The Effects of Subsidizing Employer Search*

John J. Horton
Leonard N. Stern School of Business, New York University

October 28, 2013

Abstract

Providing employers with recommendations about whom to recruit can substantially lower the number of vacancies that go unfilled. In an experiment conducted in an online labor market, employers with technical job vacancies that received recommendations had a 19.5% higher vacancy fill rate compared to the control. There is no evidence that additional matches induced by the treatment were better or worse than non-induced matches. Surprisingly, the treatment did not crowd-out non-recommended candidates. This lack of crowd-out was likely caused by a spill-over: recommendations improved the applicant pool, which increased employer screening on the extensive margin. In aggregate, increased screening outweighed the mechanical crowd-out caused by a larger applicant pool.

JEL Codes: C93, J41, D47, D21
Keywords: employer search, labor market intermediation, field experiments

1 Introduction

Both buyers and sellers in labor markets bear search costs to find each other and learn about each other's attributes. This paper reports the results of an experimental intervention designed to reduce search costs within a labor market. In the experiment, treated employers received

*All code and data is or will be available at http://john-joseph-horton.com. I am grateful to the oDesk Research and oDesk Product team members for encouragement, ideas and assistance in this project. Larry Katz, Richard Zeckhauser, Luke Stein, Jacob Leshno, Mandy Pallais, Tal Gross, Peter Coles, Ramesh Johari and especially Dana Chandler offered numerous helpful comments and suggestions. Thanks to seminar participants at HBS, Stanford MS&E, Wharton, LBS and NYU Stern. Thanks to Robin Yerkes Horton, Carolyn Yerkes, Ada Yerkes Horton, Heidi Yerkes and David Yerkes with help in preparing the manuscript. This paper previously circulated under the title "Computer-Mediated Matchmaking: Facilitating Employer Search and Screening."

recommendations about which workers to recruit for their recently-posted vacancies. Recommendations were based on statistical models of employer preferences, fit using historical data from the marketplace in which the experiment was conducted. This algorithmic approach to lowering search costs has become ubiquitous in computer-mediated product markets (Varian, 2010; Resnick and Varian, 1997; Adomavicius and Tuzhilin, 2005), but it is still rare in labor markets. However, this characterization is likely to change as computer-mediation or "wiring" of all markets continues (Autor, 2001).

Algorithmic labor market recommendations may become feasible, but it is not clear that they would improve upon what buyers can do for themselves; perhaps buyers can readily assemble a pool of candidates for their vacancies, but generating a match requires the evaluation of idiosyncratic or ineffable qualities difficult to capture in a statistical model. Further, while more information generally improves markets, recommendations do not provide more information generically—by design, they encourage a buyer to consider some workers but not others. If displacement or "crowd out" effects—which have proven important in search assistance interventions in conventional labor markets—are strong, it makes recommendation interventions less attractive from a social welfare perspective.

Motivated by these questions of effectiveness and crowd out, this analysis focuses on four questions: (1) Do firms act upon recommendations and if so, why? (2) How do recommendations affect the firm's probability of filling the vacancy? (3) How do recommendations affect the size, wages and self-evaluation of resultant matches? (4) How do the recommendations affect the outcomes of non-recommended applicants?

To help guide this analysis, I develop a simple model of employer search and screening adapted to the empirical context of this experiment. In the model, firms receive a pool of applicants whose productivity they can partially observe. Based on the "draw" of applicants, the firm can either call off the search or begin screening applicants sequentially with recall, at some cost. By paying a per-applicant evaluation cost, the firm can learn the applicant's true productivity.

When the search is complete, either the vacancy goes unfilled (if none of screened applicants exceeds the firm's reservation productivity) or the firm and the hired applicant Nash bargain over the surplus to set the wage. Knowing how this process will evolve, the firm can choose, ex ante, to recruit applicants, thereby increasing the size of their applicant pool, albeit at some recruiting cost. The introduction of algorithmic recommendations can usefully be thought of as subsidizing the cost of this ex ante recruiting.

If the experiment reduced recruiting costs, the model predicts that more firms should recruit when in the treatment. The data strongly bears out this prediction: the treatment increased the fraction of firms recruiting by 40.6%. Consistent with the model, there is a strong negative correlation between the number of organic applicants received and the firm's propensity to recruit. Other indicators that the firm has a high reservation productivity—such as requiring prior on-platform experience—are also associated with a greater propensity to recruit.

In the model, recommendations increase the probability that the firm fills the vacancy. This increase occurs via two channels: recommendations increase the chance that the firm makes a hire, conditional upon screening, and they increase the probability that the firm screens anyone at all. The treatment increased the overall fill rate in technical vacancies by 19.5%, with no detectable effect on non-technical vacancies. I provide evidence that the greater effectiveness of the treatment in technical categories is likely due to a higher marginal return to additional applicants in technical categories, which is the result of a smaller baseline number of applicants in those technical categories.

Although the model predicts that recommendations raise the value of a pool of applicants to the firm (expected surplus less evaluation costs), it does not predict that recommendations necessarily improve match quality. The reason is that recommendations can cause the firm to evaluate fewer applicants in total, which lowers expected match quality. I find no evidence that the matches induced by the treatment were any better or worse than the matches in the control group: treatment and control matches are not statistically different in terms of hours worked, total wage bill, bilateral feedback and so on. However, detecting match quality effects is difficult given that experimentally induced matches are pooled with non-induced matches in the treatment group.

In the modeling framework, there are two possible reasons that a non-recommended applicant is not hired: (a) the firm decides to hire someone else or (b) the firm hires no one at all, perhaps because the search is called off. Recommendations mechanically make the first "crowd out" explanation more likely, but if they increase the quality of the overall pool, they make the second explanation less likely. I find that the treatment did not crowd out non-recommended applicants. If anything, the recommendations treatment *increased* the hiring of non-recruited, organic applicants. This result is important, as it addresses a major concern of any search-focused intervention that helps only some participants. I present evidence that the treatment increasing employer screening on the extensive margin, which caused the apparent complementarity between organic and recruited hiring. If this finding generalizes to other markets, the existence of a meaningful extensive margin in employer screening provides an opportunity for targeted, relatively crowd-out free labor market interventions.

This experiment was conducted in a marketplace that is particularly information-rich, in that both sides have access to the universe of job-seekers and vacancies via a search interface. Comprehensive, public data—past vacancies posted/applied to, wages earned/paid, feedback earned/given, etc.—describe both workers and firms. In this environment, we might suppose that the marginal benefit of recommendations would be very low, and yet this is strongly not the case. And although conventional labor markets are relatively higher-staked, the treatment was only effective in the higher-staked, technical portion of the oDesk market.

This paper provides a rich description of the employer search and screening process: it shows that after firms post vacancies, they endogenously decide whether to recruit, quickly receive a more or less complete pool of applicants and then evaluate them serially, but ordered not by arrival time but by "on paper" quality. This description is at odds with models of employer search screening that mirror the canonical worker search models—i.e., each firm has a "reservation productivity" analogous to a reservation wage and firms screen serially without recall. There is remarkably little work in economics on how employers actually recruit and evaluate applicants. Oyer and Schaefer (2010) characterize the literature as having been "less successful at explaining how firms can find the right employees in the first place. Economists understand the broad economic forces—matching with costly search and bilateral asymmetric information... but the main models in this area treat firms as simple black-box production functions." To the extent the characterization of firms from the setting of this paper generalizes to conventional labor markets, the results suggest the need for progress in models of employer search and screening.

2 Empirical Context

During the last ten years, a number of online labor markets have emerged. In these markets, firms hire workers to perform tasks that can be done remotely, such as computer programming, graphic design, data entry, research and writing. Markets differ in their scope and focus, but common services provided by the platforms include maintaining job listings, hosting user profile pages, arbitrating disputes, certifying worker skills and maintaining feedback systems.

The experiment in this paper was conducted on oDesk, the largest of these online labor markets. On oDesk, would-be employers write job descriptions, self-categorize the nature of the work and required skills and then post the vacancies to the oDesk website. Workers learn about vacancies via electronic searches or email notifications and can submit applications. In addition to worker-initiated applications, employers can also search worker profiles and invite workers to apply. These invitations—which I use as a measure of recruiting—simply alert workers that an employer is interested in them. Workers can either ignore recruiting invitations or apply for the vacancies. After a worker submits an application, the employer can interview and hire the applicant on the terms proposed by the worker or make a counteroffer, which the worker can counter, and so on.

To work on hourly oDesk contracts, contractors must install custom tracking software on their computers. The tracking software, or "Work Diary," essentially serves as a digital punch clock that allows for remote monitoring of employees. This monitoring makes hourly contracts and hence employment relationships possible, which in turn makes the oDesk marketplace more like a traditional labor market than project-based online marketplaces where contracts are usually arm's-length and fixed price.

In the first quarter of 2012, $78 million was spent on oDesk. The 2011 wage bill was $225 million, representing 90% year-on-year growth from 2010. As of October 2012, more than 495,000 employers and 2.5 million contractors have created profiles (though a considerably smaller fraction are active on the site). Approximately 790,000 vacancies were posted in the first half of 2012. See Agrawal et al. (2013a) for additional descriptive statistics on oDesk.

Each vacancy is labeled by the posting employer with some number of required skills. Based on dollars spent, the top required skills in the marketplace are web programming, mobile applications development (e.g., iPhone and Android) and web design. Based on hours worked, the top required skills are web programming, data entry, search engine optimization and web research. The difference in the top skills based on dollars versus hours reflects a fundamental split in the marketplace between technical and non-technical work. There are highly-skilled, highly-paid contractors working in non-technical jobs, yet a stylized fact of the marketplace is that technical work tends to pay better, generate longer-lasting relationships and require greater skill.

The oDesk marketplace is not the only marketplace for online work (or IT work more generally). As such, one might worry that every vacancy we see on oDesk is simultaneously posted on several other online labor market sites *and* in the traditional market. However, survey evidence suggests that online and offline hiring are only very weak substitutes and that multi-homing of vacancies is relatively rare. When asked what they would have done with their most recent project if oDesk were not available, only 15% of employers responded that they would have made a local hire. Online employers report that they are generally deciding among (a) getting the work done online, (b) doing the work themselves and (c) not having the work done at all. The survey also found that 83% of employers said that they listed their last job vacancy on oDesk alone. This limited degree of multi-homing makes it more reasonable to discuss crowd-out effects.

There has been some research which focuses on the oDesk marketplace. Pallais (2013) shows via a field experiment that past worker experience on oDesk is an excellent predictor of being hired for subsequent work on the platform. Stanton and Thomas (2012) use oDesk data to show that agencies (which act as quasi-firms) help workers find jobs and break into the marketplace. Agrawal et al. (2013b) investigate what factors matter to employers in making selections from an applicant pool and present some evidence of statistical discrimination; the paper also supports the view of employers selecting from a more-or-less complete pool of applicants rather than serially screening.

2.1 Why do some vacancies go unfilled?

As in the conventional labor market, many vacancies on oDesk go unfilled. In an online survey of employers that had recently posted but not filled their vacancies, would-be employers were asked to state the primary reason why they did not hire anyone. In Table 1 we report the counts and percentages of employers selecting the various reasons. The survey was sent to 2,566 employers and a total of 104 relevant responses were received. Employers stating they were still in the process of evaluating candidates or had hired someone were excluded.

The usual caveat about self-reports aside, a plurality of employers claim that a lack of needed skills was the cause for not filling their vacancy (33%). Two possibilities exist: workers simply do not exist at the price employers were willing to pay or such workers do exist but failed to find

Table 1: Employers' stated reasons for not filling a recently posted vacancy on oDesk

Reason given:	Number of employers	Percent of employers
Couldn't find a worker with the right skills	34	32.69
Project cancelled/postponed	23	22.12
Workers I was interested in didn't respond	11	10.58
Couldn't find a worker at an affordable price	11	10.58
Another reason	8	7.69
Haven't spent enough time looking for a worker	8	7.69
Too much effort required to find the right worker	8	7.69
Not enough applicants	1	0.96

Notes: This table reports the survey responses of a sample of oDesk employers (*not* employers who participated in the experiment described in this paper) who did not fill their vacancies. The purpose of the survey was to learn more about the reasons some vacancies go unfilled. The survey was conducted online during the month of February 2013. An invitation to participate in the survey was sent to 2,566, employers from which 104 relevant responses were received. Respondents were asked to indicate the primary reason they did not fill their vacancy. The counts of employers and the fractions of employers are reported for each primary reason.

the employer and apply. The second explanation is consistent with a "search friction" conception of a labor market and implies that recommendations that remove frictions could raise fill rates. However, if the first explanation holds, recommendations cannot help—the vacancy will necessarily go unfilled.

