Dec 11th, 12:00 AM

The Impact of ChatGPT on the Demand for Human Content Generating and Editing Services: Evidence from an Online Labor Market

Ziqing Yuan
University of Hong Kong, zqyuan2-c@my.cityu.edu.hk

Hailiang Chen
The University of Hong Kong, chen19@hku.hk

ICIS 2023 Hyderabad, The Impact of ChatGPT on the Demand for Human Content Generating and Editing Services: Evidence from an Online Labor Market

Completed Research Paper

Ziqing Yuan
The University of Hong Kong
Pokfulam, Hong Kong
zqyuan2@hku.hk

Hailiang Chen
The University of Hong Kong
Pokfulam, Hong Kong
chen19@hku.hk

Abstract

The rise of generative AI has been a subject of debate among researchers and practitioners regarding its effect on the labor market. While some argue that it may displace jobs, others suggest it could create new opportunities and improve productivity. This study examines the impact of the ChatGPT launch on 18,130 services with 199,430 observations using a difference-in-differences approach and data from the online labor marketplace Fiverr. The findings suggest that ChatGPT had a negative effect on the demand for human content generating and editing services, with a concentration on writing services. However, there was no significant effect on the demand for editing services. The study also found that the demand for services with higher prices was more negatively affected. These results contribute to the ongoing debate on the impact of generative AI on the labor market and offer practical recommendations for service providers to navigate this new AI-driven landscape.

Keywords: Generative AI, human labor, ChatGPT, demand and supply

Introduction

The remarkable progress in natural language processing (NLP) made possible by the recent proliferation of generative artificial intelligence (AI), such as ChatGPT, has led to significant advancements in content generation and editing (McKinsey 2023). Despite the numerous potential benefits of technological advancement, they also raise concerns about the future demand for human writers and editors. Some experts posit that generative AI may lead to a reduction in demand for human labor, as businesses and individuals increasingly rely on automated solutions for their tasks (Weber and Ellis 2023). However, other experts argue that instead of replacing human writers and editors, generative AI can augment their writing skills and capabilities (Forbe 2023; Thorp 2023). In turn, writers and editors enhanced with generative AI become more productive in content creation. For instance, human writers and editors can start with a rough draft created by generative AI, and then manually refine it to meet specific needs and standards. In this way, AI-generated content could potentially boost the demand for human content generation and editing services, as businesses and individuals seek to combine the speed and efficiency of AI with the creativity and expertise of human writers and editors.

Generative AI is a subfield of artificial intelligence encompassing techniques and models designed to generate novel content such as images, text, music, and more. The fundamental concept behind generative AI entails learning patterns from existing data and utilizing this knowledge to create original content that adheres to the learned patterns. This approach allows AI systems to unleash their creativity and produce innovative outputs based on the training data, a quality that has long been viewed as a unique advantage exclusive to human labor. However, developing a proficient generative AI system often requires extensive training on large datasets to ensure its ability to perform general functions and achieve specific tasks. The significant barriers to entry in constructing a generative AI system limit its widespread influence on economic factors. Previous research may have focused on the adoption of specific generative AI systems, such as machine learning or large language models, within individual firms, examining the potential effects on productivity and overall employment. (Xue et al. 2022). The launch of ChatGPT and its immense success have significantly increased the adoption of generative AI, raising concerns on the relationship between human labor and AI systems at a more granular level. As a result, our study aims to address this gap by investigating the potential impact of ChatGPT launch on demand for human labor at the service level and exploring potential heterogeneity across various service types and prices.

This study focuses on the natural experiment presented by the launch of ChatGPT, where GPT stands for generative pre-trained transformer. On November 30, 2022, OpenAI released ChatGPT to the public for free, and it quickly gained a large user base. To estimate the impact of the ChatGPT launch on demand for human writers and editors, we obtained data on human content generating and editing services from Fiverr, a global online labor marketplace. The services in the Writing & Translation category on Fiverr are considered as the treatment group, and we selected services that may not be affected by ChatGPT or other generative AI as the control group for comparison. The demand for services is quantified based on the volume of reviews received over a period of months. Our sample consists of 18,130 services with a total of 199,430 service-month observations. Of these, 9,982 services were in the treatment group and affected by the ChatGPT launch while 8,148 services were in the control group. The dataset covers a period of 7 months before and 4 months after the ChatGPT launch.

In this study, we used a difference-in-differences (DID) approach to analyze the impact of the release of ChatGPT on the demand for human content generation and editing services. The difference-in-differences (DID) model is a widely used method in empirical research for estimating causal effects of treatments, which compares the changes in an outcome variable between two groups over two time periods, before and after treatment. Additionally, we conducted further analyses to explore the heterogeneous effects of the ChatGPT launch across service types and prices. Our findings indicate that the release of ChatGPT had a negative effect on the demand for these services. We also examined the dynamic effect of ChatGPT on labor demand by analyzing the changes in the coefficients of post-event interaction terms over time. The result shows that the negative impact of ChatGPT on demand for these services worsened over the months. It indicates that the demand for human content generating and editing services is severely and negatively affected by the automation of generative AI.

Interestingly, our subsample analysis revealed that the negative impact of the ChatGPT launch was concentrated on writing services, with no significant effect on demands for editing services. This suggests that ChatGPT may be better suited for automating tasks that involve generating new content, rather than editing existing content. We also found that the price of the service significantly moderates the impact of the ChatGPT launch on the demand for human writing and editing services, and the higher-priced services were associated with a more negative impact on demand. This finding suggests that higher-priced services may be more vulnerable to displacement by generative AI than lower-priced services.

The remainder of this paper is structured as follows. We first discuss the theoretical background that motivates and supports our study. Then we describe empirical analysis by introducing our dataset and the research design. The findings are presented next. The last section discusses the implications for both research and practice.

Theoretical background

The occupational impact of technological advancements has been a topic of significant interest in recent years, with scholars examining both automation and augmentation effects on the labor market (Acemoglu and Restrepo 2022; Agrawal et al. 2019; Brynjolfsson and Mitchell 2017; Dixon et al. 2021; Hoynes and Rothstein 2019; Van den Broek et al. 2021; Xue et al. 2022). On one hand, technological advancements can enhance workers' skills, thereby improving productivity and increasing labor demand. For example, Xue et al. (2022) studied the occupational impact of AI application in China from 2007 to 2018 and found a positive association between AI adoption and overall employment. Similarly, Dixon et al. (2021) reported positive employment effects of AI adoption at the firm level. On the other hand, new technologies can replace human workers in various jobs, even those requiring advanced cognitive skills. Sturm et al. (2021) demonstrated that investments in machine learning reduced an organization's demand for human explorative learning, and Acemoglu et al. (2020) found a negative impact of robot adoption on industry employment. Notably, these two effects are not mutually exclusive; automation can lower the cost of human labor, discouraging further automation while encouraging the creation of new tasks (Acemoglu and Restrepo 2018).

