More, but Worse: The Impact of AI Writing Assistance on the Supply and Quality of Job Posts

Emma Wiles
MIT

John J. Horton
MIT & NBER

March 20, 2024

Most recent draft here.

Abstract

We study a randomized experiment conducted on an online labor market that encouraged employers to use a Large Language Model to generate a first draft of their job post. Treated employers are 20% more likely to post the job and decrease time spent writing their job post by 40%. Among the posted jobs, treated employers receive 5% more applications. Despite the additional applications, treated employers are 18% less likely to make a hire. We find no evidence that this is driven by treated employers receiving lower-quality applicants. Moreover, despite the large increase in the number of jobs posted, there is no difference in the overall number of hires between treatment and control employers. These results imply that the treatment lowered the probability of hiring not only marginal jobs, but also among jobs that would have otherwise made a hire. We rationalize these results with a model in which employers with heterogeneous values of hiring can attract better matches by exerting effort to precisely detail required skills. We show how a technology that lowers the cost of writing and imperfectly substitutes for effort causes more posts but lowers the average hiring probability through both marginal posts (as these are less valuable) and inframarginal posts (as the technology crowds out effort). We provide evidence for these mechanisms using employer screening behavior and the embeddings of the job posts texts'.

1 Introduction

The rapid advancement of artificial intelligence technologies, particularly in the field of Large Language Models (LLMs), has sparked considerable interest and speculation on their impact on the labor market. Anecdotaly, generative AI is already being used to generate application materials like cover letters and resumes as well as job posts (Smith, 2023; Mok, 2023). Hiring is costly—beginning with the writing of the job post and followed by applicant search and screening (Barron and Bishop, 1985; Blatter et al., 2012). If generative AI proves to be effective in assisting employers with job post writing, it could lower the cost of job posting. It may also lead to more standardized, coherent, and targeted job descriptions potentially reducing information asymmetry between employers and job seekers. On the other hand, concerns arise regarding the potential homogenization of job postings, the impact on job search strategies, and the downstream matches that result. From the perspective of online labor markets and other platforms, providing or encouraging the use of generative AI could either improve the efficiency and accuracy of job postings or flood the market with informationless or homogeneous posts.

We analyze an experiment run on a large online labor market. We study the question of how providing would-be employers with LLM generated job posts impacts posting, user behavior, and hiring. A randomly selected treatment group of first time would-be employers were offered first drafts of their job posts written by generative AI. First, this experiment directly tests whether a technology which lowers the cost of posting increases the number and share of jobs posted. Second, we can see what types of jobs the treatment induces, and what types of applications they receive. Third, it allows us to test the efficiency of providing this kind of AI assistance to platforms who might consider this as a policy.

If firms find hiring to be costly in terms of time and domain knowledge, using LLMs to generate hiring materials has the potential to increase the supply of jobs. The existence of a multi billion dollar HR & recruiting industry suggests this is the case. Blatter et al. (2012) show that hiring one skilled worker costs 10 to 17 weeks of wages, and that these costs increase with the skill requirements of the position. Hiring online is also costly (despite fewer frictions), and digital platforms use recommendation systems to lower the costs of search and screening (Oestreicher-Singer and Sundararajan, 2012; Horton, 2017).

Creating the job post itself can also be costly. In search and matching markets, employers create job openings and adjust their search depending on how costly a vacancy is (Rogerson, Shimer and Wright, 2005). There are mechanisms that search and matching models abstract away from that could be important for practice, like the decision to finish posting a job once started. In writing job posts, employers have traditionally had to rely on their own expertise or outsourced this work to recruiting agencies.

In the platform on which the experiment is run, 92% of employers who have posted jobs before publish a job post that they have started. A technology which makes it easier to post a job could benefit such employers. This is especially true for first time job posters who are less familiar with norms of the platform. On this platform, only 25% of those who begin the job posting process for the first time eventually publish a post. If this intervention can lessen frictions and make it easier for employers to post, in addition to whatever resources firms might allocate themselves, platforms and other social planners might consider further expenditure to subsidize job posting, given the financial and social returns to job formation and employment.

Our first finding is that there is significant interest from employers—the employer can choose to opt in to receive an AI written first draft of the job post or opt out and write the job post themselves. 75% of employers opted in to receive the AI written first draft. We are able to track both which employers received the AI generated draft and the edits that they made to it before publishing the post.

Because we generated AI-written draft for both job posts in the treatment and control group (despite only revealing the AI-written draft to the treatment group), we can estimate a treatment effect for the similarity of the AI writing to the post that resulted. We use a measure of similarity where 1 means the two documents are identical and 0 means they share no elements. We find that job posts in the treatment group had mean similarity of 0.65 as opposed to jobs in the control group which had a similarity coefficient of 0.3.

We find that treated employers are 20% (or 6 percentage points) more likely to post a job than employers in the control group. Among those who do post a job, treated employers spend about 40% less time writing the job post than employers in the control group, on a base of 8 minutes. The distribution of length is compressed–the job posts which would have been short get longer and those which would have been very long get more compact. While the difference in mean number of words is small, this two-sided distributional shift causes the median word count to increase by 60%.

We also look downstream to how these treated job posts fared among jobseekers. Jobseekers rely on noisy signals of fit and job quality from job posts to decide whether or not to apply, which can vary by employer and job type. For example, non-native English[1] speaking employers in the control group receive fewer significantly fewer applications than native speaking employers. We find that the treatment was particularly useful to non-native English speakers—for them, treated job posts got significantly more applications. Since native English speakers saw no effect to the number of applications they received, the treatment significantly tightened the gap between the number of applications received by employers along this dimension. We also test whether applicant pools for treated job posts are lower quality on average, using a measure of quality defined by the platform based on jobseekers' prior experience. We find no evidence to support this.

Despite this increase in applicants, treated employers are 18% less likely to make a hire on their first job post. The overall share of treated employers that hire a worker is no different to the share of control employers who hire. It may have saved the employers time,

[1] We proxy for native English or not using the country that the employer reported registering from. We lassify those from Anglophone countries US, UK, Ireland, Canada, New Zealand, and Australia as native English speakers.

but access to AI written drafts resulted in no more matches.

In order to reconcile the large treatment effect to the number of job posts with no effect to the number of hires, we present a model where would-be employers decide whether to post a job with effort, post a job without effort, or not post a job at all. When writing a job description, employers can decide how much time and effort to put into carefully detailing the specifics of the tasks the job required, and the skills necessary to complete it. Therefore we model effort as something which causes the range of applications the job post induces to shrink, making it more likely for at least one application to come from a worker similar enough to what the job post requires to be worth hiring.

We introduce the AI as a technology which lowers the cost of posting a job but crowds out effort that some employers would have otherwise put into making the job post precise. If the cost of posting goes down for a subset of the job posts (the treated group with access to AI), a higher share of employers will post a job. However, the marginal jobs induced by the lower costs in the first period are ones with lower value to the employer, and causes ambiguous effects to hiring unconditional on posting. This rationalizes our otherwise surprising result that the treatment group had no more hires, despite 20% more job posts. Not only does the treatment induce lower value jobs that are less likely to hire, but it even makes the inframarginal jobs less likely to make a hire by decreasing the specificity of the posts.

To empirically test these hypotheses, we first show that employers' in the treatment group exhibited lower search effort than those in the control group. This is consistent with the hypothesis that the jobs posted in the treatment group were of lower value to the employers. Next, we embed the text of the job posts using OpenAI's "text-embedding-ada-002" model to create numerical representations of the texts. We first plot the embeddings of the job posts in the treatment and control group and show that the job posts in the treatment group are clustered closer together than the treatment group. We then calculate the cosine similarity between each job post's embeddings to show that job posts in the treatment group are on average more similar to each other than those in the control group. This is consistent with the hypothesis that the text of the job posts in the treatment group are more generic.

The main contribution of this paper is to provide early evidence on the impact of generative AI in hiring. Tambe et al. (2019) suggests that using ML algorithms for recruiting can provide new knowledge that the recruiters missed. And Van den Broek et al. (2021) shows that humans can use ML in a hybrid practice in which candidates are judged and selected by relying on a combination of ML and recruiters domain expertise. While much of the literature on AI/ML in hiring is focused on algorithmic approaches to search and screening, there are a few papers on the use of AI/ML for application materials. Early evidence suggests that using algorithmic writing assistance on resumes makes workers more likely to be hired (Wiles, Munyikwa and Horton, 2023). But there is evidence that when the use of AI in application materials is disclosed, that people perceive the applicants as less competent and warm (Weiss, Liu, Mieczkowski and Hancock, 2022).

We also contribute to a very young literature on generative AI and productivity. Across multiple domains and versions of LLMs, there is evidence of large productivity effects. Noy and Zhang (2023) find that the use of ChatGPT on writing tasks caused treated workers to take 0.8 SD less time to create work that was even higher quality than the work from the control group. When paired with GitHub Copilot, treated workers completed coding tasks 55% faster than a control group without Peng et al. (2023). We contribute to this literature by providing a case study in a real labor market where access to a LLM saves users time and increases their engagement with the platform, but that the positive results do not exist downstream of posting.

Lastly, we contribute to a literature on the role of cost of entrance to market quality. Mankiw and Whinston (1986) theorize that free entry is not always socially optimal. In one paper analyzing the effect of Chinese export subsidy program, the author finds that the subsidy made exporters worse off by polluting the market with low-quality firms (Zhao, 2018). Filippas et al. (2023) find that when subsidizing entrance to an online labor market, the financial benefits to workers outweighed the cost of the resulting increase in job search.

Our results suggest that if employers have perfect information about the technology, generative AI is weakly welfare increasing for employers. They saved time, and the technology made it possible for some jobs that would otherwise been abandoned to be posted. However, no more hires resulted from these matches, and we suggest that for most employers, the use of the AI crowded out effort that they would have otherwise exerted to make a more specific job post.

Our results suggest that workers are made worse off. The flood of job posts with no increase in hires increased search costs for workers, and makes it harder for them to tell apart good and bad jobs. They also "wasted" their time on applications—despite resulting in no more hires, job posts in the treatment group got 106,565 more applications. From a platform's perspective, the usefulness of such a tool depends on if the increase in likelihood of posting a first job posts induced by the treatment caused employers to keep coming back to the platform for future jobs. If not, the overwhelming result of the treatment was to flood the market with low-quality job posts.

The rest of the paper proceeds as follows. Section 2 describes the online labor market which serves as the focal market for this experiment. Section 3 describes the experimental design and results from the first-stage. Section 4 reports the experimental results of the treatment on job posting and hiring. Section 5 provides evidence for our proposed mechanism. In Section 6 we present a simple model that can rationalize our findings. Section 7 concludes.

2 The setting

This experiment was conducted on a large online labor market. In online labor markets, employers[2] search for and hire workers to complete jobs that can be done with only a computer and an internet connection. These markets can differ in their scope and focus, and platforms have different responsibilities they provide to employers and workers. Some common services platforms provide include soliciting and promoting job openings, hosting profile pages, processing payments, certifying worker skills, and maintaining a reputation system (Horton, 2010).

In the platform which we use as our empirical setting, employers post job openings on the platform website with job descriptions, required skills, and scope of project. First, a would-be employer gives the title of the job, for example "E-commerce website copywriter", "Web developer", or "Executive assistant." They then report the skills necessary to complete the job. Next, the employer picks the broad category the job falls into, for example, as "Administrative Support", "Data Entry", "Software Development", among others. The jobs can either be one-off projects called "fixed price jobs" or hourly jobs, in which case the employer estimates how many hours they expect the job to take.