3 Description of the experiment

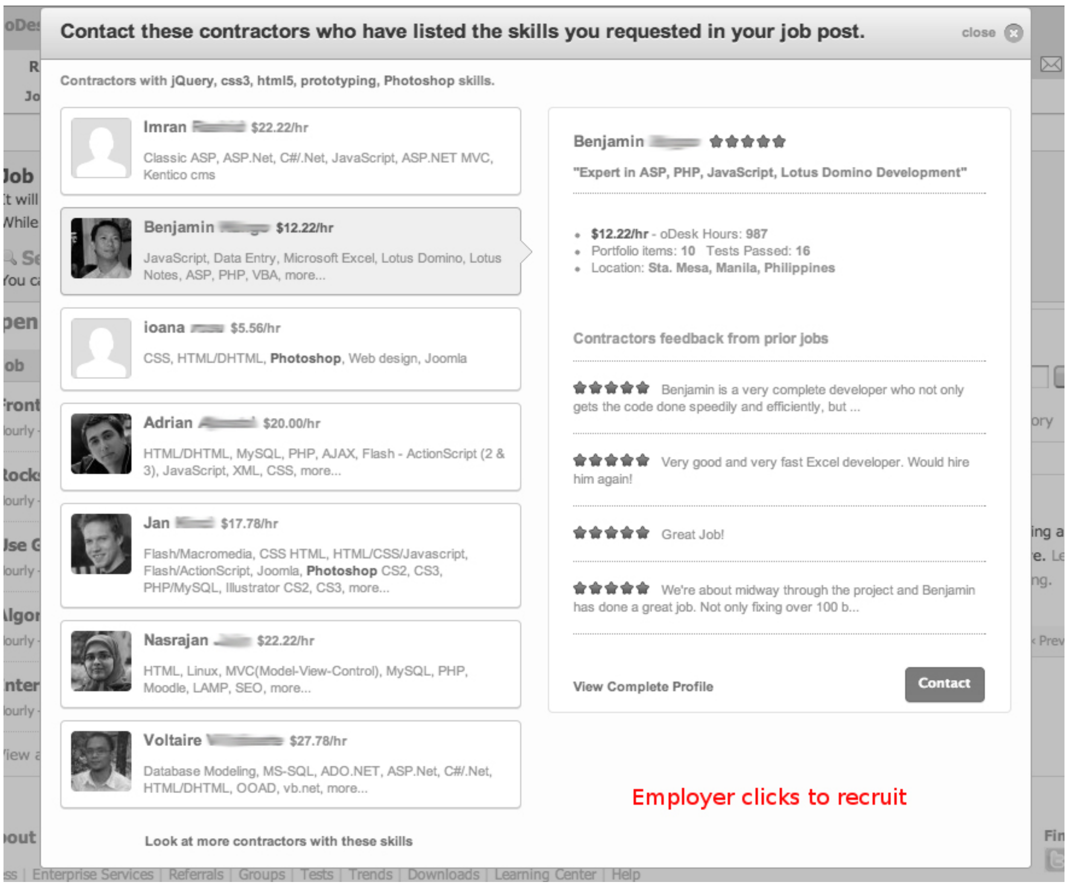
In June 2011, oDesk launched recommendations as an experimental feature. Only new employers for whom recommendations could be made were eligible for the experiment. Immediately after posting a vacancy, a treated employer was shown six recommended workers. Recommended workers were selected based upon the inferred relevance of their skills, ability and availability. Relevance was measured by the degree of overlap in the skills required for the vacancy and the skills listed by the worker in their profile. Ability was defined as a weighted sum of skill test scores, feedback ratings and past earnings. Availability was inferred from signals such as a worker recently ending a project or applying to other vacancies. Employers in the control group received the status quo experience of no recommendations. The resulting data from this experiment is based on 5,953 vacancies. Appendix A shows that randomization was effective and the experimental groups were well balanced.

Figure 1 shows the recommendation interface, which was shown in a pop-up window. After a treated employer was presented a pop-up window, they could view each recommended worker's photograph, listed skills, average feedback score and stated hourly wage. If the employer clicked on a worker's "tile" they could see the worker's country, total hours worked on the platform, portfolio size, passed skills tests and past employer evaluations. Employers in the treatment group could choose to invite any number of the recommended workers to apply for their job (including no one at all).

Once a treated employer closed the recommendations pop-up window, they experienced the same interface and opportunities as the control group. Employers in both groups could use the marketplace search tools to find and invite other, non-recommended candidates to apply to their vacancies. If an employer invited a worker to apply, the invited worker received an email notification with a link to the employer's vacancy posting. The recruited worker could then decide whether or not to apply. After a vacancy was posted, non-recruited, "organic" workers were free to find the vacancy and apply.

Unfortunately, the actual recommended candidates were not logged. In order to identify which employer recruiting invitations were experimentally induced, we can examine the time at which they were made. We know the time that an employer posted a vacancy and the time each recruiting invitation was made, down to the millisecond. Because recommendations were presented as a pop-up immediately after a vacancy was posted (and the employer could not save the results), recruiting invitations made shortly after the posting were likely to have been experimentally induced. We define "recruiting" as the employer sending one or more invitations to apply within the first hour of posting their vacancy.

Figure 1: Recommendations shown to treated employers after posting their vacancy



Notes: This figure shows the interface presented to employers in the treatment group. It displays a number of recommended workers with good on-platform reputation, skills relevant to the employer's vacancy and predicted availability for the employer's project.

4 Conceptual framework

The algorithmic recommendations described in this paper are novel—third parties helping labor market participants is not. For example, in conventional labor markets, "labor market intermediaries" verify and disseminate match-relevant information about workers, such as educational, legal and credit histories (Autor, 2008). However, they generally do not advise firms on which specific workers to recruit or hire. Services are available to firms through corporate headhunters, but they are bespoke, paid services that simply outsource what a firm would have done with its own search—the headhunter is not a disinterested third party trying to increase the overall supply of information in the labor market.

Some active labor market programs have focused on explicit job-finding assistance for workers. They tend to have positive albeit modest effects on employment probability (Kluve, 2010; Card et al., 2010). However, a perennial concern with such programs is that the benefits mainly come at the expense of those not assisted. This concern is highlighted by Gautier et al. (2012) and was recently illustrated by Crépon et al. (2013), which was a large-scale job-finding assistance program that seemed to "work" mostly by displacing non-assisted job seekers.

In this paper, the labor market intervention was demand-focused, with assistance offered to employers. This kind of assistance is relatively rare in the literature, with most of the focus being on assisting unemployed workers. Perhaps not coincidentally, there is relatively little research in economics on how employers fill vacancies and thus little guidance on how worker-finding assistance should affect outcomes. However, the extant empirical studies support characterizing the employer vacancy-filling process as a search process. Barron and Bishop (1985) find that employers with hard-to-fill vacancies or those that require more training report screening larger pools of applicants and screening each applicant more intensively. Pellizzari (2011) finds that more intensive recruitment by a sample of British employers is associated with better quality matches. The resultant matches pay more, last longer and lead to greater employer satisfaction, though the direction of causation is not clear.

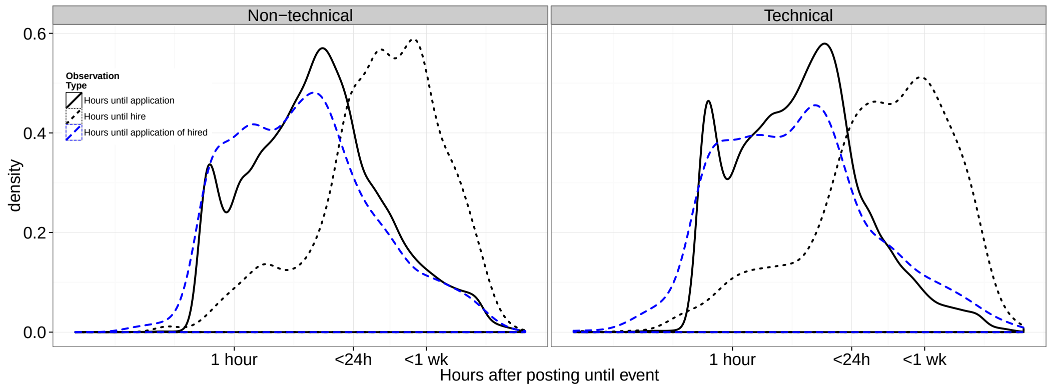
Existing models of employer search are similar to simple job search models. In these models, firms serially screen applicants without recall and hire the first applicant above their reservation ability for the position. The distribution of worker abilities is known and setting a reservation ability fully characterizes the firm's strategy. This is the modeling approach used by Burdett and Cunningham (1998); Barron et al. (1989); Barron and Bishop (1985). Below, I develop a model of employer search and screening that is also a dynamic screening model, but differs in a few important respects. Namely, firms screen a complete pool of applicants that they can order by expected productivity at no cost; to learn actual productivity the firm pays a marginal screening cost. The firm screens sequentially with recall or it can "call off" the search if the applicant pool does not look promising, i.e., if the expected value of screening the pool is below the extensive margin screening cost. Firms also have the possibility of ex ante recruiting to improve their applicant pool.

4.1 Model of employer search and screening

A firm receives a collection of applicants, A. Applicant i has productivity y_{ij}, which can be decomposed into two orthogonal components: $y_{ij} = a_i + \theta_{ij}$, where a_i is the applicant's general ability and θ_{ij} is the applicant's vacancy-specific ability. The firm can, without cost, directly observe each a_i. We might think of a_i as representing how good an applicant is "on paper" and θ_{ij} as the refinement the firm can discover from interviewing. To learn θ_{ij}, the firm must pay fixed screening cost, C_S, and a per-applicant screening cost, c. Marginal evaluation costs are the cost of performing another interview. Fixed costs are the costs of assembling a team to evaluate applicants and making an organizational decision about whom to hire, if anyone. For example, if the firm screens $x > 0$ applicants, total screening costs will be $C_S + cx$. The firm knows that θ_{ij} is drawn from a distribution with pdf $f(\cdot)$ and cdf $F(\cdot)$ and $\mathbf{E}[\theta] = 0$ and support $[-\infty, \infty]$.

The framework assumes the applicant pool be formed shortly after a vacancy is posted. This

Figure 2: Relative timing of vacancy applications and hires



Notes: This figure plots the kernel density estimates of the elapsed time until various vacancy-related events. Each curve is the distribution of hours elapsed from the time of vacancy posting. The solid line is hours until an application is made. The dotted line is hours until a worker is hired. The dashed line is hours until an application is made by the applicant that was ultimately hired.

assumption is quite realistic with respect to oDesk. It also appears to be the case in conventional markets: van Ours and Ridder (1992) "conclude that almost all vacancies are filled from a pool of applicants that is formed shortly after the posting of the vacancy." van Ommeren and Russo (2010) reach a similar conclusion. To demonstrate the plausibility of this assumption, within a category of work, we can plot the distributions of hours from job posting to application and the hours from job posting until first hire. If the firm were serially screening applicants upon arrival, the distribution of hours elapsed from job post until hire would be the same as the distribution of hours elapsed from job post until application (suitably right-shifted to account for the time it actually takes to screen each applicant). However, in Figure 2 we show that the distribution of elapsed time until hire is considerably right-shifted compared to the elapsed time until application. The modal application occurs less than 24 hours after posting, whereas the modal hire occurs nearly a week after posting. This right-shifting of hires is not due to selection—we also plot the distribution of the arrival time of the ultimately-hired applicant and show that it has approximately the same distribution as the entire pool of applicants.

Wages are determined by a bargaining game that determines how match surplus is shared.

Each worker has the same outside option, normalized to 0, as does the firm. The solution to the bargaining game is a simple sharing rule on the match surplus: the firm receives γy_i and the worker receives $(1-\gamma)y_i$. A firm will only hire a worker i if $y_i \geq 0$. We will assume that firms will not hire an applicant without screening him or her.

The firm, knowing the expected value of the screening process, can, ex ante, decide whether to recruit. Recruiting will increase the expected size of the applicant pool, but at a cost of C_R. We assume that the firm must recruit ex ante in order to have a complete pool of applicants to consider when the firm begins screening. Firms cannot draw applicants, then decide to recruit. While it is the case that the majority of all recruiting happens ex ante (the importance of ex ante recruiting can be clearly seen in the left-most spike of the bi-modal application arrival time distributions in Figure 2), this assumption of only ex ante recruiting is a simplification.

4.1.1 The firm's decision problem

The Bellman equation which describes the value of being able to hire someone immediately at productivity y and having a collection of unevaluated applicants, A, with the next-to-evaluate applicant having expected productivity a is

$$V(y,A) = \max\left\{ \underbrace{V(y, A \setminus a)}_{\text{Skips } a}, \underbrace{F(y-a)V(y, A \setminus a)}_{a+\theta \text{ worse than } y} + \underbrace{\int_{y-a}^{\infty} V(a+\theta, A \setminus a) f(\theta)\, d\theta - c}_{a+\theta \text{ better than } y} \right\}, \tag{1}$$

where $A \setminus a$ indicates A without the applicant represented by a. We will refer to $V(y,A)$ as the value of a vacancy with A unevaluated applicants when the firm has a current option of y. The firm always considers the applicant with the highest observable productivity among remaining applicants. The firm will rank applicants in descending order of observable ability and then evaluate them sequentially. This follows from the fact that $a_i \perp \theta_{ij}$, and so for any A, the applicant offering the highest probability of exceeding the firm's current option of y is the remaining applicant with the highest on paper productivity. Observe that firms do not "skip" applicants—once a firm chooses not to evaluate an applicant, it does not evaluate any further applicants. The reason is that all subsequent applicants offer lower expected productivity. Once the firm decides to stop evaluating candidates, the leftmost term in brackets in Equation 1, $V(y, A \setminus a)$, is equal to γy.

