# More Choices or Help Choosing?: Experimental Evidence on Helping Firms Hire

by

Emma van Inwegen

B.A., University of Washington (2015)

Submitted to the Department of Management
in partial fulfillment of the requirements for the degree of

Master of Science in Management Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2022



Author . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Department of Management
May 6, 2022

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
John Horton
Richard S. Leghorn (1939) Career Development Professor
Thesis Supervisor

Accepted by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Catherine Tucker
Sloan Distinguished Professor of Management
Professor, Marketing
Chair, MIT Sloan PhD Program

# More Choices or Help Choosing?: Experimental Evidence on Helping Firms Hire

by

Emma van Inwegen

Submitted to the Department of Management
on May 6, 2022, in partial fulfillment of the
requirements for the degree of
Master of Science in Management Research

## Abstract

Broadly, there are two main ways to help employers hire: (a) expand their choice set by attracting more applicants or (b) help them choose among that choice set. I report the results of an experiment where employers in a large online labor market were given hiring assistance that could take either form, based on the determination of the helper. In general, job openings with few applicants were given recruiting help, while applicants with many applicants were given selection help. All were given general advice on the hiring process. I find that employers of treated job posts were over 10% more likely to make a hire than those in the control group. While increased recruiting can potentially crowd-out other matches, I find that little if any of the experimental increase was coming at the expense of the control group. In decomposing the reasons for the increased hiring, I find evidence that both (a) and (b) were important, but with recruiting help being about three times more important than selection help. Despite assistance having a marginal cost, the hiring assistance was remarkably cost-effective and a central planner that could tax the wage bill at even just 2% could fund the intervention.

Thesis Supervisor: John Horton
Title: Richard S. Leghorn (1939) Career Development Professor

# Acknowledgments

I am incredibly grateful to my colleagues and fellow graduate students at MIT for their thoughtful comments and constant support. I am especially grateful for the mentorship of my advisor John Horton. This work is dedicated to Greg van Inwegen, from whom I got my love of learning.

# 1 Introduction

Hiring firms expend resources on both (a) creating an applicant pool and (b) assessing those applicants. This process requires the firms to decide the right level of expenditure, as well as the relative allocation of resources between applicant pool expansion and applicant pool assessment. In addition, the firm faces the make-or-buy decision on aspects of the process: should the firm do all of its own recruitment and assessment, or should it buy inputs from recruiting firms and human resource technology? For some tasks, the hiring firm clearly has a comparative advantage and must choose "make" but if some tasks are sufficiently general, than the greater productivity of some third-party at the hiring process might make "buy" the more prudent choice. In addition to whatever resources firms allocate themselves, there is a potentially a policy case for further expenditure by third-parties to subsidize match formation, given the social returns to job formation and employment.

In this paper, I consider the firm's hiring problem primarily via an experiment conducted in a large online labor market.[1] Over 80,000 treated jobs were eligible to receive assistance from an assigned human agent, while a control group of about 10,000 job posts received no assistance. Assistance was not one service applied uniformly, but rather depended on the judgment of the agent about what would be most helpful. The agent could choose from a number of ostensibly helpful actions for the employer—consulting on how to use the platform and recruit, editing a job post, proactively recruiting, reminding the employer about un-interviewed applicants, and recommending applicants. And in cases when the agent believed the job post had enough good applications for the employer to make a hire, this "help" could include nothing at all.

My primary research question is a simple one: does offering third-party assistance—regardless of the precise form it took—help employers make a hire, on average? I find that it does: treated employers were 10.4% more likely to make a hire relative to the control group.

My second research question is *why* the assistance helped. I present a simple framing of the firm's decision problem—how much to invest in recruiting versus selection—and demonstrate that a firm already optimizing its hiring process can benefit more from recruiting assistance than from help in selection. The returns to assistance of various kinds, however, depends on the firm's situation and is ultimately an empirical question. Empirically, the fact that assistance could take many forms makes this question harder to answer. A set up where the agent decides based on endogenous characteristics of the job post and applicant pool makes the experimental intervention realistic, in the sense that any real-world help is likely to take this adaptive form. This kind of intervention is similar in spirit to Bloom, Eifert, Mahajan, McKenzie, and Roberts (2013) who randomly assigned managerial consultants to firms. In my case, the assistance was not about general management, but specifically about recruiting and hiring. However, this adaptive form also makes understanding *why* a particular piece of assistance works more challenging. Further complicating matters, actually receiving assistance—getting recruiting help or having shortlisting help—is *highly* negatively correlated with the employer hiring. This is unsurprising, in that agent effort is concentrated on jobs that likely need the help, but it makes

[1] I use the terms "worker", "job-seeker," "job post," and "application" for consistency with the economics literature and not as a commentary on the legal nature of the relationships created on the platform.

credible non-experimental estimates more difficult to obtain.[2]

I first document that pool expansion was delivered. Both employers and the agent were able to recruit a worker by sending an invitation for them to apply to the job. Job posts in the treatment group received 5 more recruited workers. The treatment also induced employers to do more recruiting of their own. Employers of treated job posts recruited 2.24 more workers on their own, which is surprising given that agent-recruiting could have potentially substituted for employer recruiting. Overall, treated job posts received 2.2 additional applications.

I also document that selection help was delivered. Shortlisting is a feature of the platform which allows employers (or the agents) to flag applications they are interested in, which then brings them to the top of the list of all applications. Treated job posts also had more shortlisted applicants. Unlike recruiting, there is no evidence that the treatment stimulated more employer shortlisting. However, I do find that treated employers conducted considerably more interviews, even among applicants that applied organically i.e., were not recruited. The intervention caused employers to conduct 25% more interviews overall. This could partially reflect a spill-over effect, where the assistance sufficiently improved the quality and depth of the applicant pool that it overcame some fixed cost to doing any interviews at all. It could also simply reflect that part of the assistance reminded employers they had applicants who had not been interviewed.

I find that treatment effects were consistent for both experienced employers and ones who recently joined the platform. This is surprising given the cold-start problem and the literature on new entrants to platforms (Pallais, 2013). It is possible that firms hiring technology is optimal based on their needs, and that assistance from a generic agent would have low, or nonexistent, returns. However I find no difference in the treatment effect to the likelihood of making a hire for employers who are new to the platform or if it is their first job post.

To estimate causal effects of each action taken by an agent on a job deemed worthy of help, my approach is to model the agent's decisions about how to treat a particular job post, and then use that model to estimate local average treatment effects. In broad strokes, jobs that had many applicants when the agent first considered helping received nothing or shortlisting help, whereas those that had few applicants received recruiting help. I fit propensity score models to predict recruiting and shortlisting using the treatment group, partition the two scores into blocks and then compute treatment effects in each "cell." This approach finds effects on pool expansion and pool shortlisting where one would "expect." I then estimate overall treatment effects on hiring by cell. It shows no clear pattern of treatment effects.

As an alternative, I use a series of instrumental variables approaches to try to estimate the effects of shortlisting and recruiting. (1) I use the fact that there are some days where recruiting is more likely to be given and others where shortlisting is more likely to create an instrument for whether or not a given job would receive each type of help. Using this method I find that the treatment effects of recruiting are twice as large as those for shortlisting. (2) I use the variation in agent's proclivity to shortlist or recruit to do a residualized leniency instrumental variables approach, similar to a judge IV (Dahl, Kostøl, and Mogstad, 2014). Using this method I find that between one fifth and one third of the hiring increase is attributable to shortlisting and

[2]This is reminiscent of the famous Ashenfelter's dip that makes job training program evaluation challenging.

the rest to recruiting. Each of these approaches rely on different exogeneity assumptions, but all suggest that recruiting has a larger treatment effect than shortlisting.

This paper makes several contributions. It shows that third-party assistance in hiring—delivered to employers in a realistic manner—can lead to more jobs being filled. I believe I am the first to show this experimentally. While the setting is online, the basic process is of recruiting, hiring, and screening is similar to what is found in more conventional markets—a process that is also increasingly taking place online. It was not *ex ante* obvious that third-party assistance would help, as various active labor market policies have a mixed record at best (Card, Kluve, and Weber, 2010, 2018). If the experimental assistance had not improved hiring, a reasonable criticism would be that firms know their own requirements better than a third-party agent giving generic hiring assistance. However, I find no difference between the treatment effect to new and experienced employers. The strong negative selection of jobs into actually receiving assistance, and the multiple forms assistance can take, makes this large experiment particularly valuable.

The existence of the multi-billion dollar recruiting industry is also evidence that individual firms find recruiting assistance helpful and cost effective, particularly since they can experiment trying to hire with and without this help, and make decisions accordingly. In this experiment, much of the benefit likely comes from pool expansion, which is perhaps unsurprising given the literature on algorithmic recommendation systems (Pellizzari, 2005; Horton, 2017). To the extent recruiting firms are simply competing with each other, the social returns might not be present—it depends on how much this recruiting effort crowds-out other matches and whether the crowded-out matches are inferior (Crépon, Duflo, Gurgand, Rathelot, and Zamora, 2013). To test for this I look at the variation in fraction of job posts treated by day, and find that treatment effects are larger on days with a higher percentage of the market treated. Furthermore, I show that there are things that "work" that are not subject to this crowd-out critique—namely applicant pool processing improvements. For example, the treatment clearly induced employers to interview more candidates that had already applied organically.

The approach discussed in this paper does require human intervention, which means marginal cost is not minimal, as in other algorithmic approaches, such as Belot et al. (2018). Whether these process improvements can be delivered at scale with algorithmic approaches is an open and interesting question. Despite having a marginal cost, the ROI was remarkably high, in that a social planner that could tax the wage bill even less than 2% would find it cost-effective to provide the subsidization.

In Section 2 I describe the marketplace. In Section 3 I give details on the experiment. In Section 4 I present a conceptual framework for understanding the hiring process. In Section 5 I give the main results of the treatment to hiring outcomes. In Section 6 I explore why and which parts of the treatment caused the increase in hires. Section 7 concludes.

## 2 Empirical context

This experiment was conducted on an online labor market. In online labor markets, employers 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 the employers and workers. Some common services provided by platforms include soliciting and promoting job openings, hosting profile pages, processing payments, certifying worker skills, and maintaining a reputations system (Horton, 2010; Filippas et al., 2018).

