Computer-mediated Matchmaking: Facilitating Employer Search and Screening*

John J. Horton
oDesk Research & Harvard University

October 25, 2012

Abstract

Participants in matching markets often face high search and screening costs. An informed third party may reduce these costs by recommending matches—an increasingly easy task in computer-mediated markets. This approach to reducing friction raises questions: (a) when are recommendations effective and (b) to what extent, if any, do recommendations crowd out non-recommended, "organic" matches? We answer these questions using an experiment conducted in an online labor market in which a treatment group of employers received algorithmically-generated job candidate recommendations. Recommendations improved fill rates by nearly 17% for computer programming vacancies, but had modest or insignificant effects in other types of work. This heterogeneity was likely due to relatively high screening costs of programming categories, which in turn made recommendations more valuable. However, where fill rates did increase, it was only partly the result of employers acting upon recommendations: the treatment also increased hiring of non-recruited, organic applicants. This complementarity was caused by treated employers screening more intensely and extensively, and their additional attention "spilling over" onto organic applicants. This spill-over result suggests that in some cases recommendations can improve marketplace efficiency without making anyone worse off, despite explicit promotions of certain workers over others.

JEL Codes: C93, J64

Keywords: employer search, labor market intermediation

*Preliminary—please do not cite or distribute.

1 Introduction

Firms and workers trying to form new matches face a variety of search and screening costs. These costs might be reduced via the efforts of third-party intermediaries, which could inform workers and firms about the existence and attributes of vacancies and job-seekers or even make explicit recommendations. There are several policy-relevant reasons to be interested in platforms providing algorithmic recommendations in labor markets. In traditional labor markets, placement assistance is relatively effective, compared to other active labor market policies. Algorithmic assistance is qualitatively similar, but has several advantages, including an attractive cost structure and better scalability (Resnick et al., 2000). Even if the platform merely provides credible information, they would be playing a role known to be important in traditional markets (Autor, 2008).

On the question of effectiveness, in a recent meta-analysis of 97 active labor policy programs, Card et al. (2010) found that job search assistance programs were more likely to yield positive short-term impacts compared to public sector employment. A similar analysis by Kluve (2010) of European active labor market policies also concluded that offered services such as placement assistance and counseling were effective at raising employment probability. Recent specific studies that offer credible estimates include at least two randomized experiments where market participants received job-placement assistance (though in both cases it was workers rather than firms receiving the assistance). Gorter and Kalb (1996) report the results of a experiment where treated unemployed workers received job-finding assistance. It failed to improve their matching probability, but it did increase their application intensity. Gurgand et al. (2011) reports the results of a recent randomized experiment in France that showed that placement assistance had (modest) benefits without having crowd-out effects.

Despite the qualitative similarity to traditional placement assistance, algorithmic assistance has the advantage of being highly scalable. Once the fixed costs of creating the system are borne, the marginal costs of additional recommendations are close to zero. Further, the quality of those recommendations may improve as more data becomes available.[1] Computer-mediated markets, by necessity, capture huge amounts of data that can be used to improve the recommendation algorithms. Because they act in nearly real-time, recommendations made by the platform can also ameliorate market congestion and market thinness by ensuring that neither vacancies nor workers are over- or under-recommended. This flow-management possibility is particularly exciting as congestion and thinness are thought to be important sources of match-

[1] An emerging perspective from the machine learning community is that generally more data is better than more sophisticated algorithms (Halevy et al., 2009).

ing friction (Petrongolo and Pissarides, 2001).

Given that both firms and workers are in a sense experience goods, with the quality of the match revealed over time, labor markets are particularly beset with informational problems. As Autor (2001) notes, "[w]orkers and jobs are naturally heterogeneous, and the quality of their interaction when paired is notoriously difficult to forecast." A consequence of these informational problems is that labor market intermediaries play an important role in labor markets (Autor, 2008). It seems likely that algorithmically-assisted labor market platforms can potentially provide many of the same services that intermediaries are called upon to provide today.

To be effective, an intermediary must understand the economic forces that shape how firms approach the vacancy-filling problem. Without this understanding, the intermediary would likely provide the wrong information, make too many or too few recommendations per vacancy and subsidizing the search efforts of the wrong parties. Unfortunately, our understanding is limited, as there has been little research to date on how firms fill their vacancies. Oyer and Schaefer (2010) summarize the state of the literature:

> The literature has 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.

The main culprit for the lack of research is not that the topic is unimportant, but rather that data are limited. In this paper, we use data—both experimental and observational—from a labor market where the entire match-making process is observed. We observe the posting of a vacancy, the search efforts of the firm, the arrival of applicants, the screening of applicants and the making of job offers. In the experiment, treated employers received algorthmically-generated, by-name recommendations of candidates for their vacancies. These candidates were chosen to relevant (defined as having the skills the employer requested), high-ability (as measured by past feedback, test scores etc.) and available (as measured by that candidate recently seeking work through the platform). We investigate whether these recommendations increased employer recruiting—defined as an employer inviting workers to apply for a vacancy—and whether the recommendations increased vacancy fill rates. We also examine whether recommended applicants "crowd out" organic applicants (organic applicants are those applicants that apply without being explicitly invited to do so).

To help interpret our empirical findings, we develop a model of employer search and screening. In the model, firms obtain applicants in two ways: applicants either apply organically, which costs the firm nothing, or applicants are actively recruited. Recruiting is costly, but it increases the quantity and average quality of applicants. Regardless of application type, firms pay a fixed, per-applicant cost to screen applicants. This screening tells the firm if an applicant is a "match" for their vacancy. Screening is done in "batch" and firms must choose the number of applicants to screen ex ante. This optimal sample size approach follows the original Stigler formulation of the search problem. If one or more screened applicants are a match, then the vacancy is filled. The firm's decision problem is to decide whether to recruit, and condition upon that choice, how many applicants to screen. The model focuses on predicting how changes in fundamental parameters like cost, project value, worker quality and labor market tightness affect employer's choices and the vacancy fill rate. We view the experiment as exogenously lowering the costs of recruiting, which the model predicts should lead to more recruiting, potentially more screening and higher fill rates.

This paper contributes to three literatures. First, it adds to the labor literature a new model of employer search and screening, as well as causal evidence to how firms fill vacancies. Second, it enriches our understanding of how platform-provided information is valued by market participants and how this information affects market outcomes. Examples of similar work include the value of reputations on eBay (Resnick et al., 2000), a work history on oDesk (Pallais, 2010) and reviews on Yelp (Luca, 2011). Ours is the first study that shows how market participants value and use recommendations in a labor market context. Lastly, it is a contribution to the market design literature. The existing market design literature focuses on a strong "center" that fixes market congestion and market thinness by setting matches directly (Roth, 2008; Niederle et al., 2008). However, platforms in decentralized markets like oDesk do not have—nor do they necessarily want—this match-setting power. And yet this paper shows that a market can be tilted towards similar ends through purely informational interventions. Coles et al. (2010) is a recent example of work in this decentralized market design literature. As these kinds of marketplaces grow, opportunities for this kind of market design will increase as well.

1.1 Overview of the paper and key results

Section 2 describes the oDesk marketplace and the details of the experiment. Section 3 reviews the employer search literature and compares it with our approach and empirical setting. Our model is developed in Section 4, while the empirical analysis is presented in three sections. These empirical sections use a number of different datasets, which are described in Table 1. The actual results are summarized in Table 2, with reference to the associated model assumptions and predictions.

Section 5 describes the composition of vacancies by category of the work, illustrating large cross-category variation in how employers approach the vacancy-filling decision. We also show that recruited applicants are far more likely to be screened and ultimately hired by employers. In fact, a large fraction of organic applications are given negligible attention by employers. This finding motivates our modeling assumption that firms choose to screen only some of the applications they receive.

In Section 6 we analyze the results of the experiment. We find that providing algorithmically-generated recommendations increases employer recruiting and that this recruiting raised the fill rate substantially in high-skill categories—namely in categories that required programming ability. This heterogeneity in hiring treatment effects requires an explanation, particularly since the treatment increased recruiting for nearly all categories of vacancies. A potential theoretical explanation comes from our model, which predicts that vacancies with higher associated screening costs should experience a larger jump in fill rates in the treatment group. We find some evidence that these programming categories do in fact have higher screening costs.

We surprisingly find that the treatment increased fill rates for both recruited *and* organic applicants. We explore several hypotheses for why, but the evidence suggests a kind of screening "spill over" in the treatment group. In the treatment group, the presence of recruited candidates caused the employers to more extensively screen applicants regardless of source, to the benefit of both kinds of applicants. In terms of the attributes of the matches formed, we find no detectable difference across experimental groups in the hired worker's wages or the feedback they received from employers.

We continue our analysis of the hypothesized spill-over in Section 7, using a non-experimental data from the larger marketplace. We estimate how a recruited candidate's response to an employer's invitation affects the probability that a vacancy is filled, and if it is filled, by what kind of applicant. Because a recruited candidate's response is endogenous, we use an instrumental variables approach. As expected, a positive response to employer's invitation increases the hiring of recruited applicants. Less expected but consistent with our earlier findings, this acceptance does *not* decrease the probability that an organic applicant is hired. This implies crowd out is a secondary concern which in turn suggests recommendations can be Pareto efficient and not just Kaldor-Hicks efficient *within* the oDesk marketplace. Of course, the oDesk marketplace is not *the* labor market, yet we present evidence that few oDesk vacancies are posted anywhere other than oDesk.

In Section 8, we conclude by discussing how the increasing computer-mediation of nearly all marketplaces will expand the role for third-party intermediation, which in turn will raise new research questions. In addition to the main paper, we also include several appendices to address secondary predictions or assumptions of the model and empirical findings that are of more limited interest. In Appendix A, we show that employers on oDesk process applications in batch. In Appendix C, we exploit the fact that employers post multiple vacancies on oDesk over time to estimate the effects of project value and labor market tightness on recruiting, screening and hiring. Finally, in Appendix D we discuss some practical market design "lessons learned" that would be relevant to other market design interventions.

2 The oDesk marketplace

During the last ten years, a number of online labor markets have emerged. In these markets, firms can 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 platform include maintaining job listings, hosting user profile pages, arbitrating disputes, certifying worker skills and maintaining feedback systems. Horton (2010) presents models on pricing in an online labor market, optimal provision of user services, market specialization and competitive equilibrium between platforms.

Our experiment takes place on the largest of these online labor markets, oDesk. On oDesk, would-be employers write job descriptions, self-categorize the nature of the work and the skills required and then post these vacancies to the website. Workers learn about vacancies via electronic search. When they find an application they would like to apply for, they submit an application which generally includes a wage bid (for hourly jobs) or a total project bid (for fixed-price jobs) and a cover letter. In addition to these worker-initiated applications, employers can also search through worker profiles and invite them to apply. Employers make these invitations on a worker-by-worker basis and it is not possible for them to automatically invite massive numbers of workers. After a worker applies, the employer can interview him or her and hire at the terms proposed by the worker, or make a counter-offer, which the worker can counter, and so on.

There has been some research which focuses on the oDesk marketplace. Pallais (2010) 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 (2011) use oDesk data to show that agencies (which act as quasi-firms) help workers find jobs. Agrawal et al. (2012) investigate what factors matter to employers in making selections from an applicant pool.

2.1 Description of the experiment

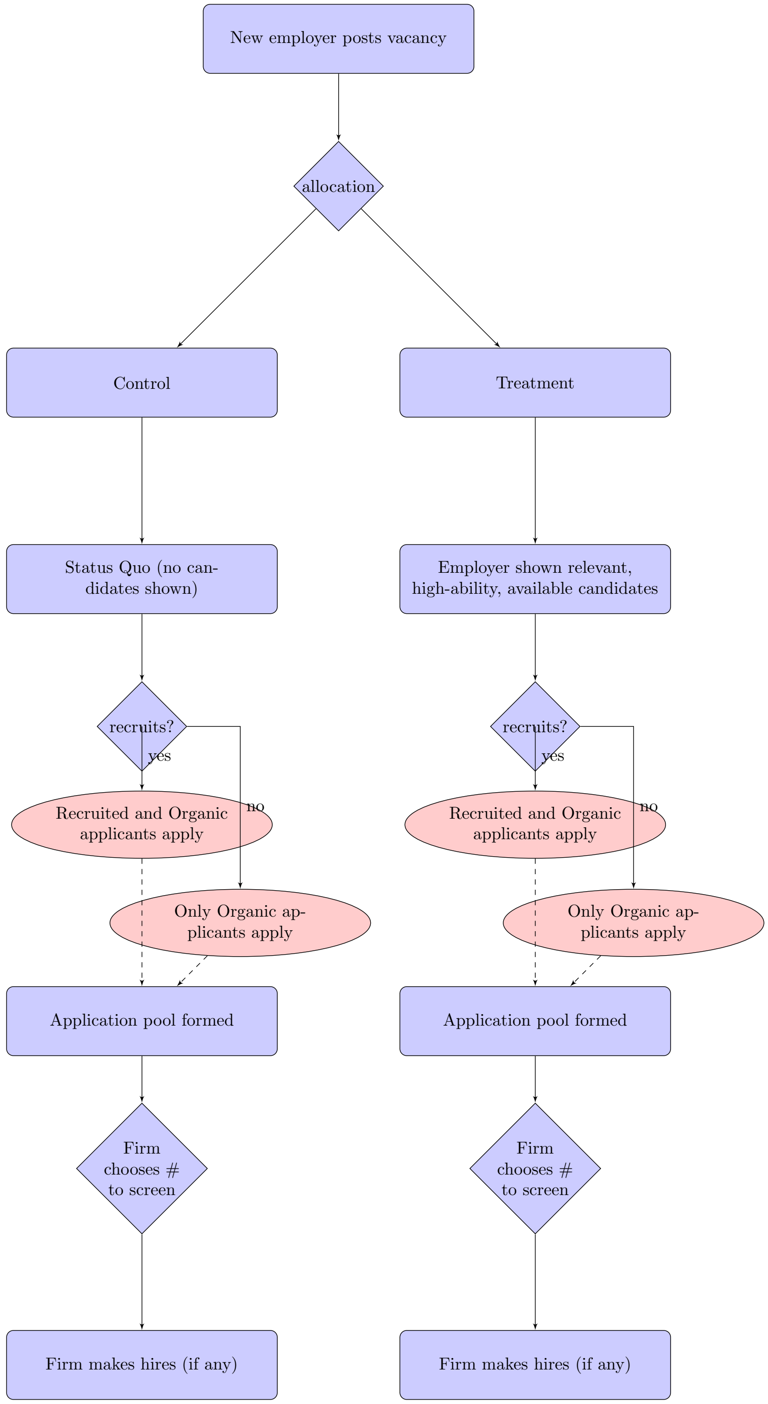
The experimental subjects were new employers that posted a vacancy during the experimental period. That period was approximately six weeks long, during the summer of 2011. Figure 1 illustrates how the experiment worked. First note that allocation occurred immediately after the employer posted their vacancy. In the treatment group, employers were shown a pop-up window with some number of recommended candidates along with their match-relevant details. Figure 2 is screen shot of the actual recommendation interface. For each recommended worker, employers could view their photograph, their listed skills, average feedback score and their stated hourly wage. If the employer clicked a worker's "tile" they could see the worker's country, total hours worked on the platform, portfolio size, passed skills tests and snippets of text from past employer evaluations. Employers in the treatment group could *choose* to invite some number (including zero) of the recommended candidates to apply for their job. Employers in the control received no recommendations. Because randomization occurred after the employer fully set a vacancies attributes (e.g., category, job description, required skills), vacancy attributes are orthogonal to treatment assignment.

It is important to note that once treated employers closed the recommendations pop-up window, they had the same interface and opportunities as the control group. This means that employers in both groups could use the existing marketplace search tools and find and invite other candidates to apply to their vacancy. If an employer invited a candidate—regardless of whether done through the "normal" channel or through the experimental channel—that candidate would receive an email notification with link to the employer's vacancy. The recruited candidate could then decide whether or not to apply.

