# Computer-Mediated Matchmaking: Facilitating Employer Search and Screening*

John J. Horton
oDesk Research & Harvard University

November 2, 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 as more markets become computer-mediated. This approach to reducing friction raises questions: When are recommendations effective? 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% among technical (e.g., computer programming) vacancies but had no effect on non-technical vacancies. This heterogeneity was likely caused by higher screening costs (which we estimate with a structural model of employer screening) and tighter markets for technical vacancies. Where fill rates did increase, however, it was only partly because employers acted upon recommendations: the treatment also increased the 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. An instrumental variables analysis of the larger marketplace confirms both the positive effect of recruited applicants on fill rates and the absence of crowd-out. Together, these results imply that, despite their smaller size, search costs do impede matching in computer-mediated markets, but they can be reduced through informational interventions. Furthermore, despite explicit promotions of certain workers over others, in some cases recommendations can improve marketplace efficiency without making anyone worse off.

JEL Codes: C93, J64

Keywords: employer search, labor market intermediation, field experiments

*All code and data is or will be available at http://john-joseph-horton.com. Acknowledgments are at the end of the paper.

# 1 Introduction

Firms and workers trying to form new matches must perform a variety of search and screening tasks. These tasks include finding each other, learning the attributes of would-be partners and forecasting the returns to a putative match. All of these tasks are information processing tasks that information technology (IT) can, in principle, make radically cheaper and easier to perform. However, using IT in this way requires high-quality, timely data about the state of the market and its participants. In the past, this kind of information was neither digitized nor centrally collected, but in just the last few years, conditions have dramatically changed. Now, an enormous amount of match-relevant labor market data is being digitized, with much of it held by private labor market intermediaries. Examples of intermediaries include online job boards such as Monster.com, SimplyHired, The Ladders and CareerBuilder; professionally-oriented social networks such as LinkedIn (which alone has more than 175 million users); and online labor markets such as oDesk, Elance and Freelancer.

The intermediaries can use this labor market data to create informational services and offer them to recruiters and/or job-seekers. In this paper, we explore the effects of perhaps the most ambitious and demanding form of an informational service: the direct recommendation of potential matches. Our setting is oDesk, the largest and fastest growing online labor market. Our main empirical results come from a randomized field experiment in which treated employers received algorithmically generated, by-name recommendations of candidates for their vacancies. The candidates were chosen to be relevant, high-ability and available to work. In this market, we as researchers observe the posting of a vacancy, the search efforts of the employer, the arrival of applicants, the screening of applicants and the making of job offers. This unprecedented ability to peek inside the firm's in situ environment and examine its decision making allows us to get much closer to identifying the causal mechanisms for the effects that we do find.

The analysis focuses on two broad questions: (a) are the recommendations effective, as measured by the vacancy fill rate, and (b) to what extent, if any, do recommendations crowd out organic matches? In sum, we want to know whether we made the pie larger and in doing so whether or not we shrunk anyone's slice.

To interpret our empirical findings, we develop a model of employer search and screening. In the model, a firm seeks to hire a single worker for a vacancy and obtains applicants in two ways: applicants either apply organically, which costs the firm nothing, or they are actively recruited by the firm. Recruiting is costly, but it increases the quantity and average quality of applicants. Firms pay a per-applicant screening cost, which is the same for all applicants. This screening tells the firm if an applicant is a match for their vacancy and is done in "batch" (as opposed to serially), with the firm choosing the batch size ex ante. If one or more screened applicants are a match, then the vacancy is filled. The firm must decide whether to recruit and, conditional upon that choice, how many applicants to screen. The model shows how reducing employer search costs via recommendations can increase fill rates and how this increase is mediated by the firm's screening costs, the quality of the recommended candidates and the tightness of the labor market.

### 1.1 Overview of the paper and key results

Section 2 describes the oDesk marketplace and Section 3 explains the details of the experiment. Section 4 reviews the employer search literature and compares it with our approach and empirical setting. Our model is developed in Section 5 and the empirical analysis is presented in two sections: Section 6 reports the experimental results, while Section 7 reports the quasi-experimental results. The results are summarized in Table 9, which contains reference to the associated model assumptions and predictions.

We find that providing algorithmically generated recommendations increased employer recruiting, albeit by a smaller amount among non-technical vacancies. This recruiting raised the fill rate substantially, but only for technical vacancies. In terms of the productivity of matches formed, we find no detectable differences across experimental groups.

The heterogeneity in treatment effects on hiring requires an explanation. One simple explanation is that technical vacancies receive fewer organic applicants, which makes recruiting more effective. Our model offers a potential theoretical explanation, which is that vacancies with higher associated screening costs in the treatment group should experience larger jumps in fill rates. A structural estimation of the employer search and screening model suggests that technical categories do in fact have higher screening costs.

For technical vacancies, the treatment increased the hiring of both recruited *and* organic applicants. Although this complementarity is surprising, it is not contra the predictions of the model: if the number and quality of recruited candidates is sufficiently high, the employer screens more candidates. By looking directly at how employers process their applicant pools, we find evidence for precisely this kind of screening "spillover."

In the experiment, employers were free to disregard the recommendations and candidates were free to ignore employer invitations. As a result, our estimate of the effect of recruited candidates on the vacancy fill rate comes from a rather indirect process. However, by using non-experimental data from the larger marketplace, we can estimate the direct effects of recruited applicants on both aggregate fill rates and crowd-out using a quasi-experimental design. We perform this analysis by exploiting idiosyncratic factors that determine whether a recruited worker applies for a vacancy but which are not conditionally dependent upon worker/employer attributes. In Section 7 we report an instrumental variables analysis in which we find that having more recruited applicants available causes an increase in the fill rate. Less expected but consistent with our experimental findings, this acceptance does *not* decrease the probability that an organic applicant is hired. This implies that crowd-out is a secondary concern, which in turn suggests that 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 B 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.

Traditional markets are different from oDesk in many ways; one of main ways they differ is that search costs are much higher. In the oDesk marketplace, the universe of vacancies and job-seekers is fully indexed, searchable and described by rich, standardized meta-data. Nevertheless, search and screening costs were sufficiently high that algorithmic matching assistance improved market efficiency. Furthermore, it improved efficiency in the high-value segment of the market where parties have greater incentives to search and screen. For this reason, it is possible that the results of our experiment underestimate the results that could be obtained by a similar intervention—were it possible—in a traditional labor market.

## 2 The oDesk marketplace

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

Horton (2010) presents models on online labor market pricing, optimal provision of user services, market specialization and competitive equilibrium between platforms.

Our experiment was conducted on oDesk, the largest of these online labor markets. On oDesk, would-be employers write job descriptions, self-categorize the nature of the work and required skills and then post the vacancies to the oDesk website. Workers learn about vacancies via electronic search or emailed notification. When a worker finds a vacancy 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 workers to apply. These invitations—which we use as a measure of recruiting—simply alert workers that an employer is interested in them. Workers can either ignore recruiting invitations or apply for the vacancies. After a worker submits an application, the employer can interview and hire the applicant on the terms proposed by the worker or make a counteroffer, which the worker can counter, and so on.

To work on hourly oDesk contracts, contractors must install custom tracking software on their computers. The tracking software, or "Work Diary," essentially serves as a digital punch clock. When the contractor is working, the software logs the count of keystrokes and mouse movements; at random intervals, the software also captures an image of the worker's computer screen. All of this captured data is sent to the oDesk servers and then made available to the employer for inspection. The advantage to the contractor of this monitoring system is that so long as they use the tracking software to log hours, they are guaranteed payment by the employer. If the employer is unhappy with the quality or pace of the contractor's work, it is incumbent upon them to offer corrective instruction or terminate the contract. This monitoring makes hourly contracts and hence employment relationships possible, which in turn makes the oDesk marketplace more like a traditional labor market than project-based online marketplaces.

oDesk itself is a private company with headquarters in Redwood City, California. It started as an enterprise software company in 2003 and launched an online marketplace in 2005. The company has received $44 million in capital investment and has approximately 100 full-time employees which are augmented by a large number of contractors hired through the marketplace. In the first quarter of 2012, $78 million was spent on oDesk. The 2011 wage bill was $225 million, representing 90% year-on-year growth from 2010. As of October 2012, more than 495,000 employers and 2.5 million contractors have created profiles (though a considerably smaller fraction are active on the site). About 790,000 vacancies were posted in the first half of 2012. The top four employer countries by headcount are the United States, Australia, Canada and the United Kingdom, and the top contractor countries by headcount are India, the Philippines, the United States and Ukraine.

Based on dollars spent, the top skills in the marketplace are web programming, mobile applications development (e.g., iPhone and Android) and web design. Based on hours worked, the top skills are web programming, data entry, search engine optimization and web research. The difference in the top skills based on dollars versus hours reflects a fundamental split in the marketplace between technical and non-technical work. In our analysis, we frequently draw a distinction between vacancies depending on the type of work. When oDesk started, it was a marketplace hiring software developers. Over time, as new countries were added and new categories created, employers began posting non-programming jobs such as data entry, customer service support, writing and design. There are highly-skilled, highly-paid contractors working in non-technical jobs, yet a stylized fact of the marketplace is that technical work tends to pay better, generate longer-lasting relationships and require greater skill.

While some oDesk employers are large (e.g., AOL, Dun & Bradstreet), a survey of oDesk employers by an independent market research firm finds that more than half of all employers consider themselves "start-ups." Their reasons for using oDesk are varied, but the main reasons appear to be the lower cost compared to traditional employment (both variable and fixed) and the ability to obtain specialized labor in small chunks. In a survey of employers, 3 out of 4 report online hiring being a "long-term strategy" rather than a stop-gap measure.

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 other online labor market sites *and* in the traditional market. However, survey evidence 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% of employers said that they listed their last project on oDesk alone. While this is 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 discussing the equity and efficiency implications of the experiment.

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 and break into the marketplace. Agrawal et al. (2012) investigate what factors matter to employers in making selections from an applicant pool and present some evidence of statistical discrimination.

# 3 Description of the experiment

In June 2011, oDesk launched an experimental feature designed to help employers fill their vacancies. After posting a vacancy, an employer was sent approximately six recommendations of relevant, available and high-ability contractors. Only new employers were eligible. After the experiment showed promising results, it became the default experience for all employers. The dataset from this experiment contains 11,414 vacancies, which in turn generated 358,324 organic and recruited applications.

