Submitted to *Management Science*
manuscript MS-0001-1922.65

The Impact of Generative AI on Job Opportunities for Junior Software Developers

Samuel Westby
Network Science Institute, Northeastern University, Boston, MA 02115, westby.s@northeastern.edu

Alicia Sasser Modestino
School of Public Policy and Urban Affairs, Northeastern University, Boston, MA 02115, a.modestino@northeastern.edu

Emerging research suggests generative AI offers large productivity gains for the software industry, yet the employment impact on less- versus more-experienced software developers remains unclear. Using the public release of Chat-GPT in November 2022 as a natural experiment, we explore the effects of generative AI on the software developer job market, documenting three new stylized facts. First, using the near-universe of online job vacancies, we find the widespread introduction of generative AI resulted in a 16.3 percent drop in the relative proportion of junior- versus senior-level software developer job vacancies. This finding holds even when accounting for month and location fixed effects, controlling for firm size, population, and industry composition, and considering other labor market trends affecting computer and mathematical workers during this period. Second, we find that the relative decline in labor demand for junior- versus senior-level software developers appears concentrated among larger firms and bigger cities while industries with moderate exposure to the software industry were more insulated from the effects of Chat-GPT. Third, we reveal alternative career pathways for potentially displaced junior software developers using an occupation similarity network to identify credible job opportunities for junior software developers that require minimal re-skilling. Our findings are consistent with both recent industry observations as well as the vast literature on skill-biased technological change, reinforcing the notion that technological disruptions do not uniformly eliminate employment but rather reshape the labor market. Policies to create apprenticeships, co-ops, and other subsidized educational opportunities could help preserve future talent pipelines.

Key words: Job vacancies, employer skill requirements, skill biased technological change, generative AI

Acknowledgments

This work was supported by the National Science Foundation [Grant #2128743]

1. Introduction

The broad economic consequences of generative artificial intelligence (AI) are just beginning to unfold. Emerging research suggests that generative AI offers heterogeneous productivity gains for low-skilled (Cui et al. 2024, Peng et al. 2023, Brynjolfsson et al. 2025, Hui et al. 2024) versus high-skilled workers (Toner-Rodgers 2024, Otis et al. 2023) even within various occupations. For example, software developers have been identified as particularly affected due to the ability of generative AI to write text and code (Eloundou et al. 2024, Demirci et al. 2025), and entry- and junior-level software developers experienced the highest productivity gains (Cui et al. 2024, Peng et al. 2023). Not coincidentally, these computer and mathematical occupations also have the highest rate of adoption of generative AI, with 43% of workers in these occupations using generative AI tools as of November 2024 (Bick et al. 2024).

Why might generative AI affect certain occupations and workers more than others? Generative AI and large language models (LLMs) like ChatGPT work by predicting the next item in a sequence. If you ask a baseball fan to finish the sentence, “I love the Boston Red ______,” they would easily respond, “Sox”. Similarly, LLMs can write code, edit papers, and draft emails by predicting the next item in a pattern based on prior patterns observed in large amounts of data. Transformer models, formalized by Vaswani et al. (2017), expanded the capacity of neural networks to model these long-range dependencies and relationships in vast amounts of data. OpenAI then fine-tuned the transformer architecture using reinforcement learning with human feedback (Ouyang et al. 2022) to align the model’s outputs with human preferences for understanding and usability. This approach led to the development of ChatGPT, a generative AI model capable of producing coherent and accurate output across a wide range of topics and in a variety of contexts. Although generative AI had been in development for more than a decade, this ”last mile” advancement led to the wide-spread and immediate adoption of ChatGPT upon its release in November 2022, reaching 100 million users in two months (Hu 2023).

Why would generative AI have a more immediate impact on software developers in particular? The tasks required for many computer and mathematical occupations, such as software developers or data scientists, have a greater degree of overlap with tasks performed by LLMs such as ChatGPT (Eloundou et al. 2024) compared to other occupations. Software developers in particular use generative AI to produce routine code, giving general guidance on solving their particular problem, learning new concepts, and brainstorming solutions during each stage of software development, including the planning, implementation, and testing phases (Khojah et al. 2024). Software developers report interacting with generative AI mainly by seeking examples, explanations, or direction rather than looking for ready-to-use code or other artifacts. They also affirm the need to critically evaluate AI responses due to concerns about precision and trustworthiness (Xiao et al. 2024), suggesting that this new technology is particularly useful as a tool to be used by workers with at least some basic knowledge of the task to be performed.

However, whether generative AI will help or hurt the employment prospects of less-experienced software developers will depend on whether it is seen as a substitute or a complement for their services and if this shift is persistent versus transitory. In general, increasing labor productivity generally benefits more highly skilled workers, who can leverage technology and innovation to become even more efficient. In contrast, low-skilled workers often experience job displacement or wage stagnation as automation and technology replace routine tasks. These countervailing forces can increase inequality between workers with different skillsets, especially within occupations where workers collaborate in teams. For example, as productivity initially rises with the introduction of generative AI, employers may substitute away from hiring lower-skilled software developers whose tasks can now be done using less labor and more technology. This may lead to greater inequality between senior and junior software developers, at least in the short term during the initial transition when adopting this new technology (Katz and Murphy 1992, Berman et al. 1998, Autor et al. 2003, Acemoglu 2025). However, as generative AI diffuses across the labor market, rising productivity may eventually reduce the overall cost of producing new software, thereby leading to an increased demand for new software products, and hence a greater need for both junior and senior software developers in the long run.

Using the public release of Chat-GPT in November 2022 as a natural experiment, we explore the effects of generative AI on the software developer job market, documenting three new stylized facts. First, using the near-universe of online job vacancies collected by Lightcast (formerly Burning Glass Technologies), we find that generative AI resulted in a large and significant reduction in the relative proportion of junior-level (requiring less than 4 years of experience) versus senior-level (requiring 4 or more years of experience) software developer job vacancies immediately after the release of ChatGPT in November 2022. Our differences-in-differences estimates show that during the 12 months after the public release of Chat-GPT in November 2022, the number of software developer job postings requiring less than four years of experience fell by 16.3 percent relative to that of more senior level positions. This finding holds even when accounting for month and location fixed effects, controlling for firm size, industry composition and population, and considering the broader labor market trends affecting all computer and mathematical workers during this time period. In the short term, this shift may put downward pressure on both the employment and wages for junior versus senior software developers, especially for new computer science graduates, as the supply of entry-level workers outstrips demand.

Second, we explore the mechanisms by which the widespread introduction of generative AI differentially impacted junior software developers. Using a shift-share analysis, we also confirm that the aggregate fall in demand for junior software developers did not result from the shifting composition of employers who were posting positions but rather from a shift in the experience requirements that employers were looking for. However, there are important nuances within specific types of jobs that suggest these effects are not uniform throughout the software sector. We also examine the heterogeneity of these effects across industry, job title, and geography. For the most part, we find that the relative decline in labor demand for junior- versus senior-level software developers appears to have occurred among larger firms and bigger cities while industries with moderate exposure to the software industry were more insulated from the introduction of ChatGPT.

Finally, we use network methods to reveal new career pathways for potentially displaced junior software developers using an occupation similarity network. Although early career software developers are skilled in many different ways, their inexperience puts them at a disadvantage in the post-LLM software developer job market. As a result, a significant portion of the 100,000+ annual computer science college graduates (National Center for Education Statistics 2022) may not find work as a software engineer upon graduation in the short term. Indeed, the Occupational Employment Statistics data show that employment growth slowed between 2022 and 2023 and actually declined in 2024. Identifying employment opportunities in other occupations that require a similar set of skills can help displaced workers apply their qualifications to new positions, boosting their potential for career advancement and income. To ensure accuracy, our model relies on the relationship between skill combinations associated with a particular worker, rather than assuming that these skillsets are just a linear sum of individual skills (Arntz et al. 2017). This approach identifies credible job opportunities for junior software developers that require minimal re-skilling such as Management Analyst, Industrial Engineer, Market Research Analyst, and Project Manager. Each of these roles require some experience with mining large datasets generated by high-frequency interactions collected by sensors, platforms, or administrative sources to generate actionable insights for manufacturers, businesses, and governments.

Our results contribute to several branches of literature on the impact of technology in the labor market. First, our findings speak to the ongoing discussion on the effects of automation and augmentation on workers (Autor 2015, Acemoglu and Restrepo 2018b). Inexperienced software developers may experience large productivity gains in routine coding tasks from generative AI (Cui et al. 2024, Peng et al. 2023), that makes them more vulnerable to displacement compared to their senior colleagues who are likely to be engaged in more creative and/or complex tasks. This is evidenced by the notable split that we observe in hiring preferences for roles requiring fewer than four years of experience compared to those requiring four or more years of experience.

While Frey and Osborne (2017) predict high “computerization” for many routine-based occupations, we find that the reality is more nuanced. Most software developers perform a mix of tasks, which lowers the overall likelihood of full-scale job displacement (Arntz et al. 2017). Moreover, large-scale technological adoption does not always reduce employment in the aggregate. For example, Graetz and Michaels (2018) found that robot adoption within industries boosted total factor productivity (TFP) without shrinking total employment — essentially growing the economic pie such that the employment loss from productivity gains was offset by increased economic activity. However, there is a transition period where less-skilled workers entering the labor market are likely to suffer some period of unemployment until the TFP gains can be realized through future economic growth.

Second, our findings also indicate that generative AI has shifted software developer hiring toward more experienced labor, in line with theories of skill-biased technological change (Autor et al. 2003, Acemoglu and Restrepo 2018a, Autor et al. 2024, Berman et al. 1998). Although organizations typically use generative AI to automate lower-level duties, this may ultimately increase the demand for more advanced problem-solving and interpersonal skills (Brynjolfsson et al. 2019) among junior developers in the long run. As a result, firms might raise their hiring criteria for entry-level candidates to seek individuals who can oversee, troubleshoot, refine, and design AI-based workflows. These trends would reduce the number of traditional coding roles that were once open to junior developers, and signal a broader labor market shift toward favoring higher-level skills for future software developer job vacancies.

Third, our results show *which* inexperienced developers are more susceptible to generative AI-driven displacement. This evidence aligns with other research on AI adoption, which shows that larger and more advanced firms often lead in embracing new technologies (Acemoglu et al. 2022, Zolas et al. 2020), which points to a broader, market-wide shift triggered by generative AI's capabilities that would affect more workers in the long-run.

2. Data, & Methods

2.1. Using Real-Time Labor Market Data

Lightcast (formerly Burning Glass Technologies) specializes in labor market analytics. They collect and maintain the near-universe of online job postings from over 50,000 sources. Each of the 440+ million unique U.S. job postings is processed and classified by relevant measures, including job title, six-digit Standard Occupation Code (SOC), North American Industry Classification System (NAICS) industry code, employer name, and location, as well as the required level of education, years of experience, and specific skillsets (e.g., Python). This allows researchers to study granular changes in occupational labor demand over time. This dataset has been used by both public and private sector organizations to inform consequential decisions such as salary rate[1], university curriculum changes[2], and regional small business investment[3] as well as researchers studying on employer skill requirements, labor market mismatch, and industry concentration (Modestino et al. 2016, 2020, Alekseeva et al. 2021).

Lightcast's dataset has been shown to accurately represent U.S. labor market demand. Research shows a strong positive correlation between Lightcast job postings and the Bureau of Labor Statistics (BLS) Job Openings and Labor Turnover Survey (JOLTS)[4] (Carnevale et al. 2014, Hershbein and Kahn 2018, Modestino et al. 2016, 2020, Cammeraat and Squicciarini 2021) in terms of geographical, occupational, and industrial trends over time. As expected, online job vacancies tend to slightly over-represent cognitive and white-collar industries and under-represent manual and blue-collar industries. Over-representation of the software and technology industries is advantageous for the context of this chapter, providing better sampling for these occupations.