If the employer posted the job publicly, then regardless of whether they recruited, other workers were free to find the vacancy and submit an application. These "organic" applicants would join whatever number of recruited candidates also chose to apply. In Figure 1, we illustrate how the employer's decision to recruit affects the applicant pool: if the employer recruited, then they can receive a mixture of both recruited and organic applicants. If the employer did not recruit, then they can only receive organic applicants.

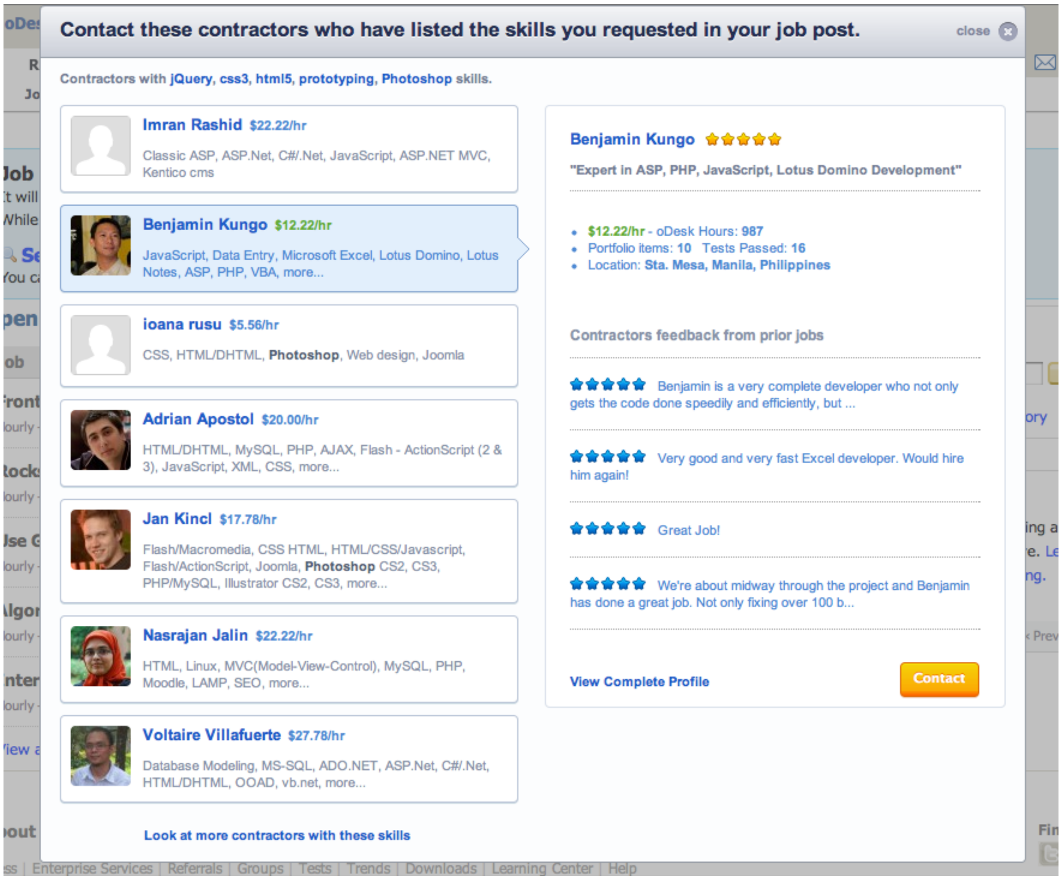
The experimental feature did not log which workers were actually recommended, so we cannot directly compare the collection of recommended candidates to those that were actually invited. However, this is a minor limitation for our purposes because (a) the unit of randomization is the vacancy, not the worker and (b) we can infer which invitations were experimentally induced by comparing the timing of those recommendation. Because of how recommendations were made, recommendation-induced invitations were made shortly after the employer

Figure 1: Employer recruiting experiment



Notes: This diagram illustrates the sequential steps of each vacancy in the experiment, by treatment and control group. Employers were first randomized into treatment and control groups, with treatment groups receiving algorithmic candidate recommendations. Both groups decided whether or not to recruit candidates. Depending on their choice, employers received recruited applicants (if they recruited) and organic applicants. Firms then choose whom to hire—if anyone—from this pool of applicants.

Figure 2: 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.

posted the vacancy. We can thus classify invitations as "early" (made within an hour posting) or "late," with early invitations being the treatment-induced invitations.

2.2 Ranking of candidates for recommendations

An algorithm determined which workers were recommended. These recommendations were determined in two steps. The first step identified skill-relevant workers while the second step ranked them by ability and inferred availability. The algorithm determined relevancy by comparing the required skills for the vacancy and the worker's profile-listed skills. Other workers attributes like test skill scores, feedback and past earnings determined the worker's ability score. The algorithm inferred availability from signals like a worker recently ending a project or apply-

Table 1: Data sources

Dataset Label	Description & Use
A-EXPERIMENT	This dataset contains 11,414 vacancies posted on oDesk by first-time employers during the summer of 2011. Employers in the treatment group received algorithmic candidate recommendations based upon the characteristics of their vacancies; employers in the control did not receive any recommendations. This dataset is used to examine the causal effect of the treatment on recruiting, screening, hiring and match quality.
B-APPLICATIONS	This dataset consists of the 358,324 organic and recruited applications to the 11,414 vacancies in A-EXPERIMENT. This dataset is used to characterize the employer screening process, particularly the differing experience of recruited and organic applicants.
C-VACANCIES-OBS	This dataset is a sample of 677,408 vacancies posted on oDesk from January 2011 to September 2012. It is used in an employer fixed-effects regression to examine other predictions of the model not addressed by the experiment, such as the effects of increased project value and labor market tightness on a firm's recruiting and screening decisions.
D-RECRUIT-RESPONSE	This dataset consists of a cross-section of 134,654 vacancies in which an employer made a single invitation to one worker. It is used to estimate the effects of a marginal recruited applicant on the vacancy fill probability and extent (if any) of crowd-out.

ing to other vacancies. If the algorithm failed to find suitable candidates, no recommendations were made.

2.3 To what extent can we treat oDesk as a self-contained marketplace?

The oDesk marketplace is clearly not "the" labor market. As such, we might worry that every vacancy we see on oDesk is simultaneously posted on several online labor market sites (e.g., Elance, Freelancer.com) *and* in the traditional market. If this is "multi-homing" is the case, then improvements in fill rates in one market are less impressive if these gains are simply coming at the expense of hires that would have been made in some other market. For example, a filled vacancy on oDesk mechanically reduces the filled vacancies on other sites. Further, statements about crowd-out and equity must be very qualified if all vacancies are in fact posted in multiple places, with both seen and unseen applicants.

Fortunately, we have some evidence on these questions from an independent market research firm that recently surveyed a random sample oDesk employers. The survey suggests that online and offline hiring are only very weakly substitutes and that multi-homing 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 said they would have made a local hire. Employers posting jobs report that they are generally deciding between (a) getting the work done online (b) doing the work themselves or (c) not having the work done at all. The survey also found that 83% said that they listed their last project on oDesk alone. While the survey evidence does show some evidence of multi-homing (both between traditional and online and between online markets), the extent of it is rather limited, making us more confident in inferring the equity and efficiency implications of the experiment.

2.4 Multiple comparisons and sub-population treatment effects

One the key results of the paper is that the treatment substantially increased the fill rate only in the software and web development categories of work, with no evidence of an overall effect across categories. A natural concern with this finding is that it stems from the multiple comparisons problem: if we examine enough sub-populations, eventually we will find some that have treatment effects by chance. However, there are several reasons why this is not a concern in this particular context. The first and strongest argument is that even with the conservative Bonferroni correction for simultaneous inference, the treatment effect on fill rates for the software and web development categories is still conventionally significant. If we use $\alpha = 0.05$ and we planned to compare treatment effects for two groups—"software and web" versus "all else", then our $1-\alpha/m$ Bonferroi-adjust confidence interval still does not include 0 effect: the p-value for the means comparison on fill rates is $p = 0.003$. As we might expect, we can also easily reject the null hypothesis of constant treatment effects across category.

The second argument is institutional, in the marketplace is thought of within the oDesk corporation as being broken into two segments. There is the "knowledge process outsourcing" segment, which includes clerical work, writing, graphic design etc. and a "technical" segment, captured almost entirely by the software and web-development categories. The different categories are heterogeneous in all most all the import primitives of any model of employer search—number of applicants, number screened, project value etc. Software and web development in particular stand apart and are qualitatively unlike other categories of work, due to the greater skill required and the associated higher wages. The natural binary character of the marketplace ex ante justifies analyzing the two "halves" of the marketplace separately.

3 Theoretical framework

Existing models of employer search are conceptually similar to simple job search models: firms serially screen applicants and hire the first one 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. Rees (1966) introduces an intensive screening margin to this framework, with firms also deciding how much effort to put into each screening. Empirical work within this framework is quite limited, but the results are largely consistent with the basic predictions of the search model. Barron and Bishop (1985) finds that employers with hard-to-fill vacancies or those that require more training report screening a larger pool of applicants and screen 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 jobs pay more, last longer and lead to greater employer satisfaction, though the direction of causation is not clear.

Despite the reasonableness of the job-search style model of employer search, we find it inadequate for our empirical setting. As such, our model differs substantially. The main substantive differences are that we assume (a) firms screen applicants in batch, (b) firms can alter the applicant ability distribution by costly recruiting and (c) screening costs are fixed and exogeneous. However, we still have a model in the tradition of Stigler (1961); there is uncertainty about the attributes of choices that can be reduced via costly effort.

3.1 Serial screening, altering the pool of applicants and the intensive screening margin

Burdett and Cunningham (1998); Barron et al. (1989); Barron and Bishop (1985) all assume that firms process applications serially whereas we assume that firms process applications in batch. In our model, conditional upon choosing whether or not to recruit, the firm's only decision is to choose a batch size. We assume batch processing because it appears to be the dominant mode of screening applicants in our setting. On oDesk, employers generally receive all the applications they will receive very quickly, and as such, the serial strategy is less appealing to employers. In Appendix A, we present evidence that oDesk employers process applications in batch. However, batch processing also seems to be commonplace and perhaps even dominant in traditional markets as well. Studying a traditional market, 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) reaches a similar conclusion.

Some findings from the existing empirical literature are easier to reconcile with a batch processing model. For example, Andrews et al. (2008) examine the duration of vacancies, finding that many vacancies simply go unfilled, with the firm canceling its search without hiring anyone. This result is hard to explain if firms face an infinitely deep pool of applicants drawn from some known ability distribution—since past screening costs are sunk, why stop looking? However, non-fulfillment is simple to understand when the applicant pool was fully (and finally) formed shortly after the vacancy was posted.

Our most consequential modeling innovation is that we assume firms can change the quality of the pool of applicants they draw from, albeit at some cost. While this feature is evident in oDesk data, it also seems consistent with the choices traditional firms make: at some cost, they can advertise, hire a recruiter, sponsor an event, etc. The options available to the firm differ in cost, the expected number of applicants they will generate and, critically, the *kind* of applicants they will generate. Making choices about these options is a large part of what firms are doing when trying to fill vacancies and our model captures this phenomena.

Existing employer search models assume that firms have a meaningful intensive screening margin, whereas our model assumes a fixed, per-applicant screening cost. While the degree of per-applicant screening certainly differs by occupation, deducing from this fact that firms decide how intensely to screen is analogous to deducing that because jobs differ in compensation, firms set wages. One reason to be skeptical of endogenous screening costs is that they are borne not just by the firm, but also by applicants—perhaps in equal proportion, as screening costs are largely interviewing costs. As such, firms will be powerless to vary the intensity of their screening procedures away from the "industry standard" without making side payments to applicants. Given the absence of side payments, it seems reasonable to model firms and workers in a competitive market as "interview onerousness"-takers.

4 Model of employer search and screening

Consider a firm that is attempting to fill a vacancy. The firm knows that it will receive A applicants, each with a match probability q distributed according to the pdf f and cdf F on the support $(0,1)$. Match probabilities are independent and identically distributed. The firm chooses some number of applicants a, with $a \leq A$ to screen. After the firm chooses a, the A applicants are rank ordered by their match probabilities and the firm screens the top a, each at a cost of screening c to the firm. Each screening is a Bernoulli trial and all trials are done simultaneously (Assumption 1). If any one of the screened applicants is a match, the firm fills the vacancy and gets $v > 0$; if no applicants are a match, the firm gets 0.

Assumption 1. *Applicants are screened in batch and firms must decide on the number of applicants to screen ex ante, before the realizations of q are known.*

For a given realization of match probabilities, $q_1, q_2, q_3 \ldots q_A$, we can define a hiring function, which is the probability that the firm makes a hire. This function is:

$$h(a) = 1 - \prod_{i=1}^{a}(1 - q_i) \tag{1}$$

Obviously the firm cannot interview fractional numbers of applicants, but having a be continuous makes the model more useful. To create a continuous hiring function, we first need to characterize q_i as a function of the number of screened applications and the applicant pool (a and A respectively) as well as the properties of f. To do this, we note that values of q_i in Equation 1 are realized order statistics of the distribution f when a sample size of A is taken. An approximation of the zth order statistic is:

$$\mathbf{E}[q_{(z)}] \approx F^{-1}\left(1 - \frac{z}{A}\right) = q(z) \tag{2}$$

where F^{-1} is the inverse cumulative density function, or quantile function. Note that when A is large and z is small, $1 - z/A \approx 1$, corresponding to a high quantile of f, giving a high q. Similarly, if $z \approx A$, the quantile of f is very low and hence q is low. Next we need to recast the product term of Equation 1 into a summation. By taking the log of the product term, and then using the approximation that $\log\left(1-y\right) \approx -y$ for small values of y, we can write:

$$\log(1-h(a)) = \int_0^a \log\left(1-q(z)\right) dz \approx \int_0^a -q(z) dz$$

which gives us

$$h(a) = 1 - \exp\left(\int_0^a -q(z) dz\right) \tag{3}$$

For an alternative derivation of Equation 3, consider the effect on the hiring probability arising from a small change in the number of screened applicants. If the firm interviews an additional $\mathrm{d}a$ applicants, then the probability that they are a match is $\mathrm{d}a q(a)$. However, the firm only "needs" this match with probability $1-h(a)$. This is the probability that the firm will not a find a match within the existing a applicants. Thus $\mathrm{d}h = (1-h(a))\, q(a)\mathrm{d}a$, which we can write as:

$$h'(a) = [1-h(a)]\, q(a) \tag{4}$$

The solution to this first-order differential equation is $h(a) = 1 - \exp\left(\int_0^a -q(z) dz\right)$, which is the same hiring function found in Equation 3. In the absence of strong tools to sort applicants, it might be reasonable to model recruited applicants as each having an identical match probability of q. Under this assumption, we have a hiring function $h(a) = 1 - \exp\left(-qa\right)$. We will refer to this hiring function as the homogeneous pool hiring function and make use of it to explore model outcomes that cannot be readily addressed with weaker assumptions.

4.1 Properties of hiring functions

As we expect, hiring functions are monotonically increasing but concave in a. The concavity of the hiring function is in a sense overdetermined. First, because we assume that the firm is ranking applicants and then drawing from the top, each successive applicant will have a lower expected match probability. Second, even if all applicants are identical, the need for only one match creates decreasing returns to additional screening.

Proposition 1. *The probability of filling a vacancy is increasing in the number of screened applicants:* $\forall a, h'(a) > 0$.

Proof. The hiring function is the probability of forming a match and thus for all $a > 0$, $h(a) \in (0,1)$. Because the support of f is $(0,1)$, for all a, $q(a) > 0$. With $h(a) < 1$ and $q(a) > 0$, by Equation 4, $h'(a) > 0$. □

Proposition 2. *The probability of filling a vacancy is concave in the number of applicants:* $\forall a, h''(a) < 0$.