The flow of subjects through the experiment—from initial posting to fulfillment or vacancy abandonment—is show in Figure 1. Note that allocation occurred immediately after the employer posted a vacancy. In the treatment group, employers were shown a pop-up window with some number of recommended candidates. Figure 2 is a screen shot of the actual recommendation interface. Recommended candidates were chosen algorithmically based upon their inferred relevance, ability and availability. Relevance was measured by the degree of overlap in the skills required for the vacancy and the skills listed by the worker in their profile. Ability was defined as a weighted sum of skill test scores, feedback ratings and past earnings. Availability was inferred from signals such as a worker recently ending a project or applying to vacancies. If the algorithm failed to find a suitable candidate, no recommendations were made.

After treated employers were presented the pop-up window, they could view each recommended worker's photograph, listed skills, average feedback score and stated hourly wage. If the employer clicked on a worker's "tile" they could see the worker's country, total hours worked on the platform, portfolio size, passed skills tests and 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. Because randomization occurred after the employer established the vacancy attributes (e.g., category, job description, required skills), vacancy attributes are independent of 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 vacancies. If an employer invited a candidate—regardless of whether the invitation was made through the "normal" channel or the experimental channel—the candidate would receive an email notification with a link to the employer's vacancy posting. The recruited candidate could then decide whether or not to apply.

If the employer posted the job publicly, other non-recruited workers were free to find the vacancy and apply. These organic applicants joined 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, they received a combination of recruited and organic applicants. If the employer did not recruit, they only received organic applicants.

An empirical fact of the marketplace is that recruited candidates are far more likely to be hired than organic candidates. In the experiment, pooling across groups, the raw per-application probability of hire for a recruited applicant was 0.08. The per-application probability of hire for an organic applicant was only 0.02. This 4x increase in hiring for recruited versus organic applicants motivates some of our modeling decisions, which we discuss in Section 5.

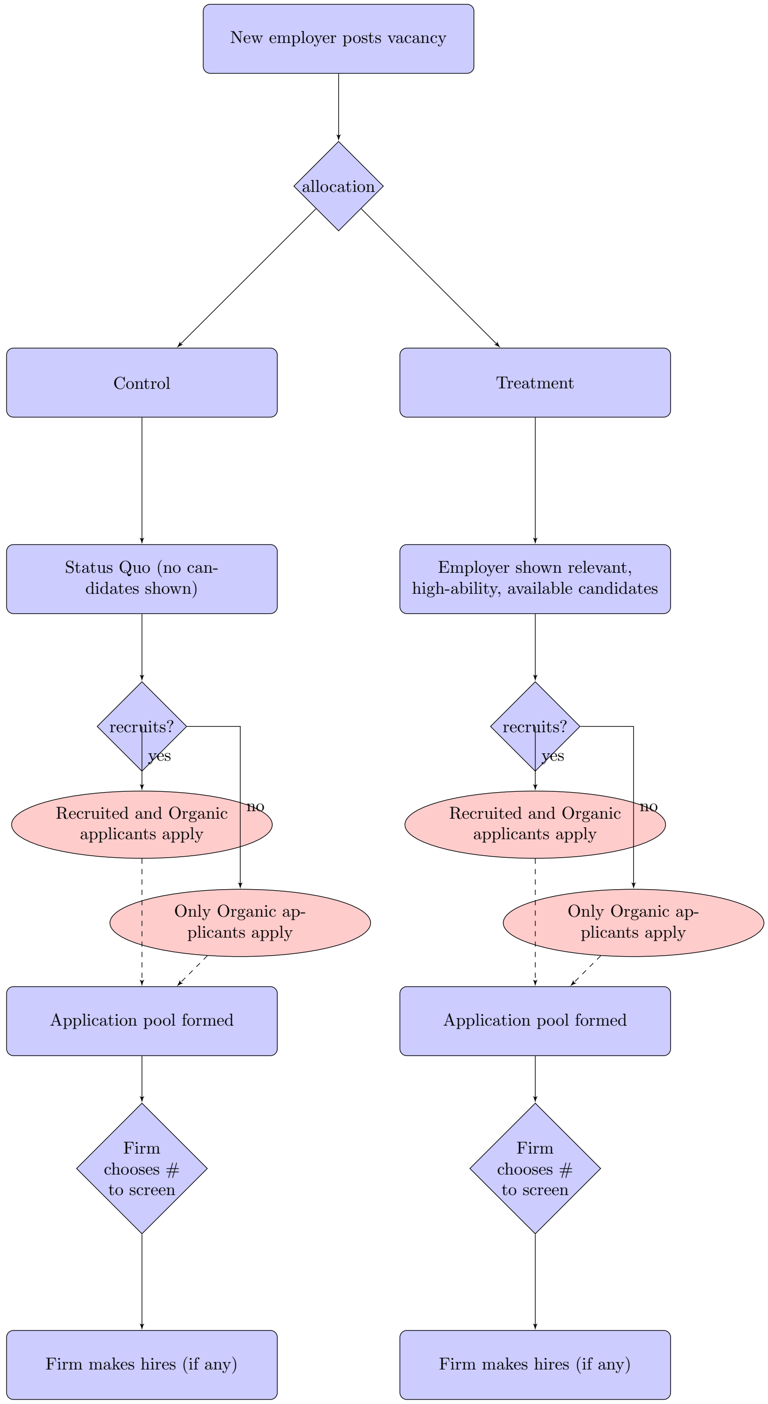
In order to identify which employer recruiting invitations were experimentally induced we can examine the time at which they were made. We know the time that an employer posted a vacancy and the time each recruiting invitation was made, down to the millisecond. Because recommendations were presented as a pop-up immediately after a vacancy was posted (and the employer could not saving the results), all recruiting invitations made shortly after the posting were likely to have been experimentally induced.

## 3.1 Summary statistics for the control group and randomization check

Table 1 presents summary statistics for vacancies in the control group by the type of work. The marketplace is roughly evenly split between technical and non-technical vacancies, with 44% of vacancies being technical. In technical vacancies, employers are more likely to recruit: ("0 > invites" column) and the mean number of recruiting invitations is greater. Technical vacancies receive fewer organic applications and are somewhat less likely to be filled. The biggest difference between the two categories is the size of the wage bill: the average wage bill for technical filled vacancies is nearly three times that of non-technical filled vacancies.

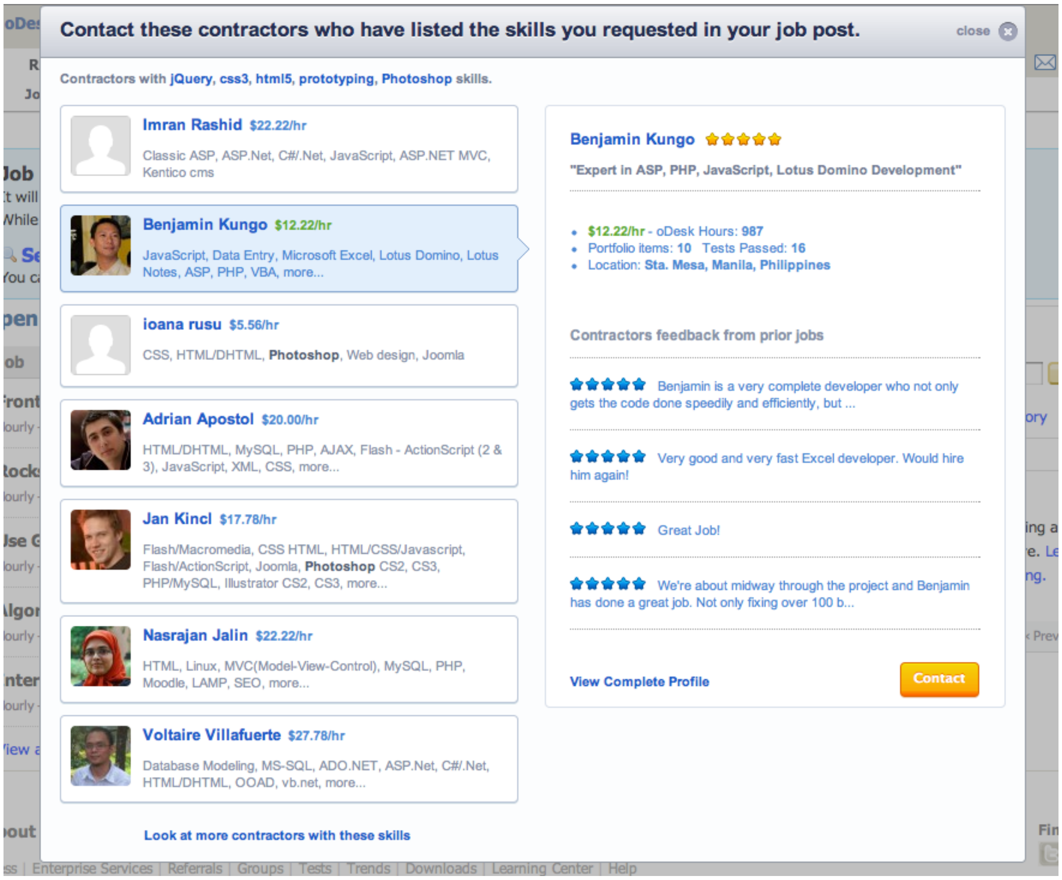
The software used to randomize subjects into groups has been tested in numerous prior experiments. However, as confirmation that the allocation was effective, in Table 2 we report three different measures of balance. First, in the top panel labeled "Observation Counts" we report raw counts of observations in the treatment and control groups. In each panel below, we show the fraction of vacancies of each type (e.g., technical, non-technical, writing, etc.) in each experimental group. In the panel "Fraction of Vacancies—High Level," vacancies are classified

Figure 1: Employer recruiting experiment




*Notes:* This diagram illustrates the flow of employer subjects through the experiment, by treatment and control group. Employers were first randomized into treatment and control groups, with treatment groups receiving algorithmically-generated candidate recommendations. Both groups decided whether or not to recruit candidates. Depending on their choice, employers received recruited applicants (if they recruited) and/or organic applicants. Firms then chose 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.

