The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market *

Xiang Hui†
Washington University in St. Louis

Oren Reshef ‡
Washington University in St. Louis

Luofeng Zhou §
New York University

October 12, 2024

Abstract

Generative artificial intelligence (AI) holds the potential to either complement workers by enhancing their productivity or substitute them. We examine the short-term effects of the recently released generative AI models (ChatGPT, DALL-E 2, and Midjourney) on the employment outcomes of freelancers on a large online platform. We find that freelancers in highly affected occupations suffer from the introduction of generative AI, experiencing reductions in both employment and earnings. We find similar effects studying the release of other image-based, generative AI models. Exploring the heterogeneity by freelancers' employment history, we do not find evidence that high-quality service, measured by their past performance and employment, moderates the adverse effects on employment. In fact, we find suggestive evidence that top freelancers are disproportionately affected by AI. These results suggest that generative AI may transform the role of human capital in the organization and reduce overall demand for workers.

Keywords: Artificial Intelligence, online labor markets, large language model, generative AI

*We thank Avi Collis, Ashvin Gandhi, Deepak Hegde, John Horton, Danielle Li, Meng Liu, Lamar Pierce, Lindsey Raymond, Dan Rock, Rob Seamans, Ananya Sen, Catherine Tucker, and participants at the NBER Summer Institute, Wharton Business & Generative AI Workshop, and NYU Economics of Strategy Workshop for very helpful comments and discussions.

†hui@wustl.edu

‡oren@wustl.edu

§lz2198@stern.nyu.edu

1 Introduction

Recent developments in generative artificial intelligence (AI) models, particularly ChatGPT and Midjourney, have dramatically improved performance compared with previous versions, leading to their rapid adoption. Users are able to employ these powerful AI tools to complete a variety of tasks without requiring specialized knowledge. Expansion of these new technologies may fundamentally transform the role of human capital in the organization (Zammuto et al., 2007; Mokyr et al., 2015), with the potential for far-ranging economic and societal effects. Therefore, a central theoretical and empirical question organizations face is whether, over time, AI technologies will fully replace human input or instead augment workers' labor and expertise (Lebovitz et al., 2022; Kellogg et al., 2020; Furman and Seamans, 2019).

The effect of generative AI on employment is theoretically ambiguous: On the one hand, AI can complement human workers by increasing their productivity (Choudhury et al., 2020; Noy and Zhang, 2023); on the other hand, it may substitute workers, leading to mass layoffs and unemployment (Acemoglu and Autor, 2011; Brynjolfsson et al., 2018a; Agrawal et al., 2019a). In addition, the introduction of AI into organizations may also differentially affect workers of different abilities and expertise (Lebovitz et al., 2022; Allen and Choudhury, 2022), ultimately altering the composition of workers and either exacerbating or mitigating wage inequality both within and across occupations. Although previous literature has studied the effects of related technologies on labor market outcomes (Autor et al., 2003; Acemoglu and Restrepo, 2020), the effects of generative AI have remained unexplored. Specifically, generative AI may be fundamentally different from previous technologies, as it incorporates the ability to learn and act in ways that mirror human behavior (Bailey et al., 2022; Furman and Teodoridis, 2020; Faraj et al., 2018; Raisch and Krakowski, 2021).

In this paper, we provide novel empirical evidence on the short-term effects of generative AI on labor market outcomes. We focus on the introduction of ChatGPT in November 2022, and of image-based AI models DALL-E 2 and Midjourney in April 2022. The main data are obtained from a large online labor market, Upwork, which matches freelancers with short-term projects. Due to the flexibility of this spot market compared with traditional, formal employment, it is a great setting to explore the short-term effects of generative AI models on employment. We begin by obtaining publicly available data on the freelancers' full employment histories on the platform. We then use a difference-in-differences research design to study the differential change in employment outcomes of freelancers working in occupations most affected by generative AI. For ChatGPT, given results from previous research (Eloundou et al., 2023), anecdotal evidence, and the fact that ChatGPT is a large language model (LLM) specifically trained to predict and generate text, we focus on writing-related services as the main affected occupations. For the image-based AI models, we focus on freelancers offering design-, image-, and art-based services.

We find that ChatGPT had a substantial adverse effect on workers' employment outcomes. Following the release of ChatGPT, freelancers in more affected occupations experienced a decrease of 2% in the number of monthly jobs and a decrease of 5.2% in monthly earnings on the platform compared with freelancers in less-affected occupations. These effects represent an economically meaningful and statistically significant (at the 1% level) reduction in employment on the platform. Moreover, they are generally at least as large as the effects of other technological and non-technological labor market shocks. We observe these adverse effects of the introduction of ChatGPT on both the extensive and intensive margins: Freelancers are 1.2% less likely to receive any employment in a given month and take 4.7% fewer jobs, conditional on employment.

We next turn to evaluate the effect of image-based generative AI. This additional analysis studies the effect of another type of generative AI, which occurred at a different time and affected a different set of highly-trained workers. It therefore allows us to further evaluate the internal validity of the research design and assess the generalizability of our main findings. Reassuringly, we observe qualitatively similar effects to those of ChatGPT in terms of both the number of monthly jobs (reduction of 2.1%) and total monthly earnings (decrease of 5.2%).

Next, we study whether freelancer quality mediates the effect of generative AI on employment outcomes. Since we do not directly observe quality in our data, we consider several measures to determine freelancer quality, including their past employment and earnings, the skill level necessitated by past jobs, past performance, and hourly rate. Across the board, we do not find evidence that being high-quality moderates the adverse effects of generative AI on a freelancer's employment outcomes. In fact, evidence is consistent with high-quality workers being disproportionately affected by the release of generative AI tools.

We interpret these findings as suggesting that generative AI models may act as a substitute for knowledge workers of all quality types, at least in the short term, effectively reducing their employment and earnings.

We discuss several alternative explanations consistent with these heterogeneous effects, such as generative AI potentially reducing the productivity gap between low-quality and high-quality workers. We conclude by reviewing the broader implications of our results on the effect of AI on modern organizations and the labor market more generally.

Our results directly contribute to the growing literature exploring the effects of AI on organizations and the labor force. A recurring theme within these studies is the theoretically ambiguous nature of the effect of AI on labor: While AI can serve as a substitute for labor (Balasubramanian et al., 2022), it also has the potential to complement workers by increasing labor productivity (Murray et al., 2021), making the net effect of AI on firms' human capital and competitive strategy unclear (Choudhury et al., 2020; Tschang and Almirall, 2021; Krakowski et al., 2023). Toward understanding the conditions under which AI might substitute or complement labor, Autor et al. (2003), Acemoglu and Restrepo (2018, 2020), and Agrawal et al. (2019a) provide frameworks for examining the implications of AI for the labor market. The effect of AI on labor dynamics, including the substitution and complementarity of labor and capital, may also have profound implications for organizational design, aligning with central themes in organizational theory (Mintzberg, 1980; Grant, 1996; Argyres and Zenger, 2012; Pfeffer and Salancik, 2015). Our results contribute to this strand of literature by providing novel empirical evidence on a theoretically ambiguous question—we find that the latest artificial intelligence technology, namely generative AI, constitutes a net substitute for workers in the short run. Additionally, previous work mainly focuses on technologies substituting low-skilled workers performing routine tasks (e.g., Autor et al., 2003; Acemoglu and Restrepo, 2018; Balasubramanian et al., 2022). In contrast, we find that high-skilled workers are not shielded from the negative effects of generative AI, suggesting that general AI may have a pervasive effect on employees and organizations.

One of the most recent and influential developments in AI is the release of new generative AI models and LLMs. Several papers have tried to predict their effects on the labor market. Particularly, Eloundou et al. (2023) and Felten et al. (2023) attempt to quantify which industries would be most affected by generative AI in the future by using a measure of compatibility between the required tasks in an industry and AI capabilities; this methodology was first seen in Felten et al. (2018), Brynjolfsson et al. (2018a), and Webb (2019). Our paper extends this recent literature by empirically testing these predictions and providing concrete, empirical evidence on the effect of generative AI on labor market outcomes in the short run.

Additionally, other authors conducted field experiments randomizing access to generative AI tools and studied their effect on worker performance and productivity (Noy and Zhang, 2023; Brynjolfsson et al., 2023; van Inwegen et al., 2023; Peng et al., 2023; Dell'Acqua et al., 2023). This set of experiments documents the direct benefits of ChatGPT in a relatively controlled setting. Our paper extends the results in previous work by studying the effect on aggregate labor market outcomes (as opposed to individual productivity) at a larger scale and in a natural setting. This setting offers the advantage of reflecting real-world conditions and behaviors, providing insights that are robust to issues such as differential adoption rates and selection.

More broadly, our results also contribute to the literature examining the effect of AI technologies on economic activities. Ours is one of the first papers to directly evaluate the effect of generative AI specifically on organizations and workers. Past work has examined the consequences of deploying AI in areas such as machine translation (Brynjolfsson et al., 2019), customer service (Luo et al., 2019; Schanke et al., 2021), sales (Cao and Zhang, 2020; Luo et al., 2021), human resources (Tambe et al., 2019), online retail (Overby et al., 2010; Sun et al., 2019; Zhang et al., 2021), international trade (Goldfarb and Trefler, 2018), data market (Tucker, 2018; Jin, 2018), and programming (Cowgill et al., 2020; Cheng et al., 2022). As a general-purpose technology, AI has the potential to drastically change productivity (Brynjolfsson et al., 2018b; Agrawal et al., 2019b), and education (Mollick and Mollick, 2023) and innovation processes (Cockburn et al., 2018; Lifshitz-Assaf et al., 2019; Masclans et al., 2023). Lastly, Yilmaz et al. (2023) find that the introduction of Google's machine translation reduces translators' employment, especially for tasks with analytical elements. Similar to this work, they find that ChatGPT serves as a substitute to human input, in their case, the number of questions and answers on Stack Exchange.