Workers find out about job openings in three ways. They can use electronic search to seek job posts in specific categories or for job openings that require specific skills. They can receive email notifications from the platform when a job is posted in a particular category. And finally, they can receive invitations from employers to apply to specific jobs.

Employers find workers in two ways. They receive organic applications from workers who find the job opening independently, or they search for workers themselves and invite specific workers to apply. Employers can search through worker "profiles." These profiles contain workers' history of work on the platform (jobs, hours, hourly rates, ratings) as well as their education history and skills. For both workers and employers, the platform verifies some of the information available to the other side of the market.

When a worker chooses to apply to a job opening, they apply with a cover letter and an hourly wage bid or a total project bid for fixed-price jobs. The employer determines whether to hire and, if so, which worker(s) to select. To complete the work on hourly jobs, workers install custom tracking software that serves as a digital punch clock. The software records not

[2] We use the terms "employer," "job opening," and "application" for consistency with the economics literature and not as a commentary on the legal nature of the relationships created on the platform.

only the time spent working but also keystroke count and mouse movements. The software also captures images of the worker's computer screen randomly. This information is all sent to the platform's servers, and made available to the employer for monitoring in real-time. At the end of the contract, both parties give a reason for ending the contract (usually that the project was completed successfully) and provide both written and numerical feedback about each other.

3 Experimental design

This experiment intercepts would-be employers at the moment they begin to post their first job. From June 7, 2023, through July 20, 2023, newly registered employers on the platform were randomly allocated into a treatment and control group.

The experimental sample includes 181,962 employers who post 50,125 openings between them. Appendix Figure 9 shows the daily allocations of employers into the treatment and control groups.

3.1 Experimental intervention at the start of posting a job

When an employer on the platform wants to post a job, they go through a series of steps. First they provide a job title, the length of time they expect the job to last, and a list of skills required or demanded of the job. After they provide this information, they report some information on their expected budget and then move on to a page where they can input a job description. For employers in the control group, here they type in their job description and then submit the job to be posted.

For employers in the treatment group, as soon as they start to post a job, they are are offered two options. They can either "get started using AI" or "I'll do it without AI." If they click on the latter button, they receive the status quo job posting experience. If they elect to "get started using AI" they are asked to describe the job they want to post in a sentence or two. See Figure 1 for a stylized version of the interface that the employers use.

As an example, after being asked to describe the job in a sentence or two, one employer wrote:

> *I need someone to generate a an Excel database showing the frequency of a search term in a list of targeted business media*

This is incorporated into a prompt, calling a popular generative AI service. [3] On the next

[3] The exact prompt is listed in Appendix A.1.

page, the employer is shown the job post as written by the AI as well as a list of required skills. In the case of the above input, the employer would be shown:

> *We are looking for an expert who can generate an Excel database that displays the frequency of a specific search term in a list of targeted business media. The ideal candidate should have the following skills:*
>
> - *Strong knowledge of Excel*
> - *Ability to work with large sets of data*
> - *Research skills*
> - *Attention to detail*
> - *Time management skills*

The page containing this draft contains the message to "take time to review your job post and make it your own." Employers are able to edit the job post however they want and they are also shown a series of options for how the AI can edit the job post for them. The options are "Make it casual", "Make it formal", "Shorten it", "Add more details", "Rewrite it". They will observe an have the option to edit a job post category which is determined by the API call as well.

3.2 Description of data used in the analysis

The dataset we use in this analysis consists of all job posts posted by employers in the experimental sample between the moment they were allocated into the experiment and August 4, 2023, 14 days after allocation ended. We construct job post level data with all posts, applications, and hires they have within 14 days of posting. Our economic outcomes of interest are 1) whether the employer eventually completed the job posting, 2) the number of applications to the job posts, and 3) whether or not anyone is ever hired for the job. We also collect the text of the job posts themselves, the amount of time the employer spent writing the post, and the count of skills required for the job. Lastly, we collect the country that the employer reported being in when they registered for the platform. We construct an imperfect definition of "native English speaker" which includes all employers who registered from the United States, Canada, United Kingdom, Ireland, Australia, or New Zealand.

We also use the output generated by the AI for both posts from treated and control employers. For employers in the treatment group, we observe the text generated each time they call the API, either through the initial job post generation or the later buttons used to have the AI edit the job post after. For employers in the control group, an API call is made

Figure 1: Stylized job post process for employers in the treatment group

using the job title, budget amount, expected length of job, and skills required. We observe the output, although the control employers do not.

3.3 Treatment take up

Of all the job posts by treated clients, 75% opted in to receive an AI generated first draft. The platform records every action and even click taken by each user to the microsecond. This helps us to see the 'first-stage' of the treatment. While opting into receiving the first draft was widely used, the personalization buttons were not. Treated employers made on average 1.2 API calls through the buttons, the first of which generated the draft job post. Employers used the "Make it casual" feature the most, 3.3% of the time. The next most used feature was "Add more details" which was used in 2% of the job posts. "Shorten it", "Make it formal", and "Rewrite it" were used on between 1 and 2% of the job posts.

Next, we calculate the Sørensen–Dice coefficient of each job post measuring the similarity between the job post an employer submitted and the one that the AI generated, regardless if the employer was in the treatment or control group. [4] It measures the proportion of common elements or features shared by two sets, relative to the total number of elements present in both sets. A dice coefficient of 1 means the job posts were identical, whereas a dice coefficient of 0 means they have nothing in common. While only treated employers who opt-in to the treatment ever receive these generated job posts, they are calculated for all job postings regardless of treatment group. We find that job posts where employers opt-ed out of the AI treatment had low levels of similarity, which matched the similarity of job posts in the control group, of around 0.3. In the treatment group the average dice coefficient is 0.65. This gives us a first stage of the treatment–job posts in the treatment group share more than double the elements with those generated by the API call than job posts in the control group.

We see an even more pronounced difference between those that opted in to the treatment compared to those that opted out, we can see that on job posts where employers opted in the dice coefficient is 0.89. This last measure gives us a magnitude for how much employers who use the AI generated first drafts are editing them before they post them.

[4] We cannot use the same prompt for job posts in the treatment and control group because the inputs from the employer are different. For job posts in the control group we generate a job post based on the title, skills, and proposed job duration that the employer inputs before writing their job description.

4 Experimental Results

Since employers were offered the choice to opt out of getting help from AI, our estimates are intent-to-treat effects for the entire experimental sample. In addition to the overall treatment effects, we present results interacting the treatment with a dummy variable for whether or not the employer registered for the platform from an anglophone country.

4.1 Treated employers were more likely to post a job

We will begin by observing that treated employers are about 20% more likely to post a job than employers in the control group. In Table 1 we show that only 25% of first time would-be employers who get to the landing page ever publish a job. This is very low compared to the employers who have posted a job before, for whom 92% who start the process for a given job publish it. There is clear room for improvement for keeping these would be employers in the hiring funnel. Treated job posts are 5 percentage points more likely to publish.

The effect of the treatment to whether an employer posts is significantly larger for employers who are native English speakers. Non native English speakers in the treatment group are 6.5 percentage points, or 24% more likely to publish than non native English speakers in the control group, while native English speakers only experience a 10% increase in likelihood of posting.

4.2 Treated employers spent less time writing the job post

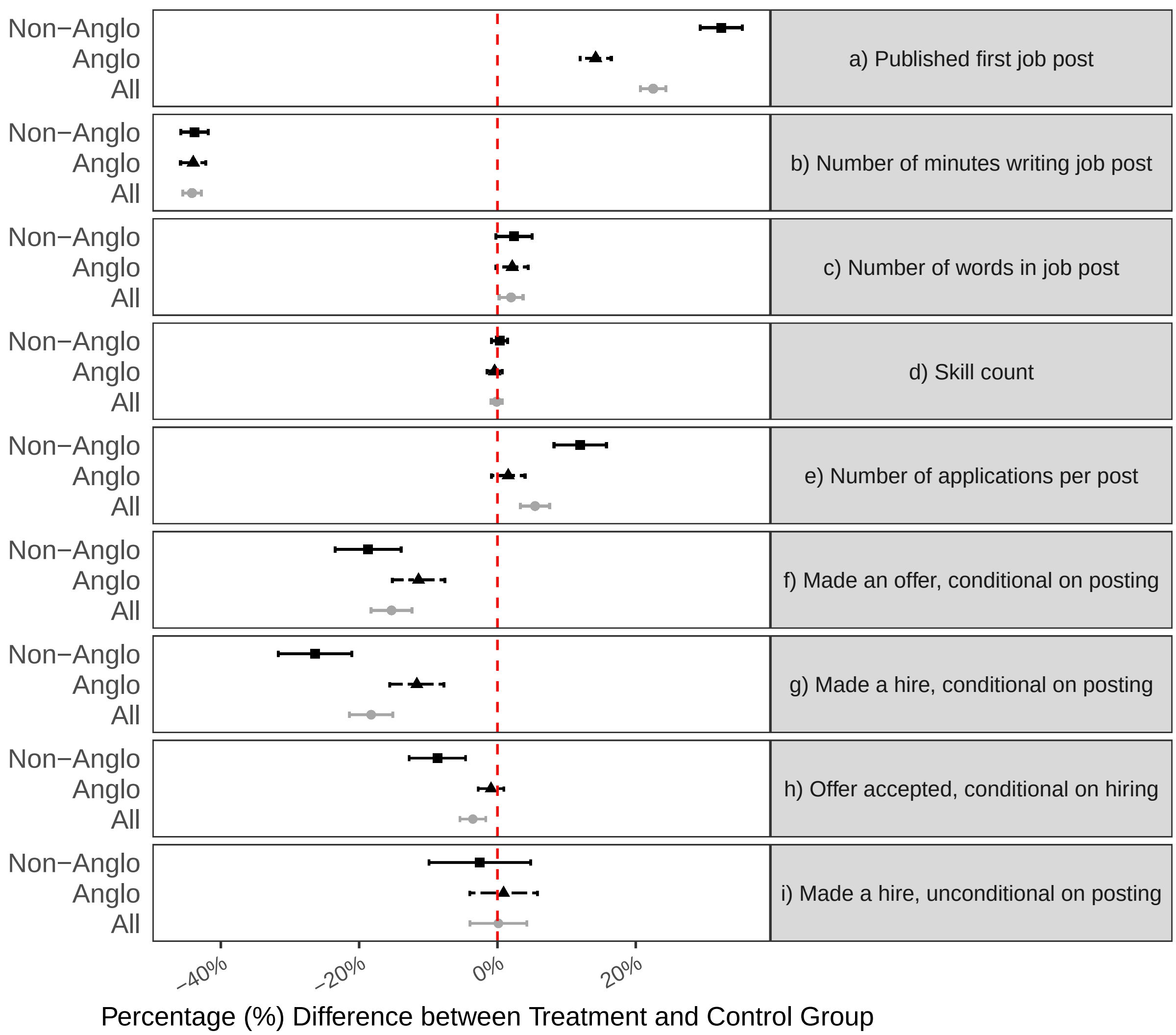
The treatment caused employers to spend less time writing the job post. The outcome in Table 2, minutes, is defined as the difference in the timestamps from when the employer first clicks on the page to post a job and when the employer finally presses submit on the job post. Column (1) shows that employers in the control group spend on average 8 minutes writing the job post while employers in the treatment group spend only 4.5 minutes.

4.3 Treated job posts were longer

On average, the treatment caused job posts to be longer. However, this masks an important distributional effect— the distribution of number of words in the treatment group was more compressed.