Table 1: Summary statistics for the control group, by type of work

| | | Discrete firm outcomes: | | Continuous firm outcomes | | |
|---|---|---|---|---|---|---|
| | N | > 0 invites | made a hire | # apps. | # invites | wage bill ($) |
| Non-technical | 3,220 | 0.179 | 0.237 | 12.38 (0.33) | 0.37 (0.02) | 1,495 (116) |
| Technical | 2,565 | 0.233 | 0.211 | 10.88 (0.22) | 0.64 (0.04) | 4,543 (429) |

*Notes:* This table presents summary statistics for the control group of the recommendations experiment. 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 column ">0 invites" shows the fraction of employers that sent any recruiting invitations. The columns " apps." and " invites" show mean number of applications received and invitations sent, respectively. For the "wage bill ($)" column, we compute a trimmed mean that excludes the top 2.5% and bottom 2.5% of filled vacancies.

Table 2: Balance of allocations across experimental groups

| | Treatment | Control | p-value | |
|---|---|---|---|---|
| *Observation Counts* | | | | |
| | 5,629 | 5,785 | 0.144 | ($\chi^2$ test) |
| *Fraction of vacancies—High-level* | | | | |
| technical | 0.429 (0.007) | 0.443 (0.007) | 0.122 | |
| non-technical | 0.571 (0.007) | 0.557 (0.007) | 0.122 | |
| *Fraction of vacancies—Type of Work* | | | | |
| 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 the recommendations experiment. The standard error for each calculated mean is in parentheses next to the 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 the p-value for a $\chi^2$ test. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

as technical and non-technical; in "Fraction of Vacancies—Type of Work" we use a finer classification for the type of work. As expected, all three balance measures show allocations consistent with random assignment ($\chi^2$ test for counts; two-sided t-test for fractions).[1]

# 4 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.

There has been little empirical research to date on how firms fill their vacancies. Oyer and Schaefer (2010) summarize the state of the literature:

[1]As subjects were added to the experiment over time, the experimental groups are not the same size or intentionally balanced/blocked.

> 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 empirical research is not that the topic is unimportant, but rather that data are scarce. 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) find that employers with hard-to-fill vacancies or those that require more training report screening larger pools 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 and we develop a new model. The main substantive differences are that we assume that (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) in that there is uncertainty about the attributes of choices that can be reduced via costly effort.

## 4.1 Batch screening, recruiting and exogeneous screening costs

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) reach 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.

## 4.2 The role of labor market intermediaries

Providing algorithmically generated recommendations in a labor market is a new phenomena; third parties assisting labor market participants is not. Labor market intermediaries have long played an important role in labor markets (Autor, 2008). Their services are often needed because 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."

In a recent meta-analysis of 97 active labor policy programs, Card et al. (2010) find that job search assistance programs were more likely to yield positive short-term impacts compared to public sector employment. A similar analysis of European active labor market policies by Kluve (2010) also concludes that offered services such as placement assistance and counseling are 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) report the results of a recent randomized experiment in France that showed that placement assistance had (modest) benefits without crowd-out effects.

# 5 Model of employer search and screening

To interpret our empirical results, we develop a simple model of employer search and screening. The goals of the model are to characterize the firm's decision of whether or not to search for candidates and predict how these choices will manifest themselves in measurable outcomes such as the number of applicants screened and the fill rate. We will leave most of the explication of the model's predictions to the appropriate point in the empirical analysis.

## 5.1 Basic model without employer search

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, $q_1, q_2, q_3 \dots q_A$, 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 $z$th 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:

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

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 d$a$ applicants, then the probability that they are a match is d$aq(a)$. 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.

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.

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

*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, $h'(a) > 0$.

Differentiating $h'(a)$ 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 (since $h'(a) > 0$) 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). □

## 5.2 Adding the possibility of employer search

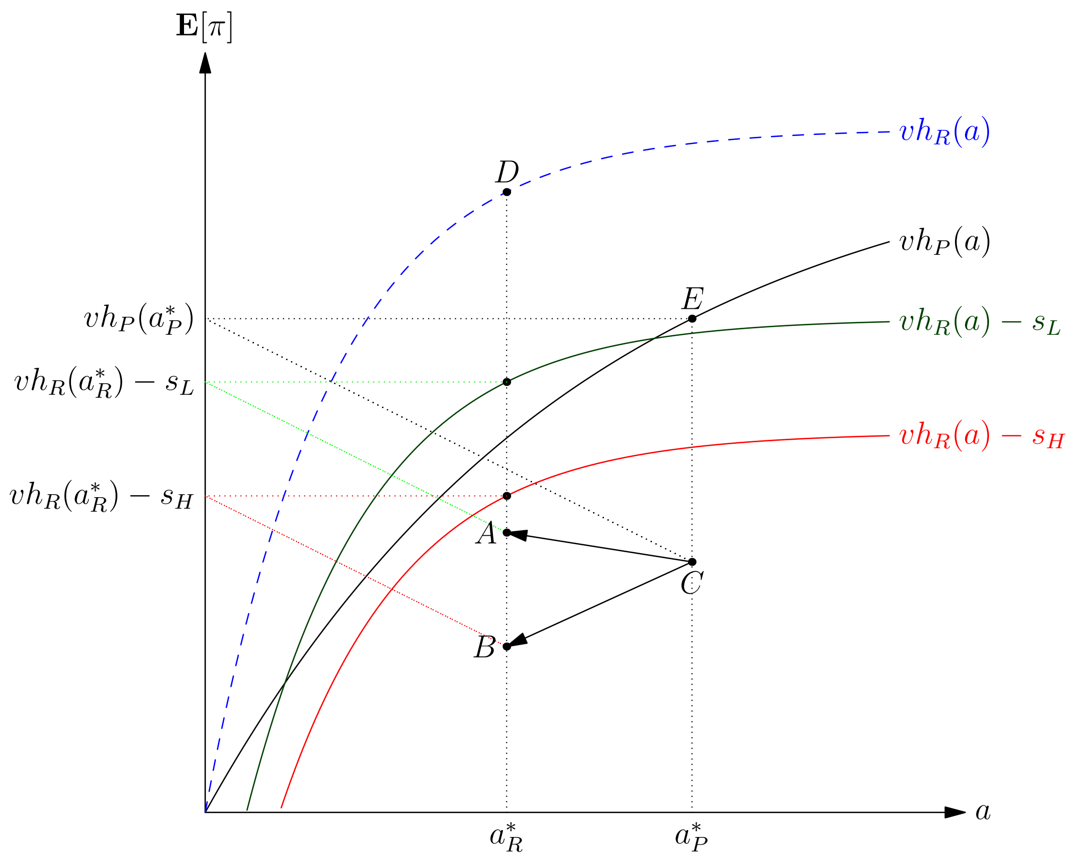
A goal of the model will be to understand when the firm will pay to recruit from a "better" applicant pool. Now we introduce the possibility that firms can recruit. We model the firm as facing a binary choice whether to pursue a *recruiting* strategy or a *passive* strategy. We represent this choice 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 "better" 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:* $\mathbf{E}[A_P|A_R, x = R] = \mathbf{E}[A_P|x = P]$.

Figure 3: The firm's vacancy-filling strategy and the number of applicants to screen




*Notes*: In this figure, the x-axis is $a$, the number of applicants the firm screens. The y-axis is the firm's expected utility, which is the pay-off that the firm gets from filling a vacancy, $v$ times the probability of filling the vacancy using strategy $x$ when evaluating $a$ applicants, which is $h_x(a)$. 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, depending upon the recruiting costs. 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.

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 Lemma 2 dominates the passive distribution.

**Lemma 2.** *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 $f_R \succ f_P$. Further, for the same number of screened applicants, the recruiting strategy offers a higher fill rate: $f_R \succ f_P$, then $h_R(a) > h_P(a)$.*

*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 $F_P = \frac{A_R}{A_R+A_P}F_P + \frac{A_P}{A_R+A_P}F_P > \frac{A_R}{A_R+A_P}G_R + \frac{A_P}{A_R+A_P}F_P = F_R$. $f_R \succ f_P$ 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 Lemma 2 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.

## 5.3 The firm's decision problem

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 $a \leq A$. 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 $c/v \leq h'(A)$, 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.

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

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 $\upsilon h_R(a)$ and $\upsilon h_p(a)$ respectively. In Figure 3, for the recruiting strategy, we draw three curves: $\upsilon h_R(a)$, $\upsilon h_R(a) - s_L$ and $\upsilon h_R(a) - s_H$. 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 $\upsilon h_P(a)$, with a single optimal number of candidates to screen, $a_P^*$.

To decide whether or not to recruit, the firm considers not just the direct costs and benefits of recruiting, but also the associated screening costs, $ca_R^*$. In Figure 3, we can start to understand the firm's decision-making by determining the utility obtainable from a passive strategy. Start from the point $(0, \upsilon h_P(a_P^*))$ (the y-intercept) and draw a line with slope $-c$ for $a_P^*$ units on the x-axis. The vertical distance covered by this line is the screening cost and the height of point $C$ is the expected utility of the passive strategy. For the recruiting strategy, we compare the expected utilities when recruiting costs are high, $s_H$, and low, $s_L$. Because recruiting costs are sunk, they do not effect the number of screening applicants, $a*$ and hence screening costs are identical. Following the same procedure we used for the passive strategy, we can draw a line with slope $-c$ to bring us to the expected utilities $A$ and $B$ when recruiting costs are $s_L$ and $s_H$, respectively. In the figure, at $s_L$, recruiting is more attractive, while at $s_H$, the passive strategy would be more attractive.

In Figure 3, we plot the point $(a_R^*, \upsilon h_R(a_R^*))$ 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 $\upsilon$ 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 we will later show in Proposition 2. However, to prove this, we need Lemma 3 that when two strategies have the same fill rate, a marginal applicant offers more benefit when the firm is pursuing a recruiting strategy.