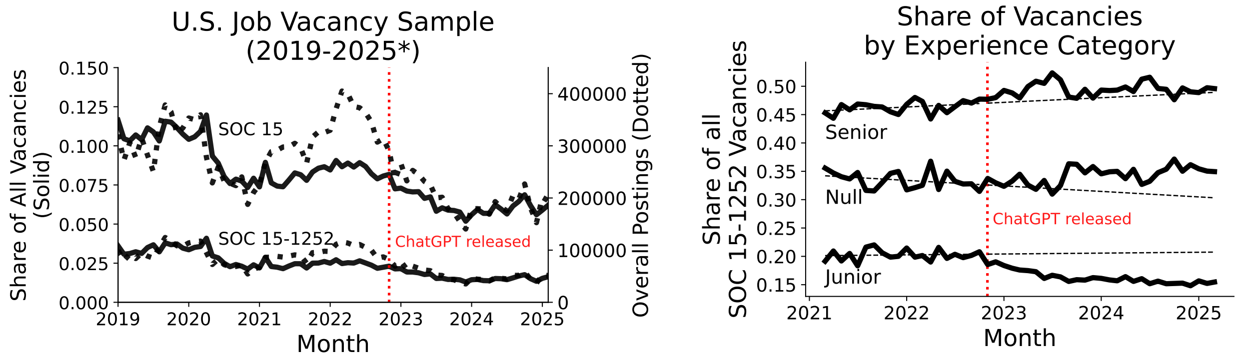
2.1.1. Sample Characteristics Our sample covers 20.3 million vacancies for Computer and Mathematical Occupations (SOC-15) and 5.7 million vacancies for software developers (SOC 15-1252) in the U.S., spanning from January 2019 to March 2025. This represents 8.3% and 2.3% respectively of the 246 million available job postings collected by Lightcast during that time period. Figure 1a shows the yearly values for these categories. As expected, we see a surge in the overall number of job postings post-pandemic (2021-2022) that eventually declines (2023-2024) yet remains above the pre-pandemic level (2019), indicating a continued strong labor market during this period. Although the number of postings for computer and mathematical occupations (SOC 15) follows a similar pattern as the aggregate economy, the drop-off in 2023 is nearly twice as steep–falling 38 percent from 2022 compared to a decrease of only 20 percent across all occupations combined–and falls below its pre-pandemic level. The fall in software engineer (SOC 15-1252) postings is even larger (-43 percent) and by 2024 is only half of the level that prevailed in 2019. Figure 1b shows that this decline was driven largely by a fall in junior-level postings. By November 2023, one year after the release of ChatGPT, the relative proportion of junior-level vacancies in SOC 15-1252 had dropped -4.7 percentage points below their January 2021-October 2022 trend line. In contrast, senior-level postings and vacancies with and no experience level specified were above their trend lines after the release of Chat-GPT by 1.6 and 3.1 percentage points respectively.

[1] `https://lightcast.io/case-study-jb-hunt`

[2] `https://lightcast.io/case-study-princeton-graduate-school`

[3] `https://lightcast.io/case-study-indy-accelerate-skills`

[4] `https://www.bls.gov/jlt/`



(a) (b)

Figure 1 **(a) Overall number and share of total U.S. job postings accounted for by Computer and Mathematical Occupations (SOC-15) versus Software Developers (SOC 15-1252) from 2019 through 2025. *The data for 2025 includes only the period from January to March. (b) Monthly posting ratios by experience category for software developers, with trend lines estimated from March 2021 to October 2022 data. Following the release of ChatGPT (the red dashed line), the proportion of postings requiring 0–3 years of experience declines sharply. This shift reflects a notable contraction in junior-level opportunities rather than an increase in senior-level demand.**

Researchers have used the required years of experience in a job vacancy to distinguish between “high-skill” and “low-skill” demand for workers (Modestino et al. 2016, 2020). For example, take the job titles “Software Engineer” and “Staff Software Engineer.” A software engineer (SE) can advance to Staff SE, so we expect Staff SE to have a higher experience requirement than SE. The Lightcast data accurately captures this nuance with Staff SE vacancies requesting 7.58 years of experience on average, while a SE vacancy only requested 4.89 years, as of February 2025.

Required education has also been used to distinguish high-skill from low-skill roles, such as whether the posting requires a college degree or asks for a particular certification. However, 93.5% of vacancies for software engineers ask for a Bachelor’s degree or do not list required education during our period of interest (one year before and after the release of ChatGPT). The opposite is true of certifications which are less common, with only 16.4% of software developer vacancies requiring at least one certification and only ten unique certifications for junior-level vacancies appearing in fifty or more vacancies in November 2024. Four of these certifications are related to security clearance, which likely do not reflect skills related to coding that we would expect to be enhanced or replaced by AI. For now, we ignore certifications in our analysis, although future research studying the impact of generative AI on the labor market may benefit from using required certifications as a measure of skill in other occupations where they are more commonly required for tasks that might be suitable for using LLMs.

2.1.2. Robustness Check: Classification of Unlabeled Experience Levels One common but little explored issue with using real-time labor market data is the high number of job postings that do not state any experience requirements, which is distinct from postings that explicitly state "no experience required". Roughly 54.7% of all Lightcast job postings in our sample, including 35.7% of software engineer postings, have no stated experience requirements. Prior researchers have either assumed these postings require no experience (Hershbein and Kahn 2018) or left these job postings as unassigned (e.g., missing) year of experience (Modestino et al. 2020) in their analyses. Neither of these approaches is ideal for our context, given the high share of postings that do not have any stated experience requirements and the importance of years of experience in our analysis.

To explore whether job postings with no years of required experience (which is different than zero years) could bias observed differences between junior and senior positions, we built a logistic regression model to classify job postings into "0-3 required years of experience" (junior) or "4+ required years of experience" (senior) categories. We refer to this subset of postings as "unlabeled" and postings with stated experience requirements as "labeled". Logistic regression was chosen specifically for its interpretability. It allows a clear and straightforward understanding of how each feature contributes to the classification outcome.

As inputs, our model takes individual vacancies with listed skills, job title, education, salary, hiring firm, month, year, and geographic indicators. To incorporate each posting's list of required skills, we use one-hot encoding for the 1000 most common skills. This method represents each skill with a unique position in a 1000-entry vector. When a job posting lists a specific skill, the vector value at that skill's position will equal one. All skills not in the job posting will have a value of zero. This common technique creates a machine-interpretable representation of the total skill landscape. The model output provides the estimated years of required experience for the job posting as either "0-3 required years of experience" or "4+ required years of experience".

The logistic regression classifier was trained on a random sample of the postings for SOC 15-1252 (Software Developers) within our period of interest, 2021-2023. Table 1 presents the results, showing precision, recall, and F1-scores by experience category. The overall accuracy on the test set is 0.74. The model has balanced discrimination between junior and senior categories (precision of 0.73 for junior and 0.75 for senior). These results validate that the logistic regression is reasonably accurate.

We conducted a sensitivity analysis to quantify any uncertainty introduced by misclassification. The number of monthly unlabeled postings ranges from 12,514 to 41,005. The proportion of estimated senior roles each month ranges from 37.7% to 49.3%. Applying a binomial proportion

Table 1 **This table reports the classification performance on experience levels, confirming that the classifier achieves a moderate level of success in distinguishing between the two experience groups.**

	Precision	**Recall**	**F1-Score**	**Support**
0-3 years experience	0.73	0.76	0.74	69477
4+ years experience	0.75	0.71	0.73	69921
Accuracy			0.74	139398
Macro Avg	0.74	0.74	0.74	139398
Weighted Avg	0.74	0.74	0.74	139398

calculation, we obtained a standard error of 0.0006. The results in a 95% confidence interval of [0.419, 0.421] under the assumption of no misclassification.

Given the classifier's 74% accuracy, we calculated monthly margins of error for our estimates of senior roles (4+ years experience) in the unlabeled set using a binomial proportion method. Across months, these margins of error range from 0.48% to 0.87%, indicating a high degree of confidence in our estimates. These calculations confirm that the uncertainty introduced by misclassification is relatively minor.

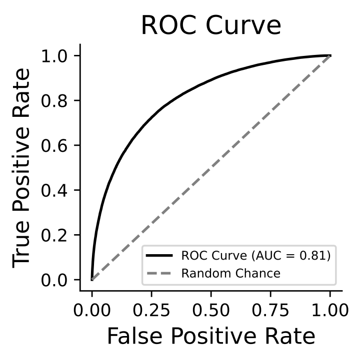


Figure 2 **The receiver operating characteristic curve for the logistic regression classifier with an AUC=0.81. This indicates the experience level classification has a satisfactory false positive and true positive rate.**

2.2. Empirical Methodology

2.2.1. Difference-in-Differences Analysis of the Introduction of ChatGPT We employ a difference-in-differences (DiD) design to estimate the impact of ChatGPT's public release (Version 3.5, November 2022) on software developer job vacancies. This well-studied causal inference technique allows us to separate the effect of an intervention for affected and comparison groups in *pre* and *post* periods assuming both groups show parallel trends in the *pre* period (Bertrand et al. 2004). Software developer vacancies requiring fewer than four years of experience (junior-level) are our affected group. Vacancies requiring four or more years of experience (senior-level roles) are our comparison group. The rapid adoption of ChatGPT beginning in November 2022 is a natural experiment that acts as our intervention. This defines the *pre* and *post* periods. We define a binary indicator, $PostChatGPT_t$, equal to 1 for months on or after November 2022, otherwise, this indicator takes a value of 0. We conduct our analysis at the Combined Statistical Area (CSA) level to account for differences in technology adoption across both place and time. Our main estimating equation is,

$$\ln Vacancies_{cjm} = \alpha + \delta\, Junior_{cjm} + \theta \left(PostChatGPT_m \times Junior_{cjm} \right) + Month_m + CSA_c + X_{cjm} + \epsilon, \tag{1}$$

where for CSA c, job title j, and month m, let ln *Vacancies* be the natural log of the number of software developer vacancies for a particular cell. *Junior* is a binary variable which takes a value of 1 for junior-level postings or 0 for senior-level postings. The coefficient on the interaction term, θ, captures the difference-in-differences (DiD) estimate of the effect of ChatGPT's release on junior-level postings relative to senior-level postings. If junior vacancies experience a similar trend from *pre-* to *post*-ChatGPT as senior vacancies, then θ will be approximately 0 or show no statistically significant difference between the two groups. If junior postings increase relative to senior postings in the *post* period, then θ will be positive and the reverse will be true if junior postings decrease relative to senior postings.

We include fixed effects for both time ($Month$) and location (CSA). Finally, we also include more aggregated controls X to account for differences in firm size (by quintile), population (by quintile), and industry composition–held constant at their *pre-ChatGPT* values (Nov. 2021 - Oct. 2022). We estimate the impact of the introduction of generative AI by examining the period one year before to one year after the public release of version 3.5 in November 2022. We report sample summary statistics in Appendix Table A.1 and control variable summary statistics in Appendix Table B.1. Standard errors are clustered at the job title level to account for potential correlation in outcomes within job titles over time.

2.2.2. Shift-Share Analysis of Experience Requirements To investigate whether the change in the required years of experience within SOC 15-1252 is driven by a shift in employer preferences for different job titles or by an increase in the experience requirements for the same roles, we employ a shift-share analysis comparing the periods *pre-* and *post-ChatGPT*.

Let E_{pre} denote the average required years of experience in SOC 15-1252 *pre-ChatGPT*, and E_{post} denote the same *post-ChatGPT*. These overall averages can be expressed as the weighted average of the experience requirements for each job title within software developers, where the weights are the proportion of each job title in the total number of job postings.

Formally, let n be the number of distinct job titles within SOC 15-1252. For each job title $i \in \{1, 2, ..., n\}$, let $s_{i,pre}$ and $s_{i,post}$ be the share of job title i in the total number of job postings *pre-* and *post-ChatGPT*, respectively. Similarly, let $e_{i,pre}$ and $e_{i,post}$ be the average required years of experience for job title i *pre-* and *post-ChatGPT*, respectively. Then, the overall average required years of experience in each period is given by:

$$E_{pre} = \sum_{i=1}^{n} s_{i,pre} \cdot e_{i,pre}; \qquad E_{post} = \sum_{i=1}^{n} s_{i,post} \cdot e_{i,post}$$

To decompose the change in the overall average required years of experience ($\Delta E = E_{post} - E_{pre}$), we construct two counterfactual scenarios:

1. **Counterfactual 1: Holding Constant the Composition of Job Titles.** We calculate the overall average required experience across all postings if the distribution of job titles remained at its *pre-ChatGPT* level, but the experience requirements for each title changed to their *post-ChatGPT* values. This counterfactual average is:

$$E_{cf1} = \sum_{i=1}^{n} s_{i,pre} \cdot e_{i,post}$$

The difference between this counterfactual and the *pre-ChatGPT* average ($E_{cf1} - E_{pre}$) indicates the change attributable to new experience demands within the same job titles.