In the platform which I use as my empirical setting, employers post job openings on the platform website with job descriptions, required skills, and scope of project. The employer then categorizes the job, 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 gives an estimate for how many hours they expect the job to take. The experimental sample only includes hourly jobs, which can take a few hours or many months of full time work to complete.

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 which 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 out about workers 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, some of the information available to the other side of the market is verified by the platform. Employers are particularly interested in past experience on the platform (Pallais, 2013), and generally are looking for signals to overcome information asymmetries (Stanton and Thomas, 2015).

When a worker chooses to apply to a job opening, they submit an application with a cover letter and an hourly wage bid or a total project bid for fixed price jobs. As the employer collects worker applications, they can choose to interview applicants and eventually make an offer. When workers make an offer, employers can make a counteroffer for the wage, however about 90% of hired workers are hired at the wage they initially proposed (Barach and Horton, 2021). To complete the work on hourly jobs, workers install custom tracking software that serves as a digital punch clock. The software records not only the time spent working, but also keystroke count and mouse movements. The software also captures images of the worker's computer screen at random intervals. 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.

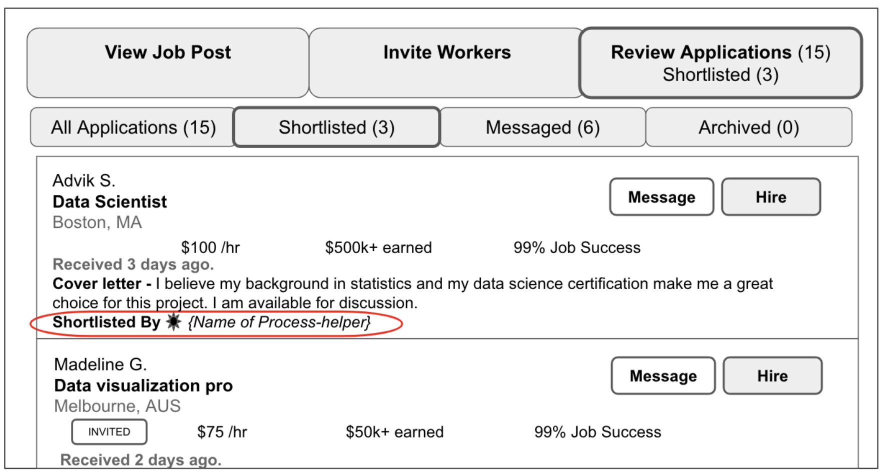
## 2.1 Hiring on the platform

The hiring process on the platform is broadly similar to the process in a conventional labor market. A would-be employer posts a job opening with a job title, job description, list of desired skills, and category of work. Once the employer submits the job opening, the platform reviews it and posts it publicly to the marketplace. Once the job begins to receive applications, the employer can view all applications.

Figure 1 shows a stylized version of the interface. In the first tab, the employer could view the job they posted. Figure 1a shows the second tab, where employers can search the platform for available workers and invite desirable candidates. Note that employers can see each workers' wage bid, name, self-reported skills, and a few pieces of platform-verified information, such as total hours worked and average feedback rating from previous projects (if any). Figure 1b shows the third tab, where employers are able to either view all of their applications or only their "shortlisted" applications. The employer can screen applications by asking applicants questions through the messaging feature on the interface and by organizing phone or video interviews. After screening, the employer decides whether or not to make an offer(s). In the control group 32% of job posts lead to a hire, suggesting there is a need for employers to search for good matches.

Figure 1: Job post manager on the platform, stylized version


(a) Inviting workers to apply for a job




(b) Proposal manager with shortlisting feature

# 3 Experiment design

In March 2020, the platform ran a six month experimental evaluation of the hiring assistance program. The platform had a version of this hiring assistance program in operation for years prior, but ran this experiment to test its efficacy. The sample included all hourly job openings over the experimental period which require 30 or more hours of work per week for at least 3 months, totaling 88,208 job posts.

After being allocated into the experiment, 90% of job posts were randomly assigned to the treatment cell and 10% to the control cell. The unit of randomization was the job post. As such, it is possible employers who posted multiple jobs over the course of the experiment could have had hiring assistance for one job but not another. Neither employers nor workers were aware of the experiment.

Table 1: Description of various actions taken by those assisting treated employers

| Overview of the action | | # of Job Posts (% of Treated Job Posts) |
|---|---|---|
| **Initial contact by Process-helper** | | |
| | After posting a job, employers received a message with an introduction by the Process-helper. The Process-helper reports their hours and availability and a short description of how they can help. | 81,234 (100%) |
| **Reply by the firm / hiring manager** | | |
| | Employers did not need to respond. However, some did. Even if they did not, the job post was still eligible for further help. | 27,619 (34%) |
| **Phone consultation with Process-helper** | | |
| | Employers could schedule a phone call with Process-helper to discuss the job post and get help identifying workers to invite. | 1,835 (2%) |
| **Worked by Recruiter/Shortlister** | | |
| | Worked jobs were jobs that a Process-helper judged if and what sort of help it needed. Jobs were worked in order of arrival when demand outstripped supply. | 48,564 (61%) |
| **Invited workers** | | |
| | For job posts with insufficient high quality applications the Recruiter/Shortlister invited on average nine workers to apply to the job. | 15,745 (19%) |
| **Recommend workers** | | |
| | For job posts with a sufficient number of high quality applications, the Recruiter/Shortlister shortlisted on average the three best fitting applicants to a prominent place on the interface. | 20,890 (26%) |

*Notes:* This table reports counts of job posts which received each type of treatment. The sample is all job posts assigned into the treatment cell of the experiment. The counts and frequencies in the last column refer to the job posts which received each piece of the treatment, although all jobs in this sample were eligible for all.

## 3.1 The nature of assistance

Employers in the treatment were eligible to receive several different forms of assistance. Table 1 lists the various kinds of assistance and uptake. Treated posts are assigned a "Process-helper" who engages directly with the employer, and a "Recruiter/Shortlister" who highlights top applications and recruited workers they believe to be a good fit for the job. The Process-helper fields questions from the employer about things like navigating the interface and applicant quality. They also send messages reminding the employer to begin inviting and later, interviewing workers to the job. Through this messaging system the employers can also share specifics of what they are looking for in a hire beyond what they had written in a job post. These are taken into account by the Recruiter/Shortlister, who uses the platform search feature to invite applications from workers who have not applied for the job but they judge to be a good fit.

Both Process-helpers and Recruiter/Shortlisters specialize in one of three categories of job posts; Web, Mobile, and Software Development, Design and Creative, or General. It might take a different skill set to hire someone for graphic design work than it does to hire someone to build a database, and so both the Process-helper and the Recruiter/Shortlisters assist in the hiring for job posts only in one category.

### 3.1.1 Help using the platform

When the platform designated a job post treated, the firm received an invitation for assistance in their hiring for that job. If the firm accepts or does nothing, the platform assigns them a Process-helper who will send the employer a greeting and an offer in the platform-provided messaging system and over email. Once an firm is assigned a Process-helper for a job post, that person helps the firm through the hiring process until a hire is made or the firm retracts the job post. The text of the initial message is in Appendix A.

### 3.1.2 General guidance

The Process-helper may edit the job post to make it more clear or attractive to workers. They also respond to any emails or messages from the firm regarding the hiring process or navigating the interface. When an employer reaches out with a question, if their designated Process-helper is unavailable they are able to speak to any available Process-helper within their job post category if they do not want to wait. If the employer does not respond and has not scheduled any interviews three days after they publish the job post, the Process-helper will send them a follow up email encouraging them to start interviewing applicants. At first contact Process-helper also offer to have a phone call with the employer where the Process-helper helps them to identify and invite top workers. Even if the employer doesn't respond to the initial message (as 66% of them don't) the Process-helper sends follow up messages as applications to the job post are submitted, reminding the employer to interview candidates.

### 3.1.3 Recommending applicants from the existing pool, or "shortlisting"

The Process-helper sends the job post to a "Recruiter/Shortlister." This Recruiter/Shortlister reads the job post to understand the needs of the job and the skills required to successfully complete the project. The Recruiter/Shortlister can also read the message history of the Process-helper and the employer to look for any particular requests that the employer has made about what they are looking for in an applicant.

The Recruiter/Shortlisters look through the available worker applications and check to see if the employer has been inviting or interviewing applicants. They take one of two actions depending on the quality and fit of those workers who applied organically and whether or not the employer is actively engaging in searching. They either (1) shortlist organic applicants or (2) invite more applicants.

If there are a sufficient number of high quality candidates, in terms of their skills, platform history, and reputation, the Recruiter/Shortlister will recommend (at the median) three of the best fitting workers' for hire. These "shortlisted" workers' applications then appear prominently in the job post's application manager page, and include the text "Shortlisted by: {Name of

Process-helper}" to distinguish from shortlisting that the employer might do themselves (see Figure 1b).

#### 3.1.4 Recruiting additional applicants

If the Recruiter/Shortlister does not believe there are good fitting organic applications, they will invite workers to apply for the job. For job posts which the Recruiter/Shortlister deem to have too few high quality applications, the Recruiter/Shortlister uses the platform search feature to invite workers. If the job then receives new applications, the Recruiter/Shortlister may follow up by shortlisting. At any point during this process, the employer can pick applicant(s) to interview or hire. If they do not choose to hire or to take the job post down within 30 days of posting, the platform takes the job post off of public view.

### 3.2 Prevalence of various kinds of assistance

About a third of treated employers engaged with the Process-helper by responding with at least one message. About a fifth of job posts received invitations to workers by the Recruiter/Shortlister, and a quarter of job posts received shortlisted candidates. Note that 59% of treated job posts never received either invited nor shortlisted workers. This reflects both the jobs that the Recruiter/Shortlister could not work in time and the job posts that the Recruiter/Shortlisters judged were likely to receive sufficient applications without help.[3]

Jobs with fewer applications were more likely to receive recruiting services and jobs with more applications were more likely to receive shortlisting. Figure 2 shows the distributions of treated job posts which received recruiting vs shortlisting. Note that about a third of jobs which received recruiting later got shortlisted candidates, so the same job may appear in both distributions.