4.1.2 The firm's optimal policy and its implications

Lemma 1. *The firm only considers the productivity of the next unevaluated applicant when deciding whether to continue evaluating applicants.*

Proof. Consider the scenario in which the firm is indifferent between evaluating the next applicant and stopping screening, i.e.,

$$\gamma y = F(y-a)V(y, A \setminus a) + \int_{y-a}^{\infty} V(a+\theta, A \setminus a) f(\theta)\, d\theta - c. \tag{2}$$

If the firm does evaluate the next applicant, it will not evaluate any further applicants (since the next candidate has a lower observable productivity and the pool of unevaluated applicants will be smaller), and so $V(a+\theta, A \setminus a) = \gamma(a+\theta)$ (where $a+\theta$ is the realized productivity of the evaluated applicant) and $V(y, A \setminus a) = \gamma y$, which implies that $y = F(y-a)y + a + \mathbf{E}[\theta|\theta > y-a] - \frac{c}{\gamma}$. Since $A \setminus a$ does not enter into the equation, changes in $A \setminus a$ cannot affect the firm's decision-making; only the values of a and y are relevant to the firm. □

Lemma 2. *The value of a vacancy is non-decreasing in the productivity of the firm's current best option.*

Proof. Let the firm have a current option of y and a collection of unevaluated applicants A. If the firm stops screening, then $\frac{\partial V(y,A)}{\partial y} = 1$. If the firm continues screening, then

$$\frac{\partial V(y,A)}{\partial y} = \frac{\partial}{\partial y}\left[F(y-a)V(y,A\setminus a) + \int_{y-a}^{\infty} V(a+\theta, A\setminus a) f(\theta)\, d\theta - c\right] = \frac{\partial V(y,A\setminus a)}{\partial y}. \tag{3}$$

Unsurprisingly, the effect of higher y when the firm has A applicants and chooses to evaluate is whatever effect the higher y has when the firm just has $A \setminus a$ applicants. Because $\frac{\partial V(y,\emptyset)}{\partial y} > 0$ for all $y > 0$, then by mathematical induction we can show that $\frac{\partial V(y,A)}{\partial y} > 0$. □

Lemma 3. *An increase in the observable productivity of any applicant in the pool increases the value of the vacancy to the firm.*

Proof. We consider an arbitrary applicant a_k of A increasing in observable productivity, where k is the applicant's position in the rank ordering of applicants. Because of Lemma 1, the firm's decision-making is unaffected for the first $k-1$ applicants. We only have to consider what happens to the value function at $V(y_k, A(k))$ where y_k is the current option after evaluating the first $k-1$ applicants and $A(k)$ is the pool of remaining applicants, i.e., $A(k) = a_k, a_{k+1}, a_{k+2}, \ldots, a_n$. The partial derivative of the value function with respect to a_k is

$$\frac{\partial}{\partial a_k}\left[F(y_k - a_k)V(y_k, A(k)\setminus a_k) + \int_{y_k-a_k}^{\infty} V(a_k+\theta, A(k)\setminus a_k) f(\theta)\, d\theta - c\right] = \int_{y_k-a_k}^{\infty} V^{(1,0)}(a_k+\theta, A(k)\setminus a_k) f(\theta)\, d\theta,$$

and by Lemma 2, $V^{(1,0)}(a+\theta, A(k)\setminus a_k) \geq 0$ and hence $\frac{\partial V(y_k,A(k))}{\partial a_k} \geq 0$. Because the value function is recursively defined, we can conclude that $\frac{\partial V(y,A)}{\partial a_k} \geq 0$. □

4.1.3 Determinants and effects of recruiting

One interpretation of the experimental intervention is that it lowered the firm's cost of recruiting, C_R. Unsurprisingly, in the model, firms are more willing to recruit when the costs of recruiting are low (Proposition 1). This recruiting in turn increases the probability that the vacancy is filled (Proposition 2). However, not all recruiting is equally effective—the gain from recruiting is declining in the firm's current best option (Proposition 3). An implication of this differential effectiveness is that firms that expect to get fewer applicants or worse applicants are more likely to recruit and that those same types of firms are likely to see more benefit from recruiting. Recruiting does not necessarily increase even the expected productivity of the realized match. This seems counter-intuitive—particularly since recruiting unambiguously increases the value of the vacancy to the employer. The reason is that a better option—which recruiting provides at some point in the firm's screening—can decrease screening on the intensive margin, which mechanically lowers expected realized productivity. However, simulations at least show that, in expectation, recommendations should increase realized match quality.

Proposition 1. *A firm's probability of recruiting is decreasing in the cost of recruiting.*

Proof. A firm recruits when the increase in the value function exceeds the cost of recruiting: $\mathbf{E}[V(0, A \cup a)] - V(0, A) > C_R$. If C_R decreases, the firm is willing to recruit for a larger collection of applicant collections, A. $\square$

Proposition 2. *Recommendations increase the expected value of a vacancy and the likelihood that it will be filled.*

Proof. Suppose a recommended applicant has observed productivity a'_k. Adding a_k to the applicant pool creates a new applicant pool $a_1, a_2, \dots a'_k, a_k, a_{k+1} \dots a_n$. This new applicant pool is as good or better than the old applicant pool at every position: $a_1 = a_1, a_2 = a_2, \dots a'_k > a_k, a_k > a_{k+1}$. Note that the new pool has one more element at the end, which can only increase the value function. The pool with the recommendations is better than the original pool at every position that is greater than or equal to the original k position, since the recommended applicant shifts every applicant down one position. By repeated application of Lemma 3, we can show that this everywhere-better applicant pool offers a higher value function than the original. If $V(y, A)$ is high, it is more likely to exceed the fixed costs of evaluation, and if the firm evaluates, it is more likely to fill a vacancy.[1] □

Proposition 3. *The return to an additional applicant is decreasing in the firm's current outside option.*

Proof. If the firm screens, the returns to the ability of the next applicant is

$$\frac{\partial V(y,A)}{\partial a} = \int_{y-a}^{\infty} \frac{\partial V(a+\theta, A \setminus a)}{\partial a} f(\theta)\, d\theta.$$

If we differentiate with respect to the firm's current option, we get $\frac{\partial^2 V(y,A)}{\partial a \partial y} = -f(y-a)\frac{\partial V(y,A\setminus a)}{\partial y}$. By Lemma 2, $\frac{\partial V(y,A\setminus a)}{\partial y} > 0$, and since $f()$ is a pdf, $\frac{\partial^2 V(y,A)}{\partial a \partial y} \leq 0$. □

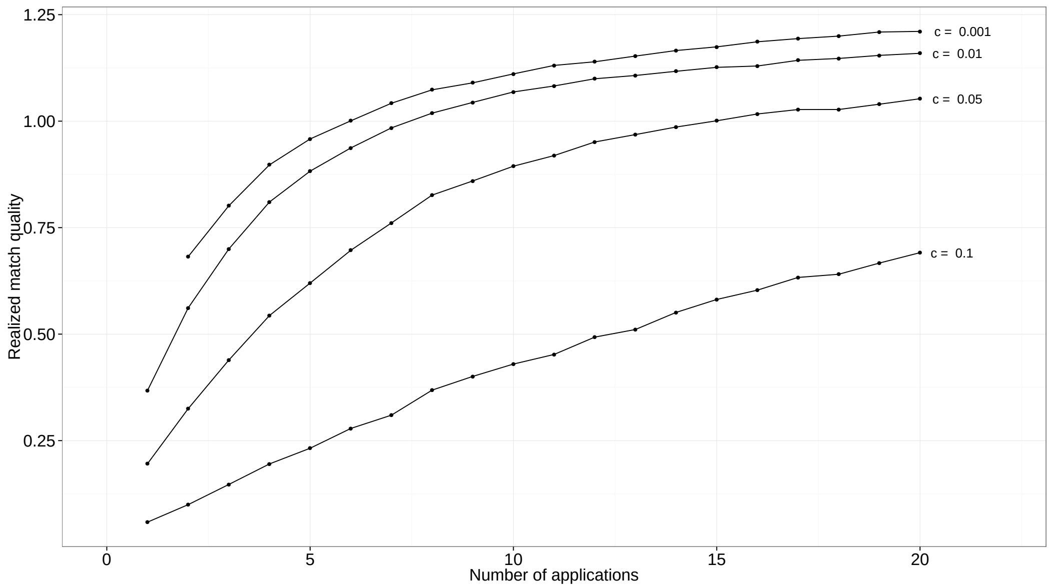
Observation 1. *A better current option does not necessarily increase the expected productivity of the hired applicant.*

We can see this most clearly with a simple example. Consider a scenario with a single unevaluated applicant with a current value of y_0. The firm screens if $\int_{y_0-a}^{\infty}(a+\theta)f(\theta)\,d\theta > c$. This condition defines a cut-off option of $\underline{y}(a)$: if y_0 is above $\underline{y}(a)$, the firm does not screen; if y_0 is equal to or below the cutoff, the firm screens. As such, an ϵ increase in y_0 can cause the firm to reduce screening, which lowers expected match quality since screening always has a positive effect on expected match quality.

Despite the possibility that recruiting lowers realized match quality, simulations suggest that, in expectation, recommendations would increase realized match quality. I assume that for each applicant, a is uniformly distributed on $[0,1]$ and that θ is uniformly distributed on $[-0.5, 0.5]$. For each iteration of the simulation, I draw a new applicant pool and have the firm screen that pool optimally, given the realized abilities and a fixed marginal cost of screening—I ignore the fixed costs of screening. Figure 3 shows the simulation results, using 10,000 simulations per scenario and taking the mean match quality. The y-axis is the mean realized match

[1] Note that while recommendations do not necessarily increase realized match productivity (as we saw from the example), this is only because recommendations can cause the firm to economize on the intensive margin; this point is not relevant when considering whether the firm screens at all.

Figure 3: Simulated realized match quality versus the number of applicants



Notes: This figure plots realized match quality versus applicant pool size assuming the firm solves thy dynamic programming problem described in Equation 1, under different assumptions about the marginal screening cost, c. In each simulation, $a \sim U[0,1]$ and $\theta \sim [-0.5, 0.5]$.

quality versus the size of the applicant pool. The various lines correspond to different values of c. We can see that across screening values, more applicants is associated with a higher mean realized match quality.

Observation 2. *An additional applicant decreases an existing applicant's expected probability of being hired if the existing applicant would have been hired with some positive probability in the absence of the additional recommended applicant.*

To see this, consider some recommended applicant a_k at position k. With some probability, all existing applicants in A have a lower expected productivity than a_k. For all applicants with productivity below a_K, the firm has a higher expected option y at position k than it would have in the absence of a_k. As such, all applicants after k are less likely to be screened and thus hired. For applicants before k, screening and hiring probability are unchanged because the optimal policy is myopic (by Lemma 1).

Observation 3. *An additional applicant can increase the probability that some other applicant is hired if there is some chance the firm would have called off the search but for the additional applicant.*

If $V(A, y) < C_S$, the firm calls off the search and no one is hired. From Lemma 3, we know that $\mathbf{E}[V(A \cup a, y)] > V(A, y)$, and so it is possible that with an additional applicant, the firm screens, in which case someone is potentially hired—including applicants other than the additional applicant.

5 Experimental results

If the treatment lowered recruiting costs, the model makes two unambiguous predictions with respect to the research questions: (1) the treatment increases recruiting, which in turn will (2) increase the vacancy fill rate. The model further predicts that firms with lower numbers of organic applicants will experience the greatest effect on fill rates from the treatment. On the other two research questions, the model makes ambiguous predictions. More applicants can lower match quality (though simulations suggest this is unlikely) and the degree of crowd-out depends upon how much it is offset by increased employer screening. The relationships between the research questions and the model predictions are summarized in Table 2.

The key experimental results are that (1) the treatment increased recruiting among all vacancies; (2) the treatment increased fill rates among technical vacancies only, which is consistent with the model prediction of greater effects for vacancies with fewer organic applicants; (3) matches induced by the treatment did not measurably differ from non-induced matches; (4) rather than crowd-out there was, if anything, complementarity between organic and recruited applicants. Because this paper reports the results from a true experiment, we can readily see most of these results simply by comparing means across the treatment and control groups. In Table 3, we report means for our main outcome variables. The top panel of the table uses all

Table 2: Research questions and the model predictions

Research Question	Model Prediction
Does the experiment increasing recruiting?	Yes (if the treatment reduced recruiting costs), by Proposition 1
Does an additional applicant increase the fill rate?	Yes, by Proposition 2. Proposition 3 suggests effects could be heterogeneous based on attributes of the recruiting employer.
Does an additional applicant increase the match quality?	Ambiguous—Observation 1 showed that recommendations do not necessarily increase match quality, but simulations suggest that expected match quality is higher when there are more applicants.
Does recruiting crowd out non-recommended applicants?	Ambiguous—Observation 2 gives conditions under which crowd-out is implied, but Observation 3 explains an alternate channel that can increase hiring of non-recommended candidates.

vacancies as the sample. The middle and bottom panels use the same outcomes but the samples are restricted to non-technical and technical vacancies, respectively. The remainder of this section elaborates on these results. Every regression in all tables was estimated using ordinary least squares, and the standard errors are heteroscedasticy-robust and calculated using the HC3 method proposed by MacKinnon and White (1985), unless otherwise noted.

5.1 Employer recruiting

The treatment had a strong positive effect on employer recruiting. The top panel of Table 3 shows that 15.2% of employers in the control group recruited, while 21.4% in the treatment group did.[2] In Table 4, we report several regressions where the outcome variable are measures of employer recruiting. Column (1) simply recapitulates what is shown in the means compari-

[2]Note that although employers could make any number of invitations, we use the extensive margin measure of recruiting, i.e., an indicator for any recruiting at all. Most employers send only one recruiting invitation conditional upon sending any, so little information is lost by truncating the count. Furthermore small number of employers send a very large number of invitations, making the standard error of the mean count of invitations problematically large.

