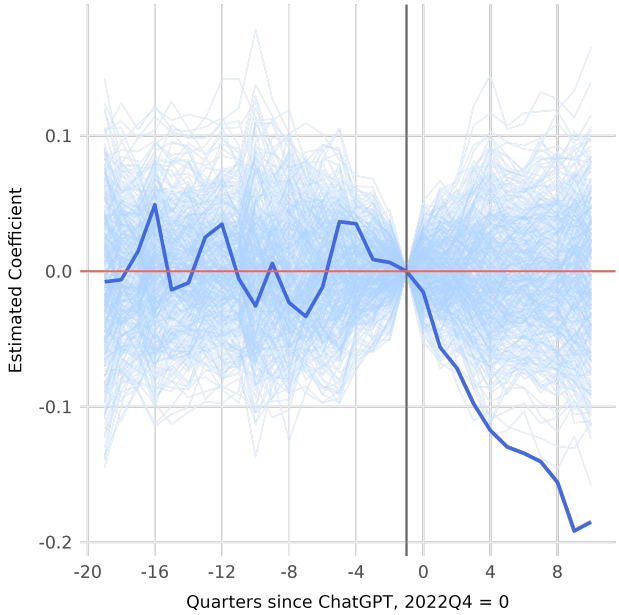


Figure 6: Placebo tests: Event study

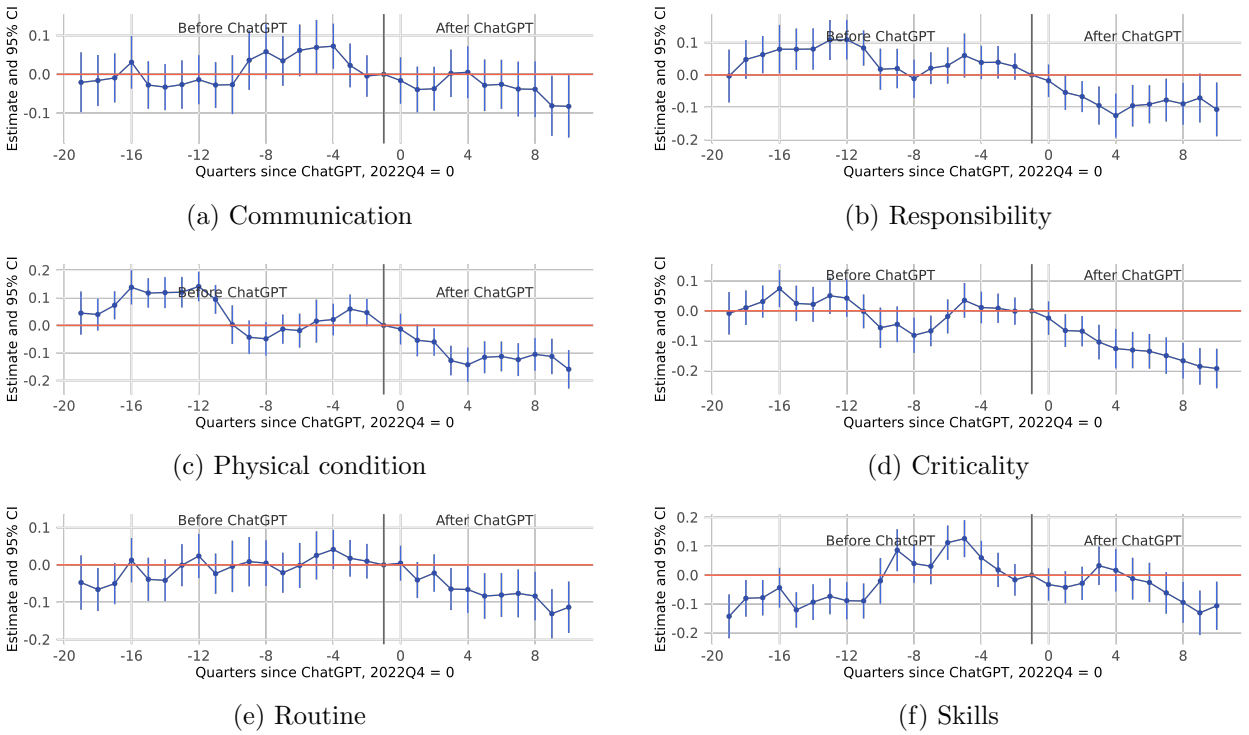
Note: The dark blue line indicates the baseline coefficients as in Figure 4, and the light blue lines show the coefficients from 400 placebo tests with randomly assigning half of all occupations to a "pseudo-treatment" group and estimated with specification in Equation 2.

5.1.2 Alternative AI Substitution Measures

We examine whether our baseline results remain robust when using alternative AI substitution measures. We first use each of the six individual components that comprise the aggregate AI-substitution score, and then adopt another relevant occupation characteristic provided by O*NET: interpersonal skills requirement. It is expected that an occupation with higher requirements for interpersonal skills is less likely to be replaced by GenAI in the short term.

As detailed previously, Pizzinelli et al. (2023) identifies six occupational dimensions that are relevant for the likelihood of AI displacement: communication, responsibility, physical conditions, criticality, routine, and skill requirements. Five of the six dimensions are related to the work context of an occupation, with the last one being about the skill zone (see Pizzinelli et al. 2023 for details). Occupations that require higher levels of communication, responsibility, physical presence, and specialized skills, while being less routine, are likely to have a lower risk of being replaced by AI. We classify occupations into treatment and control groups using the median score for each dimension. Occupations with above-median scores are placed in the treatment group, while those

with below-median scores are assigned to the control group.



(a) Communication

(b) Responsibility

(c) Physical condition

(d) Criticality

(e) Routine

(f) Skills

Figure 7: Robustness check: define treatment using the sub-components of the AI-substitution score

Note: The figures plot the point estimates and 95% confidence intervals based on the event study specification in Equation 2. The reference period is the quarter before ChatGPT first launched (Q=-1). Dots represent quarterly gaps in job postings between the treatment group and the control groups. The treatment groups are defined below median on Communication, Responsibility, Physical conditions, Criticality, or Skill requirements, or above median on Routine task intensity from O*NET. Regression unit is at the state-industry-occupation-quarter level, with zero counts included. The specification controls for state-industry-occupation FE, state-industry-quarter FE, as well as a series of dummies, based on the decile ranking of GenAI exposure index from Gmyrek, Berg, and Bescond (2023), interacting with quarterly dummies.

Figure 7 presents the results across all six dimensions. The job displacement effect of ChatGPT is found across almost all dimensions of the AI substitution score. However, it is most pronounced when considering the following factors to identify the treatment group: physical condition, criticality, responsibility, and routine. This finding suggests that these four work-context features seem to be more relevant to determining an occupation's vulnerability to AI substitution. Communication requirements produce weaker and largely insignificant effects when used to classify the treatment group. Since many early adopters utilized ChatGPT to enhance communication (Bick, Blandin, and Deming 2024), this aspect may not effectively identify the treatment group as initially expected. Additionally, it is possible that ChatGPT is employed to bridge the skill gap, leading to a

delayed effect when the skill score is used for grouping.