**Lemma 3.** *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 $h'_R(a_R) > h'_P(a_P)$. □

Table 3: Means comparison of outcomes across treatment and control groups

| | Treatment | Control | p-value | |
|---|---|---|---|---|
| *Outcomes for All Vacancies* | | | | |
| Number of recruited candidates | 1.018 (0.051) | 0.951 (0.049) | 0.345 | |
| Number of applications received | 13.936 (0.306) | 14.146 (0.304) | 0.624 | |
| Recruited early | 0.172 (0.005) | 0.132 (0.005) | <0.001 | *** |
| Filled vacancy | 0.230 (0.006) | 0.225 (0.006) | 0.496 | |
| Hired an organic applicant | 0.212 (0.005) | 0.210 (0.005) | 0.784 | |
| Hired an early recruited applicant | 0.038 (0.003) | 0.035 (0.002) | 0.404 | |
| *Outcomes for Technical Vacancies* | | | | |
| Number of recruited candidates | 1.099 (0.067) | 1.165 (0.084) | 0.532 | |
| Number of applications received | 11.921 (0.278) | 12.301 (0.295) | 0.342 | |
| Recruited early | 0.203 (0.008) | 0.151 (0.007) | <0.001 | *** |
| Filled vacancy | 0.246 (0.009) | 0.211 (0.008) | 0.003 | ** |
| Hired an organic applicant | 0.219 (0.008) | 0.195 (0.008) | 0.036 | * |
| Hired an early recruited applicant | 0.048 (0.004) | 0.037 (0.004) | 0.047 | * |

*Notes:* This table reports statistics by treatment and control group from the recommendations experiment. The standard error for the mean is in parentheses, next to the estimate. The reported p-values are the for two-sided t-tests of the null hypothesis of no difference in means across groups. Note that for the hiring measures, "Hired and organic applicant" and "Hired an early recruited applicant" do not sum to the total hiring measure "Filled vacancy" because employers can make multiple hires. Significance indicators: $p \le 0.05 : *$, $.p \le 0.01 : **$ and $p \le .001 : ***$.

# 6 Experimental results

Because vacancies were randomly assigned to experimental groups, we can begin our analysis by simply comparing treatment and control group means. Table 3 reports these means, standard errors and the p-value for a two-sided t-test of the null hypothesis of no difference across experimental groups. The top panel of the table (labeled "Outcomes for All Vacancies") uses all vacancies as the sample, while the bottom panel is restricted to technical vacancies (labeled "Outcomes for Technical Vacancies").

## 6.1 Effects of algorithmic recommendations on employer recruiting

If the treatment lowers the recruiting cost, $s$, then we expect more employers to recruit. Proposition 1 shows that this is a prediction of the model: lower recruiting cost increases the expected utility of recruiting.

**Proposition 1.** *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* $\partial\pi(x = P)/\partial s = 0$.

*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 derivative with respect to $s$, we have $\frac{\partial\pi(x=R)}{\partial s} = -1$ and $\frac{\partial\pi(x=P)}{\partial s} = 0$. $\square$

In Table 3, we examine whether the Proposition 1 prediction is borne out in the data. For each vacancy, we compute the mean number of recruiting invitations sent. The means do not differ significantly across experimental groups using the full sample or when restricting the sample to just technical vacancies. However, for both samples, the fraction of employers recruiting is substantially higher in the treatment group. The left panel of Figure 4 plots the fraction of employers recruiting by vacancy type (technical and non-technical) and experimental group. The figure makes it clear that although the treatment increased recruiting for non-technical vacancies, the increase was both relatively and absolutely smaller. Despite the apparent differences between technical and non-technical vacancy treatment effects, we cannot reject the null hypothesis of constant effects (Wald test, $p = 0.283$).

The treatment increased the fraction of employers engaged in recruiting but it did not significantly affect the mean number of recruiting invitations sent. This suggests that the treatment worked primarily by inducing employers that otherwise would not have recruited to send at least one invitation, but that this increase was "lost" in computing the mean. This is plausible since the mean number of recruiting invitations has a high variance—some employers send very large numbers, even though most send none or one. An alternative outcome measure an indicator for whether the employer sent more than *one* invitation. Unlike the overall mean, this measure has a suitably small variance but still provides insight into how the treatment affected the distribution of invitations that were sent. In Table 4, Column (2), we regress this more-then-one indicator on the treatment indicator. Compared to Column (1) (which simply re-capitulates the previous means comparison using the more-than-zero indicator), the Column (2) estimated treatment effect is nearly halved, though it is still significant. This suggests that the treatment mostly worked on the extensive margin but perhaps had some intensive effects as well.

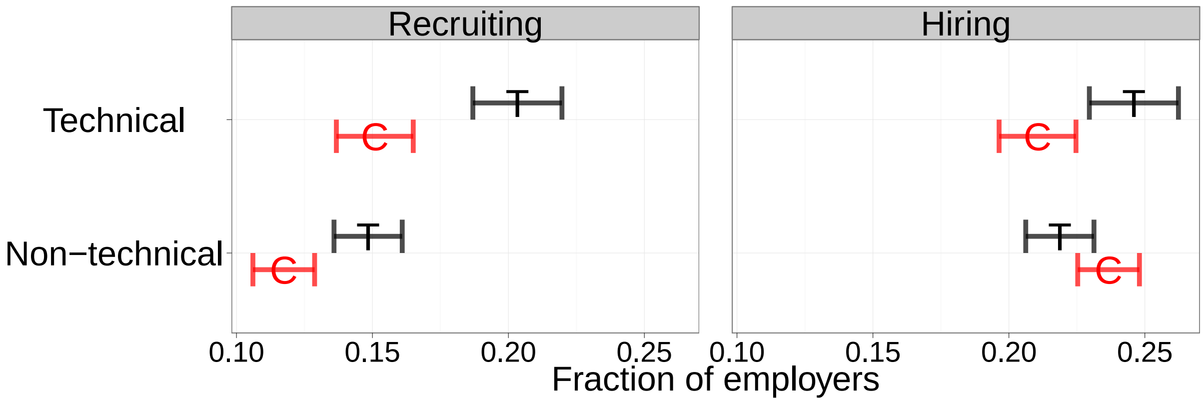
Both control and treatment employers were free to recruit non-recommended candidates at any time after posting their vacancies. Because recommendations were made immediately after an employer posted a vacancy, the treatment should affect the fraction of "early" recruits but not the fraction of "late" recruits. We define invitations as "early" or "late" based on whether they were sent within an hour of the posting. Comparing the treatment's effects on early and late recruiting behavior helps us test two hypotheses for how the treatment may have worked. First, we can see whether the treatment alerted employers to the very possibility of recruiting, regardless of whether they followed the recommendations. Second, we can see whether

Table 4: 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*: The dependent variable in each of these OLS regressions are various measures of employer recruiting behavior from the recommendations experiment. This table reports the results of OLS regressions. The dependent variables are measures of employer recruiting; the main independent variable is the treatment indicator. The dependent variable in Column (1) is whether the employer sent any recruiting invitations; in Column (2) it is whether they sent more than one recruiting invitation. In Columns (3) and (4) the dependent variable is whether the employer sent any "early" (first hour after posting a vacancy) or "late" recruiting invitations, respectively. In Column (4), the regression includes controls for whether the vacancy-posting employer required a worker to have prior experience, the log of the number of characters in the employer's job description and fixed effects for each category of work (not shown). The standard errors are robust to heteroscedasticity. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

Figure 4: Employer recruiting and hiring by vacancy type (technical versus non-technical) and treatment group




*Notes:* This plot shows the mean fraction of employers recruiting (as measured by sending an early recruiting invitation) in the left panel and making a hire in the right panel. The data come from the recommendations experiment. 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).

recommendation-induced invitations substituted for recruiting that would have occurred anyway, albeit at a later date. In Table 4, Column (3), we regress an indicator for whether the employer made any early recruiting invitations on the treatment indicator, and in Column (4) we run the same regression but use an indicator for any late recruiting invitations. In Column (5), we add several pre-treatment variables, including the length of the job description (log of the characters) and an indicator for whether the employer required applying organic applicants to have some number of hours worked on the oDesk platform. Both measures have a strong positive correlation with recruiting, which suggests they might proxy for a higher value project or a perhaps one like to receive fewer organic applicants. The treatment increased early recruiting but had no effect on late recruiting, implying that the treatment worked by directly providing the employer access to new candidates that otherwise would not have been recruited.

## 6.2 Effects of algorithmic recommendations on vacancy fill rates

The treatment increased employer recruiting, but the goal of the experiment was to increase fill rates. We define a "fill" as the employer hiring at least one worker and spending some amount of money against their vacancy. Given that the treatment increased recruiting—and thus according to the model, the number and likely quality of applicants—we expect fills to increase

Table 5: Effects of treatment on hiring of applicants

| | Employer hired: | | | |
|---|---|---|---|---|
| | All Vacancies | Technical Vacancies Only | | |
| | (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 technical vacancies. The data from the recommendations experiment. The standard errors are robust to heteroscedasticity. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

in the treatment group. The model does make this prediction, with Proposition 2 proving that a switch to the recruiting strategy increases fill rates in expectation.

**Proposition 2.** *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 Lemma 1. If $a_R^* < a_P^*$, then consider an $a'$ such that $h_P(a') = h_R(a_R^*)$. Because $h_R'(a_R^*) = h_p'(a_P^*) = c$, by Lemma 3, $h_P'(a') < c$. Because $h_P(\cdot)$ 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 see whether Proposition 2 holds in the data, we start again with our means comparison table, Table 3. In both the top and bottom panels of the table, there is a row labeled "Filled vacancy." For the full sample, there is no meaningful difference in means, but for the technical vacancies, the fill rate is substantially higher in the treatment. The right panel of Figure 4 shows the fraction of filled vacancies by treatment group and by vacancy type. As we already saw in the means table, The treatment had a positive effect on fills for technical vacancies, but for non-technical vacancies, the fill rate was actually lower. A Wald test confirms what is visually evident: we can easily reject a null hypothesis of homogeneous treatment effects ($p = 0.001$).