Proof. Differentiating $h'(a)$ with from Equation 4 with respect to a and applying the chain rule, we have $h''(a) = [1 - h(a)]\, q'(a) - q(a)h'(a)$. The $q(a)h'(a)$ term is positive (by Proposition 1) and $1 - h(a) > 0$, so if $q'(a) < 0$, then $h''(a) < 0$. We can see that $q'(a) < 0$ by applying the inverse function theorem to Equation 2, which gives us $q'(a) = -\frac{1}{Af\left(F^{-1}\left(1-\frac{a}{A}\right)\right)} \leq 0$ (since F^{-1} and f are always positive). □

4.2 Introducing the possibility of employer recruiting

A goal of the model will be to understand when the firm will pay to recruit from a "better" applicant pool. Now we introduce the possibility that firms can recruit. We model the firm as facing a binary choice whether to pursue a *recruiting* strategy or a *passive* strategy. We represent this choice as the firm selecting x, where $x \in \{R, P\}$. Recruiting has two effects: it additively increases the number of applicants the firm can consider and it increases the quality of the pool of applicants. By "quality" we mean that the distribution of match probabilities for recruited applicants first-order stochastically dominates the same distribution for organic applicants (Assumption 2).

Assumption 2. *Recruited applicants are better than passive applicants, meaning that the distribution of match probabilities for recruited applicants first order stochastically dominates ($\succ$) the match probabilities for organic applicants:* $G_R \succ F_P$.

We assume that the firm gets A_R applicants drawn from a distribution with pdf g_R and cdf G_R. Thus, a recruiting firm will face $A_P + A_R$ applicants (Assumption 3), with A_R drawn from G_R and A_P drawn from F_P.

Assumption 3. *The quantity of applicants received from recruiting has no effect on the expected number of passive applicants:* $\mathbf{E}[A_P|A_R, x = R] = \mathbf{E}[A_P|x = P]$.

The match probabilities of applicants in the recruiting sample pool are distributed according to a mixture of g_R and f_P, which we show in Proposition 3 dominates the passive distribution.

Proposition 3. *If the recruiting strategy attracts any applicants, then the resultant mixture distribution of applicants is better than the distribution that the firm would have obtained from the passive strategy: If* $g_R \succ f_P$ *and* $A_R > 0$, *then* $f_R \succ f_P$.

Proof. If $g_R \succ f_P$, then $\forall q, F_P(q) > G_R(q)$. Dropping the function argument and writing F_R as a mixture distribution, we have $F_P = \frac{A_R}{A_R+A_P}F_P + \frac{A_P}{A_R+A_P}F_P > \frac{A_R}{A_R+A_P}G_R + \frac{A_P}{A_R+A_P}F_P = F_R$. □

Proposition 4. *For the same number of screened applicants a, if $f_R \succ f_P$, then $h_R(a) > h_P(a)$.*

Proof. $f_R \succ f_P$ implies $\int_0^a q_R(z)dz > \int_0^a q_P(z)dz$ since each quantile of f_R is greater than the corresponding quantile in f_P. This in turn implies that $\exp(-\int_0^a q_R(z)dz) < \exp(-\int_0^a q_P(z)dz)$ and hence by Equation 3, $h_R(a) > h_P(a)$. □

It is important to note that Proposition 4 does not imply that the firm first screens all the recruited applicants and then the organic applicants, creating a kinked hiring function. Although recruited applicants come from a better pool, they do not necessarily strictly dominate passive applicants. Also, firms do not care about source when rank-ordering them, as there is no inference left to be made: the firm observes the realized q for each of the applicants.

4.2.1 Firm's decision problem with costly recruiting

Choosing whether or not to recruit determines the firm's hiring function. For a given hiring function, the firm's optimization problem is selecting the profit-maximizing number of applicants to screen, given that each screening has a cost c. The optimal number of applicants to screen is $a^* = \arg\max_a vh(a) - ca$, subject to $a \le A$. If the problem has an interior solution, then the employer chooses an a such that the marginal benefit equals the marginal cost, or $vh'(a^*) = c$. There are two corner solutions: if $c/v > h'(0)$, then the firm screens no one and if $c/v \le h'(A)$, then $a^* = A$.

Pursuing a recruiting strategy has a lump sum cost, $s > 0$. The firm simply selects either the passive or the recruiting strategy depending on which one offers the higher pay-off. Equation 5 states the firm's decision problem.

$$\arg\max_x \pi(x) \quad \text{where} \quad \pi(x) = \begin{cases} vh_R(a_R^*) - ca_R^* - s, & \text{if } x = R \text{ (firm recruits)} \\ vh_P(a_P^*) - ca_P^*, & \text{if } x = P \end{cases} \tag{5}$$

4.3 Model predictions about recruiting and screening

The predictions of the model are fairly straightforward and we can illustrate several of them in Figure 3. The firm's expected gross benefit from the recruiting and passive strategies as a function of the number of applicants screened are $vh_R(a)$ and $vh_p(a)$ respectively. In Figure 3, for the recruiting strategy, we draw three curves: $vh_R(a)$, $vh_R(a) - s_L$ and $vh_r(a) - s_H$. As the recruiting cost is a lump sum, the three curves have the same slopes at all points and the optimal number of applicants, a_R^*, is the same in all three cases. For the passive strategy, the single curve is $vh_P(a)$, with a single optimal point a_P^*.

In Figure 3, we plot the point $(a_R^*, vh_R(a_R^*))$ as D and the point $(a_P^*, h_P(a_P^*))$ as point E. As we can see, the point D is higher than point E, implying that $h_R(a_R^*) > h_P(a_P^*)$ (the v cancels out), i.e., the recruiting strategy offers a higher fill probability than the passive strategy. This turns out to be a general prediction of the model, as shown by Proposition 6. However, before proving this, we first show in Proposition 5 that when two strategies have the same fill rate, a marginal applicant offers more benefit when the firm is pursuing a recruiting strategy.

Proposition 5. *If two strategies yield the same fill rate, then the effect of a marginal applicant is greater in the pool that first order stochastically dominates the other: If $f_R \succ f_P$ and $\exists a_R, a_P | h_R(a_R) = h_P(a_P)$, then $h'_R(a_R) > h'_P(a_P)$.*

Proof. By assumption, the two strategies have the same fill rate, so $1 - h_R(a_R) = 1 - h_P(a_P)$. Because $f_R \succ f_P$, it takes fewer applicants to obtain the equal fill rates, so $a_R < a_P$, which implies that $q_R(a_R) > q_P(a_P)$ (by $f_R \succ f_P$). Thus, $q_R(a_R)(1 - h_R(a_R)) > q_P(a_P)(1 - h_P(a_P))$, which by Equation 4 implies that $h'_R(a_R) > h'_P(a_P)$. □

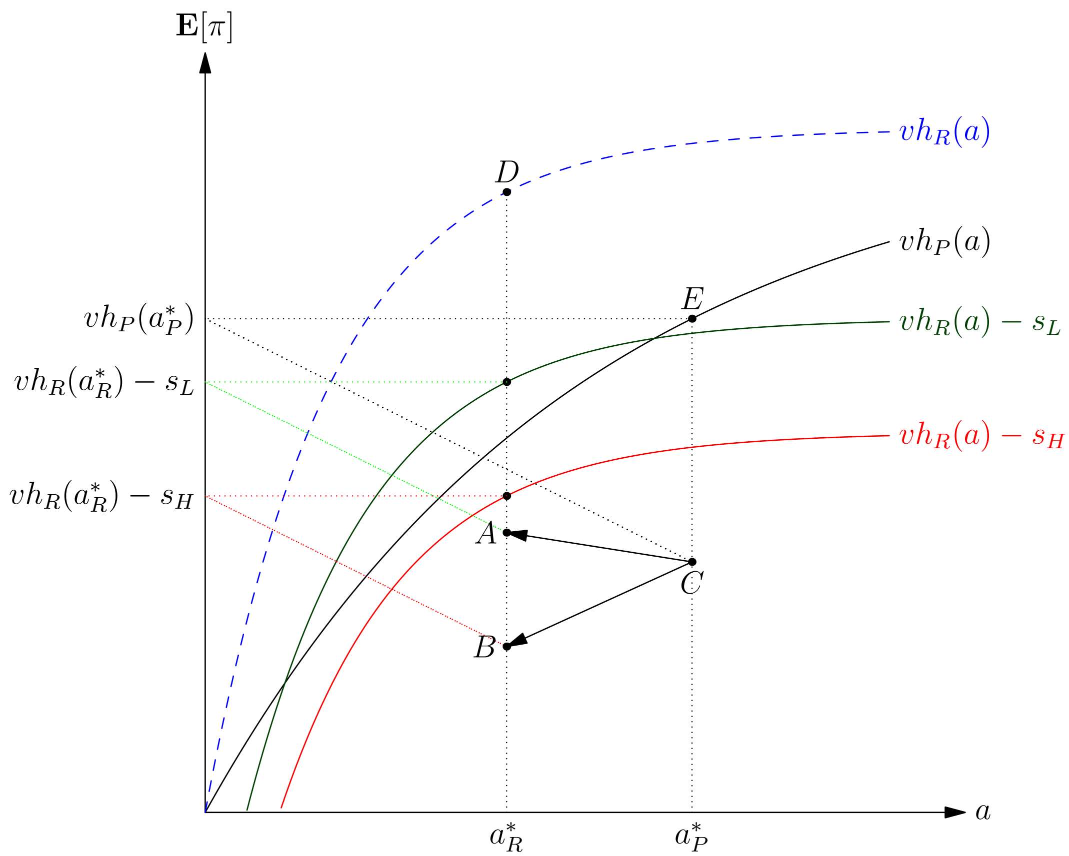
Proposition 6. *Fill rates are higher when the employer recruits: $h_R(a_R^*) > h_P(a_P^*)$.*

Proof. If $a_R^* > a_P^*$, then $h_R(a_R^*) > h_P(a_P^*)$ by Proposition 4 and Proposition 1. If $a_R^* < a_P^*$, then consider $a' | h_P(a') = h_R(a_R^*)$. Because $h'_R(a_R^*) = h'_p(a_P^*) = c$, by Proposition 5, $h'_p(a') < c$, and since $h'_p(\cdot)$ is concave, $a_P^* < a'$ which implies that $h'_p(a_P^*) < h_R(a_R^*)$, again by the concavity of $h_P(\cdot)$. □

To decide whether to recruit, the firm considers not just the direct costs and benefits of recruiting, but also the associated screening costs, ca_R^*. We can illustrate these costs in Figure 3: start from the point $(0, vh_P(a_P^*))$ (the y-intercept) and draw a line with slope $-c$ for a_P^* units on the x-axis. In Figure 3, the height of point C is the expected utility of the passive strategy, while points A and B are the expected utilities of the recruiting strategy when lump sum recruiting costs are s_L and s_H, respectively. In this figure, at s_L, recruiting is more attractive, while at s_H, the passive strategy would be more attractive. In Proposition 7, we show—as we strongly suspect—that greater recruiting costs decrease the expected utility of recruiting.

Proposition 7. *An increase in recruiting costs decreases the value of recruiting and does not affect the value of the passive strategy: $\partial\pi(x = R)/\partial s < 0$ and $\partial\pi(x = P)/\partial s = 0$.*

Figure 3: The firm's decision to recruit or passively accept applicants



Notes: In this figure, the x-axis is a, the number of applicants the firm screens. The y-axis is the firm's expected value, which is the value that the firm gets from filling a vacancy, v times $h_x(a)$, which is the probability of filling the vacancy using strategy x when evaluating a applicants. Per-applicant screening costs are c and the "recruiting" strategy costs a lump sum of s to pursue. Point D marks the point on the recruiting hiring function curve ($vh_R(a)$) where the slope is c, corresponding to the profit-maximizing choice of a^* when an interior solution is optimal. The other two recruiting curves have the same a^* because they only differ from $vh_R(a)$ by a constant s. Point C is the net benefit of passive strategy, while points A and B are the net benefits of the recruiting strategy when lump sum recruiting costs are s_L and s_H respectively.

Proof. The expected utility of the recruiting strategy is $\pi(x=R) = vh_R(a_R^*) - ca_R^* - s$, while the value of the passive strategy is $\pi(x=P) = vh_P(a_P^*) - ca_P^*$. Taking the partial with respect to s, we have $\frac{\partial\pi(x=R)}{\partial s} = -1$ and $\frac{\partial\pi(x=P)}{\partial s} = 0$. $\square$

If v increases, the rewards to filling a vacancy are greater and we should expect more applicants to be screened. Proposition 8 shows that this is the case.

Proposition 8. *The optimal number of applicants is increasing in the project value for both the recruiting and passive strategies: $\partial a^*/\partial v > 0$ where $a^* = a_R^*$ or $a^* = a_P^*$.*

Proof. At the optimum value, $vh'(a^*) \equiv c$ (dropping the subscripts for R and P). Differentiating with respect to v and solving for $\frac{\partial a^*}{\partial v}$, we have $\frac{\partial a^*}{\partial v} = -\frac{h'(a^*)}{vh''(a^*)}$ and since $\forall a, h'(a) > 0 \land h''(a) < 0$, $\frac{\partial a^*}{\partial v} > 0$. $\square$

The number of screened applicants increases with project value regardless of strategy, but Proposition 9 shows that the effect is larger when the firm recruits.

Proposition 9. *An increase in project value, v, increases the appeal of the recruiting strategy more than it increases the appeal of the passive strategy.*

Proof. The partial derivative of the effect of an increase in v on expected utility is:

$$\begin{aligned}\frac{\partial\pi(x=R;v)}{\partial v} &= h_R\left(a_R^*\right) + \left(vh_R'(a_R^*) - c\right)\frac{\partial a_R^*}{\partial v} \\ &= h_R\left(a_R^*\right) \qquad \text{(by the envelope theorem)}\end{aligned} \tag{6}$$

By the same argument, $\frac{\partial\pi(x=P;v)}{\partial v} = h_P(a_P^*)$, and so by Proposition 6, $\frac{\partial\pi(x=R;v)}{\partial v} > \frac{\partial\pi(x=P;v)}{\partial v}$. $\square$

4.3.1 Effects of screening costs on recruiting and fill rates

Increased screening costs lower a^* and hence lower fill rates. However, small changes in screening costs have ambiguous effects on the relative attractiveness of recruiting. The reason is simple: by the envelope theorem, the effects on profits from a small change in c is a^* and the number of applicants screened under recruiting strategy, a_R^*, can be more or less than under the passive strategy, a_P^*. If $a_R^* > a_P^*$, then an increase in c makes recruiting relatively less attractive, but it also potentially makes the passive strategy no longer rational, making the recruiting strategy the only viable strategy.

Proposition 10. *A small increase in screening costs increases the relative attractiveness of recruiting if $a_R^* < a_P^*$.*

Proof. By the envelope theorem, the partial derivatives of the expected utility for the recruiting and passive strategies with respect to c are: $\frac{\partial\pi(x=R;c)}{\partial c} = -a_R^*$ $\frac{\partial\pi(x=P;c)}{\partial c} = -a_P^*$ □

We know that the recruiting strategy offers a higher fill rate, but how does this gap change if c increases? We are interested in this question in part because it might allow us to explain cross-category differences. It seems natural to assume that high-skilled, higher-valued positions have larger screening costs. If we assume that these differences are the only way that categories differ and that comparative static predictions hold generally, then we can connect by-category treatment effects on fill rates to category screening costs. If $h(a_R^*) - h(a_P^*)$ increases in c, then we should expect a bigger jump in fill rates from the treatment in high screening cost categories and vice-versa otherwise. Proposition 11 shows that the gap in fill rates depends on the relative change in the number of screened applicants with respect to c and that the gap is increasing with the homogeneous pool hiring function. However, it does not establish whether this is a general property of hiring functions.

Proposition 11. *The gap between the hiring probability is increasing in c if $\partial a_R^*/\partial a_c < \partial a_P^*/\partial c$. With the homogeneous hiring function, the gap in fill rates between the two strategies is increasing in the costs of screening.*

Proof. The gap is $h_R(a_R^*) - h_P(a_P^*)$.

$$\frac{\partial}{\partial c}[h_R(a_R^*) - h_P(a_P^*)] = h_R'(a_R^*)\frac{\partial a_R^*}{\partial c} + h_P'(a_P^*)\frac{\partial a_P^*}{\partial c} = c\left(\frac{\partial a_R^*}{\partial c} - \frac{\partial a_P^*}{\partial c}\right)$$

With the homogeneous pool hiring function, $\frac{\partial a_R^*}{\partial c} = -\frac{1}{cq_R}$ and $\frac{\partial a_P^*}{\partial c} = -\frac{1}{cq_P}$. Since $q_R > q_P$, by Proposition 11, $\frac{\partial}{\partial c}[h_R(a_R^*) - h_P(a_P^*)] > 0$. □

4.3.2 Effects of the applicant pool size on recruiting and screening

By Assumption 3, a change in the number of applicants obtainable through recruiting, A_R, has no effect on the expected payoff from the passive strategy: $\frac{\partial\pi(x=P)}{\delta A_R} = 0$. However, it improves the recruiting strategy both by making more and, on average, better applicants available.