2. **Counterfactual 2: Holding Constant Experience Requirements.** We calculate the overall average required experience across postings if the experience requirements for each job title remained at their *pre-ChatGPT* level, but the distribution of job titles changed to its *post-ChatGPT* values. This counterfactual average is:

$$E_{cf2} = \sum_{i=1}^{n} s_{i,post} \cdot e_{i,pre}$$

The difference between this counterfactual and the *pre-ChatGPT* average ($E_{cf2} - E_{pre}$) indicates the change attributable to a shift in the mix of job titles being posted.

By comparing these counterfactuals to the actual change in the overall average required years of experience, we can assess the relative contributions of changes in job title composition versus changes in experience requirements for specific job titles. Our analysis focuses on which counterfactual better approximates the true post-ChatGPT average, E_{post}.

2.3. Skills Analysis for Job Displacement

Given the observed decline in the overall number of software job postings (Figure 1a), and the potential shift in the ratio of senior to junior software developer vacancies, we might expect to see relatively fewer job openings for recent computer science graduates. If this is the case, then a significant number of these computer science graduates could be excluded from consideration for, or potentially displaced from, junior software developer roles–at least in the short term. Even if the increased productivity from AI eventually leads to greater economic growth and a total increase in the demand for software engineers—both junior and senior–where these displaced workers initially find employment upon graduating into a post-ChatGPT labor market may result in long-lasting labor market scarring (Kahn 2010). To address this issue, we compare the existing skill sets associated with pre-ChatGPT software developer job postings to those of other vacancies throughout the economy to construct a network of occupations connected by their similarities. This network could serve as a recommender system for recent computer science graduates to target related occupations in their job search and/or make appropriate career transitions if they are laid off.

Occupation Co-Occurrence Network Researchers have used skill co-occurrence networks to show the value of diverse skill sets (Anderson 2017, Stephany and Teutloff 2024), assess the growing cognitive-physical skill polarization (Alabdulkareem et al. 2018), and to predict worker mobility (Frank et al. 2024). A skill co-occurrence network maintains the information between skills. Nodes are individual skills, and edges are the probability of co-occurrence of skills along job vacancies or job applicants. We adapt the skill co-occurrence network methodology to create an Occupational Co-Occurrence Network.

As the dataset and network density grow, we must prune away edges to maintain the network's usefulness. Following Alabdulkareem et al. (2018), we start with the frequencies of each skill for each occupation. This is the raw number of vacancies for an occupation in Lightcast asking for a specific skill divided by the total number of vacancies for that occupation. Creating edges between every occupation with co-occurring skills would result in a network too dense for useful analysis. Instead, we calculate the revealed comparative advantage of each skill for each occupation. This is similar to the method used in Hidalgo et al. (2007) to form a network of products in international trade:

$$rca(\text{occupation, skill}) = \frac{\text{Frequency of a skill in a specific occupation}}{\text{Frequency of a skill across all occupations}}$$

Common skills receive low scores while more distinct skills receive high scores. Next, skills are "effectively used", $e(\text{occupation}, \text{skill}) = 1$, when $rca(\text{occupation}, \text{skill}) > 1$. Otherwise $e(\text{occupation}, \text{skills}) = 0$. Each occupation is left with its set of defining skills. Finally, we calculate the conditional probability that two occupations o and o' share skills s,

$$\Theta(o, o') = \frac{\sum_s e(s, o)e(s, o')}{\max\left(\sum_s e(s, o), \sum_s e(s, o')\right)}$$

Let occupations be nodes and $\Theta(o, o')$ denote edge weights in our occupational co-occurrence network. We have 734 nodes representing each unique SOC 6 occupation and 263,551 edges with non-zero θ values. Figure 3 shows how the majority of edge weights are small, with only 8.4% (22,012) exceeding 0.10 and only 0.58% (1,519) exceeding 0.20. This means different occupations share only a few effectively used skills and that only a minority of occupations are similar, suggesting that our model will be able to make useful recommendations for junior software developers seeking to apply their skills in other occupations.

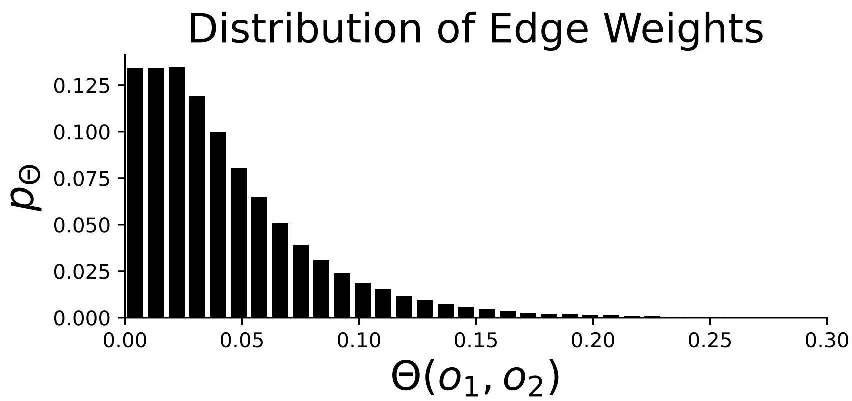


Figure 3 Most edge weights in the occupational co-occurrence network are small, indicating few occupations have high conditional co-occurrence probability. Only a few occupation pairs are truly similar, which is essential to provide targeted career transition recommendations.

3. Impact of Generative AI on Junior Software Developer Hiring

Immediately following ChatGPT's introduction in November 2022, Figure 1(b) showed that employer demand fell sharply for less experienced software developers. The magnitude of this abrupt change was on par with that observed during the onset of the COVID-19 pandemic. Junior-level positions experienced a pronounced decline, deviating from their pre-ChatGPT hiring trends. Statistical analyses confirm that this shift was a meaningful departure from the prior trend, revealing significant reductions in junior software developer vacancies post-ChatGPT. Senior-level hiring also declined, but at a rate similar to that which was occurring before the public release of ChatGPT. These recent trends set the stage for our more detailed exploration of how generative AI may have reshaped hiring patterns within the software development field.

3.1. Trends in the Ratio of Senior to Junior Software Developer Vacancies

The relative trajectory of senior versus junior software developer job postings shows a rapid increase coinciding with the introduction of ChatGPT in November 2022. As shown in Figure 4, almost immediately upon the public release of ChatGPT version 3.5, there is an abrupt and clear increase in the ratio of senior (four or more years of required experience) relative to junior (fewer than four years of required experience) software developer vacancies. In comparison, a similar rapid disruption of comparable magnitude also appeared from March to April 2020 during the onset of the COVID-19 pandemic. During that prior period, the pandemic-induced recession led to a large drop in vacancies across the entire economy, not just for software developers. Prior research has shown that employers tend to increase skill requirements for both education and experience in response to the greater availability of skilled workers during business cycle downturns (Modestino et al. 2020). This opportunistic response also tends to reverse itself as the labor market recovers, with employers lowering skill requirements as workers become more scarce, particularly for tasks that can be learned on the job (Modestino et al. 2016). However, unlike the subsequent recovery after the disruption caused by the COVID-19 pandemic, the job market for junior software developers has not yet recovered after the introduction of generative AI.

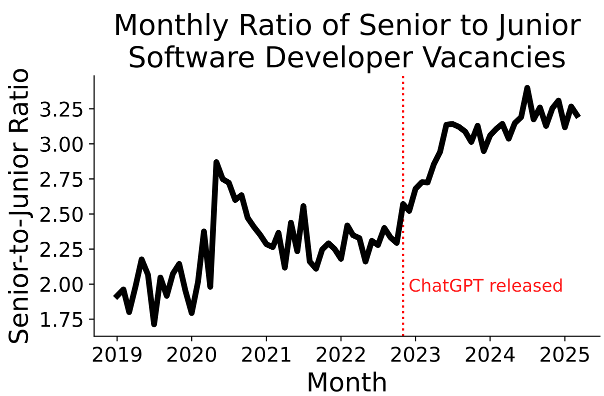


Figure 4 Monthly ratio of senior to junior software development job vacancies from 2017 to March 2025. The ratio sharply increases following the public release of ChatGPT 3.5 in late 2022, marked by the red dashed line. This spike in demand for senior roles is similar to the earlier disruption seen during the onset of the COVID-19 pandemic in 2020 which caused a sharp contraction in labor demand across the U.S. economy.

Our statistical analysis confirms that the ratio of junior to senior job postings is significantly different in the pre- versus post-LLM eras. In our difference-in-differences regression, the interaction term for $ShareJunior \times PostChatGPT$ is large, negative, and significant ($\theta = -3.490$, $p = 0.023$, Model 1, Table 2), indicating that the public release of Chat-GPT reduced the number of junior software job vacancies by 16.3 percent relative to more senior roles over the subsequent 12 month period. This sharp decline in the relative number of postings for junior versus senior software developers after ChatGPT's public release aligns with long-standing theories on skill-biased technological change (Katz and Murphy 1992, Autor et al. 2003, Acemoglu and Restrepo 2018a). As generative AI improves the productivity of lower-skilled software developers more than experienced developers (Peng et al. 2023, Cui et al. 2024), firms appear to respond by rapidly changing their hiring strategy (Kinder et al. 2024). In November 2022, employer demand for inexperienced developers drops as some portion of their labor is likely substituted by this new technology. The increased productivity associated with the introduction of Chat-GPT on routine coding tasks decreases the need to recruit junior software developers. The relative magnitude of this shift parallels that of the COVID-19 labor market shock, suggesting that generative AI has had a meaningful impact on hiring for junior software developers.

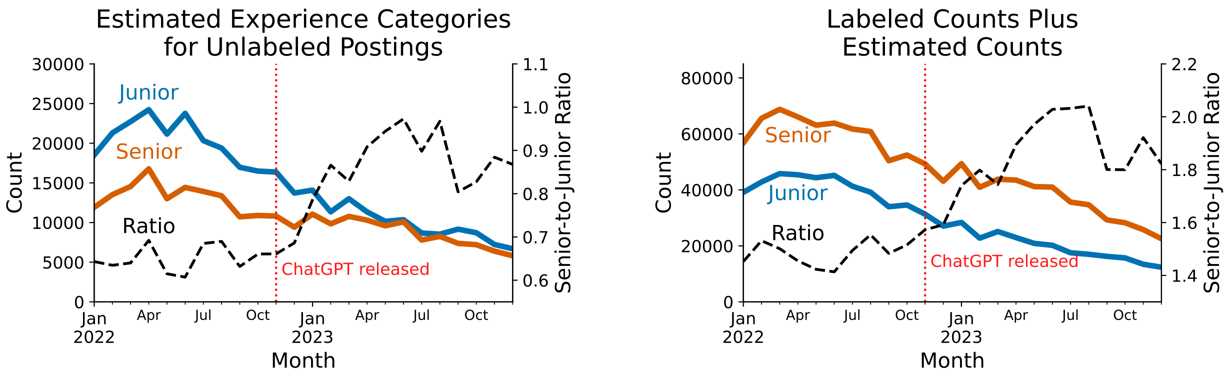
Table 2 Difference-in-differences regression estimates of the impact of ChatGPT's public release on junior relative to senior software developer vacancies. Observations are based on a total of 575,397 underlying individual job vacancies aggregated to 54,456 job title x month x location cells. Across all specifications, the PostChatGPT x Junior interaction is negative and statistically significant, indicating a meaningful decline in junior relative to senior software developer job postings after the widespread introduction of generative AI.