Table 3: Outcome means by experimental groups

	Treatment mean: $\bar{X}_{TRT}$	Control mean: $\bar{X}_{CTL}$	Difference in means: $\bar{X}_{TRT} - \bar{X}_{CTL}$	p-value	
Outcomes for All Vacancies					
Recruited	0.214 (0.008)	0.152 (0.007)	0.062 (0.010)	<0.001	***
Filled vacancy	0.307 (0.009)	0.287 (0.008)	0.020 (0.012)	0.092	†
Hired an organic applicant	0.284 (0.008)	0.277 (0.008)	0.007 (0.012)	0.547	
Hired recruited applicant	0.051 (0.004)	0.041 (0.004)	0.009 (0.005)	0.083	†
Wage bill > $500	0.075 (0.005)	0.060 (0.004)	0.015 (0.007)	0.025	*
Outcomes for Non-Technical Vacancies					
Recruited	0.210 (0.011)	0.160 (0.010)	0.050 (0.014)	0.001	***
Filled vacancy	0.355 (0.013)	0.362 (0.013)	-0.007 (0.018)	0.699	
Hired an organic applicant	0.333 (0.013)	0.345 (0.012)	-0.012 (0.018)	0.509	
Hired recruited applicant	0.049 (0.006)	0.051 (0.006)	-0.002 (0.008)	0.803	
Wage bill > $500	0.063 (0.006)	0.053 (0.006)	0.010 (0.009)	0.249	
Outcomes for Technical Vacancies					
Recruited	0.218 (0.011)	0.145 (0.009)	0.073 (0.014)	<0.001	***
Filled vacancy	0.262 (0.011)	0.219 (0.010)	0.043 (0.015)	0.006	**
Hired an organic applicant	0.237 (0.011)	0.215 (0.010)	0.022 (0.015)	0.143	
Hired recruited applicant	0.053 (0.006)	0.033 (0.004)	0.020 (0.007)	0.006	**
Wage bill > $500	0.086 (0.007)	0.067 (0.006)	0.019 (0.010)	0.046	*

Notes: This table reports outcome means and standard errors across experimental groups from the recommendations experiment. In the experiment, the treatment group of employers received algorithmically-generated recommendations of candidates for their vacancies while the control group did not. The unit of randomization was the posted vacancy. The standard error for each calculated mean is in parentheses next to the estimate. The top panel means are calculated using all observations. The middle panel uses only non-technical vacancies, while the bottom panel uses only technical vacancies. Reported p-values are the for two-sided t-tests of the null hypothesis of no difference in means across groups. "Recruited" is an indicator for whether the employer made an invitation to a worker within 1 hour of posting their job. "Filled vacancy" is an indicator for whether the employer hired a worker and spent more than $1. A "recruited applicant" is one one that applied to the vacancy after being recruited by the employer, whereas an organic applicant is one who applied without prompting by the employer. The "wage bill" is the total amount spent by the employer against that vacancy. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

son table. In Column (2) we interact the treatment indicator with an indicator for whether the vacancy was technical. Although the treatment had a larger effect among technical vacancies, this difference is not significant. The baseline levels of recruiting were similar across vacancy types. Figure 4 shows this graphically, plotting control and treatment levels of recruiting for technical and non-technical vacancies separately, as well as levels for all vacancies pooled together.

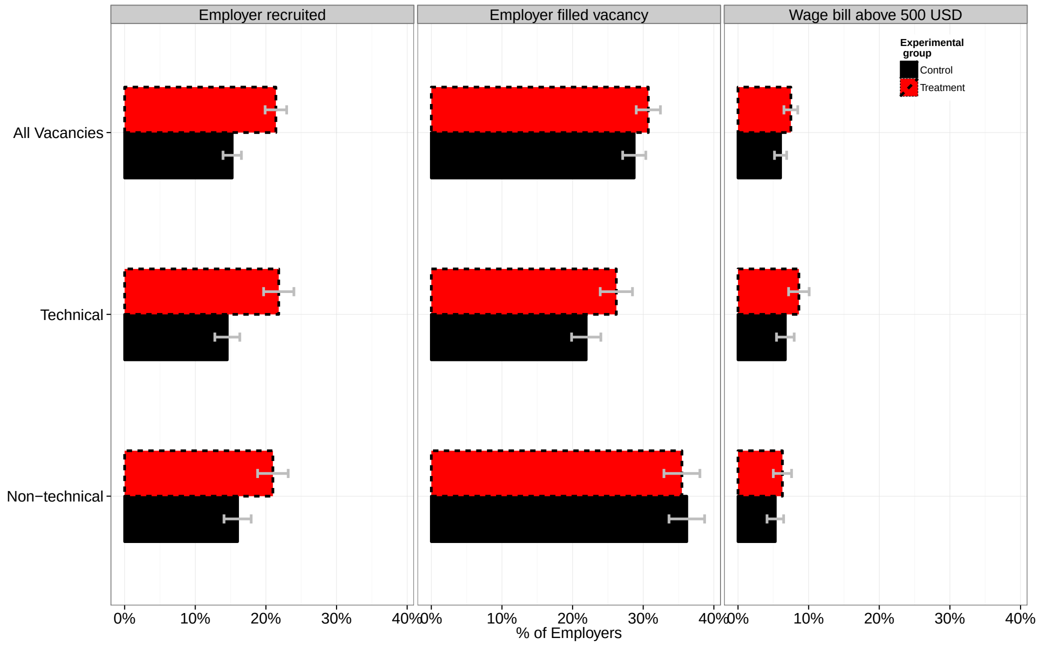
Table 4: Effects of the recommendations treatment on employer recruiting

	Employer recruited workers to apply:			
	(1) *DV*= Any (within 1st hour)	(2) *DV*= Any (within 1st hour)	(3) *DV*= Any (after 1st hour)	(4) *DV*= Any (within first hour)
Intercept	0.152***	0.160***	0.128***	0.331***
	(0.007)	(0.010)	(0.006)	(0.030)
Assigned to treatment	0.062***	0.050***	−0.004	0.050***
	(0.010)	(0.014)	(0.009)	(0.014)
Technical vacancy		−0.015		−0.024
		(0.013)		(0.014)
Treatment × Technical vacancy		0.023		0.020
		(0.020)		(0.020)
Log (number of organic applications + 1)				−0.055***
				(0.004)
Log (job post length in characters + 1)				−0.008
				(0.005)
Employer wanted prior oDesk experience				0.044**
				(0.015)
N	5,953	5,953	5,953	5,953
R-squared	0.006	0.007	0.000	0.038

Notes: This table reports several regressions where the outcome variable are measures of employer recruiting, using data from the recommendations experiment. Each regression was estimated using OLS and robust standard errors are reported. The dependent variable in Column (1), (2) and (4) is whether the employer recruited any workers (within one hour of posting their vacancy) to apply for their job. The Column (2) regression is the same as (1) but includes an indicator for whether the vacancy was for a "technical" job (i.e., required computer programming). This indicator is also interacted with the treatment indicator. In Column (3) the dependent variable is whether the employer sent any recruiting invitations later than hour after posting their vacancy. In Column (4), several pre-treatment controls are included in the regression, including details about the number of organic applications eventually received. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

In addition to recommending applicants, the treatment may have also simply alerted employers to the possibility of recruiting. Recall that employers were new to oDesk. However, the evidence does not support this interpretation. In Table 4, Column (3), we regress an indicator for whether the employer made any "late" (i.e., after the first hour of posting) invitations on the treatment indicator. The coefficient on the treatment indicator is negative, small in magnitude

Figure 4: Effects by experiment group and vacancy type



Notes: This figure shows means of various outcomes measures in the treatment and control groups (illustrated by color and bar outline), by vacancy type. Each error bar is a 95% confidence interval, using the $\sqrt{p(1-p)N^{-1}}$ approximation for the standard error of a proportion (where p is the proportion and N is the sample size). In the left panel labeled "Employer recruited" the outcome measure is whether the employer sent at least one invitation within the first hour of posting. In the middle panel labeled "Employer fill vacancy" the outcome measure is whether the employer filled their vacancy, as measured by spending at least some money against that vacancy. In the right panel labeled "Employer total wage bill > \$500" the outcome measure is whether the employer spent more than \$500 against that vacancy.

and insignificant—contrary to what we would expect if the treatment was informing employers about the potential for recruiting. Comparing intercepts across Columns (1) and (3), we can see that a slight majority of recruiting is ex ante (i.e., happens within the first hour of job posting).

5.1.1 Worker responsiveness to employer recruiting invitations

After employers send recruiting invitations, the receiving workers decide whether to apply. Table 5, Column (1) reports a regression where the outcome variable is an indicator for whether one or more recruited applicants applied. Reassuringly, the treatment substantially increased the probability that an employer had one or more recruited applicants in their pool. Invitation acceptance does nor appear to be mediated by the vacancy type, nor does the vacancy type seem to matter.

In addition to aggregate numbers of employers with recruited applicants in their pool, we can also look at invitation response rates by treatment group. Invitation response rates provide a measure of recommendation quality. If the treatment was only providing poor recommendations that employers were acting on without much consideration, then we would expect a lower invitee response rate in the treatment group. In Table 5, Column (2), we regress an indicator for whether a *recruiting* employer had any accepted recruiting invitations. We can see that there is no appreciable difference in invitation acceptance rates by treatment status nor by vacancy type.

One concern with any marketplace experiment is that SUTVA is violated. This possibility is discussed in Appendix A. There is little reason to worry about the experiment having market-moving effects on price, entry etc., given that only a small fraction of total vacancies received recommendations, but there is some concern and evidence that the treatment depleted the supply of "recruitable" workers.

5.2 Employer hiring

The model predicts that recommendations should increase the vacancy fill rate. In Table 3, we can see that overall, the fraction of employers hiring was higher in the treatment. The overall increase was about 2% points, or a little less than 7% improvement from the control. This effect is significant at the 10% level, but not at the 5% level. However, the effect of the treatment on fill rates is much more pronounced among technical vacancies. The "Employer hired" panel in Figure 4 shows that technical vacancies are wholly responsible for the positive overall effect—the control group fill rate is slightly lower in the treated group among non-technical vacancies.

Although the technical and non-technical distinction is central on oDesk and it is reason-

Table 5: Worker responsiveness to recruiting invitations

	Employer has at least one recruited applicant in applicant pool:			
	(1)	(2) Conditional upon recruiting	(3)	(4) Conditional upon recruiting
Intercept	0.108***	0.674***	0.478***	0.415
	(0.008)	(0.031)	(0.120)	(0.441)
Assigned to treatment	0.042***	0.040	−0.179	0.223
	(0.012)	(0.041)	(0.179)	(0.581)
Technical vacancy	−0.007	0.020	−0.011	0.448
	(0.011)	(0.043)	(0.159)	(0.621)
Treatment × Technical vacancy	0.015	−0.009	0.384	−0.438
	(0.017)	(0.056)	(0.245)	(0.807)
Week			−0.015**	0.011
			(0.005)	(0.018)
Treatment × Week			0.009	−0.008
			(0.007)	(0.024)
Technical × Week			0.000	−0.018
			(0.006)	(0.026)
Treatment × Technical × Week			−0.015	0.018
			(0.010)	(0.033)
N	5,953	1,086	5,953	1,086
R-squared	0.006	0.002	0.012	0.002

Notes: This table reports several regressions where outcome varaiables are measures of worker responsiveness to the employer invitations, using data from the recommendations experiment. The unit of randomization was the posted vacancy. In Columns (1) and (3), the outcome variable is whether the employer has any recruited applicants apply to their vacany. This dependent variable is zero both for employers that did not recruit and for employers that recruited but whose invitations were rejected. In Columns (2) and (4), the outcome variable is the same, but the sample is restrcited to only those employers that recruited. Each regression was estimated using OLS and robust standard errors are reported. The regressions reported in Columns (3) and (4) include controls for time (the week the vacancy was posted) and their interactions with the vacancy type and treatment. Significance indicators: $p \le 0.05 : *$, $.p \le 0.01 : **$ and $p \le .001 : ***$..

able to condition treatment effects on vacancy type, there appears to be an economic basis for the differential effects rather than something inherent in the nature of the work. In Table 6, Column (1), we regress an indicator for whether the vacancy was filled on the treatment indicator interacted with the number of organic applicants the vacancy ultimately received. This is a somewhat problematic regression, in that the number of organic applicants could in principle have been affected by the treatment. However, the regression of counts on the treatment is not significant. We see that the treatment indicator is not strong and positive, while the treatment and application count interaction term is strongly negative. This suggests that technical/non-technical heterogeneity in treatment effects reflects differences in organic application counts, which in turn affect the marginal return to another applicant (recall Proposition 3).

The overall positive effect of the treatment on fill rates—at least for technical vacancies—implies that the treatment did not simply cause a 1-for-1 swap of counter-factually hired organic applicants with experimentally induced recruited applicants. Had this occurred, the fill rate would have remained the same. However, the overall rise in fill rates does not imply no crowd-out. We can look for evidence of crowd-out by decomposing the "fill" measure into indicators for whether a vacancy was filled by a recruited or organic applicant.

To begin our analysis of crowd-out, we first note that the effect on fills among technical vacancies shown by Column (2)—a 4.3% point increase—is a large effect. In fact, it is too large to be explained solely be the direct channel of increased recruiting—to obtain an effect that large, hiring of organic applicants also had to increase. To see why, recall that the treatment increased recruiting by 7.3% points among technical vacancies. Technical vacancies in the control group had a baseline recruiting rate of 15% and a fill-from-recruits rate of 3.8%. If we assume that the marginal increase in recruits from the treatment "converted" to hires at the same rate (roughly 1/4) as the "natural" recruits, the effect size should be $(1/4)0.07 \approx 1.8\%$. This is quite close to the effect the treatment has on the hiring of recruits, which is a 2.0% point increase (see the Column (3) regression where the outcome is an indicator for hiring a recruited applicant). The additional increase in the fill rate is due to increased hiring of organic applicants, which is shown in Column (4).[3] This effect is not conventionally significant—it is quite plausible that the treatment effect on the hiring of organics was simply due to chance. However, the relevant null

[3]Note that these two measures of hiring do not precisely sum to the overall fill rate because a small number of employers make multiple hires.