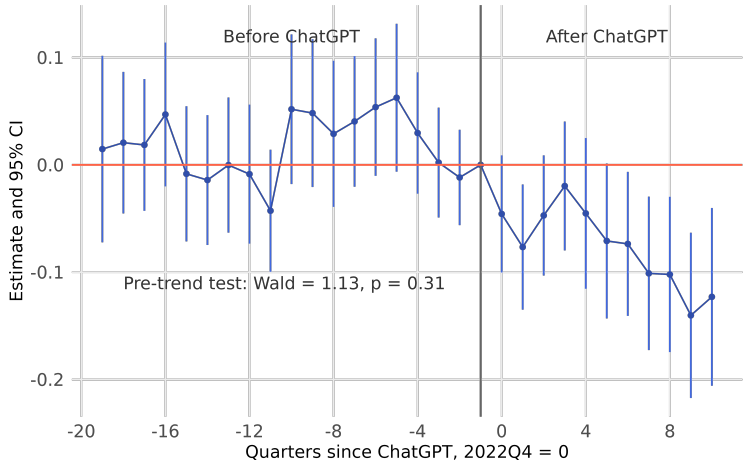


Figure 8: Robustness check: Interpersonal skills requirement

Note: The figures plot the point estimates and 95% confidence intervals based on the event study specification in Equation 2. The reference period is the quarter before ChatGPT first launched (Q=-1). Dots represent quarterly gaps in job postings between the treatment group and the control groups. The treatment groups are defined below median on interpersonal skills required from O*NET. Regression unit is at the state-industry-occupation-quarter level, with zero counts included. The specification controls for state-industry-occupation FE, state-industry-quarter FE, as well as a series of dummies, based on the decile ranking of GenAI exposure index from Gmyrek, Berg, and Bescond (2023), interacting with quarterly dummies.

We further test our findings using interpersonal skill requirements as an alternative measure of AI-substitution vulnerability. Drawing on Gmyrek, Berg, and Bescond (2023), we employ O*NET's interpersonal skills work context score, which measures the extent to which occupations require face-to-face interactions, relationship building, and social coordination abilities. It is expected that occupations requiring minimal interpersonal engagement are more likely to be displaced by AI. Therefore, we define our treatment group as occupations with below-median interpersonal skills scores where human social interaction is less essential to performance. Figure 8 confirms our baseline finding that job postings for occupations requiring limited interpersonal skills declined significantly following the launch of ChatGPT, with effects persisting and intensifying over time. In summary, these robustness checks support our main conclusions while highlighting that certain occupational features, such as responsibility and criticality, play a more significant role in determining the risk of AI substitution for an occupation.

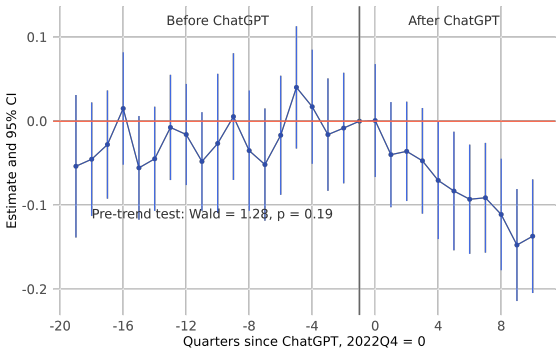
5.2 Alternative AI Exposure Measurements as Controls

Our baseline specification uses the GenAI exposure index from Gmyrek, Berg, and Bescond (2023) to construct the GenAI exposure decile-quarter fixed effects. To ensure our findings are not artifacts of this specific control choice, we conduct additional robustness checks using alternative exposure measures and model specifications. Specifically, we first used two alternative AI exposure scores as controls: the GenAI exposure score provided by Gmyrek et al. (2025) and AI Occupational Exposure Index (AIOE) from Felten, Raj, and Seamans (2021). The former is an updated version of Gmyrek, Berg, and Bescond (2023), where the automation risks experienced by workers are integrated into the new exposure measure. The latter is one of the first AI exposure measures and has been employed in many empirical studies (Acemoglu et al. 2022; Fossen and Sorgner 2021; Goldfarb, Taska, and Teodoridis 2023). This AIOE index links 10 AI applications to 52 O*NET labor abilities by surveying 2,000 workers on Amazon Mechanical Turk, creating an application-ability relatedness matrix that is aggregated into occupation-level exposure scores. We also show results without controlling for the GenAI exposure decile-quarter fixed effects, allowing us to assess whether our conditioning strategy is necessary for identification.

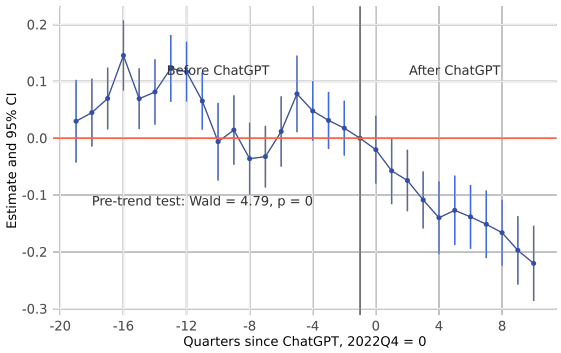
Figure 9 presents the results based on Equation (2) across all three different control specifications. The patterns remain largely consistent with our baseline findings. In each case, we observe significant declines in job postings for occupations with above-median AI-substitution scores following the launch of ChatGPT, with effects intensifying over time. When using Gmyrek et al. (2025)'s exposure score to construct the control dummies, the negative effect becomes slightly smaller in magnitude, with statistical significance starting from the second year. This is largely because the updated GenAI exposure index integrated the automation potential felt by workers after the launch of ChatGPT into the exposure score. Therefore, it may dilute the treatment effect identified when conditioning on Gmyrek et al. (2025)'s scores.

When using AIOE or no exposure controls, the fall in job postings seems to start earlier than the release date of ChatGPT, and it shows an overall downward long-term trend. As a result, these two specifications failed the pre-trend test. This is potentially due to the fact that AIOE precedes the launch of ChatGPT and is not specifically about GenAI. These results confirm that proper conditioning on GenAI-specific exposure is crucial for causal identification. Otherwise, the

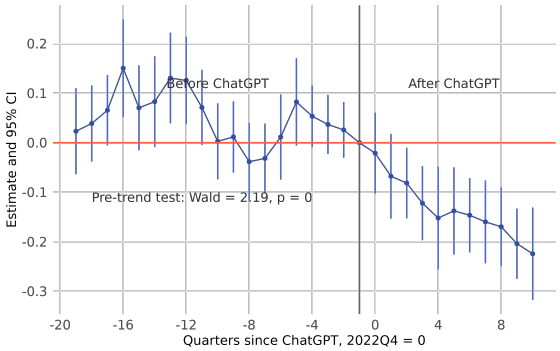
treatment effects identified by the corresponding DiD specification, such as in Columns (2) and (5) of Table 1, may capture the long-term divergent trends between the treatment and control groups, potentially overestimating GenAI's impact on labor demand. In contrast, our baseline model specification is more appropriate than alternative models by distinguishing GenAI impacts from general automation effects.



(a) Conditional on ILO(2025) GenAI exposure deciles



(b) Conditional on AIOE deciles



(c) Not conditional on GenAI exposure deciles

Figure 9: Robustness check: Using alternative AI exposure scores as controls or without controls

Note: The figures plot the point estimates and 95% confidence intervals based on the event study specification in Equation 2. The reference period is the quarter before ChatGPT first launched (Q=-1). Dots represent quarterly gaps in job postings between the treatment group and the control groups. The treatment groups are defined below median on interpersonal skills required from O*NET. Regression unit is at the state-industry-occupation-quarter level, with zero counts included. The specification controls for state-industry-occupation FE, state-industry-quarter FE. In panel (a) and (b), additional controls are included, which is a series of dummies indicating decile ranking of GenAI exposure index from Gmyrek et al. (2025) for panel (a), and AI exposure index from Felten, Raj, and Seamans (2021), interacting with quarterly dummies.