2 Setting

ChatGPT, released by OpenAI, is a powerful large language model (LLM) consisting of 175 billion parameters and trained on large datasets of text. The trained data and flexibility of the models enable the learning of the statistical relationships between words and phrases, which in turn allow the model to generate text that is both natural and informative. Thanks to the models' advanced capabilities in natural language generation, LLMs can perform tasks typically considered high-skilled or creative. As a result, even occupations previously thought to be safe from automation, namely those requiring complex problem-solving ability and creativity, could be affected by this new technology. Immediate use cases of ChatGPT include content writing, copyediting, answering questions, and language translation. Besides its capability, ChatGPT is also free to use and accessible to the public. Perhaps unsurprisingly, ChatGPT's popularity grew rapidly, reaching 100 million users since its launch in November 2022. At the time of writing, generative AI is still in a phase of rapid development, with recent advancements such as GPT-4 and Bard.

To further evaluate the internal validity of our design and assess the generalizability of our findings, we also analyze another type of generative AI, which took place at a different time and affected a different set of highly trained workers. Specifically, we study the effect of the release of DALL-E 2, also developed by OpenAI, and Midjourney, which produce high-quality images. Similarly to how ChatGPT was trained on text, DALL-E 2 and and Midjourney were trained on millions of images using cutting-edge machine learning techniques in order to respond to natural language prompts. The main difference is that, rather than generating a text outputs, the two models produce high-quality images. Both models were released around the same time, with DALL-E 2 released in April 2022, and Midjourney in July.

We study the effects of generative AI on labor market outcomes in the context of Upwork, one of the largest online labor markets in the world. The platform matches employers with independent freelancers for small- to medium-size tasks. The services on the platform are typically remote jobs, ranging from data entry and graphic design to software development and marketing (Horton, 2010). A typical workflow starts with the creation of a job posting by an employer. The job posting includes a description of the job, the category of the service (e.g., writing or administrative support), the required skills or qualifications for the job, and the expected outputs and timeline. Once a job is posted, workers may apply to it by submitting a proposal or by responding to the employer's invitation to apply for the job (Barach and Horton, 2021).

Upwork is a good setting to study the short-term effects of generative AI on labor outcomes, for several reasons. First, we can obtain the freelancers' complete work histories on the platform and observe relevant freelancer attributes (or control for time-invariant unobserved differences). Second, because Upwork is a relatively short-term, spot labor market, employers are able to renegotiate contracts and rehire workers frequently, allowing for more flexibility compared with traditional, formal employment. Thus, while generative AI is still nascent to replace full-time jobs, its effect in an online labor market may already be detectable.[1]

3 Empirical Framework

3.1 Data

We obtained our primary data set from Upwork's freelancers API. We scraped all available information on all searchable freelancers' profile pages,[2] including their occupations, qualifications, and skills. We

[1]A potential limitation is that we do not observe off-platform employment and thus cannot say whether a reduction in employment on the platform represents a reduction in total employment or is merely substituting other forms of employment. However, Horton (2017) provides evidence that multi-homing is rare and that online and offline hiring are not very substitutable. In addition, we are unable to evaluate whether high-skill workers in the context of Upwork are also highly skilled in the general population.

[2]We could get data only on freelancers available through the API. A subset of freelancers are not "searchable" and thus do not appear in our data. Unfortunately, we do not have an estimate of the share of non-searchable freelancers. In addition, in order to reduce bias from fraudulent jobs, which are jobs generated to improve a freelancer's employment history, we omit freelancers whose income from the first job on the platform is less than 5% or more than 500% the average of their first five jobs, which is approximately the 10th and 90th percentiles. We discuss our data sample restrictions in detail in Appendix B.

also observe metrics on past employment on Upwork, such as total past earnings, number of previous jobs, success rate of completed jobs, and the reviews left by previous employers. Notably, we are able to observe only a snapshot of the freelancer pool in April 2023, and are unable to observe changes to a freelancer's profile page over time.[3]

We restrict our attention to the period from January 2022 through April 2023. Employment outcomes are aggregated to the freelancer-month level by the job start date; we focus on monthly number of jobs and monthly income.[4] We winsorize both outcomes at the 99% level. Because our research design uses variation across occupations, we restrict our attention to freelancers offering service in a single occupation (approximately 85% of the freelancers). Our final data set consists of approximately 80,000 freelancers.[5] Table 1 presents the summary statistics for the variables used in the main analyses by text-based and image-based AI tools. We note that, on average, a freelancer starts a job once every two to three months, for an average monthly pay of around $200. There exists, however, substantial variation across both months and freelancers.

Our secondary data set, which we use to provide descriptive motivation for our research design, consists of data obtained from Google Trends. The API provides relative interest in specific search terms on Google. For each one of the search terms, we obtain the change in interest in the United States over our sample period. Notably, Google allows only a comparison across five search terms at a time and always normalizes the results relative to the highest value. For this reason, when comparing across many search terms, we ran multiple queries, always including the highest-value search term.

3.2 Empirical Strategy

The release of ChatGPT in November 30, 2022,[6] provides a shock to both the salience and accessibility of generative AI.

[3]For example, our sample includes only freelancers with active profiles at the time of our data collection; we do not observe terminated accounts. Similarly, we are unable to observe changes made to freelancers' profile descriptions or qualifications over time.

[4]We also observe the total number of hours for a subset of jobs. This measure, however, is poorly populated and is available only for hourly jobs and not for fixed-price jobs.

[5]More specifically, the sample for the text-based AI analysis comprises 79,322 freelancers, and the sample for the analysis of image-based AI comprises 80,993 freelancers.

[6]https://openai.com/blog/chatgpt

Our primary goal is to estimate the direct short-term effect of this new technology on employment outcomes. Naturally, the release of ChatGPT has the potential to disrupt multiple industries and affect a large share of the workforce in the long term, similar to what has been observed for pre-ChatGPT AI technologies (Acemoglu et al., 2022). Looking at the short run, and even immediate run, we focus on occupations that are most prominently affected by the current rather than future capabilities of the model. One such susceptible type of industry is writing, and the affected occupations include content writing, editing, and proofreading. We chose to focus on this industry for the following reasons.

First, as an LLM, dialogue-based AI, ChatGPT is specifically trained using a large amount of text to predict and generate new text. ChatGPT particularly excels at understanding and generating human-like responses to a wide range of queries. Second, though ChatGPT has the potential to transform multiple industries in the long term (Eloundou et al., 2023)—similarly to other general-purpose technologies—the tangibility of tasks in the writing industry allows for a straightforward application of ChatGPT's text-generation capabilities even in the short run. For example, users who are looking for copyediting services can easily paste the text to be copyedited into the ChatGPT prompt and immediately evaluate the tool's output. In fact, one of the main stated goals of the developers of ChatGPT "is to improve their [the models'] ability to understand and generate natural language text, particularly in more complex and nuanced scenarios" (OpenAI, 2023).

Additionally, previous research has offered strong evidence that writing-related occupations are among the most affected. According to Eloundou et al. (2023), writing is consistently ranked among the top five occupations most exposed to ChatGPT, across different measurements based on AI models and on human experts.[7] Finally, the relevance of ChatGPT to writing-related tasks is often highlighted in popular media, which is perhaps more accessible to casual ChatGPT users.[8]

To more directly examine the interest in ChatGPT over time and by application type, we take advantage of Google Trends data. Google Trends provides access to all search queries entered in Google's

[7]Measurements are based on either human experts' opinions or GPT-4's ratings, and they evaluate how exposed an occupation is to either ChatGPT or ChatGPT-related software. Specifically, these measurements are Human α, Human β, Human ξ, Model α, Model β, and Model ξ in Eloundou et al. (2023).

[8]For instance, https://www.businessinsider.com/chatgpt-jobs-at-risk-replacement-artificial-intelligence-ai-labor-trends-2023-02#media-jobs-advertising-content-creation-technical-writing-journalism-2 and https://www.nytimes.com/2022/12/06/opinion/chatgpt-ai-skilled-jobs-automation.html.

search engine together with a relative volume index by geographies and over time. This tool allows researchers to observe Google users' relative interests in various topics. As presented in Panel A of Figure 1, we begin by examining the relative interest in the search terms "GPT" and "AI" over time. We document a sharp increase in the relative interest in ChatGPT-related queries following its launch. This observation is consistent with the rapid growth in ChatGPT's user base.

We then examine the relative interest in various occupational domains combined with the keyword "ChatGPT" as captured by Google search queries. We enter domains as they appear on the platform and combine them with the keyword "GPT" to obtain the search term (e.g., "GPT Translation" or "GPT Mobile Development"). We expect that users running a Google search for a combination of the term "GPT" and a specific task are interested in using ChatGPT to help with that task or intend to learn about how ChatGPT could help them with that task. Either way, a relative increase in the co-occurrence of the search term "GPT" used together with a term denoting a specific occupation suggests a higher interest in using ChatGPT in that occupation. The results, presented in Figure 1, Panel B, provide descriptive evidence to support our research design and the choice of writing-related jobs for our analysis. "GPT Writing" is by far the most commonly searched term compared with "GPT" and other domains, such as "GPT translation" or "GPT software development." This suggests that the public's attention and interest are skewed toward the implementation of AI technologies in the writing industry. We interpret these patterns as suggestive evidence consistent with the notion that while multiple industries are affected by the introduction of ChatGPT, its short-term implications are much more pronounced in writing-related tasks.

Similarly to how ChatGPT disrupted the writing industry, the release of image-based AI models, namely DALL-E 2 and Midjourney, disrupted the design and art industries. One key advantage of examining the release of image-based AI models is a clear definition of a treatment group and lack of spillovers to other job categories. This is because occupations not revolving around images are unlikely to be directly or even indirectly affected by the release of DALL-E 2 and Midjourney. In particular, we define the affected group as freelancers who offer design, image editing, and art services.[9] Notably, we omit the treated text-based occupations from our image-based analysis, and vice

[9] We restrict the treated group to occupations regarding *image*-based design, as opposed to video or audio editing.

versa.