Proposition 12. *An increase in the number of passive applicants, A_P, has a greater effect on expected utility of the passive strategy than of the recruiting strategy:* $\frac{\partial\pi(x=P)}{\delta A_P} \geq \frac{\partial\pi(x=R)}{\delta A_P} > 0$

Proof. At the optimum for each strategy, $h_R'(a_R^*) = h_P'(a_P^*)$. Because of Proposition 6 combined with Equation 4, we know that $q_R(a_R^*) > q_P(a_P^*)$. Because $q_R(a_R^*) > q_P(a_P^*)$, an additional applicant drawn from F_P has a greater chance of being included in the a_P^* than in a_R^*. In other words, a new applicant drawn from F_P (or any distribution) is more likely to be inframarginal (and hence expected utility raising) when $x = P$ than when $x = R$ because the q-threshold is lower in the $x = P$ case. $\square$

5 Descriptive statistics on vacancies

We introduce the main categories of work, highlighting the substantial cross-category variation the extent of employer recruiting, the number of organic applications and the sizes of projects. We then demonstrate that drawing a distinction between A, the number of applicants received, and a, the number of applicants screened, is well-supported by the data. Next we show that recruited applicants are far more likely to be viewed and hired by employers, motivating our characterization of recruited applicants as being "better" (as revealed by employer actions) than non-recruited applicants.

5.1 Characteristics of the vacancies by category of work

In Table 3, we present a number of summary statistics for vacancies in the control group, by the type of work. The type of work corresponds to the category the employer chose to list the vacancy in, with the exception of *misc.*, which contains several very small categories pooled together. The *web* category (short for web development) is the largest category by far. It had over twice as many vacancies during the experimental period as the next largest category, *admin*, which consists of relatively low-skilled work like data entry. The *software* category contains jobs that are often similar to *web* and could easily be placed in either category.

There is substantial cross-category variation in nearly every important vacancy attribute or outcome. For example, whether the firm recruits (the column "> 0 invites") varies from a high of 25% of employers posting a *web* vacancy to only 10% in *admin*. There is also large by-category variation in the number of applicants a vacancy receives, ranging from a high of 21 in *admin* to a little more than 7 in the *software* category. Job sizes—as measured by the average wage bill for filled vacancies—range from about $4.8K in *software* to only $743 in *writing* (which also includes translation). The *web* category is quite similar to *software* in terms of project size. In fact, there is a large gap between these two relatively high-skilled programming categories and the rest of the categories.

Table 2: Summary of model predictions and empirical findings

	True?	Note
Model assumptions		
A1: Applicants screened in batch	yes	Applicants before and after hired applicant equally likely to be viewed
A2: Recruited applicants are "better"	yes	Recruited applicants are more likely to be hired
A3: Num. organic applicants independent of recruiting	weak yes	Treatment did not affect number of organic applicants
Model predictions		
P1: Fill rate is increasing in number of applicants	weak yes	True cross-sections and in employer-specific FE model
P2: Hiring function is concave in number of screened applicants	-	-
P3: *Pool* from recruiting is "better"	yes	More likely to be viewed & hired
P4: Hiring function with recruits dominates passive hiring function	-	-
P5: Marginal effect of organic application lower when fill rates are the same	-	-
P6: Fill rates higher when firm recruit	mixed	Only detectable in high-skilled software and web development categories
P7: Lowered recruiting costs increases recruiting	strong yes	Very strong experimental evidence in favor
P8: Increased screening in project value	weak yes	True in cross-sections and with employer-specific FE model
P9: Increased recruiting in project value	weak yes	True in cross-sections and with employer-specific FE model
P10: Change in screening costs has ambiguous effects on appeal of recruiting	-	-
P11: High screening costs imply greater fill rate boost from recruiting	weak yes	If technical fields have higher screening costs, this would explain higher fill rates in treated software categories
P12: An increase in expected passive applicants makes recruiting relatively less attractive	moderate yes	True in cross-sections and with employer-specific FE model

Notes: This table summarizes the empirical evidence from the experiment and from the observational data with respect to the assumptions and predictions of the model.

Table 3: Summary statistics for the control group, by type of work

		Discrete employer outcomes:			Continuous employer outcomes		
	N	> 0 invites	> 2 invites	made a hire	# apps.	# invites	wage bill ($)
admin	854	0.107	0.046	0.184	21.30 (1.13)	0.18 (0.03)	1,822 (312)
misc	418	0.141	0.048	0.146	9.28 (0.86)	0.23 (0.04)	2,762 (668)
software dev.	678	0.193	0.097	0.130	7.14 (0.32)	0.58 (0.07)	4,884 (1,292)
design	621	0.206	0.063	0.316	13.90 (0.61)	0.34 (0.03)	1,316 (179)
writing	693	0.209	0.081	0.309	7.75 (0.33)	0.48 (0.05)	743 (110)
sales	634	0.243	0.096	0.211	9.16 (0.52)	0.67 (0.08)	2,422 (514)
web dev.	1,887	0.247	0.108	0.240	12.35 (0.28)	0.67 (0.04)	4,358 (430)

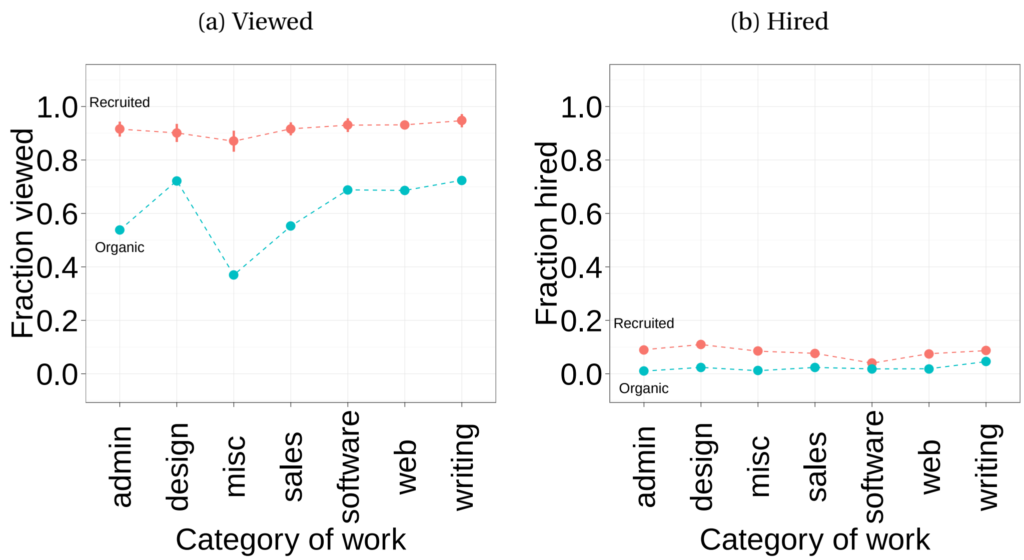
Notes: This table presents summary statistics for the control group for the experiment, which is a subset of dataset A-EXPERIMENT, described in Table 1. The "N" column is the number of observations in that category of work. All other columns contain the mean value for the variable in the control group. If the variable is not binary, then the standard error is reported in parentheses. The columns ">0 invites" and ">2 invites" show the fraction of employers sending any invitations or more than two invitations respectively. The column "apps." means the number of organic applications the vacancy received. For the "wage bill ($)" column, compute a trimmed mean that excludes the top 2.5% and bottom 2.5% of filled vacancies.

5.2 The outcomes of recruited and organic applications

In the model we assume that through costly recruiting the firm can improve the *quality* (as measured by per-worker match probability) of received applicants. If these applicants are better, then we should expect to see employers being more likely to consider their applications and ultimately hire them. We can define viewing as a binary indicator for whether the employer clicked on the worker's application (which reveals the worker's cover letter) sometime after they applied. Using data from B-APPLICATIONS(all the applications sent to vacancies in the experiment), in Figure 4, we show the mean fraction of applicants viewed and hired by application source and category.

Recruited applicants are viewed approximately 95% of the time across categories, while organic applicants average only about a 60% viewing percentage. They are also far more likely to be hired, with an average hire probability of nearly 10%, compared to a 2% hire probability among organic applicants. These two patterns support our modeling assumption that the pool of recruited candidates is "better" than organic candidates, as well as our drawing of a distinction between A, the number of applicants and a, the number of screened applicants.

Figure 4: Employer viewing of application, interviewing of applicants and hiring of applicants by category of work and application type (recruited versus organic)



Notes:: The two panels in this figure show the fraction of applications from dataset B-APPLICATIONS (described in Table 1) that were (a) viewed by employers and (b) hired by employer. The category of work is on the x-axis and the type of applicant is labeled on the line plots. Each point estimate is within a 95% confidence interval. However, these error bars are imperceptible because the number of applications is so large.

6 Effects of reducing recruiting costs via recommendations

In this section, we analyze the main results of the experiment. We first show that the randomization was effective by comparing the composition of vacancies in the treatment and control groups. We show that the treatment increased recruiting regardless of the category of the vacancy, but was only effective at substantially raising fill rates in the software and web development categories. In the software and web development categories, we decompose the hiring event into the source—recruited or organic—of the hired applicant. We find that the treatment raised the hire rate for both recruited applicants *and* organic applicants.

6.1 Randomization check

In Table 4, in the top panel labeled "Observation Counts" we report raw counts of observations in the treatment and control groups. As expected, the counts are consistent with a random assignment process (χ^2 test).[2] Below, the panel labeled "Fraction of Vacancies" we compare the fraction of vacancies in the different work categories, by treatment and control group. In the far right column, we report the p-value for a two-sided t-test for the null hypotheses of no difference in means. Also consistent with random assignment, the composition of vacancies is well-balanced across the groups.

6.2 Comparison of group means

Given the apparently effective randomization, we can begin the analysis by simply comparing treatment and control group means for the outcomes of interest. In the top panel of Table 5 labeled "Outcomes—All Categories" we compare the treatment and control groups on several post-treatment outcomes. The first two rows are the number of invitations sent by an employer ("Number of Invitations") and the number of organic applications received ("Number of applications"). There is no statistically significant difference across the two groups. A lack of difference in the mean number of organic applications is consistent with Assumption 3, which is that employer recruiting has no effect on the number of received organic applicants. However, both of these outcome measures have a high variance that makes it difficult to detect differences.

The next two rows report the fraction of employers sending at least one invitation and the fraction sending an "early" invitation, respectively. The treatment increased both measures and the effect is highly significant. Remaining outcomes are all different measures of hiring:

[2] As subjects were added to the experiment over time, allocation is done at random, at each moment, which is why the groups are not the same size.

Table 4: Check of balance across experimental groups

	Treatment	Control	p-value	
Observation Counts				
	5,629	5,785	0.144	(χ^2 test)
Fraction of vacancies				
admin	0.147 (0.005)	0.148 (0.005)	0.915	
misc	0.077 (0.004)	0.072 (0.003)	0.325	
software	0.114 (0.004)	0.117 (0.004)	0.538	
design	0.112 (0.004)	0.107 (0.004)	0.400	
writing	0.126 (0.004)	0.120 (0.004)	0.316	
sales	0.109 (0.004)	0.110 (0.004)	0.881	
web	0.316 (0.006)	0.326 (0.006)	0.222	

Notes: This table reports the count of observations (top panel) and fraction of observations in each category (bottom panel) by treatment and control group from A-EXPERIMENT, described in Table 1. The standard error for the mean calculation is in parentheses next to the mean estimate. The p-value is the for a two-sided t-test against the null hypothesis of no difference in means across the treatment and control groups, except for the top "Observation Counts" panel which reports a χ^2 test. Significance indicators: $p \le 0.05 : *$, $.p \le 0.01 : **$ and $p \le .001 : ***$.

whether the employer hired a recruited/invited applicant, an organic applicant and an early invited applicant. There is no statistically significant difference across the groups.

The bottom panel labeled "Outcomes—Software+Web" mirrors the middle panel but restricted to to the two programming categories, *web* and *software*. As our earlier discussion highlighted, vacancies in these two programming categories are comparatively high-skilled and high-valued and as such, are worth examining separately. As in the middle panel, the two recruiting measures are significantly different. However, unlike the middle panel, the hiring outcomes measures are positive and in most cases significant.

6.3 Effects of treatment on employer recruiting

If the treatment lowered the costs of recruiting, s, then by Proposition 7, recruiting should have increased in the treatment group. As we saw in Table 3, the level of recruiting differs strongly by category. A natural question is whether the treatment affected vacancies in different categories differently. Figure 5, Panel A, plots the mean fraction of employers recruiting, by treatment and control groups. As expected, recruiting is category-dependent: there is relatively little recruiting in *admin* (on the order of 10%) whereas it is closer to 30% in the case of *web*. In all categories, the fraction of employers recruiting was higher in the treatment group than in the control group,

Table 5: Means comparison of outcomes across treatment and control groups

	Treatment	Control	p-value	
Outcomes - All categories				
Number of invitations	1.018 (0.051)	0.951 (0.049)	0.345	
Number of applications	13.936 (0.306)	14.146 (0.304)	0.624	
Made invites	0.234 (0.006)	0.203 (0.005)	<0.001	***
Made early (1st hour) invites	0.172 (0.005)	0.132 (0.005)	<0.001	***
Filled vacancy	0.230 (0.006)	0.225 (0.006)	0.496	
Hired from invitation	0.065 (0.003)	0.066 (0.003)	0.885	
Hired form organic application	0.212 (0.005)	0.210 (0.005)	0.784	
Hired from early invite	0.038 (0.003)	0.035 (0.002)	0.404	
Outcomes - Software + Web				
Number of invitations	1.099 (0.067)	1.165 (0.084)	0.532	
Number of applications	11.921 (0.278)	12.301 (0.295)	0.342	
Made invites	0.274 (0.009)	0.233 (0.009)	0.001	***
Made early (1st hour) invites	0.203 (0.008)	0.151 (0.007)	<0.001	***
Filled vacancy	0.246 (0.009)	0.211 (0.008)	0.003	**
Hired from invitation	0.077 (0.005)	0.069 (0.005)	0.302	
Hired form organic application	0.219 (0.008)	0.195 (0.008)	0.036	*
Hired from early invite	0.048 (0.004)	0.037 (0.004)	0.047	*

Notes: This table reports statics by pre-treatment by treatment and control group from A-EXPERIMENT, described in Table 1. The standard error for the mean calculation is in parentheses next to the mean estimate. The p-value is the for a two-sided t-test against the null hypothesis of no difference in means across the treatment and control groups. Significance indicators: $p \le 0.05 : *$, $.p \le 0.01 : **$ and $p \le .001 : * * *$.