To provide standard errors for the treatment effect, in Table 5 we report regressions in which the outcome variables are various measures of hiring. Confirming what we learned from the means comparison, Column (1) shows the that treatment effect for the full sample is positive but not significant. Restricting our attention to technical vacancies, in Column (2) we see that the treatment effect was about 3.5% points—a large and statistically significant difference. Given the "uptake" of recommendations in the treatment group, the implied treatment-on-the-treated (TOT) effect is about 65%, with a standard error of 22%, assuming uptake is perfectly estimated.[2]

In any experiment, when treatment effects are only significant for a sub-population—as they are here—a natural concern is that the "result" is due to the multiple comparisons problem: if we examine enough sub-populations but use the usual cutoffs for statistical significance, eventually we will find sub-population(s) that have significant treatment effects purely by chance. However, this is not a concern in our setting. Even with the conservative Bonferroni correction for simultaneous inference, the treatment effect on fill rates for technical vacancies is still conventionally significant. If we use $\alpha = 0.05$ and we planned to compare treatment effects for two vacancy populations ($m = 2$)—technical and non-technical—then our $1 - \alpha/m$ Bonferroni-adjusted confidence interval still does not contain an estimate of no effect, as the p-value for the means comparison on fill rates is $p = 0.003$.

### 6.3 Why are algorithmic recommendations only effective in technical categories of work?

The treatment only increased fill rates for technical vacancies—a natural question is why this is the case, particularly since the treatment somewhat improved recruiting for non-technical vacancies. Several hypotheses seem plausible—none of which are mutually exclusive:

- The recommendation algorithm was simply more effective for technical vacancies.
- The lower average number of organic applicants, $A_p$, in technical vacancies made recruited applicants more valuable.
- Technical vacancies have higher screening costs and high screening costs make recruiting—which can lower the number of candidates that must be screened—more effective.

The first hypothesis—that technical recommendations were better recommendations—is simple and intuitively appealing, but the data are not very supportive. One way to measure

[2]Although this estimate of the TOT seems very large (albeit given the confidence interval, much smaller values are plausible), we will see in Section 7 than an IV estimate suggests that the marginal value of a recruited candidate is very high.

the quality of recommendations is to see how often they lead to invitations that are then acted upon by the recruited candidate. The idea is that the recruited candidates would know a good match when presented with one and would be more likely to accept, i.e., apply to the vacancy. This measure of recommendation quality is better than looking at hires directly, since crowd-out concern is less important. To elaborate, because the treatment increase the quantity of recruited candidates, even if their quality is better than what would be obtained with the control, the greater number of alternatives pushes down the hire probability. Using this recommendation response measure, we find all null effects (not reported here): there is no difference in the invited-candidate response rate across experimental groups or across treatment groups, suggesting that experimentally-induced recruits were similar on this dimension to "natural" recruits, regardless of vacancy type.

The second hypothesis is that because technical vacancies receive fewer organic applications on average, they receive a bigger benefit from the treatment. However, technical vacancies valuing recruits more is not sufficient to explain a heterogeneous treatment. This greater valuation should create a baseline higher level of recruiting, but what we care about is the sign of $\frac{\partial}{\partial A_p}\left[h_R(a_R) - h_P(a_P)\right]$. We can show that this partial derivative reduces to $c\left(\frac{\partial a_R^*}{A_P} - \frac{\partial a_P^*}{A_P}\right)$. This means that the effect of $A_p$ on the change in fill rates in turn depends on how $a_P^* - a_R^*$ changes with $A_p$. Characterizing this derivative is not straightforward, as the effect of more applicants on each $a^*$ depends on two effects moving in opposite directions. Increasing the size of the pool has two effects: it increases the quality of the marginal applicant at $a^*$ and it increases the probability that a match will already have been found with $a^*$. The first effect pushes towards a larger $a^*$, while the second effect pushes towards a smaller $a^*$—the actual direction of the effect depends upon which effect dominates.

It is theoretically ambiguous as to whether the gap in fill rates increases in $A_P$. However, in Appendix B we show empirically that a decreasing gap is more probable. The analysis examines the effects of $A_p$ on recruiting, screening and hiring by exploiting the fact that employers post multiple vacancies over time. By observing identical employers posting different kinds of vacancies over time, we can net out employer-specific effects. We find that greater $A_p$ tends to suppress recruiting and increase screening. If the decreased recruiting also lowers $a_R^*$ by making fewer candidates available to be evaluated—then it seems likely that $\frac{\partial a_R^*}{A_P} - \frac{\partial a_P^*}{A_P} < 0$. This would in turn imply that a lower $A_p$ tends to increase the gain in fill rate from moving from a passive to a recruiting strategy. If true, this would be consistent with our finding of a larger hiring treatment effect for technical vacancies—since $A_P$ is lower, on average, for technical vacancies.

Finally, we consider the effects that per-applicant screening costs, $c$, might have on relative treatment effects. As with $A_p$, the size of the gain from switching to a recruiting strategy ends up depending on the effect $c$ has on the number of screened applications under the recruiting and passive strategies. However, unlike in the case of $A_p$, the prediction is clearer: Proposition 3 predicts that vacancies with higher screening costs experience a greater increase in fill rates when switching from a passive to a recruiting strategy, subject to some likely-to-be-met constraints on $A_p$. Yet for this prediction to explain why technical vacancies have higher fill rates under the treatment, we need to confirm that technical vacancies have higher screening costs.

It seems plausible that screening costs are higher for technical vacancies. Presumably technical jobs require a larger body of knowledge that must be assessed during screening and the project may be more complex, requiring a more in-depth understanding of both employer and worker attributes and abilities. However, we do not directly observe screening costs—but we do observe their consequence. These costs—combined with the quantity and quality of received applicants—determine the number of evaluated applicants, which in turn determines whether or not the employer filled their vacancy. We can use this causal dependency to estimate a structural model whose latent parameters are the screening costs and applicant match probability. These two parameters can be estimated separately for technical and non-technical vacancies, allowing us to compare relative screening costs.

**Proposition 3.** *If the number of applicants is sufficiently large and assuming both strategies have an interior solution, then as per-applicant screening costs increase, the fill rate gap between the recruiting and passive strategy increases:* $\forall A > \bar{A}, \frac{\partial}{\partial c}[h_R(a_R^*) - h_P(a_P^*)] > 0$.

*Proof in Appendix C.*

### 6.3.1 Structural estimate of employer screening costs

Let us assume that we have $k$ vacancy categories. We assume that within a category, all vacancies have the same screening-cost to value ratio, $\frac{c_k}{v_k}$ and match probability, $q_K$. Our hiring function is thus simply $h_k(a) = 1 - \exp(-q_k a)$. Employers optimally choosing some number of candidates to evaluate gives us one moment condition:

$$\mathbf{E}[h'(a_i^*; q_k) - c_k / v_k] = 0 \tag{6}$$

Note that because we observe whether 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 $y = 1$

Table 6: Estimate of the ratio of screening costs to project value and category-specific match probability

| | $a^*$ = Interviews | | $a^*$ = Applications | |
|---|---|---|---|---|
| | Non-Tech. | Tech. | Non-Tech. | Tech. |
| | (1) | (2) | (3) | (4) |
| Screening cost / Project Value | 0.06514*** | 0.08091*** | 0.01502*** | 0.01784*** |
| | (0.00078) | (0.00171) | (0.00011) | (0.00025) |
| Match probability | 0.0905*** | 0.11344*** | 0.02141*** | 0.02239*** |
| | (0.00468) | (0.0101) | (0.00057) | (0.00171) |

*Notes:* This table reports estimates $\frac{c_k}{v_k}$ and $q_k$ from a structural model of employer hiring, using pool data from the recommendations experiment. 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.

indicates a filled 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 turn gives us one moment condition based on the score: $\mathbf{E}[L'(q_k)] = 0$. The vector of moment conditions is thus:

$$g\left(a^*, y; \frac{c_k}{v_k}, q_k\right) = \begin{bmatrix} h'(a^*) - c_k / v_k \\ \frac{\partial \log L(q_k)}{\partial q_k} \end{bmatrix} \tag{7}$$

One difficulty we face is that evaluations ($a^*$ in the model) are not unambiguously measured. 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. Using the number of applicants seems counter to our $a^*$ versus $A$ modeling distinction, but not if employers immediately close a job once they have received $a^*$, even though they would have received $A_p$ (particularly since we are assuming uniform $q_k$). As far as which evaluation proxy is more reasonable, "interviews" is probably an under-count of true evaluations, while "applications" received is probably an over-count. We estimate the model using both methods. We restrict the sample to vacancies where at least one evaluation was observed, since doing otherwise can make the likelihood function unbounded.

Table 6 reports $\frac{\hat{c}_k}{v_k}$ and $\hat{q}_k$. The two-stage generalized method of moments (GMM) is used

(Hansen, 1982).[3] Separate estimates are presented for technical and non-technical vacancies. Standard errors are computed using the weight matrix of the second step of the two-stage GMM.

When using the total "applications" measure, the estimated $c/v$ and $q$ are much lower than they are when using the "interviews" measure, regardless of vacancy type. This is expected since using applications to proxy for $a^*$ makes it look like the firm evaluated many more candidates (suggesting evaluation is cheap) but made the same number of hires (suggesting low per-applicant match probabilities). Comparing costs across vacancy types, we can see that the estimated $c/v$ ratio is far higher for technical vacancies: the ratio is 24% higher for technical vacancies using the interviews measure and 15% higher for technical vacancies using the total applications measure. This is the pattern we would expect if the logic of Proposition 3 holds and heterogeneity in screening costs is driving the larger treatment effect for technical vacancies.

## 6.4 Did recruited applicants crowd-out organic applicants?

From the perspective of the organic applicant, the experimentally-induced increase in recruited applicants—close competitors offering a substitute service—was probably viewed as a purely negative development. However, among technical vacancies, the treatment raised the overall fill rate. This means that the treatment did not simply cause a 1-for-1 swap out of counterfactually hired organic applicants with experimentally-induced recruited applicants. Had this occurred, the fill rate would have remained the same. However, the overall rise in fill rates does not mean there was no displacement. We can look for evidence of displacement by decomposing the "fill" measure into indicators for whether the vacancy was filled by an organic or a recruited applicant.

As before, we can begin our analysis by comparing means: in Table 3, under the "Filled Vacancy" row we have two sub-rows: "Hired an organic applicant" and "Hired an early recruited applicant."[4] Recall than an early recruited applicant is likely to have been an experimentally-induced applicant. Unsurprisingly, given the lack of overall treatment effect on the fill rate, in the full sample there is no difference in either measure. However, for technical vacancies, we can see that both recruited *and* organic applicants were more likely to be hired. To obtain standard errors on the treatment effects, in Table 5 we report regressions where the outcome variables are whether the employer hired a recruited candidate, Column (3), or an organic candidate, Column (4). In both cases, the treatment has a large and significant effect. Greater hiring

[3]The optimization step was done using Mathematica's NMinimize function.