Our analysis of image-based models focuses on two key releases: DALL-E 2, in April 2022, and Midjourney, in July 2022. To be conservative, we code the post-period as post-April 2022. Relative interest in those models over time can again be assessed using Google Trends data. The combined interest in both DALL-E 2 and Midjourney can be inferred from Panel A of Figure 1. However, since the interest in ChatGPT is substantially larger, it is difficult to assess the changes visually. Appendix Figure A1 provides a more detailed view, showing a notable increase in the relative interest in DALL-E 2 and Midjourney following their respective releases. Accordingly, we also observe a similar sharp rise in Google users' interest in AI image-related searches.

3.3 Research Design

The main specification uses the panel structure of the data to employ a difference-in-differences research design. In particular, we estimate the following specification:

$$Y_{it} = \beta \text{Post}_t \times \text{T}_i + X_{it} + \alpha_i + \delta_t + \varepsilon_{it}, \tag{1}$$

where i and t are indices for freelancer and month, respectively. In the specification we use for studying the effect of text-based generative AI, $Post_t$ is an indicator for post-April 2022 for image-based AI and November 2022 for text-based AI. T_i denotes the occupations most affected by generative AI, e.g., writing, proofreading, and copyediting for ChatGPT; and design- and image-focused occupations for Midjourney and DALL-E 2. X_{it} is a vector of time-varying freelancer-level characteristics, which includes the average feedback score (and an indicator for whether a score is missing), and inverse hyperbolic sine (IHS)-transformed past monthly income and number of previous jobs completed.[10] Y_{it} denotes our main outcomes of interest, monthly number of jobs and income, which we also decompose into the extensive and intensive margins. In particular, we first estimate the total effect on monthly

[10]As presented in Appendix Table A1, we experiment with alternative control variables and clustering. Across the board, the results remain qualitatively similar and statistically significant.

number of jobs and income, which are IHS-transformed; second, on the intensive margin, as measured by the IHS-transformed number of jobs and income conditional on working; and third, on the extensive margin, as measured by the indicator for whether the freelancer accepted a new job at a given month. All regressions also include freelancer and month fixed effects. Finally, we cluster the standard errors for the error term ε_{it} at the occupation level, corresponding to the level at which treatment is assigned (Bertrand et al., 2004; Abadie et al., 2023).

The key identifying assumption of the difference-in-differences design is that freelancers operating in treated and untreated occupations would have evolved similarly absent the release of ChatGPT.[11] To evaluate this assumption, we estimate a more flexible difference-in-differences event-study specification of the following form:

$$Y_{it} = \sum_{j=T_0}^{-1} \beta_j Pre_j \times T_i + \sum_{k=1}^{T_1} \beta_k Post_k \times T_i + X_{it} + \alpha_i + \delta_t + \varepsilon_{it}, \tag{2}$$

where Pre_j and $Post_k$ are dummy variables equal to 1 when an observation is j months before or k months after the addition of the label, and 0 otherwise.

4 Results

We begin by estimating the effect of ChatGPT on the employment and compensation of freelancers in affected occupations, as described in Equation 1. Our main results are presented in Panel A of Table 2. Across specifications, we find significant negative effects of the release of ChatGPT on freelancer employment on the platform. Columns (1) and (2) present the total effect on monthly employment and earnings of freelancers in writing-related occupations, respectively, relative to freelancers in less affected occupations. Following the release of ChatGPT, the monthly number of jobs on the platform for freelancers in more affected occupations decreased by 2% (s.e.=0.004), and total monthly com-

[11] Another key assumption is the stable unit treatment value assumption (SUTVA), which is not directly testable in our setting. We note that while ChatGPT can potentially affect all categories to a varying extent, this is not the case for image-based models, which are applicable to a much narrower and well-defined set of tasks. To the extent that control units are affected by ChatGPT to a lesser degree, we think of our estimates as a lower bound of the true effect of the new technology.

pensation decreased by 5.2% (s.e.=0.016).

In columns (3) through (5), we decompose the total effect into the extensive and intensive margins. Column (3) presents the effect on the linear probability that a freelancer receives at least one job in a given month, which we interpret as the extensive margin of working on the platform. Following the release of ChatGPT, the probability that a freelancer offering writing-related services starts a new job in a given month decreased by 1.2 percentage points, which is approximately a 10% drop compared to baseline employment. In columns (4) and (5), we estimate the effect on the number of jobs and monthly income, conditional on receiving at least one job that month, respectively. The number of jobs decreased by approximately 4.7% (s.e.=0.011), which is substantially larger than the total effect. The effect on income is approximately the same size, a decrease of 5.1%, though it is much noisier (s.e.=0.037).

Similarly, Panel B presents the estimates of the effect of the release of image-based AI on employment. The estimates largely echo those in Panel A: We find statistically significant reductions in both the monthly number of jobs and revenue generated on the platform. Reassuringly, though the estimates represent the effect at a different point in time and for a different set of workers, relative to the effect of ChatGPT on affected occupations, the point estimates are almost identical: Specifically, the monthly number of jobs dropped by 2.1% (s.e.= 0.005) in affected occupations, and total monthly compensation decreased by 5.2% (s.e.=0.019).

In addition, to alleviate concerns regarding unobserved secular trends in freelancers' employment driving the main results, in Figure 2 and Figure 3 we present the event time for our main specifications, as detailed in Equation 2. We observe similar pre-trends between freelancers in eventually treated and untreated occupations prior to the release of ChatGPT. This suggestive evidence assuages concerns regarding any violations of the parallel trends assumption. In both figures, we see sharp, persistent, and growing decreases in both the monthly number of jobs (Panel A) and monthly compensation (Panel B) on the platform following the release of generative AI models, suggesting a causal relationship.

The estimated effect sizes, about 2%–5% decreases in the number of jobs and income, are largely consistent with those of concurrent research on the effect of artificial intelligence on employment in similar settings. For example, Liu et al. (2023) and Yilmaz et al. (2023) find average reductions of 3.87% and 4.81%, respectively. Thus, even in the nascent stage of generative AI, its estimated effect on the labor market is generally at least as large as the effect of other shocks to labor markets (Cengiz et al., 2019; Lipsitz and Starr, 2022; Farber et al., 2021) and other disruptive technologies (Acemoglu and Restrepo, 2020; Autor et al., 2003; Feigenbaum and Gross, 2024).

To assess the magnitudes of these effects, we provide a simple back-of-the-envelope calculation to evaluate the potential effect of generative AI on the revenue of the platform. According to Upwork's 2023 10K filing, the company's revenue in 2023 from both clients (i.e., employers) and freelancers was more than 689 million dollars.[12] Our estimate suggests a drop of 5.2% in transactions in affected occupations. Assuming generative AI obtains the capabilities to substitute *all* jobs on the platform, Upwork could lose approximately 34 million dollars per year in revenue.[13] Extrapolating these effects to the entire labor market is beyond the scope of this paper. However, even a significantly smaller effect in overall employment is expected to have meaningful ramifications to the economy.

4.1 Robustness of Main Results

We conduct several tests to evaluate the robustness of the main results. The first set of analyses examines the robustness of the estimates to alternative specifications. The results are presented in Appendix Tables A1 and A2 for text-based and image-based generative AI models, respectively. We begin by presenting estimation results without any control variables. The estimates remain largely the same magnitude but are somewhat noisier. In addition, due to the small number of clusters, we also use wild bootstrap (Cameron et al., 2008) to calculate the t-statistics of the estimates. Alternatively, we cluster standard errors at the industry level, rather than at the occupation level. In Appendix Table A3, we also examine the robustness of the main analysis to alternative empirical specifications and transformations of the dependent variable, such as logit, Poisson, and negative binomial, as well as using the log transformation instead of IHS. The main results remain generally robust to these

[12] https://investors.upwork.com/static-files/a413c452-40f0-4dd9-9c88-1add6fd05366

[13] This figure may be over- or underestimating the effect of generative AI on online labor platforms in the long run. On the one hand, advances in generative AI may lead to replacement of workers in more occupations, as well as the ability of generative AI to substitute a larger share of tasks in each occupation. On the other hand, the emergence of a new class of freelancers specializing in using and engaging with generative AI systems may partially offset the adverse effect of generative AI on employment.

alternative procedures and specifications.

A related concern is that there exist unobserved differences between freelancers in affected and unaffected occupations that are driving the main results, regardless of the introduction of generative AI. We thus conduct an additional analysis in which we match freelancers based on pre-release characteristics to account for observable and potentially unobservable differences. Specifically, we use a coarsened exact matching (CEM) algorithm, where we coarsely match on the freelancer's past employment, past income, education, badges, success rate, and hourly rate on the platform. As presented in Table A4, for both of our main specifications, using the matched sample improves the precision of, and generally strengthens, the main results.

Another potential concern is that the estimated effects stem from disparities in seasonal trends across professions that might be misconstrued as the effects of generative AI. While this concern is alleviated by the fact that we find consistent evidence across two different worker groups at two different times, we provide an additional piece of evidence to address this potential alternative explanation. Specifically, we evaluate changes in outcomes following the introduction of generative AI using a year-by-year comparison. In particular, our control group consists of prior-year outcomes *for the same set of treated freelancers* in the corresponding calendar months. This design ensures that the control group has the same seasonal trends, for instance, due to summer vacation or holidays. The design is similar in spirit to previous papers studying online activity, as well as the effect of AI, such as Troncoso et al. (2023); Goldberg et al. (2023); Burtch et al. (2023). As presented in Appendix Table A5, the main results are robust to this alternative specification and often larger in magnitude.