We start by looking at the mean. In Table 3 we see that job posts in the control group are on average 98 words long. Treated jobs are about 100 words long. However, if you look at the CDF of job post length in Figure 3 you see that the median job in the control group

Figure 2: Experimental Estimates



Notes: This analysis looks at the effect of being assigned treatment on outcomes for employers in the experimental sample. The x-axis is the percentage difference in the mean outcome between employers in the treated group and the control group. The outcome a) published first job post is a 0 if the employer never posts a job after allocations and 1 if they do. The outcomes b), c), d), e), f), and g) are all conditional on the employer posting a job. The outcome h) offer accepted is conditional on the employer posting a job and making an offer. The outcome i) made a hire is unconditional on posting a job, it is 0 if the employer doesn't hire anyone after allocations and 1 if they do. A 95% confidence interval based on standard errors calculated using the delta method is plotted around each estimate. The experimental sample is of employers who posted a job between June 7st and July 20th, 2023, with N = 181,962. Regression details on the number of jobs posted can be found in Table 1, on minutes in Table 2, on number of words in Table 3, and on skill count in Table 4. Regression details on the number of applications can be found in Table 5, on offers in Appendix Table 13, on hires in Table 6, on offers accepted in Table 7, and on hires unconditional on posting in Table 8.

Table 1: Effects of generative AI on employer proclivity to post jobs

	Dependent variable:	
	Indicator for if first job is posted	
	(1)	(2)
GenAI Treatment Assigned (Trt)	0.056***	0.065***
	(0.002)	(0.003)
Anglophone		0.109***
		(0.003)
Anglophone X Trt		−0.021***
		(0.004)
Constant	0.248***	0.200***
	(0.001)	(0.002)
Observations	181,962	181,962
R^2	0.004	0.016

Notes: This table analyzes the effect of the treatment on the number of jobs the employer posts over the experimental period. Likelihood of completing first job post is a binary variable for the job post that the employer was working on when they were allocated into the experiment. "Anglophone" is 1 if the employer registers from an anglophone country, defined as the United States, Canada, United Kingdom, Ireland, or New Zealand. The sample is made up of all employers in the experimental sample. Significance indicators: $p \leq 0.10 : *$, $p \leq 0.05 : **$ and $p \leq .01 : ***$.

Table 2: Effects of generative AI on length of time employer worked on job post

	Dependent variable:	
	Minutes writing job post	
	(1)	(2)
GenAI Treatment Assigned (Trt)	−3.581***	−3.302***
	(0.071)	(0.105)
Anglophone		1.033***
		(0.108)
Anglophone X Trt		−0.467***
		(0.143)
Constant	8.107***	7.537***
	(0.054)	(0.080)
Observations	38,841	38,841
R^2	0.061	0.064

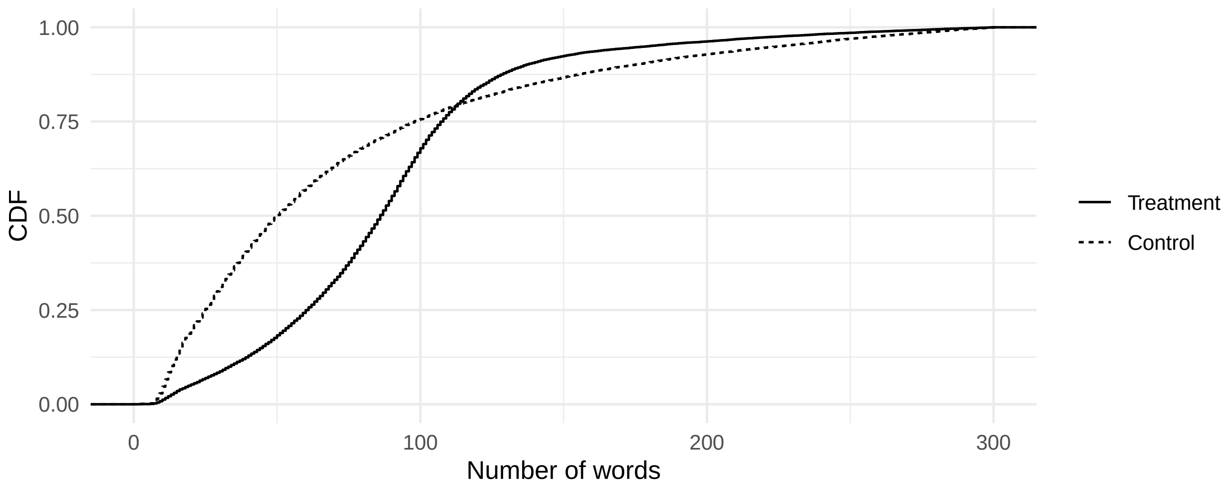
Notes: This table analyzes the effect of the treatment on the number of minutes the employer spent working on the job post. Minutes is the difference in the timestamps from when the employer starts the job post till the timestamp when they publicly post it. "Anglophone" is 1 if the employer registers from an anglophone country, defined as the United States, Canada, United Kingdom, Ireland, or New Zealand. The sample is conditioned on employers who posted a job. Significance indicators: $p \leq 0.10 : *$, $p \leq 0.05 : **$ and $p \leq .01 : ***$.

Table 3: Effects of generative AI on length of job post

	Dependent variable:		
	Number of words in job post		
	OLS		Quantile Regression
	(1)	(2)	(3)
GenAI Treatment Assigned (Trt)	1.942**	2.253*	32.000***
	(0.861)	(1.260)	(0.555)
Anglophone		8.208***	
		(1.282)	
Anglophone X Trt		−0.066	
		(1.726)	
Constant	98.509***	94.010***	56.000***
	(0.639)	(0.949)	(0.505)
Comparing	Means	Means	Medians
Observations	50,125	50,125	50,125
R^2	0.0001	0.002	

Notes: This table analyzes the effect of the treatment on the number of words in the job posts. In Columns (1) and (2) are Ordinary Least Squares models. Column (3) compares the median number of words in job posts. "Anglophone" is 1 if the employer registers from an anglophone country, defined as the United States, Canada, United Kingdom, Ireland, or New Zealand. The sample is conditional on jobs which were posted. Significance indicators: $p \leq 0.10 : *$, $p \leq 0.05 : **$ and $p \leq .01 : ***$.

Figure 3: Cumulative distribution function of the number of words in job posts



Notes: CDF of the number of words in employers job posts, by treatment status.

had 56 words, while the median number of words on job posts in the treatment group was 88.

4.4 Treated job posts listed different skill requirements

Every job post on the platform contains a list of skill requirements that jobseekers use to see if they are a good fit for a particular job. On average, there was no difference in the number of skills listed on each job post between the treatment and the control group, although this masks substantial heterogeneity. Table 4 shows jobs in more technical categories (Design, Software) saw more skills listed on the job posts, while less technical categories (Admin, Writing) had fewer. This shows that the treatment had effects on how skills were conveyed on the job posts, but these effects were heterogeneous and not straightforward to summarise in a uniform way.

Table 4: Effects of generative AI on the number of skills requested by a job post

	Dependent variable:					
	Number of skills requested in job post					
	(1)	(2)	(3)	(4)	(5)	(6)
GenAI Treatment Assigned (Trt)	0.027	0.050^{*}	0.255^{***}	0.298^{***}	-0.218^{***}	-0.438^{***}
	(0.020)	(0.029)	(0.041)	(0.036)	(0.068)	(0.069)
Anglophone		0.080^{***}				
		(0.030)				
Anglophone X Trt		−0.040				
		(0.040)				
Constant	4.788^{***}	4.745^{***}	4.814^{***}	4.468^{***}	4.870^{***}	4.842^{***}
	(0.015)	(0.022)	(0.030)	(0.027)	(0.051)	(0.054)
Category	All	All	Design	Software	Writing	Admin
Observations	50,125	50,125	11,918	13,100	4,068	4,558
R^2	0.00003	0.0002	0.003	0.005	0.003	0.009

Notes: This table analyzes the effect of the treatment on the number of skills requested in a job post. For jobs in the control group and those who opt-out of receiving an AI-written first draft, the skills are listed by the would-be employer as part of writing the job post. For jobs which get the post drafted by AI, the skills are pulled from the API call, although they can be overidden by the employer. "Anglophone" is 1 if the employer registers from an anglophone country, defined as the United States, Canada, United Kingdom, Ireland, or New Zealand. The sample is conditional on jobs which were posted. Significance indicators: $p \leq 0.10 : *$, $p \leq 0.05 : **$ and $p \leq .01 : ***$.

4.5 Treated job posts received more applications

Table 5: Effects of generative AI on number of applications a job post received

	Dependent variable:		
	Total apps		
	(1)	(2)	(3)
GenAI Treatment Assigned (Trt)	0.889***	1.754***	1.182***
	(0.170)	(0.249)	(0.058)
Anglophone		3.109***	
		(0.253)	
Anglophone X Trt		−1.477***	
		(0.341)	
Constant	16.361***	14.657***	4.051***
	(0.126)	(0.188)	(0.041)
Observations	50,125	50,125	181,962
R^2	0.001	0.005	0.002

Notes: This table analyzes the effect of the treatment on the number of applications the employer recieves within 14 days of posting a job. "Anglophone" is 1 if the employer registers from an anglophone country, defined as the United States, Canada, United Kingdom, Ireland, or New Zealand. In Columns(1) and (2) the sample is made up of all employers who post a job post, and in Column (3) the sample is unconditional on whether or not the employer posted a job. Significance indicators: $p \leq 0.10 : *$, $p \leq 0.05 : **$ and $p \leq .01 : ***$.

The treatment induced 28% more applications in the treatment group than in the control group. Job posts in the control group received 369,020 applications overall, while job posts in the treatment group received 475,585, a difference of 106,565 applications .

In Table 5 we will break down the effect of the treatment to the number of applications per job post—conditional on the employer publishing a job post. In Column (1) we see that job posts in the control group jobs received 16 applications on average. Across all job posts, treated jobs received almost 1 additional application. However, this masks significant heterogeneity by employer. In Column (2) we interact the treatment with a dummy variable for whether or not the employer was a native English speaker. First we notice that on average, workers prefer to apply to jobs from native English speaking employers—on average, job applications from native English speakers receive three more applications than those who were not. As for the effect of the treatment, there is no significant treatment effect to native English employers, while the treatment induces 1.75 additional applicants for jobs posted by non native English speaking employers. While the treatment did not entirely close the gap between number of applications received by employers along this dimension, it significantly tightened it.

In Column (3) we look at the effect to applications unconditional on whether or not the employer posted a job. As we would expect looking at the entire sample, we see employers in the treatment group receive more applications— both because they are more likely to post a job and because the jobs posted in the treatment group draw more applications.

4.6 Treated employers job posts were less likely to make an offer, conditional on posting a job

Table 6: Effects of generative AI on number of hires

	Dependent variable:	
	Hire, conditional on posting a job	
	(1)	(2)
GenAI Treatment Assigned (Trt)	-0.035^{***}	-0.035^{***}
	(0.003)	(0.005)
Anglophone		0.111^{***}
		(0.005)
Anglophone X Trt		0.006
		(0.007)
Constant	0.192^{***}	0.131^{***}
	(0.003)	(0.004)
Observations	50,125	50,125
R^2	0.002	0.025

Notes: This table analyzes the effect of the treatment on if the employer makes a hire. Hire, conditional on post is 1 if the job post that the employer was working on when they were allocated into the experiment makes a hire within 14 days. The sample is made up of all employers in the experimental sample. Significance indicators: $p \leq 0.10: *$, $p \leq 0.05: **$ and $p \leq .01: ***$.