	Dependent variable: Log Number of Postings				
	(1)	(2)	(3)	(4)	(5)
Junior	-0.760***	-0.800***	-0.760***	-0.760***	-0.763***
	(0.074)	(0.070)	(0.077)	(0.069)	(0.067)
PostChatGPT × Junior	-0.151***	-0.167***	-0.151***	-0.152***	-0.154***
	(0.027)	(0.029)	(0.028)	(0.028)	(0.028)
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes
Job Title Fixed Effects		Yes			
CSA Fixed Effects			Yes		
Population Quintile Controls				Yes	Yes
Firm Size Quintile Controls				Yes	Yes
Industry Mix Controls					Yes
Observations	54456	54456	54456	54456	54456
Adjusted R^2	0.168	0.305	0.192	0.170	0.178

Note: *p<0.1; **p<0.05; ***p<0.01
Each coefficient listed is from a separate regression as specified by equation 1 in the text. Each regression uses the first sample listed in Appendix Table A.1 that aggregated data from all job postings into Job Title × Month x CSA cells containing at least 5 total postings. Cells that are missing a CSA code or are located in Guam or Puerto Rico are excluded from the analysis. Baseline controls include pre Chat-GPT characteristics for each cell including the average firm size across job titles by quintile, the average population size across by quintile, and the industry mix by two-digit NAICS code.

3.2. Accounting for Vacancies with No Required Experience

To account for potential measurement error, we use our fitted logistic regression model to estimate the experience categories for job vacancies that have no required experience. This is because job vacancies that do not explicitly state experience requirements (e.g., management positions that recruit base on skills-based hiring) may differ in important ways from job vacancies explicitly



(a) (b)

Figure 5 (a) Counts of previously unlabeled job postings categorized by experience using a logistical model. The senior-to-junior ratio sharply increases immediately after the public release of ChatGPT version 3.5 that is similar to what is observed using only postings that explicitly list years of experience. (b) Adding the estimated counts of previously unlabeled postings to those that explicitly list experience requirements does not change the timing nor the magnitude of the observed shift between senior and junior vacancies shown in the prior figure.

listing no required experience (e.g., zero years for an entry-level software developer). Figure 5a, shows that our logistical model labels anywhere from 30.3% to 40.6% of software developer postings with no required experience as experienced software developer jobs. Categorizing these previously non-labeled postings we find a clear spike in the estimated senior-to-junior vacancy ratio after the public release of Chat-GPT in November 2022 that is similar to what we documented using only the explicitly experience-labeled vacancies. Adding these previously unlabeled vacancies to our prior counts of experience-labeled vacancies does not change the timing nor reduce the magnitude of the hiring shift from junior to senior postings beginning with the widespread introduction of generative AI (Figure 5b).

The broad consistency of these results confirms that the observed contraction in junior software developer roles that coincides with the introduction of generative AI, does indeed reflect a change in employer demand, rather than a change in how employers write their job vacancies. Employers may have strategically avoided labeling "entry-level" thresholds while still screening for more experienced candidates. By reframing or leaving experience fields blank, firms might seek to recruit candidates with higher skill sets without admitting to a change in job requirements. The observed similarity in the post-ChatGPT shift in the relative demand for senior versus junior software developers, even after imputing experience levels for "unlabeled" vacancies, confirms that the hiring shift is does not reflect changes in how employers phrase or omit experience requirements when writing job descriptions.

3.3. Comparison to other Computer and Mathematical Occupations

Next, we compare job posting trends for software developers to the rest of the computer and mathematical occupations (SOC 15) to test whether junior software developers were disproportionately impacted by the widespread introduction of generative AI. This is important for two reasons. First, although other computer and mathematical roles also require skills such as problem-solving or programming, software development is particularly exposed to LLMs' capacity to produce and debug code (Eloundou et al. 2024, Cui et al. 2024) and we want to accurately capture those specific impacts. Second, during the subsequent recovery from the COVID-19 recession, postings for computer-related occupations trended downward after the prior surge in demand for workers, including software developers, who could help build, launch, and maintain online systems for business that pivoted to conducting more of their business remotely, often using online tools to automate ordering, tracking, billing, and shipping processes. Although all workers in computer and mathematical occupations, including software developers, were impacted by this widespread decline in aggregate demand post-COVID, if software developers were additionally impacted by the public release of Chat-GPT in November 2022, then we should expect to see a steeper decline in the number of junior software developer postings compared to entry level positions in other computer and mathematical occupations.

Figure 6 compares the trend in software developer job postings to that observed for all other computer and mathematical occupations in SOC 15 (excluding software developers). Specifically, we examine the ratio of job postings in SOC 15-1252 to all other SOC 15 postings for junior, senior, and null experience levels. We find that junior-level postings (those requiring fewer than 4 years of experience) show a relatively larger decrease post-ChatGPT among software developers compared to other computer and mathematical workers the ratio that is observed for senior-level postings. These trends suggest that the relative decline in junior versus senior software developer postings that we previously documented immediately after the release of ChatGPT does not simply reflect a broader decline in aggregate demand for all junior positions in the computer and mathematical sector.

To better quantify the impact of generative AI on junior software developers relative to entry-level positions in other computer and mathematical occupations, we use a difference-in-difference-in-differences (DiDiD) approach as specified in Equation (2):

$$\begin{aligned}\ln \mathit{Vacancies} = &\alpha + \beta\ \mathit{Junior} + \delta\ \mathit{Software} + \xi\ \left(\mathit{Junior}_t \times \mathit{Software}\right) + \\ &\theta\ \left(\mathit{PostChatGPT}_t \times \mathit{Junior}\right) + \zeta\ \left(\mathit{PostChatGPT}_t \times \mathit{Software}\right) + \\ &\eta\ \left(\mathit{PostChatGPT}_t \times \mathit{Junior} \times \mathit{Software}\right) + \mathit{Month} + CSA + X + \epsilon,\end{aligned} \tag{2}$$

where ln *Vacancies* is the log number of vacancies for each cell. There are three binary indicators. *Junior* equals 1 for cells that require three or fewer years of experience, *Software* equals 1 for

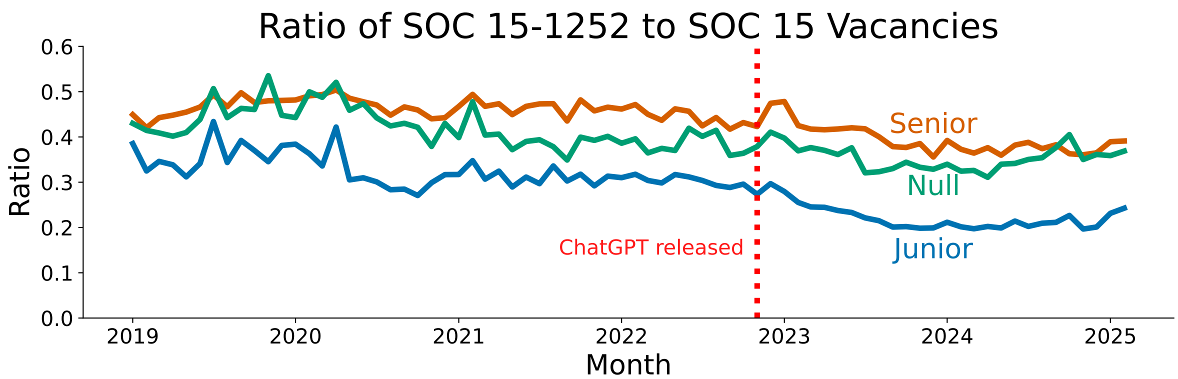


Figure 6 Ratio of SOC 15-1252 to SOC 15 job postings for junior, senior, and null (aero experience) experience levels. Junior-level postings show a relatively larger decrease in this ratio compared to senior and null postings after the public release of ChatGPT, suggesting a differential impact of generative AI on software developers compared to all other workers in computer and mathematical occupations.

software developer positions in SOC 15-1252, and $PostChatGPT$ equals 1 for cells with job postings on or after November 2022. As in Equation (1), the variable $Month$ captures time fixed effects and CSA captures fixed effects by geography denoted by Combined Statistical Areas. Finally, we include X to control for firm size quintile, population size quintil, and industry mix held constant at their *pre-ChatGPT* aggregate values (Nov. 2021 - Oct. 2022). Comparing trends one year before to one year after the release of ChatGPT version 3.5 in November 2022, the coefficient η estimates the differential change over time in the log number of job vacancies for junior versus senior positions among software developers (SOC 15-1252) relative to the change over time in log number of job vacancies for junior versus senior positions among all other computer and mathematical occupations (SOC 15). Thus, η captures the unique effect of ChatGPT on junior software roles beyond the effects of ChatGPT on junior roles in other Computer and Mathematical Occupations.

Table 3 shows that η is indeed negative and significant, confirming that junior software developers were more impacted by the widespread introduction of generative AI than entry level workers in other computer and mathematical occupations. Moreover, the magnitude of the coefficient is relatively similar to that from our prior DiD estimate in Table 2, indicating that the public release of Chat-GPT reduced the relative number of junior versus senior software developer postings by 12.5 percent during the subsequent 12 months compared to the broad aggregate trends affecting entry-level positions across other computer and mathematical occupations. Our finding that junior developers bear the brunt of this new technology's impacts on the demand for labor echoes what Arntz et al. (2017) call "task reconfiguration." Technology displaces certain tasks or routines while leaving other tasks intact or even creating new ones. Based on industry reports, the task overlap between junior software developers and generative AI is higher than the rest of SOC 15 (Kinder et al. 2024), which supports these findings.

Table 3 Difference-in-difference-in-differences regression for junior software developer postings (SOC 15-1252) versus those in all other computer and mathematical occupations (SOC 15). The negative and significant coefficient on the triple-interaction term indicates a relatively larger decline for junior versus senior software developers vacancies compared to junior versus senior roles in other related occupations. PostChatGPT .

	Dependent variable: Log Number of Postings				
	(1)	(2)	(3)	(4)	(5)
Junior	-0.379***	-0.389***	-0.376***	-0.358***	-0.353***
	(0.081)	(0.084)	(0.083)	(0.080)	(0.080)
Software	0.534***	0.418***	0.667***	0.524***	0.516***
	(0.112)	(0.105)	(0.145)	(0.111)	(0.116)
PostChatGPT × Junior	-0.155***	-0.153***	-0.155***	-0.159***	-0.161***
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
PostChatGPT × Software	-0.090***	-0.100**	-0.112***	-0.096***	-0.098***
	(0.033)	(0.039)	(0.039)	(0.033)	(0.033)
PostChatGPT × Junior × Software	-0.136**	-0.140**	-0.138**	-0.122**	-0.118**
	(0.058)	(0.059)	(0.059)	(0.056)	(0.056)
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes
Job Title Fixed Effects		Yes			
CSA Fixed Effects			Yes		
Population Quintile Controls				Yes	Yes
Firm Size Quintile Controls				Yes	Yes
Industry Mix Controls					Yes
Observations	66572	66572	66572	66572	66572
Adjusted R^2	0.106	0.186	0.168	0.110	0.114

Note: *p<0.1; **p<0.05; ***p<0.01
Each coefficient listed is from a separate regression as specified by equation 2 in the text. Each regression uses the first sample listed in Appendix Table A.1 that aggregated data from all job postings into Job Title × Month x CSA cells containing at least 5 total postings. Cells that are missing a CSA code or are located in Guam or Puerto Rico are excluded from the analysis. Baseline controls include pre Chat-GPT characteristics for each cell including the average firm size across job titles by quintile, the average population size across by quintile, and the industry mix by two-digit NAICS code.

4. Heterogeneity of Generative AI Impacts on Software Developers

To better understand the mechanisms by which the widespread introduction of generative AI differentially impacted junior software developers, we explore the heterogeneity of these effects in three ways. First, we conduct a shift-share analysis across job titles within the software developer occupation (SOC 15-1252) to test whether the observed shift away from junior and towards senior level postings for software engineers is due to changing experience requirements *within* junior job software titles or the changing composition of job postings *across* junior and senior software job titles. Second, we estimate impacts by industry and geography to better assess whether this shift away from junior software engineering roles is concentrated in certain economic sectors (e.g., information or professional and scientific services) or regions of the country (e.g., Silicon Valley or Austin, TX) where software development is more prevalent and/or expanding. Third, we examine which specific skills are in greater demand for junior software developer positions to determine whether acquiring new skills to work with generative AI will make entry level workers more attractive in the post-Chat-GPT labor market.