Table 6: Treatment effects on employer recruiting behavior

	Invitations made by employer:				
	Any	> 1	Any early	Any late	Any early
	(1)	(2)	(3)	(4)	(5)
Intercept	0.203***	0.116***	0.132***	0.101***	0.016
	(0.005)	(0.004)	(0.004)	(0.004)	(0.016)
Treatment	0.031***	0.017**	0.040***	−0.008	0.040***
	(0.008)	(0.006)	(0.007)	(0.006)	(0.007)
Job description length (log)					0.010***
					(0.003)
Firm required prior exper.					0.047***
					(0.013)
Category Fixed Effects	No	No	No	No	Yes
N	11,414	11,414	11,414	11,414	11,410
R-squared	0.001	0.001	0.003	0.000	0.015

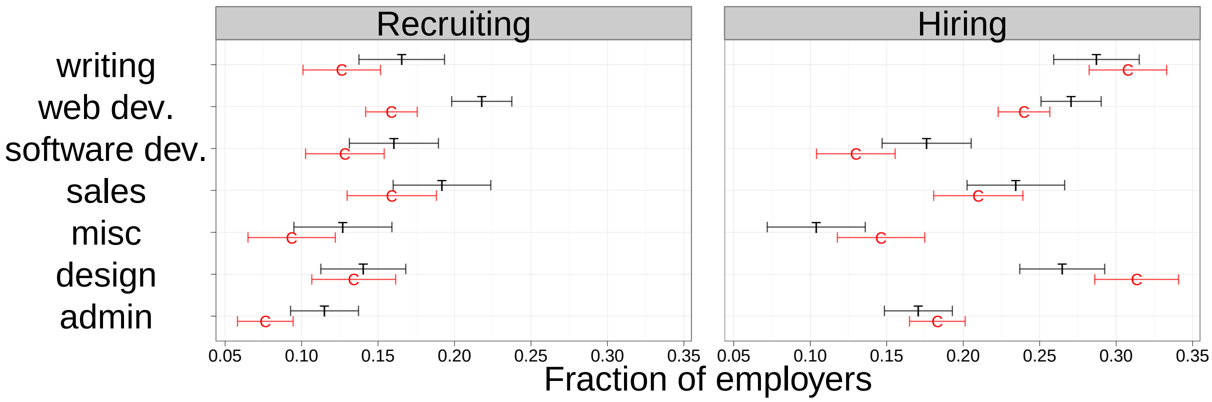
Notes: This table reports the results of OLS regressions of employer recruiting on the treatment indicator as well as various controls. The dataset is A-EXPERIMENT, described in Table 1. The dependent variable in Column (1) is whether the employer sent any invitations; in Column (2) it is whether they sent more than 1 invitation. In Columns (3) and (4) the dependent variable is whether the employer sent any "early" (first hour after posting their vacancy) or "late" invitations, respectively. The standard errors are robust to heteroscedasticity. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

though in many categories, the difference is quite small. Despite this heterogeneity, we cannot reject the null hypothesis of constant treatment effects across category (Wald test, $p = 0.67$).

In Table 6, Column (1), we report a regression where the outcome variable is whether the employer made any invitations (invites > 0) and the independent variable is the treatment indicator. The treatment had a strong, positive effect, increasing the fraction of employers making at least one invitation from a base of 20% to a little more than 23%. In Column (2), we regress an indicator for whether the employer sent more than one invitation. The effect is nearly halved, suggesting that much of the effect of the treatment came from inducing employers that otherwise would have sent no invitations to send at least one invitation.

We can more precisely characterize employer behavior by classifying whether invitations as coming early or late. Early invitations should be affected by the treatment but late invitations should not: recall that recommendations were made immediately after the employer posted the vacancy. In Column (3) we regress an indicator for whether the employer made any early invites on the treatment indicator. It shows the treatment effect is strong and positive. In contrast and as expected, when the outcome is whether the employer send any late invitations (in

Figure 5: Employer recruiting and hiring by category and treatment group



Notes: This plot shows the mean fraction of employers recruiting (as measured by sending an early invitation) in the left panel and making a hire in the right panel. Estimates are by category of work (listed vertically) and by treatment assignment, indicated by 95% confidence intervals marked with either "C" or "T" (for control and treatment respectively). The dataset used in this plot is A-EXPERIMENT, described in Table 1.

Column (4)) the effect of treatment is negative and insignificant. These findings suggest the treatment worked directly by providing candidates the employer could invite, rather than just alerting employers to the possibility of recruiting. The absence of treatment effects on late invites also suggests that early recruiting was not simply "moving up" late recruiting that would have happened anyway.

In Column (5) of Table 6, we add two pre-randomization attributes of the vacancy: (1) the actual textual length of the employer's job description (2) whether or not the employer required organic applicants to have some number of on-platform hours before applying. The length of the job descriptive is positively correlated with recruiting. One possible interpretation is that a longer job description proxies for a more valuable job (higher v); another interpretation is that it is a more complex, harder-fill project (lower A_P). Under either interpretation, Propositions 9 and 12 predict the finding of a positive correlation.

Requiring prior experience is also strongly positively correlated with propensity to recruit. Employers could have many reasons for requiring applicants to have on-platform experience, but regardless, the end result is presumably a smaller applicant pool. If A_p is smaller, the positive correlation between requiring prior hours and propensity to recruit is consistent with Proposition 12.

Table 7: Effects of treatment on hiring of applicants

	Employer hired:			
	All Categories	Sofware + Web		
	(1) Anyone	(2) Anyone	(3) Early recruit	(4) Organic applicant
Intercept	0.225***	0.211***	0.037***	0.195***
	(0.005)	(0.008)	(0.004)	(0.008)
Treatment	0.005	0.035**	0.011*	0.024*
	(0.008)	(0.012)	(0.006)	(0.012)
N	11,414	4,980	4,980	4,980
R-squared	0.000	0.002	0.001	0.001

Notes: The dependent variable in each of these OLS regressions is whether or not the employer hired a worker or a particular type: in Columns (1) and (2) the indicator is for hiring anyone at all, while in Column (3) is it is for hiring an early recruit and in Column (4), it is hiring an organic applicant. Regressions in Columns (1) is for all vacancies, while the remainder are just those vacancies in the software and web development categories. The dataset is A-EXPERIMENT, described in Table 1. The standard errors are robust to heteroscedasticity. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : * * *$.

6.4 Effects of treatment on hiring

We have strong evidence that the treatment expanded employer recruiting. We now examine whether this increased recruiting had any effect on vacancy fill rates. Proposition 6 predicts that firms induced to switch to the recruiting strategy should have higher fill rates. In Table 7, Column (1), we report the results of a regression of an indicator for whether the employer filled their vacancy (as measured paying some amount of wages) on the treatment indicator. Although positive, the treatment effect is not significant. However, the effect of the treatment seems to depend strongly on the category of work. In Figure 5, Panel B, we plot the fraction of employers filling their vacancy, by treatment group and by category. In many of the categories, the experimental groups do not differ, yet in *software* and *web*, the treatment group outpaces the control. A Wald test confirms our visual inspection: we can reject a null hypothesis that the treatment effect does not differ across category ($p = 0.023$).

Motivated by this finding of heterogeneous treatment effects, In Column (2) of Table 7, we regress the same indicator for hiring on the treatment indicator, but with the data restricted to vacancies in the two programming categories. The resultant estimate is large and significant, with the treatment raising the fill rate from a baseline of about 21% by 3.5% points.

6.5 Why are treatment effects on fill rates heterogeneous?

We would like to explain why the treatment was only effective at raising fill rates in the software and web development categories. One possible explanation is based upon category-specific heterogeneity in relative screenings costs: if screening costs are relatively high in these categories, then Proposition 11 can explain the observed pattern: Proposition 11 predicts that, all else equal the jump in fill rates from switching to a recruiting strategy is higher when screening costs are high. Of course, we do not directly observe screening costs. However, we do observe the consequence of screening costs (and the quantity and quality of received applicants), as these factors determine the number of evaluated applicants, which in turn determines whether or not the employer filled their vacancy. Using several assumptions about the data generating process, we can use our model of employer search to structurally estimate category-specific screening costs and applicant quality. We can then compare the estimated screening costs across category and see whether—consistent with our heterogeneous treatment effect result—screening costs are relatively higher in the programming categories.

One problem is that "evaluation" is not unambiguously measured in the data. However, we can proxy for a^* in several ways. One approach is to simply use the count of the number of interviews conducted; another is to use the number of applicants received. Using the number of applicants received seems counter to our a^* versus A modeling distinction, but not if we believe that employers immediately close a job once they have received a^* (even though they would have received A_p in total). A third approach is to simply use the number of applicants viewed by the employer, though this is surely an over-count. We restrict the sample to observations where at least one evaluation—regardless of how measured—was observed.[3]

To generate an estimate-able model, we need to make several consequential assumptions. First, we pool observations (regardless of treatment assignment) into two groups, indexed by k: the "Programming" group, consisting of the *software dev.*and *web dev.* categories and "Other" consisting of all other categories. We assume that all projects within a category have the same screening-cost to value ratio, $\frac{c_k}{v_k}$ and that all applicants have homogeneous quality q_K, which makes the hiring function $h_k(a) = 1 - \exp(-q_k a)$.

To estimate the model, we will generate two moment conditions. First note that because we observe whether or not a vacancy fills and because the model gives the probability of fill as a function of q and a^*, we can directly write the likelihood function (where $y = 1$ indicates a filled

[3]The optimization step was done using Mathematica's NMinimize function, using the default settings.

vacancy and $y = 0$ an unfilled vacancy).

$$L(q_k) = \prod_{i=1}^{N_k} h(a_i^*; q_k)^{y_i} \left(1 - h(a_i^*; q_k)\right)^{1-y_i}$$

which in term gives us one moment condition based on the score: $\mathbf{E}[L'(q_k)] = 0$. Our second moment condition comes from the employer's optimizing behavior with respect to the number of applicants to evaluate:

$$\mathbf{E}[h'(a_i^*; q_k) - c_k / v_k] = 0 \tag{7}$$

The vector of moment conditions is thus:

$$g\left(a^*, y; \frac{c_k}{v_k}, q_k\right) = \begin{bmatrix} h'(a) - c_k / v_k \\ \frac{\partial \log L(q_k)}{\partial q_k} \end{bmatrix} \tag{8}$$

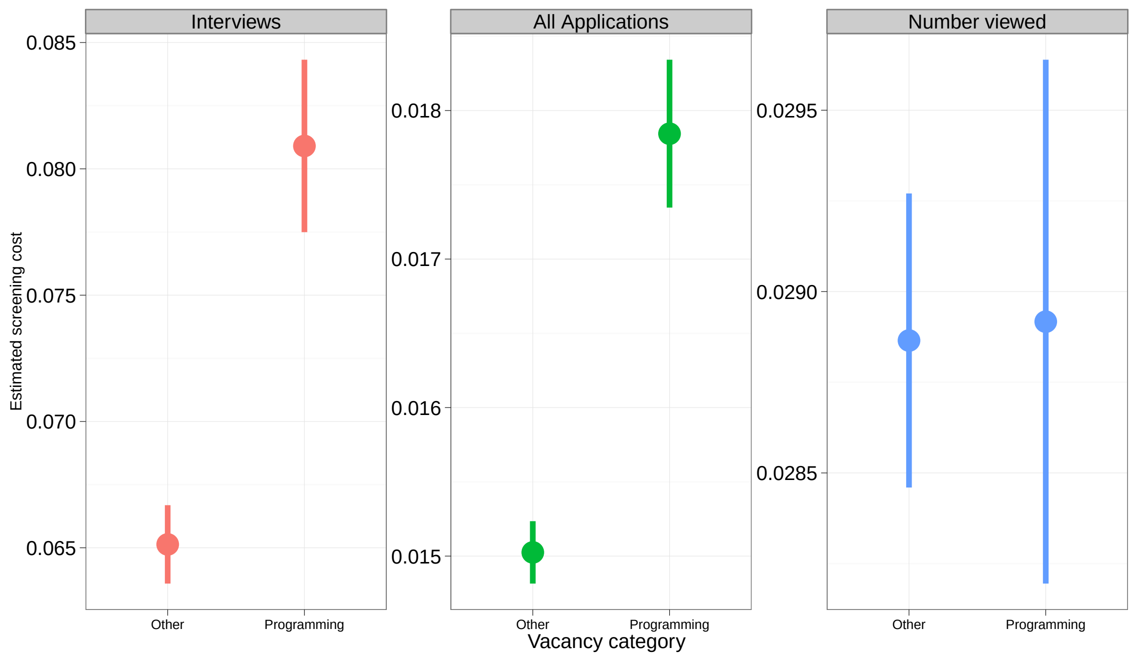
We can use the generalized method of moments (GMM) to estimate $\frac{c_k}{v_k}$ and q_K (Hansen, 1982). In Figure 6, we plot $\frac{\hat{c}_k}{v_k}$ using three different measures of the number of evaluated candidates. A 95% CI is shown for each point estimate (with standard errors computed using the weight matrix of the second step of the two-stage GMM). In the left and middle panels, the c/v ratio is far higher in the programming group (0.065 versus 0.080 using the interviews measure and 0.015 to 0.018 using the total applications measure); in the third panel, there is no meaningful difference in the two estimates.

Given that the left panel, "Interviews" proxy is probably the best measure of a^*, this estimation exercise suggests that screening costs may be higher in the programming categories, though the persuasiveness of this evidence is limited by the strong assumptions needed to obtain it. A more realistic approach might be to allow for per-vacancy heterogeneity in screening costs and project values (which are not directly observed), though obviously this would greatly increase the complexity of estimation and would likely make it more difficult to compare screening costs across categories.

6.6 Decomposition of the hiring effect into type of applicant hired

We know that *software* and *web* vacancies in the treatment group were more likely to be filled, but the precise causal mechanism is unclear. The model predicts that that a recruiting employer faces a more attractive slate of candidates and is more likely to make a hire (Proposition 6). In Column (2) or Table 7, the dependent variable was whether the employer hired anyone. In Column (3), the dependent variable in the regression is whether the employer hired a recruited

Figure 6: Estimate of the ratio of screening costs to project values



Notes: This plots shows the estimate $\frac{c_k}{v_k}$ (using data from A-EXPERIMENT) to estimate a structural model of employer hiring. The estimates were obtained using two-stage GMM. Moment conditions were based on the (a) the likelihood function for the observed fill rates and (b) the assumption that employer's evaluate candidates until the marginal benefit equals the marginal cost.

applicant (restricted to early invites). As expected, the treatment is positive and significant. In Column (4), the outcome measure is whether the employer hired an *organic* applicant. Given that most employers are only looking to hire one applicant, this result is surprising. The treatment increased hiring among organic applicants by an even larger amount.

Despite the attractiveness of a "substitutes" framing of organic and recruited applicants, there are several plausible sources of complementarity, including one that is a direct prediction of the model. If $c > v h_P(0)$, then $a_P^* = 0$, meaning that no applicants are screened. If $v h_R(0) > c$, then $a_R > 0$, and so long as $\sup f_P > \inf g_R$, i.e., the best passive applicant is better than the worst active applicant, a non-zero number of passive applicants will get screened *only* when the firm recruits. Because our treatment did induce more employers to recruit, more hires from organic applicants is plausible.[4] Even without special constraints placed on c, perhaps organic applicants benefit from "spill over" attention. As we saw in Figure 4, firms are far more likely to view recruited applicants. It is possible that in the treatment, the greater enthusiasm for viewing of recruited applicants meant more visits to the site, longer screening sessions etc. that benefited organic applicants as well. Yang and Ghose (2010) provides a similar example of "obvious" substitutes actually being complements in an electronic commerce setting. They found that paid and organic search results in a search engine were complements, each stimulating more clicks on the other "side."

6.7 Micro-foundations for spill-overs

Employers are emailed a notice when recruited candidates accept their invitations. Perhaps these emailed notifications—which only occur in response to the invited candidate accepting—trigger a screening session that also benefits organic applicants. We know from Figure 4 that employers are far more likely to view the applications of recruited applicants. If some of this attention spills over onto organic applicants, then applicants arriving shortly *before* recruited applicants should receive more position compared to those arriving shortly after.

Of course, any detectable effect would be confounded with time. And if the employer makes a hire (which is publicly visible to would-be applicants), then the applicants arriving immediately before and immediately after could be quite different. Furthermore, vacancies getting positive responses could be systematically different from those receiving negative responses. We can address these concerns in several ways. First, we can restrict our attention to the win-

[4]This story begs the question of why an employer would bother posting a vacancy if they planned not to screen any applicants. One possibility is that v is a random variable prone to idiosyncratic negative shocks. For example, employers can easily multi-home, so a promising offer from another site could explain a post-and-abandon outcome on a any particular site.

dow around recruited applicants that responded positively but were not hired. Second, we can include very flexible controls for arrival time by havingan arrival rank specific effect. We can also include a vacancy specific effect. Finally, we can compare the effects of arriving recruited applicants to the effects of the arrival of a counter-factual "placebo" group of invited but declining candidates. In other words, we assume that recruited but declining candidates actually did apply and had arrival patterns similar to the actually arriving applicants.