6 Heterogeneity

Heterogeneity by education and experience requirements. This section first examines GenAI's heterogeneous effects on job postings along two key dimensions: (1) educational requirements, and (2) experience requirements. Within the same occupation, GenAI may affect workers with different education levels in systematically different ways.[22] Similarly, its effects may differ between less experienced and more seasoned workers. On the one hand, GenAI can enable less experienced workers to perform tasks previously requiring extensive training or specialized knowledge, potentially reducing firms' demand for more experienced workers. On the other hand, experienced workers may leverage GenAI to automate routine tasks typically assigned to junior staff, potentially displacing entry-level positions. Recent evidence suggests the latter effect dominates, with entry-level and less-experienced workers bearing a disproportionate displacement effect induced by GenAI (Beane 2024; Brynjolfsson, Chandar, and Chen 2025; Lichtinger and Hosseini Maasoum 2025).[23]

Using the detailed information provided by Lightcast, we classify job postings into discrete categories based on the minimum requirements specified in job descriptions. Educational requirements are grouped into high school completion, college degrees, and advanced degrees (master's, professional, or doctoral degree). Experience requirements fall into three tiers: entry-level positions (0-2 years), mid-level positions (3-5 years), and senior positions (6+ years). To estimate the heterogeneous effect, we apply the following specification:

$$arsinh(Y_{siogt}) = \sum_e \alpha_{1e} D_o \times \mathbf{1}_t^{Post\,1yr} \times \mathbf{1}_g^e + \sum_e \alpha_{2e} D_o \times \mathbf{1}_t^{Post\,2+yr} \times \mathbf{1}_g^e + \chi_{j(o)t} + \theta_{sigt} + \gamma_{siog} + u_{siogt}, \tag{4}$$

where the subscript g indexes different job characteristics (i.e., education requirements and experience levels). The superscript e is used to construct dummy variables, $\mathbf{1}_g^e$, which take the value of one if observation g belongs to group e, and zero otherwise. This allows us to estimate separate coefficients for each group. The coefficients $\{\alpha_{1e}\}$ and $\{\alpha_{2e}\}$ reveal how GenAI's displacement effects differ among particular types of positions, in the first year and over the longer term, respectively.

[22] For example, in software development, a PhD graduate may focus on designing novel algorithms, while a college graduate in the same firm may be tasked with implementing and deploying them.

[23] Financial Times (2025). "A white-collar world without juniors?"

Figure 10 illustrates the effect of GenAI on job postings by educational and experience requirements.[24] The results reveal a clear skill gradient in GenAI's labor market impact for the first-year effect, with positions requiring only high school completion experiencing the sharpest decline. These job postings dropped 12% in occupations with above-median AI substitution vulnerability relative to the control group. College-degree positions showed a more moderate decline of 6%, while advanced-degree positions actually increased by 5%, suggesting possible complementarity effects or increased demand for more skilled professionals.

However, the longer-term effect reveals an intensification and broadening of GenAI's displacement impact. All educational categories eventually experienced significant declines, and the magnitude shifted to a U-shaped curve related to skill requirements. College-degree positions showed the steepest long-term decline at 18%, followed by high school positions at 15%, while advanced-degree positions declined by only 7%. This pattern suggests that while GenAI initially spared higher-skill positions, its evolving capabilities eventually affected middle-skill jobs most severely.

The experience dimension reveals that early-career workers are more affected by GenAI's displacement effect. This finding supports the argument that GenAI's knowledge transfer capabilities particularly threaten positions where learning through experience traditionally provides value. Entry-level positions (0-2 years) experienced the most severe immediate impact, declining 10% in the first year. Mid-career positions (3-5 years) showed moderate displacement at 7%, while senior positions (6+ years) showed no discernible short-term effect. Over the longer term, displacement effects intensified across all experience categories while maintaining their hierarchical structure. Entry-level and mid-level positions show declines of 19% and 14%, respectively, while senior positions experienced a 10% decline. This suggests that GenAI complements rather than substitutes the expertise and judgment accumulated through workers' years of experience. These findings corroborate emerging evidence indicating that young, early-career workers represent the "canaries in the coal mine" for GenAI displacement (Brynjolfsson, Chandar, and Chen 2025).

[24]The regression results are presented in Tables A2 and A3. To ensure our findings are not driven by sample selection, we include job postings that lack explicit education or experience requirements in the first two columns, and exclude such postings in the last two columns. The inclusion of postings without specified requirements yields very similar results to those in Columns (3) and (4), confirming that our findings are robust to sample composition.

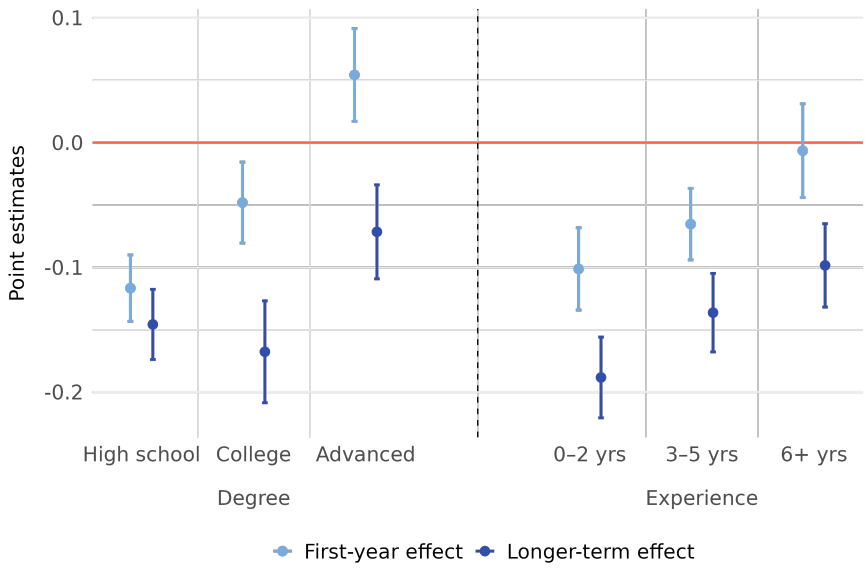


Figure 10: Heterogeneity by education and by experience

Note: This figure shows estimates from regressions based on Equation 4, presented separately by education and by experience. Advanced degrees include master's, professional, and doctoral degrees. "yrs" is short for "years". Job postings without education or experience requirements are excluded. The dots represent the estimated effects, while the spikes indicate the corresponding 95 percent confidence intervals. The light blue dots and spikes show the effect within the first year after ChatGPT's release, while the dark blue dots and spikes show longer-term effects till June 2025. Full regression results are presented in the first column of Tables A2 andA3, with additional specifications provided as robustness checks in the other columns.

Heterogeneity by industry. Next, we analyze GenAI's labor displacement effects across 20 two-digit NAICS industries to capture its sectoral variation. Specifically, we interact the industry dummies with the previous set of DiD interaction terms, as shown below:

$$arsinh(Y_{siot}) = \sum_{\iota} \alpha_{1\iota} D_o \times \mathbf{1}_t^{Post\,1yr} \times \mathbf{1}_i^{\iota} + \sum_{\iota} \alpha_{2\iota} D_o \times \mathbf{1}_t^{Post\,2+yr} \times \mathbf{1}_i^{\iota} + \chi_{j(o)t} + \theta_{sit} + \gamma_{sio} + u_{siot}. \quad (5)$$

where the superscript ι is used in dummy variables to indicate an industry, so that dummies $\mathbf{1}_i^{\iota}$ take the value of one if observation i belongs to industry ι and zero otherwise.

Figure 11 illustrates both the short-term and long-term effects of GenAI by industry. Overall, six industries experienced significant negative effects within the first year: administrative and support services, professional services, management of companies, real estate, other services, and healthcare. Eleven industries showed no significant short-term effects, spanning diverse sectors from education and construction to information technology, hospitality, and manufacturing. Three industries—retail trade, finance, and transportation—initially showed positive effects.