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

of organic applicants in the treatment suggests that the recruited applicants were actually gross complements. While this may seem counter-intuitive, this kind of result has been found before in contextually similar settings. Yang and Ghose (2010) found that paid and organic search engine results—similarly "obvious" substitutes—were in fact complements, each stimulating more clicks on the other "side."

Despite the attractiveness of the substitutes framing of organic and recruited applicants, pure substitutability is not a prediction of the model. In fact, the model predicts that in some cases, organic candidates are *only* hired when the firm recruits. To see why, consider the simple case where an interior solution to $a^*$ exists only in the recruiting case, but not in the passive case: if $c > v h_P(0)$, then $a_P^* = 0$, meaning that no applicants are screened, but if $v h_R(0) > c$, then $a_R > 0$, and so long as the best organic applicant is better than the worst recruited applicant, a non-zero number of organic applicants will get screened in expectation (and thus have a chance of being hired) *only* when the firm recruits.[5]

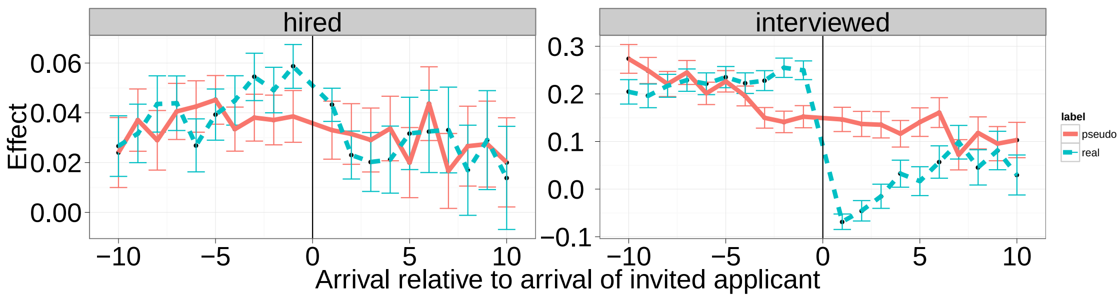
#### 6.4.1 Micro-foundations for spill-overs

The model provides a high-level explanation for organic/recruited complementarity: informally, the presence of recruited applicants causes more extensive screening which benefits organic applicants as well. Because of our precise data on the timing of when events occur in the marketplace, we can see whether this explanation is "micro-founded" in the oDesk context. To do this, we will compare the outcomes of applicants that "arrived" right before and right after the arrival of a recruited applicant. We suspect that applicants that arrived before (and hence were available to be viewed) got more attention than applicants arriving immediately after. Employers are emailed a notice when recruited candidates accept their recruiting invitations. These emailed notifications potentially triggered a screening session that we hypothesize benefited the organic applicants.

Of course, the outcomes of applicants arriving before and after some point in time are inherently confounded with time. We can partially deal with this problem by including time and arrival rank specific controls in a regression, but we also construct a placebo group of vacancies. For the placebo group, we find vacancies in which an employer sent a recruiting invitation but that the recruited worker rejected. We counter-factually assume that the recruited but declining candidates actually did apply and applied after the same elapsed time as those that actually did apply.

[5] This story raises the question of why an employer would bother posting a vacancy if they planned not to screen any applicants, though one possibility is that $v$ is a random variable prone to idiosyncratic shocks.

Figure 5: 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 8. 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 recruiting invitation. The estimates of $\beta_D$ were computed using a multi-level model, using pooled observations of all the applications made to vacancies from the recommendation experiment.

Let $2D$ be the number of applicants around the arrival of a recruited applicant that we are interested in. 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: $d_{ij} = (r_{ij} - z_j) \cdot \{|r_{ij} - z_j| <= D \wedge r_{ij} \neq z_j\}$ 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 $\gamma$ be a vacancy-specific effect and $\alpha(r_{ij})$ be an arrival-rank specific effect.

$$y_{ij} = \beta_d \cdot \{d = d_{ij}\} + \gamma_j + \alpha(r_{ij}) + \epsilon_{ij} \tag{8}$$

We can estimate Equation 8 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 recruited candidate that declined the employer's invitation. The counter-factual, predicted arrival is made by drawing simulated arrival times from a fitted linear model.

In Figure 5, we plot $\hat{\beta}_{-D}$ to $\hat{\beta}_D$, fit with a multi-level model, using all applications to vacancies from the recommendation experiment, pooled across both the treatment and control groups. The dashed line connects estimates for actual arrivals, while the solid line connects estimates for the placebo group. In the left panel, the outcome is hiring, while in the right panel, the outcome is interviewing. We can see that the four applicants arriving immediately before the recruited applicant were more likely to be hired and 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 substantial difference in hiring is not in the post arrival period, but in the noticeable bump of "extra" hiring these applicants obtain in the pre-arrival period.

## 6.5 Effects of algorithmic recommendations on match attributes

The explicit goal of the experiment was to raise the quantity of matches, yet we ultimately care about both quantity and quality. Assessing match quality is difficult, in that we only observe measures of match quality for filled vacancies and we cannot easily separate organic from experimentally-induced matches. Further, the model has little to offer in interpretation—it is silent on how wages are determined and it implicitly assumes that matches are homogeneous. These concerns notwithstanding, we have at least two ways to measure match quality. One is the wages of hired applicants for hourly contracts; another is the employer-provided feedback given to the hired applicant at the end of the relationship. When a contract ends, employers publicly rate workers on a 1-5 "star" scale.

Hired wages can be thought of as proxying for the productivity of the match and feedback as proxying for the employer's surplus. Obviously wages-as-marginal-productivity is making a strong assumption about bargaining and feedback-as-surplus-proxy is more of a psychological than an economic assumption. Furthermore, feedback is inherently subjective and endogeneous, in that the rating is given by the same employer that also made the hiring and recruiting decisions. As on many electronic commerce sites with a bi-lateral feedback mechanism, there are strong incentives for parties to give positive feedback. On oDesk, a little more than 50% of all feedback scores are perfect, 5-star scores. With these caveats in mind—both the inherent difficulty of measuring match quality and the problems with our measures—we still compare these measures across experimental groups.

In Table 7, in Columns (1), (2) and (3), the outcomes are whether the mean hires wage exceeded \$3/hour, \$9/hour and \$12/hour, respectively. The cutoffs approximately correspond to the 25th, 50th and 75th percentiles of the wage distribution for hired candidates. We can see that regardless of threshold, the treatment effect is positive, but small and insignificant. In Column (4), the outcome is whether the mean feedback for hired applicants was a perfect 5 stars. As with the wage regressions, the coefficient on the treatment indicator is positive but small and insignificant. Using both the feedback measure and the wage measure, we find no evidence that matches formed in the treatment group were any worse than those formed in the control group.

Table 7: Effects of algorithmic recommendations on match attributes for technical vacancies

| | Hired hourly wage exceeded: | | | 5-star feedback?: |
|---|---|---|---|---|
| | \$3/hour | \$9/hour | \$12/hour | Yes |
| | (1) | (2) | (3) | (4) |
| Intercept | 0.148*** | 0.107*** | 0.070*** | 0.097*** |
| | (0.007) | (0.006) | (0.005) | (0.006) |
| Treatment | 0.015 | 0.011 | 0.006 | 0.011 |
| | (0.010) | (0.009) | (0.007) | (0.009) |
| N | 4,980 | 4,980 | 4,980 | 4,980 |
| R-squared | 0.000 | 0.000 | 0.000 | 0.000 |

*Notes*: The dependent variable Columns (1)—(3) in each of these OLS regressions is whether the mean wage for applicants hired for that vacancy exceeded some threshold. In Columns (1), (2) and (3), the threshold wages are \$3, \$9 and \$12 per hour, respectively. In Columns (4), the outcome variable is whether the mean employer provided feedback was equal to 5 stars. The data from the recommendations experiment. The standard errors are robust to heteroscedasticity. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

# 7 Quasi-experimental approach to the effect of recruited applicants on fill rates and crowd-out

There are three main findings from the recommendation experiment: (1) lowering the cost of recruiting increases recruiting (2) increased recruiting raises fill rates for some types of vacancies and (3) recruited and organic applicants are complements. However, some of these conclusions rest upon imprecisely estimated treatment effects. Further, the finding that increased recruiting improved fill rates—a central conclusion of the analysis—hints at, but does not show the causal mechanism. Presumably the reason the treatment increased fills was that the employers that acted upon the recommendation were able to hire applicants they otherwise would not have been exposed to. However, the experiment does not show this directly—it just shows that employers eligible to get recommendations also had, on average, a higher fill rate. We cannot justifiably look at only those employers that recruited in response to the recommendations, since we do not know who their counter-parts would have been in the control.

There is a more direct way to assess the effects of a marginal recruited applicant on fill rates: we could restrict our sample to recruiting employers and then exogenously vary the recruited candidate response rate—such as by intercepting and discarding some fraction of positive responses. This would give us the causal effect of having an additional recruited applicant in the pool of applicants—this estimate would be similar in interpretation to the TOT we calculated earlier. Unfortunately, directly manipulating recruited candidate response rates is infeasible. However, if we could find factors that (a) influence response rates but (b) are unrelated to other factors affecting vacancy outcomes, then we could obtain the same estimate we would have obtained experimentally. In this section, we describe factors that meet qualifications (a) and (b) and thus let us perform an instrumental variables analysis that recovers the causal effect of an additional recruited applicant on fill rates and crowd-out.

## 7.1 Instrumenting for a recruited candidates response to an invitation

Workers are inherently supply constrained. Their capacity for taking on additional work—and hence their willingness to accept recruiting invitations—ebbs and flows depending upon how many other projects they are working on at a moment in time. As a result, their recruiting response probability varies predictably in response to external factors—namely how many recent recruiting invitations they have received and accepted. We will show that a worker's per-recruiting invitation response probability is decreasing in the number of recruiting invitations recently received, but increasing in the number accepted. The number of invitations received and accepted by a worker certainly has a stochastic component, as recruiting employers do not coordinate. However, the number of invitations is also not as good as randomly assigned—how heavily a worker is recruited is likely correlated with other attributes of the worker. For example, a sought-after worker is presumably sought-after because they are good. This is problematic for our purposes, because a good worker is also more likely to be hired. However, if we can control for a worker's baseline propensity to be recruited and accept recruiting invitations, then the remaining variation in response probability is the variation we want.