The mixed evidence highlights the complex relationship between technological advancements and human labor, which is influenced by multiple factors such as the education level of labor (skill-biased technological advancement) (Autor et al. 1998; Zhang et al. 2023) and the routineness of job tasks (routine-biased technological advancement) (Autor et al. 2003; Zhang et al. 2023). Research indicates that automation technologies increase demand for high-education workers while displacing low-education labor workers (Autor et al. 1998). Additionally, technology has been found to reduce labor input for routine tasks and enhance input for nonroutine cognitive tasks (Autor et al. 2003). As tasks become more programmable, middle-education workers risk being replaced by technologies, such as robotics (Acemoglu and Restrepo 2022). Lysyakov and Viswanathan (2022) examined designers' responses to the introduction of an AI logo design system on a crowdsourcing platform, noting that high-skilled designers can avoid competition with AI by delivering more sophisticated designs. In recent years, the technology deskilling effect has also been observed, with Xue et al. (2022) finding a positive association between AI applications and the employment of low-educated workers without college degrees at the firm level. Autor et al. (2022) discovered an increase in demand for low-skilled jobs, but not for high-skilled or middle-skilled positions. Zhang et al. (2023) demonstrated that the relationship between IT and labor varies based on education levels and industry characteristics; IT generally complements high-education labor and substitutes for low-education labor. For middle-education labor, the relationship is contingent upon the industry's routine intensity and AI exposure.

As large language models and natural language processing continue to advance, the impact of AI adoption on the labor market has garnered significant attention in recent years (Acemoglu and Restrepo 2022; Agrawal et al. 2019; Brynjolfsson and Mitchell 2017; Dixon et al. 2021; Hoynes and Rothstein 2019; Van den Broek et al. 2021; Xue et al. 2022). Most studies have concentrated on the adoption of AI or robotics in the workplace, investigating their effects on potential employee displacement (Dixon et al. 2021; Xue et al. 2022), employee attitudes and behaviors (Li et al. 2019; Tong et al. 2021), customer behaviors (Wang et al. 2023), and the characteristics of jobs most susceptible to automation through AI techniques (Bresnahan et al. 2002). Generative AI, which differs from traditional AI systems that focus on tasks such as pattern detection, decision-making, data classification, and fraud detection, exhibits the ability to create new content like text, images, music, and videos. Historically, the adoption of these systems has been limited to large companies due to high implementation and maintenance costs. However, with the release of ChatGPT, the barriers to entry for using generative AI have been reduced, making it more accessible and affordable to a broader range of users. The growing adoption of ChatGPT presents the potential to introduce new forms of competition and fundamentally transform the way content-generating businesses operate. Given the complex relationship between technological advancements and human labor, it is essential to examine the impact of generative AI on the demand for human labor and explore the potential heterogeneity.

Effects of ChatGPT launch on the Labor Market

To comprehend the potential effects of ChatGPT launch on the labor market, we employ the theoretical lens of the elasticity of substitution and complementarity. This concept posits that two input factors for creating a good or service are considered complements (substitutes) if an increase in the input of one factor results in an increase (decrease) in the demand of another factor. To investigate the effects of ChatGPT launch on the labor market, it is essential to examine the relationship between two input factors: human labor and ChatGPT.

On the one hand, the increased accessibility of ChatGPT may lead to the displacement of human labor in content generation services. ChatGPT has the capability to produce high-quality, human-like content at a significantly faster pace and lower cost compared to human labor (McKinsey 2023). Assuming all other factors remain constant, businesses and individuals seeking to save time and minimize costs may increasingly rely on AI-generated content, consequently leading to reduced demand for human services, in accordance with the law of demand. Furthermore, the negative effects of generative AI on the labor market are likely to become more pronounced as the technology advances rapidly, enabling it to automate a wider range of writing and editing tasks across various fields. Consequently, businesses may find it increasingly cost-effective to employ ChatGPT as an alternative to human labor.

On the other hand, the assumption that all other factors remain constant may not accurately represent real-world scenarios. AI-generated content might not outperform or equal human-created content in terms of quality, particularly when input data quality is low, due to generative AI's dependence on data input. Under such circumstances, businesses may still necessitate human involvement to ensure content accuracy, grammatical correctness, and adherence to ethical or legal standards. Furthermore, technological advancements can contribute to human capital accumulation, as human workers can leverage generative AI tools, such as ChatGPT, to augment their productivity. For instance, human writers might use AI-generated content as a starting point or rough draft, refining and editing it to meet specific needs and standards. In this scenario, ChatGPT could expedite the process of translating ideas into text, allowing human labor to focus on other tasks, such as proofreading or ideas planning. This potential synergy could benefit human workers, as businesses and individuals seek to combine AI's speed and efficiency with human expertise in writing and editing. However, it is crucial to acknowledge that the adoption of ChatGPT may diminish inequality among human workers (Noy and Zhang 2023) and intensify competition between them. Consequently, the launch of ChatGPT could potentially boost demand for human labor only if the demand for content generation demand surpasses the expansion in labor supply. As these effects are not mutually exclusive, it is imperative to investigate the effects of ChatGPT launch on the labor market and explore niches where human expertise continues to be in demand.

Service Types

The potential for AI-generated content to surpass or equal human-created content in terms of quality is a critical factor in examining the interplay between generative AI and human labor. This relationship may be contingent upon the tasks involved and the inherent capabilities and constraints of the technology. One of the strengths of generative AI is its high level of accuracy in grammar and wording, efficiency in content generation, and ability to produce novel ideas (Davenport and Mittal 2022). However, generative AI has some limitations as well. For instance, it is sensitive to the prompts fed into it and may require human involvement in trying alternative prompts before settling on the content. Additionally, AI-generated content is advised to be edited and evaluated carefully, as generative AI models are known to hallucinate and produce content that may not be factually accurate; they may also lack the ability to fully understand the nuances of contextual knowledge and expertise necessary to ensure that the information generated is credible.