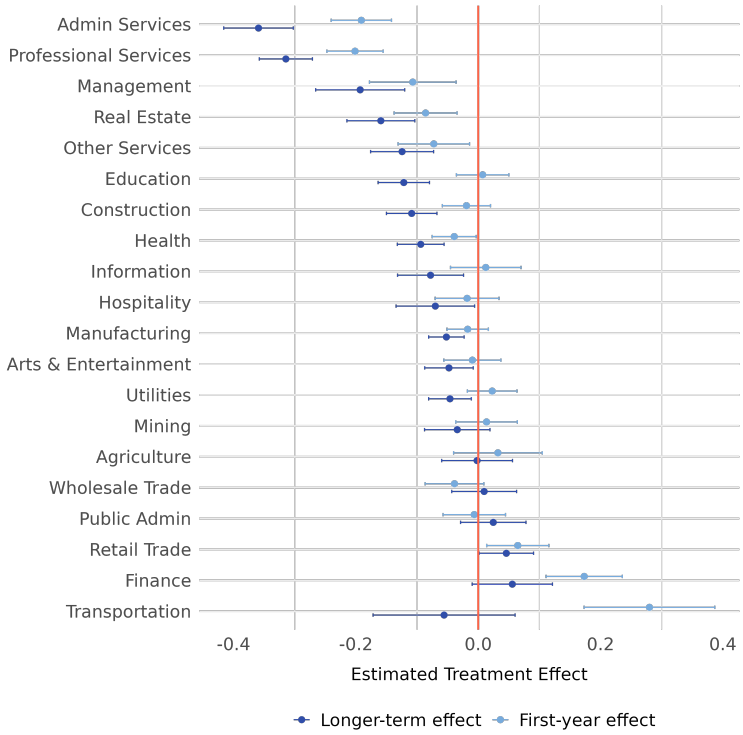


Figure 11: Heterogeneity by industry

Note: This figure shows estimates from regressions based on Equation 5, presented separately by industry. "yrs" is short for "years". The dots represent the estimated effects, while the spikes indicate the corresponding 95 percent confidence intervals. The light blue dots and spikes show the effect within the first year after ChatGPT's release, while the dark blue dots and spikes show longer-term effects till June 2025.

The dynamic analysis reveals a striking convergence toward negative effects across most industries. Thirteen of the twenty industries ultimately experienced significant declines in job postings, with the effect intensifying in previously affected sectors and emerging in previously resistant ones. Administrative services, professional services, management of companies, real estate, other services, and education showed particularly pronounced long-term effects. These industries also tend to have high GenAI adoption rates and high occupational exposure, as indicated by existing studies.[25]

Lastly, seven industries maintained insignificant effects over the full period. They are transportation, finance, public administration, wholesale trade, agriculture, mining, and utilities. These sectors likely possess characteristics that limit GenAI's applicability, including physical task requirements, regulatory constraints, or specialized domain knowledge requirements. The results

[25] See, e.g., Felten, Raj, and Seamans (2021), Gmyrek, Berg, and Bescond (2023), and Bick, Blandin, and Deming (2024)

support the arguments made by Korinek (2024) that in the presence of GenAI, technical and social barriers, along with fundamental human-centric aspects of labor, can maintain some demand for human involvement in the economy, at least in the short to medium term.

7 Conclusion

This paper presents the first comprehensive analysis of GenAI's impact on U.S. labor demand, utilizing real-time online job posting data. Taking the public release of ChatGPT in November 2022 as an exogenous technology shock, we apply both the DiD and event study methods to estimate the impacts of GenAI on job postings across occupations. It is expected that occupations with higher AI substitution vulnerability within the same level of GenAI exposure are more likely to be negatively affected by the launch of ChatGPT than those with lower AI substitution vulnerability. Therefore, we compare labor demand for occupations with above-median and below-median AI-substitution vulnerability conditional on similar levels of GenAI exposure. This model setup allows us to isolate the labor-displacing effects of GenAI while controlling for the effect of GenAI adoption over time and other confounding factors. Our dynamic analysis tracks how these impacts evolve over time as GenAI capabilities improve and adoption deepens.

We document large, statistically significant, and intensifying negative impacts of GenAI on job postings for occupations with higher AI-substitution vulnerability, when controlling for the GenAI exposure level. Following ChatGPT's launch, we find an average 12% decline in postings for occupations with above-median substitution scores between late 2022 and June 2025. This displacement effect grew from approximately 6% in the first year to 18% in the third year, suggesting that GenAI's labor market impacts compound as the technology matures and diffuses.

Heterogeneity analyses reveal systematic variation in GenAI's effects across worker characteristics and industries. Entry-level positions requiring neither advanced degrees nor extensive work experience face disproportionate displacement, with particularly severe reductions in professional services and administrative support industries. This pattern indicates that GenAI's democratization of specialized knowledge most directly threatens traditional pathways for early-career skill development. These findings suggest that the advancements of GenAI tools may help explain some of the trends observed in the U.S. labor market, such as the slowdown in hiring among white-collar

workers and in industries with high GenAI adoption rates. Moreover, our results also correspond with the deteriorating employment prospects for recent college graduates.[26].

Early evidence shows that GenAI's impact on labor demand differs from the impact of past technologies. Disruptions from previous technologies often unfolded gradually over decades, giving workers, firms, and policymakers time to adapt. By contrast, GenAI has led to a nearly 20% decline in demand for vulnerable occupations within only three years of its adoption. The decline is more notable for candidates without advanced degrees and early-career workers, and is most acute in administrative support (40%) and professional services (30%). Moreover, unlike earlier technologies that primarily displaced routine physical jobs, GenAI seems to affect the demand for traditionally white-collar jobs more.

Our findings carry profound policy implications. First, education and training systems need to adapt to prepare students for the era of AI. This may include more emphasis on building creative problem-solving skills, emotional intelligence, interpersonal communication skills, as well as digital or AI literacy. Second, the decline of traditional learn-by-doing opportunities in entry-level positions highlights the need for strong support for early-career development. Examples of such support programs include apprenticeship programs, internship programs, and entrepreneurship or skill training programs. Third, policymakers should also consider expanding social protection systems to support workers during this transition. Lastly, redistributive policies should be considered if such job displacement effects become more permanent (Korinek 2024).

The rapid adoption of GenAI and its labor market implications opens broad avenues for future research (Brynjolfsson, Korinek, and Agrawal 2025). Comparative studies across countries—particularly across developing countries—are currently lacking. They can shed light on what country-specific factors can shape the impact of GenAI on the labor market. Beyond labor demand, future work can also assess the impact on wages, income distribution, and skill requirements. More importantly, greater efforts should be made to study how GenAI may generate entirely new types of jobs and enhance productivity in ways that could increase labor demand.

[26]Data from the Current Population Survey reveal troubling patterns for young workers aged 23-27. College graduates in this cohort experienced unemployment rates averaging 4.59% in 2025, representing a 1.34 percentage point increase since 2019. This magnitude far exceeds increases for other demographic groups: non-college workers in the same age range saw only a 0.47 percentage point rise, while older college graduates experienced a 0.38 percentage point increase. These disparities suggest that GenAI's labor market effects are not merely cyclical but reflect structural changes in how the economy absorbs newly educated talent (Federal Reserve 2025 (link))

References

Acemoglu, Daron and David Autor (2011). "Lectures in labor economics". *Manuscript. http://economics. mit. edu/files* 4689, p. 22.

Acemoglu, Daron, David Autor, Jonathon Hazell, and Pascual Restrepo (2022). "Artificial intelligence and jobs: Evidence from online vacancies". *Journal of Labor Economics* 40.S1, S293–S340.

Acemoglu, Daron and Pascual Restrepo (2019). "Automation and new tasks: How technology displaces and reinstates labor". *Journal of economic perspectives* 33.2, pp. 3–30.

— (2020). "Robots and jobs: Evidence from US labor markets". *Journal of political economy* 128.6, pp. 2188–2244.

Autor, David H (2015). "Why are there still so many jobs? The history and future of workplace automation". *Journal of economic perspectives* 29.3, pp. 3–30.