Despite the large increase in employers propensity to post a job, and despite the increase in applications to those jobs, treated employers who post jobs are actually significantly less likely to make an offer or hire.

In Column (1) of Table 13 we can see that a posted job has around a 20% chance of making an offer on average. Treated jobs are 3 percentage points less likely to make an offer. These results are generally consistent for both hires and offers[5], so results to hiring are not overall driven by the employer making offers that are not accepted.

[5] See regression details on whether or not an employer made a hire in Table 6

4.7 Treated non-native English speakers experienced more rejections after making an offer

Table 7: Effects of generative AI on the share of offers that are accepted

	Dependent variable:	
	Offer accepted	
	(1)	(2)
GenAI Treatment Assigned (Trt)	-0.029^{***}	-0.060^{***}
	(0.008)	(0.013)
Anglophone		0.178^{***}
		(0.011)
Anglophone X Trt		0.052^{***}
		(0.016)
Constant	0.802^{***}	0.688^{***}
	(0.006)	(0.009)
Observations	10,996	10,996
R^2	0.001	0.060

Notes: This table analyzes the effect of the treatment on the share of offers that are accepted. Offer accepted is 0 if an offer is made which does not lead to a hire and 1 if it does lead to a hire. The sample is made up of all employers in the experimental sample who post a job and make at least one offer. Significance indicators: $p \leq 0.10: *$, $p \leq 0.05: **$ and $p \leq .01: ***$.

Most offers are accepted. In our sample, 80% of job offers result in a hire. Table 7 Column (1) shows that workers given offers by employers in the treatment group were 3 percentage points less likely to accept. However, we can see from Column (2) that this result is driven entirely by non-native English speaking employers, for whom the treated group are 6 percentage points less likely to have an offer accepted.

4.8 Treated employers were no more likely to make a hire

One puzzle of this experiment is that despite the 20% increase in job posting, there is no overall increase in hires. Unconditional on whether or not the employers post a job, the likelihood of hiring in the control group is only 5% as we can see from Column (1) of Table 8. Among this entire experimental sample, treated employers are no less likely to make a hire.

These results imply that either none of the marginal jobs induced by the treatment made a hire, or the treatment actually made the inframarginal jobs worse.

Table 8: Effects of generative AI on number of hires, unconditional on posting a job

	Dependent variable:	
	Hire, unconditional on posting a job	
	(1)	(2)
GenAI Treatment Assigned (Trt)	0.0001	−0.001
	(0.001)	(0.001)
Anglophone		0.049***
		(0.001)
Anglophone X Trt		0.001
		(0.002)
Constant	0.048***	0.026***
	(0.001)	(0.001)
Observations	181,962	181,962
R^2	0.00000	0.013

Notes: This table analyzes the effect of the treatment on if employer makes an offer. Hire, unconditional on post is 1 if the employer makes any hire within 14 days of being allocated into the experiment. The sample is made up of all employers in the experimental sample. Significance indicators: $p \leq 0.10 : *$, $p \leq 0.05 : **$ and $p \leq .01 : ***$.

5 Mechanisms

In Table 6 we showed that treated employers were 3 percentage points less likely to make a hire. In this section we provide evidence of underlying mechanisms. We first provide evidence that this effect is driven by the job posts induced by the treatment being on the margin of the cost benefit trade off. We provide evidence for this by showing that employers of treated job posts exhibit lower employer screening efforts. We next provide evidence that the job posts in the treatment group were more generic than those in the control group, as measured by both the text of the job posts and the similarity of their applicants. Lastly, we show that while the applicant pools were more similar for jobs in the treatment group, the issue is the applicant's fit, not their quality.

5.1 Employers exhibited lower search effort

Employers of treated job posts exhibited less search and screening efforts than those in the control group. Table 9 Column (1) shows that employers invite fewer would-be applicants to apply to treated jobs. In Column (2) the outcome is the number of applicants an employer puts on their short list. Employers shortlist fewer applicants to treated jobs. And in Column (3) the outcome is the number of interviews initiated by the employer, defined as a direct message from the employer to the applicant. Employers of treated job posts also interview fewer applicants. The magnitudes of these effects are all small but statistically significant, suggesting that the treatment induces more job posts, but that these job posts are relatively less beneficial to employers.

Table 9: Effects of generative AI on employer behavior

	Dependent variable:		
	Number of invites	Number of shortlists	Number of interviews
	(1)	(2)	(3)
GenAI Treatment Assigned	-0.008^{***}	-0.005^{***}	-0.251^{***}
	(0.002)	(0.001)	(0.029)
Constant	0.103^{***}	0.032^{***}	1.596^{***}
	(0.001)	(0.001)	(0.021)
Observations	50,125	50,125	50,125
R^2	0.0004	0.001	0.002

Notes: This table analyzes the impact of the treatment on employer behavior. Number of invites is the number of times a would-be employer reached out to a potential applicant and invited them to apply. Number of shortlists is the number of applications an employer put on their short list of potential hires. And number of interviews is defined as a 1 if the employer direct messaged a jobseeker after receiving their application. The sample is conditioned on employers who posted a job which received at least one application. Significance indicators: $p \le 0.10 : *$, $p \le 0.05 : **$ and $p \le .01 : ***$.

5.2 Treated job posts were more "generic"

Treated job posts had more generic text than job posts in the control group. To do this we use cosine similarity, which measures the distance between two texts language and content. To do this we first get the embeddings for each job post using OpenAI's model "text-embedding-ada-00'. These embeddings are high-dimensional vectors that codify the semantic attributes and content of the job descriptions, transforming the text into a numerical format that captures underlying meanings and themes. In Table 10 our outcome of interest is the mean of the cosine similarities between the embedding of job post i and the embeddings for each other job post $-i$. A cosine similarity of 1 means the texts are identical, and a cosine similarity of 0 means they are completely orthogonal. We give the definition of cosine similarity in Equation 1.

$$\cos(\mathbf{A},\mathbf{B}) = \frac{\mathbf{A}\mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n}\mathbf{A}_i\mathbf{B}_i}{\sqrt{\sum_{i=1}^{n}(\mathbf{A}_i)^2}\sqrt{\sum_{i=1}^{n}(\mathbf{B}_i)^2}} \tag{1}$$

We find that job posts in the treatment group were on average closer to mean job post

Table 10: Mean cosine similarity of job posts by treatment cell

	Dependent variable:	
	Mean cosine similarity	Rank
	(1)	(2)
GenAI Treatment Assigned	0.014***	−6,647.600***
	(0.0002)	(98.939)
Constant	0.753***	20,179.000***
	(0.0002)	(73.492)
Observations	33,022	33,022
R^2	0.107	0.120

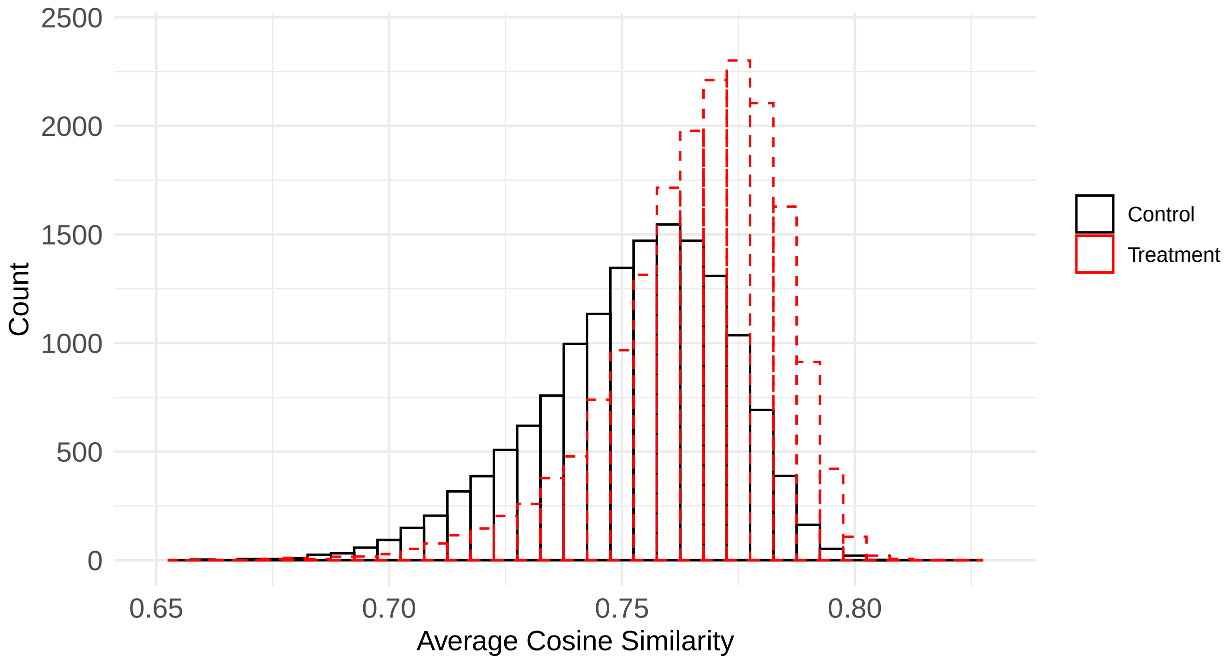
Notes: This table analyzes the effect of the treatment on how different job posts are from each other. For each job post we get the embeddings using OpenAI's 'text-embedding-ada-002' model, we then create a matrix of the cosine similarity between each job post and each other job post in the experiment. Then for each job post we take the mean of all of the cosine similarities, as a proxy for how generic a job post is. The outcome in column (1) is the mean cosine similarity between the ego and all other job posts in the experiment. The outcome in column (2) is the rank of those job posts in descending order. The sample consists of the subset of the experimental sample which post a job, and randomization occurs at the job post (and employer) level. Significance indicators: $p \leq 0.10 : *$, $p \leq 0.05 : **$ and $p \leq .01 : ***$.

embedding than job posts in the control group. The treatment effect is small, only 0.014 on a base of 0.75. However, the range of average cosine similarities is very narrow— the lowest average cosine similarity is 0.67 while the highest is 0.81[6]. This is because despite the fact that these job posts can be in very different industries, they are all job posts. This treatment effect covers 10% of the distance between the minimum and maximum average cosine similarity. Figure 4 shows how much this shifts the distribution of average cosine similarities.

Given the high dimensional nature of the embeddings, we cannot directly visualize them. To this end, we apply Principal Component Analysis (PCA) to reduce the dimensionality of the embeddings to two principal components, allowing us to visualize the embeddings in a 2D space. This reduction preserves as much of the variance in the data as possible in 2D. We plot the 2D embeddings for the treatment and control group in Figure 5. While the principal components themselves are not directly interpretability due to their composite nature, they still can facilitate a visual comparison of the job postings' embeddings. Most notably, the treatment appears to cause a shift in the distribution along the first principal component. We plot the 2D embeddings for the control group in Figure 5a. We investigate the treatment group in Figure 5b. Here we break down the job posts in the treatment group into those

[6] See Appendix Section A.2.1 to see examples of job posts with average cosine similarities at the min, max, and mean average cosine similarity.

