Why the Referential Treatment? Evidence from Field Experiments on Referrals

Amanda Pallais and Emily Glassberg Sands*
Harvard University and NBER; Harvard University

November 10, 2013

Abstract

This paper presents the results of three field experiments in an online marketplace designed to determine whether referred workers perform better than non-referred workers and, if so, why. We first hired experienced workers and asked them to refer other workers; we then hired all referred and non-referred applicants that met our basic criteria. We find that referred workers have significantly better observable characteristics than non-referred workers, but perform substantially better and have less turnover even conditional on these characteristics. We then consider three potential explanations for this performance difference: selection (referred workers would perform better than non-referred workers even if they had not been referred), peer influence (referred workers perform better because they believe their performance will affect their referrer's employment outcomes or their relationship with their referrer) and team production (referred workers perform better because they are working with their referrer). We find that team production is important: referred workers are much more productive when working with their own referrer than with someone else's referrer. However, we also find a large role for selection: referred workers perform better and persist longer even at jobs to which they are not referred at a firm where their referrers do not work. A referral contains information about worker quality that is not present in the workers' resume. This information is more valuable (the referral performs better on average) when the referrer is more productive and when the referral is closer with her referrer.

*We would like to thank Kehinde Ajai, David Autor, Felipe Barrera-Osorio, Raj Chetty, Melissa Dell, David Deming, Itzik Fadlon, Adam Guren, John Friedman, Roland Fryer, Edward Glaeser, Claudia Goldin, Josh Goodman, Rick Hornbeck, Lisa Kahn, Lawrence Katz, Ben Schoefer, Sarah Turner, Marty West, and seminar participants at Berkeley, Kellogg, Harvard, Princeton, RAND, University of British Columbia, and NBER Summer Institute 2013 Labor Studies for their many helpful comments and suggestions. We would like to thank John Horton and the oDesk Corporation for help running the experiment. Sophie Wang provided excellent research assistance. Financial support from the Lab for Economic Applications and Policy at Harvard is gratefully acknowledged. Amanda Pallais's email address is apallais@fas.harvard.edu and Emily Glassberg Sands's email address is esands@fas.harvard.edu.

1 Introduction

A large empirical literature has shown that many workers find jobs through networks (e.g., Bewley, 1999; Ioannides and Datcher Loury, 2004; Granovetter, 1995). A consensus estimate is that at least half of jobs are found through informal contacts (Topa, 2011). Theoretical literature (e.g., Calvo-Armengol and Jackson, 2004; Montgomery, 1991) suggests that the use of referrals may disadvantage workers without labor market connections; consistent with this, empirical findings show that applicants who are not referred by current employees are much less likely than referred applicants to receive an offer (e.g., Fernandez and Weinberg, 1997; Brown et al., 2012; Peterson et al., 2000).

Yet the prevalence of referrals suggests that firms likely benefit from their use. Existing empirical work finds that referred workers have less turnover than non-referred workers (e.g., Brown et al., 2012; Holzer, 1987; Simon and Warner, 1992; Datcher, 1983). It remains divided, however, on whether referred workers are more productive. A few studies directly compare the performance of referred and non-referred workers working at the same or very similar firms: Castilla (2005) finds that referred workers perform better, Blau (1990) finds that they perform worse, and Burks et al. (2013) finds that referred workers perform better, but only on a few metrics. Other papers use wages or promotion rates to proxy for the performance of referred and non-referred workers; their findings are similarly mixed (e.g., Dustmann et al., 2011, Simon and Warner, 1992; Brown et al., 2012; Pistaferri, 1999; Bentolila et al., 2010).

In this paper, we present the results of three field experiments in an online marketplace. We first analyze whether referred workers perform better and have lower turnover than non-referred workers.[1] Then, finding that referred workers do perform better than non-referred workers, we examine why.