4.2 Heterogeneous Treatment Effects

Having documented robust evidence of the adverse effects of the introduction of generative AI on employment outcomes, we now turn to explore which type of workers are most affected by the new technology. We are particularly interested in understanding whether freelancers' quality or experience can mitigate (or potentially exacerbate) the effects of generative AI on employment outcomes. Our granular, freelancer-level data allow us to examine the heterogeneity in treatment effects by the observable attributes of workers performing similar jobs and tasks.

Previous work suggests that high-quality and high-skilled suppliers are generally less threatened by adverse market shocks (e.g., Syverson, 2004; Reshef, 2023) and, in particular, labor market shocks (e.g., Autor et al., 2003; Acemoglu and Restrepo, 2018; Balasubramanian et al., 2022). If that is the case here, then we would expect generative AI to mostly hinder the performance of relatively low-quality freelancers. On the other hand, several recent papers suggest that similar technologies differentially benefit low-skilled workers by disproportionately increasing their productivity and performance (e.g., Noy and Zhang, 2023; Brynjolfsson et al., 2023). Understanding treatment effect heterogeneity by worker quality can thus shed light on whether generative AI mainly affects specific types of workers and the potential effects on the levels of inequality across workers.

To this end, we interact our main estimator, $Post_t \times T_i$, with several measures of freelancers' past performance: the number of past jobs on the platform, total past income on the platform, and skill level necessitated by past jobs.[14] We define these heterogeneity measures based on their November 2022 levels, right before ChatGPT was launched, and on their April 2022 levels, before the launch of DALL-E 2. In addition, we also test for heterogeneity by the freelancers' attributes that are most salient on their profile page: whether they received a "Top Rated" badge from Upwork, their past success rate (as detailed by past employers), and their stated hourly rate. Notably, we observe only these three measures at the time of collecting the data and cannot assess whether they changed over our sample period.

The results are presented in Table 3 and Table 4. In both tables, Panel A presents the heterogeneous treatment effects on the monthly number of jobs, and Panel B presents the effects on monthly income. In general, for both types of freelancers and across the various measures of freelancer quality, we can generally rule out that freelancer quality moderates the adverse effects of generative AI on employment. For example, observing column 2 in Table 3 we find that a 1% increase in freelancers' earnings leads to an additional 0.5% (s.e.=0.002) reduction in the number of jobs and a 1.7% (s.e.= 0.009) reduction in monthly income for text-based services. Similarly, as presented in Table 4, for

[14]For each job posting, the employer must indicate the desired level of freelancer experience: entry, intermediate, or expert. We code these levels as 1, 2, and 3, respectively. For each worker we calculate the average desired experience level for all jobs performed on the platform prior to the introduction of ChatGPT. We interpret this measure as an additional proxy for a freelancer's skill level.

image-based services we estimate reductions of 1.1% (s.e.= 0.002) and 3.3% (s.e.= 0.008) in the number of monthly jobs and income, respectively, for each 1% increase in prior earnings on the platform. Across the board, we mostly observe negative point estimates on the interaction terms. These effects, however, are often not statistically significant, especially for monthly income, which tends to be substantially noisier than the monthly number of jobs.

We interpret the heterogeneity results as suggesting that offering high-quality service does not mitigate the adverse effects of generative AI on employment. Though the results are slightly weaker, they could be interpreted as suggestive evidence that high-quality workers are disproportionately hurt by the adverse effects of generative AI. There are several potential explanations consistent with this interpretation: First, there may be a myriad of unobserved factors that are correlated with freelancer quality and the effect of AI. For instance, high-quality freelancers may engage in fundamentally different types of jobs, particularly ones that are more easily substituted by generative AI.

A second potential explanation is that low-quality freelancers are more likely to adopt the new technology, thus differentially benefiting from generative AI. This explanation is consistent with evidence from Noy and Zhang, 2023; Brynjolfsson et al., 2023. This supply-side explanation suggests that the benefits to low-quality sellers somewhat moderate the adverse effects of generative AI. Finally, we generally find that high-quality freelancers tend to charge higher average job prices[15] than lower-quality freelancers but did not lower these prices significantly after the introduction of generative AI to the market. If AI serves as a cheap substitute to labor and high-quality sellers continue to charge higher prices on average, then we would expect them to be most harmed by the new technology. Unfortunately, since our data do not allow us to directly observe job postings, wages, or the adoption of AI, we are unable to tease out the relative importance of these mechanisms. We leave this work to future research.

[15] We obtain this estimate by dividing monthly total income by monthly number of jobs. We note that this rough proxy does not take into account jobs without a posted price, nor does it differentiate between fixed-cost and per-hour jobs.

5 Discussion and Conclusion

This paper studies the short-term effects of generative AI on labor outcomes by estimating the effect of ChatGPT, as well as DALL-E 2 and Midjourney, on the employment of workers in a large online labor market. We find that freelancers who offer services in occupations most affected by AI experienced reductions in both employment and earnings. The release of ChatGPT led to a 2% drop in the number of jobs on the platform and a 5.2% drop in monthly earnings. We find a similar reduction in the employment outcomes of freelancers offering design and image-editing services following the introduction of image-focused generative AI. The results are robust to several alternative tests, including matching procedures and year-by-year comparisons. In addition, we find that offering high-quality service does not mitigate the negative effect of AI on freelancers, and in fact present suggestive evidence that high-quality freelancers are disproportionately hurt by the release of generative AI.

Our paper has several implications for policymakers and business leaders. Aside from the many benefits of AI technology, it may also have substantial economic and societal ramifications. The results suggest that in the short run, employers, rather than workers, are better positioned to capture the value generated by rapid AI advancement, as the new technology replaces human labor across various skill levels. On the other hand, workers—in particular freelancers, who lack formal protection against adverse labor shocks—are likely to disproportionately experience immediate and powerful effects.

These findings give scope for public policies, especially in industries highly vulnerable to AI disruption, or vocational training programs for skills needed in this new era, such as harnessing the power of generative AI. Finally, as generative AI capabilities continue to develop, business leaders need to strategically plan for the age of AI (Bailey, 2022). Successful planning involves anticipating required skills, balancing AI and human labor, and designing career pathways to cultivate and promote the right talents. Big changes in management practices and corporate culture may also be needed to adapt to AI integration in everyday operations.

Our work has several limitations. The online freelancer market we study typically involves short-term, unbundled jobs, allowing for easier labor adjustments relative to traditional employment. While the setting is useful for studying the short-term effects of generative AI, it is different from traditional employment settings where tasks are integrated within firm boundaries. Therefore, our estimates may not quantitatively extrapolate to traditional labor markets, which typically involve longer-term employment, bundled tasks, and stronger firm incentives to hire skilled workers. Additionally, we cannot distinguish between demand and supply effects in our estimates. Identifying the drivers and effects of AI adoption within and across organizations is crucial for a more holistic understanding of AI effect on the labor market.

The paper focuses on the short-term effects of generative AI by comparing labor market outcomes between "more affected" and "less affected" occupations. We envision several potential long-run spillover effects on both types of occupations that warrant future research. As technology improves, it will likely be adopted in additional sectors and affect even initially unaffected sectors, further altering the labor-technology relationship (Berg et al., 2023). Our estimates thus suggest that these effects could grow and become more pervasive in the long run. In contrast, there may be other mitigating factors in the long run, such as adjustments by different parties. For example, companies may switch focus to complement generative AI and adapt to the new technological environment. Similarly, workers may strategically switch jobs or acquire new skills to mitigate some of the adverse effects of generative AI on labor outcomes.

Lastly, reduced costs in one industry can also boost economic activity and innovation in others. For example, Brynjolfsson et al. (2019) show that AI-based machine translation on eBay, while potentially lowering translators' earnings, increases international trade for small businesses. This finding demonstrates how increased efficiency in one sector can lead to lower costs, higher output, more transactions, and potentially increased employment and earnings in complementary sectors. Finally, more transactions and potentially lower prices could enhance consumer welfare. Assessing the total welfare implications of generative AI for all stakeholders is a promising direction for future research.

References

ABADIE, A., S. ATHEY, G. W. IMBENS, AND J. M. WOOLDRIDGE (2023): "When should you adjust standard errors for clustering?" *The Quarterly Journal of Economics*, 138, 1–35.

ACEMOGLU, D. AND D. AUTOR (2011): "Skills, tasks and technologies: Implications for employment and earnings," in *Handbook of labor economics*, Elsevier, vol. 4, 1043–1171.

ACEMOGLU, D., D. AUTOR, J. HAZELL, AND P. RESTREPO (2022): "Artificial intelligence and jobs: evidence from online vacancies," *Journal of Labor Economics*, 40, S293–S340.

ACEMOGLU, D. AND P. RESTREPO (2018): "Artificial intelligence, automation, and work," in *The economics of artificial intelligence: An agenda*, University of Chicago Press, 197–236.

——— (2020): "Robots and jobs: Evidence from US labor markets," *Journal of political economy*, 128, 2188–2244.

AGRAWAL, A., J. S. GANS, AND A. GOLDFARB (2019a): "Artificial intelligence: the ambiguous labor market impact of automating prediction," *Journal of Economic Perspectives*, 33, 31–50.

——— (2019b): "Exploring the impact of artificial intelligence: Prediction versus judgment," *Information Economics and Policy*, 47, 1–6.

ALLEN, R. AND P. CHOUDHURY (2022): "Algorithm-augmented work and domain experience: The countervailing forces of ability and aversion," *Organization Science*, 33, 149–169.

ARGYRES, N. S. AND T. R. ZENGER (2012): "Capabilities, transaction costs, and firm boundaries," *Organization Science*, 23, 1643–1657.

AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): "The skill content of recent technological change: An empirical exploration," *The Quarterly journal of economics*, 118, 1279–1333.

BAILEY, D. E. (2022): "Emerging technologies at work: Policy ideas to address negative consequences for work, workers, and society," *ILR Review*, 75, 527–551.