Figure 4: Average cosine similarity by treatment status



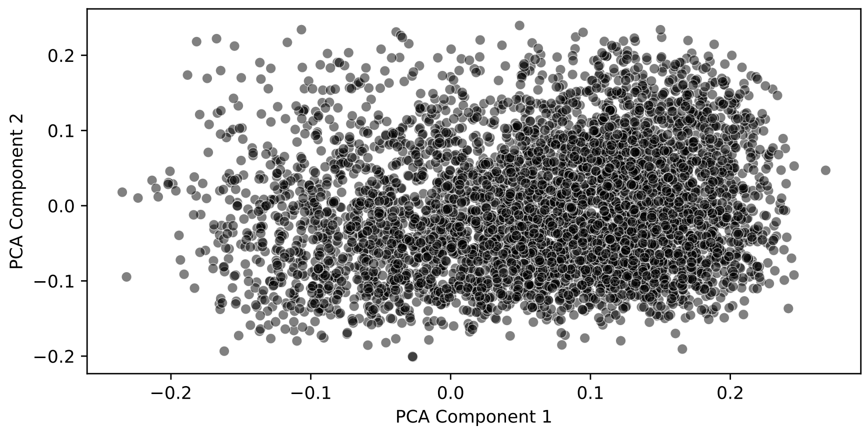
Notes: This plot shows the average cosine similarity in the treatment and control cells for all employers in the sample who posted a job.

that opted-in to receive the AI written first draft, plotted in blue, and those that opted-out, plotted in red. For 75% of the job posts the employer opted-in to receive the draft, and therefore the vast majority of embeddings are in blue. While the red embeddings for those that opted-out are placed more uniformly across the distribution of the first component, the ones that opted-in are clustered to the right.

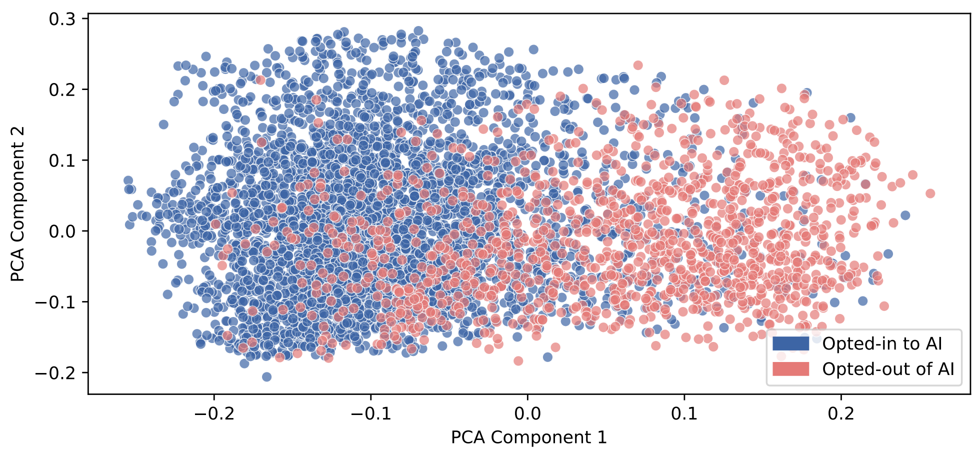
5.3 Treated job posts had a higher fraction of their applications in common with other job posts

We might imagine that if the treated job posts are in fact more generic, that they get applicants who are more similar to other job posts in the experiment. We then create a two dimensional matrix of job posts where the m by nth entry in the matrix is the cosine similarity of the n and mth job post in terms of the applications they received. For each job post i we then take the mean across all other job posts $-i$. This gives each job post an average measure of application overlap. In Table 11 we show that the mean share of applications in common is higher for jobs in the treatment group. In the control group, a job post's cosine similarity in terms of the applications it shares with other job posts in the experiment is 0.054. In the treatment group, it is 0.060. In both cases, the cosine similarity is very low

Figure 5: Embeddings of Job Posts Reduced to 2 Dimensions



(a) Embeddings of Job Posts Reduced to 2 Dimensions for Control Group



(b) Embeddings of Job Posts Reduced to 2 Dimensions for Treatment Group

Notes: These plots shows the job posts' embeddings reduced to two dimensions. We use OpenAI's "text-embedding-ada-002" model to turn the text of job posts into embeddings, and then use PCA to reduce the dimensionality of the embeddings into two dimensions. We then take a random sample of 5,000 job posts in the treatment group and 5,000 job posts in the control group, for ease of visualization.

Table 11: Average share of applications in common with other job posts, times 100

	Dependent variable:
	Mean share of apps in common
GenAI Treatment Assigned	0.005***
	(0.001)
Constant	0.055***
	(0.001)
Observations	47,931
R^2	0.001

Notes: This table analyzes the effect of the treatment on how many applications job posts share with other job posts. We construct a matrix of all job posts, where the m by nth element is the fraction of the mth job posts' applications which come from a freelancer who also applies to the nth job. For each job post we take the mean of this measure across all other job posts. This is the independent variable. The sample is conditioned on employers who posted a job post which received at least one application. Significance indicators: $p \leq 0.10 : *$, $p \leq 0.05 : **$ and $p \leq .01 : ***$.

but in the treatment group the similarity between job posts is higher. However, this might be because treated job posts get a larger number of applications overall.

5.4 Applicant pools were not worse overall

First, we show that this effect is not driven by a worse applicant pool. It is possible that the job posts induced by the treatment had lower interest from applicants. However, we've already shown that treated jobs actually received more applications, in Table 5. Now we show that those applicants are no worse on average. When an employer collects applications for a job post, the platform recommends some applicants based on their wages, ratings, and employment history on the platform. In Table 12 the outcome of interest is the share of a jobs' applications which came recommended from the platform. Jobs in the treatment group saw a larger share of their applications come from recommended applicants.

Table 12: Effects of generative AI on quality of applicant pool

	Dependent variable:	
	Share of apps recommended	Number of recommended apps
	(1)	(2)
GenAI Treatment Assigned	0.009***	0.534***
	(0.003)	(0.078)
Constant	0.296***	5.007***
	(0.002)	(0.058)
Observations	50,125	50,125
R^2	0.0002	0.001

Notes: This table analyzes the impact of the treatment on the quality of a jobs' applicant pool. Share of apps recommended is the mean of applications a job post receives which the platform flags as recommended. Number of recommended apps is the number of application a job post receives which are recommended. The sample is conditioned on employers who posted a job which received at least one application. Significance indicators: $p \leq 0.10 : *$, $p \leq 0.05 : **$ and $p \leq .01 : ***$.

6 Conceptual Framework

There is a unit mass of would-be employers ("employers") considering posting one job each. There are two periods. In period 1, employers decide whether to post the job and whether to exert effort to be as specific as possible about the skill requirements and the details of the job. If they post the job, then in period 2 they receive applications and decide whether to hire a worker. If they do not post the job, nothing happens in period 2.

6.1 Period 2: The decision to hire

We first describe period 2. Each job j is defined by a location on a Hotelling line, $\theta_j \in (\underline{\theta}, \overline{\theta})$, which reflects the type of skills needed to complete the job. If the employer exerted effort in period 1, they receive N applications, with skills $\{\theta_i\}_{i=1}^N$, drawn iid from $U[\theta_j - \gamma, \theta_j + \gamma]$, where $\gamma > 0$ is a parameter that captures the fact that the employer cannot perfectly describe the skills needed in the job post.[7] If the employer did not exert effort in period 1, they instead receive N applications drawn iid from $U[\theta_j - \rho\gamma, \theta_j + \rho\gamma]$, where $\rho > 1$ captures the fact that exerting no effort to specify the skills required results in a vague job post and thus draws applicants with a wider–and less relevant–set of skills.

Intuitively, exerting effort shrinks the support of the distribution of applicant skills and makes it more likely that the employer will receive an application close to θ_j. An employer

[7] We define $\underline{\theta}$ and $\overline{\theta}$ such that this and subsequent ranges of applications are always interior to $(\underline{\theta}, \overline{\theta})$.

is able to fill the job iff at least one application is within distance $m > 0$ of θ_j. If the employer is unable to fill the job–because they did not receive any application within distance $m > 0$ of θ_j, they receive period 2 utility of 0.

If the employer has at least one such application, they can choose whether to make a hire. If they make a hire, they receive value $v_j \sim G$ from completing the job and pay wage w.[8] They also must pay idiosyncratic utility cost $\epsilon_j \sim U[0,1]$, which reflects various hiring costs like search and screening. Therefore, conditional on being able to hire, they will hire iff $v_j - w - \epsilon_j \geq 0$.

6.2 Period 1: The decision to post

In period 1, employers decide both whether to post the job, $p \in \{0,1\}$ and, if they do post, whether to exert effort, $e \in \{0,1\}$. Posting incurs cost $c > 0$ and effort incurs cost $c_e > 0$. They know v_j, but do not know ϵ_j nor whether they will receive an application sufficiently close to θ_j to be able to hire, so must form expectations over these objects when making their period 1 decisions. In particular, their utility if they post is given by

$$U(p=1,e) = \pi(e,v_j)\left(v_j - w - \mathbb{E}[\epsilon_j | v_j - w - \epsilon_j \geq 0]\right) - c - ec_e,$$

where $\pi(e,v_j)$ is the probability of hiring, which happens if they are able to hire and ϵ_j is sufficiently low relative to v_j. If they do not post, they receive utility 0.

We now compute the objects $\mathbb{E}[\epsilon_j | v_j - w - \epsilon_j \geq 0]$ and $\pi(e,v_j)$. Since $\epsilon_j \sim U[0,1]$, we can write $\mathbb{E}[\epsilon_j | v_j - w - \epsilon_j \geq 0] = (v_j - w)/2$.[9] To obtain $\pi(e,v_j)$, note that this is given by Pr(at least one application is within distance m of $\theta_j | e) \cdot \Pr(v_j - w - \epsilon_j \geq 0)$. The latter term is just $v_j - w$. For the former term, denote an application as θ_i. Assume for now that $e = 1$. Then, this probability can be written as $\Pr_{e=1}(\min_i |\theta_j - \theta_i| < m) = 1 - \Pr_{e=1}(|\theta_j - \theta_i| > m)^N$. Since $\theta_i \sim U[\theta_j - \gamma, \theta_j + \gamma]$, this is $1 - (1 - \frac{m}{\gamma})^N$. Figure 6 shows the intuition for this: the probability of not being able to hire is simply the probability that all N draws fall outside of the shaded area, each of which occurs with probability $1 - \frac{2m}{2\gamma}$. If instead the employer did not exert effort in period 1, then this probability falls to $1 - (1 - \frac{m}{\rho\gamma})^N < 1 - (1 - \frac{m}{\gamma})^N$. Intuitively, if the support from which applications are drawn is wider, the probability of receiving an application within distance m of θ_j is lower.

[8] We assume an exogenous and fixed wage because our experiment only affects a small subset of the market.

[9] We assume for simplicity that $v_j - w \in (0,1)$. This is not a substantively important assumption–it merely simplifies the algebra. More generally, we could write $\epsilon_j \sim U[0, \overline{v} - w]$ where $\overline{v}$ is the upper bound of v_j.