### 3.3 Internal validity and delivery of the assistance

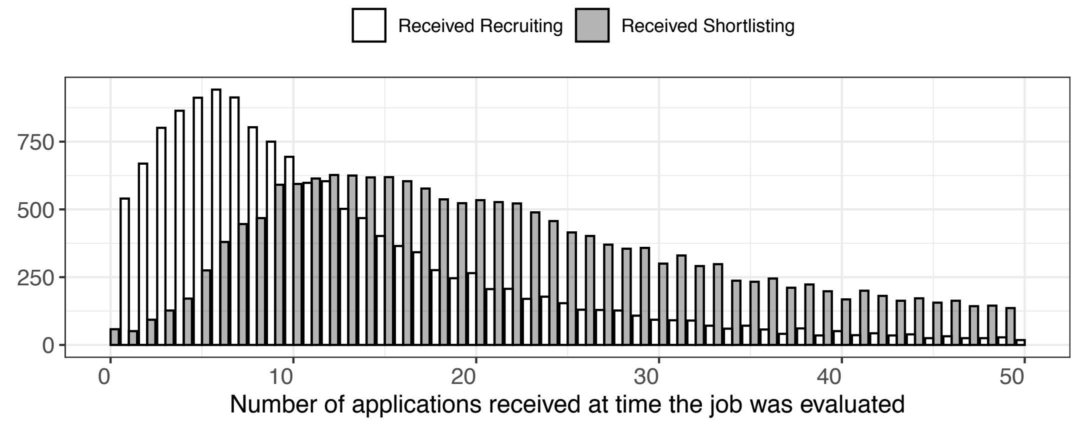
On average 459 job openings were allocated to the treatment and 47 to the control on a given day. The weekly count of job posts allocated to each cell over the course of the experiment is shown in the top panel of Figure 3. The y-axis is counts of jobs posts on a log scale. These allocations track closely—regressing treatment on week dummy variables gives an F-statistic of 0.87 and a p-value of 0.65.

To assess the effectiveness of randomization, in Appendix B I report the mean values and t-tests for various pre-randomization attributes of the job post, and I obtain excellent balance on these covariates. This is unsurprising, as the software used to allocate job posts into treatment groups has been used for many experiments and has proven reliable.

Figure 3 also shows that the assistance was delivered. The second panel from the top shows that treated job posts were almost three times as likely to shortlist at least one applicant. The third panel shows the average number of applications per job post over the course of the

[3] Appendix Section C gives details on the substantial differences between jobs that received help and those that did not; namely, that jobs which received shortlisting or recruiting by the Recruiter/Shortlisters are "worse" on all observables.

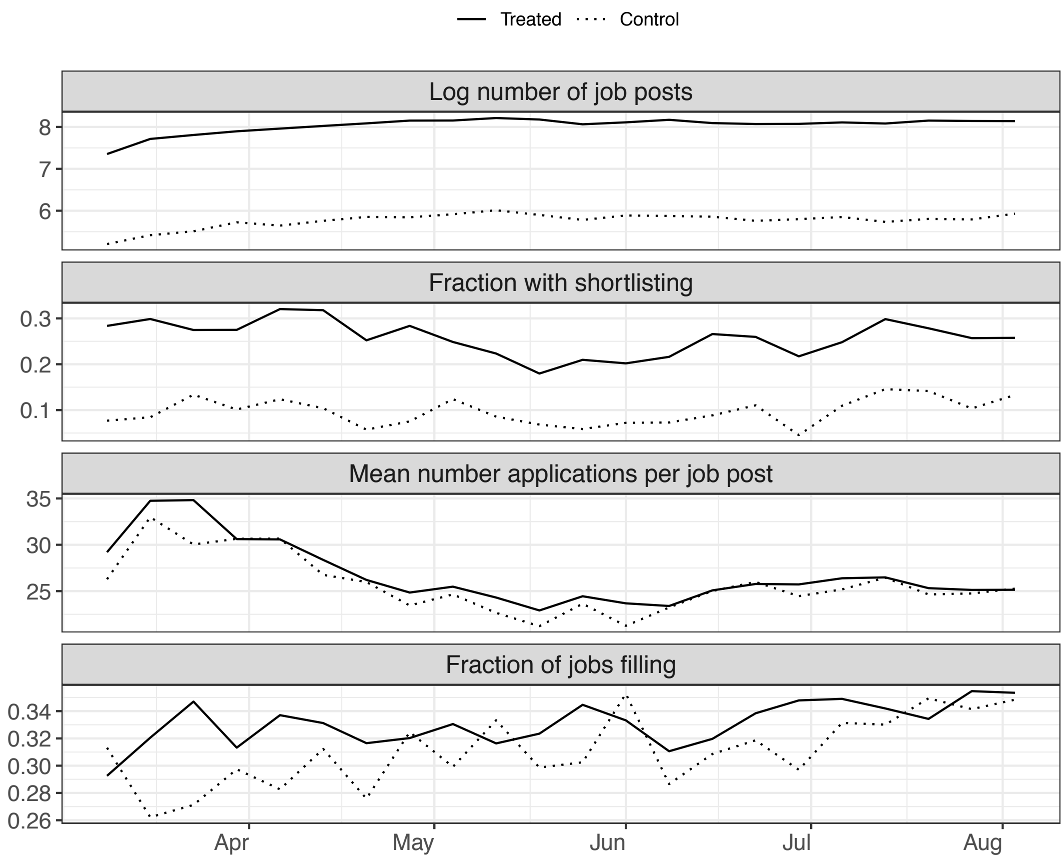
Figure 2: The distribution of applications received at the time the Recruiter/Shortlister first assessed a job post, by what action the Recruiter/Shortlister then took (recruiting or shortlisting)




*Notes:* This sample is job posts in the experimental sample which are assigned treatment. The x-axis is the number of applications a given job post has at the time a Process-helper first sees the job. The solid bars are the density of job posts which receive recruiting help, while the dashed bars are the density of job posts which receive shortlisting.

experiment, with treated job posts consistently receiving more applications. Previewing a main result, the finally panel shows the fraction of jobs which made a hire. Even in this week-by-week breakdown, it is visually clear that treated job posts were consistently more likely to make a hire.

Figure 3: Allocations to the treatment and control cells per week, along with mean number of applications, job-post fill rates and job-post shortlisting over time




*Notes:* These plots show posting and hiring outcomes by treatment cell for each week of the experimental period. The first facet shows the log number of job posts. The second shows the fraction of jobs which shortlist at least one applicant. The third shows the mean number of applications per job post, and the fourth shows the fraction of the posted jobs which led to a hire.

# 4 Conceptual framework

What are the returns to an employer from having more applicants versus being able to better assess the applicants they have? Although in the experiment there are multiple actions that can be taken, in broad strokes, they fall into two categories: (1) recruiting more applicants and (b) helping asses applicants. From a whole system perspective, increasing the screening technology is attractive because it means applications are used more efficiently. Increasing applicant counts on a per-job basis—while helping with a hire—likely increases crowd-out and increases application costs. And there is a concern that it just draws applications away from other job posts.

Consider an employer receiving a collection of $A$ applicants to create some output, of value $v$. Workers have two types for each specific job, $H$ and $L$, with fraction of $H$ types in the population being $\theta$. Workers do not know their own type, and all bid $w$. Only $H$ types can create the output if hired; if a $L$ type is hired, they still have to be paid but the employer does not get $v$. The employer cannot observe the type directly, but instead receives a binary signal about how well fit an applicant is for the job, $s \in \{0, 1\}$, through the screening process. High-type applicants signal their type with some $p > 0$, $\Pr\{s = 1|H\} = p$ . There are no false positives, i.e., $\Pr\{s = 1|L\} = 0$. The employer will not hire without a positive signal, as $\theta v - w < 0$.

The probability the employer fills their job is thus

$$\begin{aligned} h(p, A) &= \Pr\{\text{Hires}\} \\ &= 1 - (1 - p\theta)^A \\ &\approx 1 - \exp(-p\theta A). \end{aligned} \tag{1}$$

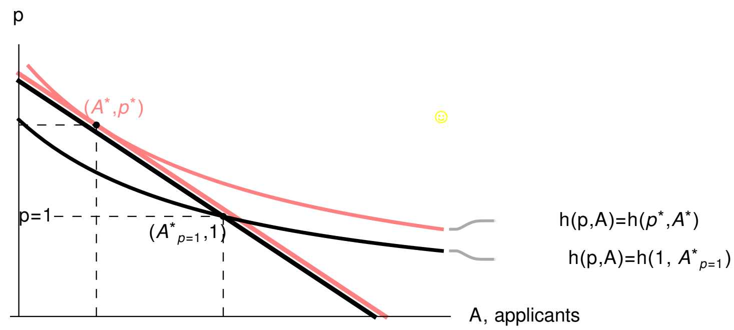
The returns to better screening technology and more applicants are both positive. And each has positive cross-partials—an additional applicant is worth more, all else equal, when the screening technology is good; the marginal benefit of improved screening technology is increasing in the number of applicants it can be used on. However, improving either is not free. Suppose it costs the employer $c_A A$ to have $A$ applicants and $c_p p$ to have a screening technology of quality $p$. Ignoring integer problems in $A$ and the $[0, 1]$ bound on $p$, at an interior solution for optimal investment in both, I have

$$\left| \frac{\partial h(p, A)/\partial A}{\partial h(p, A)/\partial p} \right| = \frac{c_A}{c_p} = \frac{p}{A}. \tag{2}$$

or simply that the marginal increase in hiring from improving either the pool or the screening technology equals marginal cost. This interior solution also equalizes expenditure—the amount spent on screening technology equals the amount spent on recruiting. Figure 4 shows this solution in pink, with an implied budget line equal to $p^* c_p + A^* c_A$, which is tangent to the iso-hire curve at the optimal solution, $h(p, A) = h(p^*, A^*)$.

Of course, an interior solution is not always feasible, as $p$ is bounded above by 1. If at the optimal level of total investment, $B^*$, if it is the case that $c_p < B - c_A A$, then the optimal

Figure 4: Optimal expenditure on screening and recruiting, with and without constraints on screening probability




$p^* > 1$, which is impossible, and so the employer has to solve a constrained optimized problem. This constrained optimization problem is illustrated in black, in Figure 4. Note that at this constrained solution with $p^* = 1$, I can have greater effort on recruiting (higher $A^*$) because this is the margin that is still available. The iso-hiring curve shifts in relative to the infeasible curve—it necessarily does so because the employer always had the option of this $(A^*, p^*)$. Note also that the implied budget line is (a) no longer tangent to the optimum point, and (b) is shifted in, as the optimal expenditure on recruiting is lower, given the $p^* = 1$ constraint. The hiring firm does not actually have a true budget, but rather the spending on hiring is endogenous.