4.1. Shift Share Analysis by Job Title

The Lightcast data contain a Titles Taxonomy which is a proprietary system that standardizes over 20 million unique job titles into a more manageable set of about 75,000 standardized titles. Within the Software Developer occupation (SOC 15-1252), there are more than 200 Lightcast job titles that contain at least 50 or more unique job postings per month on average. These include titles that that fall primarily within either the junior (e.g., Software Engineer) or senior (e.g., Principal Software Engineer) roles as well as those that have a mix of junior and senior roles (e.g., Software Development Engineer)–all of which involve researching, designing, and developing computer and network software or specialized utility programs. To varying degrees workers in these roles analyze user needs and develop software solutions, applying principles and techniques of computer science, engineering, and mathematical analysis to update software or enhance existing software capabilities. Many of these positions involving working as part of a team, either with computer hardware engineers to integrate hardware and software systems or with other software engineers to develop specifications, performance requirements, and maintain databases within an application area.

Given this heterogeneity across job titles within the software developer occupation, employers have some degree of flexibility to reconfigure tasks and roles within teams when adopting new processes or technologies such as generative AI. Thus, it's an empirical question as to whether the shift in software developer requirements from requiring less than four years experience to four or more years reflects raising the bar for junior entry-level workers for whom Chat-GPT might be complementary to their existing skills or eliminating those positions entirely in favor of substituting towards senior software engineers who can use Chat-GPT to replace more junior team members on projects.

We find that the overall move toward increased experience requirements appears consistent across job titles. In Q3 2022, the average required experience for a job title was 5.30 years. By Q3 2023, it had increased to 5.64 years, a shift of 0.34 years. Our shift-share calculations indicate that the largest contributor (0.35 years) is the within-job-title effect, which reflects more employers increasing experience requirements within each job title. Only 0.01 years is attributable to a change in the mix of job titles. This pattern underscores that while hiring managers may still post the same titles, they increasingly prefer candidates with greater experience.

Figure 7 provides a closer look within the Lightcast Titles Taxonomy groupings to capture changes in experience both within and between specific job titles, revealing a more nuanced transformation within the software developer labor market. Largely junior roles such as Software Engineer exhibit large increases in experience requirements post-Chat GPT due to increasing years of experience requested *within* the underlying specific job titles. In contrast, more senior roles such as Principal Software Engineers actually show decreasing experience requirements, largely *across* specific job titles, perhaps because these workers are complementary to the adoption of Chat-GPT. Finally, titles that have a mix of junior and senior roles, such as Software Development Engineers, manifest both phenomena simultaneously occurring with increasing experience requirements within some job titles and decreasing experience requirements across others.

Although prior research suggests that when a new technology increases the marginal productivity of lower-skilled workers, employers may respond by reducing overall hiring for those roles due to the potential for substitution and increased efficiency (Autor 2015, Acemoglu and Restrepo 2018b), this does not appear to entirely be the case among junior software developers. Our shift-share analysis indicates that employers are seeking candidates with more experience *within* job titles associated with more junior roles. This suggests that employers may be shifting tasks across roles in response to generative AI, similar to other technological shifts that have transformed job requirements rather than entirely eliminated jobs (Arntz et al. 2017, Peng et al. 2023). In effect, the introduction of Chat-GPT has raised the bar for recent computer science graduates who may need to re-train by acquiring the skills needed to effectively use generative AI to gain entry into the profession.

4.2. Differential Impacts by Industry, Firm Size, and Geography

Given that new technology tends to diffuse rapidly but unevenly, we explore whether the public release of Chat-GPT in November 2022 affected the relative demand for junior versus senior software engineers across industries, firms of different sizes and geographies. One might imagine that the initial shift in experience requirements for software developers was more concentrated in certain economic sectors (e.g., information or professional and scientific services), firms that would benefit more from automating certain processes (e.g., larger firms) or regions of the country where software development is more prevalent and/or has been rapidly expanding (e.g., large CSAs). To estimate heterogeneous impacts separately across industry, firm size, and geography, We restrict our analyses to the period of November 2021 through October 2023, centering on ChatGPT's release in November 2022, and use the fully saturated model shown in Equation (3).

$$Ratio_t = \alpha + \beta PostChatGPT_t + \sum_{n}^{5} (\delta S_n + \theta_n (PostChatGPT_t \times S_n)) + \tau_t, \tag{3}$$

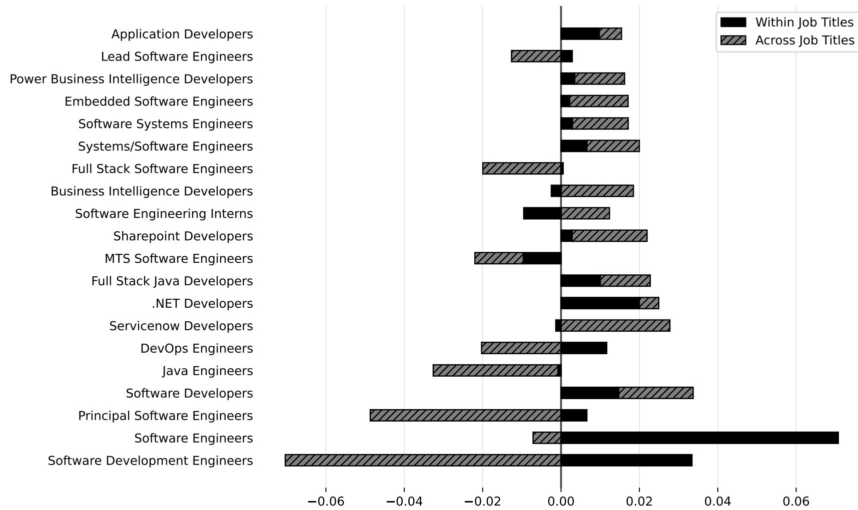


Figure 7 Decomposing the change in average years of experience pre- versus post-Chat GPT occurring within versus between job titles. Our shift-share analysis shows that junior roles such as Software Engineer exhibit increasing experience requirements within job titles compared to senior roles such as Principal Software Engineer that show decreasing experience requirements across job titles.

where $Ratio_t$ is the junior-to-senior ratio of job postings in month t, $PostChatGPT_t$ is 1 for any month on or after November 2022, S_n represents the specific subgroups within each area (e.g., NAICS industry groups[5], firm size quintiles, population quintiles) and τ_t are month-specific fixed effects. Overall. We find that while some industries are affected more than others, the increase in demand for more experienced software developers is fairly widespread across firms and CSAs of different sizes. We summarize these findings below and refer the reader to the Supplementary Materials for the individual estimates.

Our industry-level regressions reveal that although all industries experienced a general contraction in junior versus senior software developer vacancies starting in mid-2022, some industries were affected more than others (see Table D.1 in the Supplementary Materials). Perhaps not surprisingly, both Information (NAICS 51) and Professional, Scientific and Technical Services (NAICS 54)–industries that employ a disproportionately large share of software developers displayed post-Chat-GPT impacts that were twice as large as other sectors. However, Retail Trade (NAICS 44-45) and Accommodation and Food Services (NAICS 72)–industries that had rapidly expanded their use of software applications during the pandemic–also exhibited out-sized impacts. In contrast, the one industry that did not demonstrate any significant impacts during the post-period was Management of Companies and Enterprises (NAICS 55), perhaps because these are companies

[5] We drop industries that do not have at least 100 software developer vacancies per month. This eliminates Agriculture, Forestry, Fishing and Hunting (NAICS 11), Mining, Quarrying, and Oil and Gas Extraction (NAICS 21), and Arts, Entertainment, and Recreation (NAICS 71).

that primarily engage in influencing management decisions or actively managing other enterprises, including strategic or organizational planning which may have distinctive regulatory requirements, business models, or risk tolerances that mitigate or delay the impact of generative AI on their software hiring.

We check for heterogeneity across firm sizes by quintile using the total number of postings a company has made in Q3 2022 (pre-ChatGPT)). Only firms in the largest quintile show a significant impact of Chat-GPT on the ratio of junior versus senior level postings (see Table D.2 in the Supplementary Materials). This result aligns with prior research that finds larger firms are often first to adopt new technologies (Acemoglu et al. 2022, Zolas et al. 2020). More resource-rich organizations typically have the capacity and foresight to implement novel technologies like generative AI. Smaller firms might face different labor market realities. For instance, they may not have the same flexibility as larger firms to substitute junior labor with AI-augmented junior labor. In such cases, even with increased productivity from generative AI, the demand for junior developers might remain relatively stable or even increase among smaller firms might if they are constrained by development budgets in the short-term. If generative AI makes their existing junior-level developers more productive, this could translate into higher demand for their services where workers may be able to undertake more challenging projects in-house. This could explain the divergence in senior-to-junior ratios between differently sized firms.

We repeat our heterogeneity investigation across combined statistical areas (CSAs), excluding those with fewer than 100 monthly software developer postings was excluded from our analysis. We find no significant differential impact of Chat-GPT on any of the five population quintiles (see Table D.3 in the Supplementary materials). This suggests that the impact of more widespread use of generative AI on the relative demand for junior versus senior software developers is indeed driven by large cities. However, these effects appear to have a gradient whereby the impacts grow with population, suggesting that perhaps Chat-GPT, like many other technological advancements, will not necessarily remain confined to large, economically dominant cities, but quickly spread to smaller regions as well Zolas et al. (2020).

Table 4 quantifies the magnitude of these differential impacts, explicitly testing for significant differences across sub-groups within each area. For example, the first column tests whether the observed differences by firm size are statistically significant and confirms that the largest firms do display a 9.7 percent larger impact of Chat-GPT on the relative demand for junior versus senior software developers. Similarly, the second column shows that the largest CSAs also show a 6.7 percent larger impact compared to other places. Finally, to test the differential impacts by industry we categorize sectors into three groups based on the number of software developer postings per industry:

- **Light** - Management of Companies and Enterprises (NAICS 55), Accommodation and Food Services (NAICS 92), and Public Administration (NAICS 72);
- **Medium** - Construction (NAICS 23), Wholesale Trade (NAICS 42), Transportation and Warehousing (NAICS 48-49), Real Estate and Rental and Leasing (NAICS 53), Educational Services (NAICS 61), Health Care and Social Assistance (NAICS 62), and Other Services (except Public Administration) (NAICS 81).
- **Heavy** - Manufacturing (NAICS 31), Retail Trade (NAICS 44-45), Information (NAICS 51), Fiance and Insurance (NAICS 52), Professional, Scientific, and Technical Services (NAICS 54),Administrative, and Support and Waste Management and Remediation Services (NAICS 56).

Surprisingly, we find that the impacts of Chat-GPT by industry appear to be nonlinear with the greatest impacts being felt by industries with the fewest and the most software developer postings. This suggests that perhaps generative AI does indeed have a substitution effect for junior software developers who are perhaps displaced in smaller firms by others workers who can make use of Chat-GPT to assist with simple coding tasks as well as larger firms where Chat-GPT can play the role of a junior team member alongside more experienced software developers.

4.3. Changes in Required Skills

Isolating the top 100 skills for junior software developers, only three increased in demand (relative to seniors) since the public release of ChatGPT version 3.5: (1) problem solving, (2) interpersonal communication, and (3) mathematics. The increase in employer demand for these particular skillsets aligns with routine coding tasks often done by entry-level software developers being replaced by ChatGPT, such that firms require these junior workers to do more higher-order thinking (e.g., problem-solving, mathematics) as well as work more often in teams with senior software developers (e.g., interpersonal communication).

Thus, employers may be screening more on general cognitive or social aptitudes, as these qualities remain less likely to be automated by generative AI (Autor et al. 2024, Cui et al. 2024). This is consistent with the findings of Berman et al. (1998), who discuss how new technologies often reduce reliance on certain "production-focused" skills while increasing the premium on more conceptual and collaborative abilities. If employers believe coding tasks can be handled by AI tools, then a competitive junior candidate is one who can conceptualize system architectures, communicate effectively within teams, and troubleshoot problems as they arise.