To give an example of how we would use idiosyncratic variation, consider two observationally identical workers, A and B, that, on average, receive exactly one application each morning for jobs that last one day. Assume that they only accept one invitation each day and that if they receive more than one, they choose among them at random. On a particular day, suppose A gets 2 invitations and B gets 1, with each invitation made by a different firm. The firms themselves do not know how many other invitations a worker received, but they do know the average number. In this scenario, one of A's would-be employers will get crowded out and their vacancy will have one fewer recruited applicant. We can then compare the outcomes of the unlucky A-recruiting firm to the outcomes of the lucky B-recruiting firm. Identification comes from the fact that A and B are comparable but that on a particular day, their response rates varied for idiosyncratic reasons that the firms could not foresee or condition upon.

For our actual instruments, we will use the daily count of invitations—both received and accepted—while controlling for *by-week* counts of the same quantities. The identifying assumption then becomes that the employer may condition on the weekly recruiting invitation counts (or any other observable factors correlated with those counts) but not the remaining daily count variation. On oDesk, recruiting invitation counts and responses are observable to employers. 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—a contention that out instrument diagnostic tests support.

For the sample, we use a large number of vacancies where the employer made just one early recruiting invitation. By using only employers sending a single, early invitation, we avoid potentially selecting our sample on an outcome (e.g., employers whose early invitations are rebuffed might alter the number of late invitations they send). The data has 134,654 vacancy observations.

## 7.2 IV results

In Table 8 we present the instrumental variables results. The endogeneous regressor in each IV regression is whether the invited worker accepted the employer's recruiting invitation ("Accepts invite"). Column (1) reports the first stage regression (coefficients on the month and category dummies are not shown). The first stage is strong, with an F-statistic over 40; each of the excluded instruments are highly significant and have the predicted signs. Each of the second-stage regressions easily passes the Sargan over-identifying restrictions test.

In Columns (2) and (3), the outcome variable is whether the vacancy was filled. Column (2) reports the OLS regression, with the excluded instruments included as regressors. In Column (3), the coefficient on "Accepts invite" is both large and statistically significant: the LATE on fill rate from a positive response from the recruited candidate is nearly 40%. Comparing the OLS and IV coefficients, we see that the OLS estimate is considerably smaller, suggesting that workers more likely to accept invites are adversely selected with respect to hire-probability.

In Column (4), the dependent variable is whether the firm hired an organic applicant and in Column (5) it is whether the firm hired *the* recruited applicant. As expected, the coefficient on "Accepts invite" is large and significant in Column (5), but remarkably, it is positive (albeit not significant) in Column (4). What makes this result remarkable is that despite "Accepts invite" having a huge positive effect on the overall fill rate and the probability of hiring the recruited candidate, it still does not obviously crowd-out organic applicants.

The IV analysis adds two pieces to our understanding of how vacancies are filled. First, it

Table 8: Effects of positive recruiting invitation response on fill rates

| | First Stage | Employer hired: | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | Accept Inv. | Anyone | Anyone | Organic | Recruit |
| Accept Invite | | 0.205*** | 0.388*** | 0.156 | 0.346*** |
| | | (0.003) | (0.111) | (0.113) | (0.090) |
| Log applications | 0.058*** | 0.037*** | 0.026*** | 0.050*** | 0.002 |
| | (0.001) | (0.001) | (0.007) | (0.007) | (0.005) |
| Invites (7d) | -0.010*** | -0.000 | 0.002 | 0.004* | -0.003* |
| | (0.000) | (0.000) | (0.001) | (0.002) | (0.001) |
| Invites (24h) | -0.018*** | -0.003 | | | |
| | (0.002) | (0.002) | | | |
| Accepted invites (7d) | 0.005*** | -0.004*** | -0.005*** | -0.003** | -0.003*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Accepted invites (24h) | 0.019*** | 0.004 | | | |
| | (0.003) | (0.003) | | | |
| Intercept | 0.431*** | 0.352*** | 0.273*** | 0.156** | 0.048 |
| | (0.016) | (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 | 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 (3)), an organic application (Column (4)) and a recruited applicant (Column (4)). Column (1) reports the first stage regression (category and month fixed effects not shown), which is the same for all three regressions. Column (2) reports the OLS estimate (category and month fixed effects not shown). The endogenous regressor is "Accepts invite" which is an indicator for whether or not the invited candidate accepted the recruiting invitation. The data used for these regressions is a sample of 134,654 vacancies in which the associated employer sent one early recruiting invitation. Each regression has two excluded instruments: the number of recruiting invitations and the number of accepted recruiting 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 recruiting invitations received and accepted that same *week* by the invited candidate. For each regression, we report the F-statistic for the first stage of the 2SLS. 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 $p \leq .001 : ***$.

Table 9: 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: Lowered recruiting costs increases recruiting | strong yes | Very strong experimental evidence in favor |
| P2: Fill rates higher when firms recruit | yes but ... | Only true in technical vacancies |
| P3: High screening costs imply greater fill rate boost from recruiting | weak yes | Technical fields seem to have higher screening costs and greater treatment effects on fill rate |
| P4: An increase in expected passive applicants makes recruiting relatively less attractive | moderate yes | True in cross-sections and with employer-specific FE model |
| P5: Increased screening and recruiting in project value | weak 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.

documents the large effect positive that additional recruited applicants can have on vacancy fill rates, making our finding of positive treatment effects on fill rates in the recruiting experiment more explicable. Second, we now have an additional example where the "substitutes" characterization of organic and recruited applicants is incorrect.

## 8 Discussion and conclusion

We demonstrate that recommendations are both acted upon by employers and effective at raising fill rates, at least for technical vacancies. Surprisingly, recommendations do not cause crowd-out effects because organic, non-recruited applicants and recruited applicants appear to be complements. These results show that at least in one labor market context, reducing frictions does not entail an equity versus efficiency trade-off: both were improved. Furthermore, these results come from a context where search frictions are already very low, suggesting that bigger gains are possible in traditional, more friction-prone markets.

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 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. This paper shows that even without this match-setting power, markets can be tilted towards desirable ends through purely informational interventions. Coles et al. (2010) is a recent example of work in this decentralized market design approach.

## 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[6] 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 may submit 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 this information is not available to them.

With the rise of professional networking sites such as LinkedIn, we are now witnessing the unprecedented digitization of the supply side of the labor market. Individuals can create public profiles and list everything that would normally be found on a traditional resume plus many other pieces of data that would not be, such as their professional connections. The importance

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

of this development is not just the volume and accessibility of the data, but that much of it is verified. Verification comes not from a centralized credentialing authority, but by a person's connections vouching for various stated facts and attributes.

LinkedIn in particular has already achieved impressive penetration: according to LinkedIn, as of March 12, 2012, over 160 million people had created profiles.[7] In some industries, a LinkedIn profile is expected of all applicants.[8] In online labor markets, the amount of information available 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 described with rich, match-relevant meta-data that is computer-accessible. If this digitization and instrumentation continues, algorithmic approaches to the labor market matching problem are likely to become commonplace.

## 8.2 Future work

There are a number of directions for future work. One obvious challenge is to improve the quality of labor-related recommender systems. Better machine learning techniques and more and better data are likely to improve recommendation quality.

Aside from the machine learning challenges, there are also fundamentally economic challenges that will require more research. We show that vacancies receive all the applications they will ever receive very quickly. A natural question is why this 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 is unknown. A satisfactory theory of how vacancies accrue applicants as a function of market policies (such as application quotas) and how these policies ultimately affect marketplace efficiency would have great practical application.

Economics also has a role to play in setting platform policies and mechanisms that are information-revealing and incentive compatible. On the mechanism design side, the platforms would like workers to truthfully report their availability for work, yet getting them to do so proves difficult in practice. The strategic issue is that workers have close to free dis-

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

[8] 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."

posal on job offers and hence have little incentive to opt-out of being recommended, despite the externalities this imposes on other workers with the availability to take on further projects. This reporting of availability problem is important because knowing a worker's true availability would allow the platform to ameliorate market congestion and thinness. Because they can act in nearly real-time, the platform could ensure that neither vacancies nor workers are over- or under-recommended—*if* they could predict availability. This flow-management possibility is particularly exciting as congestion and thinness are thought to be important sources of matching friction in labor markets (Petrongolo and Pissarides, 2001).

A key practical problem for a platform is deciding how to allocate visibility via recommendations and position in search to maximize some objective function. This objective function captures the platform's preferences over total matches formed and match quality. The optimization problem is dynamic and the platform cares about factors other than the instantaneous match rate. For example, real platforms care about the integration of new users and equity among platform participants. Currently, there is no satisfactory answer to this visibility allocation question.

### 8.3 Acknowledgments

Thanks to Karim Lakhani, Misiek Piskorski, Erik Brynjolfsson and other participants at the Digital Business Seminar at Harvard Business School for helpful comments on earlier versions of this work. I am grateful to to the oDesk Research and oDesk Product teams members, particularly Yannis Antonellis, Panos Ipeirotis, Ramesh Johari, Hayden Brown, Michael Levinson, Panagiotis Papadimitriou and especially Odysseas Tsaltos for encouragement, ideas and assistance. Larry Katz, Richard Zeckhauser, Luke Stein, Jacob Leshno, Mandy Pallais, Tal Gross, Peter Coles and especially Dana Chandler offered numerous helpful comments and suggestions. Thanks to Robin Yerkes Horton, Carolyn Yerkes, Ada Yerkes Horton and David Yerkes with help in preparing the manuscript.

In preparing this paper, I used numerous open source packages and tools. I am especially grateful for the tools created by Bates and Sarkar (2007), Wickham (2009) and Rossini et al. (2004).