Workers in all of our experiments were recruited the same way. We first hired experienced workers and solicited referrals from them; we then hired their referrals and all non-referred applicants who met our basic wage criteria. Whereas most of the existing empirical literature compares the performance of referred and non-referred *hires* (where hires are of course a subset of all applicants), our design facilitates comparisons between referred and non-referred *applicants*. Our motivation for including all applicants is simple: suppose a firm knows that referred applicants are on average more productive than non-referred applicants; a rational firm would incorporate this information into its hiring decisions such that, in order to be

[1] There are other reasons firms might benefit from hiring referrals. For example, referrals might decrease the cost of recruiting or be a perk to existing (referring) workers. In this paper, we abstract away from these other potential benefits and focus on productivity and turnover differences between referred and non-referred workers.

hired, a non-referred applicant would have to look relatively better on other characteristics. Because of such differential selection of referred and non-referred workers into employment, *hired* referred workers may not perform any better than *hired* non-referred workers, even when referrals provide positive information about worker quality.

In all three experiments, we find that referrals contain information about worker quality that is not contained in workers' resumes. Referred workers performed better and stayed longer than non-referred workers. They also had better observable characteristics. But even controlling for these characteristics, they outperformed and outlasted their non-referred counterparts.[2]

The heart of this paper lies in assessing three potential explanations for performance and turnover differences between referred and non-referred workers.[3] The first explanation, *selection*, says that a referred worker would perform better and stay longer at the firm even if she had not been referred. This may be, for example, because high-ability referring workers also have high-ability friends (e.g., Montgomery, 1991; Granovetter, 1995; Rees, 1966), or because workers have information about their friends' abilities and persistence and select relatively productive and persistent contacts to refer (e.g., Beaman and Magruder, 2012; Fernandez et al., 2000; Rees, 1996).[4]

The second and third explanations, in contrast, emphasize how on-the-job interactions with their referrers might affect the productivity and turnover of referred workers. In the second, *peer influence*, a referred worker exerts more effort and stays on at the firm longer when working at the same firm as her referrer because she believes her performance and persistence will affect her friend's employment outcomes and/or their relationship. Consistent with this explanation, Heath's (2013) model suggests that referred workers work hard because if they perform poorly, the firm will punish their referrers through lower wages and Kugler's (2003) model assumes referrers directly exert peer pressure on their referrals to perform well.[5]

In the third explanation, *team production*, a referred worker performs better and may

[2] The online marketplace in which these experiments takes place is a unique setting in that we see workers' entire resumes. Because interviews are relatively uncommon and workers and employers do not meet face-to-face, we observe most characteristics that an actual employer would observe when making its hiring decisions.

[3] There is little empirical evidence on this question. A notable exception is the concurrent paper Burks et al. (2013), which uses firms' administrative data to address this question. We return to this paper in the conclusion, where we reconcile our findings with its findings.

[4] Our experiments were designed so that referring workers had no information about the job itself at the time they submitted their referrals. Our results thus speak to selection on, for example, general observable and unobservable characteristics, and not to selection on match quality (i.e., workers referring friends who would be a good fit for the particular job).

[5] This explanation is related to group lending in microfinance wherein a worker's peers may pressure the worker to repay the loan (e.g., Bryan et al., 2012).

enjoy her job more when working directly with her referrer. While this explanation for referrals' positive performance has not been emphasized to the same extent in the economics literature, general research on team production indicates it may be an important benefit of referrals. Bandiera et al.'s (2012) model, for example, suggests that when working in teams with their friends, workers both receive more utility and are less likely to free-ride. Furthermore, Bandiera et al. (2005) finds that workers are more able to cooperate with their teammates when their teammates are friends; and Costa and Kahn (2003, 2010) find that Civil War soldiers were less likely to desert and were more resilient to job-related stress when more of their unit was from their own birthplace.

In testing these three explanations for why referred workers outperform and outlast non-referred workers, we find that selection is important. Even when working at a different firm to which they were not referred and with which their referrers were not affiliated, the referred workers in our experiment exhibited higher performance and lower turnover than their non-referred counterparts.

On-the-job interactions between referrers and their referrals, especially in team tasks, are also important in explaining the performance differential. While we see only limited evidence of peer influence on the job, we find substantial evidence of team production. A referred worker performed much better when working with her own referrer than when working with an equally qualified worker she did not previously know. When working with her own referrer, she also worked longer on the task and was more likely to report wanting to continue on that team.