Surprisingly, it does not appear that learning any particular artificial intelligence/machine learning (AI/ML) skills will make inexperienced software developers more employable. Between October 2022 and October 2024, the monthly share of software developer vacancies asking for AI/ML skills[6]

[6] See Section E of the Supplementary Materials for a full list of AI/ML skills as defined by Lightcast.

Table 4 Differences-in-Differences Estimates of Chat-GPT Impacts across Industry, Firm Size, and Geography. Formally testing for heterogeneity across subgroups within each area confirms that junior software developer roles in larger firms and bigger cities showed greater impacts while industries with moderate demand for software developers demonstrated the smallest impacts.

Dependent variable: Log Number of Postings	*Coefficient/SE on PostChatGPT x Junior*		
	Firm Size	**CSA Population**	**Industry Type**
Quintiles 1-4	-0.136**	-0.199**	
Standard error	(0.03)	(0.039)	
Number in sub-group	6198	29081	
Quintile 5	-0.229*	-0.264*	
Standard error	(0.05)	(0.069)	
Number in sub-group	28247	5364	
Difference	-0.093**	-0.065**	
Standard error	(0.043)	(0.038)	
SDE Light			-0.258*
Standard error			(0.076)
Number in sub-group			198
SDE Medium			-0.039**
Standard error			(0.048)
Number in sub-group			1495
SDE Heavy			-0.218**
Standard error			(0.042)
Number in sub-group			32752
Difference (Medium-Light)			0.219*
Standard error			(0.089)
Difference (Heavy-Medium)			-0.179*
Standard error			(0.08)
Difference (Heavy-Light)			0.04*
Standard error			(0.082)
Month Fixed Effects	Yes	Yes	Yes
Population Quintile Controls	Yes	Yes	Yes
Firm Size Quintile Controls	Yes	Yes	Yes
Industry Mix Controls	Yes	Yes	Yes
Observations	34445	34445	34445
Adjusted R^2	0.032	0.029	0.032

*Note: We run separate regressions for each group listed and report the coefficients on the interactions of the main PostChatGPT effect with a set of dummy variables that fully specify the sample according to the subgroup of interest. Each regression also includes the main subgroup effect as well as the full set of controls for the pre-ChatGPT characteristics listed in Table B.1. Each regression uses the first sample listed in Appendix Table A.1 that aggregated data from all job postings into Job Title × Month x CSA cells containing at least 5 total postings. Cells that are missing a CSA code or are located in Guam or Puerto Rico are excluded from the analysis. Baseline controls include pre Chat-GPT characteristics for each cell including the average firm size across job titles by quintile, the average population size across by quintile, and the industry mix by two-digit NAICS code. Statistical significance is indicated by * p¡0.10, ** p¡0.05, *** p¡0.01.*

rose from 10.5% to 13.1%, suggesting some increase in relative employer demand for AI/ML expertise amid a decline in the monthly count of vacancies. During this same period, the number of junior-level vacancies requiring at least one AI/ML skill dropped by 86.1% (from 1,915 to 1,029) while senior-level AI/ML vacancies only declined by 10.3% (from 4,338 to 3,934). This finding aligns with prior research showing an "escalator" effect (Autor et al. 2006), where technological progress raises the demand for high-skill workers, often leaving middle-skill workers behind.

Consequently, the market's preference for seasoned AI/ML developers exemplifies how cutting-edge tech amplifies the skill gap. AI/ML projects in particular often demand a higher level of knowledge and ability. The rapid evolution of the field means practitioners must constantly relearn and update best practices. This is easier for experts than novices. Aspiring AI/ML developers may need to consider alternative pathways to gain experience. Academic research, contributing to open-source AI/ML projects, or advanced degrees can all bridge the experience gap. This trend also emphasizes the importance of educational institutions maintaining updated curricula.

5. Shifting Occupations and Potential Transitions

To better understand where displaced software developers might transition, we used the Occupational Co-Occurrence Network described in Section 2.3, focusing on occupations with high skill similarity (Θ) to software developers. Table 5 lists the 30 junior-level occupations that we identified as most similar to junior-level software developers (SOC 15-1252) along with each occupation's share of total vacancies, its change in junior vacancies since November 2022, and the corresponding change in junior-to-senior hiring ratios.

Nine of the top ten most similar occupations are within a related subset of Computer Occupations (the SOC 15-1200 series). While nearly all computer-related roles had some reduction in junior hiring, few experienced as sever of a vacancy contraction or a junior-to-senior ratio shift as pronounced as software developers. For instance, Data Scientists (SOC 15-2051) showed only moderate declines in the total number of junior vacancies (-29%) compared to software developers (-49%) alongside smaller decrease in the junior-to-senior ratio (-11% versus -33% respectively). This suggests weaker displacement pressures for junior data scientists. Other high-similarity occupations (e.g., Computer Systems Analysts, SOC 15-1211 and Web and Digital Interface Designers (SOC 15-1255) appear to have maintained a balanced junior-to-senior hiring proportion, indicating they may be more accessible to early-career software developers seeking comparable work.

These findings highlight plausible transitions for junior software developers affected by the generative AI shift. The Occupational Co-Occurrence Network shows that many computer-related roles still require overlapping skill sets (e.g., coding, database management, or technical problem-solving). By measuring the overlap of skill sets between occupations, this network accounts for

Table 5 Possible transition occupations identified by our Co-Occurrence Network as similar to junior software developers in terms of skill requirements. Θ represents the probability that jobs share a set of effectively used skills. Total employment share, change in the number of junior-level vacancies and the change in the ratio of senior to junior vacancies indicate the feasibility of junior software developers transitioning to a related occupation.

Occupation	SOC	Θ	Market Share	Vacancy Change	Ratio Change
Software Developers	*15-1252*	*1*	*0.82%*	*−49%*	*33%*
Computer Occupations, All O...	15-1299	0.409	0.78%	−39%	−26%
Data Scientists	15-2051	0.312	0.52%	−29%	−11%
Database Administrators	15-1242	0.286	0.17%	−37%	−11%
Computer User Support Speci...	15-1232	0.283	0.73%	−38%	−13%
Computer Systems Analysts	15-1211	0.261	0.28%	−22%	3%
Computer Network Architects	15-1241	0.26	0.15%	−34%	−27%
Network and Computer System...	15-1244	0.246	0.14%	−37%	2%
Web Developers	15-1254	0.236	0.11%	−45%	−3%
Management Analysts	13-1111	0.232	0.39%	−36%	1%
Software Quality Assurance ...	15-1253	0.226	0.08%	−53%	−10%
Database Architects	15-1243	0.215	0.16%	−35%	−20%
Industrial Engineers	17-2112	0.21	0.34%	−30%	10%
Computer Programmers	15-1251	0.185	0.06%	−36%	24%
Marketing Managers	11-2021	0.18	0.36%	−32%	−28%
Web and Digital Interface D...	15-1255	0.176	0.1%	−46%	−2%
Operations Research Analyst...	15-2031	0.171	0.12%	−40%	−32%
Information Security Analys...	15-1212	0.171	0.09%	−35%	−6%
Mechanical Engineers	17-2141	0.16	0.23%	−29%	−17%
Market Research Analysts an...	13-1161	0.154	0.59%	−30%	−3%
Project Management Speciali...	13-1082	0.152	0.75%	−10%	−19%
Financial and Investment An...	13-2051	0.145	0.57%	−24%	−9%
Electrical Engineers	17-2071	0.143	0.17%	−34%	−33%
Architectural and Engineeri...	11-9041	0.141	0.07%	−30%	$3-6\%$
Financial Risk Specialists	13-2054	0.128	0.12%	−15%	−4%
Business Operations Special...	13-1199	0.128	0.07%	−26%	34%
Electronics Engineers, Exce...	17-2072	0.124	0.07%	−15%	9%
Logisticians	13-1081	0.124	0.4%	−32%	−15%
Buyers and Purchasing Agent...	13-1028	0.121	0.46%	−19%	−3%
First-Line Supervisors of P...	51-1011	0.12	0.47%	−25%	−9%
Detectives and Criminal Inv...	33-3021	0.116	0.09%	2%	−7%
Managers, All Other	11-9199	0.116	0.47%	−17%	−17%
Commercial and Industrial D...	27-1021	0.115	0.06%	−31%	−12%
Production, Planning, and E...	43-5061	0.114	0.52%	−20%	−23%
Computer Hardware Engineers	17-2061	0.113	0.02%	−44%	−47%

Note: The change in job postings and the junior-to-senior ratio is calculated during the period from September 2022 (pre-ChatGPT) through January 2025

the transferability of skills across different roles (Arntz et al. 2017). This means displaced junior software developers could transition into these adjacent roles without extensive retraining, rather than exiting the tech sector altogether.

6. Conclusion and Policy Implications

The public release of ChatGPT version 3.5 provides a real-time case study of how generative AI can reshape the labor market. Our analysis shows a distinct shift from junior to senior software developer roles. This finding is consistent with both recent industry observations as well as the vast literature on skill-biased technological change (Katz and Murphy 1992, Autor et al. 2003, Acemoglu and Restrepo 2018a). We document how the rapid adoption of generative AI is correlated with measurable shifts in employer demand. Our findings reinforce prior research suggesting that technological disruptions do not uniformly eliminate employment but rather reshape the labor market by altering what is most valuable (Autor et al. 2024).

It remains to be seen whether this shift is a one-time tranisition to a ”new normal” or indicative of a long-term restructuring of the software development occupation. Future research could probe whether displaced entry-level software developers migrate to adjacent tech roles such as testing, tech support, or design. A longer-term follow-up of this cohort of workers could confirm theories of long-lasting “scarring” effects from a weak job market (Kahn 2010). Incorporating more detailed worker-level data (e.g., resumes or LinkedIn profiles) would clarify how novice developers pivot to other roles and whether there are pathways that mitigate the scarring.

That said, this research must be interpreted in light of several limitations. First, our reliance on online job vacancy data may not fully capture positions that are advertised through alternative channels. As a result, the true scale of hiring shifts could be either under-or overestimated, although the Lightcast data tend to have more accurate sampling for professional and technical occupations compared to service sector or public sector positions. Second, we measure postings rather than actual hires, such that the employment rate of junior candidates might deviate from the trends implied by our analysis of job vacancies. Third, our data do not cover the qualitative aspects of the hiring process that could also shed light on how firms might adjust junior software developer roles once a candidate is onboarded. Finally, our study does not fully capture long-term adjustments. Over an extended horizon, firms might relax experience requirements once they understand how generative-AI can augment junior software developer roles. Conversely, new types of entry-level roles may emerge and programs that train entry-level workers may adapt by responding to the increased demand for interpersonal, communication, and problem-solving skills that go beyond acquiring rote coding knowledge..

Overall, by studying the ways in which AI can alter entry-level opportunities and skill requirements, this paper offers a window into how technology reshapes the workforce. Firms, educational institutions, and public agencies should consider how these findings can guide long-term strategic decisions to prepare the next generation of workers. Simply obtaining an advanced degree, such as a Master's or Ph.D., may not insulate workers from these broader industry contractions, especially since they only accounted for 2.46% of the vacancies as of 2024.

Instead, policies and educational strategies should address both emerging skill requirements and structural changes in hiring. In the short-run, recent computer science graduates have fewer on-ramps to establish themselves within the software developer occupation. Research confirms that this scarcity of entry-level roles can lead to "scarring" effects for young workers Kahn (2010) and Oreopoulos et al. (2012) show that graduating into a weak labor market depresses earnings and stunts professional growth. These effects can persist for decades with talented graduates ending up underemployed or working outside of their field of study, making it harder for them to compete when labor market conditions improve.