Table 6: Effects of the recommendations treatment on hiring

	All Vacancies	Technical Vacancies Only		
	(1)	(2)	(3)	(4)
	DV= Employer hired	*DV=* Employer hired	*DV=* Employer hired **recruited** applicant	*DV=* Employer hired **organic** applicant
Intercept	0.253***	0.219***	0.033***	0.215***
	(0.010)	(0.010)	(0.004)	(0.010)
Assigned to treatment	0.037**	0.043**	0.020**	0.022
	(0.014)	(0.015)	(0.007)	(0.015)
Num. Organic Apps	0.002***			
	(0.000)			
Treatment × Num. Organic Apps	−0.001*			
	(0.001)			
N	5,953	3,081	3,081	3,081
R-squared	0.008	0.002	0.002	0.001

Notes: This table reports several regressions where the outcome variables are measures of employer hiring, using data from the recommendations experiment. Each regression was estimated using OLS, and robust standard errors are reported. The key independent variable across across regressions is the indicator for whether the vacancy was assigned to the treatment group. The dependent variable in each of these regressions is whether or not the employer hired a worker of a particular type. In Columns (1) and (2) the indicator is for hiring anyone at all. In Column (1) the sample is all vacancies, while in Column (2) the sample consists of only technical vacancies. In Column (3) the outcome is whether the employer hired a recruited applicant, while in Column (4) the outcome is whether the employer hired an organic applicant. In both Columns (3) and (4) the sample is restricted to technical vacancies. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

hypotheses for the crowd-out effect is not that it is zero—we have a priori reasons to suppose it would be negative, notwithstanding the possibility of the treatment stimulating screening on the intensive margin.

5.2.1 Employer screening

As an intermediate step between receiving an applicant pool and hiring, the firm screens applicants. While the model does make predictions about screening, unfortunately analyzing screening does not permit us to directly test whether recommendations are increasing screening on the extensive margin. There are countervailing but not separately identifiable effects.

That limitation aside, it is still useful to consider whether gross crowd-out (or complementarity) is detectable at the screening stage. We have two measures—both imperfect—of screening: whether an employer "views" an application (i.e., the employer clicks on their application within the applicant tracking system) and whether an employer formally contacts a recruited applicant for further discussion.[4]

To start, let us decompose control vacancies into four groups: those who (1) hire an organic applicant (2) hire a recruited applicant (3) hire no one but will hire under recommendations (4) hire no one regardless of treatment. Under the treatment, crowd-out of organic applicants comes from vacancies in group 1 hiring a recruited applicant instead; complementarity comes from vacancies in 3 being tilted to screening and then hiring an organic applicant. If we have some measure of screening, crowd-out is a bit more complex: recommendations can crowd-out organically screened applications in both 1 and 2 and increase screening in 3.

In Table 7, Column (1), we report a regression of the per-vacancy fraction of all applicants viewed on the treatment indicator. We can see from the intercept that a large fraction of applicants are completely ignored. Among technical vacancies, a substantially greater fraction of applicants are viewed. We will return to this point later, as it provides an explanation for why the recommendations experiment was only effective among technical vacancies—the marginal application is apparently worth more to employers. The treatment has no significant effect on the overall level of screening.

In Column (2) we restrict the sample to only organic applicants (the sample size is small, reflecting that not all vacancies receive organic applicants). If organic crowd-out was occurring because of the treatment, presumably the fraction of organic applications viewed would be substantially lower in the treatment. However, the treatment has no significant effect on the overall level of screening, though the sign of the point estimate of the treatment effect for tech-

[4]For a complex but uninteresting oDesk technical reason, we cannot separately look at interviewing for recruited applicants: in the database, all recruited applicants that accept an invitation are recorded as having been interviewed.

nical vacancies is negative. In Column (3), instead of the fraction, we use as our outcome the count of organic applicants screened. As with the fraction of applicants, there are no significant effects, though in this case the point estimate of the treatment effect for technical vacancies is negative.

Given the overall lack of treatment effects on screening, one might wonder if crowd-out is really a concern. In Column (4) we report a regression where the outcome variable is the fraction of organic applicants interviewed, but with the sample restricted to only those employers that sent a single recruiting invitation. One of the regressors is an indicator for whether that single recruited worker accepted. The coefficient on this regressor is strongly negative: The baseline interview rate is 25.4% for organic applicants when the recruited worker rejects the invitation—it falls to only 18.8% when that recruited worker accepts. Of course, whether or not the recruited worker responds is clearly endogenous, but a substantial amount of the variation in the response rate is also presumably driven by idiosyncratic factors (e.g., is the invited worker busy, on vacation etc.). To the extent that a substantial fraction of this variation is as good as random, this regression does show that recruits can crowd-out organic applicants.

5.2.2 How much crowd-out should we have expected?

To predict how much crowd-out should have occurred, we need to know several unknowns. Among those employers that hired an organic applicant in the control, what fraction would have recruited had they been in the treatment and what fraction would have instead hired an organic applicant? We also need to know the same two parameters for employers in the control that did not fill their vacancies. Unfortunately, these parameters are not identified by the experiment. However, we can apply several heuristics and make simplifying assumptions to roughly estimate the expected magnitude of crowd-out.

We can partition control group employers into three groups: those that (1) hired from or-

Table 7: Effects of the recommendations treatment on measures of employer screening

	DV = Fraction of applications viewed by employer:		*DV* = "X" of organic applications interviewed by employer:	
	(1) All applicants	(2) Organic applicants	(3) "X" = Count	(4) "X" = Fraction (among those recruiting just 1)
Intercept	0.637*** (0.010)	0.620*** (0.010)	1.928*** (0.093)	0.223*** (0.028)
Assigned to treatment	0.002 (0.014)	0.007 (0.014)	0.151 (0.154)	0.006 (0.034)
Technical vacancy	0.056*** (0.013)	0.063*** (0.014)	−0.004 (0.124)	0.052 (0.036)
Treatment × Technical vacancy	−0.003 (0.019)	−0.019 (0.020)	−0.094 (0.197)	0.036 (0.052)
Recruited accepted				−0.075** (0.027)
N	5,526	5,170	5,953	448
R-squared	0.006	0.006	0.000	0.033

Notes: This table reports several regressions where the outcome variables are measures of employer screening of applicants. In Columns (1) and (2) the outcome varaible is the fraction of applicants that were "viewed" by employers, in the sense that the employer looked at their application, however briefly. In Column (1) the dependent variable is the fraction of all applications, while in Column (2) the dependent variable is the fraction of organic applications viewed. Each regression was estimated using OLS, and robust standard errors are reported. The sample size is smaller in Column (2) because not all vacancies receive organic applications. Columns (3) and (4) report regressions where the outcome variables are measures of employer interviewing of organic applicants. In Column (3), the outcome is the count of applicants, while in Column (4) the outcome if the fraction of organic applicants interviewed. In Column (4) the sample is restricted to vacancies in which the posting employer sent only one recruiting invitations. The predictor "Recruited accepted" is an indicator for whether or not that single recruiting invitation was successful—i.e., did the invited worker apply. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

ganic applicants, (2) did not hire at all and (3) hired from a recruit. Let us assume homogeneous treatment effects across all employer groups. In the control, the baseline hiring rate of recruited candidates was negligible, so we will assume that all increases in the hiring of recruits came from employers segments 1 and 2—those that would have not filled their vacancies or hired an organic candidate. We will also assume that all vacancies convert recruits into hires at the same rate—approximately 1/4 (taken from technical vacancy control group).

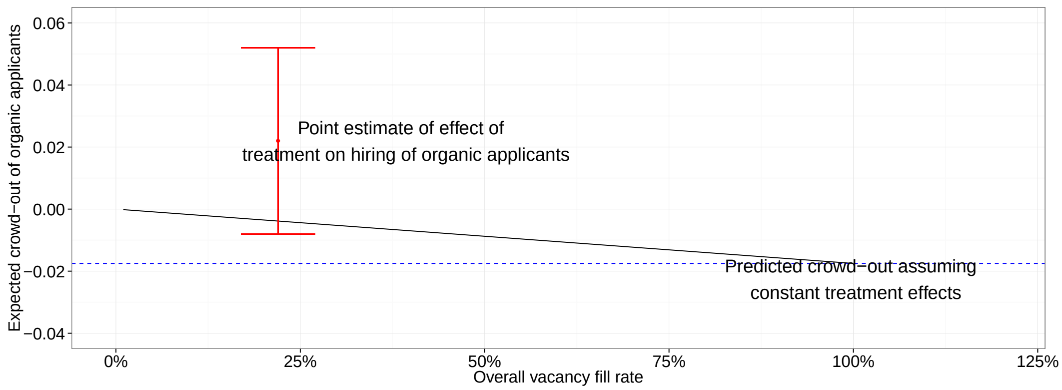
Using the 7% point effect of the treatment on recruiting, we plot the implied crowd-out of organic applicants as a function of the overall vacancy fill rate Figure 5. Note that as the overall fill rate increase, we will have more crowd-out, as less of the gains in hiring come from otherwise unfilled vacancies. Overlaid on the plot is the point estimate for the effect of the treatment on the hiring of organic applicants, as well as a 95% CI. This CI is plotted at the overall vacancy fill rate in the control group. We can see that the CI for the actual crowd-out includes the predicted crowd-out from our estimate, but barely.

Under these assumptions, crowd-out was unlikely to be very important in this setting—only 1/5 of vacancies are filled, which means that there is a substantial pool of unfilled vacancies for the recommendations to "work" on without displacing anyone. If there was a higher baseline vacancy fill rate, eventually crowd-out effects would dominate any complementarity effects predicted by the model. However, if some substantial fraction of vacancies were posted by "tire kickers" that were impervious to the treatment, then the implied crowd-out would be greater—there would be fewer unfilled vacancies capable of being filled via the intervention. The most extreme assumption is that all unfilled vacancies remaining in the treatment are posted by tire-kickers, in which case the vacancy fill rate of 24% is an upper-bound. Under this assumption, the predicted crowd-out would be far more substantial—we plot this amount as a dashed blue line on Figure 5. The CI does not contain the point estimate.

5.3 Match attributes

We find no evidence that matches induced by the treatment are better or worse than "natural" matches. None of the measures that could be interpreted as measuring match quality are conventionally significant. However, there are two caveats to these negative results. First, we can measure only certain outcomes for matches that actually occurred, and given the low over-

Figure 5: Predicted crowd-out of non-recruited organic applicants as a function of vacancy fill rate, versus actual observed effects



Notes: This figure shows the predicted relationship between the overall vacancy fill rate and the amount of crowd-out of organic applicants. The prediction is predicated on there being constant uptake of recruited and constant treatment effects across all vacancy outcome "types" in the control group. The black, downward sloping line shows the point prediction of crowd-out versus the overall vacancy fill rate. At the observed overall vacancy fill rate in the control group, the point estimate and 95% for the empirical crowd-out effect is plotted. The dashed horizontal line illustrates the predicted crowd-out when all remaining unfilled vacancies in the treatment are unfillable and all gains in fill rate had to come from existing organic applicants.

all fill rate, we have considerably less statistical power when talking about match outcomes. Second, we cannot separately compare induced and non-induced matches—we can compare only some mixture of induced and non-induced matches in the treatment to the non-induced matches found in the control.

Although a match has many attributes that might be thought of as measures of quality, perhaps the most important is the surplus that it generates. We do not directly observe match surplus, but under the assumptions of the model, the wage bill is proportional to the surplus. This allows us, at least in principle, to claim that the treatment's effect on the wage bill provides information about the treatment's effect on surplus.

One measure of the treatment's effect on the wage bill is an indicator for whether the wage bill exceeded some threshold. This avoids the conditional-on-positive problem and lets us examine all vacancies. It also lets us test whether the treatment raised fill rates merely by creating degenerate matches that, say, paid only $1 and were very short-lived. This possibility is not the case: recall from Table 3 that vacancies in the treatment were significantly more likely to have a wage bill exceeding \$500. In fact, even though the overall fill rate was not significant, this \$500 wage bill measure was significant overall. Among technical vacancies, the effect is smaller than the effect on overall fill rate, though the confidence intervals overlap substantially.

As we noted, the wage bill is a problematic outcome variable, in that we observe it only if a match occurs: any "treatment effect" estimated with a selected sample of matches will presumably capture both selection effects and also whatever causal effect the treatment has on average match quality for employers that would have filled their vacancy regardless of the treatment. To clarify the problem, it is useful to partition employers into the Imbens and Angrist (1994) types: there are (1) "always-taker" employers that fill their vacancies regardless of assignment, (2) "never-takers" that never fill their vacancies and the (3) "compliers" that fill their vacancies when in the treatment group but not when in the control group. We assume there are no "deniers" that fill their vacancies when in the control group but not when in the treatment group.

We obtain an estimate of the fraction of always-takers from the fill rate in the control group; we also obtain the mean wage bill for always-takers from this group. However, in the treatment, all we observe is the mean wage bill for always-takers under the treatment—which could be influenced by the treatment—co-mingled with the compliers under the treatment. Intuitively, we cannot decompose higher wage bill matches in the treatment into the contributions from always-takers when in the treatment and the contribution's from compliers, as we never get to observe them separately.