Writing and editing are two crucial tasks of human content generation, and they require different skills and have distinct focuses. Writing typically involves the creation of new content, such as articles, blog posts, or product descriptions. It requires a high level of creativity, originality, and expertise in a particular subject matter. Writers must be able to develop compelling ideas, craft engaging narratives, and deliver information in a way that is both informative and entertaining. On the other hand, editing is the process of refining existing content. It involves ensuring that the content is factually accurate and grammatically correct and meets certain standards. Editors must have an eye for detail and be able to identify errors in grammar, punctuation, spelling, and syntax. They must also be able to provide constructive feedback to writers in order to improve the overall quality of the content. In sum, writing requires more creativity and originality, while editing requires attention to detail and the ability to refine existing content to meet specific standards. In this regard, the release of generative AI may have a more detrimental effect on writing services than on editing services, given its proficiency in creativity and originality but the limitation in ensuring the accuracy and credibility of generated information.

Service Price

Service price is a significant factor in studying the impact of generative AI on the labor market because it affects the demand for their services. However, it is ad-hoc unclear how the price moderates the impact of generative AI on demand for content generation services. On the one hand, low-priced services tend to be more standardized and repetitive in nature, making them more susceptible to automation. Generative AI can quickly generate content that meets basic requirements for these tasks, such as simple product descriptions or blog posts, at a fraction of the cost of hiring human workers. In contrast, higher-priced services often involve more complex and specialized tasks that require a higher level of expertise and a deep understanding of contextual knowledge, which makes it more difficult for generative AI to replicate. As a result, low-priced services may be less vulnerable to displacement by generative AI.

On the other hand, higher-priced services may be more vulnerable to being displaced by generative AI than lower-priced services because businesses and individuals may be more motivated to save costs by automating these services. When the price of human writing and editing services is higher, the cost savings from using generative AI can be more substantial. As a result, businesses and individuals may be more likely to switch to generative AI to reduce costs, leading to a greater negative impact on demand for human services. Conversely, lower-priced services may be less vulnerable to being displaced by generative AI because the cost savings from automation are relatively small. Moreover, lower service prices may help to maintain or increase demand for human services, as buyers may see these services as a more cost-effective alternative to AI-generated content.

Research Methodology

Empirical Context: Release of the ChatGPT

We study the impact of AI-generated content on the online labor market by investigating the consequences of the ChatGPT release on the demand for human content generating and editing services. On November 30, 2022, OpenAI released an advanced artificial intelligence language model, ChatGPT, to the public for free (OpenAI 2022). ChatGPT has proved to be a powerful tool for content creation in a variety of contexts, including traditional article writing, grammar and spelling checking, social media posts, product descriptions, and so on (Davenport and Mittal 2022). It can generate content on a wide range of topics, such as poems, business proposals, social media posts, academic writings, and so on, and can be trained to generate content in a specific style or tone and customized for different audiences. While ChatGPT can efficiently generate high-quality content, it has been criticized for its potential to spread misinformation, and it may require additional human editing or refinement to meet specific needs and standards (Forbe 2022). The launch of ChatGPT was a significant exogenous shock, as it was unlikely anticipated by the related human content creators and editors and garnered a significant number of users immediately following its release, amassing one million users within the first five days of launch. In January 2023, two months after launch, ChatGPT is estimated to have reached 100 million monthly active users.

Data and Sample

The dataset used in this study was collected from Fiverr, an online labor marketplace for freelance services that offer a wide range of services, such as web design, virtual assistance, content writing, and editing. Fiverr operates on a global scale and has registered users from approximately 230 countries and territories. As of December 31, 2022, the platform had 4.3 million active buyers. To use Fiverr, users need to register and create a user profile, including the profile image, country of origin, and languages spoken. Sellers can list their services, which must be placed in one of the platform's predefined categories and subcategories. Fiverr has 10 categories and 243 subcategories for services available.

Buyers can search for services within a subcategory on Fiverr and are presented with a ranked list of up to 960 services. Each page displays 48 services, with a maximum of 20 pages available to browse. Clicking on a service displays detailed information, such as the service image, package, price, reviews, ratings, and more. A screenshot of a service page on Fiverr is presented in the appendix a. Buyers can leave reviews within 10 days after the service's order is completed, and they cannot add new reviews after that time. Reviews cannot be removed unless they violate the platform's standards. It's worth noting that Fiverr's terms of service prohibit sellers from promoting their services on other third-party advertising platforms.

Our unit of analysis is at the service-month level. As the historical order information is not public on Fiverr, we use the number of reviews as a proxy of consumer demand, consistent with prior studies (Foerderer et al. 2018; Ye et al. 2011). Using the number of reviews is appropriate for our study for three reasons. Firstly, a review can only be given after the order is completed. Secondly, reviews cannot be removed by the seller. Thirdly, buyers cannot add new reviews for past orders. Since only reviews published within the past year on Fiverr are at the monthly level, our study period covers 4 months after and 7 months before the event, i.e., April 2022 to March 2023.

This study investigates how generative AI affects the labor market, specifically in human content generation and editing services. ChatGPT launch is used as an exogenous shock to compare demand for affected services with those not affected by generative AI models. To define the affected and unaffected services, the study identifies 33 subcategories of Writing & Translation on Fiverr as impacted by ChatGPT and selects 33 other subcategories involving large human involvement (e.g., modeling, acting, or coaching), physical handcrafts (e.g., creating or painting miniatures), and sensitive data and privacy (e.g., filing tax or financial consulting) as not impacted by AI-generated content tools. The full list of these 66 subcategories is included in the appendix b and c. Due to the limitations of the Fiverr platform, which only displays a maximum of 960 services within 20 pages for each subcategory, we collected up to 960 services for each subcategory, resulting in a preliminary sample of 32,980 services. To ensure the reliability of our sample, we only included services published before April 2022, and we excluded services without any reviews during our study period. The dataset used for analysis includes 18,130 services, out of which 9,982 are treated and 8,148 are control services. For each service, we have collected the following data. Table 1 reports the summary statistics for 18,130 services over our study period, i.e., April 2022 to March 2023.

- $Review_{i,t}$. The number of reviews of services *i* at time *t*.
- $Treat_i$. Binary variable, which is 1 if the service *i* is in the treated service subcategories, and 0 otherwise.
- $After_t$. Binary variable indicating whether the ChatGPT is released at time *t*.
- $Price_i$. The starting price of service *i*.

Variable	#Obs.	Mean	Std. Dev.	Min	Median	Max
$Log(Review_{i,t})$	199,430	0.647	0.802	0	0	4.564
$After_t$	199,430	0.364	0.481	0	0	1
$Treat_i$	199,430	0.551	0.497	0	1	1
$Log(Price_{i,t})$	199,430	5.217	1.171	3.738	5.106	11.151

Table 1. Summary Statistics

Model-free Evidence

The effect of the ChatGPT launch on the treatment and control groups can be observed through model-free evidence, specifically by examining the trend of the number of reviews over months for both groups. Figure 1 displays the number of reviews for the treatment and control groups from April 2022 to March 2023, with the shaded area representing the period during which the treatment was effective.