## 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.

**Bates, D. and D. Sarkar**, "lme4: Linear mixed-effects models using S4 classes," 2007.

**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.

**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.

**Rossini, A.J., R.M. Heiberger, R.A. Sparapani, M. Maechler, and K. Hornik**, "Emacs speaks statistics: A multiplatform, multipackage development environment for statistical analysis," *Journal of Computational and Graphical Statistics*, 2004, *13* (1), 247–261.

**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.

**Wickham, H.**, *ggplot2*, Springer, 2009.

**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. This batch versus serial characterization is presumably a false dichotomy, with real firms doing both or some amalgam of strategies (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.

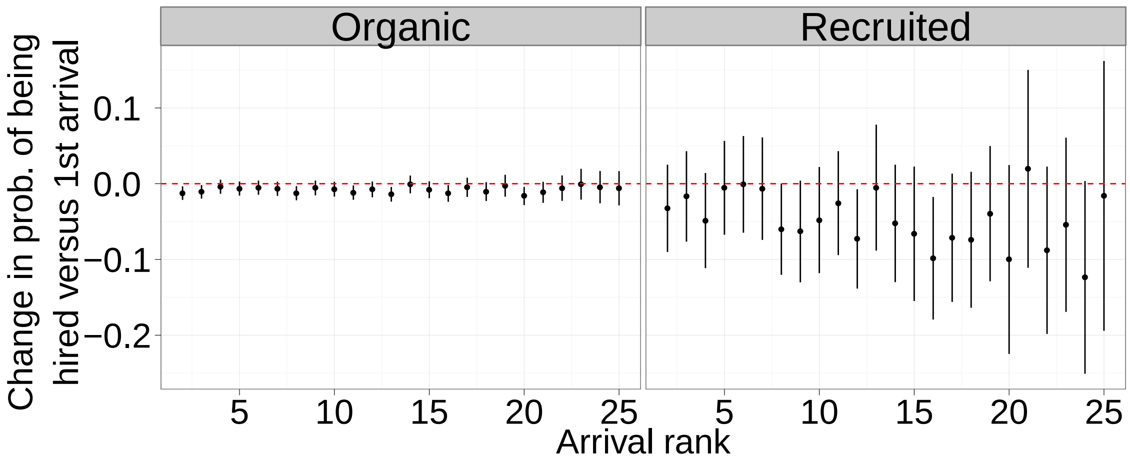
## A.1 Arrival rank does not predict of hiring

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 $i$th applicant. We cluster standard errors at the vacancy level. The regression we estimate is:

$$\text{hired}_{ij} = \sum_{k=1}^{C} \beta_k \cdot 1\{rank(i) = k\} + l_j + \epsilon_{ij} \tag{9}$$

Figure 6, we plot $\beta - \hat{\beta}_k$ for $k = 2 \ldots 16$. We can see that none of the coefficients are distinguishable from zero.

Figure 6: Estimated probability of being hired as function of arrival rank




*Notes:* These plots shows the estimated $\hat{\beta}_k$ coefficients from Equation 9 and a 95% confidence interval for the point estimate. The dependent variable is whether or not an 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 dataset consists of all applications made to vacancies in the recommendation experiment, restricted to those vacancies that received less than 25 applications and more than 5 applications. These coefficients are the mean differences 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 10: 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 all the applications sent to vacancies in the recommendation experiment, but 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 not less likely to be viewed by the 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 application. Here "viewed" means opening up the application in the web interface, similar to how one would open an email. The independent variable is dummy variable $\text{after}_{ij}$, which indicates that an applicant arrived immediately after the hired applicant. We also include a vacancy-specific fixed effect. The regression equation is:

$$\text{viewed}_{ij} = \beta \cdot \text{after}_{ij} + l_j + \epsilon_{ij} \tag{10}$$

In Table 10, Column (1), we report the estimate of Equation 10. The coefficient on $\text{after}_{ij}$ is only slightly negative and the 95% confidence interval easily includes 0, implying no difference 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 Effects of the value of filling a vacancy and labor market tightness on recruiting

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 (Proposition 5). A tight labor market (smaller $A_p$) predicts more recruiting (Proposition 4). 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.

**Proposition 4.** *An increase in the number of passive applicants, $A_P$, has a greater positive 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 in Appendix C*

**Proposition 5.** *The optimal number of applicants is increasing in the value of filling the vacancy for both the recruiting and passive strategies: $\partial a^* / \partial v > 0$ where $a^* = a^*_R$ or $a^* = a^*_P$. However, an increase in value increases the appeal of the recruiting strategy more than it increases the appeal of the passive strategy.*

*Proof in Appendix C*

Different projects presumably vary in their value and their ability to 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 variation 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. We do not have random assignment of projects to employers, 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.

We construct a dataset of 677,408 vacancies posted on oDesk from January 2011 to September 2012. We use this data to estimate Equation 11 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, $\delta_j$, and fixed effects for the category of work, $\gamma_{ij}$.

$$y_{ij} = \beta_A \log A_P^{ij} + \beta_H \log H_{ij} + \beta_V \log V_{ij} + \delta^j + \gamma_{ij} + \epsilon_{ij} \tag{11}$$

## B.1 Effects of project value and labor market tightness of probability of recruiting

In Table 11, we report the results of estimating Equation 11 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 realized wage bill and estimate $\beta_V$.

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 significant in all regressions except the Column (4)—though the magnitude of the point estimate is largest in this regression. The actual wage bill regressor, $\beta_Y$ 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. By using an employer-specific fixed effect, we can net out this strategic effect.

In the Column (4) regression, we the dataset is restricted to filled vacancies. In this regres-

Table 11: Project value, organic applications, wage bill and the probability of recruiting

| | Employer sent early recruiting 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 the full sample of vacancies (677,408 ) vacancies posted on oDesk from January 2011 to September 2012, 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 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: $p \le 0.05 : *$, .$p \le 0.01 : **$ and $p \le .001 : ***$.

sion, 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 the wage bill as in some sense exogenous, then this regression is more reasonable.

In every specification of Table 11, 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.

## B.2 Effects of project value and labor market tightness on number of screened applicants

In Table 12, we regress the number of screened *organic* applicants on the same set of regressors used in Equation 11. We use this outcome 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. The results imply that of the two effects of more applicants—greater per-applicant quality but greater likelihood of finding a match with fewer applicants—the "substitution" effect dominates: when more organic applicants are available, employers screen more applicants.

## B.3 Effects of labor market tightness of hires

In Table 13, 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.

Table 12: 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 of organic applicants conducted by an employer. In each regression, we include the log of the project size in hours (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 the full sample of vacancies (677,408 ) vacancies posted on oDesk from January 2011 to September 2012, 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 sample consists of vacancies with one or more organic applications; 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 sample restriction allows us to include the ultimate wage bill for the project as a regressor, though 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 $p \le .001 : ***$.

Table 13: 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 the full sample of vacancies (677,408 ) vacancies posted on oDesk from January 2011 to September 2012. 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: $p \le 0.05 : *$, $.p \le 0.01 : **$ and $p \le .001 : ***$.

# C Technical Proofs

## C.1 Proposition 3

*Proof.* The increase in fill rate from switching from the passive to the recruiting strategy is $h_R(a_R^*) - h_P(a_P^*)$.

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

The gap between the hiring probability is increasing in $c$ if $\partial a_R^* / \partial a_c < \partial a_P^* / \partial c$. The rate of change in $a^*$ due to changes in screening costs actually depends upon the curvature of the hiring function: regardless of strategy, the first order condition is $v h'(a) = c$ and so $a'(c) = \frac{1}{h''(a)}$. If the gap is increasing in $c$, then $a_R' > a_P'$ or $|a_R'| < |a_P'|$ (since the number of screened candidates is decreasing in $c$), which implies that $|h_R''| > |h_P''|$.

We now obtain an expression for $h''$ in terms of the underlying distribution of worker match probabilities and the number of applicants. Starting with the equation $h'(a) = (1 - h(a))q(a)$, we can differentiate with respect to $a$ and get $h'' = q - qh = q' - (q'h + qh') = (1 - h)q' - qh'$.

Multiplying through by $q/q$ and replacing $h' = c$, we have $h'' = c(q'/q - q)$.

For $|h''_R| > |h''_P|$, it must be the case that $|q'_R/q_R| + |q_R| > |q'_P/q_P| + |q_P|$, or $|q_R| - |q_P| > |q'_P/q_P| - |q'_R/q_R|$. Let $\Delta = |q_R| - |q_P|$.

We know that $q_R > q_P$, but we have not placed any conditions of $q'_R$ or $q'_P$, which will be needed for the inequality to hold. If $|q'_R/q_R| \geq |q'_P/q_P|$, then the inequality holds. In the case where $|q'_P/q_P| > |q'_R/q_R|$, for a given value of $|q'_P/q_P|$, $|q'_P/q_P| - |q'_R/q_R|$ would be maximized if $|q'_R/q_R| = 0$. In this extreme case, the inequality would hold if $\Delta > |q'_P/q_P|$. Recall that $q' = -\frac{1}{Af(q)}$, and so $|q'/q| = \frac{1}{qAf(q)}$. As such, the inequality holds so long as $\bar{A} > \frac{1}{q\Delta f(q)}$. □

## C.2 Proposition 4

*Proof.* At the optimum for each strategy, $h'_R(a^*_R) = h'_P(a^*_P)$. Because of Proposition 2 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 $a^*_P$ evaluated applicants than in $a^*_R$ evaluated applicants. 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. □

## C.3 Proposition 5

*Proof.* At the optimum number of screened applicants, $vh'(a^*) = c$ (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 $\forall a, h'(a) > 0 \land h''(a) < 0$, $\frac{\partial a^*}{\partial v} > 0$. The partial derivative of the effect of an increase in $v$ on expected utility is:

$$\begin{aligned} \frac{\partial \pi(x=R;v)}{\partial v} &= h_R\left(a^*_R\right) + \left(vh'_R(a^*_R) - c\right)\frac{\partial a^*_R}{\partial v} \\ &= h_R\left(a^*_R\right) \qquad \text{(by the envelope theorem)} \end{aligned} \tag{12}$$

By the same argument, $\frac{\partial \pi(x=P;v)}{\partial v} = h_P(a^*_P)$, and so by Proposition 2, $\frac{\partial \pi(x=R;v)}{\partial v} > \frac{\partial \pi(x=P;v)}{\partial v}$. □