6.7.1 Effect of applying immediately before and immediately after a recruited candidate

Let $2D$ be the number of applicants around the arrival of a recruited applicant that we are interested. For each applicant, we can define a categorical variable d_{ij} that is the relative rank of that applicant vis-a-vis the recruited applicant: $d_{ij} = (r_{ij} - z_j)1 \cdot \{|r_{ij} - z_j| <= D \wedge r_{ij} \neq z_j$ where z_j is the arrival rank of the recruited applicant. Let y_{ij} be some outcome of interest such as hiring or interviewing. Let γ be a vacancy-specific effect and $\alpha(r_{ij})$ be an arrival-rank specific effect. We can estimate Equation 9 to obtain the estimated marginal effect of being in relative position d_{ij} on the outcome of interest. For the associated placebo regression, we replace d_{ij} with d'_{ij}, where d'_{ij} is based on the predicted counter-factual arrival, z'_j of the negatively responding non-recruited candidate. The predicted arrival is made using a linear model fit with the positive-responding applicants.

$$y_{ij} = \beta_d \cdot \{d = d_{ij}\} + \gamma_j + \alpha(r_{ij}) + \epsilon_{ij} \tag{9}$$

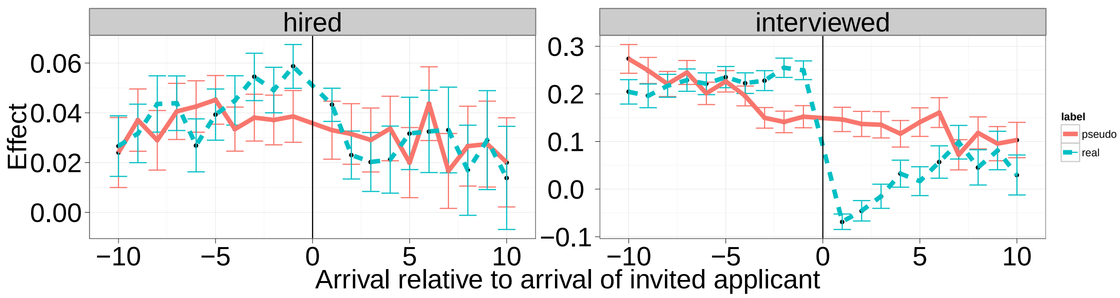
In Figure 7, we plot $\hat{\beta}_{-D}$ to $\hat{\beta}_D$, fit with a multi-level model, using data from B-APPLICATIONS, pooled across both the treatment and control groups. The dashed line is for the window around actual arrivals, while the solid line is for the placebo window. In the left panel, the outcome is hiring, while in the right panel, the outcome is interviewing.

The four applicants arriving immediately before the recruited applicant were both more likely to be hired and more likely to be interviewed. Applicants arriving after the arrival seem to suffer a dramatic fall-off vis-a-vis the placebo group, whereas in hiring, the placebo and actual groups have similar outcomes. The real difference in hiring is not in the post arrival period, but in the noticeable bump of "extra" attention of these applicants.

6.8 Effects of treatment on match attributes

In addition simply examining whether or not a match was formed, we also have some indicators for the attributes of the formed match. In particular, we can examine the wages of hired

Figure 7: The effects of applicant arrival position relative to the arrival of a recruited applicant (actual or placebo) on interviewing and hiring



Notes:: In this figure we plot the coefficients from estimates of Equation 9. The dashed line connects the estimates for the marginal effect of position relative to the arrival of a recruited applicant on the probability of hiring (left panel) or interviewing (right panel). The solid line connects the same estimates, but for the counter-factual "arrival" of recruited applicants that rejected the employer's invitation. The estimates of β_D were computed using a multi-level model, using the pooled data from B-APPLICATIONS.

applicants and the feedback they received from employers.

6.8.1 Wages of hired applicants

The model of employer screening and hiring is silent on how wages are determined. In Table 8, we regress an indicator for whether the mean hired wage for that vacancy exceeded \$3/hour, \$9/hour and \$12/hour in Columns (1), (2) and (3), respectively. We use this quantile indicator approach because for vacancies that were not filled, we do not observed wages. We use the mean hired wage because an employer can hire multiple workers. As we can see, the treatment has no discernible effect on the realized wages.

6.8.2 Employer-provided feedback to hired applicants

We know that the treatment affected recruiting and we have reasonably strong evidence that it increased hiring. A natural question is whether these treatment-induced matches were good matches. One proxy for match quality is employer feedback to workers following the completion of a project. When a project ends, employers publicly rate workers on 1-5 point "star" scale. This measure of match quality is not ideal. Feedback is inherently subjective and endogeneous, in that rating is given by the same firm that also made the hiring decision. The feedback scores are also compressed: as on many electronic commerce sites with a feedback

Table 8: Effects of treatment on wages

	Hired hourly wage exceeded:		
	(1) $3/hour	(2) $9/hour	(3) $12/hour
Intercept	0.126***	0.073***	0.047***
	(0.004)	(0.003)	(0.003)
Treatment	−0.003	−0.002	−0.003
	(0.006)	(0.005)	(0.004)
N	11,414	11,414	11,414
R-squared	0.000	0.000	0.000

Notes: The dependent variable in each of these OLS regressions is whether or not the employer hired a worker or a particular type: in Columns (1) and (2) the indicator is for hiring anyone at all, while in Column (3) is it is for hiring an early recruit and in Column (4), it is hiring an organic applicant. Regressions in Columns (1) is for all vacancies, while the remainder are just those vacancies in the software and web development categories. The dataset is A-EXPERIMENT, described in Table 1. The standard errors are robust to heteroscedasticity. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

mechanism, there are strong incentives for parties to give positive feedback regardless of the outcome of the project. A little more than 50% of all feedback scores are perfect, 5-star feedback scores. A final complication is that we only have a feedback measure if a hire was actually made. For this reason, as an outcome variable we use an indicator for whether a 5-star feedback was given for a particular vacancy. This indicator is zero if (a) some feedback other than 5.0 was given or (b) no one was hired.

In Table 9, Column (1), we use the entire sample of vacancies and find no treatment effect. In Column (2), we estimate the model using only vacancies that were filled. As in Column (1), there are no treatment effects. We find no evidence that matches formed in the treatment group were better or worse than those formed in the control group.

7 Crowd-out and spill-over in the larger marketplace

From our experimental results and the analysis of hiring and our analysis in Section 6.7, we suspect that organic applicants receive a spill-over of attention from recruiting employers. Since this result is counter-intuitive but could have practical importance, we would like additional confirmation. In the analysis that follows, we estimate the effect that a recruited worker's pos-

Table 9: Effects of treatment assignment on the feedback of the created firm-worker matches

	Employer gave 5-star feedback?:		
	All vacancies	Filled vacancies only	
	Yes	Yes	Yes
	(1)	(2)	(3)
Intercept	0.093^{***}	0.351^{***}	0.369^{***}
	(0.004)	(0.029)	(0.030)
Treatment	0.002	−0.007	−0.007
	(0.005)	(0.019)	(0.019)
Hired from invitation			-0.076^{***}
			(0.022)
Category Specific FE	No	Yes	Yes
N	11,414	2,599	2,599
R-squared	0.000	0.013	0.017

Notes: This table reports the results of OLS regressions where the dependent variable is whether or not the employer posting the vacancy gave 5-star feedback to a worker hired for their vacancy. The independent variables are the treatment indicator and category-specific controls. Note that the outcome is zero for both employers who gave less than 5-star feedback *and* for employers that failed to fill their vacancies. In Columns (2) and (3), we restrict the sample to only those employers filling their vacancies. The data used in all regressions is A-EXPERIMENT, described in Table 1. Standard errors are robust to heteroscedasticity. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

itive response to an employer invitation has on vacancy fill rate. We use a sample of vacancies where the employer made just one invitation. The response to this invitation determines whether the employer has a recruited applicant in their applicant pool. Because the response to this invitation is endogenous, we use an instrumental variables strategy. The instruments are constructed from factors that affect recruited worker response rates but are difficult for employers to condition upon directly. Unsurprisingly, we find that a positive response to an invitation increases the fill rate. However, when the outcome is whether the employer hired an *organic applicant*, the effect of a positive response to the recruitment invitation is also positive, albeit not significant. If crowd-out from recruited applicants were important, we would expect a strongly negative effect.

7.1 Empirical framework for estimating the effect of recruited applicants

In the actual marketplace, a worker's decision to accept an employer's recruitment invitation is clearly endogeneous. As such, the number of recruited applicants in an employer's pool of applicants is also endogeneous. Although the experiment manipulates the number of recruited applicants available to an employer, it does so indirectly. A more direct would be to manipulate the number of recruited applicants that accept an employer's recruiting invitation. Manipulating invitation responses is not feasible, but many recruited candidates choose not to apply for factors that are partly unrelated to the attributes of the inviting employer and their vacancy. These unrelated factors could potentially serve as instruments for the invited worker's response to the invitation. A valid instrument would strongly affect the recruited candidate's probability of applying and yet satisfy the exclusion restriction.

Workers can only work on so many projects. Their capacity—and hence willingness to accept invitations—ebbs and flows as projects end and new projects begin. Because of this flow, we suspect that their response probability is time-varying. We hypothesize that a worker's per-invitation response probability is decreasing in the number of invitations received and increasing in the number of other invitations accepted. The count of invitations received and responded to positively is endogenous, but presumably the precise counts have a stochastic component. Our identifying assumption is that conditional upon weekly counts and responses, *daily* counts and responses are exogeneous. In other words, we use the "residual" variation in daily counts as the source of variation in positive response probability. We are assuming that employers do not condition their choices upon this residual. On oDesk, invitation counts and responses are observable to employers. It is possible that an employer could condition their recruiting choice on these counts or other more static worker factors correlated with these counts.

However, it seems unlikely that employers perceive and act upon information conveyed by the daily count that is not already conveyed by the weekly count.

7.2 Effects of recruitment response on vacancy outcomes

In Table 10 we present our instrumental variables analysis. The endogeneous regression is whether the invited worker accepted the employer's invitation. The dependent variables are whether the employer hired anyone (Column (2)), an organic applicant (Column (3)) and a recruited applicant (Column (4)). In Column (1), we present the first stage regression but without showing the coefficients on the month and category dummies. The first stage implies we have strong instruments. The coefficients are highly significant and have the predicted signs; the F-statistic for the model is over 40. Each of the second-stage regressions suggest the exclusion restriction is met, with each regression readily passing the Sargan overidentifying restrictions test.

Column (2) implies that the acceptance of a recruiting invitation raises the fill probability by nearly 40% points. Most of this effect is due to the hiring of that recruited applicant, as we see from Column (4), where the coefficient on invitation acceptance is smaller but of similar magnitude. However, the surprising result is that the coefficient on invitation acceptance in Column (5) is not negative. The dependent variable in Column (5) is whether the firm hired an *organic* applicant. We now have an additional piece of evidence that the "substitutes" framing of organic and recruited applicants is incorrect.

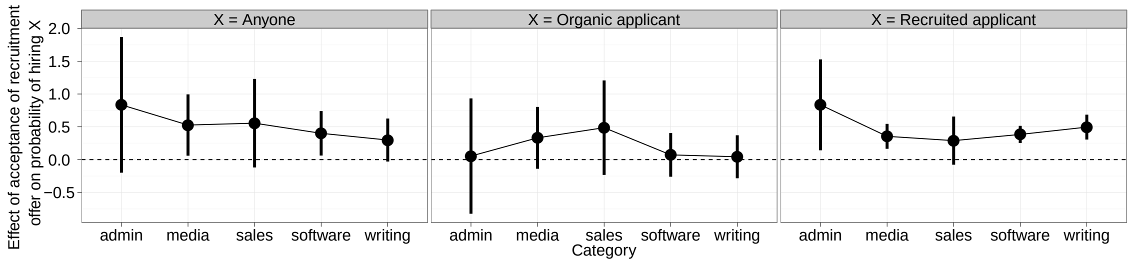
Given the size of D-RECRUIT-RESPONSE, we can estimate the effects of an accepted invitation by category. In Figure 8 we plot the coefficient on the endogeneous "Accept invite" regressor for the main works categories for the three outcome measures. In the left panel, we report the coefficient on following an IV regression when the outcome variable is whether the firm hired anyone. Across the categories, we can see the effect is positive and on the order of about 45%, which is close to our pooled estimate from Table 10. It is significant or nearly significant in all categories. The *admin* measure is far more imprecise since there are relatively few recruiting employers in that category. In the right panel, the outcome is whether the recruiting firm hired a recruited applicant. Unsurprisingly, the effect is positive ans significant across the categories. In the middle panel, the outcome variable is whether the firm hired an organic applicant. In all cases, the effect is positive but insignificant, consistent with our Table 10 estimates.

Table 10: Effects of positive invitation response on fill rates

	First Stage	Employer hires (IV):		
	(1)	(2)	(3)	(4)
	Accept Invite	Anyone	Organic Applicant	Recruited Applicant
Accept Invite		0.388***	0.156	0.346***
		(0.111)	(0.113)	(0.090)
Log applications	0.058***	0.026***	0.050***	0.002
	(0.001)	(0.007)	(0.007)	(0.005)
Invites (7d)	-0.010***	0.002	0.004*	-0.003*
	(0.000)	(0.001)	(0.002)	(0.001)
Invites (24h)	-0.018***			
	(0.002)			
Accepted invites (7d)	0.005***	-0.005***	-0.003**	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Accepted invites (24h)	0.019***			
	(0.003)			
Intercept	0.431***	0.273***	0.156**	0.048
	(0.016)	(0.051)	(0.052)	(0.041)
F-statistic (1st stage)		41.50	41.50	41.50
Sargan p-value		0.764	0.199	0.788
N	128,067	128,067	128,067	128,067

Notes: This table reports the results of three 2SLS regressions in which the outcome variables are whether or not the employer hired any at all (Column (2)), an organic application (Column (3)) and a recruited applicant (Column (4)). Column (1) contains part of the first stage regression (category and month fixed effects not shown), which is the same for all three regressions. The endogenous regressor is "Accept invite" which is an indicator for whether or not the invited candidate accepted the invitation. The data used for these regressions in D-RECRUIT-RESPONSE, described in Table 1. Each regression has two excluded instruments: the number of invitations and the number of accepted invitations the invited candidate received that same day from different employers. The regression includes controls for the category of work, month, contract structure, log number of organic applications to the vacancy and counts of the number of invitations received and accepted that same *week* by the invited candidate. For each regression, we report the F-statistic the first stage of the 2SLS, which is strong across the regressions. Because we have two instruments, we can also report the p-value for the Sargan test of the over-identifying restrictions. In all cases, we fail to reject the null hypothesis of valid instruments. Standard errors are robust to heteroscedasticity. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

Figure 8: IV strategy for estimating effect of recruited applicants on vacancy outcomes



Notes:: These three plots show the by-category estimated causal effect on hiring outcomes when a recruited worker accepts a recruitment invitation. The hiring outcomes are whether the firm hired anyone (left panel), an organic applicant (middle panel) or a recruited applicant (right panel). The dataset is D-RECRUIT-RESPONSE, described in Table 1. For each point estimate, a 95% confidence error is shown, based on standard errors robust to heteroscedasticity.

8 Discussion and conclusion

This paper demonstrates that recommendations are both (a) acted upon by employers and (b) effective at raising the match probability, at least for the skills-heavy categories of software development. Surprisingly, they seem to have little crowd-out effect and in some cases, seem to work as complements to organic, non-recruited applicants. The effectiveness of these recommendations is surprising given the information-rich context on online labor markets. In the oDesk marketplace, the universe of vacancies and job-seekers is fully indexed, searchable and described by rich standardized meta-data. Online labor markets have radically lower search costs compared to traditional labor markets, nevertheless, search costs still impede matching.