Connecting this framing back to the experiment, so long as a firm is not at the $p^* = 1$ corner solution, the fill rate can improve from either recruiting assistance or shortlisting assistance. If the firm is at $p^* = 1$, then recruiting assistance can help, though marginal returns might be low because of the firm's endogenous choice of relatively large $A^*$. A firm with a higher $v$—a greater pay-off to hiring—will organically invest more in both recruiting and screening—making the marginal returns to assistance on hiring lower, but the financial returns would be higher given the greater pay-off to hiring.

# 5 Results

I begin by examining whether being assigned to the treatment cell impacted whether the employer made a hire. I estimate

$$y_i = \beta_0 + \beta_1 \textsc{Assigned}_i + \epsilon_i,$$

where $y_i$ is the outcome of interest for job $i$, $\textsc{Assigned}_i$ is an indicator for whether or not the job was assigned to be in the treated group. The estimate of this regression is reported in Column (1) of Table 2. Jobs assigned to the treatment were 0.02 percentage points more likely to fill, or 5.3% more likely relative to the control group.

This Column (1) likely understates the benefits. One aspect of the assistance program is that sometimes the number of job posts assigned to be treated was too large for all of them to receive the treatment. A job post appears in a queue to Process-helpers and Recruiter/Shortlisters based

Table 2: Effects of treatment assignment on whether an employer hires, the number of applicants for the job, the number of workers invited to the job, and the number of shortlisted candidates

| | *Dependent variable:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Hired | | Num Applications | | Num Invites | | Num Shortlists | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Treatment Assigned | 0.02*** | | 1.14*** | | 2.57*** | | 0.64*** | |
| | (0.01) | | (0.28) | | (0.20) | | (0.04) | |
| Treatment Received [IV] | | 0.03*** | | 2.24*** | | 5.06*** | | 1.26*** |
| | | (0.01) | | (0.55) | | (0.40) | | (0.04) |
| Constant | 0.32*** | 0.32*** | 24.47*** | 24.25*** | 3.85*** | 3.36*** | 1.35*** | 1.22*** |
| | (0.01) | (0.01) | (0.27) | (0.32) | (0.08) | (0.10) | (0.03) | (0.03) |
| Observations | 88,224 | 88,224 | 88,224 | 88,224 | 88,224 | 88,224 | 88,224 | 88,224 |
| Adjusted $R^2$ | 0.0001 | -0.01 | 0.0002 | 0.02 | 0.0002 | 0.001 | 0.002 | 0.03 |

*Notes:* This table reports effects of treatment to whether or not a job post filled, its number of applications, invitations, and shortlists. Sample is experimental sample of high value posts from March 13, 2020 to August 31, 2020. OLS specification regresses outcome on treatment assignment. IV specification uses treatment assignment as an instrument for employers receiving the treatment. Heteroskedasticity-robust (Eicker–Huber–White) standard errors are used in all specifications. Significance indicators: $p \leq 0.05$ : *, $p \leq 0.01$ : **, and $p \leq .001$ : ***.

on the time it is posted and they have to work jobs in the order they appear in the queue. The jobs that the Recruiter/Shortlisters could not get to in time were considered not worked, as they were neither evaluated nor helped. As I am predominantly interested in the effect of *receiving* the treatment, in my preferred specification I will use treatment assignment as an instrument for being evaluated (and hence potentially helped). Conditional on being assigned to the treatment, which job posts were evaluated was primarily a function of how many job posts were added to the program that week.

Whether a job was evaluated was recorded, and so I can construct a variable RECEIVED. I then estimate the following IV regression,

$$\textsc{Received}_i = \gamma_0 + \gamma_1 \textsc{Assigned}_i$$
$$y_i = \alpha_1 + \beta_1 \widehat{\textsc{Received}}_i + \varepsilon_i$$

where $y_i$ is the outcome of interest for job $i$, $\textsc{Received}_i$ is an indicator for whether or not the job $i$ was worked by a Recruiter/Shortlister. A job being "worked" simply means a Recruiter/Shortlister went over it and decided whether or not to provide additional assistance, it does not require them to have actually given any assistance.

Treated job posts were 10.4% more likely to make a hire. In Column (2), note that job posts in the control group had a base fill rate of 0.32% and the treatment caused an increase of 0.03 percentage points, or about a 10.4% increase in likelihood of filling the job.

## 5.1 Effects on the size of the applicant pool

Treated job posts received 9.2% more applications. In Table 2 Column's (3) and (4), the outcome is the number of applications the job post received. Job posts in the control group received 24 applications, on average. The treatment caused job posts to receive 2.2 additional applications.

Job posts receiving the treatment sent 5 more invitations to workers. In Table 2, Columns (5) and (6), the outcome is the number of invitations sent to workers, either those sent by the employer or by the Recruiter/Shortlister. In the control group, the employers send on average 3 invitations, while employers in the treated group's job posts sent 5 more.

Job posts which received Recruiter/Shortlister invited workers induced applications from at least one of the invited workers 90% of the time. This increase in the pool size presumably could explain part of the increase in whether a job opening is filled. However, it seems improbable that this is the sole reason if recruited applicants were "average." A 9.2% increase in the number of applications leading to an 10.4% increase in hiring seems improbable if the added applicant is one more "draw" from the kinds of job applications the job would receive organically. As such, it seems likely that the marginal proposal induced by the treatment is of particular interest to the firm. I show in Appendix Section D that in fact workers recruited by the Recruiter/Shortlisters have a history of more hires, higher earnings, and have gotten more recruiting invitations prior to the experiment than the rest of the applicant pool.

## 5.2 Effects to narrowing the choice set

Treated job posts which received the treatment shortlisted one additional applicant, double the shortlisting done in the control group. In Table 2 Columns (7) and (8) the outcome is the number of applicants shortlisted by either the employer or the Recruiter/Shortlister. Treated jobs shortlisted 1.26 more applicants, much smaller than the expected 3 more applicants if all job posts which received help got shortlisting. However, many treated job posts received recruiting and no shortlisting at all, hence the small difference between the average number of shortlisted applicants in the treatment and control group.

## 5.3 Effect of the treatment on employer behavior

I now examine how employers of treated job posts changed their behavior in response to the treatment. In Table 3 I use the same OLS and IV regressions as in Section 5. Because I am most interested in the effect of the treatment on the treated, I will focus on the IV estimates in this section.

Almost half of the increase in total invitations to workers came from the employers. In Column (2) I show that employers of treated job posts invited 2.24 more workers to apply for the job than in the control group, a 62% increase. These worker invitations were sent by the employers themselves, not by the Process-helpers. Recall that Table 2 Column (4) showed that in total treated job posts sent 5 more invitations to workers. Therefore half of the effect to recruiting intensity came from the employers themselves.

Employers of treated job posts did not do any more shortlisting than those in the control group. In Table 3 Column (4) I show that employers in the control group shortlisted on average just over one application per job post. Employers in the treated group behaved the same.

Employers in the treated group gave more interviews than those in the control group. In Column (6) the outcome of interest is the number of interviews employers gave to workers

Table 3: Effects of treatment assignment to the number of employer shortlists, invites, and interviews

| | *Dependent variable:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Invites | | Shortlists | | Organic Interviews | | All Interviews | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Treatment Assigned | 1.14*** | | −0.03 | | 0.12*** | | 0.35*** | |
| | (0.20) | | (0.04) | | (0.04) | | (0.05) | |
| Treatment Received [IV] | | 2.24*** | | −0.05 | | 0.24*** | | 0.68*** |
| | | (0.40) | | (0.07) | | (0.07) | | (0.09) |
| Constant | 3.82*** | 3.60*** | 1.21*** | 1.22*** | 1.82*** | 1.80*** | 2.77*** | 2.70*** |
| | (0.08) | (0.10) | (0.03) | (0.04) | (0.04) | (0.04) | (0.04) | (0.05) |
| Observations | 88,224 | 88,224 | 88,224 | 88,224 | 88,224 | 88,224 | 88,224 | 88,224 |
| Adjusted $R^2$ | 0.0000 | 0.0000 | -0.0000 | -0.001 | 0.0001 | 0.004 | 0.0004 | 0.01 |

*Notes:* This table reports effects of treatment to employer invites, shortlists, and interviews. Organic interviews are for workers who applied withough being invited, and all interviews sums organic interviews and interviews from invited workers. Sample is experimental sample of high value posts from March 13, 2020 to August 31, 2020. OLS specification regresses outcome on treatment assignment. IV specification uses treatment assignment as an instrument for employers receiving the treatment. Heteroskedasticity-robust (Eicker–Huber–White) standard errors are used in all specifications. Significance indicators: $p \leq 0.05$ : *, $p \leq 0.01$ : **, and $p \leq .001$ : ***.

who applied without being invited ("organically".) Employers in the treated group interviewed 13% more workers who applied organically than in the control group. In Column (8) the outcome is the number of interviews employers gave to workers invited by themselves or by the Recruiter/Shortlister. Employers in the treated group interviewed 25% more invited applicants than in the control group. An increase in interviews could be simply due to the larger applicant pool, however I also see an increase in employer interviews to applicants who were not invited by the Recruiter/Shortlister. This effect could be a result of the messages from the Process-helpers reminding employers to conduct interviews. It could also be a spillover effect of the larger applicant pool as employers interviewed more workers from Recruiter/Shortlister invited applications.