In the long-run, creating apprenticeships, fellowships, co-ops, and subsidized internships is one feasible path for preserving the future talent pipeline. These types of "Learn and Earn" programs also provide a natural feedback loop between firms and educational institutions to ensure students are keeping pace with the latest skill requirements, especially in occupations where technology advances rapidly, such as computer science. Lerman (2019) emphasizes that apprenticeships equip entry-level workers with practical skills and resilience, helping new hires learn competencies that might be overlooked in conventional academic curricula and reducing the risk of deficits that might have a persistent effect on employability. Applying network methods to real-time data also can help education and training institutions focus on skill complementarity when designing curriculum and helping displaced workers mitigate long-term career disruptions caused by technological change (Kahn 2010).

References

Acemoglu D (2025) The simple macroeconomics of ai. *Economic Policy* 40(121):13–58.

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

Acemoglu D, Restrepo P (2018a) Artificial intelligence, automation, and work. *The economics of artificial intelligence: An agenda*, 197–236 (University of Chicago Press).

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

Alabdulkareem A, Frank MR, Sun L, AlShebli B, Hidalgo C, Rahwan I (2018) Unpacking the polarization of workplace skills. *Science advances* 4(7):eaao6030.

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

Anderson KA (2017) Skill networks and measures of complex human capital. *Proceedings of the National Academy of Sciences* 114(48):12720–12724.

Arntz M, Gregory T, Zierahn U (2017) Revisiting the risk of automation. *Economics Letters* 159:157–160.

Autor D, Chin C, Salomons A, Seegmiller B (2024) New frontiers: The origins and content of new work, 1940–2018*. *The Quarterly Journal of Economics* 139(3):1399–1465, URL `http://dx.doi.org/10.1093/qje/qjae008`.

Autor DH (2015) Why are there still so many jobs? the history and future of workplace automation. *Journal of Economic Perspectives* 29(3):3–30, URL `http://dx.doi.org/10.1257/jep.29.3.3`.

Autor DH, Katz LF, Kearney MS (2006) The polarization of the us labor market. *American economic review* 96(2):189–194.

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

Berman E, Bound J, Machin S (1998) Implications of skill-biased technological change: international evidence. *The quarterly journal of economics* 113(4):1245–1279.

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? *The Quarterly journal of economics* 119(1):249–275.

Bick A, Blandin A, Deming DJ (2024) The rapid adoption of generative ai. Working Paper 32966, National Bureau of Economic Research, URL `http://dx.doi.org/10.3386/w32966`, accessed: 2025-05-10.

Brynjolfsson E, Li D, Raymond L (2025) Generative ai at work. *The Quarterly Journal of Economics* 140(2):889–942, URL `http://dx.doi.org/10.1093/qje/qjae044`.

Brynjolfsson E, Rock D, Syverson C (2019) Artificial intelligence and the modern productivity paradox. *The economics of artificial intelligence: An agenda* 23(2019):23–57.

Cammeraat E, Squicciarini M (2021) Burning glass technologies' data use in policy-relevant analysis: An occupation-level assessment. Technical Report 2021/05, OECD Publishing, URL https://www.oecd.org/en/publications/burning-glass-technologies-data-use-in-policy-relevant-analysis_cd75c3e7-en.html, accessed: 2025-05-10.

Carnevale AP, Jayasundera T, Repnikov D (2014) Understanding online job ads data. Technical report, Georgetown University Center on Education and the Workforce.

Cui ZK, Demirer M, Jaffe S, Musolff L, Peng S, Salz T (2024) The effects of generative ai on high skilled work: Evidence from three field experiments with software developers. *Available at SSRN 4945566* .

Demirci O, Hannane J, Zhu X (2025) Who is ai replacing? the impact of generative ai on online freelancing platforms. *Management Science* (Forthcoming).

Eloundou T, Manning S, Mishkin P, Rock D (2024) Gpts are gpts: Labor market impact potential of llms. *Science* 384(6702):1306–1308.

Frank MR, Moro E, South T, Rutherford A, Pentland A, Taska B, Rahwan I (2024) Network constraints on worker mobility. *Nature Cities* 1(1):94–104.

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

Graetz G, Michaels G (2018) Robots at work. *Review of economics and statistics* 100(5):753–768.

Hershbein B, Kahn LB (2018) Do recessions accelerate routine-biased technological change? evidence from vacancy postings. *American Economic Review* 108(7):1737–1772.

Hidalgo CA, Klinger B, Barabási AL, Hausmann R (2007) The product space conditions the development of nations. *Science* 317(5837):482–487.

Hu K (2023) Chatgpt sets record for fastest-growing user base – analyst note. *Reuters* URL https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/, accessed: 2025-05-10.

Hui X, Reshef O, Zhou L (2024) The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market. *Organization Science* 35(6):1977–1989.

Kahn LB (2010) The long-term labor market consequences of graduating from college in a bad economy. *Labour economics* 17(2):303–316.

Katz LF, Murphy KM (1992) Changes in relative wages, 1963–1987: supply and demand factors. *The quarterly journal of economics* 107(1):35–78.

Khojah R, Mohamad M, Leitner P, de Oliveira Neto FG (2024) Beyond code generation: An observational study of chatgpt usage in software engineering practice. *Proceedings of the ACM on Software Engineering* 1(FSE):1819–1840.

Kinder M, de Souza Briggs X, Liu S, Muro M (2024) Generative ai, the american worker, and the future of work. *Brookings Institution* URL `https://www.brookings.edu/articles/generative-ai-the-american-worker-and-the-future-of-work/`, accessed: 2025-05-10.

Lerman R (2019) Do firms benefit from apprenticeship investments? *IZA World of Labor* URL `http://dx.doi.org/10.15185/izawol.55`.

Modestino AS, Shoag D, Ballance J (2016) Downskilling: changes in employer skill requirements over the business cycle. *Labour Economics* 41:333–347.

Modestino AS, Shoag D, Ballance J (2020) Upskilling: Do employers demand greater skill when workers are plentiful? *Review of Economics and Statistics* 102(4):793–805.

National Center for Education Statistics (2022) Integrated Postsecondary Education Data System (IPEDS), Completions Survey: 2022. URL `https://nces.ed.gov/ipeds/`, u.S. Department of Education. Data retrieved from the IPEDS Data Center.

Oreopoulos P, Von Wachter T, Heisz A (2012) The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics* 4(1):1–29.

Otis NG, Clarke R, Koning R (2023) The uneven impact of generative ai on entrepreneurial performance. Working Paper 24-042, Harvard Business School, Cambridge, MA.

Ouyang L, Wu J, Jiang X, Almeida D, Wainwright C, Mishkin P, Zhang C, Agarwal S, Slama K, Ray A, et al. (2022) Training language models to follow instructions with human feedback. *Advances in neural information processing systems* 35:27730–27744.

Peng S, Kalliamvakou E, Cihon P, Demirer M (2023) The impact of ai on developer productivity: Evidence from github copilot. *arXiv preprint arXiv:2302.06590* URL `https://arxiv.org/abs/2302.06590`.

Stephany F, Teutloff O (2024) What is the price of a skill? the value of complementarity. *Research Policy* 53(1):104898.

Toner-Rodgers A (2024) Artificial intelligence, scientific discovery, and product innovation. *arXiv preprint arXiv:2412.17866* URL `https://arxiv.org/pdf/2412.17866`.

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Łukasz Kaiser, Polosukhin I (2017) Attention is all you need. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)*, 5998–6008.

Xiao T, Treude C, Hata H, Matsumoto K (2024) Devgpt: Studying developer-chatgpt conversations. *Proceedings of the 21st International Conference on Mining Software Repositories*, 227–230 (New York, NY, USA: Association for Computing Machinery), ISBN 9798400705878, URL `http://dx.doi.org/10.1145/3643991.3648400`.

Zolas N, Kroff Z, Brynjolfsson E, McElheran K, Beede DN, Buffington C, Goldschlag N, Foster L, Dinlersoz E (2020) Advanced technologies adoption and use by us firms: Evidence from the annual business survey. Working Paper 28290, National Bureau of Economic Research, URL `http://dx.doi.org/10.3386/w28290`.

Supplementary Materials

A. Sample Characteristics

Employer skill requirements are constructed using online job posting data provided by Lightcast. This sample uses data from all 5.7 million job postings for software developers aggregated into CSA x software job title cells containing at least 5 total postings. The second sample uses data from the subset of job postings that identify employer name and contain at least 5 total postings in each employer × software job title × CSA cell within a given month. The third sample includes only postings for which a given employer × software job title × CSA cell is observed at least once before November 2022 and once on or after November 2022. In all three of the above samples, we exclude postings that are missing a CSA code or are located in Guam or Puerto Rico.

Table A.1 Summary Statistics for Employer Skill Requirements.

	Oct. 21	Oct. 22	Oct. 23	Difference (*Oct. 22 - Oct. 21*)	Difference (*Oct. 23 - Oct. 22*)	Total
Sample 1: Cross-sectional sample of all job postings aggregated by CSA × software job title cells						
Conceptual number of observations	46000	46000	46000			1104000
Actual number of observations	3177	2795	1485	-382	-1310	95953
Total postings per CSA×title	54969	45473	18923	-9496	-26550	1555766
Mean	17.3	16.27	12.74	-1.03	-3.53	
Std. dev	31.57	29.94	16.61			
Percent of job postings requesting:						
Bachelor's degree or higher	53.1	55.1	54.3	1.9	-0.8	
No education listed	41.3	41.8	41.9	0.5	0.1	
Four or more years of experience	47.8	49.6	50.9	1.8	1.3	
Three or fewer years of experience	20.9	20.2	14.5	-0.8	-5.6	
No experience listed	31.2	30.2	34.6	-1.0	4.3	
Sample 2: Cross-sectional sample of postings with non-missing employer names						
Total number of job postings	49821	40052	17561	-9769	-22491	1431527
Percent of job postings requiring:						
Bachelor's degree or higher	52.5	56.4	54.7	3.9	-1.6	
Four or more years of experience	46.6	48.8	50.9	2.2	2.1	
Three or fewer years of experience	21.2	20.8	14.7	-0.4	-6.1	
Sample 3: Panel sample of repeated employer × title × CSA observations						
Total number of job postings	28274	29221	10483	947	-18738	868622
Percent of job postings requiring:						
Bachelor's degree or higher	53.6	58.0	57.3	4.5	-0.8	
Four or more years of experience	45.8	48.5	50.9	2.7	2.4	
Three or fewer years of experience	24.0	21.7	14.7	-2.3	-7.0	

B. Independent Variables

Below we provide summary statistics for each of our independent variables used in the regression presented in the text. The population quintile is generated using data from the Census for 2022. The firm size quintile is generated as a firm's total postings for any job title from November 2021 to October 2022. The industry mix is by NAICS code for October 2022.

Table B.1 Summary Statistics for Independent Variables.

	Quintile 1	*Quintile 2*	*Quintile 3*	*Quintile 4*	*Quintile 5*
Share CSA population as of Oct. 2022	19.4	20.3	24.3	16.4	19.6
Share firm size as of Oct. 2022	1.3	2.7	5.0	10.9	80.2
Industry mix as of Oct. 2022					
NAICS 11: Agriculture, Forestry, Fishing, and Hunting				0.0	
NAICS 21: Mining				0.0	
NAICS 22: Utilities				0.3	
NAICS 23: Construction				0.2	
NAICS 31: Manufacturing				17.4	
NAICS 42: Wholesale Trade				1.0	
NAICS 44: Retail Trade				8.3	
NAICS 48: Transportation and Warehousing				0.3	
NAICS 51: Information				4.3	
NAICS 52: Finance and Insurance				9.5	
NAICS 53: Real Estate and Rental and Leasing				0.1	
NAICS 54: Professional, Scientific, and Technical Services				18.4	
NAICS 55: Management of Companies and Enterprises				0.1	
NAICS 56: Admin, Support, Waste Management and Remediation Services-				38.0	
NAICS 61: Education Services				0.2	
NAICS 62: Health Care and Social Assistance				0.4	
NAICS 71: Arts, Entertainment, and Recreation				0.0	
NAICS 72: Accommodation and Food Services				0.0	
NAICS 81: Other Services (except Public Administration)				0.0	
NAICS 92: Public Administration				0.2	
NAICS 99: Undefined				1.1	

C. Robustness Checks

Below we provide several robustness checks to the main specification presented in the text. These include (1) using the level of postings, (2) restricting the sample of postings to either those with employer names and/or those that we can observe repeatedly over time, and (3) separate estimates of the impact of Chat-GPT on postings with no listed experience, 0-3 years of experience, and 4 or more years of experience.