8.1 Digitization of the supply side of the labor market

Given the effectiveness of recommendations, it is puzzling that algorithmic matching assistance is not already commonplace in labor settings. One potential explanation is that the data needed for these kinds of recommendations (particularly on the supply-side of the market) were or are missing. Online job boards[5] have been suggesting vacancies to workers, but the boards have limited data to inform those recommendations. They can only condition their recommendations on whatever search terms and perhaps geographic and/or salary constraints a job-seeker

[5] Examples include Monster.com, CareerBuilder.com, Indeeed and Simply Hired. There are also a large number of more specialized job listing sites.

may enter in a relatively brief search session. Online job boards cannot condition their recommendations on a worker's employment history, educational background, skills, current employment status, professional connections, certifications, personality, test scores or other match-relevant factors for the simple reason that information was not available.

With the rise of professional networking sites such as LinkedIn, we are now witnessing the unprecedented data collection and digitization of the supply side of the labor market. Individuals can create public profiles and list their education, professional credentials, skills, current and past work experiences and, critically, their other professional connections. According to LinkedIn, as of March 12, 2012, over 160 million people had created profiles.[6] and in In some industries, a LinkedIn profile is expected of all applicants.[7] In online labor markets, the amount of information is even greater, as it includes detailed, verified and searchable information about wages, hours worked and project outcomes. oDesk is small compared to the traditional market and narrow—the only work that is possible is work that can be done online. However, we believe it approximates where traditional labor markets appear to be heading, at least in terms of both sides of the marketplace being digitized and described with rich, match-relevant meta-data. If this digitization and instrumentation continues to grow, algorithmic approaches to the labor market matching problem are likely to become commonplace.

8.2 Future work

There are a number of directions future work could take. One obvious challenge is to improve the quality of labor-related recommender systems. Better machine learning techniques and more and better data may improve recommendation quality. However, there are also fundamentally economic challenges, such as ensuring recommendations avoid congestion and enabling thickness and setting polices and mechanisms that are incentive compatible.

A key practical problem for a platform is deciding how to allocate visibility via recommendations and position in search, etc., to maximize some objective function. This function captures the platform's preferences over total matches formed, match quality, integration of new users and equity among platform participants. Currently, there is no satisfactory answer to this allocation question and research is needed. On the mechanism design side, the platforms would like workers to truthfully report their availability for work, yet this proves difficult in practice. The strategic issue is that workers have close to free disposal on job offers and hence

[6] Source: http://press.linkedin.com/about, accessed on July 20, 2012.

[7] One corporate recruiter—interviewed by the author—from a Silicon Valley tech start-up commented on the role of LinkedIn, reporting: "I'd say it is close to 100% (and certainly 100% for viable candidates). I can't think of an example of someone who I have screened who didn't have a profile on LinkedIn."

have little incentive to opt-out of being recommended, despite the externalities this imposes on other workers with actual availability to take on further projects. In the category of strategic issues, the model ignores how recruiting affects wage negotiations, and yet recruiting is clearly "type" revealing the model to a savvy recruited worker. If firms foresee this effect of recruiting, then the market may get an inefficient amount of employer recruiting. Incorporating wage-determination into the model would improve its realism and perhaps generate other testable predictions.

We show that vacancies receive all the applications they will ever get very quickly. A natural question is why that is the case and what effects this has on efficiency. On the one hand, vacancies can theoretically be filled quickly, but this also puts a great deal of pressure on workers and firms to find each other in a short amount of time. Under this compressed schedule, presumably match quality deteriorates, but the trade-off between matching speed and quality and the welfare implications is unknown. A satisfactory theory of how vacancies accrue applicants as a function of market policies (such as quotas) and parameters and these factors ultimately determine marketplace efficiency would potentially have great practical application.

References

Agrawal, Ajay, Nico Lacetera, and Elizabeth Lyons, "How Do Online Platforms Flatten Markets for Contract Labor?," *Working Paper*, 2012.

Andrews, M. J., Steve Bradley, D. Stott, and Richard Upward, "Successful Employer Search? An Empirical Analysis of Vacancy Duration Using Micro Data," *Economica*, 2008, *75* (299), 455–480.

Autor, David, "The Economics of Labor Market Intermediation: An Analytic Framework," 2008.

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

Barron, John and John Bishop, "Extensive Search, Intensive Search, and Hiring Costs: New Evidence on Employer Hiring Activity," *Economic Inquiry*, 1985, *23* (3), 363–82.

Barron, John M, 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, D., J. Kluve, and A. Weber, “Active Labour Market Policy Evaluations: A Meta-Analysis*,” *The Economic Journal*, 2010, *120* (548), F452–F477.

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

Gorter, C. and G.R.J. Kalb, “Estimating the effect of counseling and monitoring the unemployed using a job search model,” *Journal of Human Resources*, 1996, pp. 590–610.

Gurgand, B.C.E.D.M., R. Rathelot, and P. Zamora, “Do labor market policies have displacement effect? Evidence from a clustered randomized experiment,” 2011.

Halevy, A., P. Norvig, and F. Pereira, “The unreasonable effectiveness of data,” *Intelligent Systems, IEEE*, 2009, *24* (2), 8–12.

Hansen, L.P., “Large sample properties of generalized method of moments estimators,” *Econometrica: Journal of the Econometric Society*, 1982, pp. 1029–1054.

Horton, J., “Online labor markets,” *Internet and Network Economics*, 2010, pp. 515–522.

Kluve, J., “The effectiveness of European active labor market programs,” *Labour Economics*, 2010, *17* (6), 904–918.

Luca, M., “Reviews, reputation, and revenue: The case of Yelp. com,” *Harvard Business School Working paper 12*, 2011, *16*, 12–016.

Niederle, Muriel, Alvin E. Roth, and Tayfun Sönmez, “Matching and Market Design,” *The New Palgrave Dictionary of Economics. Palgrave Macmillan, Basingstoke*, 2008.

Ommeren, Jos Van and Giovanni Russo, “Firm Recruitment Behaviour: Sequential or Non-Sequential Search?,” *SSRN eLibrary*, 2010.

Oyer, Paul and Scott Schaefer, “Personnel Economics: Hiring and Incentives,” May 2010, (15977).

Pallais, A., “Inefficient hiring in entry-level labor markets,” Technical Report, V Working paper 2010.

Pellizzari, M., "Employers' Search and the Efficiency of Matching," *British Journal of Industrial Relations*, 2011, *49* (1), 25–53.

Petrongolo, B. and C.A. Pissarides, "Looking into the black box: A survey of the matching function," *Journal of Economic Literature*, 2001, *39* (2), 390–431.

Rees, A., "Information networks in labor markets," *The American Economic Review*, 1966, *56* (1/2), 559–566.

Resnick, P., K. Kuwabara, R. Zeckhauser, and E. Friedman, "Reputation systems," *Communications of the ACM*, 2000, *43* (12), 45–48.

Roth, Alvin E., "What Have We Learned from Market Design?," *The Economic Journal*, 2008, *118* (527), 285–310.

Stanton, C. and C. Thomas, "Landing the First Job: The Value of Intermediaries in Online Hiring," 2011.

Stigler, G.J., "The economics of information," *The Journal of Political Economy*, 1961, *69* (3), 213–225.

van Ours, Jan and Geert Ridder, "Vacancies and the Recruitment of New Employees," *Journal of Labor Economics*, 1992, *10* (2), 138–55.

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

A Evidence of batch processing by employers

Assumption 1 is that employers process applications in batch. By assuming batch processing rather than serial processing, the employer's decision problem becomes determining a sample size. With serial processing, if screening is simply a Bernoulli trial, then there is no firm decision problem: the firm just hires the first screened worker that proves to be a match. The batch versus serial is presumably a false dichotomy, with real firms doing both or some amalgam of the strategy (e.g., screen a series of micro-batches). However, at least within the oDesk marketplace, batch processing seems to predominate.

We present two pieces of evidence that employers process applications in batch. First, we show that the precise order that applicants "arrive" to a vacancy is irrelevant to their probability of being hired, conditional upon arriving before the last applicant to be viewed. Second, we show that the applicant arriving immediately after the ultimately hired applicant is no less likely to be viewed than the applicant arriving immediately before. We restrict our attention to vacancies were the employer (a) made one and only one hire against that vacancy (b) there was at least one applicant that arrived before the ultimately hired applicant and at least one applicant that arrived immediately after that same hired applicant and (c) the vacancy attracted more than 5 but fewer than 25 organic applicants.

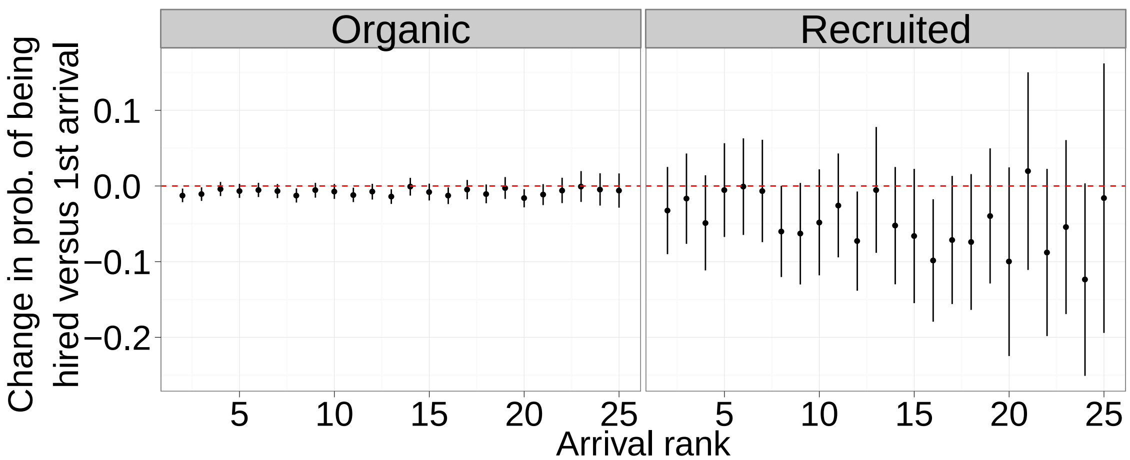
A.1 Arrival rank does not predict of hiring

For the first method, the observations are all of the individual applications to the vacancy. The outcome variable is a binary indicator for whether or not that particular application lead to a hire. We include a vacancy-specific fixed-effect l_j and we include an indicator for the rank of the ith applicant. We cluster standard errors at the vacancy level. The regression we estimate is:

$$\text{hired}_{ij} = \sum_{k=1}^{C} \beta_k \cdot 1\{rank(i) = k\} + l_j + \epsilon_{ij} \tag{10}$$

Figure 9, we plot $\beta - \hat{\beta}_k$ for $k = 2 \ldots 16$. We can see that none of the coefficients are distinguishable from zero.

Figure 9: Estimated probability of being hired as function of arrival rank



Notes: These plots shows the estimated $\hat{\beta}_k$ coefficients from Equation 10 and a 95% confidence interval for the point estimate. The dependent variable is whether or notan applicant in that arrival position was hired; the plotted estimates are the coefficients on a collection of indicators for the precise arrival rank of each applicant. The data is based upon B-APPLICATIONS(described in Table 1), but is restricted to those vacancies that received less than 25 applications and more than 1 application. These coefficients are mean difference in probability of being hired compared to the first applicant; k indexes their arrival order. The regression estimating these equations includes a vacancy-specific fixed effect and standard errors are clustered at the vacancy level. The left panel shows the estimates for organic applicants, while the right panel shows the estimates for recruited applicants.

Table 11: Employer viewing of applications by arrival position relative to eventually-hired applicant

	Employer viewed application?	
	(1)	(2)
	Yes	Yes
Intercept	0.862***	0.849***
	(0.008)	(0.009)
Application after hired worker	-0.024	-0.000
	(0.015)	(0.012)
Vacancy-Specific FE	Yes	No
N	3,638	3,638
R-squared	0.799	2.58e-09

Notes: This table reports the results of regressions in which the dependent variable is whether or not an employer viewed an applicant's application. Here "viewed" means opening up the application in the web interface, similar to how one would open an email. The dataset is selected sample of dataset B-APPLICATIONSfrom Table 1, consisting only of those vacancies where one and only one applicant was hired and an applicant arrived immediately before and immediately after the eventually hired applicant. The dataset consists of the applications immediately before and immediately after the hired applicant. The important independent variable is an indicator for the "after" application. In Column (1), we include a vacancy-specific effect, while in Column (2) we perform OLS. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

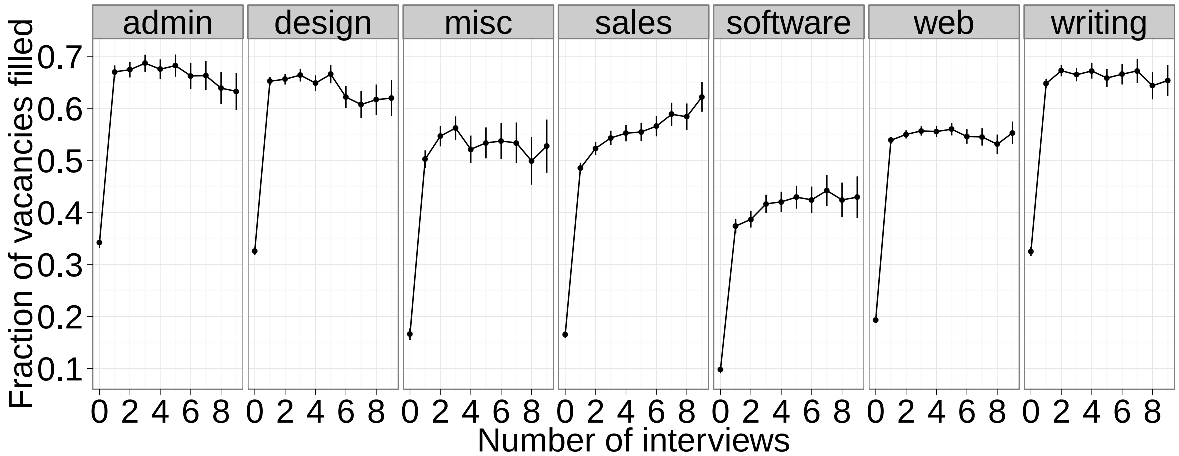
A.2 Applicant arriving after eventually-hired applicant no less likely to be viewed by employer

For the second method, we restrict out sample to cases where there was at least one applicant that arrived before the ultimately hired applicant and at least one applicant that arrived immediately after that same hired applicant. As such, we have two observations per vacancy. We then estimate a regression where the outcome variable is whether or not the employer viewed that particular equation; the independent variable is dummy variable after_{ij}, which indicates that an applicant arrived immediately after the hired applicant. We also include a vacancy-specific fixed effect. The regression equation is

$$\text{viewed}_{ij} = \beta \cdot \text{after}_{ij} + l_j + \epsilon_{ij} \tag{11}$$

In Table 11, Column (1), we report the estimate of Equation 11. The coefficient on after_{ij} is only slightly negative and the 95% confidence interval readily includes 0, implying no difference

Figure 10: Fill probability versus number of interviews



Notes: This panel plots the estimated conditional probability of a vacancy being filled versus the number of interviewed applicants using dataset C-VACANCIES-OBS, described in Table 1. A 95% confidence interval is shown for each point estimate.

in viewing rates between the before and after application. If firms process applications serially, hiring the first applicant exceeding some threshold, then we should (counter-factually) observe no applications being processed *after* the eventually hired applicant applies.

B The shape of the hiring function

Proposition 1 & 2 make predictions about the shape of the hiring function—namely that the function is increasing and concave in a. This hiring function is vacancy specific: it depends on the distribution of match probabilities for that employer's particular vacancy as well as other factors like the number of applications received and whether or not the employer recruited. We obviously do not get to observe the full hiring function for any particular vacancy—at best, we observe a single point on that curve and that point is clearly endogeneously determined. However, it is still potentially informative to see the empirical relationship between number of applications screened by an employer and the probability of filling their vacancy. In Figure 10, we plot the estimated probability of fill as a function of the applicants screened for a large number of vacancies posted on oDesk in the last year.

In Figure 10, several patterns are clear. First, the large increase in fill probability goes from interviewing 0 to interviewing 1 applicant. Additional applicants seem to be associated with higher fill rates, though no additional screening has the same 0-1 effect. One interesting feature is that a fair amount of hiring occurs without any interviews, contra our assumption that $h(0) = 0$. This is most likely because when workers apply, they are making a binding wage offer and for more experienced workers, there is often sufficient information available in their application to justify hiring without a formal interview. In a sense, the formal interview that we capture in the data is probably an underestimate of true evaluations in the sense of the model.