As shown in Figure 1, there is a clear difference in the trend of the number of reviews between the treatment and control groups after the release of ChatGPT. Prior to the release of ChatGPT, the number of reviews per month for both groups was parallel, with the gap between the two groups remaining stable. However, following the release of ChatGPT, the number of reviews for the treatment group decreased, while the trend for the control group remained relatively stable. This suggests that the release of ChatGPT had a negative impact on the treatment group, leading to a decrease in demand for their services as indicated by the fewer reviews left by buyers. In contrast, the control group appeared to be unaffected by the release of ChatGPT. In summary, the model-free evidence provided by Figure 1 indicates that the ChatGPT launch had a negative impact on the treatment group, while demand for services in the control group remained relatively stable.

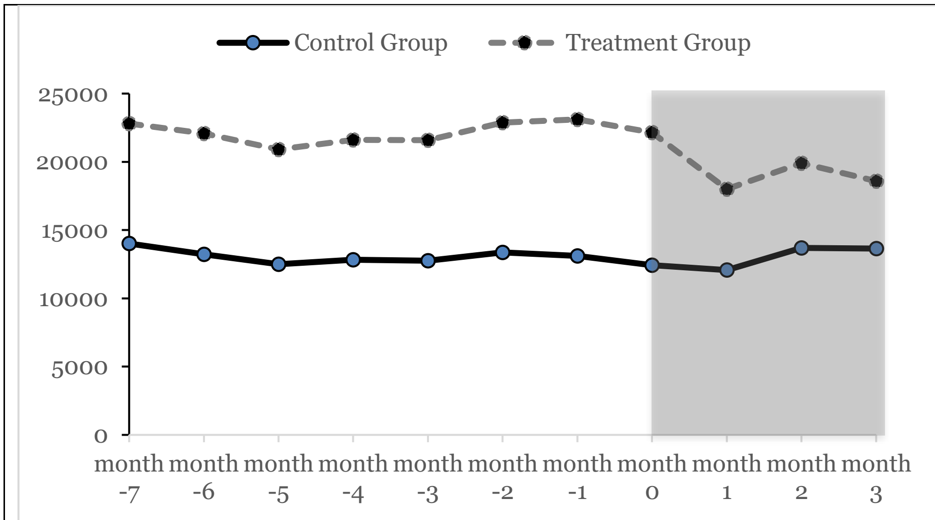


Figure 1. Model-free Evidence: Number of reviews over Months

Difference-in-differences Model

The main analysis employed a difference-in-differences (DID) model, which compares the changes in an outcome variable between two groups over two time periods, before and after treatment. The basic idea behind DID is that if the treatment has no effect, the differences observed in the two periods should be statistically the same. However, if the treatment has a significant effect, a statistically significant difference, either positive or negative, would be observed. To estimate the effect of the ChatGPT launch on the demand for human content generating and editing services, we use the following specification:

$$Log(Review_{i,t}) = Treat_i \times After_t + \mathrm{X}_i + \mathrm{T}_t + \varepsilon_{it} \quad (1)$$

Where $Log(Review_{i,t})$ is the natural logarithm of the number of reviews of service i in month t. $Treat_i \times After_t$ is the main variable of interest. Its coefficient estimate reveals the effect of the ChatGPT launch. X_i is the service fixed effect, which controls for time-invariant service characteristics; T_t is the time fixed effect, which accounts for seasonality and other time shocks; and ε_{it} is the error term. $Treat_i$ and $After_t$ are not listed in the model because they are absorbed into the service and month fixed effects. We cluster heteroskedasticity-robust standard errors at the service level to account for potential correlations.

The Impact of ChatGPT Launch

We present the results of the difference-in-differences (DID) model estimating the effect of the ChatGPT launch on demand for human content generating and editing services. Table 2 shows the regression results using the specification outlined in Equation (1). The coefficient of interest, $After_t * Treat_i$, which measures the effect of ChatGPT launch on demand, is statistically significant and negative (-0.047, p<0.01). Specifically, the launch of ChatGPT led to a decrease of approximately 4.59% ($e^{-0.047} - 1$) in reviews for these services, suggesting that the ChatGPT launch had a negative impact on the demand for human content generating and editing services.

	(1)
Variable	$Log(Review_{i,t})$
$\boldsymbol{After_t * Treat_i}$	**-0.047** (0.006)**
Constant	0.695** (0.004)

Service Fixed Effects	Included
Month Fixed Effects	Included
R-squared	0.0087
#Obs.	199,430

Table 2. Main Analysis

Notes: Standard errors are reported in parentheses.
* and ** denote significance at 5% and 1% levels, respectively.

Editing vs. Writing Services

To better understand the effects of ChatGPT, we conducted a subsample analysis on two specific subcategories of services: editing services and writing services. The rationale behind this analysis is to examine whether the effect of the ChatGPT launch varies across different types of services within the content generation and editing industry. To conduct the subsample analysis, we first selected two subcategories of services: editing services and writing services. Editing services include services related to proofreading, editing, and formatting of documents, while writing services include services related to the creation of original written content such as articles, blog posts, and academic papers. To categorize services, we rely on keyword search. In particular, we search the titles of all the services in the treatment group to determine whether they include the keyword "edit" or not. If the title of a service contains editing, edit, or edited, then we classify the service as editing services, otherwise the service in the treatment group is a writing service.

Table 3 presents the results of the subsample analysis. The first column reports the coefficient estimate of $Treat_i \times After_t$ for the writing services subsample, while the second column reports the coefficient estimate for the editing services subsample. The coefficient estimate of $Treat_i \times After_t$ for the writing services subsample is negative and statistically significant (-0.054, p<0.01), indicating that the launch of ChatGPT had led to a decrease of approximately 5.26% ($e^{-0.054} - 1$) in reviews for writing services. In contrast, the coefficient estimate for the editing services subsample is positive but statistically insignificant, indicating that the launch of ChatGPT had no significant effect on demand for editing services. To validate the robustness of this finding, we also conducted the subsample analysis on 33 subcategories of human writing and editing services. The results, as shown in the appendix c, also support the finding, as the release of ChatGPT does not significantly affect the demand for services in the Booking Editing and Proofreading & Editing. A caveat is that there can be both writing and editing services under a specific subcategory. Therefore, the results in the appendix c should be interpreted with caution.