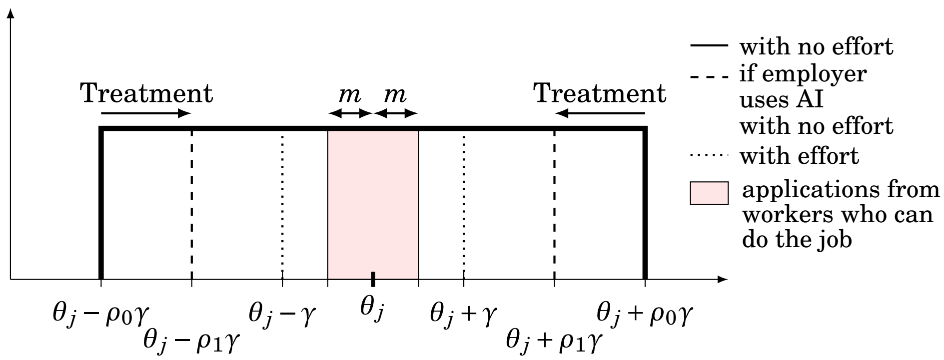


Figure 6: Stylized version of the distribution of applications job post j receives, and effect of the treatment

Thus, plugging these objects in and simplifying, period 1 utility of posting is given by

$$U(p=1,e) = \frac{1}{2}\left[\left(1-\left(1-\frac{m}{(1+e(\rho-1))\gamma}\right)^N\right)\right](v_j-w)^2 - c - ec_e.$$

Note that effort and value of the job are complements: $\partial^2 U/\partial e \partial v_j > 0$. Intuitively, if v_j is higher, then the return to effort in terms of increased likelihood of finding a suitable applicant is also higher.

The employer can choose one of three sets of actions: not post, post without effort, and post with effort. Their choice will be governed by v_j, as shown in Figure 7.[10] For $v_j < \underline{v}_l$, they will not post, where $\underline{v}_l$ is the unique value of v_j such that $U(p=1,e=0;v_j)=0$. Intuitively, if the value of the job is low, it is not worthwhile for the employer to pay the posting cost c. For $v_j \in (\underline{v}_l, \underline{v}_h)$, they will post the job and not exert effort. Intuitively, for these workers the value of the job is high enough to justify the posting cost c, but not so high that the incremental gain from exerting effort to shrink the application pool exceeds the effort cost c_e. Finally, for $v_j > \underline{v}_h$, employers will post the job and exert effort. Intuitively, for very valuable jobs, the increased hiring probability from exerting effort is sufficient to justify the effort cost c_e.

[10]This depiction imposes a technical assumption that the first threshold for v_j is for the employers to post without effort, and the second threshold is that they will post with effort. This assumption is required for the effort choice to have bite: because effort and value are complements, if even the employer on the margin of posting preferred to exert effort, then all employers that post would exert effort (in which case the decision over effort would be irrelevant for the model). This assumptions holds when c_e is sufficiently large–i.e., effort is costly enough that at least some employers that post prefer not to exert effort.

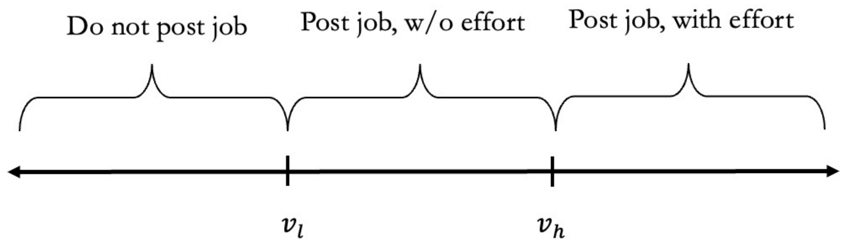


Figure 7: Possible values of v_j and what action the employer takes

6.3 Treatment

We now introduce a technology (AI) that does two things. First, it lowers the cost of posting a job from c_0 to c_1, where $0 < c_1 < c_0$. Intuitively, AI writing software allows employers to spend less time writing a job post. Second, it shrinks the support of the application distribution when the employer does not exert effort by lowering ρ from ρ_0 to ρ_1, where $1 < \rho_1 < \rho_0$. Intuitively, AI writing software clarifies key elements of the job post if the employer's original post was vague, but is still not as precise as the employer would be if they exerted effort to clearly specify the skills required.

Both of these effects cause $\underline{v}_l$ to shift left. The lower cost of posting induces a previously-marginal employer to post as the cost has decreased. As the marginal employer was not exerting effort, the shift in ρ also increases their likelihood of being able to hire and thus further increases the return to posting. Intuitively, the cost of posting has decreased and the probability of hiring has increased, both of which cause employers with lower v_j to post who otherwise would not have.

The reduction in ρ causes $\underline{v}_h$ to shift right. For an employer who was previously indifferent between exerting effort or not, the technology increases the probability that they will be able to hire if they do not exert effort, and thus they now prefer to not exert effort. Employers who have a very high value of v_j will still exert effort as $\rho_1 > 1$–i.e., the incremental hiring probability is still worthwhile paying the effort cost for for very valuable jobs.

Treatment causes changes in the share of jobs that get posted, the likelihood of making a hire conditional on posting, and the unconditional likelihood of making a hire. We can see this in Figure 8, which shows that treatment causes a change for three groups. First, those with $v_j \in (\underline{v}_l^1, \underline{v}_l^0)$ post a job in treatment but not in control. These marginal jobs are less likely to hire than the inframarginal jobs because they are less valuable ($v_j < \underline{v}_l^0$) and so require even lower draws of the period 2 hiring cost ϵ_j.[11] Thus, for these jobs, the share that get posted increases, the probability of hiring conditional on posting decreases, and the

[11]The probability that a job j posted without effort hires is $(1-(1-\frac{m}{\rho\gamma})^N)(v_j - w)$. As v_j is for these marginal jobs is lower than v_j for all inframarginal jobs, this probability decreases.

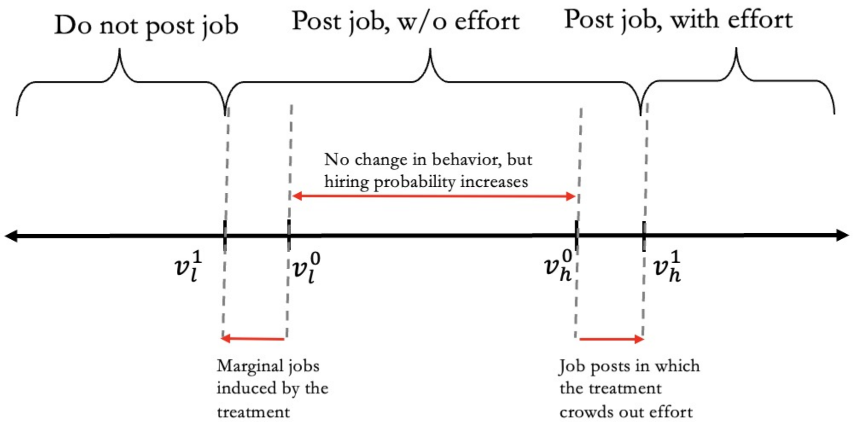


Figure 8: Impact of AI treatment on possible values of v_j and what action the employer takes

unconditional probability of hiring increases.

Second, those with $v_j \in (\underline{v}_l^0, \underline{v}_h^0)$ do not change their behavior–they post without effort in both treatment and control–but their probability of hiring increases as the shift in ρ from the technology increases their probability of finding a suitable applicant.

Third, those with $v_j \in (v_h^0, v_h^1)$ exert effort in control but not treatment. This does not affect the probability of posting because these jobs are always posted. It does reduce the probability for these jobs of making a hire, because the reduction in effort lowers the probability of finding a application with θ_i sufficiently close to θ_j. Thus, for these jobs, the share that gets posted is unaffected, and both the conditional and unconditional probability of hiring decreases.

Combining the previous three ranges of v_j, the model predicts that treatment increases the share of jobs that get posted. The effect to the probability of hiring conditional on posting is ambiguous, and will decrease if the effects to the first and third regions dominate the effects to the second. The effect of treatment on the unconditional probability of hiring is ambiguous. On the one hand, the increase in posted jobs increases the probability of a hire. On the other hand, the probability of hiring conditional on posting a job is lower for both marginal jobs (as they are less valuable) and inframarginal jobs (as some of them stop exerting effort). The net effect to the unconditional probability of hiring depends on which force dominates, which depends on the relative masses of v_j in the two regions as well as the various parameters.[12]

[12]Formally, the effect to the unconditional probability of hiring is given by $\int_{\underline{v}l^1}^{\underline{v}_l^0}(1-(1-\frac{m}{\rho_1\gamma})^N)(v-w)dG(v)+$ $\int \underline{v}l^{0\underline{v}_h^0}(v-w)\left((1-\frac{m}{\rho_0\gamma})^N-(1-\frac{m}{\rho_1\gamma})^N\right)dG(v)-\int \underline{v}_h^{0\,\underline{v}_h^1}(v-w)\left((1-\frac{m}{\rho_1\gamma})^N-(1-\frac{m}{\gamma})^N\right)dG(v)$. The first two terms are positive and the third term is negative. This object could be either positive or negative. For example, if the mass of v in the first two ranges is small relative to the mass of v in the third range, this expression will be

6.4 Welfare

The treatment unambiguously increases employer welfare as they can always choose to ignore the technology. Marginals post jobs who not otherwise have done so, all inframarginals benefit from lower posting costs, and some inframarginal benefit from substituting costly effort towards using the technology instead.

7 Conclusion

We show that job posting is costly– in an experiment run on an online labor market, treated employers who were offered to have an LLM write the first draft of their job post were 20% more likely to post a job.

We find the treatment benefited would-be employers. Treated employers spent over 40% less time to write their job posts, and those resulting job posts received at least as many applications with no worse applicant pools. These positive effects were significantly larger for employers who are not native English speakers. Nonetheless, treated job posts were less likely to make a hire.

Despite the large increase in job posts, the treatment group saw no more hires. We rationalize these results with a model where the treatment induces more job posts, but these marginal job posts are relatively less valuable to employers, and therefore less likely to result in a hire. Additionally, for the inframarginal job posts, the use of AI crowds out effort that employers would have put in themselves—resulting in what are likely more generic job posts.

negative (and vice versa).

References

Barron, John M. and John Bishop, "Extensive Search, Intensive Search, and Hiring Costs: New Evidence on Employer Hiring Activity," *Economic Inquiry*, 1985, *23* (3), 363–382.

Blatter, Marc, Samuel Muehlemann, and Samuel Schenker, "The costs of hiring skilled workers," *European Economic Review*, 2012, *56* (1), 20–35.

den Broek, Elmira Van, Anastasia Sergeeva, and Marleen Huysman, "When the Machine Meets the Expert: An Ethnography of Developing AI for Hiring.," *MIS quarterly*, 2021, *45* (3).

Filippas, Apostolos, Andrey Fradkin, and John J Horton, "Subsidizing Job Search Considered Harmful: Evidence from a Field Experiment (Preliminary)," 2023.

Horton, John J., "Online labor markets," *Internet and Network Economics: 6th International Workshop, WINE 2010, Stanford, CA, USA, December 13-17, 2010. Proceedings*, 2010.

_, "The effects of algorithmic labor market recommendations: evidence from a field experiment," *Journal of Labor Economics*, 2017, *35* (2), 345–385.

Mankiw, N Gregory and Michael D Whinston, "Free entry and social inefficiency," *The RAND Journal of Economics*, 1986, pp. 48–58.

Mok, Aaron, "Don't expect ChatGPT to help you land your next job," Oct 2023. Accessed: 2024-02-22.

Noy, Shakked and Whitney Zhang, "Experimental evidence on the productivity effects of generative artificial intelligence," *Available at SSRN 4375283*, 2023.

Oestreicher-Singer, Gal and Arun Sundararajan, "The visible hand? demand effects of recommendation networks in electronic markets," *Management Science*, 2012, *58* (11), 1963–1981.

Peng, Sida, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer, "The impact of ai on developer productivity: Evidence from github copilot," *arXiv preprint arXiv:2302.06590*, 2023.