C Effects of project value and labor market tightness

The model makes several predictions about how project value and labor market tightness affect recruiting, screening and hiring. A high project value (larger v) implies more screening and more recruiting (Propositions 9 and 8). A tight labor market (smaller A_p) predicts more recruiting (Proposition 12), though the effect on screening depends on additional factors (Proposition **??**). These are central predictions of the model and yet we have no direct exogeneous variation in either v or A_P. In fact, we cannot even directly observe the value an employer would receive from successfully filling a vacancy, though there are several measurable attributes that are likely to proxy for value.

Different projects presumably vary in their value and their ability attract applicants because of both the nature of the required work and the identity of the employer. Variation driven by the nature of the work is variation we are interested in; variation due to employer attributes is what we would like to net out. In an ideal experiment, we would randomly assign heterogeneous projects to employers and observe outcomes. Of course, employers are not randomly assigned projects. However, many employers used oDesk repeatedly, often for very different kinds of work. We can use this source of variation to estimate the relationship between project characteristics and employer decisions and outcomes. This still is not random assignment and the estimates are not causal, but this approach allows us to net out constant employer-specific effects. A fixed-effects regression will expose the model—however imperfectly—to at least some empirical scrutiny.

To estimate the fixed-effects regression, we construct a dataset C-VACANCIES-OBS consisting of a large number of vacancies posted on oDesk. The dataset is described in Table 1. We estimate Equation 12 where y_{ij} is an indicator for various outcomes, which we will describe. The independent variables are the number of organic applicants, A_P, the number of estimates project hours, H_{ij} and, in some regressions where we focus only on filled vacancies, the realized total wage bill for the project, V_{ij}. We also include employer-specific fixed effects, δ_j, and fixed-effects for the category of work, γ_{ij}.

$$y_{ij} = \beta_A \log A_P^{ij} + \beta_H \log H_{ij} + \beta_V \log V_{ij} + \delta^j + \gamma_{ij} + \epsilon_{ij} \tag{12}$$

C.1 Effects on recruiting and screening

In Table 12, we report the results of estimating Equation 12 where the outcome y_{ij} is whether the firm chooses to recruit, $1 \cdot \{x_{ij} = R\}$. Column (1) reports the OLS estimate without employer-specific fixed effects. In Column (2) we add the employer-specific fixed effect and in Column (3) we restrict the dataset to only those vacancies receiving five or more organic applicants. In Column (4) we estimate the model using only filled vacancies, which allows us to include the log total wage bill and estimate β_V.

In every specification with employer-specific fixed-effects, recruiting is increasing in the proxies for project value. The log hours estimate regressor is positive and significantly in all regressions except the Column (4) (though the magnitude of the point estimate is largest in this regression). The actual wage bill regressor, β_Y is strongly significant and positive. In Column (1), the sign on estimated hours is negative, implying that in the cross-section, employers stating long-duration projects are *less* likely to recruit.

The sign reversal of the coefficient on the log estimated hours in Column (1) could potentially reflect some fraction of employers strategically manipulating their hours estimate when posting vacancies. Employers stating a duration have an incentive to imply that projects will last for a long time, as this might induce more relationship-specific investment from workers, who think they are working in the shadow a potentially longer-term relationship.

In the Column (4) regression, we include only filled vacancies. In this regression, the coefficient on the total wage bill for the eventual project is positive and highly significant. Of course, the direction of causality is particularly problematic in this regression, given that the actual wage bill is realized *after* the employer has made the recruiting decision, which presumably affects the wage bill. However, if we think of the market as competitive and this wage bill as in some sense exogenous, then this regression is more reasonable.

In every specification, the log number of organic applicants is negatively correlated with employer recruiting. Comparing Columns (2) and (3), restricting the vacancies to those that receive five or more organic applications does not seem to dramatically change the coefficient. The absence of a drop-off suggests that even with a fairly larger number of applicants, the returns to a marginal applicant are still high.

Table 12: Project value, organic applications, wage bill and the probability of recruiting

	Employer sent early invitations?			
	All vacancies			Filled vacancies only
	(1)	(2)	(3)	(4)
	Yes	Yes	Yes	Yes
Log estimated hours	-0.0073***	0.0027**	0.0032*	0.0049
	(0.0005)	(0.0010)	(0.0013)	(0.0031)
Log number of organic applications	-0.0227***	-0.0103***	-0.0088**	-0.0119*
	(0.0010)	(0.0017)	(0.0033)	(0.0051)
Log dollars spent				0.0073**
				(0.0024)
Employer-Specific FE	No	Yes	Yes	Yes
N	210,678	210,678	148,081	50,319
R-squared	0.00843	0.567	0.602	0.717

Notes: In the regressions reported in this table, the dependent variable is whether the employer recruited candidates to their vacancy. In each regression, we include the log of the project size (as estimated by the employer at the time of posting) and the log of the number of organic applications received. The dataset used for the regressions is C-VACANCIES-OBS, described in Table 1, but restricted to vacancies that receive at least one organic application. Column (1) is an OLS regression; the remaining columns all include an employer-specific fixed effect. In the employers-specific fixed effect regressions, standard errors are clustered at the employer level. All other standard errors are robust to heteroscedasticity. In Column (2) the number of organic applicants is one or more. In Column (2) the sample is restricted to vacancies with five or more organic applicants. In Columns (1), (2) and (3), the sample is all vacancies; in Column (4), we restrict our attention to only those vacancies wherean applicant was hired and money was spent. This allows us to include the ultimate wage bill for the project as a regressor. A caveat is that this wage bill measure can obviously be affected by the recruiting choice. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

Table 13: Project value, organic applications, wage bill and the number of screened applicants

	Number of screened applicants			
	All vacancies			Filled vacancies only
	(1)	(2)	(3)	(4)
Log estimated hours	0.1707***	0.0869***	0.1062***	0.1805***
	(0.0068)	(0.0134)	(0.0191)	(0.0422)
Log number of organic applications	1.3835***	1.6248***	2.2045***	1.4758***
	(0.0135)	(0.0438)	(0.0944)	(0.0819)
Log dollars spent				0.0911***
				(0.0252)
Employer-Specific FE	No	Yes	Yes	Yes
N	210,678	210,678	148,081	50,319
R-squared	0.0890	0.633	0.677	0.699

Notes: In the regressions reported in this table, the dependent variable is the number of interviews conducted by an employer. In each regression, we include the log of the project size (as estimated by the employer at the time of posting) and the log of the number of organic applications received. The dataset used for the regressions is C-VACANCIES-OBS, described in Table 1, but restricted to vacancies receiving at least one organic applicant. Column (1) is an OLS regression; the remaining columns all include an employer specific fixed effect. In the employers-specific fixed effect regressions, standard errors are clustered at the employer level. All other standard errors are robust to heteroscedasticity. In Column (2) the number of organic applications is one or more; In Column (2) the sample is restricted to vacancies with five or more organic applications. In Columns (1), (2) and (3), the sample is all vacancies; in Column (4), we restrict our attention to only those vacancies where an applicant was hired and money was spent. This allows us to include the ultimate wage bill for the project as a regressor. A caveat is that this wage bill measure can obviously be affected by the recruiting choice. Significance indicators: $p \le 0.05 : *$, $.p \le 0.01 : **$ and $p \le .001 : ***$.

C.1.1 Screening

In Table 13, we regress the number of screened *organic* applicants on the same set of regressors used in Equation 12. We use this measure because recruited applicants are automatically classified as being interviewed. Across all specifications, we can see that the number of screenings is strongly increasing in proxies for project value and the number of organic applications, as the model predicts. In terms of Proposition **??**, the results imply that of the two effects of more applicants—greater per-applicant quality but greater likelihood of finding a match fewer applicants—the “substitution” effect dominates: when more organic applicants are available, employer screens more applicants.

Table 14: Organic applications and the vacancy fill rate

	Vacancy filled?		
	(1)	(2)	(3)
Log estimated hours	-0.0327***	-0.0178***	-0.0186***
	(0.0005)	(0.0013)	(0.0016)
Log number of organic applications	0.0659***	0.0462***	0.0597***
	(0.0011)	(0.0020)	(0.0040)
Employer-Specific FE	No	Yes	Yes
N	210,678	210,678	148,081
R-squared	0.0522	0.538	0.568

Notes: In the regressions reported in this table, the dependent variable is an indicator for whether or not the firm hired any anyone. In each regression, we include the log of the project size (as estimated by the employer at the time of posting) and the log of the number of organic applications received. The dataset used for the regressions is C-VACANCIES-OBS, described in Table 1, but restricted to vacancies receiving at least one organic applicant. Column (1) is an OLS regression; the remaining columns all include an employer specific fixed effect. In the employers-specific fixed effect regressions, standard errors are clustered at the employer level. In Column (2) the number of organic applicants is 1 or more; In Column (2) the sample is restricted to vacancies with 5 or more organic applicants. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

C.1.2 Hires

In Table 14, we report regressions where the outcome variable is whether the employer filled their vacancy. Unlike in the previous regressions of this type, we cannot examine the effect of wage bill on fill rates because of the mechanical relationship between the two. Across specifications, the fill probability is decreasing in the size of the project (as measured by estimated hours) but increasing in the number of organic applications. The negative coefficient on the estimated hours suggests that if the model is correct, this is not a *ceteris paribus* estimate of project value on fill rates. Presumably higher-value projects might also be harder to fill, i.e., match probabilities and project value are negatively correlated. This result highlights one of the limitations of this approach: even though we are able to control for employer-specific effects, we do not actually obtain exogeneous variation in the different independent variables.

D Market design considerations

Earlier, we compared hired wage rate the feedback across the experimental groups to asses the quality of the induced matches, finding no significant difference. Because invited worker are not compelled to accept recruiting invitations, one measure of recommendation quality is invitation response rate. If the platform is making bad recommendations, i.e., only superficially relevant or not conditioned on the worker's availability, then we should see a lower invitation response rate to treatment-induced recommendations. In Table 15, Column (1), we regress an employer's invitation response rates to an early invitation, *conditional* upon sending any invitations at all. As we know that the treatment increases the number of early invitations sent, this regression is merely summarizing the data and does not have a straightforward causal interpretation. In Column (1), we can see that average response rate is about 4% points lower in the treatment group that in the control. This difference is almost significant at the 5% level.

There are several possible explanations. Probably the simplest is that the effect is due to chance, especially since it is not even significant at the conventional level. Another is that the algorithmically-selected candidates have lower match-specific productivity than the candidates that the employers would have selected themselves. A more prosaic explanation is that the same workers were recommended multiple times and that the experiment simply exhausted their availability.

Although the rank-ordering of workers on oDesk is dynamic, the re-ordering is not dramatic over short periods of time. For vacancies with common skills, the same workers could have been recommended multiple times. Given that workers are inherently supply-constrained, these excess invitations were not valuable and thus more frequently turned down. We can partially test this idea by seeing if the response rate in the treatment group declines over time. In Column (2) of Table 15, we regress the response rate on the treatment, but we also interact the treatment indicator with the date the vacancy was posted (normalized to be in $[0,1]$). With this interaction, the coefficient on the treatment becomes positive (though not significant) and the coefficient on the treatment and date interaction is negative, consistent with an "exhaustion" explanation.

While it does not appear that the treatment-induced invitations were adversely selected in terms of response rate, we would like to know if the response rate is in fact sensitive to any economically meaningful factors. Although it is merely observational/correlational, we can check whether employers who send many (and presumably less discriminating) invitations experienced a lower response rate. In Column (3), we add a regressor that is the number of early invitations sent by an employer. We can see that positive response rate is strongly negatively correlated with the number of early invitations.

Table 15: The response rate to employer "early invite" recruitment by treatment group

	Fraction on recruited workers accepting invitation:		
	(1)	(2)	(3)
Intercept	0.509^{***}	0.505^{***}	0.554^{***}
	(0.015)	(0.024)	(0.017)
Treatment	−0.036	0.012	−0.037
	(0.020)	(0.032)	(0.020)
Date posted (normalized)		0.012	
		(0.052)	
Treatment × Date posted		−0.136	
		(0.071)	
Number of early invites			-0.013^{***}
			(0.002)
N	1,733	1,733	1,733
R-squared	0.002	0.006	0.024

Notes: In this table, the outcome variable in all cases is the fraction of invited candidates that actually applied to the employer's vacancy. "Date posted" is the elapsed time from when a vacancy was posted and when the experiment began a variable normalized to lie in $[0,1]$. For example, a vacancy posted on the first day of the experiment would have a value of 0, and vacancy posted on the last day would have a value of 1. In Column (3), we include the *count* of early invites as a predictor. The standard errors robust to heteroscedasticity. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

D.1 Do invitations crowd-out organic applications?

In Sections 6 and 7, we analyzed whether or not the treatment crowded out organic *hires*. We found that they did not and, if anything, the treatment seem to stimulate the hiring of organic applicants in some categories. Here we examine whether the treatment had any affect on the number of organic applications *received*. Aside from our market design interest in this question, it is also relevant to the model. Assumption 3 is that the quantity of organic applicants, A_P does not depend on whether or not the employer recruits.

In traditional markets, it seems unlikely that workers have a very precise estimate of the size of the competition generally. On oDesk and in many other online labor markets, workers have perfect, real-time information about the size of the queue and the employer's recruiting efforts. Assuming workers do observe employer recruiting behavior, the effect on the decision-making of would-be applicants is theoretically ambiguous. On the one hand, more applicants—recruited or not—means more competition. On the other hand, a recruiting employer is an employer with a high-value project (assuming the workers have at least some bargaining power).[8]

In Table 16, we investigate whether employer recruiting affected would-be organic applicant decision-making. As our outcome variable, we use an indicator for whether count of organic applicants exceeded the median number of applications for vacancies in that category. We use this indicator because the number of organic applicants is highly skewed (it appears log-normal conditional upon being positive) but also highly category-dependent. In Column (1), we present a descriptive regression where we regress the median-clearing outcome on the count of early invites. The coefficient in positive but small and insignificant.

In Column (2), we use the number of *accepted* early invites, which is *positively* correlated with the number of organic applications beating the median, contra a crowd-out story. Of course, the number of early invites and the number of accepted early invites are endogenously determined. In Column (3), we use the treatment indicator—which we know affects both early invites and accepted early invites—as our predictor. The effect is negative but very small and insignificant. Another approach is to use the treatment indicator as an instrument for accepted early invites. Doing this in Column (4), we find that the local average treatment effect for this group is negative and large—consistent with crowd-out—but the effect is highly imprecise and the 95% confidence interval comfortably includes zero. With the data available, we cannot conclude anything more definitive on the question of whether invited applicants crowd out organic

[8] Although the simple model of employer recruiting had no real strategic component vis-a-vis wages, it is easy to imagine the direction it would take if workers could condition their wage offer on the employer's recruiting choices—there would a slice of employers that would find it attractive to recruit but would be induced not to do so by the threat of worker wage demands in a separating equilibrium.

Table 16: Effect of treatment on application counts

	Organic applicant count > median # for that category?:		
	(1)	(2)	(3)
Intercept	0.344***	0.342***	0.348***
	(0.005)	(0.005)	(0.006)
Early invites	0.002		
	(0.002)		
Accepted Early invites		0.013**	
		(0.005)	
Treatment			−0.006
			(0.009)
N	11,414	11,414	11,414
R-squared	0.000	0.001	0.000

Notes: Each regression reported in this table has the same outcome variable: an indicator for whether or not the number of organic applicants to that vacancy exceeded the median number of organic applications for that category. "Early invites" is the count of invitations sent out by the employer. "Accepted Early Invites" is the number of invitations that received a positive response from the recipient. Columns (1) through (3) are OLS, while Column (4) is a 2SLS regression using treatment assignment as an instrument for the number of early invites. The standard errors are robust to heteroscedasticity. Significance indicators: $p \le 0.05 : *$, $.p \le 0.01 : **$ and $p \le .001 : ***$.

applicants.