Babina, Tania, Anastassia Fedyk, Alex X He, and James Hodson (2023). *Firm investments in artificial intelligence technologies and changes in workforce composition*. Vol. 31325. National Bureau of Economic Research.

Baker, Andrew, Brantly Callaway, Scott Cunningham, Andrew Goodman-Bacon, and Pedro HC Sant'Anna (2025). "Difference-in-differences designs: A practitioner's guide". *arXiv preprint arXiv:2503.13323*.

Beane, Matt (2024). *The Skill Code: How to Save Human Ability in an Age of Intelligent Machines*. HarperCollins.

Benitez-Rueda, Miguel and Eric Parrado (2024). *Mirror, Mirror on the Wall: Which Jobs Will AI Replace After All?: A New Index of Occupational Exposure*. Tech. rep. IDB.

Bick, Alexander, Adam Blandin, and David J Deming (2024). *The rapid adoption of generative ai*. Tech. rep. National Bureau of Economic Research.

Bonfiglioli, Alessandra, Rosario Crinò, Gino Gancia, and Ioannis Papadakis (2025). "Artificial intelligence and jobs: evidence from US commuting zones". *Economic Policy* 40.121, pp. 145–194.

Bonney, Kathryn, Cory Breaux, Cathy Buffington, Emin Dinlersoz, Lucia S Foster, Nathan Goldschlag, John C Haltiwanger, Zachary Kroff, and Keith Savage (2024). *Tracking Firm Use of AI in Real Time: A Snapshot from the Business Trends and Outlook Survey*. Tech. rep. National Bureau of Economic Research.

Braxton, J Carter and Bledi Taska (2023). "Technological change and the consequences of job loss". *American Economic Review* 113.2, pp. 279–316.

Brynjolfsson, Erik, Bharat Chandar, and Ruyu Chen (2025). *Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence*. Tech. rep.

Brynjolfsson, Erik, Anton Korinek, and Ajay K Agrawal (2025). "A Research Agenda for the Economics of Transformative AI".

Brynjolfsson, Erik, Danielle Li, and Lindsey Raymond (2025). "Generative AI at work". *The Quarterly Journal of Economics* 140.2, pp. 889–942.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson (2017). *Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics*. Tech. rep. National Bureau of Economic Research.

Chandar, Bharat (2025). "Tracking Employment Changes in AI-Exposed Jobs". *Available at SSRN 5384519*.

Chen, Jiafeng and Jonathan Roth (2024). "Logs with zeros? Some problems and solutions". *The Quarterly Journal of Economics* 139.2, pp. 891–936.

Cui, Zheyuan Kevin, Mert Demirer, Sonia Jaffe, Leon Musolff, Sida Peng, and Tobias Salz (2024). "The effects of generative ai on high skilled work: Evidence from three field experiments with software developers". *Available at SSRN 4945566*.

Dell'Acqua, Fabrizio, Edward McFowland III, Ethan R Mollick, Hila Lifshitz-Assaf, Katherine Kellogg, Saran Rajendran, Lisa Krayer, François Candelon, and Karim R Lakhani (2023). "Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality". *Harvard Business School Technology & Operations Mgt. Unit Working Paper* 24-013.

Deming, David J and Kadeem Noray (2020). "Earnings dynamics, changing job skills, and STEM careers". *The Quarterly Journal of Economics* 135.4, pp. 1965–2005.

Demirci, Ozge, Jonas Hannane, and Xinrong Zhu (2025). "Who is AI replacing? The impact of generative AI on online freelancing platforms". *Management Science*.

Eisfeldt, Andrea L, Gregor Schubert, Miao Ben Zhang, and Bledi Taska (2023). "The labor impact of generative AI on firm values". *Available at SSRN 4436627*.

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock (2024). "GPTs are GPTs: Labor market impact potential of LLMs". *Science* 384.6702, pp. 1306–1308.

Felten, Edward, Manav Raj, and Robert Seamans (2021). "Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses". *Strategic Management Journal* 42.12, pp. 2195–2217.

Felten, Edward W, Manav Raj, and Robert Seamans (2018). "A method to link advances in artificial intelligence to occupational abilities". *AEA Papers and proceedings* 108, pp. 54–57.

— (2019). "The occupational impact of artificial intelligence: Labor, skills, and polarization". *NYU Stern School of Business*.

— (2023). "Occupational Heterogeneity in Exposure to Generative AI". *SSRN Electronic Journal*.

Forsythe, Eliza, Lisa B Kahn, Fabian Lange, and David Wiczer (2022). "Where have all the workers gone? Recalls, retirements, and reallocation in the COVID recovery". *Labour Economics* 78, p. 102251.

Fossen, Frank and Alina Sorgner (2021). "Digitalization of work and entry into entrepreneurship". *Journal of Business Research* 125 (C). Publisher: Elsevier, pp. 548–563.

Frey, Carl Benedikt and Michael A. Osborne (Jan. 1, 2017). "The future of employment: How susceptible are jobs to computerisation?" *Technological Forecasting and Social Change* 114, pp. 254–280.

Gmyrek, Pawel, Janine Berg, and David Bescond (2023). "Generative AI and jobs: A global analysis of potential effects on job quantity and quality". *ILO Working Paper* 96.

Gmyrek, Paweł, Janine Berg, Karol Kamiński, Filip Konopczyński, Agnieszka Ładna, Balint Nafradi, Konrad Rosłaniec, and Marek Troszyński (2025). *Generative AI and jobs: A refined global index of occupational exposure*. 140. ILO Working Paper.

Goldfarb, Avi, Bledi Taska, and Florenta Teodoridis (2023). "Could machine learning be a general purpose technology? a comparison of emerging technologies using data from online job postings". *Research Policy* 52.1, p. 104653.

Hampole, Menaka, Dimitris Papanikolaou, Lawrence DW Schmidt, and Bryan Seegmiller (2025). *Artificial intelligence and the labor market*. Tech. rep. National Bureau of Economic Research.

Handa, Kunal, Alex Tamkin, Miles McCain, Saffron Huang, Esin Durmus, Sarah Heck, Jared Mueller, Jerry Hong, Stuart Ritchie, Tim Belonax, et al. (2025). "Which Economic Tasks are Performed with AI? Evidence from Millions of Claude Conversations". *arXiv preprint arXiv:2503.04761*.

Hershbein, Brad and Lisa B Kahn (2018). "Do recessions accelerate routine-biased technological change? Evidence from vacancy postings". *American Economic Review* 108.7, pp. 1737–1772.

Hershbein, Brad and Bryan A Stuart (2024). "The evolution of local labor markets after recessions". *American Economic Journal: Applied Economics* 16.3, pp. 399–435.

Hui, Xiang, Oren Reshef, and Luofeng Zhou (2024). "The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market". *Organization Science* 35.6, pp. 1977–1989.

Humlum, Anders and Emilie Vestergaard (2025). *Large language models, small labor market effects*. Tech. rep. National Bureau of Economic Research.

Jiang, Wei, Junyoung Park, Rachel Jiqiu Xiao, and Shen Zhang (2025). *AI and the Extended Workday: Productivity, Contracting Efficiency, and Distribution of Rents*. Tech. rep. National Bureau of Economic Research.

Korinek, Anton (2024). *Economic policy challenges for the age of AI*. Tech. rep. National Bureau of Economic Research.

Lichtinger, Guy and Seyed Mahdi Hosseini Maasoum (2025). "Generative AI as Seniority-Biased Technological Change: Evidence from US Resume and Job Posting Data". *Available at SSRN*.

Liu, Yan, Jingyun Huang, and He Wang (2025). *Who on Earth Is Using Generative AI? Tracking Global Trends and Shifts in 2025*. Tech. rep.