These limitations aside, it is still useful to perform the conditional-on-positive regression. If we were to find, say, a strongly negative association between the treatment and the wage bill, it would rule out the possibility that always-taker matches were better when in the treatment group and complier-matches were better than always-taker matches when in the control.

In Column (1), Table 8, we regress the log wage bill on the treatment indicator and its interaction with the technical indicator. The coefficients on the treatment indicator indicate a

5% higher wage bill and the treatment/technical interaction term adds another 10% but imprecisely estimated—the 95% confidence interval easily includes zero for both parameters. As expected, technical vacancies have a far larger wage bill than non-technical vacancies. From this regression, we can conclude only that we do not have strong evidence on the effect of the treatment on match quality. However, if our prior was that induced, "marginal" matches would be of lower value, this is some evidence against that hypotheses.

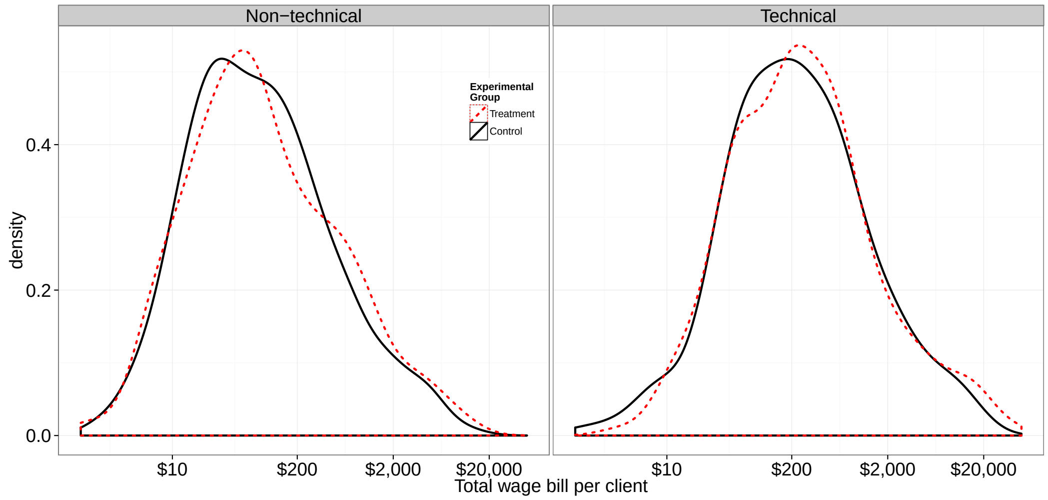
Table 8: Wage bill, hours worked and wage rate by experimental group

	All Vacancies	Hourly Vacancies Only		
	(1) *DV*= Log wage bill	(2) *DV*= Log wage bill	(3) *DV*= Log hours	(4) *DV*= Log wage
Intercept	4.404*** (0.074)	4.840*** (0.102)	3.659*** (0.117)	1.659*** (0.054)
Assigned to treatment	0.053 (0.109)	0.209 (0.149)	0.250 (0.169)	−0.008 (0.075)
Technical vacancy	1.014*** (0.118)	1.108*** (0.152)	0.255 (0.158)	0.740*** (0.076)
Treatment × Technical vacancy	0.101 (0.167)	−0.097 (0.223)	−0.346 (0.230)	0.155 (0.102)
N	1,757	1,017	1,017	1,017
R-squared	0.086	0.084	0.003	0.198

Notes: This table reports several regressions where the outcome variables are measures of match surplus—namely the wage bill, hours worked and wage of hired applicant. The sample for these regressions is vacancies from the recruting experiment that *filled* their vacancies. As such, these regressions cannot be interpreted causally, as they are a selected sample—we know from Table 6 that the treatment affects the quantity of matches formed. Each regression was estimated using OLS, and robust standard errors are reported. In Column (1) the dependent variable is the log of the total wage bill for the associated vacancy for all vacancies that fill. In Column (2) the dependent variable is also the log wage bill, but the sample is restricted to hourly vacancies. Columns (3) and (4) are the log hours worked and the log wage rates respectively, restricted to hourly vacancies (this data is only available for hourly vacancies). Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

Figure 6 shows the kernel density estimates of the log wage bill for vacancies by type and experimental group. We can see that the modal peaks are right-shifted for technical vacancies, with the modal vacancy being about \$200 compared to less than \$100 among non-technical vacancies. Among non-technical vacancies, the treatment and control groups appear to have very

Figure 6: Wage bill distributions by treatment and segment



Notes: This figure shows the kernel density estimates of the distribution of employer wage bill (conditional upon being positive), by segment and experimental group. The bandwidth for the kernel density estimate is calculated using Silverman's rule of thumb (Silverman, 1986) and the gaussian kernel is used. The left panel shows results for non-technical vacancies, while the right panel shows the results for technical vacancies (e.g., computer programming). The x-axis is on a log scale.

similar distributions. For technical vacancies, despite the evidence of no statistically significant effects, there does appear to be some movement at the tails: there are fewer very low wage bill matches and more very high wage bill matches.

For hourly jobs, we observe more than simply the wage bill—we can also observe the hours worked and the wage. Within an hourly match, the wage bill is partially mechanically related to the wage rate and hours worked, but not entirely so, as employers can and do pay bonuses (and workers can make refunds). If we assume multiplicative errors, i.e., $y = wh\epsilon_w\epsilon_h$, then $\log y = \log w + \log h + \eta$, where $\eta = \log \epsilon_h + \log \epsilon_w$. Let us assume that the combined treatment effect/selection effect on the wage bill can be decomposed into an effect on hours, τ_h, and an effect on wages, τ_w. The treatment effect on the wage bill, τ_y, will satisfy $(1+\tau_y) \approx (1+\tau_w)(1+\tau_h)$.

In Table 8, in Columns (2), (3) and (4), we obtain estimates for τ_y, τ_h and τ_w, respectively, using only hourly vacancies. Neither the treatment indicator nor the treatment/technical interaction are conventionally significant, either for technical or non-technical vacancies. The sign of coefficients for technical vacancies—negative for hours worked, positive for wages—is some weak evidence for the notion that recommendations might have led to the swap of more productive (and hence higher wage) individuals.

5.3.1 Subjective attributes of formed matches

The oDesk context affords other more qualitative measures of the quality of a match. In particular, when an employer or worker dissolves a relationship, they *may* give a reason for ending it, selected from a drop-down list. One reason is the option "job was completed successfully," which is a attractive proxy for match quality. Another qualitative measure is the feedback employers and workers give to each other at the conclusion of a contract. Table 9 contains several regressions where the outcome variables are these different measures of match quality.

In Table 9, Column (1), we regress an indicator for "success" on the treatment indicator and its interaction with the indicator for a technical vacancy. We can see that slightly less than 70% of vacancies are rated as having been completely successfully. The coefficient on the treatment is positive but close to zero among both technical and non-technical vacancies (with a slightly lower point estimate in the technical case). In Column (2) the outcome variable is the worker's feedback (1 to 5 stars) on the employer, and in Column (3) the outcome variable the employer's feedback on the worker. This feedback is not required and so is not universally available for matches. In both cases, neither the treatment group nor the vacancy type had a detectable effect on the ratings.

Collectively, the results suggest no substantive differences in the matches formed across the treatment and control groups: the differences in the success measure and the feedback measure are small and statistically insignificant and there is no consistent patterns across the measures.

Table 9: Qualitative measures of match attributes by experimental group

	(1) *DV*= Employer rated relationship a "Success"	(2) *DV*= Feedback by worker on employer (1-5 stars)	(3) *DV*= Feedback by employer on worker (1-5 stars)
Intercept	0.698*** (0.020)	4.664*** (0.042)	4.342*** (0.047)
Assigned to treatment	0.027 (0.028)	−0.025 (0.062)	0.075 (0.069)
Technical vacancy	−0.015 (0.032)	0.015 (0.068)	0.088 (0.081)
Treatment × Technical vacancy	−0.034 (0.045)	0.072 (0.095)	−0.105 (0.113)
N	1,768	1,513	1,421
R-squared	0.002	0.001	0.001

Notes: This table reports several regressions where the outcome variables are measures of match quality—namely the employer's report of whether the job was completed successfully and the bilateral five-start feedback the two sides gave each other at the conclusion of the job. The sample for these regressions is vacancies from the recruting experiment that *filled* their vacancies. As such, these regressions cannot be interpreted causally, as they are a selected sample—we know from Table 6 that the treatment affects the quantity of matches formed. Each regression was estimated using OLS, and robust standard errors are reported. In Column (1) the dependent variable is an indicator for whether the employer, when ending the contract, said the reason was that "the job was completed sucessfully." In Columns (1) and (2) the dependent variables are the post-contract feedback given by each side (out of one to five stars): Column (2) is the worker's feedback of the employer, while Column (3) is the worker's feedback of the employer. Significance indicators: $p \le 0.05 : *$, $.p \le 0.01 : **$ and $p \le .001 : ***$.

6 Discussion

In this section, I discuss the main results and consider possible alternative explanations. Though not every alternative can be ruled out, the preponderance of the evidence suggests the following: (1) recommendations were effective because they increased the pool of applicants; (2) recommendations were more effective among technical vacancies because technical vacancies receive fewer applicants and the marginal applicants have a higher return; and (3) lack of crowd-out was due to the presence of recruits, increasing the probability that an employer would evaluate anyone at all, which benefited organic applicants.

6.1 What is the causal mechanism for the effectiveness of recommendations?

The recommendation treatment was effective at increasing fill rates, but the experiment cannot tell us precisely why it was effective. The model provides one mechanism: the treatment increased the pool of would-be applicants, and treated employers were exposed to workers that they would not have encountered if they had counter-factually been assigned to the control group. There are other potential mechanisms to consider which are discussed below, yet none explain the pattern of results as well as the simple story presented above.

One alternative possibility is that the recommendations simply alerted employers to the possibility of recruiting. The sample did consist of new employers relatively unfamiliar with oDesk who may not have appreciated this possibility. But recall from Table 4, Column (3), that there was no increase in "late" recruiting. This is contrary to what we would expect if the experiment mainly worked by increasing knowledge of recruiting among employers. Another possibility is that the recommendations changed employer's beliefs about worker's latent characteristics; the treatment could have made them believe the pool of "realized" applicants was better and therefore encouraged hiring. This is unlikely for several reasons. First, if the treatment caused employers to believe that their pool was more productive than the control, then they would, on average, be relatively disappointed by realized performance, and yet we find no evidence of this; treatment and control matches look very similar with respect to all of our measures of match quality. Second, the treatment was introduced via a pop-up window that users that saw for seconds or minutes. And as we saw from Figure 2, most employers made decisions a week later once their pool was complete, making it doubtful that many employers even remembered the treatment.

6.2 Why were recommendations effective only among technical vacancies?

The treatment increased fill rates only among technical vacancies. While the cause may have been that that technical algorithmic recommendations were simply of higher quality, several pieces of evidence suggest a more economic explanation.

In the model, additional applicants offer decreasing marginal returns. When a vacancy already receives many applications, an additional applicant is more likely to be infra-marginal. As such, all else being equal, a vacancy receiving fewer organic applicants should achieve a relatively larger increase in fill rate compared to a vacancy with few applicants. On oDesk, technical vacancies receive significantly fewer applicants than non-technical vacancies—about 8 fewer than the baseline level of 22 applicants in non-technical vacancies. Perhaps the strongest evidence for the greater returns to additional applicants comes from our analysis in Table 7: recall that employers with technical vacancies viewed and interviewed about 10% more of their organic applicants than employers with non-technical vacancies did.

At the micro-level of individual observations, the number of organic applications a vacancy receives is strongly negatively correlated with whether or not the employer recruits: employers that recruited went on to receive substantially fewer applicants (recall Column 4 from Table 4). This is at least suggestive that employers recruit when they know they will receive fewer organic applicants. This association could be mechanical if recruiting causes jobs to fill quickly—employers that recruit will seem to have received fewer applicants. However, if we restrict the sample to employers that recruited but had no workers actually accept their invitations, the pattern still exists.

One alternative explanation for the greater effectiveness of technical recommendations is that they could simply be of better quality. But this seems unlikely. If we thought technical recommendations were simply better, then we might find that invitations sent by employers on the basis of those recommendations would have a higher positive response rate. As shown in Table 5, the response rate to technical invitations in the treatment and control groups are nearly identical and the response rate to non-technical invitations is the same in treatment and the control. This lack of a difference in differences is some evidence of treatment-induced invitations being comparable to non-induced invitations, regardless of work category.

6.3 Micro-Evidence of an extensive screening margin

The treatment increased the hiring of recruited applicants but did not decrease the hiring of non-recruited, organic applicants. The absence of crowd-out is surprising, but finding complementarity when the expectation is substitution is not unheard of in other electronic commerce settings (which generally have fine-grained enough data to detect these relationships). Ghose and Yang (2010) found that paid and organic search engine results—similarly "obvious" substitutes—were in fact complements, each stimulating more clicks on the other "side."[5]

The model provides an explanation for complementarity even when the employer is only looking to make one hire: the presence of recruits—which are generally better on observables than organic applicants—makes screening more attractive and hence more likely. This added screening benefits both recruited and organic applicants: an employer is less likely to call off their search when recruits are present, and this effect offsets crowd-out.