DV = $Log(Review_{i,t})$	Writing	Editing
Variable	(1)	(2)
$After_t * Treat_i$	**-0.054** (0.007)**	**0.011 (0.016)**
Constant	0.682** (0.004)	0.653** (0.005)
Service Fixed Effects	Included	Included
Month Fixed Effects	Included	Included
R-squared	0.0061	0.0232
#Obs.	187,352	101,706

Table 3. Subsample Analysis – Service's Category

Notes: Standard errors are reported in parentheses.
* and ** denote significance at 5% and 1% levels, respectively.

Price of the Service

To explore the moderating impact of service price on the effects of ChatGPT on demand for human content editing and writing services, we conducted subsample analyses for both high- and low-priced services by dividing the sample based on the median price value. Additionally, we used a three-way difference-in-differences (DID) analysis to confirm our findings. This method compares the differences in an outcome variable between two groups during two time periods, before and after treatment, while considering a third variable that moderates the treatment effect. In our case, the third variable is service price. Our regression model is specified as follows:

$$Log(Review_{i,t}) = Treat_i \times After_t + Treat_i \times After_t \times Log(Price_{i,t}) + \mathrm{X}_i + \mathrm{T}_t + \varepsilon_{it} \qquad (2)$$

Where $Log(Price_{i,t})$ is the natural logarithm of the price of service i. The coefficient estimate of the main variable of interest, $Treat_i \times After_t \times Log(Price_{i,t})$, reveals the moderating effect of price on the treatment effect.

The moderating effect of price was estimated and the results are presented in Table 4. Column (1) and (2) show the outcomes of the subsample analysis on high and low-priced services, respectively. Column (3) presents the result of the three-way difference-in-differences (DID) model. In Column (1), the coefficient estimate of $Treat_i \times After_t$ is -0.034 and statistically significant at the 5% level, indicating the ChatGPT launch has led to a decrease of 3.34% on demand for low-priced services. For high-priced services, the coefficient estimate of $Treat_i \times After_t$ is -0.058 and statistically significant at the 1% level in Column (2), indicating the demand for high-price services drops 5.64% after the ChatGPT launch. These results suggest that the negative impact of ChatGPT launch on demand is more pronounced for higher-priced services. Moreover, the result of the three-way difference-in-differences (DID) model in Column (3) provides consistent evidence that the coefficient estimate of $Treat_i \times After_t \times Log(Price_{i,t})$ is negative and statistically significant (-0.027, p<0.01). This indicates that higher service prices are associated with a more negative impact of ChatGPT launch on demand for the service.

DV = $Log(Review_{i,t})$	Low Price	High Price	Full Sample
Variable	(1)	(2)	(3)
$After_t * Treat_i$	-0.034* (0.009)	-0.058** (0.009)	0.095** (0.020)
$After_t * Treat_i * Log(Price_{i,t})$			-0.027** (0.003)
Constant	0.743** (0.005)	0.637** (0.006)	0.695** (0.004)
Service Fixed Effects	Included	Included	Included
Month Fixed Effects	Included	Included	Included
R-squared	0.0098	0.0105	0.0004
#Obs.	108,647	90,783	199,430

Table 4. Subsample Analysis – Service's Price

Notes: Standard errors are reported in parentheses.
 * and ** denote significance at 5% and 1% levels, respectively.

Falsification Test

To further validate that the decrease in demand for editing and writing services is indeed driven by the impact of the ChatGPT launch, we conducted a falsification test by separating reviews based on the countries of the buyers. We estimated the impact of the ChatGPT launch on demand for buyers in countries where ChatGPT is available and where it is not separately. We hypothesize that the demand from buyers in countries where ChatGPT is not available should remain largely unchanged or be less affected by the

ChatGPT launch. In our sample, there are 44 countries and territories where ChatGPT is not available and 191 countries and territories where it is available according to the list of countries in OpenAI.

We conducted a similar DID analysis as shown in Equation (1), but this time, instead of using $Log(Review_{i,t})$ as the outcome variable, we used $Log(Review\_GPT_{i,t})$ and $Log(Review\_nonGPT_{i,t})$ as the proxy demand of buyers in countries where is available or not, respectively. $Log(Review\_GPT_{i,t})$ represents the number of reviews from buyers in countries where ChatGPT is available, and $Log(Review\_nonGPT_{i,t})$ represents the number of reviews from buyers in countries where ChatGPT is not available.

Table 6 presents the results of this falsification test. Column (1) shows the original DID results using $Log(Review_{i,t})$, while Columns (2) and (3) show the results using $Log(Review\_GPT_{i,t})$ and $Log(Review\_nonGPT_{i,t})$, respectively. In Column (2), the coefficient estimate of $After_t * Treat_i$ is negative and statistically significant at the 1% level, indicating a significant negative impact of the ChatGPT launch on the number of reviews in countries where ChatGPT is available. Moreover, the coefficient estimates of $After_t * Treat_i$ between Columns (1) and (2) are very similar, indicating that the decrease in demand is likely driven by buyers in countries where ChatGPT is available. The results in Column (3) show that although the coefficient estimate of $After_t * Treat_i$ is also negative and statistically significant, the effect size is relatively small (a drop of 0.4% only). This result is consistent with the fact that some users in countries where ChatGPT is not available may still access this service through VPN or other techniques. Overall, the results of this falsification test support the main findings of our analysis that the decrease in demand for editing and writing services is mainly driven by the release of ChatGPT.

	(1)	(2)	(3)
Variable	$Log(Review_{i,t})$	$Log(Review\_GPT_{i,t})$	$Log(Review\_nonGPT_{i,t})$
$After_t * Treat_i$	**-0.047**** **(0.006)**	**-0.046**** **(0.006)**	**-0.004**** **(0.002)**
Constant	0.695** (0.004)	0.679** (0.004)	0.040** (0.001)
Service Fixed Effects	Included	Included	Included
Month Fixed Effects	Included	Included	Included
R-squared	0.0087	0.0080	0.0100
#Obs.	199,430	199,430	199,430

Table 5. Falsification Test

Notes: Standard errors are reported in parentheses.
* and ** denote significance at 5% and 1% levels, respectively.

Robustness Check: Relative Time Model

To assess the validity of the parallel trend assumption in our DID model, we conducted a robustness check using a relative time model. The relative time model is a variant of the DID model. But it excludes $After_t * Treat_i$ and instead includes a series of monthly dummies and the interaction terms between $Treat_i$ and those monthly dummies. The monthly dummies are denoted by $Month_t^j$, where the subscript t indicates the month t and superscript j represents the number of months relative to the event date (November 30, 2022), ranging from -6 to -1 for pre-event observations and from 0 to 3 for post-event observations. This model captures the underlying time trend and estimates the effect of the ChatGPT launch by comparing the relative change in demand for editing and writing services before and after the launch.