Table C.1 Difference-in-Differences Analysis using Posting Levels.

	Dependent variable: Number of Postings			
	(1)	(2)	(3)	(4)
Share Junior	6.747***	5.901***	1.414	1.414
	(1.474)	(1.434)	(1.148)	(1.148)
PostChatGPT * Share Junior	-7.090***	-6.823***	-3.751***	-3.751***
	(1.610)	(1.606)	(1.393)	(1.393)
Month Fixed Effects	Yes	Yes	Yes	Yes
Job Title Fixed Effects		Yes	Yes	Yes
Firm Fixed Effects			Yes	Yes
CSA Fixed Effects				Yes
Observations	35179	35179	35179	35179
R^2	0.004	0.024	0.180	0.180
Adjusted R^2	0.003	0.018	0.129	0.129

*Note: Each coefficient listed is from a separate regression as specified by equation 1 in the text. Each regression uses the first sample listed in Appendix Table A.1 that aggregated data from all job postings into Job Title × Month x CSA cells containing at least 5 total postings. Cells that are missing a CSA code or are located in Guam or Puerto Rico are excluded from the analysis. Baseline controls include pre Chat-GPT characteristics for each cell including the average firm size across job titles by quintile, the average population size across by quintile, and the industry mix by two-digit NAICS code. Statistical significance is indicated by * p¡0.10, ** p¡0.05, *** p¡0.01.*

Table C.2 Difference-in-Differences Analysis using Restricted Samples.

	Share of Postings Requiring 4 or More Years of Experience			
Panel A: Cross-sectional sample of job postings with non-missing employer names (N=589,292)				
	(1)	(2)	(3)	(4)
Share Junior	6.747***	5.901***	1.414	1.414
	(1.474)	(1.434)	(1.148)	(1.148)
PostChatGPT * Share Junior	-7.090***	-6.823***	-3.751***	-3.751***
	(1.610)	(1.606)	(1.393)	(1.393)
Observations	35179	35179	35179	35179
Adjusted R^2	0.003	0.018	0.129	0.129
Panel B: Panel sample of repeated employer × job title × state observations (N=141,164)				
Share Junior	7.287***	3.611***	2.018***	2.311***
	(1.472)	(0.800)	(0.672)	(0.679)
PostChatGPT * Share Junior	-8.735***	-4.941***	-2.878***	-3.159***
	(1.535)	(0.945)	(0.824)	(0.858)
Observations	14287	14287	14287	14287
Adjusted R^2	0.010	0.071	0.088	0.100
Month Fixed Effects	Yes	Yes	Yes	Yes
Job Title Fixed Effects		Yes	Yes	Yes
Firm Fixed Effects			Yes	Yes
CSA Fixed Effects				Yes

*Note: Each coefficient listed is from a separate regression as specified by equation 1 in the text. Each regression uses the first sample listed in Appendix Table A.1 that aggregated data from all job postings into Job Title × Month x CSA cells containing at least 5 total postings. Cells that are missing a CSA code or are located in Guam or Puerto Rico are excluded from the analysis. Baseline controls include pre Chat-GPT characteristics for each cell including the average firm size across job titles by quintile, the average population size across by quintile, and the industry mix by two-digit NAICS code. Statistical significance is indicated by * p¡0.10, ** p¡0.05, *** p¡0.01.*

Table C.3 Difference in Difference Estimates by Skill Level.

	No Experience	> 0 but ≤ 3 Years	≥ 4 Years
	(1)	(2)	(3)
PostChatGPT	-0.015***	-0.026***	0.041***
	(0.003)	(0.002)	(0.003)
Population Quintile Controls	Yes	Yes	Yes
Firm Size Quintile Controls	Yes	Yes	Yes
Industry Mix Controls	Yes	Yes	Yes
Observations	47840	47840	47840
Adjusted R^2	0.040	0.043	0.024

*Note: Each coefficient listed is from a separate regression as specified by equation 1 in the text. Each regression uses the first sample listed in Appendix Table A.1 that aggregated data from all job postings into Job Title × Month x CSA cells containing at least 5 total postings. Cells that are missing a CSA code or are located in Guam or Puerto Rico are excluded from the analysis. Baseline controls include pre Chat-GPT characteristics for each cell including the average firm size across job titles by quintile, the average population size across by quintile, and the industry mix by two-digit NAICS code. Statistical significance is indicated by * p¡0.10, ** p¡0.05, *** p¡0.01.*

D. Hetergeneous Impacts by Industry, Firm Size, and Geography

In this section we report estimates exploring heterogeneity by industry, firm size, and geography as specified by equation 3 in the text.

Table D.1 Differences-in-Differences Estimates of Chat-GPT Impacts by Industry.

	Dependent variable: ratio of junior to senior vacancies	
	Coefficient	Std Error
Coefficient on PostChatGPT x Junior x		
Construction	-0.705**	(0.319)
Manufacturing	-0.779**	(0.319)
Wholesale Trade	-0.805**	(0.319)
Retail Trade	-1.584***	(0.319)
Transportation and Warehousing	-0.658**	(0.319)
Information	-1.534***	(0.319)
Finance and Insurance	-0.934***	(0.319)
Real Estate, Rental, and Leasing	-1.428***	(0.319)
Professional, Scientific, and Technical Services	-1.766***	(0.319)
Management of Companies and Enterprises	-0.187	(0.319)
Admin. Support, Waste Mgmt., and Remediation Services	-0.904***	(0.319)
Educational Services	-1.015***	(0.319)
Health Care and Social Assistance	-0.857***	(0.319)
Accommodation and Food Services	-2.337***	(0.319)
Other Services (except Public Administration)	-1.115***	(0.319)
Public Administration	-0.769**	(0.319)
Controlling for:		
NAICS FE	Yes	
Month FE	Yes	
Observations	1170894	
Groups	480	
R^2	0.614	
Adjusted R^2	0.597	

*Note: Each coefficient listed is from a separate regression as specified by equation 3 in the text. Each regression uses the first sample listed in Appendix Table A.1 that aggregated data from all job postings into Job Title × Month x CSA cells containing at least 5 total postings. Cells that are missing a CSA code or are located in Guam or Puerto Rico are excluded from the analysis. Baseline controls include pre Chat-GPT characteristics for each cell including the average firm size across job titles by quintile, the average population size across by quintile, and the industry mix by two-digit NAICS code. Statistical significance is indicated by * p¡0.10, ** p¡0.05, *** p¡0.01.*

Table D.2 Differences-in-Differences Estimates of Chat-GPT Impacts by Firm Size.

Dependent variable: Ratio of junior to senior vacancies		
	Coefficient	Std Error
Coefficient on PostChatGPT x Junior x		
Firm Size Quintile Q1	-0.005	(0.121)
Firm Size Quintile Q2	-0.015	(0.121)
Firm Size Quintile Q3	-0.297**	(0.121)
Firm Size Quintile Q4	-0.007	(0.121)
Firm Size Quintile Q5	-0.480***	(0.121)
Controlling for:		
Quintile FE	Yes	
Month FE	Yes	
Observations	853538	
Groups	125	
Adjusted R^2	0.853	

*Note: Each coefficient listed is from a separate regression as specified by equation 3 in the text. Each regression uses the first sample listed in Appendix Table A.1 that aggregated data from all job postings into Job Title × Month x CSA cells containing at least 5 total postings. Cells that are missing a CSA code or are located in Guam or Puerto Rico are excluded from the analysis. Baseline controls include pre Chat-GPT characteristics for each cell including the average firm size across job titles by quintile, the average population size across by quintile, and the industry mix by two-digit NAICS code. Statistical significance is indicated by * p¡0.10, ** p¡0.05, *** p¡0.01.*

Table D.3 Differences-in-Differences Estimates of Chat-GPT Impacts by Geography.

	Dependent variable: ratio of senior to junior vacancies	
Coefficient on PostChatGPT x Junior x		
	Coefficient	Std. Error
Population Quintile 1	-0.014	(0.122)
Population Quintile 2	-0.021	(0.122)
Population Quintile 3	-0.106	(0.122)
Population Quintile 4	-0.134	(0.122)
Population Quintile 5	-0.251**	(0.122)
Controlling for:		
Quintile FE	Yes	
Month FE	Yes	
Observations	956596	
Groups	120	
Adjusted R^2	0.822	

*Note: Each coefficient listed is from a separate regression as specified by equation 3 in the text. Each regression uses the first sample listed in Appendix Table A.1 that aggregated data from all job postings into Job Title × Month x CSA cells containing at least 5 total postings. Cells that are missing a CSA code or are located in Guam or Puerto Rico are excluded from the analysis. Baseline controls include pre Chat-GPT characteristics for each cell including the average firm size across job titles by quintile, the average population size across by quintile, and the industry mix by two-digit NAICS code. Statistical significance is indicated by * p¡0.10, ** p¡0.05, *** p¡0.01.*

E. Lightcast's AI/ML Skill Clusters

There are a total of 110 skills in Lightcast's Artificial Intelligence and Machine Learning skill cluster that we used in our analysis of skill requirements. These include the following: 3D Reconstruction, Activity Recognition, AdaBoost (Adaptive Boosting), Adversarial Machine Learning, AIOps (Artificial Intelligence For IT Operations), Amazon Textract, Apache Mahout, Apache MXNet, Apache SINGA, Apache Spark, Applications Of Artificial Intelligence, Artificial Intelligence, Artificial Intelligence Development, Artificial Intelligence Markup Language (AIML), Artificial Intelligence Systems, Artificial Neural Networks, Association Rule Learning, Autoencoders, Automated Machine Learning, AWS SageMaker, Azure Cognitive Services, Azure Machine Learning, Baidu, Boosting, Caffe, Caffe2, Chatbot, Classification And Regression Tree (CART), Cognitive Automation, Cognitive Computing, Cognitive Robotics, Collaborative Filtering, Computational Intelligence, Confusion Matrix, Convolutional Neural Networks, Cortana, Cudnn, Dask (Software), Deep Learning, Deep Learning Methods, Deeplearning4j, Dialog Systems, Dlib (C++ Library), Ensemble Methods, Evolutionary Acquisition Of Neural Topologies, Expert Systems, Fast.ai, Feature Engineering, Feature Extraction, Feature Learning, Feature Selection, Game Ai, General-Purpose Computing On Graphics Processing Units, Genetic Algorithm, Gesture Recognition, Google AutoML, Gradient Boosting, H2O.ai, Hidden Markov Model, Inference Engine, Intelligent Agent, Intelligent Control, Intelligent Systems, Intelligent Virtual Assistant, Interactive Kiosk, IPSoft Amelia, Kaldi, Keras (Neural Network Library), Kernel Methods, Knowledge-Based Configuration, Knowledge-Based Systems, Kubeflow, Long Short-Term Memory (LSTM), Machine Learning, Machine Learning Algorithms, Machine Learning Methods, Meta Learning, Microsoft Cognitive Toolkit (CNTK), Microsoft LUIS, MLflow, MLOps (Machine Learning Operations), mlpack (C++ Library), Multi-Agent Systems, Nvidia Jetson, Objective Function, OmniPage, Open Neural Network Exchange (ONNX), OpenAI Gym, OpenCV, OpenVINO, PaddlePaddle, Pydata, PyTorch (Machine Learning Library), Random Forest Algorithm, Reasoning Systems, Recommender Systems, Recurrent Neural Network (RNN), Reinforcement Learning, Scikit-learn (Machine Learning Library), Semi-Supervised Learning, Seq2Seq, Sorting Algorithm, Speech Recognition Software, Speech Synthesis, Supervised Learning, Support Vector Machine, TensorFlow, Test Datasets, Text-To-Speech, Theano (Software), Torch (Machine Learning), Training Datasets, Transfer Learning, Unsupervised Learning, Voice Assistant Technology, Voice Interaction, Voice User Interface, Vowpal Wabbit, Watson Conversation, Watson Studio, Xgboost