This extensive screening margin explanation has testable empirical implications. The arrival of a recruited applicant pushes the value of a vacancy over the extensive margin screening costs, which triggers a screening session. If this screening session occurs shortly after the recruited applicant arrives then it should create a pattern in the data—namely that organic applicants arriving *before* a recruited candidate arrives should be at an advantage compared to candidates arriving immediately after. To test this implication, we can compare applicants on outcomes such as hiring, being interviewed by the employer and being viewed by the employer (i.e., did the employer "open" their application).

[5]In another example, Oestreicher-Singer and Sundararajan (2012) found that co-purchase links to other products greatly amplifies demand for both linked products, even though the linked products are sometimes substitutes (though complements is probably more common).

We cannot simply compare mean outcomes by relative position for several reasons. First, for purely mechanical reasons, later applicants are less likely to be hired, interviewed and screened. This might be particularly important for applicants arriving after the recruited candidate, who is far more likely to be hired than his or her organic counter-parts. However, we address these issues in two ways: by restricting attention to vacancies where the recruited applicant was *not* hired and by including specific effects for each arrival rank position.

Using vacancies that sent only one recruiting invitation and received a positive response from the invited applicant, we estimate a series of regressions of the form

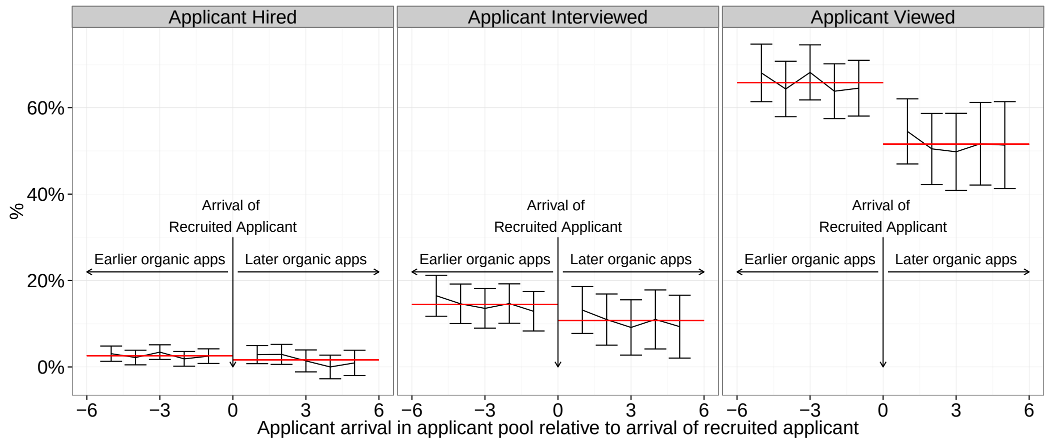
$$Y_{ij} = \beta_0 + \sum_{i=-10}^{10} \beta_i \text{RELPOS}_{ij} + \sum_i \alpha_i \text{ACTUALPOS}_{ij} + \epsilon_{ij} \tag{4}$$

where Y_{ij} is the outcome for the ith applicant to the jth vacancy. We restrict our attention to only vacancies where the recruited applicant arrived at the 5th position or later, since only those vacancies will have substantial numbers of applicants already in the pool when the recruited applicant arrives.

The indicator RELPOS is the applicant's position relative to the arrival of the recruited applicant, while ACTUALPOS is the actual position. For example, if the recruited applicant was the 4th to arrive, then the first applicant would have a relative position of −3 and an actual position of 1. Although this regression is descriptive, the source of variation in our independent variables should suggest the effects are causal. The precise arrival of the recruited candidate presumably depends largely on idiosyncratic factors like time zone differences, the email checking habits of the worker and so on. If the precise arrival rank of the recruited applicant is largely driven by idiosyncratic factors, then it is as good as randomly assigned.

In Figure 7, I plot the coefficients of RELPOS for outcomes of hiring, interviewing and viewing. For all three outcomes, applicants arriving immediately before the recruited applicant fare better: they are slightly more likely to be hired and are substantially more likely to be inter-

Figure 7: Estimated effects of arriving in the applicant pool relative to the arrival of the recruited applicant



Notes: This figure plots the positional coefficients for several applicant outcomes relative to the arrival in the applicant pool of the recruited worker. The x-axis is an applicant's arrival position relative to that of the recruited applicant. The figure illustrates how non-recruited applicants who happened to have applied immediately before a recruited applicant arrived benefitted in terms of increased viewing, interviewing and hiring. The is the claimed mechanism for the lack of crowd-out in the experiment, which otherwise should simply displace non-recruited workers. The sample consists of all applications made to vacancies in the experiment where the employer sent only one recruiting invitation which was accepted but in which that worker was not ultimately hired. Each regression is based on Equation 4 and is a multi-level model with relative position fixed-effects and vacancy-specific random effects.

viewed and viewed by the employer.

7 Conclusion

In this paper, I demonstrate that recommendations are both acted upon by employers and also effective at raising vacancy fill rates. Surprisingly, recommendations do not crowd-out non-recommended applicants. These results show that at least in one labor market context, reducing frictions does not entail an equity versus efficiency trade-off: both were improved. These results come from a context where search frictions are already very low, suggesting that bigger gains are possible in traditional, more friction-prone markets. As such, the results suggest that policy makers should take greater interest in demand-side active labor market policies that could replicate what was accomplished on oDesk.

This study raises a number of further questions. The study highlights—but does not address—the unique challenges posed by recommending supply-constrained sellers. A concern with any automated recommendation system is over-recommending workers—unlike digital goods or most physical goods sold online, workers can quickly "stock out." This problem can be partly ameliorated by incorporating real-time visibility signals in the recommendation algorithms, but this will cause less-attractive applicants—all else equal—to be recommended. A study examining how severe this problem is—and examining the optimal allocation of "visibility" across the marketplace—would be quite useful in marketplaces with supply-constrained sellers.

This study is in part about market design. The existing market design literature focuses on a center that fixes market congestion and market thinness by setting matches directly (Roth, 2008; Niederle et al., 2008). Less attention has been paid to the market design decisions made by platforms with powerful centers that nevertheless do not have—nor necessarily want—match-setting power.[6] However, with growing computer-mediation, many more markets will have the strong-but-not-omnipotent center and each will have many consequential market design decisions to make. This paper shows that even without match-setting power, platforms can tilt the market toward desirable ends through purely informational interventions.

[6] Coles et al. (2010) is an exception.

References

Adomavicius, Gediminas and Alexander Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *Knowledge and Data Engineering, IEEE Transactions on*, 2005, *17* (6), 734–749.

Agrawal, Ajay K., John Horton, Nico Lacetera, and Elizabeth Lyons, "Digitization and the Contract Labor Market: A Research Agenda ," in "Economics of Digitization," University of Chicago Press, 2013.

__ , Nicola Lacetera, and Elizabeth Lyons, "Does information help or hinder job applicants from less developed countries in online markets?," January 2013, (NBER Working Paper 18720).

Autor, David H., "Wiring the labor market," *The Journal of Economic Perspectives*, 2001, *15* (1), 25–40.

__ , "The economics of labor market intermediation: An analytic framework," September 2008, (14348).

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.

__ , Dan A. Black, and Mark A. Loewenstein, "Job matching and on-the-job training," *Journal of Labor Economics*, January 1989, *7* (1), 1–19.

Burdett, Kenneth and Elizabeth J. Cunningham, "Toward a theory of vacancies," *Journal of Labor Economics*, July 1998, *16* (3), 445–78.

Card, David, Jochen Kluve, and Andrea Weber, "Active labour market policy evaluations: A meta-analysis," *The Economic Journal*, 2010, *120* (548), F452–F477.

Coles, Peter, John Cawley, Phillip B. Levine, Muriel Niederle, Alvin E. Roth, and John J. Siegfried, "The job market for new economists: A market design perspective," *The Journal of Economic Perspectives*, 2010, *24* (4), 187–206.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora, "Do labor market policies have displacement effects? Evidence from a clustered randomized experiment," *The Quarterly Journal of Economics*, 2013, *128* (2), 531–580.

Gautier, Pieter, Paul Muller, Bas van der Klaauw, Michael Rosholm, and Michael Svarer, "Estimating equilibrium effects of job search assistance," IZA Discussion Papers 6748, Institute for the Study of Labor (IZA) July 2012.

Ghose, Anindya and Sha Yang, "Analyzing the relationship between organic and sponsored search advertising: Positive, negative, or zero interdependence?," *Marketing Science*, 2010, *29* (4), 602–623.

Imbens, Guido W. and Joshua D. Angrist, "Identification and estimation of local average treatment effects," *Econometrica*, 1994, *62* (2), 467–475.

Kluve, Jochen, "The effectiveness of european active labor market programs," *Labour Economics*, 2010, *17* (6), 904–918.

MacKinnon, James G. and Halbert White, "Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties," *Journal of Econometrics*, 1985, *29* (3), 305–325.

Niederle, Muriel, Alvin E. Roth, and Tayfun Sönmez, "Matching and market design," *The New Palgrave Dictionary of Economics. Palgrave Macmillan, Basingstoke*, 2008.

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

Oyer, Paul and Scott Schaefer, "Personnel economics: Hiring and incentives," May 2010, (15977).

Pallais, Amanda, "Inefficient hiring in entry-level labor markets," *NBER Working Paper 18917*, March 2013, (18917).

Pellizzari, Michele, "Employers' search and the efficiency of matching," *British Journal of Industrial Relations*, 2011, *49* (1), 25–53.

Resnick, Paul and Hal R. Varian, "Recommender systems," *Communications of the ACM*, 1997, *40* (3), 56–58.

Roth, Alvin E., "What have we learned from market design?," *The Economic Journal*, 2008, *118* (527), 285–310.

Silverman, Bernard W., *Density estimation for statistics and data analysis*, Vol. 26, CRC press, 1986.

Stanton, Christopher and Catherine Thomas, "Landing the first job: The value of intermediaries in online hiring," *Available at SSRN 1862109*, 2012.

van Ommeren, Jos and Giovanni Russo, "Firm recruitment behaviour: Sequential or non-sequential search?," *SSRN eLibrary*, 2010.

van Ours, Jan and Geert Ridder, "Vacancies and the recruitment of new employees," *Journal of Labor Economics*, 1992, *10* (2), 138–155.

Varian, Hal R., "Computer mediated transactions," *American Economic Review*, 2010, *100* (2), 1–10.

A Randomization

Vacancies were randomly assigned to the treatment or control group after being posted to the marketplace. As such, pre-assignment covariates describing vacancy should be balanced across the groups; the count of observations across groups should also be consistent with a random process. In Table 10 we confirm this is the case. In the top panel, we report the count of observations and the p-value for a χ^2 test. We cannot reject the null hypothesis of an identical assignment probability across the groups. In the lower panel, we report the fraction of vacancies of different "types" (by category of work) by experimental group. We also report two non-category covariates—an indicator for whether the length of job description exceeded the median and whether the employer required that applicants have prior on-platform work experience. For each covariate, we report the mean and standard error for each group; we also report the difference in means, the standard error for that difference and the p-value for a two-sided t-test. We find good balance across the collection of covariates chosen, with none of the t-test p-values being conventionally significant.

Table 10: Pre-treatment covariate means by experimental group

	Treatment mean: $\bar{X}_{TRT}$	Control mean: $\bar{X}_{CTL}$	Difference in means: $\bar{X}_{TRT} - \bar{X}_{CTL}$	p-value
Observation Counts				
	2,900	3,053		0.109
Type of work				
Technical (1 if yes, 0 otherwise)	0.512 (0.009)	0.523 (0.009)	-0.011 (0.013)	0.409
Non-Technical	0.488 (0.009)	0.477 (0.009)	0.011 (0.013)	0.409
Type of work—(more detailed)				
Admin	0.078 (0.005)	0.080 (0.005)	-0.002 (0.007)	0.776
Writing	0.126 (0.006)	0.120 (0.006)	0.006 (0.009)	0.508
Web	0.388 (0.009)	0.399 (0.009)	-0.011 (0.013)	0.384
Design	0.109 (0.006)	0.111 (0.006)	-0.002 (0.008)	0.796
Software	0.124 (0.006)	0.124 (0.006)	0.000 (0.009)	0.970
Vacancy attributes				
Job description length > median	0.563 (0.009)	0.574 (0.009)	-0.011 (0.013)	0.402
Required prior oDesk experience	0.125 (0.006)	0.113 (0.006)	0.012 (0.008)	0.159

Notes: This table reports data from a recommendation experiment conducted on the oDesk platform. In the experiment, the treated group of employers received algorithmically-generated recommendations of candidates for their vacancies while the control group did not. The unit of randomization was the posted vacancy. The top panel, labeled "Observation Counts," is the number of vacancies per experimental group. The "Observation Counts" panel p-value is for a two-sided test of equal group assignment probabilities; the the p-value was calculated by simulation, using the fact that 13,259 potential employers were allocated, but only 5,953 generated usable vacancies for assignment. In the lower panels, for each variable, we report the mean and standard error for various pre-treatment covariates, as well standard error for the cross-group differences. The p-value column is for a two-sided t-test against the null hypothesis of no difference in means across the treatment and control groups. In the middle panels ("Type of Work" and "Type of work—(mode detailed)"), the covariantes are indicators for whether the vacancy was for a particular type of work. In the bottom panel, the two covariates are (1) an indicator for whether the employer required that applicants have prior oDesk experience and (2) whether the count of text characters in the job description was greater than the median count for the pooled sample of all job descriptions in the experiment. Significance indicators: $p \leq 0.05 : *$, .$p \leq 0.01 : **$ and $p \leq .001 : ***$.