Rogerson, Richard, Robert Shimer, and Randall Wright, "Search-theoretic models of the labor market: A survey," *Journal of economic literature*, 2005, *43* (4), 959–988.

Smith, Morgan, "ChatGPT can help you write a standout CV in seconds, job experts say: It's 'the ultimate resume-writing cheat code'," 2023. Accessed: 2024-02-22.

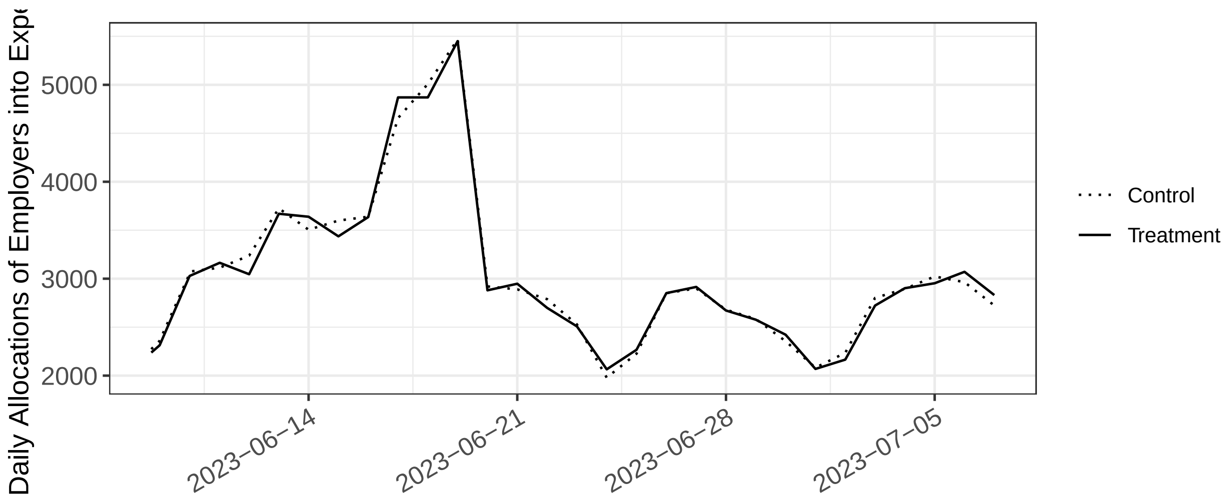
Tambe, Prasanna, Peter Cappelli, and Valery Yakubovich, "Artificial intelligence in human resources management: Challenges and a path forward," *California Management Review*, 2019, *61* (4), 15–42.

Weiss, Daphne, Sunny X. Liu, Hannah Mieczkowski, and Jeffrey T. Hancock, "Effects of using Artificial Intelligence on interpersonal perceptions of job applicants," *Cyberpsychology, Behavior, and Social Networking*, 2022, *25* (3), 163–168.

Wiles, Emma, Zanele Munyikwa, and John Horton, "Algorithmic Writing Assistance on Jobseekers' Resumes Increases Hires," January 2023, (30886).

Zhao, Yingyan, "Your (country's) reputation precedes you: Information asymmetry, externalities and the quality of exports," *Unpublished Paper, George Washington University*, 2018.

Figure 9: Daily allocations of employers into experimental cells



Notes: This plot shows the daily allocations into the treatment and control cells for the experimental sample of 181,962 employers.

A Appendix

A.1 Additional tables and figures

Table 13: Effects of generative AI on number of hires

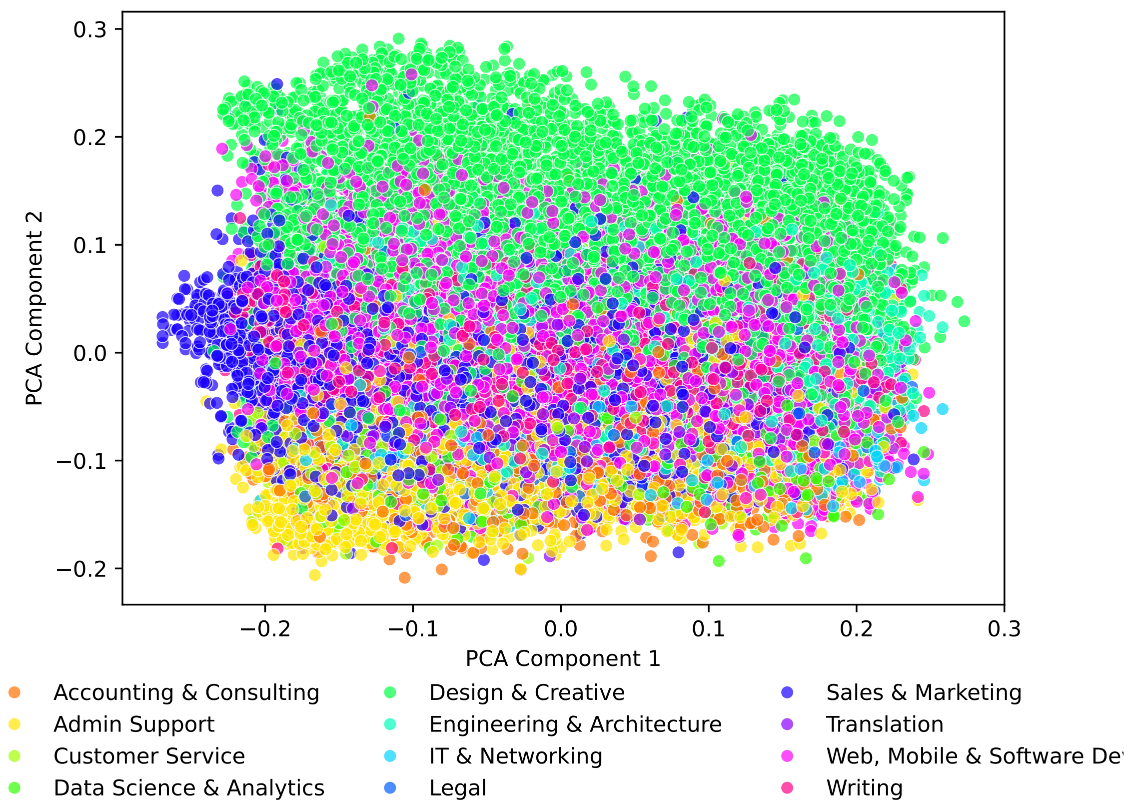
	Dependent variable:	
	Offer, conditional on posting a job	
	(1)	(2)
GenAI Treatment Assigned (Trt)	-0.034^{***}	-0.033^{***}
	(0.004)	(0.005)
Anglophone		0.078^{***}
		(0.005)
Anglophone X Trt		0.004
		(0.007)
Constant	0.220^{***}	0.177^{***}
	(0.003)	(0.004)
Observations	50,125	50,125
R^2	0.002	0.012

Notes: This table analyzes the effect of the treatment on if employer makes an offer. Ever offer, unconditional on post is 1 if the employer makes any offer within 14 days of being allocated into the experiment. Offer, conditional on post is 1 if the job post that the employer was working on when they were allocated into the experiment makes an offer within 14 days. Number of hires is the number of distinct contracts that form as a result of the job post. The sample is made up of all employers in the experimental sample. Significance indicators: $p \leq 0.10 : *$, $p \leq 0.05 : **$ and $p \leq .01 : ***$.

```
<basicSystemPrompt> You are a(n) [platform] client posting a job.

<basicUserPrompt> Based on the following job requirements, write:
# Title
# Detailed job description:
## Around 100 words in length
## List relevant skills with bullet-points
# Choose the most relevant size. Choose one of: 'small', 'medium', or'large'
# Choose the most relevant duration. Choose one of: 'under 1 month', '1 to 3
    ↪  months', '3 to 6 months', or 'more than 6 months'
# Choose the most relevant expertise level. Choose one of: 'entry', '
    ↪ intermediate', or 'expert'

Respond with JSON! Keys should be ONLY 'title', 'description', 'size', '
    ↪ duration', 'expertise'.

Requirements: """ {{requirements}}
```

A.2 Additional tables and figures

Figure 10: Embeddings of Job Posts Reduced to 2 Dimensions, by Job Category



Notes: This plot shows the text of all job posts reduced to two dimensions. We use OpenAI's "text-embedding-ada-002" model to turn the text of job posts into embeddings, and then use PCA to reduce the dimensionality of the embeddings into two dimensions.

A.2.1 Range of average cosine similarities of job posts

The interpretation of the 0.014 effect of the AI treatment on the average cosine similarity of job posts does not have an obvious interpretation. A higher average cosine similarity means a job post is more similar to other job posts in the experiment, a lower average cosine similarity means a job post is more unique. The range of average cosine similarities amongst these job posts is 0.67 to 0.81, with a mean of 0.75. The following are examples of job posts at each of these points.

Most unique: Job post with average cosine similarity of 0.67

What needs to be done: in `main.cpp` file, I am calculating the force for all the particles. I want to optimize it. There are two linked lists in `int main()`. I want you to use them (`list_cell` and `list_particle`) basically, you have to fill the `list_cell` with `list_particle`'s ID. and calculate force (in `compute_acceleration`) only for the particles which are in the same cell and neighbor cells, not for all. That's all. I would like to understand how you did it. Here is the code: (link removed) Run the project: get into the build folder...in terminal type `Make` then: `./md 100 10 0.01` (particles, Time, delta Time) For visualization, just download Paraview and open VTK file.

Mean uniqueness: Job post with average cosine similarity of 0.75

I'm looking for a logo for a small health and wellness company. I am a naturopathic doctor with an emphasis on weight loss counseling. I'm looking to have a design within the next month. I'm looking for a simple and clean logo that is modern yet whimsical and will transfer easily between Instagram, Facebook, a website, and other business materials like treatment plans, recipe books, etc.

Most generic: Job post with average cosine similarity of 0.81

We are looking for a skilled professional to assist us in creating a website and driving business growth. The ideal candidate will have expertise in web development and marketing strategies. The responsibilities include designing and developing a user-friendly and visually appealing website that aligns with our brand image and business objectives. The candidate should also possess knowledge of SEO techniques, social media marketing, and content creation to drive organic traffic and increase conversions. Excellent communication and project management skills are essential for effectively collaborating with team members and delivering satisfactory results within the specified timeframe.

B Second experiment to understand selection into receiving the AI generated first draft

In the previous experiment, employers could choose to opt out of receiving the AI generated job posts. Since these employers were all posting on the platform for the first time, we are not able to investigate which types of employers are selecting to receive help from AI. In order to investigate this selection, we look to another experiment run by the platform, this time run on a sample of employers who'd previously posted at least one job on the platform.

From April 20, 2023, through June 6, 2023, returning employers on the platform who posted a job were randomly allocated into a treatment and control group. The sample included all employers who had ever posted a job on the platform before. For treated employers, any job they post beginning at the time they are allocated into the experiment is considered treated.