## 5.4 Are these results driven by crowd-out?

It is possible that successfully recruiting an applicant for one job leads them to not apply for some other job. This "crowd-out" would attenuate the fill rate effect up as one more application to a treated job could take away an application to a control job. In that case the program is merely changing which jobs fill, instead of increasing the number of contracts. One argument against the fill rate effect being driven by crowd out is the effects to shortlisting and interviewing. Better screening of a treated job post should not impact hiring outcomes in the control group. Furthermore, the concern with crowd-out is generally about experiments which take place on a small percentage of the population which then get rolled out across the entire population, only to find the treatment effects disappear. Therefore, in Table 4 I collapse the data to the daily level and calculate the fraction of job posts which get the treatment and the fraction of job posts which eventually make a hire for each day. As fill rates are significantly higher on days where a higher fraction of jobs are treated, I can be less concerned that crowd out is driving the treatment effects. This effect does not seem to be driven by the number of jobs posted that day

Table 4: Effects of higher dose of treatment on daily fraction of jobs which made at least one hire

| | Fraction of jobs which made a hire | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Fraction treated | $0.23^{**}$ | $0.20^{**}$ | $0.21^{**}$ |
| | (0.10) | (0.10) | (0.10) |
| Number of Job Posts | | $-0.0002^{***}$ | 0.0000 |
| | | (0.0000) | (0.0001) |
| Software job | | | $-0.06^{***}$ |
| | | | (0.01) |
| Design job | | | 0.01 |
| | | | (0.02) |
| Constant | 0.14 | $0.20^{**}$ | $0.17^{**}$ |
| | (0.09) | (0.09) | (0.09) |
| Observations | 522 | 522 | 522 |
| Adjusted $R^2$ | 0.01 | 0.05 | 0.11 |

*Notes:* This table reports the relationship between the fraction of job posts treated and the fraction filled per day. In Column (2) we control for the number of job posts each day. And in Column (3) we control for job category. The leave out group is the general category, and we included dummy controls for the Software category and the Design category. Data are collapsed to the daily level. Significance indicators: $p \leq 0.05$ : *, $p \leq 0.01$ : **, and $p \leq .001$ : ***.

or by job category, as seen in Columns (2) and (3) respectively.

## 5.5 Measuring return on investment

Table 5: Effects of treatment on total wagebill in first 30 and 90 days after posting

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Wagebill, 30 day | | Wagebill, 90 day | |
| | (1) | (2) | (3) | (4) |
| Treatment Assigned | $39.98^{***}$ | | $83.49^{***}$ | |
| | (9.55) | | (30.26) | |
| Treatment Received [IV] | | $78.60^{***}$ | | $164.14^{***}$ |
| | | (18.80) | | (30.26) |
| Constant | $243.17^{***}$ | $235.45^{***}$ | $707.17^{***}$ | $691.04^{***}$ |
| | (8.85) | (10.59) | (28.42) | (28.42) |
| Observations | 88,224 | 88,224 | 88,224 | 88,224 |
| Adjusted $R^2$ | 0.0001 | -0.01 | 0.0001 | -0.003 |

*Notes:* This table reports effects of treatment to total wagebill. Wagebill is the total spend on all hires in the first 30 or 90 days after posting, respectively. Wagebill is considered to be zero for job posts which never make a hire. Sample is experimental sample of high value posts from March 13, 2020 to August 31, 2020. OLS specification regresses outcome on treatment assignment. IV specification uses treatment assignment as an instrument for employers receiving the treatment. Heteroskedasticity-robust (Eicker–Huber–White) standard errors are used in all specifications. Significance indicators: $p \leq 0.05$ : *, $p \leq 0.01$ : **, and $p \leq .001$ : ***.

Treated jobs spent \$78 more than control jobs within 30 days of posting a job, and spent

$164 more within 90 days of posting. In Table 5 I observe the effects of treatment to the total wagebill using the same OLS and IV specification as in Section 5. In Columns (1) and (2) the outcome is dollars paid to the hire(s) in the first 30 days after posting the job. In Columns (3) and (4) the outcome is dollars paid to hire(s) in the first 90 days after posting.

In order to calculate the return on investment of this program, I use this unconditional impact to wagebill in first 90 days in order to include both the effects to number and size of contracts. The platform reports that this program costs them $3 per job post. I will use the OLS estimate in calculating ROI because I care about the total effect of the program, not just its effect on those that receive treatment. Therefore, the per job post ROI necessary to break even is $\frac{83.49 * x - 3}{3} = 0$ where x is the tax rate. Therefore, any tax authority implementing this program would get positive ROI within 90 days of job posting with a tax on wagebill of any larger than 1.2%.

From a social planners perspective, the ROI of this program would include the total wagebill. This would be $\frac{83.49 - 3}{3} = 2863\%$ with zero crowd out effects.

# 6 Why does the treatment increase hiring?

I now to turn to a harder empirical question, which is why the treatment "worked" to increase the fill rate.

I will try to separate the effects of shortlisting and recruiting. The fundamental problem is that when a Recruiter/Shortlister receives a (treated) job post, they will decide whether or not to do shortlisting, inviting, or neither based on their opinion of the job post, its applicant pool, and the employers behavior. Those in the treatment receiving some kind of assistance were highly negatively selected with respect to their baseline probability of filling a job. In Appendix C, I show that if the job post had sufficient applications and the Recruiter/Shortlister saw that the employer has been inviting and shortlisting workers on their own, they will not provide assistance. These proactive employers with many applicants are precisely those with a high baseline probability of hiring.

I show this negative selection in Table 6. In Columns (1) and (2) I regress the number of applications and whether or not the job filled respectively, on whether or not the job post received shortlisting help from the Recruiter/Shortlister. Treated job posts who receive shortlists are 13 percentage points less likely to hire than those in the control group and those in the treatment group which do not receive shortlists from the Recruiter/Shortlisters. They also receive over 12 more applications than the rest of the sample, due to the fact that the primary metric for receiving shortlisting help over recruiting help is having a a lot of applications at the time of the decision. For the same reason in Column (3) I find that treated workers who received recruiting help have 6 fewer applications. In Column (4) I show that treated jobs which receive recruiting help are 9 percentage points less likely to make a hire. In order to untangle the treatment effects of the help with the negative selection, I will attempt to compare like to like by predicting which jobs in the control group would have received recruiting or shortlisting had they been treated.

Table 6: Negative selection of job posts which receive shortlisting and recruiting

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Num Apps | Any Hire? | Num Apps | Any Hire? |
| | (1) | (2) | (3) | (4) |
| Received shortlists | 12.63*** | −0.13*** | | |
| | (0.003) | (0.003) | | |
| Received recruiting | | | −6.17*** | −0.09*** |
| | | | (0.004) | (0.004) |
| Constant | 22.49*** | 0.36*** | 26.59*** | 0.35*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Observations | 88,256 | 88,256 | 88,256 | 88,256 |
| Adjusted $R^2$ | 0.04 | 0.01 | 0.01 | 0.005 |

*Notes:* The table reports effects of shortlisting and inviting to number of applications and whether or not a job post makes a hire. It uses the entire experimental sample. Heteroskedasticity-robust (Eicker–Huber–White) standard errors are used in all specifications. Significance indicators: $p \leq 0.05$ : *, $p \leq 0.01$ : **, and $p \leq .001$ : ***.

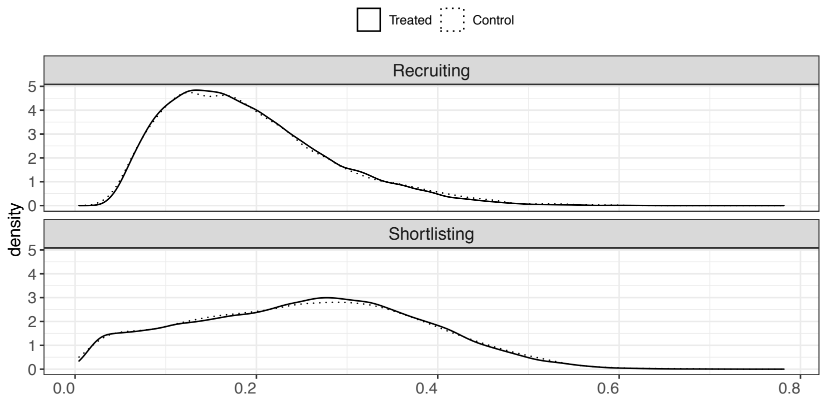
## 6.1 Modeling choice

I fit a logit model where the predictors are the number of applications the job received at the mean time a Recruiter/Shortlister would evaluate a job. One challenge is that while for treated jobs I know exactly when the Recruiter/Shortlister first sees the job post, for jobs in the control group there is no such time. Therefore I use the average length of time before a Recruiter/Shortlister begins working on a job post in the treated group to determine "number of applications before treatment" (and other time variant variables) for jobs in the control group. This number should not be affected by the treatment assignment. Indeed, Figure 5 shows the distribution of propensity scores for recruiting and shortlisting help. There is no evidence of a systematic difference between the treatment and control groups. Other predictors I include are number of interview requests the employer sends by this time, and time invariant characteristics like job category.

I partition the unit square recursively to keep equal counts of observations per cell. Figure 6 shows estimated by-cell treatment effects. For each facet, the x axis is the probability that a job post would have received shortlisting, based on the logit model described above. And the y axis is the probability that a job post would have received recruiting. In the first facet, the outcome estimated for each propensity score cell is the fraction of observations in the cell which receive recruiting. There are larger treatment effects in cells where the propensity score for recruiting is high, and smaller ones in cells where the propensity score for recruiting is low, hence the dark purple in the upper left hand quadrant. Similarly in the second facet, for which the outcome is the fraction of the cell that receives shortlisting, the larger treatment effects are all in the bottom right hand quadrant. And in the third quadrant, the largest effects to the fraction of cells which receive recruiting and shortlisting, the effects are where you would expect, the upper righthand quadrant.