Table 6 shows the results of the relative time model. The estimate coefficients of the 6 pre-event interaction terms with $Treat_i$ are all statistically insignificant, indicating no significant change in demand for services in the treatment and control groups before the ChatGPT launch. However, the estimated coefficients of post-event interaction terms between monthly dummies and $Treat_i$ are all statistically significantly negative (except month 0), suggesting a significant negative impact of the ChatGPT launch on demand for editing and writing services. It is well noted that the coefficient estimates of post-event interaction terms exhibit a decreasing trend over time, ranging from -0.047 to -0.077. This suggests that the impact of ChatGPT on demand for editing and writing services becomes increasingly negative as time progresses, implying a dynamic effect. Although the estimated coefficient of the interaction term involving month 0 is not significant, the adoption of ChatGPT in replacing human content editing and writing services may take some time. Overall, these results confirm the findings of our original DID model and show that the negative impact of the ChatGPT launch on demand for editing and writing services is not driven by pre-existing trends in demand.

	(1)
Variable	$Log(Review_{i,t})$
$Month_t^{-6}$	-0.028** (0.007)
$Month_t^{-5}$	-0.056** (0.007)
$Month_t^{-4}$	-0.045** (0.008)
$Month_t^{-3}$	-0.054** (0.008)
$Month_t^{-2}$	-0.033** (0.008)
$Month_t^{-1}$	-0.041** (0.008)
$Month_t^{0}$	-0.060** (0.008)
$Month_t^{1}$	-0.073** (0.008)
$Month_t^{2}$	-0.025** (0.008)
$Month_t^{3}$	-0.031** (0.009)
$Month_t^{-6} * Treat_i$	-0.000 (0.009)
$Month_t^{-5} * Treat_i$	-0.004 (0.010)
$Month_t^{-4} * Treat_i$	-0.001 (0.010)
$Month_t^{-3} * Treat_i$	0.010 (0.011)
$Month_t^{-2} * Treat_i$	0.006 (0.011)
$Month_t^{-1} * Treat_i$	0.022 (0.011)
$Month_t^{0} * Treat_i$	0.012 (0.012)
$Month_t^{1} * Treat_i$	-0.047** (0.012)

$Month_t^2 * Treat_i$	-0.058** (0.012)
$Month_t^3 * Treat_i$	-0.077** (0.012)
Constant	0.695** (0.004)
Service Fixed Effects	Included
Month Fixed Effects	Included
R-squared	0.0087
#Obs.	199,430

Table 6. Parallel Trend Test: Relative Time Model

Notes: Standard errors are reported in parentheses.
* and ** denote significance at 5% and 1% levels, respectively.

Discussion

This paper aims to contribute to the understanding of the impact of AI-generated content on the labor market, focusing on the effect of the ChatGPT launch on demand for writing and editing services. To achieve this objective, we employ a Difference-in-Differences (DID) model using a rich dataset of demand for content generation and editing services before and after the ChatGPT launch. Our analysis shows that the launch of ChatGPT had a negative impact on the demand for human content generation and editing services, with a significant decrease in demand for writing services. Interestingly, the launch of ChatGPT had no significant effect on demands for editing services. Furthermore, our results indicate that higher service prices are associated with a more negative impact of the ChatGPT launch on demand for the service. Finally, we conduct a falsification test to support the main findings of our analysis that the decrease in demand for editing and writing services is mainly driven by the release of ChatGPT.

The subsample analysis suggests that the effect of the ChatGPT launch on demand for content generation and editing services varies across different subcategories of services. First, writing services may be more easily replaced by ChatGPT compared to editing services. Writing involves generating new content from scratch, while editing involves refining and improving existing content. ChatGPT may be better suited to tasks that involve generating new content rather than editing, which requires a deeper understanding of the context and purpose of the content. Also, ChatGPT is accused of generating plausible-sounding but nonsensical content, including referencing a scientific study that does not exist. In this regard, users may question the validity of the AI-generated content and take extra efforts to verify it, which may lower the benefits of replacing human editing services with AI editing services. Although some users may find ChatGPT good for content editing, it is also possible that some users may rely on ChatGPT or similar AI-generated content tools to create new content and hire human labor to edit the AI-generated content. Thus, we did not observe a significant impact on demands for editing services.

Following the launch of ChatGPT, we observed that the demand for specific writing services, including "Ad Copy", "Case Studies", "Job Descriptions", "Podcast Writing", "Other", "Research & Summaries", "Resume Writing", "Scriptwriting", "Technical Writing", and "Translation", did not experience substantial impact. There are several factors contributing to these observations. Firstly, the insignificant results for services in the "Job Descriptions" and "Podcast Writing" categories can be attributed to the small sample size. Secondly, a considerable proportion of services under the "Resume Writing" and "Other" categories primarily focus on content editing rather than content generation. Thirdly, writing services necessitating a high degree of domain-specific knowledge, such as those within "Research & Summaries", "Case Studies", "Technical Writing", and "Translation" categories, remain relatively unaffected by the introduction of ChatGPT. This is likely due to the inherent complexity of these tasks, which cannot be easily replicated by AI-generated content. Lastly, demand for writing services targeted at mass media, including "Podcast Writing" and "Scriptwriting", has also been minimally impacted. As the validity and quality of content in these domains is of paramount importance, ChatGPT is not yet capable of supplanting human expertise in providing such services.

Finally, the impact of the ChatGPT launch on the demand for content generation and editing services is significantly influenced by service prices. Higher service prices may make it more difficult for businesses and individuals to justify the cost of human content generation and editing services when compared to the speed and efficiency of ChatGPT.

Our study contributes to the growing body of literature on the impact of AI applications, with a focus on the effect of generative AI on content generation and editing services. While previous studies have investigated the effects of traditional AI and robotics on employment rates, industry job displacement, and employee behavior within companies, our research specifically examines the impact of generative AI on task-level demand for content generation and editing services. Furthermore, our findings contribute to the ongoing debate on complementarity versus substitutability between AI and human labor. Unlike studies that suggest AI and human labor can be complementary, we provide evidence that generative AI is substituting for human labor in this particular market. We also identified the differential impact of the ChatGPT launch on writing and editing services, shedding light on the comparative advantages and disadvantages of generative AI and human labor in these two specific tasks. Finally, our study also highlights the significant role played by the cost of human labor in the competition with generative AI, as higher service prices are associated with a greater negative impact of ChatGPT launch on demand for services. In summary, our study provides valuable insights into the impact of generative AI on labor markets.