Liu, Yan and He Wang (Aug. 2024). *Who on Earth Is Using Generative AI?* Policy Research Working Paper 10870. Washington, DC: World Bank, pp. 1–56.

Moretti, Enrico (2010). *Local labor markets*. Tech. rep. National Bureau of Economic Research.

Narayanan, Arvind and Sayash Kapoor (2025). "AI as normal technology". *Knight First Amend. Inst.*

Noy, Shakked and Whitney Zhang (2023). "Experimental evidence on the productivity effects of generative artificial intelligence". *Science* 381.6654, pp. 187–192.

Pizzinelli, Carlo, Augustus J Panton, Ms Marina Mendes Tavares, Mauro Cazzaniga, and Longji Li (2023). *Labor market exposure to AI: Cross-country differences and distributional implications*. International Monetary Fund.

Qiao, Dandan, Huaxia Rui, and Qian Xiong (2023). "AI and jobs: has the inflection point arrived? Evidence from an online labor platform". *arXiv preprint arXiv:2312.04180*.

Teutloff, Ole, Johanna Einsiedler, Otto Kässi, Fabian Braesemann, Pamela Mishkin, and R Maria del Rio-Chanona (2025). "Winners and losers of generative AI: Early Evidence of Shifts in Freelancer Demand". *Journal of Economic Behavior & Organization*, p. 106845.

Tomlinson, Kiran, Sonia Jaffe, Will Wang, Scott Counts, and Siddharth Suri (2025). "Working with AI: Measuring the Occupational Implications of Generative AI". *arXiv preprint arXiv:2507.07935*.

Topel, Robert H (1986). "Local labor markets". *Journal of Political economy* 94.3, Part 2, S111–S143.

Webb, Michael (2019). "The impact of artificial intelligence on the labor market". *Available at SSRN 3482150*.

World Bank (June 2025a). *Future Jobs: Robots, Artificial Intelligence, and Digital Platforms in East Asia and Pacific*. Full report. Washington, DC: World Bank.

— (2025b). *South Asia Development Update, October 2025*. Washington, DC: World Bank.

Appendix: Additional Figures and Tables

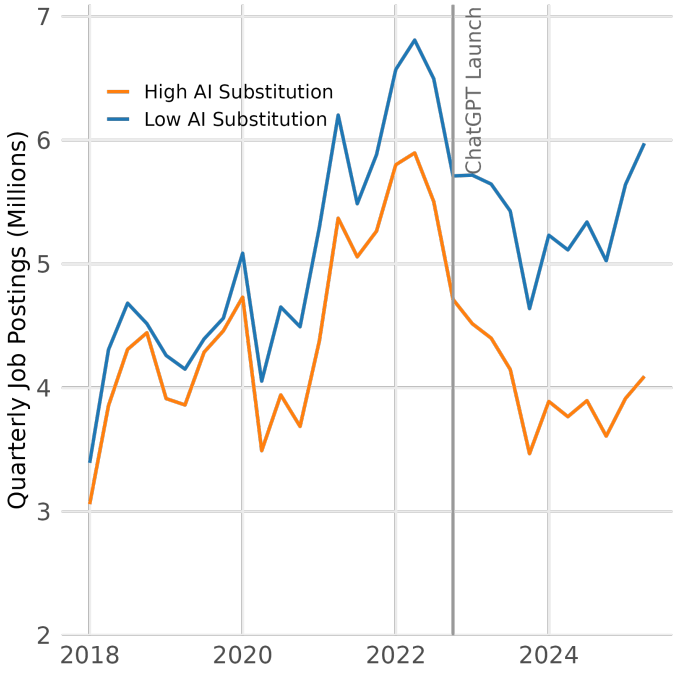


Figure A1: Job postings trends for occupations with high vs. low AI-substitution score

Note: This figure shows online job posting trends in LightCast for occupations above and below the median AI-substitution score. The AI-substitution score is defined as the inverse of the AI-complementarity measure in Pizzinelli et al. (2023), constructed along six occupational dimensions: communication, responsibility, physical conditions, criticality, routine, and skill requirements.

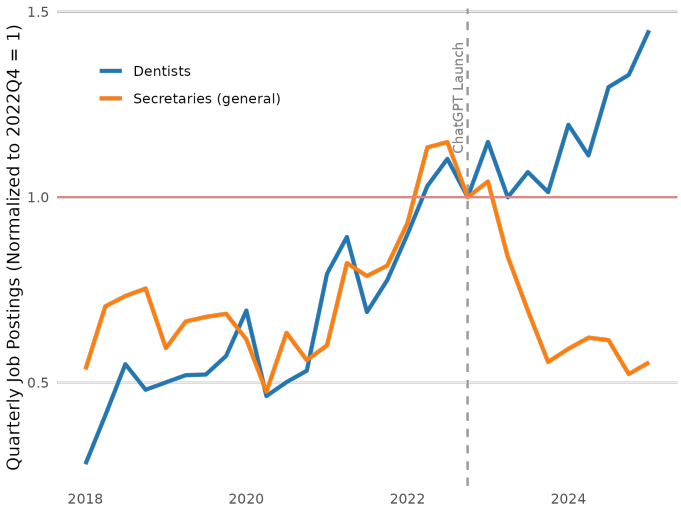


Figure A2: Job postings trends for two typical occupations: Dentists v.s. Secretaries (general)

Note: This figure shows online job posting trends in LightCast for two occupations Dentists v.s. Secretaries (general), which lie below and above the median AI-substitution score, respectively. Values are normalized to 1 in the quarter of ChatGPT's launch (2022Q4).

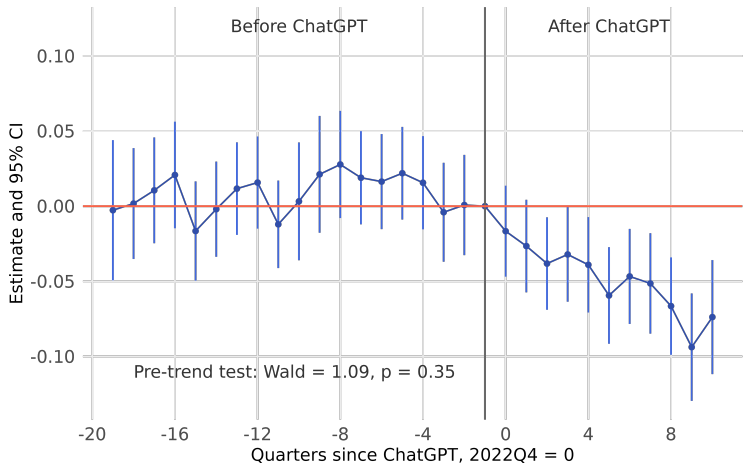


Figure A3: Robustness: non-weighted regression

Note: The figure plots the point estimates and 95% confidence intervals based on the event study specification in Equation 2, without using job postings by state-industry-occupation in 2017 as the sample weights. The reference period is the quarter before ChatGPT first launched (Q=-1). Dots represent quarterly gaps in job postings between the treatment group and the control groups. The treatment group is defined as occupations with above-median AI substitution score (or the inverse of the AI-complementary score in Pizzinelli et al. (2023)). Regression unit is at the state-industry-occupation-quarter level, with zero counts included. The specification controls for state-industry-occupation FE, state-industry-quarter FE, as well as a series of dummies, based on the decile ranking of GenAI exposure index from Gmyrek, Berg, and Bescond (2023), interacting with quarterly dummies.