Now that I have made a case that this strategy picks up true effects, I apply it to my outcome

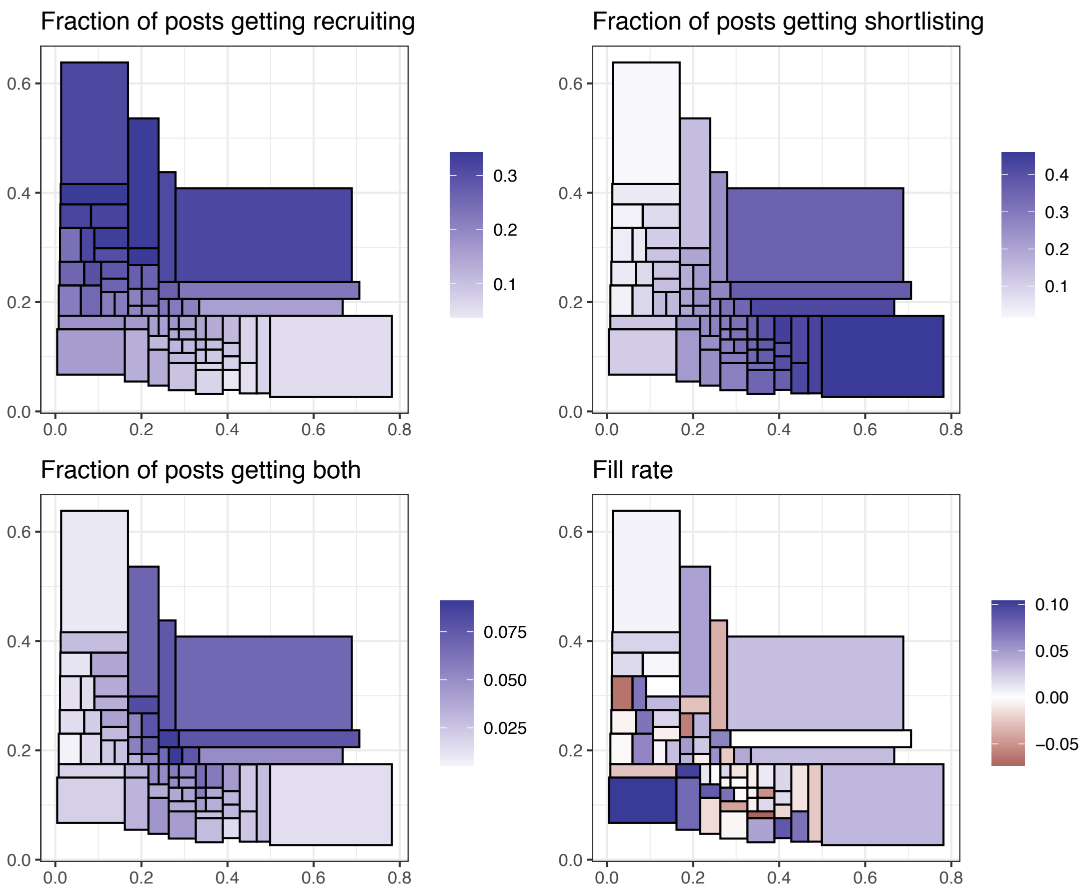
Figure 5: Distribution of propensity scores for receiving recruiting or shortlisting, by treatment assignment




*Notes:* This plot shows the distribution of estimated propensity scores for recruiting and shortlisting, for both the treatment and control group.

of interest, fill rate. The treatment effects to fill rate are scattered somewhat randomly, with no obvious interpretation. If there were larger treatment effects in the upper left hand side of the quardrant this would be evidence that cells more likely to receive recruiting have larger treatment effects. And if the same thing happened in the lower righthand quadrant, it would be evidence that cells more likely to receive shortlisting are the ones generating the largest treatment effects. However, I do not observe any particular pattern, leading us to believe treatment effects to exist and vary across the distributions of propensity scores.

Figure 6: Treatment effects by propensity score "cells"




*Notes:* This plot shows by-cell estimates of treatment effects on whether the job in the cell received recruiting or shortlisting help. For each facet the x axis is the propensity to receive shortlisting help, and the y axis is the propensity to receive recruiting help. Job posts bucketed into cells. The size of each cell is weighted by the number of job posts in each cell. Results summarized in Appendix Table E.2.

## 6.2 Catholic school IV

There is no perfect way to separate the effects of these two types of help, as they were endogenously determined and heavily negatively selected. I will try a variety of strategies that are imperfect for different reasons, and hope they point in the same direction.

First, for each type of help I will try an instrumental variables approach inspired by (Hoxby, 1994; Evans and Schwab, 1995; Neal, 1997), who use the fraction of catholics in a county as an instrument for whether or not a student went to catholic school. Instead, I will instrument whether or not an individual job post gets recruiting help with the fraction of jobs posted in each six hour block, other than that observation, which get recruiting help. There is a first stage because there are some Recruiter/Shortlisters and some periods of the day where Recruiter/Shortlisters are more likely to do shortlisting or recruiting. Therefore, I can use the fraction of jobs in each job post $i$'s category and time block (hereafter, category block) which get recruiting help as an instrument for whether or not job $i$ gets recruiting. It is useful to define the object *fraction of category block which gets recruiting*, which I will call $\Pi_{ri}$ and define for each job post as follows

$$\Pi_{ri} = \frac{R_{-i} - r_i}{n - 1}$$

where $R_{-i}$ is the number of job posts in the same time block and job category as $i$ that receive recruiting help *except for i*. $r_i$ is whether or not job post $i$ receives recruiting help, and $n-1$ is the number of job posts posted in the same time block and category (hereafter, category block) as $i$, except for $i$. I create a comparable object $\Pi_{si}$ for shortlisting.

Now, before using this object as an instrument, I must consider the exclusion restriction. In periods of time when Recruiter/Shortlisters are more likely to be recruiting, they might also be more likely to be shortlisting. This would affect hiring outcomes and invalidate the instrument. I handle this two ways. First, I control for the fraction of job posts in each category block that get worked at all. Then, for each $\Pi_{hi}$ I net out the effect of each treatment on each other. For recruiting, I subtract $\Pi_{ri}$ by the predicted effect of shortlisting on recruiting for that job post. I then use this netted out version I can call $\Pi_{ri*}$ as an instrument for whether or not an observation got recruiting help.

Table 7 reports the results of this exercise on the sample of all job posts which were assigned treatment. In all specifications I use $\Pi_{ri*}$ as an instrument for whether or not a job post gets recruiting, and control for the fraction of jobs worked at all in each category block. All specifications also include week fixed effects. In Column (1) I find a large positive effect of recruiting. However, in Column (2) I do not find a positive effect of shortlisting. In Columns (3) and (4) I control for the control group's fill rate for each category block. In Column (3) I am using the fraction getting recruiting in the job posts category block as an instrument for getting recruiting. Since it is possible for job posts to receive both types of treatment, in this specification I add a control for whether or not that job post received shortlisting. Similarly in Column (4) I am using the fraction getting shortlisting in the job posts category block as an instrument for getting shortlisting, and I control for whether or not the job post received recruiting help. I find that adding these controls slightly moderate the effect to the effect of recruiting, but the results are similar.

Table 7: Effects of shortlisting and recruiting on fill rate

| | Hiring Effect | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Shortlisting | | | $-0.14^{***}$ | |
| | | | (0.004) | |
| Recruiting | | | | $-0.09^{***}$ |
| | | | | (0.005) |
| Fraction worked | $-0.05^{***}$ | $-0.02^{*}$ | $-0.07^{***}$ | $-0.04^{**}$ |
| | (0.01) | (0.01) | (0.02) | (0.02) |
| Control group fill rate | | | $0.02^{***}$ | $0.02^{***}$ |
| | | | (0.01) | (0.01) |
| Recruiting [IV] | $0.21^{***}$ | | $0.15^{***}$ | |
| | (0.05) | | (0.05) | |
| Shortlisting [IV] | | $-0.23^{***}$ | | $-0.26^{***}$ |
| | | (0.04) | | (0.04) |
| Observations | 79,070 | 79,070 | 64,692 | 64,692 |

*Notes:* This analysis is run on all treated job posts in the experimental sample. In Columns (1) and (3), we use the fraction of job posts in each category block which receive recruiting, netting out the predicted effect of shortlisting on recruiting as an instrument for getting recruiting. In Columns (2) and (4) we use the fraction of job posts in each category block which receive shortlisting, netting out the predicted effect of recruiting on shortlisting as an instrument for getting shortlisting. In all specifications we control for fraction of job posts in each category block which were worked. In Column (3) we control for whether or not the job post received shortlisting, and in Column (4) we control for whether or not the job post received recruiting. In Columns (3) and (4) we also control for the fill rate in the control group for that job post's category block. Heteroskedasticity-robust (Eicker–Huber–White) standard errors are used in all specifications. Significance indicators: $p \leq 0.05$ : *, $p \leq 0.01$ : **, and $p \leq .001$ : ***.

## 6.3 Aggregated residualized leniency (judge IV) approach

In this section I will try another instrumental variables approach, inspired by what is often referred to as a judge leniency IV. Inspired by (Dahl, Kostøl, and Mogstad, 2014) and others, I define an aggregate measure of "leniency" to recruiting as the share of job posts assigned to each Process-helper. I define this measure as $\Gamma_j^r = \frac{R_j}{n_j}$, where $i$ is the job post in question and $j$ is the Process-helper assigned to job post $i$. $R_j$ is the number of job posts a particular Process-helper give recruiting help to, and $n_j$ is the number of jobs they work. I create the equivalent object $\Gamma_j^s$ for shortlisting.

Then, for each Process-helper $j$ I use a logit model to regress a binary indicator for whether or not a job post will be assigned $j$. I use this model to generate a prediction for the likelihood that job post $i$ will be assigned Process-helper$j$, which we shall call $\mu_{ij}$. Then I multiply $\mu_{ij}$ times that Process-helper's $\Gamma_{ij}^r$, and repeat this exercise for all Process-helpers and sum over each to get an aggregate measure of leniency or exposure to recruiting. If there are $M$ Process-helpers, then the aggregated residualized leniency instrument is defined as:

$$\Omega^r = \sum_{m=1}^{M} \mu_{ij} * \Gamma_j^r$$

and multiply this object by a binary indicator for treatment as an instrument for getting recruiting help. I regress whether or not a job fills on the probability a job receives recruiting, using $\Omega^r$ multiplied by an indicator for whether the job is in the treatment group, as an instrument for whether or not the job post gets recruiting help.

Table 8 reports the effects of shortlisting and recruiting to whether a job post fills using this specification. The entire experimental sample is included in this analysis. Column (1) uses the proclivity of each helper to recruit as an instrument for whether or not a job received recruiting help. It controls for the probability that it received recruiting help. Probabilities of receiving each type of help are generated by a logit using attributes of jobs in the treatment group to predict the likelihood that a given job in the control group receives each type of treatment. Column (2) reports the same specification for shortlisting.

This analysis relies on the assumption that the only effect a Process-helper has on hiring outcomes is whether or not they recruit (or shortlist.) The problem with this is that Process-helpers proclivity to recruit is correlated with their proclivity to shortlist (and vice versa) because there are some Process-helpers who just work a larger fraction of their jobs than others. Remember that demand for Process-helpers outstrips supply, and that they only get to on average half of the job posts they are assigned. To deal with this, I control for the fraction of jobs which each Process-helper works. I report the results in Table 8 and find that under these assumptions recruiting has almost three times the effect that shortlisting does.