Our study also has several implications for service providers in the human content generation and editing industry. Firstly, due to the negative impact of the ChatGPT launch on the demand for writing services, providers should adapt to market changes by diversifying their service offerings or finding ways to differentiate themselves from automated content generation tools. Secondly, as the negative impact of the ChatGPT launch is primarily concentrated on writing services rather than editing services, providers may consider shifting their focus to editing services to lessen the impact of generative AI. Lastly, our finding that higher service prices are associated with a more negative impact on demand for the service highlights the need for providers to carefully consider their pricing strategies and find ways to justify their prices in comparison to automated services.

We acknowledged that our investigation is subject to certain limitations. Firstly, the dataset utilized in the analysis spans only one year, which may not provide a comprehensive understanding of long-term trends and potential effects. A more extended period of observation could yield different insights and conclusions. Secondly, the investigation is centered solely on the launch of ChatGPT, which may not accurately represent the broader landscape of generative AI tools. Including additional AI tools in the analysis could offer a more holistic view of the impact on content generation services. Thirdly, the study's focus on an online labor marketplace may not fully capture the diverse range of platforms and channels through which content generation services are offered and consumed. A more diverse sample of marketplaces and platforms could strengthen the generalizability of the findings. Lastly, the use of the number of reviews as a proxy for demand may not be an entirely accurate measure, as it may not account for potential discrepancies between the number of reviews and the actual demand for services. Employing alternative metrics or triangulating multiple indicators of demand could enhance the validity of the study's conclusions.

References

Acemoglu, D., and Restrepo, P. 2018. "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American economic review* (108:6), pp. 1488-1542

Acemoglu, D., and Restrepo, P. 2022. "Tasks, Automation, and the Rise in U.S. Wage Inequality," *Econometrica* (90:5), pp. 1973-2016.

Agrawal, A., Gans, J. S., and Goldfarb, A. 2019. "Exploring the Impact of Artificial Intelligence: Prediction Versus Judgment," *Information Economics and Policy* (47), pp. 1-6.

Autor, D., Mindell, D. A., and Reynolds, E. B. 2022. "Why the Future of Ai Is the Future of Work," in: *MIT sloan*.

Autor, D. H., Katz, L. F., and Krueger, A. B. 1998. "Computing Inequality: Have Computers Changed the Labor Market?," *The Quarterly journal of economics* (113:4), pp. 1169-1213.

Autor, D. H., Levy, F., and Murnane, R. J. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics* (118:4), pp. 1279-1333.

Bresnahan, T. F., Brynjolfsson, E., and Hitt, L. M. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence," *The quarterly journal of economics* (117:1), pp. 339-376.

Brynjolfsson, E., and Mitchell, T. 2017. "What Can Machine Learning Do Workforce Implications," *Science* (358:6370), p. 6.

Davenport, T. H., and Mittal, N. 2022. "How Generative Ai Is Changing Creative Work," in: *Harvard Business Review*.

Dixon, J., Hong, B., and Wu, L. 2021. "The Robot Revolution: Managerial and Employment Consequences for Firms," *Management Science* (67:9), pp. 5586-5605.

Foerderer, J., Kude, T., Mithas, S., and Heinzl, A. 2018. "Does Platform Owner's Entry Crowd out Innovation? Evidence from Google Photos," *Information Systems Research* (29:2), pp. 444-460.

Forbe. 2022. "Ai Will Impact the Labor Market, but Workers Should Embrace the Technology, Not Fear It."

Forbe. 2023. "Generative Ai and the Future of Creative Jobs." Retrieved 2023-04-04, from https://www.forbes.com/sites/gilpress/2023/03/06/generative-ai-and-the-future-of-creative-jobs/

Hoynes, H., and Rothstein, J. 2019. "Universal Basic Income in the United States and Advanced Countries," *Annual Review of Economics* (11), pp. 929-958.

Li, J., Bonn, M. A., and Ye, B. H. 2019. "Hotel Employee's Artificial Intelligence and Robotics Awareness and Its Impact on Turnover Intention: The Moderating Roles of Perceived Organizational Support and Competitive Psychological Climate," *Tourism Management* (73), pp. 172-181.

Lysyakov, M., and Viswanathan, S. 2022. "Threatened by Ai: Analyzing Users' Responses to the Introduction of Ai in a Crowd-Sourcing Platform," *Information Systems Research*).

McKinsey. 2023. "What Is Generative Ai?", from https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai#/

Noy, S., and Zhang, W. 2023. "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence," *Science* (381:6654), pp. 187-192.

OpenAI. 2022. "Introducing Chatgpt." OpenAI.

Sturm, T., Gerlacha, J., Pumplun, L., Mesbah, N., Peters, F., Tauchert, C., Nan, N., and Buxmann, P. 2021. "Coordinating Human and Machine Learning for Effective Organization Learning," *MIS Quarterly* (45:3), pp. 1581-1602.

Thorp, H. H. 2023. "Chatgpt Is Fun, but Not an Author," *Science* (379:6630), p. 5.

Tong, S., Jia, N., Luo, X., and Fang, Z. 2021. "The Janus Face of Artificial Intelligence Feedback: Deployment Versus Disclosure Effects on Employee Performance," *Strategic Management Journal* (42:9), pp. 1600-1631.

Van den Broek, E., Sergeeva, A., and Huysman Vrije, M. 2021. "When the Machine Meets the Expert: An Ethnography of Developing Ai for Hiring," *MIS Quarterly* (45:3), pp. 1557-1580.

Wang, L., Huang, N., Hong, Y., Liu, L., Guo, X., and Chen, G. 2023. "Voice-Based Ai in Call Center Customer Service: A Natural Field Experiment," *Production and Operations Management* (32:4), pp. 1002-1018.