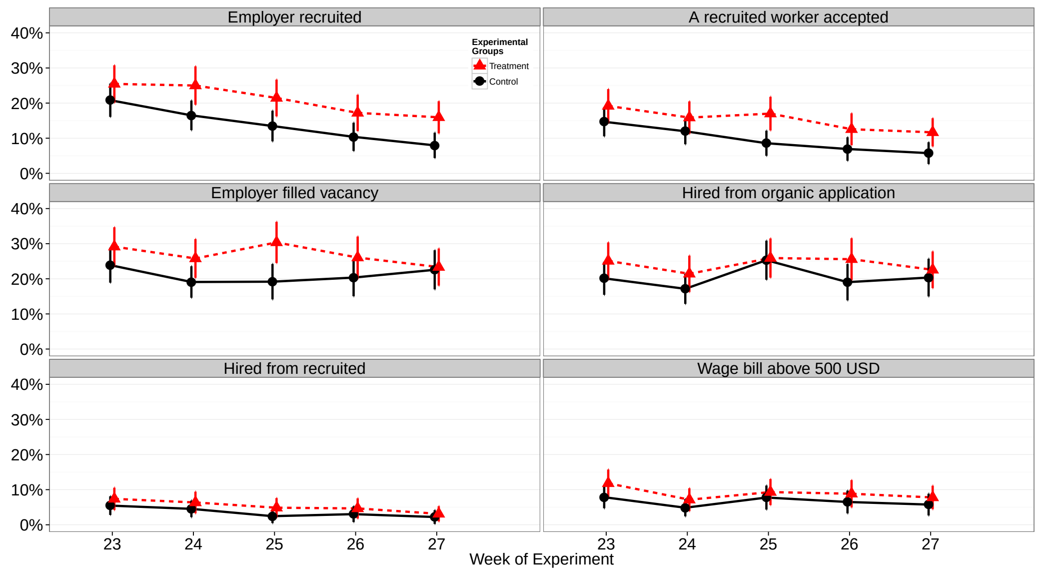
B For Online Publication: SUTVA and Stock-out

One concern with a field experiment is that the treatment cells interact with each other in some way, violating the stable unit treatment value (SUTVA) assumption required for causal inference. Any sufficiently large experiment conducted in a marketplace risks SUTVA problems. In our experimental setting, "deep" SUTVA concerns (say, moving the market or changing the incentive to entry or exit) are probably unwarranted. First, this experiment was conducted a year after Pallais (2013) (who did find evidence of market-moving effects), and the market was at least twice as large. Further, unlike her experiment, vacancies associated with all types of work were eligible. And because only new employers were eligible for the experiment and only some of these were allocated, only about 6% of the vacancies posted during the experimental period were actually assigned to the experiment (and only 1/2 of those were in the active treatment).

A more serious concern is that the experiment depleted the pool of "recruitable" workers, either absolutely (if control workers are recruiting from the same pool as the algorithmic recommendations experiment) or simply for treated employers. To test whether depletion is happening, in Figure 8 we plot several of our key experimental outcome measures but calculated at the level of the week of the experiment. In the upper left panel labeled "Employer recruited" we plot the fraction of employers recruiting by each experimental group by week of allocation. Although the treatment group always has a substantially higher fraction of employers recruiting, there is a noticeable decline in both groups over time, with seemingly the same slope in each group. In the panel to the right labeled "A recruited worker accepted" we can also see a decline over time in both groups.

To confirm this graphical analysis, we return to the recruiting response rate regressions found in Table 5. In Column (3), we regress an indicator for whether the vacancy received any recruited applicants, but we also allow for the possibility of a linear time trend in the vacancy week. The Column (4) regression is the same, but restricted to only those employers that re-

Figure 8: Effects of treatment by week of employer allocation



Notes: This figure reports several key experimental outcomes, by treatment and control groups, by week. For each each week and outcome, the point estimate is plotted, as well as a 95% CI for that estimate. The week on the x-axis refers to the Julian week of 2011—not the experimental week. The "Employer recruited" measure is whether the employer sent any recruiting invitations within 1 hour of posting their vacancy. The "Employer filled vacancy" measure is whether the employer filled their vacancy, as measured by spending at least some money against that vacancy. The "A recruited worker accepted" is whether the employer has any applicants that had been recruited. The "Hired from recruited" and "Hired from organic application" are whether the employer hired a recruited or an organic applicant, respectively. The "Wage bill above 500 USD" is whether the employer spent at least $500 against the vacancy.

cruited. As expected from the plots, in Column (3) the coefficient on the allocation week is negative and highly significant. The economic magnitude is substantial—over the five weeks of the experiment, the fraction of employer receiving at least one recruited applicant declined nearly 8% in the control group, among non-technical vacancies. This fall-off in the treated group is even is stronger, though not significantly different from the trend found in the control group. From Column (4), we can see that conditional upon recruiting, there are no obvious time trends, suggesting that most of the reduction in received recruits is from employers recruiting less and not from a reduction in worker acceptance rates. Given the short window of the experiment, declining employer uptake of recommendations is some evidence for the over-recommendation hypothesis.

C For Online Publication: Codebook

This section contains the name of each variable used in the paper, as well as its variable name in the data file accompanying this paper.

Wage bill above 500 USD (`above_med`). Indicator for whether the total wage bill associated with the vacancy was above $500.

Accepted early invites (`accepted_early_invites`). Number of accepted early invites.

Accepted invites (`accepted_invites`). Indicator for whether any recruited candidates accepted an invitation.

A recruited worker accepted (`any_accepted_early_invites`). Indicator for whether any early recruited candidates accepted an invitation.

Duration label (`duration_label`). Choice made by their employer about how long the job would last.

Early invites (`early_invites`). Count of early invitations sent.

Employer identifier (`employer`). Employer's unique oDesk identifier plus a random string.

Group (`group`). This is a categorical variable for the experimental group.

Employer filled vacancy (`has_charges`). Indicator for whether or not the employer spent any money against this vacancy.

Employer recruited (`has_early_invites`). This is an indicator for whether the employer recruited in the first hour after posting.

Has invites (`has_invites`). This is an indicator for whether the employer recruited at all.

Hired from recruited (`hired_from_early_invite`). This is an indicator for whether the employer hired any of the early invited candidates.

Hired from invites (`hired_from_invite`). This is an indicator for whether the employer hired any of the invited candidates.

Organic applications (`applications`). Count of organic applications received by the vacancy.

Invitations sent (`invites`). Count of invitations sent by the employer.

Number hired (`num_hires`). Count of number of workers hired against this vacancy.

Mean wage of recruits (`mean_recruited_wage`). Mean wage of the recruited candidates to hourly jobs.

Mean wage of applicants (`mean_applied_wage`). Mean wage of organic applicants to hourly jobs.

Invite response rate (`invite_response_rate`). Positive response rate of applicants.

Hires from organic applicants (`num_hires_from_org_app`). Count of hires from organic applicants.

Hired from organic application (`hired_from_org_app`). Hired from an organic application.

Hires from invited applicants (`num_hires_from_invites`). Hires from invited applicants.

Number of hires from early invites (`num_hires_from_early_invites`). Hires from early invited applicants.

Number hired (`num_hires`). Count of number of workers hired against this vacancy.

Original opening date (`original_opening_date`). Date the original opening was posted on the site.

Segment (`segment`). Indicator for whether the vacancy was technical or non-technical.

Employer's signup date (`signup_date`). Date the employer signed up for oDesk.

Job description length (`job_desc_length`). Length of the job description (in characters).

Number viewed (`num_viewed`). The count of applicants that were actually viewed by the employer.

Table 11: Summary statistics for all continuous variables in the dataset

Variable	Levels	n	Min	q_1	$\widetilde{x}$	$\bar{x}$	q_3	Max	s	#NA
Accepted early invites	cntrl	3053	0	0.0	0.0	0.2	0	32	1.0	0
	treat	2900	0	0.0	0.0	0.3	0	15	1.0	0
	all	5953	0	0.0	0.0	0.3	0	32	1.0	0
Accepted invites	cntrl	3053	0	0.0	0.0	0.4	0	32	1.6	0
	treat	2900	0	0.0	0.0	0.6	0	124	3.1	0
	all	5953	0	0.0	0.0	0.5	0	124	2.4	0
Early invites	cntrl	3053	0	0.0	0.0	0.5	0	45	2.2	0
	treat	2900	0	0.0	0.0	0.8	0	39	2.6	0
	all	5953	0	0.0	0.0	0.6	0	45	2.4	0
Invitations sent	cntrl	3053	0	0.0	0.0	1.2	0	117	4.8	0
	treat	2900	0	0.0	0.0	1.6	1	564	12.5	0
	all	5953	0	0.0	0.0	1.4	1	564	9.4	0
Number hired	cntrl	3053	0	0.0	0.0	0.5	1	19	1.0	0
	treat	2900	0	0.0	0.0	0.5	1	17	0.9	0
	all	5953	0	0.0	0.0	0.5	1	19	0.9	0
Invite response rate	cntrl	749	0	0.0	0.5	0.5	1	1	0.4	2304
	treat	859	0	0.0	0.5	0.5	1	1	0.4	2041
	all	1608	0	0.0	0.5	0.5	1	1	0.4	4345
Hires from organic applicants	cntrl	3053	0	0.0	0.0	0.4	1	16	0.8	0
	treat	2900	0	0.0	0.0	0.4	1	17	0.8	0
	all	5953	0	0.0	0.0	0.4	1	17	0.8	0
Hires from invited applicants	cntrl	3053	0	0.0	0.0	0.1	0	7	0.4	0
	treat	2900	0	0.0	0.0	0.1	0	13	0.4	0
	all	5953	0	0.0	0.0	0.1	0	13	0.4	0
Number of hires from early invites	cntrl	3053	0	0.0	0.0	0.0	0	7	0.3	0
	treat	2900	0	0.0	0.0	0.1	0	3	0.2	0
	all	5953	0	0.0	0.0	0.1	0	7	0.3	0
Number hired	cntrl	3053	0	0.0	0.0	0.5	1	19	1.0	0
	treat	2900	0	0.0	0.0	0.5	1	17	0.9	0
	all	5953	0	0.0	0.0	0.5	1	19	0.9	0
Job description length	cntrl	3053	3	236.0	460.0	683.8	877	5000	697.5	0
	treat	2900	2	235.0	464.0	685.8	891	4919	702.3	0
	all	5953	2	236.0	461.0	684.8	884	5000	699.8	0
Number viewed	cntrl	3053	0	2.0	6.0	8.5	11	82	9.8	0
	treat	2900	0	2.0	5.0	8.7	11	306	12.4	0
	all	5953	0	2.0	6.0	8.6	11	306	11.1	0

D For Online Publication: Descriptive Statistics

In Table 11 I report summary statistics for continuous variables used in the analysis; Table 12 is the corresponding table for the discrete variables.

Table 12: Summary statistics for all discrete variables in the dataset

Variable	**Levels**	$\mathbf{n}_{\text{cntrl}}$	$\%_{\text{cntrl}}$	$\sum\%_{\text{cntrl}}$	$\mathbf{n}_{\text{treat}}$	$\%_{\text{treat}}$	$\sum\%_{\text{treat}}$	$\mathbf{n}_{\text{all}}$	$\%_{\text{all}}$	$\sum\%_{\text{all}}$
Wage bill above 500 USD	FALSE	2869	94.0	94.0	2683	92.5	92.5	5552	93.3	93.3
	TRUE	184	6.0	100.0	217	7.5	100.0	401	6.7	100.0
	all	3053	100.0		2900	100.0		5953	100.0	
A recruited worker accepted	FALSE	2735	89.6	89.6	2453	84.6	84.6	5188	87.2	87.2
	TRUE	318	10.4	100.0	447	15.4	100.0	765	12.8	100.0
	all	3053	100.0		2900	100.0		5953	100.0	
Group	cntrl	3053	100.0	100.0	0	0.0	0.0	3053	51.3	51.3
	treat	0	0.0	100.0	2900	100.0	100.0	2900	48.7	100.0
	all	3053	100.0		2900	100.0		5953	100.0	
Employer filled vacancy	FALSE	2176	71.3	71.3	2009	69.3	69.3	4185	70.3	70.3
	TRUE	877	28.7	100.0	891	30.7	100.0	1768	29.7	100.0
	all	3053	100.0		2900	100.0		5953	100.0	
Employer recruited	FALSE	2588	84.8	84.8	2279	78.6	78.6	4867	81.8	81.8
	TRUE	465	15.2	100.0	621	21.4	100.0	1086	18.2	100.0
	all	3053	100.0		2900	100.0		5953	100.0	
Has invites	FALSE	2304	75.5	75.5	2041	70.4	70.4	4345	73.0	73.0
	TRUE	749	24.5	100.0	859	29.6	100.0	1608	27.0	100.0
	all	3053	100.0		2900	100.0		5953	100.0	
Hired from recruited	FALSE	2927	95.9	95.9	2753	94.9	94.9	5680	95.4	95.4
	TRUE	126	4.1	100.0	147	5.1	100.0	273	4.6	100.0
	all	3053	100.0		2900	100.0		5953	100.0	
Hired from invites	FALSE	2801	91.8	91.8	2644	91.2	91.2	5445	91.5	91.5
	TRUE	252	8.2	100.0	256	8.8	100.0	508	8.5	100.0
	all	3053	100.0		2900	100.0		5953	100.0	
Hired from organic application	FALSE	2208	72.3	72.3	2077	71.6	71.6	4285	72.0	72.0
	TRUE	845	27.7	100.0	823	28.4	100.0	1668	28.0	100.0
	all	3053	100.0		2900	100.0		5953	100.0	