Table 8: Effects of shortlisting and recruiting on fill rate using residualized leniency approach

| | Hired | |
|---|---|---|
| | (1) | (2) |
| Pr(Recruiting) | $-0.66^{***}$ | |
| | (0.03) | |
| Recruiting [IV] | $0.11^{***}$ | |
| | (0.03) | |
| Pr(Shortlisting) | | $0.11^{***}$ |
| | | (0.02) |
| Shortlisting [IV] | | $0.06^{***}$ |
| | | (0.02) |
| Constant | $0.44^{***}$ | $0.29^{***}$ |
| | (0.004) | (0.004) |
| Observations | 88,057 | 88,057 |

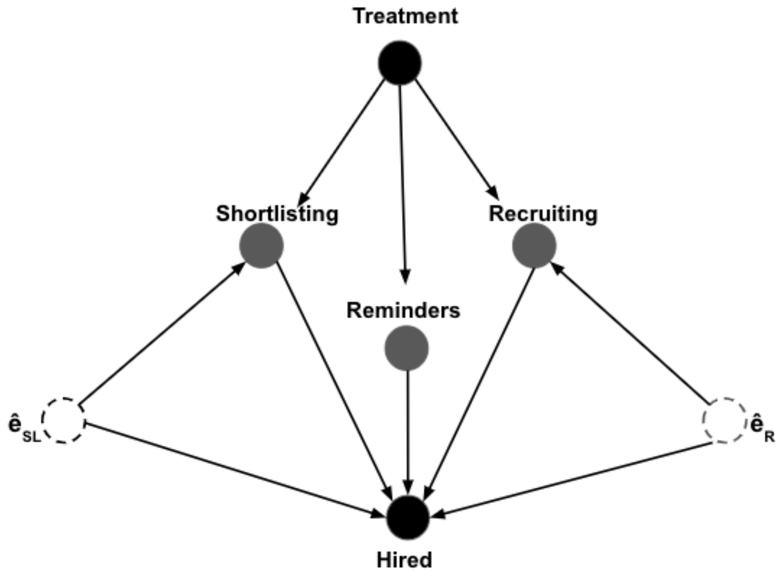
*Notes:* This table reports the effects of shortlisting and recruiting to whether a job post fills. The entire experimental sample is included in this analysis. Column (1) uses the proclivity of each helper to recruit as an instrument for whether or not a job received recruiting help. The instrument is calculated as a treatment indicator times the probability a given job post would have been helped by a particular helper times that helpers proclivity to recruit, summed over all helpers. It controls for the probability that it received recruiting help. Probabilities of receiving each type of help are generated by a logit using attributes of jobs in the treatment group to predict the likelihood that a given job in the control group receives each type of treatment. Column (2) reports the same specification for shortlisting. Heteroskedasticity-robust (Eicker–Huber–White) standard errors are used in all specifications. Significance indicators: $p \leq 0.05$ : *, $p \leq 0.01$ : **, and $p \leq .001$ : ***.

### 6.4 Decomposing the reasons

Figure 7 shows a DAG illustrating a proposed causal mechanism for the effects of the different types of help on fill rate. I will go through each proposed causal link one by one and describe my evidence for the existence and direction of each link. First, all treated job posts receive an email and message offering help and reminding the employer to start inviting and interviewing workers. The results to employers sending more invitations and interviewing more workers suggests that this causal link exists even if it is small (Table 3.) The main experimental estimates show that there is a causal link between treatment and shortlisting and recruiting (Table 2.) Treated job posts get sent more invitations and more applications shortlisted by the Recruiter/Shortlisters. For both shortlisting and recruiting, the there is some unobserved characteristic $\hat{e_{SL}}$ and $\hat{e_R}$ that affects both the probability of that action being taken, as well as whether the job opening leads to a hire (Appendix Table C.1.) This unobserved characteristic is what causes the negative selection into getting either type of help , and is why not to believe results simply regressing hiring probability on whether or not a Recruiter/Shortlister took the recruiting or shortlisting action on behalf of the job post(Table 6.)

I show a causal link between shortlisting and likelihood of making a hire. Results from the aggregated residualized leniency (judge IV) approach suggest that shortlisting has a small positive effect on hiring (Table 8.) These results suggest that between one fifth and one third of the total effects of the Recruiter/Shortlisters come from shortlisting. And lastly, I show a consistently positive causal link between recruiting and likelihood of making a hire. In both the

Figure 7: DAG




*Notes:* This DAG shows the causal links between the types of treatment and whether or not a job makes a hire.

judge IV (Table 8) and the catholic school IV (Table 7) the effect of recruiting is at least double the effect of shortlisting.

# 7 Conclusion

I analyze the results of an experiment which randomly assigned hiring assistance to employers posting jobs on a large online labor market. Hiring assistance took multiple forms including reminder messages from the helpers, inviting workers to the job, and shortlisting the top applicants. Whether inviting or shortlisting took place was endogenously determined by the helper. I find that hiring assistance significantly effects the likelihood that a firm makes a hire. I find that not only do the helpers invite workers on the employers behalf, but that treated employers invite more workers on their own. Treated employers also conduct more interviews, and even conduct more interviews of applicants that were not invited by the helpers. These results suggest that the reminder messages from the helpers lowered the cost of inviting and interviewing workers by reminding employers to act.

Lastly, I tease apart the mechanism underlying the value of help to hiring. For a given hiring budget, firms must allocate their resources between expanding their applicant pool (recruiting) and selecting among their current applicants (shortlisting and interviewing.) The results suggest that both are beneficial, but that increasing the supply of applicants has a higher payoff when the firm has a satisfactory screening technology.

These results have implications to both employers and to platforms. First to employers, I find that hiring assistance significantly effects the likelihood that a firm makes a hire. Second, I find that the most effective type of help to employers involves increasing their applicant pool with positively selected applicants. And to platforms, I find that helping employers with the hiring process is a cost effective way to increase the number of contracts on the platform as well as contract size.

After the conclusion of this experiment, the platform implemented the hiring assistance program as described in the experimental design. Starting in April 2022 they removed the Process-helpers role, leaving the hiring assistance to only take the form of recruiting and shortlisting, with no human for the employer to interact with. This will provide us with a natural experiment to understand how much of the treatment effect was caused by the human-in-the-loop responsibilities of the Process-helper. I will report these results in future drafts.

## References


M. A. Barach and J. J. Horton. How do employers use compensation history? evidence from a field experiment. *Journal of Labor Economics*, 39(1):193–218, 2021.

M. Belot, P. Kircher, and P. Muller. Providing Advice to Jobseekers at Low Cost: An Experimental Study on Online Advice. *The Review of Economic Studies*, 86(4):1411–1447, 10 2018.

N. Bloom, B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts. Does management matter? evidence from india. *The Quarterly Journal of Economics*, 128(1):1–51, 2013.

D. Card, J. Kluve, and A. Weber. Active labour market policy evaluations: A meta-analysis. *The economic journal*, 120(548):F452–F477, 2010.

D. Card, J. Kluve, and A. Weber. What works? a meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, 16(3):894–931, 2018.

B. Crépon, E. Duflo, M. Gurgand, R. Rathelot, and P. Zamora. Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment *. *The Quarterly Journal of Economics*, 128(2):531–580, 04 2013. ISSN 0033-5533. doi: 10.1093/qje/qjt001. URL https://doi.org/10.1093/qje/qjt001.

G. B. Dahl, A. R. Kostøl, and M. Mogstad. Family welfare cultures. *The Quarterly Journal of Economics*, 129(4):1711–1752, 2014.

W. N. Evans and R. M. Schwab. Finishing high school and starting college: Do catholic schools make a difference? *The Quarterly Journal of Economics*, 110(4):941–974, 1995.

A. Filippas, J. J. Horton, and J. Golden. Reputation inflation. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 483–484, 2018.

J. J. Horton. Online labor markets. *Internet and network economics*, pages 515–522, 2010.

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

C. M. Hoxby. Do private schools provide competition for public schools?, 1994.

D. Neal. The effects of catholic secondary schooling on educational achievement. *Journal of Labor Economics*, 15(1, Part 1):98–123, 1997.

A. Pallais. Inefficient hiring in entry-level labor markets. *The American Economic Review*, 104 (11), March 2013. URL http://www.nber.org/papers/w18917.

M. Pellizzari. Employers' Search and the Efficiency of Matching. *Institute for the Study of Labor Discussion Paper No. 1862*, (1862), 11 2005. URL http://dx.doi.org/10.2139/ssrn.866907.

A. E. Roth. Marketplaces, markets, and market design. *American Economic Review*, 108(7): 1609–58, 2018.

C. T. Stanton and C. Thomas. Landing the first job: The value of intermediaries in online hiring. *The Review of Economic Studies*, 83(2):810–854, 2015.

# Appendix

## A Message

Hi {employer first name}, I'm your [Process-helper]. Chat with me about any {platform} questions or issues — I'm here to help! I'm online {Mon—Fri 8:00 AM —4:00 PM Pacific Time}. If I'm offline when you contact me, I'll reply as soon as I'm back. Freelancers will begin submitting proposals for you {job opening title} job. In the meanwhile, you can invite specific freelancers to submit proposals {at this link} and I'll start connecting you with top talent for this job. Do you have any questions as you get started?

## B Internal Validity

To assess the effectiveness of randomization, in Table B.1 I report the mean values and t-tests for various pre-randomization attributes of the job post, and obtain excellent balance on these covariates.

Table B.1: Balance Table

| Variable | Control | Treatment | Difference | p-value |
|---|---|---|---|---|
| First Job Post | 0.215 | 0.213 | -0.002 | 0.670 |
| Software | 0.544 | 0.546 | 0.002 | 0.692 |
| Administrative | 0.099 | 0.102 | 0.003 | 0.423 |
| Low Skill | 0.081 | 0.083 | 0.003 | 0.411 |
| High Skill | 0.511 | 0.511 | 0.0001 | 0.993 |
| Large Company | 0.008 | 0.006 | -0.002 | 0.128 |
| Num Skills Required | 8.210 | 8.421 | 0.211 | 0.020 |

*Notes:* This table reports group summary statistics for several job post attributes in the experiment, as well as a t-test comparing those means.