BAILEY, D. E., S. FARAJ, P. J. HINDS, P. M. LEONARDI, AND G. VON KROGH (2022): "We are all theorists of technology now: A relational perspective on emerging technology and organizing," *Organization Science*, 33, 1–18.

BALASUBRAMANIAN, N., Y. YE, AND M. XU (2022): "Substituting human decision-making with machine learning: Implications for organizational learning," *Academy of Management Review*, 47, 448–465.

BARACH, M. A. AND J. J. HORTON (2021): "How do employers use compensation history? Evidence from a field experiment," *Journal of Labor Economics*, 39, 193–218.

BERG, J., M. RAJ, AND R. SEAMANS (2023): "Capturing Value from Artificial Intelligence," *Academy of Management Discoveries*.

BERTRAND, M., E. DUFLO, AND S. MULLAINATHAN (2004): "How much should we trust differences-in-differences estimates?" *The Quarterly journal of economics*, 119, 249–275.

BRYNJOLFSSON, E., X. HUI, AND M. LIU (2019): "Does machine translation affect international trade? Evidence from a large digital platform," *Management Science*, 65, 5449–5460.

BRYNJOLFSSON, E., D. LI, AND L. R. RAYMOND (2023): "Generative AI at work," Tech. rep., National Bureau of Economic Research.

BRYNJOLFSSON, E., T. MITCHELL, AND D. ROCK (2018a): "What can machines learn and what does it mean for occupations and the economy?" in *AEA papers and proceedings*, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, vol. 108, 43–47.

BRYNJOLFSSON, E., D. ROCK, AND C. SYVERSON (2018b): "Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics," in *The economics of artificial intelligence: An agenda*, University of Chicago Press, 23–57.

BURTCH, G., D. LEE, AND Z. CHEN (2023): "The Consequences of Generative AI for UGC and Online Community Engagement," *Available at SSRN 4521754*.

CAMERON, A. C., J. B. GELBACH, AND D. L. MILLER (2008): "Bootstrap-based improvements for inference with clustered errors," *The review of economics and statistics*, 90, 414–427.

CAO, X. AND D. ZHANG (2020): "The impact of forced intervention on AI adoption," *Available at SSRN 3640862*.

CENGIZ, D., A. DUBE, A. LINDNER, AND B. ZIPPERER (2019): "The effect of minimum wages on low-wage jobs," *The Quarterly Journal of Economics*, 134, 1405–1454.

CHENG, Z., D. LEE, AND P. TAMBE (2022): "Innovae: Generative ai for understanding patents and innovation," *Available at SSRN 3868599*.

CHOUDHURY, P., E. STARR, AND R. AGARWAL (2020): "Machine learning and human capital complementarities: Experimental evidence on bias mitigation," *Strategic Management Journal*, 41, 1381–1411.

COCKBURN, I. M., R. HENDERSON, AND S. STERN (2018): "The impact of artificial intelligence on innovation: An exploratory analysis," in *The economics of artificial intelligence: An agenda*, University of Chicago Press, 115–146.

COWGILL, B., F. DELL'ACQUA, S. DENG, D. HSU, N. VERMA, AND A. CHAINTREAU (2020): "Biased programmers? Or biased data? A field experiment in operationalizing AI ethics," in *Proceedings of the 21st ACM Conference on Economics and Computation*, 679–681.

DELL'ACQUA, F., E. MCFOWLAND, E. R. MOLLICK, H. LIFSHITZ-ASSAF, K. KELLOGG, S. RAJENDRAN, L. KRAYER, F. CANDELON, AND K. R. LAKHANI (2023): "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality," *Harvard Business School Technology & Operations Mgt. Unit Working Paper*.

ELOUNDOU, T., S. MANNING, P. MISHKIN, AND D. ROCK (2023): "Gpts are gpts: An early look at the labor market impact potential of large language models," *arXiv preprint arXiv:2303.10130*.

FARAJ, S., S. PACHIDI, AND K. SAYEGH (2018): "Working and organizing in the age of the learning algorithm," *Information and Organization*, 28, 62–70.

FARBER, H. S., D. HERBST, I. KUZIEMKO, AND S. NAIDU (2021): "Unions and inequality over the twentieth century: New evidence from survey data," *The Quarterly Journal of Economics*, 136, 1325–1385.

FEIGENBAUM, J. AND D. P. GROSS (2024): "Answering the Call of Automation: How the Labor Market Adjusted to Mechanizing Telephone Operation," *The Quarterly Journal of Economics*, qjae005.

FELTEN, E. W., M. RAJ, AND R. SEAMANS (2018): "A method to link advances in artificial intelligence to occupational abilities," in *AEA Papers and Proceedings*, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, vol. 108, 54–57.

——— (2023): "Occupational heterogeneity in exposure to generative ai," *Available at SSRN 4414065*.

FURMAN, J. AND R. SEAMANS (2019): "AI and the Economy," *Innovation policy and the economy*, 19, 161–191.

FURMAN, J. L. AND F. TEODORIDIS (2020): "Automation, research technology, and researchers' trajectories: Evidence from computer science and electrical engineering," *Organization Science*, 31, 330–354.

GOLDBERG, S., G. JOHNSON, AND S. SHRIVER (2023): "Regulating privacy online: An economic evaluation of the GDPR," *American Economic Journal: Microeconomics, Forthcoming*.

GOLDFARB, A. AND D. TREFLER (2018): "Artificial intelligence and international trade," in *The* *economics of artificial intelligence: an agenda*, University of Chicago Press, 463–492.

GRANT, R. M. (1996): "Toward a knowledge-based theory of the firm," *Strategic management journal*, 17, 109–122.

HORTON, J. J. (2010): "Online labor markets," in *Internet and Network Economics: 6th International Workshop, WINE 2010, Stanford, CA, USA, December 13-17, 2010. Proceedings 6*, Springer, 515–522.

——— (2017): "Price floors and employer preferences: Evidence from a minimum wage experiment," *Available at SSRN 2898827*.

JIN, G. Z. (2018): "Artificial intelligence and consumer privacy," in *The economics of artificial intelligence: An agenda*, University of Chicago Press, 439–462.

KELLOGG, K. C., M. A. VALENTINE, AND A. CHRISTIN (2020): "Algorithms at work: The new contested terrain of control," *Academy of Management Annals*, 14, 366–410.

KRAKOWSKI, S., J. LUGER, AND S. RAISCH (2023): "Artificial intelligence and the changing sources of competitive advantage," *Strategic Management Journal*, 44, 1425–1452.

LEBOVITZ, S., H. LIFSHITZ-ASSAF, AND N. LEVINA (2022): "To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis," *Organization Science*, 33, 126–148.

LIFSHITZ-ASSAF, H., F. NG, A. KITTUR, AND R. KRAUT (2019): "Using Technology to Augment Professionals, Instead of Replacing Them, for Innovative Problem Solving," *Working Paper*.

LIPSITZ, M. AND E. STARR (2022): "Low-wage workers and the enforceability of noncompete agreements," *Management Science*, 68, 143–170.

LIU, J., X. XU, Y. LI, AND Y. TAN (2023): "" Generate" the Future of Work through AI: Empirical Evidence from Online Labor Markets," *arXiv preprint arXiv:2308.05201*.

LUO, X., M. S. QIN, Z. FANG, AND Z. QU (2021): "Artificial intelligence coaches for sales agents: Caveats and solutions," *Journal of Marketing*, 85, 14–32.

LUO, X., S. TONG, Z. FANG, AND Z. QU (2019): "Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases," *Marketing Science*, 38, 937–947.

MASCLANS, R., S. HASAN, AND W. M. COHEN (2023): "The Commercial Potential of Science and Its Realization: Evidence From a Measure Using a Large Language Model," *Available at SSRN 4481928*.

MINTZBERG, H. (1980): "Structure in 5's: A Synthesis of the Research on Organization Design," *Management science*, 26, 322–341.

MOKYR, J., C. VICKERS, AND N. L. ZIEBARTH (2015): "The history of technological anxiety and the future of economic growth: Is this time different?" *Journal of economic perspectives*, 29, 31–50.

MOLLICK, E. AND L. MOLLICK (2023): "Assigning AI: Seven Approaches for Students, with Prompts," *arXiv preprint arXiv:2306.10052*.

MURRAY, A., J. RHYMER, AND D. G. SIRMON (2021): "Humans and technology: Forms of conjoined agency in organizations," *Academy of Management Review*, 46, 552–571.

NOY, S. AND W. ZHANG (2023): "Experimental evidence on the productivity effects of generative artificial intelligence," *Available at SSRN 4375283*.

OPENAI (2023): "Gpt-4 technical report," *Technical report*.

OVERBY, E., S. A. SLAUGHTER, AND B. KONSYNSKI (2010): "Research commentary—the design, use, and consequences of virtual processes," *Information Systems Research*, 21, 700–710.

PENG, S., E. KALLIAMVAKOU, P. CIHON, AND M. DEMIRER (2023): "The impact of ai on developer productivity: Evidence from github copilot," *arXiv preprint arXiv:2302.06590*.

PFEFFER, J. AND G. SALANCIK (2015): "External control of organizations—Resource dependence perspective," in *Organizational Behavior 2*, Routledge, 373–388.

RAISCH, S. AND S. KRAKOWSKI (2021): "Artificial intelligence and management: The automation–augmentation paradox," *Academy of management review*, 46, 192–210.

RESHEF, O. (2023): "Smaller slices of a growing pie: The effects of entry in platform markets," *American Economic Journal: Microeconomics*, 15, 183–207.

SCHANKE, S., G. BURTCH, AND G. RAY (2021): "Estimating the impact of "humanizing" customer service chatbots," *Information Systems Research*, 32, 736–751.

SUN, C., Z. J. SHI, X. LIU, A. GHOSE, X. LI, AND F. XIONG (2019): "The effect of voice ai on consumer purchase and search behavior," *NYU Stern School of Business*.

SYVERSON, C. (2004): "Market structure and productivity: A concrete example," *Journal of political Economy*, 112, 1181–1222.