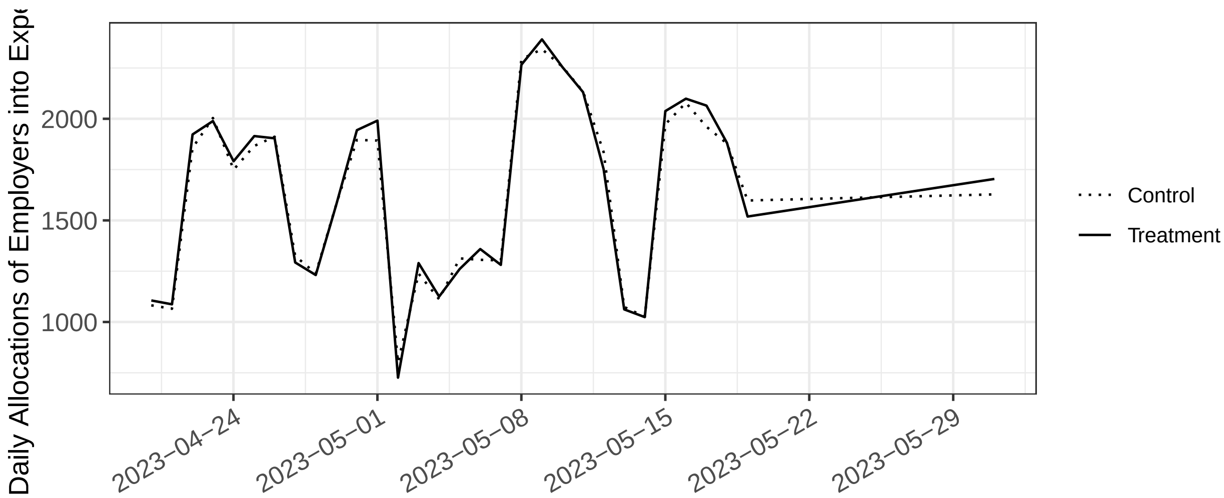
The experimental sample includes 101,601 employers who post 164,382 openings between them. Appendix Figure 11 shows the daily allocations of employers into the treatment and control groups. Table B reports pre-experiment attributes of these employers, and shows the sample of employers was well balanced in terms of the employers experience on the platform.

Table 14: Pre-randomization employer attributes by treatment status

Variable	Control	Treatment	Difference	P_value
From English-speaking country	0.59	0.59	-0.002	0.61
US-based	0.42	0.41	-0.01	0.09
Years since platform registration	3.13	3.13	-0.003	0.89
Num posts, year before allocation	5.49	5.54	0.06	0.25
Num hires, year before allocation	3.40	3.43	0.03	0.36
Hourly wagebill, year before allocation	64,416.28	68,908.11	4,491.83	0.57
FP wagebill, year before allocation	44,184.90	51,323.58	7,138.68	0.67
Total hours demanded, year before allocation	3,925.41	3,241.50	-683.91	0.45
Mean hourly wages, year before allocation	9.81	9.92	0.12	0.39

Notes: This table shows the difference between treatment and control workers for means of pre-experiment covariates, as well as a t-test comparing the difference between those means. Age is defined as the number of years between the employers' registration date and when they were allocated into the experiment. Mean hourly wages is conditional on the employer having made an hourly hire in the year prior to the experiment.

Figure 11: Daily allocations of employers into experimental cells, pilot experiment



Notes: This plot shows the daily allocations into the treatment and control cells for the experimental sample of 101,601 employers.

B.1 Description of data used in the analysis

The dataset we use in this analysis consists of all job posts posted by employers in the experimental sample between the moment they were allocated into the experiment and June 6, 2023 when allocation stopped. We construct job post-level data with all posts, applications, and hires they have within 14 days of posting. While in general we are interested in many outcomes related to posting and hiring, for these purposes we primarily want to 1) show that the take-up in this experiment was comparable with the main experiment and then 2) use the employer histories to understand if there is non-random selection into treatment.

B.2 Experimental intervention at the job description writing stage of job posting

When an employer on the platform wants to post a job, they go through a series of steps. First they provide a job title, the length of time they expect the job to last, and a list of skills required or demanded of the job. After they provide this information, they report some information on their expected budget and then move on to a page where they can input a job description. For employers in the control group, here they type in their job description and then submit the job to be posted.

For employers in the treatment group, after they input the basic information about the job and complete the budget step, they encounter an additional page that asks if they'd like help describing their job. If they select "yes" they have the option to click "Generate job post." The information they have entered so far is incorporated into a prompt, calling a popular generative AI service. The exact prompt is listed below.

```
# Given a job title of '{{title}}'
# Given a job length of '{{duration}}'
# Given job skills of {{skillNames}}
# Write a detailed job description, without a title
# Ask the candidate to submit a proposal
# The candidate should describe how they can help with the project
# The candidate should include some links to past completed projects
```

If the employer is not interested in the service, they click a button that says "I'll write it myself," and they are sent to the basic page employers in the control group would see.

C First stage

Most employers used the AI-generated job posts at least once. The platform records every action and even click taken by each user to the microsecond. Appendix Table15 helps us to see the 'first-stage' of the treatment. Of all employers in the treatment group, 53% opted-in to having the generative AI write their first job post. Of employers who made it through this stage, 62% opted-in. Of the employers who opted in, 78% edited the proposed job description, meaning 22% of employers posted the job without changing anything themselves.

Table 15: Treatment take up

Opted In	Count	Percent	Edited job after opting in
Yes	27,192	53%	78%
No	16,707	33%	NA
Never got this far	7,081	14%	NA

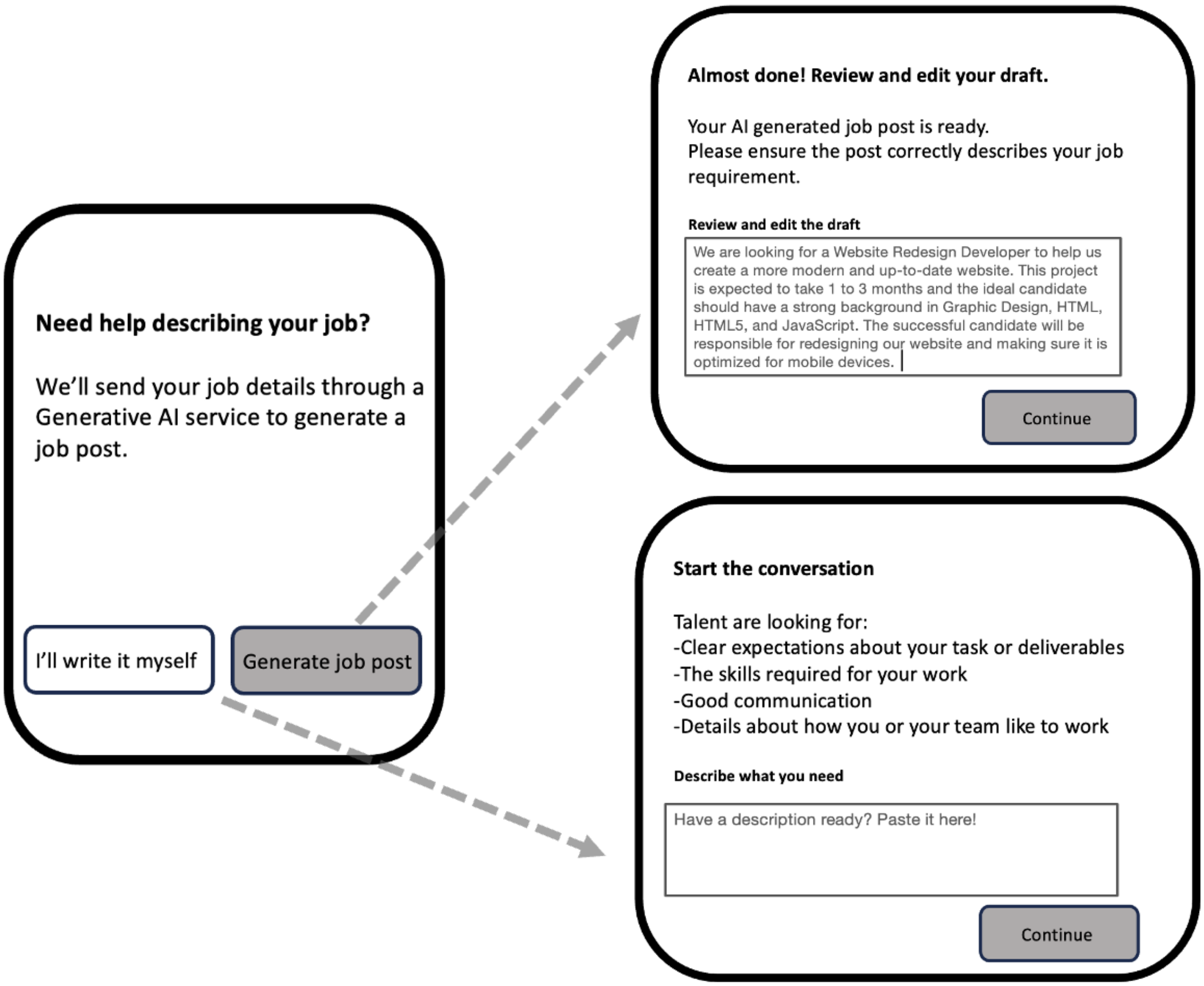
Notes: This table provides summary statistics on employers in the treatment group. "Opting in" means the employer chose to have GenAI generate at least one job post for them. Some employers drop off the job post process before getting to that step, these are labeled 'Never got this far.'

D Results

D.1 Treated employers were more likely to post a job

Treated employers were 10% more likely to post a job. In Table 16 Column (1) we see that on this sample of returning employers, 92% who start a job post end up finishing it. This 10%

Figure 12: The “describe your job post” page in the job post process for the treatment group



Notes: This is a stylized version of the page of the job post process where employers write their job post for employers in the treatment group. For employers in the control group, they only see the bottom page titled “Start the conversation.”

increase is only about half of the size of the treatment effect we saw in the main experiment, which may be because it was run in April 2023, when employers may have been less familiar with LLMs.

Table 16: Effects of generative AI on employer proclivity to post jobs

	Dependent variable:	
	Indicator for if first job is posted	Number of job postings
	(1)	(2)
GenAI Treatment Assigned	0.024^{***}	0.056^{***}
	(0.002)	(0.012)
Constant	0.921^{***}	1.570^{***}
	(0.001)	(0.008)
Observations	101,601	101,601
R^2	0.002	0.0002

Notes: This table analyzes the effect of the treatment on the number of jobs the employer posts over the experimental period. Likelihood of completing first job post is a binary variable for the job post that the employer was working on when they were allocated into the experiment. The sample is made up of all employers in the experimental sample. Number of job posts excludes any spam postings. Significance indicators: $p \leq 0.10 : *$, $p \leq 0.05 : **$ and $p \leq .01 : ***$.

D.2 Employers who opt-in to treatment are slightly positively selected

In Table 17 we compare employers who opted in to the treatment with those who opted out on pre-experiment platform experience. We find that the employers who opted in to receive the AI-written draft are slightly positive selected on observables. This suggests that negative treatment effects to likelihood of hiring in the first experiment are unlikely to be due to the selection of "worse" employers taking up the treatment.

Table 17: Selection into opt-ing into the treatment, from the treatment group

	Dependent variable:			
	Hourly earnings	Fixed price earnings	Hours demanded	Hourly wages
	(1)	(2)	(3)	(4)
Opted-In to GenAI	17,871.920*	18,859.110	1,043.654**	0.177
	(9,304.224)	(34,388.340)	(437.399)	(0.204)
Constant	60,942.970***	44,858.690*	2,844.011***	9.947***
	(7,322.737)	(27,064.780)	(344.248)	(0.160)
Observations	43,899	43,899	43,899	43,899
R^2	0.0001	0.00001	0.0001	0.00002

Notes: This table compares pre-experiment observable characteristics of employers in the treatment group who opt-ed in to the treatment to those who opt-ed out of it. Earnings, hours, and hourly rates are averages calculated from the month prior to when they were allocated into the experiment. Significance indicators: $p \leq 0.10 : *$, $p \leq 0.05 : **$ and $p \leq .01 : ***$.