One concern with using job post as the unit of analysis is that employers who are unable to fill a job might post the job opening on other online labor market sites or in the conventional market. (Roth, 2018) Since the treatment appears to make employers more likely to find a satisfactory worker on the platform, then treated workers may be less likely to post openings or hire in a different marketplace. Employers in the control group might be more likely to hire outside of the platform, and since I do not observe those hires, the estimates will be biased upward. However, survey evidence from the platform suggests that online and offline hiring are only very weak substitutes and that multi-homing of job openings is rare. Online employers reported that they generally decide between hiring a worker online, doing the work themselves, and not having the work done at all. Only 15% of employers surveyed said reported if the platform were not available, they would have made a hire locally. And furthermore, 83% of the employers reported that their most recent job posting only on this platform and not on other online labor marketplaces.

# C How do jobs which receive recruiting or shortlisting help compare to those that don't?

Recruiter/Shortlisters endogenously determine whether or not to provide shortlisting or recruiting help to the job posts that they work. I want to consider how this subset of job posts which go on to receive the help differ from the subset which do not. I use data on the observable characteristics of the job posts to estimate the determinants of receiving help. I restrict the sample to only include the 48,556 job posts which the Recruiter/Shortlister marks as "worked," and then estimate the following logistic model:

$$\text{helped}_i = \alpha + \sum_j \beta_j X_{ij} + \epsilon_i$$

where $\text{helped}_i$ equals 1 if the Recruiter/Shortlister ever shortlists or invites workers, and zero otherwise. I display the results in in Table C.1, with the vector of X's in the first Column and each independent variable's coefficient and standard error in the second Column. Number of Applications is the number of applications the job post receives prior to the Recruiter/Shortlister getting to the job post and Number of Recommended Applications isolates the total number of applications to only those that the platforms' internal algorithm marks as recommended. Projected Value is the size of the job in terms of expected number of hours and wages. And Employer Shortlists and Employer Invites are the number of shortlists and invites that the employer has done prior to the Recruiter/Shortlister working the job, respectively. I also include other job level characteristics. I find that jobs which the Recruiter/Shortlisters chose to work have fewer applications and shortlists at the time they start working on it. They are also more likely to be in the field of graphic design, and to require more skills.

# D How do recruited applicants compare to organic applicants?

Given the positive effect that additional applications have on the likelihood of a job filling, it is conceivable that the fill rate effect is driven by the effect to additional applications randomly drawn from the distribution. Alternately, it is possible that the new recruits are positively selected. This is perhaps even likely given the intentions of the recruiting help was to invite "good fitting applicants" to the job post. I can do back of the envelope math to test this. Recall that the effect of treatment on fill rate is approximately 10%. Given the size of the increase in invited applicants, is this increase in the fill rate about what we would expect? 37% of invited workers end up applying to the job they are invited to, and therefore one additional invite should generate 0.37 new applications to a job. In Table 2 Column (2), see that the effect of the treatment is 5 new invitations. Multiplying this by 0.37, gives 1.85, which is the number of new applications expected given 5 new invites. In Table 2 Column (4), the treatment generates 2.2 new applications, very close to the 1.85 expected. This means that the new recruits are no more likely to apply than any other worker invited by the employer.

Table C.1: Which worked jobs get help from Recruiter/Shortlisters?

| | *Dependent variable:* |
|---|---|
| | Job post gets at least one invite or shortlist |
| Number of Applications | $-0.003^{***}$ (0.001) |
| Number of Recommended Applications | $-0.291^{***}$ (0.005) |
| Projected Value | $-0.068^{***}$ (0.002) |
| Administrative Job | 0.054 (0.040) |
| Software Job | $-0.021$ (0.026) |
| Design Job | $0.304^{***}$ (0.061) |
| Employer Shortlists | $-0.002$ (0.001) |
| Employer Invites | $0.00003^{***}$ (0.00000) |
| Employer First Job Post | 0.043 (0.026) |
| Num Skills Req | $0.003^{**}$ (0.001) |
| Low Skill Job | $-0.018$ (0.044) |
| High Skill Job | 0.005 (0.022) |
| Constant | $1.363^{***}$ (0.033) |
| Observations | 48,556 |

*Notes:* This table compares job posts which get worked by Recruiter/Shortlisters but do not receive recruiting or shortlisting help, with those that are worked and do receive at least one type of help. The sample is all treated jobs which are worked by a Recruiter/Shortlister. Standard errors are in parentheses. Significance indicators: $p \leq 0.05$ : *, $p \leq 0.01$ : **, and $p \leq .001$ : ***.

Next, I want to understand the relationship between number of applications and likelihood of a hire. When I regress whether or not a job filled on the number of applications the job received, there is an effect of size 0.0013 percentage points on a base of 0.30, a 0.43% effect. If this effect is causal, one randomly drawn new application should correspond to a fill rate by 0.43%. Table 2 Column (2) shows that treated jobs get 2.2 more applications, therefore I would expect a treatment effect of 0.95% if inviting workers is the only thing causing the fill rate effect. Since this is only a tenth of the treatment effect I find from this experiment, it implies that the marginal invitation is not for a worker pulled randomly from the distribution, but instead is for ones more likely to be hired.

Therefore, in Table D.1 I compare Recruiter/Shortlister invited applications to all applications to jobs in the experimental sample. Column (1) looks at the wage bid on each application to a treated job, with a sample of over 2 million applications. It shows that invited workers bid almost a dollar higher than workers who are not invited. This could be due to the fact that these workers charge higher rates generally. It could also be that workers interpreted being invited to the job as a positive signal that they're likely to get the job, and raised their bids. However, I also look at the profiles of the workers that applied to see if their "profile hourly wage" is higher.

The remaining columns look at the sample of 820,268 unique workers who applied to any treated job over the experimental period. Column (2) shows that the invited workers had hourly rates posted on their profile of almost $3 more than ones that do not get invited by Recruiter/Shortlisters. Column (3) shows that the Recruiter/Shortlister invited workers had

Table D.1: Comparison of workers recruited by Recruiter/Shortlisters to other applications of treated job posts

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Wage bid | Profile rate | Total earned | Hours worked | Previous invites |
| | (1) | (2) | (3) | (4) | (5) |
| Recruited worker | $0.87^{***}$ | $2.77^{***}$ | $28,237.48^{***}$ | $1,520.18^{***}$ | $285.76^{***}$ |
| | (0.27) | (0.66) | (1,167.42) | (54.83) | (7.86) |
| Constant | $33.10^{***}$ | $35.58^{***}$ | $16,132.44^{***}$ | $551.54^{***}$ | $73.18^{***}$ |
| | (0.05) | (0.05) | (97.08) | (2.12) | (0.27) |
| Observations | 2,039,682 | 820,268 | 820,268 | 820,268 | 820,268 |
| Adjusted $R^2$ | 0.0000 | 0.0000 | 0.001 | 0.003 | 0.01 |

*Notes:* The first column reports the wage bids for recruited applications, versus all other received applications. It uses all applications to jobs in the treated group of the experimental sample. Columns (2) - (5) reports the work history on the platform of recruited workers as compared to workers who applied without being recruited. Heteroskedasticity-robust (Eicker–Huber–White) standard errors are used in all specifications. Significance indicators: $p \leq 0.05$ : *, $p \leq 0.01$ : **, and $p \leq .001$ : ***.

much higher earnings on the platform, and Column (4) shows that they have worked significantly more hours on the platform. Finally, Column (5) shows that they have received significantly more invites from employers in the past. Along all dimensions I see that the Recruiter/Shortlisters are inviting workers that have much higher hiring success on the platform than a random draw from the distribution.

# E Appendix Tables and Figures

Table E.1: Effects of treatment on whether or not a job filled for new employers

| | *Dependent variable:* | |
|---|---|---|
| | Hired | |
| | (1) | (2) |
| Treatment Assigned | $0.02^{***}$ | 0.01 |
| | (0.01) | (0.01) |
| First job post | $-0.14^{***}$ | |
| | (0.01) | |
| Treatment Assigned * First job post | $-0.01$ | |
| | (0.01) | |
| Age in weeks | | 0.0001 |
| | | (0.01) |
| Treatment Assigned * Age | | 0.0000 |
| | | (0.01) |
| Constant | $0.35^{***}$ | $0.38^{***}$ |
| | (0.01) | (0.01) |
| Observations | 88,224 | 44,735 |
| Adjusted $R^2$ | 0.02 | 0.001 |

*Notes:* This table reports effects of treatment whether or not a job filled, by age of employer on the platform. First job post is a binary variable for whether or not it is the employers's first job post on the platform. Age is the number of weeks since the time the employer registered for the platform. Sample is experimental sample of high value posts from March 13, 2020 to August 31, 2020. OLS specification regresses outcome on treatment assignment. IV specification uses treatment assignment as an instrument for employers receiving the treatment. Heteroskedasticity-robust (Eicker–Huber–White) standard errors are used in all specifications. Significance indicators: $p \leq 0.05$ : *, $p \leq 0.01$ : **, and $p \leq .001$ : ***.

Table E.2: Effects of shortlisting and recruiting on fill rate treatment effect

| | Hiring effect | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Fraction of jobs which get only recruiting | 0.01 | 0.004 | −0.01 |
| | (0.12) | (0.12) | (0.12) |
| Fraction of jobs which get only shortlisting | −0.18 | −0.01 | −0.03 |
| | (0.16) | (0.09) | (0.09) |
| Fraction of jobs which get both | −0.32 | −0.15 | |
| | (0.27) | (0.24) | |
| Fraction of jobs worked | 0.18 | | |
| | (0.15) | | |
| Constant | −0.05 | 0.02 | 0.02 |
| | (0.07) | (0.04) | (0.04) |
| Observations | 64 | 64 | 64 |
| Adjusted $R^2$ | -0.03 | -0.04 | -0.03 |

*Notes:* This table is a summary of the by-cell treatment effects detailed in Figure 6. Job posts are bucketed into cells based on their propensity to get shortlisting and recruiting help. The size of each cell is weighted by the number of job posts in each cell. Significance indicators: $p \leq 0.05$ : *, $p \leq 0.01$ : **, and $p \leq .001$ : ***.