TAMBE, P., P. CAPPELLI, AND V. YAKUBOVICH (2019): "Artificial intelligence in human resources management: Challenges and a path forward," *California Management Review*, 61, 15–42.

TRONCOSO, I., R. FU, N. MALIK, AND D. PROSERPIO (2023): "Algorithm failures and consumers' response: Evidence from Zillow," *Available at SSRN 4520172*.

TSCHANG, F. T. AND E. ALMIRALL (2021): "Artificial intelligence as augmenting automation: Implications for employment," *Academy of Management Perspectives*, 35, 642–659.

TUCKER, C. (2018): "Privacy, algorithms, and artificial intelligence," in *The economics of artificial intelligence: An agenda*, University of Chicago Press, 423–437.

VAN INWEGEN, E., Z. T. MUNYIKWA, AND J. J. HORTON (2023): "Algorithmic writing assistance on jobseekers' resumes increases hires," Tech. rep., National Bureau of Economic Research.

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

YILMAZ, E. D., I. NAUMOVSKA, AND V. A. AGGARWAL (2023): "AI-Driven Labor Substitution: Evidence from Google Translate and ChatGPT," *Available at SSRN*.

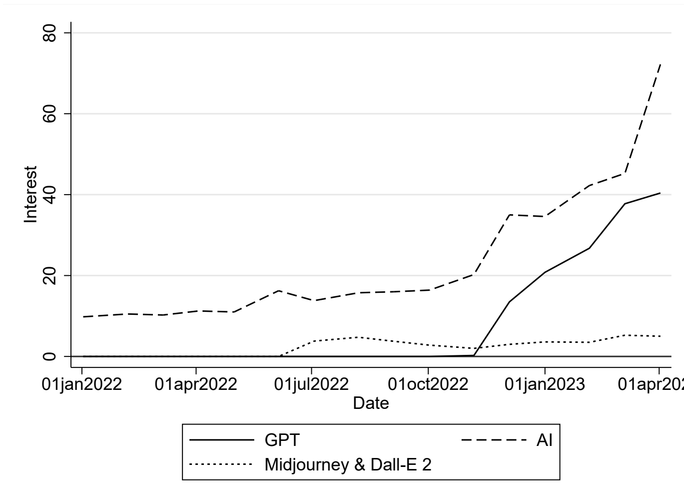
ZAMMUTO, R. F., T. L. GRIFFITH, A. MAJCHRZAK, D. J. DOUGHERTY, AND S. FARAJ (2007): "Information technology and the changing fabric of organization," *Organization science*, 18, 749–762.

ZHANG, M., T. SUN, L. LUO, AND J. GOLDEN (2021): "Consumer and AI Co-creation: When and Why Human Participation Improves AI Creation." *USC Marshall School of Business Research Paper Sponsored by iORB, No. Forthcoming*.

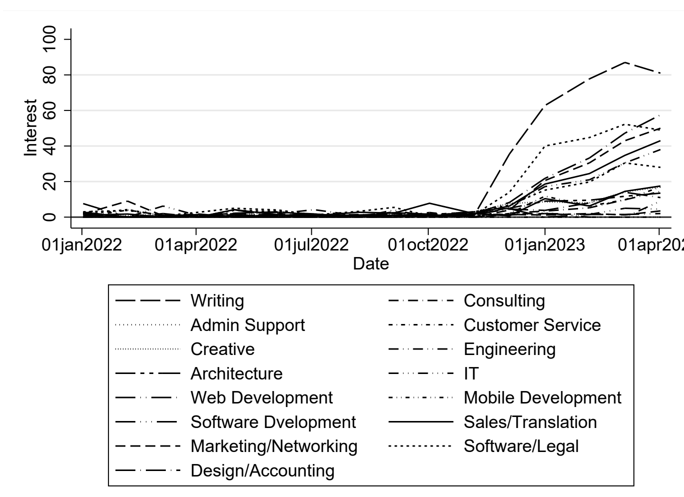
Figures

Figure 1: Descriptive Evidence on Interest in ChatGPT Over Time and Across Occupations

Panel A: Relative Interest in ChatGPT and AI



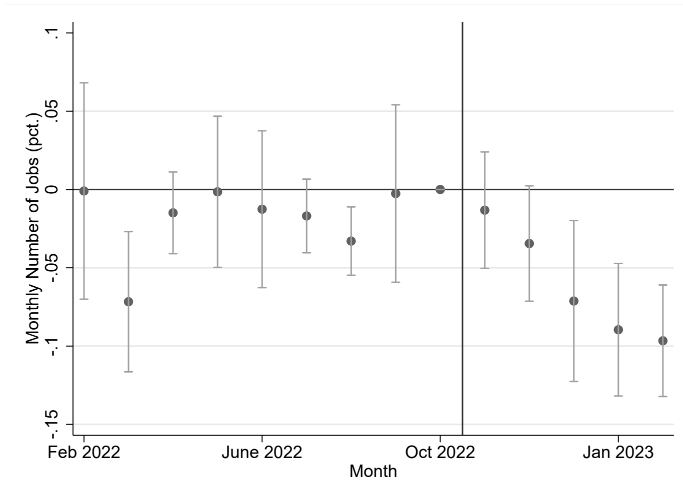
Panel B: Relative Interest in ChatGPT by Types of Occupation



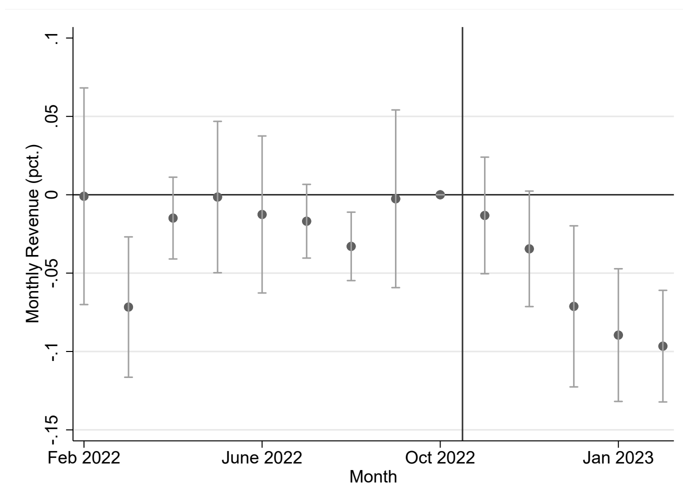
Notes: This figure displays the relative interest in separate search terms on Google Trends. Panel A presents relative interest in four separate search terms, and combines the last two terms. Panel B presents the relative interest over time in nineteen separate search terms, where each term represents a domain on the platform and was searched together with the word “GPT.” For example, “Writing” refers to interest in the search term “GPT Writing.”

Figure 2: The Effect of Text-based Generative AI on Freelancers' Employment on the Platform

Panel A: Monthly Number of Jobs



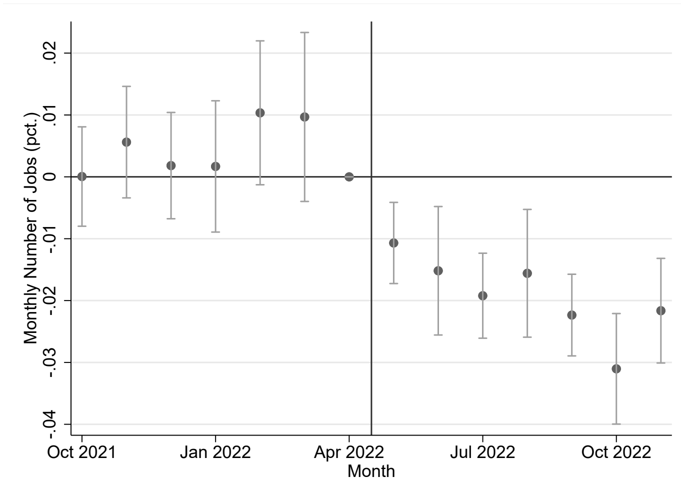
Panel B: Monthly Revenue



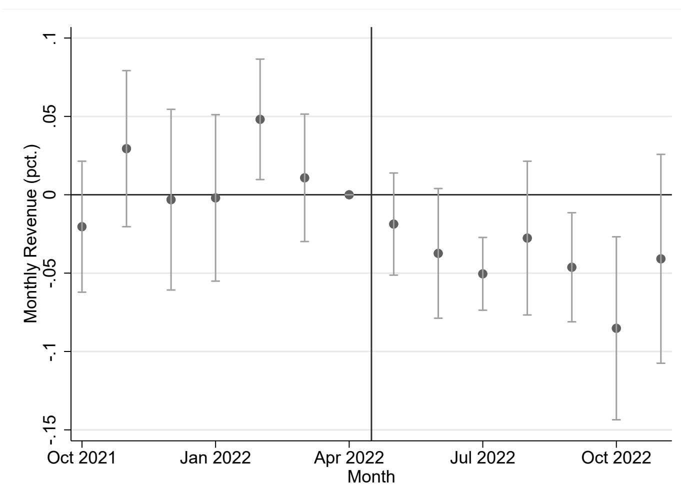
Notes: The figure displays results for monthly jobs and revenue, where the estimate between treatment and control occupations is allowed to vary for each month around the launch of ChatGPT (see description around Equation 2). The omitted category is October 2022. Each panel also reports 95% confidence intervals. Standard errors are clustered at the occupation level.

Figure 3: The Effect of Image-based Generative AI on Freelancers' Employment on the Platform

Panel A: Monthly Number of Jobs



Panel B: Monthly Revenue



Notes: This figure displays results for monthly jobs and revenue, where the estimate between treatment and control occupations is allowed to vary for each month around the launch of DALL-E 2 and Midjourney (see description around Equation 2). The omitted category is April 2022. Each panel also reports 95% confidence intervals. Standard errors are clustered at the occupation level.