All three field experiments (the "individual," "supplemental," and "team" experiments) took place on oDesk, the largest online labor market, with over 2.5 million workers (Horton, 2013) and 35 million hours billed in 2012 (oDesk Corporation, 2013). In each of the two main experiments (the individual and team experiments), we first hired moderately- and highly-qualified workers with experience in the marketplace. Once hired, these workers completed a short task unrelated to the experimental tasks and were then asked to refer other workers to our firm; 40 percent provided a referral. We then simultaneously invited the referred workers and a random sample of non-referred workers to apply. We hired all referred and non-referred applicants who met our basic wage criteria and asked the three types of hires (referrers, referred workers, and non-referred workers) to perform either an individual task (in the individual experiment) or a team task (in the team experiment). In the supplemental experiment, we re-hired the referred and non-referred workers from the individual experiment for a different individual task at a different firm.

The individual experiment distinguishes between selection and peer influence. All referred and non-referred workers in this experiment completed a short task every other day over the

course of twelve days; their task was to test an airline flight website by answering a few questions about the flights listed on the site.[6] Referrers were simultaneously completing a different task over the same period. At the end of the twelve days, all workers were asked to indicate if they would like to continue with the firm.

Before beginning the task, referrers were randomized, along with their referrals, into one of two treatment groups. Treatment 1 was designed to maximize peer influence. In this treatment, the referrer was sent an update about her referral's performance after each day of work and the referred worker knew her referrer would be receiving these updates. We also told each referrer in this treatment that she was being considered for a promotion and implied that her referral's performance and willingness to continue working for us would influence her likelihood of promotion. Treatment 2, in contrast, was designed to minimize peer influence. In this treatment, the referred worker was told her referrer would never know how she performed. Though each referrer here still knew she was being considered for a promotion, she was told explicitly that she would be judged on her own merits and *not* on the performance of her referral. The experiment was designed so that the difference in referred worker performance between Treatments 1 and 2 would identify the effects of peer influence while the difference in performance between non-monitored (Treatment 2) referred workers and non-referred workers would identify the effects of selection.

We find in the individual experiment that referred workers performed much better than non-referred workers and had substantially lower turnover. For example, non-monitored (Treatment 2) referred workers submitted work 76 percent of days, while non-referred workers submitted a response only 63 percent of days. Moreover, 95 percent of non-monitored referred workers indicated they wanted to continue the task as compared to only 73 percent of non-referred workers. These effects remain large and significant when we control for the differences in observable characteristics between referred and non-referred workers. Not only are referred workers' superior performance and persistence not predicted by observable characteristics, referred workers perform better on the last day of the contract than would be predicted by their performance on all of the prior days and they persist longer than would be predicted by their job performance on the entire contract. These results suggest that the referral provides information to employers that could not easily be obtained through observables or initial job performance.

We do not find evidence that peer influence had large effects on workers' productivity or persistence. Relative to non-monitored (Treatment 2) referred workers, monitored (Treatment 1) referred workers performed slightly better, but the difference was not statistically

[6] The tasks for all three experiments were chosen to be similar to tasks that are common on oDesk. In particular, many jobs on oDesk require visiting websites and answering questions about them.

significant. Moreover, monitored referred workers were, if anything, slightly *less* likely to want to continue working for the firm, perhaps because they disliked being monitored.

Since even non-monitored (Treatment 2) referred workers may have faced some subtle peer influence in the individual experiment, we ran a supplemental experiment four months later to isolate the effects of selection. Here, we made job offers to all referred and non-referred workers from a new firm that had no affiliation with the firm from the individual experiment; importantly, this new firm had no contact with any of the referrers. The task in the supplemental experiment, moreover, was designed to be credibly different from that in the individual experiment, though it similarly measured both diligence over time and willingness to stay on at the firm.

The supplemental experiment provides the strongest evidence that selection is a key driver in the superior performance and persistence of referred workers. Even at a firm to which they had not been referred and at which their referrers did not work, referred workers exhibited substantially better performance and lower turnover than non-referred workers. The effects are generally large and significant regardless of whether we restrict attention to workers who accepted our offer of employment.

Our third experiment, the team experiment, isolates the effect of team production. Workers were recruited for this experiment exactly as they were in the individual experiment except that any workers involved in the individual experiment were excluded. The task was to work with an assigned partner to create a single, shared slogan for a public service announcement (PSA). Each of the two partners was given a different information sheet containing a distinct criterion for the slogan (e.g., be written in all capital letters, be exactly three words long). We asked the partners to use the chat box provided on the site to discuss their slogan and then to each submit the same slogan, which should have satisfied the criteria on both information sheets.