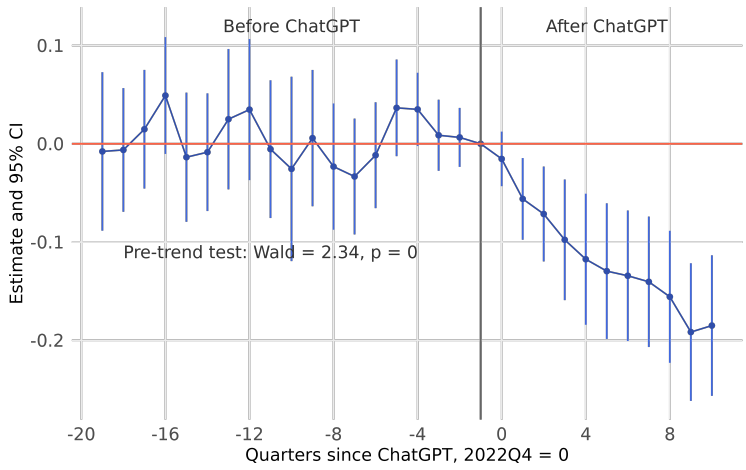


Figure A4: Robustness: SE clustered at occupational level

Note: The figure plots the point estimates and 95% confidence intervals based on the event study specification in Equation 2, while standard errors are clustered at occupation level. The reference period is the quarter before ChatGPT first launched (Q=-1). Dots represent quarterly gaps in job postings between the treatment group and the control groups. The treatment group is defined as occupations with above-median AI substitution score (or the inverse of the AI-complementary score in Pizzinelli et al. (2023)). Regression unit is at the state-industry-occupation-quarter level, with zero counts included. The specification controls for state-industry-occupation FE, state-industry-quarter FE, as well as a series of dummies, based on the decile ranking of GenAI exposure index from Gmyrek, Berg, and Bescond (2023), interacting with quarterly dummies.

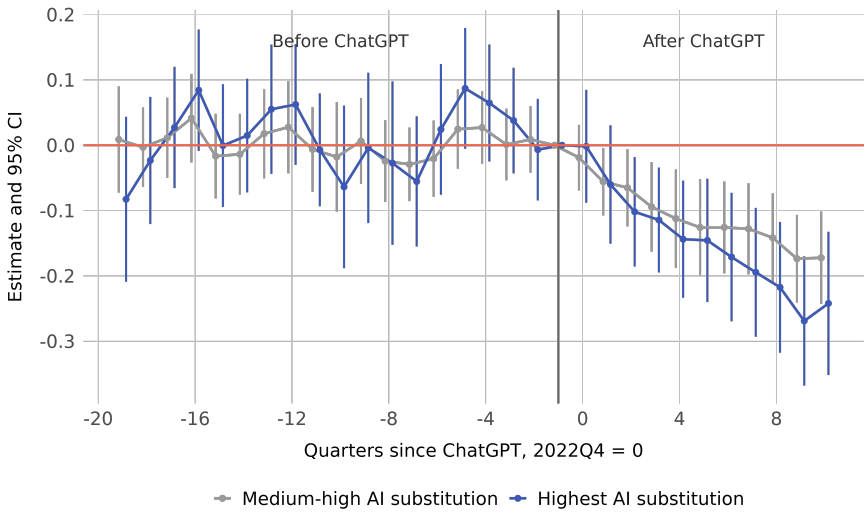


Figure A5: Robustness check: Other cutoff points to define the treatment group

Note: The figure plots the point estimates and 95% confidence intervals based on the event study specification in Equation 2, while we implement the event study with two treatment groups. The two treatment groups are defined by AI-substitution score: the medium-high substitution group includes the 50 to 90 percentile range, and the highest substitution group covers the 90 to 100 percentile range. The reference period is the quarter before ChatGPT first launched (Q=-1). Dots represent quarterly gaps in job postings between the treatment group and the control groups. The treatment group is defined as occupations with above-median AI substitution score (or the inverse of the AI-complementary score in Pizzinelli et al. (2023)). Regression unit is at the state-industry-occupation-quarter level, with zero counts included. The specification controls for state-industry-occupation FE, state-industry-quarter FE, as well as a series of dummies, based on the decile ranking of GenAI exposure index from Gmyrek, Berg, and Bescond (2023), interacting with quarterly dummies.

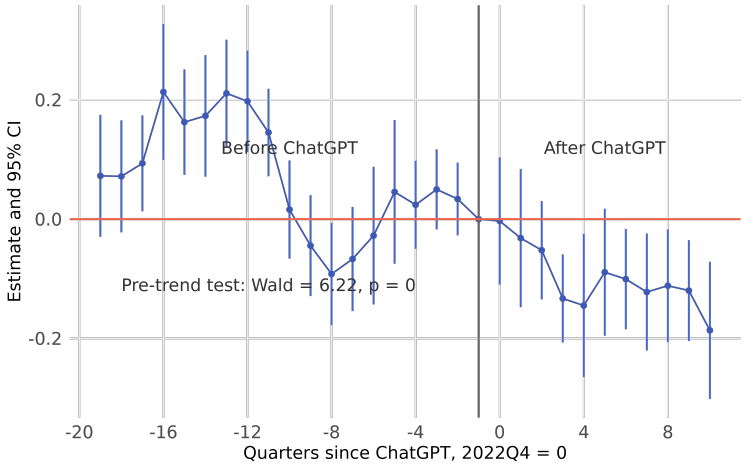


Figure A6: Robusteness check: using high GBB GenAI index to define the treatment group

Note: The figure plots the point estimates and 95% confidence intervals based on the event study specification in Equation 2. The reference period is the quarter before ChatGPT first launched (Q=-1). Dots represent quarterly gaps in job postings between the treatment group and the control groups. The treatment group is defined as occupations with above-median GenAI exposure index from Gmyrek, Berg, and Bescond (2023) (GBB). Regression unit is at the state-industry-occupation-quarter level, with zero counts included. The specification controls for state-industry-occupation FE, state-industry-quarter FE.

Table A1: Job Postings (Millions) by Occupation, Industry, Education, and Experience Requirements, 2018Q1–2025Q2