Tables

Table 1: Descriptive Statistics

	Text-based Sample				Image-based Sample			
	Mean	SD	P5	P95	Mean	SD	P5	P95
No. of Freelancers	79,322	-	-	-	80,983	-	-	-
US-based	0.16	0.37	0	1	0.14	0.35	0	1
Top-rated Badge	0.34	0.47	0	1	0.34	0.47	0	1
Past Job Success Rate	0.62	0.46	0	1	0.62	0.46	0	1
Hourly Rate	28.3	29.6	5	75	27.5	28.7	5	75
Worked in Month	0.18	0.38	0	1	0.24	0.43	0	1
Monthly Number of Jobs	0.35	1.29	0	2	0.42	1.35	0	2
Monthly Earnings	180.7	1623.2	0	570	229.6	1964.8	0	750
Mean Monthly Score	4.81	0.61	3.9	5	4.83	0.58	4	5
Job Skill level	2.12	0.62	1	3	2.12	0.61	1	3

Notes: This table provides the summary statistics of the variables used in the main analysis.

Table 2: The Effect of Generative AI on Freelancers' Employment on the Platform

	Total		Extensive Margin	Intensive Margin	
	(1) Num. of Jobs	(2) Income	(3) Worked	(4) Num. of Jobs	(5) Income
	Panel A: Text-based Generative AI				
Post × Treatment	-0.020***	-0.052***	-0.012***	-0.047***	-0.051
	(0.004)	(0.016)	(0.002)	(0.011)	(0.033)
Observations	1236052	1236052	1236052	171541	174697
# of Clusters	195	195	195	193	193
Adjusted R^2	0.41	0.34	0.32	0.30	0.46
	Panel B: Image-based Generative AI				
Post × Treatment	-0.021***	-0.052***	-0.014***	-0.018*	-0.028
	(0.005)	(0.019)	(0.003)	(0.009)	(0.019)
Observations	1498086	1498086	1498086	326468	292498
# of Clusters	192	192	192	192	191
Adjusted R^2	0.66	0.32	0.38	0.65	0.46

Notes: This table presents OLS regressions relating freelancers' employment on the platform to the introduction of text-based (Panel A) and image-based (Panel B) AI in a difference-in-differences design (see description around Equation 1). *Treatment* is an indicator of whether an occupation is substantially affected by the introduction of generative AI. The unit of observation is the freelancer-month. The dependent variables are the total monthly number of jobs and income (columns 1 and 2), an indicator for whether the freelancer had at least one job that month (column 3), and the monthly number of jobs and income conditional on working (columns 4 and 5). In columns 1, 2, 4, and 5, the dependent variables are inverse hyperbolic sine transformed. All regressions include freelancer and month fixed effects. Standard errors are in parentheses and are clustered at the occupation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: The Effect of Text-based generative AI on Freelancers' Employment on the Platform: Heterogeneity by Freelancer Attributes and Performance

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Monthly Number of Jobs					
Post × Treatment	-0.010	0.008	-0.009	-0.017***	0.020	-0.006***
	(0.010)	(0.017)	(0.013)	(0.003)	(0.014)	(0.002)
Post × Treat. × Var.	-0.002	-0.005**	-0.012**	-0.024***	-0.010***	-0.028***
	(0.003)	(0.002)	(0.006)	(0.007)	(0.003)	(0.007)
	Panel B: Monthly Income					
Post × Treatment	0.006	0.028	-0.026	-0.054***	0.131*	0.001
	(0.037)	(0.069)	(0.041)	(0.010)	(0.067)	(0.009)
Post × Treat. × Var.	-0.015	-0.017*	-0.027	-0.087***	-0.043**	-0.107***
	(0.013)	(0.009)	(0.017)	(0.033)	(0.018)	(0.030)
Observations	1225843	1225843	611375	1236052	1236052	1236052
# of Clusters	195	195	193	195	195	195
Heterogeneity By:	Past Jobs	Past Income	Skill Level	Badge	Hourly Rate	Success Rate

Notes: This table presents OLS regressions relating freelancers' employment on the platform to the introduction of ChatGPT, examining heterogeneity by freelancers' attributes and past experiences. *Treatment* is an indicator for whether an occupation is substantially affected by the introduction of ChatGPT, and *Var.* is the heterogeneity variable of interest, as indicated below each column. The unit of observation is freelancer-month. The dependent variable in Panel A is total monthly number of jobs, and in Panel B, total monthly income. Both dependent variables and heterogeneity variables are inverse hyperbolic sine transformed. All regressions include freelancer and month fixed effects. Standard errors are in parentheses and are clustered at the occupation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: The Effect of Image-based generative AI on Freelancers' Employment on the Platform: Heterogeneity by Freelancer Attributes and Performance

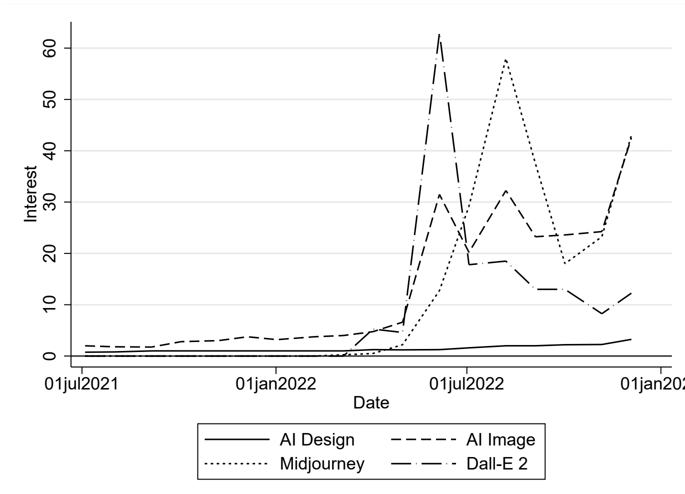
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Monthly Number of Jobs					
Post × T	-0.011	0.054***	0.026	-0.017***	-0.002	-0.008*
	(0.007)	(0.017)	(0.022)	(0.005)	(0.019)	(0.004)
Post × T × Var.	0.004	-0.011***	-0.031***	-0.005	-0.005	-0.025***
	(0.002)	(0.002)	(0.009)	(0.005)	(0.005)	(0.005)
	Panel B: Monthly Income					
Post × T	-0.026	0.171***	0.107	-0.035**	-0.001	-0.006
	(0.025)	(0.058)	(0.083)	(0.014)	(0.071)	(0.010)
Post × T × Var.	0.007	-0.033***	-0.088**	-0.039	-0.014	-0.097***
	(0.011)	(0.008)	(0.040)	(0.033)	(0.019)	(0.037)
Observations	1452788	1452788	704658	1498086	1498086	1498086
# of Clusters	192	192	190	192	192	192
Heterogeneity By:	Past Jobs	Past Income	Skill Level	Badge	Hourly Rate	Success Rate

Notes: This table presents OLS regressions relating freelancers' employment on the platform to the introduction of DALL-E 2 and Midjourney, examining heterogeneity by freelancers' attributes and past experiences. *Treatment* is an indicator for whether an occupation is substantially affected by the introduction of image-based AI, and *Var.* is the heterogeneity variable of interest, as indicated below each column. The unit of observation is the freelancer-month. The dependent variable in Panel A is total monthly number of jobs, and in Panel B, total monthly income. Both dependent variables and heterogeneity variable are inverse hyperbolic sine transformed. All regressions include freelancer and month fixed effects. Standard errors are in parentheses and are clustered at the occupation-level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Online Appendix

A Additional Figures and Tables

Figure A1: Descriptive Evidence on Interest in Image-based Generative AI Over Time



Notes: This figure displays the relative interest in four separate search terms on Google Trends in the period around the introduction of DALL-E 2 and Midjourney.

Table A1: The Effect of Text-based Generative AI on Freelancers' Employment on the Platform: Robustness to Alternative Specifications

	Total		Extensive Margin	Intensive Margin	
	(1) Num. of Jobs	(2) Income	(3) Worked	(4) Num. of Jobs	(5) Income
	Panel A: Without control Variables				
Post × Treatment	-0.015***	-0.035	-0.009***	-0.047***	-0.053
	(0.006)	(0.023)	(0.003)	(0.009)	(0.033)
Observations	1236052	1236052	1236052	171541	174697
# of Clusters	195	195	195	193	193
Adjusted R^2	0.34	0.29	0.27	0.28	0.45
	Panel B: With Wild Bootstrap				
Post × Treatment	-0.020***	-0.052***	-0.012***	-0.047***	-0.051
	[-5.53]	[-3.25]	[-5.19]	[-4.36]	[-1.53]
Observations	1236052	1236052	1236052	171541	174697
# of Clusters	195	195	195	193	193
Adjusted R^2	0.41	0.34	0.32	0.30	0.46
	Panel C: Clustered at the Industry-level				
Post × Treatment	-0.020***	-0.052**	-0.012***	-0.047***	-0.051
	(0.004)	(0.022)	(0.003)	(0.011)	(0.046)
Observations	1236052	1236052	1236052	171541	174697
# of Clusters	59	59	59	59	59
Adjusted R^2	0.41	0.34	0.32	0.30	0.46

Notes: This table presents OLS regressions relating freelancers' employment on the platform to the introduction of ChatGPT in a difference-in-differences design (see description around Equation 1). *Treatment* is an indicator for whether an occupation is substantially affected by the introduction of ChatGPT. The unit of observation is the freelancer-month. The dependent variables are the total monthly number of jobs and income (columns 1 and 2), an indicator for whether the freelancer had at least one job that month (Column 3), the monthly number of jobs, and income conditional on working (columns 4 and 5). In columns 1, 2, 4 and 5, the dependent variables are inverse hyperbolic sine transformed. All regressions include freelancer and month fixed effects. In Panel A we do not control for time-varying occupation-level characteristics. In Panel B, the t-statistic using wild bootstrap (Cameron et al., 2008) are in brackets. Standard errors are in parentheses and are clustered at the occupation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A2: The Effect of Image-based Generative AI on Freelancers' Employment on the Platform: Robustness to Alternative Specifications