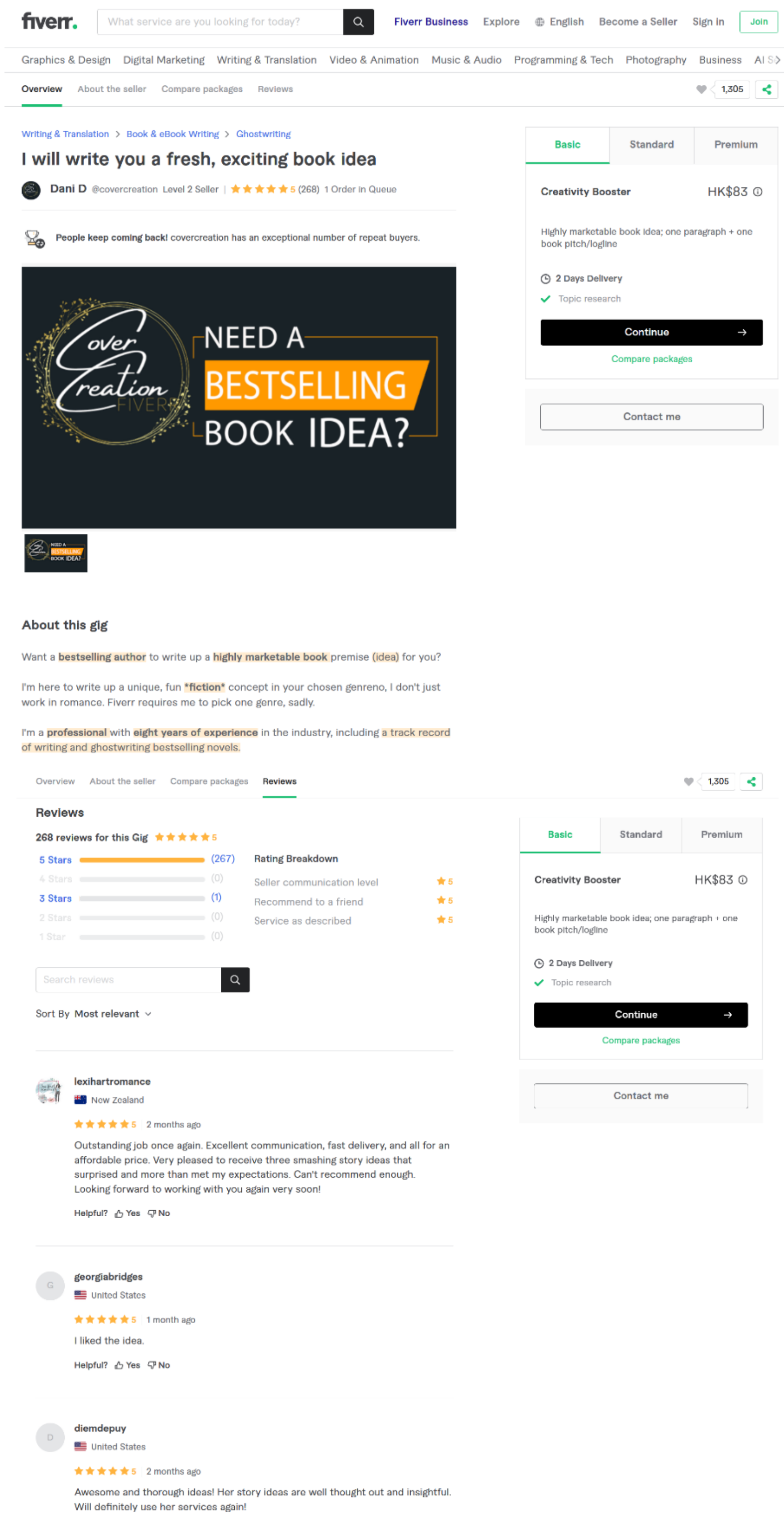
Weber, L., and Ellis, L. 2023. "The Jobs Most Exposed to Chatgpt," in: *The Wall Street Journal*.

Xue, M., Cao, X., Feng, X., Gu, B., and Zhang, Y. 2022. "Is College Education Less Necessary with Ai? Evidence from Firm-Level Labor Structure Changes," *Journal of Management Information Systems* (39:3), pp. 865-905.

Ye, Q., Law, R., Gu, B., and Chen, W. 2011. "The Influence of User-Generated Content on Traveler Behavior: An Empirical Investigation on the Effects of E-Word-of-Mouth to Hotel Online Bookings," *Computers in Human Behavior* (27:2), pp. 634-639.

Zhang, D., Peng, G., Yao, Y., and Browning, T. R. 2023. "Is a College Education Still Enough? The It-Labor Relationship with Education Level, Task Routineness, and Artificial Intelligence," *Information Systems Research* (Ahead in press).

Appendix



Appendix a. Screenshots of Service Pages and Reviews

Online Tutoring
Life Coaching
Career Counselling
Game Coaching
Gaming
Arts & Crafts
Astrology & Psychics
Modelling & Acting
Fitness
Nutrition
Wellness
Traveling
Puzzle & Game Creation
Styling & Beauty
Cosmetics Formulation
Family & Genealogy
Collectibles
Data Entry
Product Photographers
Portrait Photographers
Lifestyle & Fashion Photographers
Real Estate Photographers
Event Photographers
Food Photographers
Aerial Photographers
Photography Advice
Virtual Assistant
Financial Consulting
Electronics Engineering
Applications & Registrations
Event Management
Tax Consulting
Online Music Lessons

Appendix b. Subcategories of Services in the Control Group

Subcategories	Number of Services	β	Standard Errors
Ad Copy	146	-0.056	0.032
Articles & Blog Posts	542	-0.060*	0.026
Beta Reading	402	0.006	0.018
Book & eBook Writing	447	-0.075**	0.022
Book Editing	499	-0.009	0.018
Brand Voice & Tone	144	-0.104**	0.030
Business Names & Slogans	316	-0.061**	0.021
Case Studies	88	-0.013	0.039
Cover Letters	136	-0.101**	0.038
Creative Writing	544	-.089**	0.020
eLearning Content Development	139	-0.073*	0.034
Email Copy	280	-0.090**	0.023
Grant Writing	140	-0.106**	0.027
Job Descriptions	23	-0.152	0.091
LinkedIn Profiles	129	-0.078**	0.033
Other	379	-0.016	0.017
Podcast Writing	77	-0.008	0.058
Press Releases	328	-0.106**	0.024
Product Descriptions	472	-0.058**	0.020
Proofreading & Editing	598	0.035	0.023
Research & Summaries	347	0.019	0.023
Resume Writing	423	-0.035	0.030
Sales Copy	304	-0.101**	0.021
Scriptwriting	484	0.003	0.023
Social Media Copy	194	-0.119**	0.028
Speechwriting	175	-0.189**	0.032
Technical Writing	381	-0.013	0.018
Transcription	506	-0.058**	0.018
Translation	666	0.035	0.020
UX Writing	17	-0.164**	0.054
Website Content	557	-0.093**	0.021
White Papers	101	-0.165**	0.036
Writing Advice	19	-0.071	0.061

Appendix c. Summary of Our Subsample Analysis on Different Service Subcategories in the Treatment Group

Notes: β indicates the coefficient estimate for $After_t * Treat_i$.
* and ** denote significance at 5% and 1% levels, respectively.