	Total		By Occupation (ISCO-08 One-Digit)								
			Managers	Professionals	Technicians and Associate Professionals	Clerical Support Workers	Services and Sales Workers	Skilled Agricultural, Forestry and Fishery Workers	Craft and Related Trades Workers	Plant and Machine Operators and Assemblers	Elementary Occupations
Total job postings (millions)	284.5		31.5	87.0	54.5	22.1	45.8	1.48	13.5	14.1	14.7
Percent		100.0%	11.1%	30.6%	19.2%	7.8%	16.1%	0.5%	4.7%	4.9%	5.2%
By Industry (NAICS Two-digit)											
Agriculture, Forestry, Fishing and Hunting	0.4	0.2%	0.05	0.07	0.09	0.03	0.05	0.04	0.03	0.03	0.04
Mining, Quarrying, and Oil and Gas Extraction	0.6	0.2%	0.05	0.14	0.13	0.03	0.02	0.002	0.07	0.10	0.03
Utilities	1.1	0.4%	0.13	0.37	0.24	0.07	0.03	0.01	0.12	0.03	0.05
Construction	5.9	2.1%	0.71	1.05	1.40	0.36	0.37	0.06	1.10	0.35	0.46
Manufacturing	17.3	6.1%	2.54	5.18	3.28	1.03	1.24	0.04	1.64	1.39	0.96
Wholesale Trade	7.5	2.6%	0.94	1.39	1.48	0.65	1.03	0.03	0.50	0.81	0.68
Retail Trade	25.5	9.0%	2.55	2.13	2.90	3.02	10.4	0.04	1.46	1.16	1.81
Transportation and Warehousing	6.8	2.4%	0.56	0.51	0.55	0.79	0.50	0.01	0.38	2.82	0.66
Information	6.2	2.2%	1.08	2.50	1.24	0.30	0.68	0.01	0.19	0.06	0.10
Finance and Insurance	12.2	4.3%	1.80	4.74	3.20	1.79	0.43	0.01	0.10	0.07	0.05
Real Estate and Rental and Leasing	4.6	1.6%	0.62	0.63	1.42	0.38	0.41	0.07	0.26	0.25	0.52
Professional, Scientific, and Technical Services	25.5	9.0%	2.74	12.7	5.21	1.91	1.54	0.07	0.58	0.42	0.34
Management of Companies and Enterprises	0.9	0.3%	0.12	0.26	0.18	0.10	0.08	0.002	0.06	0.06	0.04
Administrative and Support Services	34.8	12.2%	2.34	14.1	6.04	2.35	3.59	0.31	2.07	2.05	1.91
Educational Services	11.7	4.1%	1.67	5.61	1.99	0.69	1.32	0.06	0.16	0.11	0.14
Health Care and Social Assistance	35.3	12.4%	2.94	14.3	9.03	1.67	6.30	0.05	0.24	0.25	0.50
Arts, Entertainment, and Recreation	1.9	0.7%	0.28	0.29	0.45	0.19	0.49	0.03	0.04	0.04	0.10
Accommodation and Food Services	14.6	5.1%	3.13	0.65	0.81	1.21	5.42	0.06	0.27	0.56	2.53
Other Services (except Public Administration)	5.5	1.9%	0.63	1.06	1.04	0.44	1.17	0.03	0.64	0.23	0.24
Public Administration	4.7	1.7%	0.53	1.71	1.19	0.36	0.47	0.04	0.16	0.08	0.18
Not classified	61.6	21.6%	6.09	17.5	12.6	4.72	10.2	0.50	3.43	3.17	3.32
By Education Attainment Required											
High School	71.2	25.0%	6.02	9.61	18.4	8.87	15.6	0.40	4.46	3.41	4.39
College	28.1	9.9%	5.39	15.5	5.06	1.04	0.71	0.03	0.20	0.08	0.06
Advanced	18.2	6.4%	3.09	11.4	1.79	0.47	0.91	0.02	0.20	0.11	0.15
Not Mentioned	167.1	58.7%	17.0	50.4	29.2	11.7	28.6	1.03	8.64	10.5	10.1
By Years of Experience Required											
0-2 years	68.1	23.9%	5.71	19.4	15.3	6.73	10.8	0.32	2.94	3.69	3.24
3-5 years	43.7	15.4%	8.61	18.1	9.21	1.92	2.17	0.10	2.22	0.71	0.66
6+ years	16.9	5.9%	4.81	8.97	2.10	0.27	0.25	0.01	0.31	0.10	0.07
Not Mentioned	155.8	54.8%	12.4	40.4	27.9	13.1	32.6	1.04	8.03	9.57	10.7

Table A2: ChatGPT effect on labor demand, by education requirement

Dependent variable	arsinh(# job postings)			
	(1)	(2)	(3)	(4)
Treated×Short (High school)			-0.117***	-0.118***
			(0.014)	(0.014)
Treated×Short×High school	-0.117***	-0.118***		
	(0.014)	(0.014)		
Treated×Short×College degree	-0.048**	-0.050**	0.069***	0.069***
	(0.017)	(0.017)	(0.017)	(0.017)
Treated×Short×Advanced degree	0.054**	0.053**	0.171***	0.171***
	(0.019)	(0.019)	(0.022)	(0.022)
Treated×Short×Edu requirement not mentioned		-0.056***		0.063***
		(0.014)		(0.015)
Treated×Long (High school baseline)			-0.146***	-0.155***
			(0.014)	(0.014)
Treated×Long×High school	-0.146***	-0.155***		
	(0.014)	(0.014)		
Treated×Long×College degree	-0.168***	-0.177***	-0.022	-0.022
	(0.021)	(0.021)	(0.021)	(0.021)
Treated×Long×Advanced degree	-0.072***	-0.081***	0.074***	0.074***
	(0.019)	(0.019)	(0.02)	(0.02)
Treated×Long×Edu requirement not mentioned		-0.128***		0.027*
		(0.014)		(0.013)
Education × State × Industry × Occupation FE	Yes	Yes	Yes	Yes
Education × State × Industry × Year-quarter FE	Yes	Yes	Yes	Yes
Occupational Dencitle FE × Year-quarter FE	Yes	Yes	Yes	Yes
Num.Obs.	20,261,790	27,015,720	20,261,790	27,015,720
R2	0.943	0.956	0.943	0.956

Note: Robust standard errors clustered at the occupation-quarter level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Regression unit is at the education-state-industry-occupation-quarter level, with zero counts included. The dependent variable is the inverse hyperbolic sine (arsinh) transformation of job postings, which approximates the logarithmic transformation and equals zero when job postings are zero, allowing coefficients to be interpreted as approximate log changes in job postings. Treatment is a dummy indicating the treatment group, defined as occupations with above-median AI substitution score (or the AI complementary score in Pizzinelli et al. (2023)). "Short" is dummy variable equal to 1 for quarters in the first year after ChatGPT's launch, including 2022Q4, and "Long" is the second and the third year after ChatGPT's launch, 2023Q4-2025Q2. "Occupational Decile FE" is a series of dummies based on the decile ranking of GenAI exposure index from Gmyrek, Berg, and Bescond (2023). Advanced degree includes master's, professional, and doctoral degree.

Table A3: ChatGPT effect on labor demand, by experience requirement

Dependent variable	arsinh(# job postings)			
	(1)	(2)	(3)	(4)
Treated × Short (0-2 yrs exp as ref group)			-0.101***	-0.098***
			(0.017)	(0.017)
Treated × Short × 0-2 years experience	-0.101***	-0.098***		
	(0.017)	(0.017)		
Treated × Short × 3-5yrs experience	-0.065***	-0.062***	0.036**	0.036**
	(0.015)	(0.015)	(0.013)	(0.013)
Treated × Short × 6+yrs experience	-0.006	-0.003	0.095***	0.095***
	(0.019)	(0.019)	(0.021)	(0.021)
Treated × Short × Exp requirement not mentioned		-0.077***		0.020
		(0.015)		(0.012)
Treated × Long (0-2 yrs exp as ref group)			-0.188***	-0.188***
			(0.017)	(0.016)
Treated × Long × 0-2 years experience	-0.188***	-0.188***		
	(0.017)	(0.016)		
Treated × Long × 3-5yrs experience	-0.136***	-0.136***	0.052***	0.052***
	(0.016)	(0.016)	(0.012)	(0.012)
Treated × Long × 6+yrs experience	-0.098***	-0.098***	0.090***	0.090***
	(0.017)	(0.017)	(0.018)	(0.018)
Treated × Long × Exp requirement not mentioned		-0.139***		0.048***
		(0.014)		(0.011)
Experience × State × Industry × Occupation FE	Yes	Yes	Yes	Yes
Experience × State × Industry × Year-quarter FE	Yes	Yes	Yes	Yes
Occupational Dencile FE × Year-quarter FE	Yes	Yes	Yes	Yes
Num.Obs.	20,261,790	27,015,720	20,261,790	27,015,720
R2	0.952	0.96	0.952	0.96

Note: Robust standard errors clustered at the occupation-quarter level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Regression unit is at the education-state-industry-occupation-quarter level, with zero counts included. The dependent variable is the inverse hyperbolic sine (arsinh) transformation of job postings, which approximates the logarithmic transformation and equals zero when job postings are zero, allowing coefficients to be interpreted as approximate log changes in job postings. Treatment is a dummy indicating the treatment group, defined as occupations with above-median AI substitution score (or the AI complementary score in Pizzinelli et al. (2023)). "Short" is dummy variable equal to 1 for quarters in the first year after ChatGPT's launch, including 2022Q4, and "Long" is the second and the third year after ChatGPT's launch, 2023Q4-2025Q2. "Occupational Decile FE" is a series of dummies based on the decile ranking of GenAI exposure index from Gmyrek, Berg, and Bescond (2023).