	Total		Extensive Margin	Intensive Margin	
	(1)	(2)	(3)	(4)	(5)
	Num. of Jobs	Income	Worked	Num. of Jobs	Income
	Panel A: Without control Variables				
Post × Treatment	-0.032***	-0.022	-0.008***	-0.055***	-0.020
	(0.006)	(0.024)	(0.003)	(0.012)	(0.018)
Observations	1498086	1498086	1498086	326468	292498
# of Clusters	192	192	192	192	191
Adjusted R^2	0.59	0.28	0.35	0.54	0.46
	Panel B: With Wild Bootstrap				
Post × Treatment	-0.021***	-0.052***	-0.014***	-0.018*	-0.028
	[-4.62]	[-2.71]	[-4.83]	-[-1.94]	[-1.50]
Observations	1498086	1498086	1498086	326468	292498
# of Clusters	192	192	192	192	191
Adjusted R^2	0.66	0.32	0.38	0.65	0.46
	Panel C: Clustered at the Industry-level				
Post × Treatment	-0.021***	-0.052**	-0.014***	-0.018***	-0.028
	(0.004)	(0.022)	(0.002)	(0.005)	(0.017)
Observations	1498086	1498086	1498086	326468	292498
# of Clusters	60	60	60	60	60
Adjusted R^2	0.66	0.32	0.38	0.65	0.46

Notes: This table presents OLS regressions relating freelancers' employment on the platform to the introduction of DALL-E 2 and Midjourney in a difference-in-differences design (see description around Equation 1). *Treatment* is an indicator for whether an occupation is substantially affected by the introduction of ChatGPT. The unit of observation is the freelancer-month. The dependent variables are the total monthly number of jobs and income (columns 1 and 2), an indicator for whether the freelancer had at least one job that month (column 3), the monthly number of jobs, and income conditional on working (columns 4 and 5). In columns 1, 2, 4 and 5, the dependent variables are inverse hyperbolic sine transformed. All regressions include freelancer and month fixed effects. In Panel A we do not control for time-varying occupation-level characteristics. In Panel B, the t-statistic using wild bootstrap (Cameron et al., 2008) are in brackets. Standard errors are in parentheses and are clustered at the occupation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3: The Effect of Generative AI on Freelancers' Employment on the Platform: Robustness to Alternative Empirical Specifications and Transformations of the Dependent Variable

	(1) Worked	(2) Num. of Jobs	(3) Num. of Jobs	(4) Num. of Jobs	(5) Income
	Panel A: Text-based Generative AI				
Post × Treatment	-0.196***	-0.154***	-0.149***	-0.015***	-0.043***
	(0.039)	(0.060)	(0.016)	(0.003)	(0.015)
Observations	793474	1236052	797348	1236052	1236052
# of Clusters	193	214	-	195	195
Adj./ Pseudo R^2	0.13	-	-	0.41	0.34
Specification	Logit	Poisson	Neg. Bin.	Log + 1	Log + 1
	Panel B: Image-based Generative AI				
Post × Treatment	-0.109***	-0.053	-0.028***	-0.016***	-0.042**
	(0.032)	(0.314)	(0.010)	(0.004)	(0.018)
Observations	1489768	1498596	1498086	1498086	1498086
# of Clusters	192	214	-	192	192
Adj./ Pseudo R^2	0.29	-	-	0.68	0.32
Specification	Logit	Poisson	Neg. Bin.	Log + 1	Log + 1

Notes: This table presents ten regressions results relating freelancers' employment on the platform to the introduction of ChatGPT in a difference-in-differences design (see description around Equation 1). *Treatment* is an indicator for whether an occupation is substantially affected by the introduction of ChatGPT. The unit of observation is the freelancer-month. The dependent variables are described in column headers, and the empirical specifications are presented on the last row of each panel. All regressions include freelancer fixed effects. Standard errors are in parentheses and are clustered at the occupation level, except for the negative binomial regression. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A4: The Effect of Generative AI on Freelancers' Employment on the Platform: Matched Sample

	Total		Extensive Margin	Intensive Margin	
	(1) Num. of Jobs	(2) Income	(3) Worked	(4) Num. of Jobs	(5) Income
	Panel A: Text-based Generative AI				
Post × Treatment	-0.022***	-0.074***	-0.012***	-0.047**	-0.083**
	(0.004)	(0.015)	(0.002)	(0.019)	(0.032)
Observations	1304312	1304312	1304312	225979	188564
# of Clusters	195	195	195	193	193
Adjusted R^2	0.77	0.32	0.46	0.74	0.46
	Panel B: Image-based Generative AI				
Post × Treatment	-0.019***	-0.048***	-0.010***	-0.035*	-0.070**
	(0.004)	(0.015)	(0.002)	(0.020)	(0.034)
Observations	1300269	1300269	1300269	223510	186097
# of Clusters	195	195	195	193	193
Adjusted R^2	0.78	0.32	0.45	0.75	0.45

Notes: This table presents OLS regressions relating freelancers' employment on the platform to the introduction of text-based (Panel A) and image-based (Panel B) AI, in a difference-in-differences design (see description around Equation 1). *Treatment* is an indicator for whether an occupation is substantially affected by the introduction of ChatGPT. Results are based on the matched sample of freelancers, where we match coarsely on the freelancers' past employment, past income, education, badges, success rate, and hourly rate on the platform. The unit of observation is the freelancer-month. The dependent variables are the total monthly number of jobs and income (columns 1 and 2), an indicator for whether the freelancer had at least one job that month (column 3), the monthly number of jobs, and income conditional on working (columns and 5). In columns 1, 2, 4, and 5, the dependent variables are inverse hyperbolic sine transformed. All regressions include freelancer and month fixed effects. Standard errors are in parentheses and are clustered at the occupation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A5: The Effect of Generative AI on Freelancers' Employment on the Platform: Using the Prior Year as Control

	Total		Extensive Margin	Intensive Margin	
	(1) Num. of Jobs	(2) Income	(3) Worked	(4) Num. of Jobs	(5) Income
	Panel A: Text-based Generative AI				
Post × Treatment	-0.015***	-0.095***	-0.003	-0.022**	-0.317***
	(0.004)	(0.018)	(0.003)	(0.010)	(0.030)
Observations	251052	251052	251052	26396	43939
# of Clusters	11746	11746	11746	7183	8189
Adjusted R^2	0.39	0.32	0.43	0.10	0.42
	Panel B: Image-based Generative AI				
Post × Treatment	-0.043***	-0.170***	-0.035***	-0.014	-0.093***
	(0.005)	(0.020)	(0.003)	(0.009)	(0.022)
Observations	252575	252575	252575	38334	64367
# of Clusters	12889	12889	12889	9351	10316
Adjusted R^2	0.37	0.35	0.41	0.11	0.49

Notes: This table presents OLS regressions relating freelancers' employment on the platform to the introduction of text-based (Panel A) and image-based (Panel B) AI, in a difference-in-differences design (see description around Equation 1). Only ultimately treated occupations are included in the analysis. *Treatment* is an indicator comparing freelancers' monthly outcomes in a given month to their performance in the same month the previous year. The unit of observation is the freelancer-month. The dependent variables are the total monthly number of jobs and income (columns 1 and 2), an indicator for whether the freelancer had at least one job that month (column 3), the monthly number of jobs, and income conditional on working (columns and 5). In columns 1, 2, 4, and 5, the dependent variables are inverse hyperbolic sine transformed. All regressions include freelancer and month fixed effects. Standard errors are in parentheses and are clustered at the freelancer level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Data

Our main data source is Upwork, a leading international platform for freelancers. We have used three Upwork APIs to create our panel. First, we queried the search API to collect a full list of freelancers for each occupation. Second, according to the ID assigned by the platform to each freelancer, we used freelancer API to collect the freelancer's profile page with information that includes a self-introduction, summary statistics about the freelancer's earnings and total hours spent on jobs, rating badges by Upwork, and employment history. The employment history data for each freelancer include job IDs, compensation, length of employment, hours worked, job titles, and employer feedback. Third, to complement employment history information with ex ante information, we further collected job posting data via Upwork's job posting API to obtain the following information: employer identifier, job description, required skills, expected occupation, expected hours or wage, and expected employee level.

B.1 Freelancer Collection

Upwork has a three-layer classification system of occupations, and we followed the finest level, in total 215 occupations, to collect a list of freelancers. Some freelancers work in multiple occupations, most of which are in the same industry. Freelancers may set their profiles to be private; our methodology does not allow us to collect data on these freelancers, because their profiles are not searchable. We obtained a panel of 568,647 freelancers across 215 occupations.

B.2 Freelancer Profile and Employment History Collection

For each freelancer ID, we queried the freelancer API to retrieve information on the freelancer's homepage, where the freelancers display self-introductions and list their skills. Upwork reports job success rates and assigns tags in the form of badges on the freelancer's homepage to indicate that freelancer's excellence. Upwork also provides a summary of each freelancer's total hours, total earnings, and total number of jobs on the platform. All the above information is as of the collection date. In addition, we also collected the employment history of each freelancer with job ID, job title, job description, feedback, total hours, and earnings. We kept only freelancers who work in a single layer-two occupation (i.e., subindustry). We also removed freelancers whose first job compensation exceeds 500% or is below 5% of the average compensation of the first five jobs. These two thresholds were determined based on 10th and 90th percentiles of the empirical distribution. In total, our sample has approximately 80,000 freelancers with at least one job completed on the platform.

B.3 Job Posting Collection

For each employment record, we further supplemented it using job posting API to get more information about job postings, including employers' expectations on hours, compensation, employee skills, and level of expertise. About half the job postings are private, meaning that we do not have access to additional information about these jobs.