Workers in the team experiment completed three of these PSA tasks, each with a different partner. Every referrer participated exactly once in each of three different team types: a Type A team in which she was paired with her own referral, a Type B team in which she was paired with someone else's referral, and a Type C team in which she was paired with a non-referred worker. We measured performance in each pairing and, after all three PSA tasks had been completed, also asked workers with which partner(s) they would want to work again.

As in the individual experiment, referred workers outperformed non-referred workers, even when both types of workers were assigned partners they did not previously know. When each working with someone else's referrer, the average referred worker (on a Type B team) was 31 percentage points more likely than the average non-referred worker (on a

Type C team) to submit any slogan and 14 percentage points more likely to submit the same slogan as her partner.

A referred worker performed better still when paired with her own referrer. Though she was only slightly more likely to actually complete the task when matched with her own referrer than with someone else's referrer, she was more than twice as likely to submit the same slogan as her partner and nearly twice as likely to have that slogan also meet both criteria. She worked longer on the task when paired with her referrer and was much more likely to report wanting to work again with her referral than with her other partners. These results suggest team production is an important benefit of hiring referrals.

Across the three experiments we find that referrals provide (positive) information about worker performance on top of workers' observable characteristics; but not all referrals are created equal. We find that those referred by high-performing workers perform particularly well themselves. Part of this can be explained by a tendency among referrers to refer workers with observable characteristics similar to their own: referrers with stronger resumes on average provide referrals with stronger resumes. But even controlling for workers' observable characteristics, those referred by high-performers tended to perform better themselves.

We also explore the relationship between the strength of the referrer-referral tie and the performance of the referred worker. At the time of referral, we asked the referrer three questions about her relationship with her referral: how well she knows her referral, how many friends they have in common, and how often they interact.[7] When a worker refers someone with whom she is not as close (a weak tie), she tends to refer someone who looks better on paper. Nevertheless, it is the worker who has a stronger tie to her referrer who performs better, even before controlling for observable characteristics.

Finally, we use our experimental data to show that if we had only compared the performance of those referred and non-referred applicants whom employers had actually chosen to hire, we could have obtained misleading results about the information contained in a referral. We first simulate which of our applicants employers would choose to hire if they observed both resumes and referral status, and knew the relationship between resumes and referral status and performance. Because referred workers performed much better than non-referred workers conditional on observable characteristics, employers would hire relatively few non-referred workers. The non-referred workers hired would be positively selected on observables. Then, we compare the actual performance of the workers hired in this scenario. We find that, even though the referral contained important information about worker quality, the *hired* referred and non-referred workers would not perform significantly differently

[7] One caveat to this analysis is that these relationship characteristics are self-reported by the referrer before the referral has been hired.

from each other.

The rest of the paper is structured as follows: Section 2 describes the online marketplace, explains the sample selection for our experiments, and provides descriptive statistics about our sample. The three subsequent sections describe the design and results from the individual experiment (Section 3), the supplemental experiment (Section 4), and the team experiment (Section 5). Section 6 analyzes how referrers' performance and the referrer-referral relationship predict the performance of the referred worker. Section 7 shows that the comparison of referred and non-referred workers' performance could be biased if we only observed the performance of workers employers chose to hire. Section 8 concludes.

2 Experimental Context and Recruitment Design

2.1 Online Marketplace

oDesk is an online marketplace where employers, mostly from the United States, hire independent contractors from all over the world for tasks such ranging from software development to administrative support. These workers are accustomed to developing long-standing relationships with firms. Prior to our experiment, the average job taken on by the referrers in our experiment paid $237 and lasted 81 working hours.

oDesk workers create online resumes with their skills and qualifications. Resumes typically include previous oDesk jobs, a one-to-five feedback score from these jobs, and a suggested hourly wage. Many also list other qualifications such as degrees held and oDesk tests passed. Figure 1 shows a sample oDesk resume.[8]

Employers, meanwhile, post job listings. They choose whether they want any oDesk worker to be able to apply and/or whether they want to invite a specific set workers to apply. They also decide which of the two oDesk job types they want to offer: hourly jobs or fixed-wage jobs. Hourly jobs are the more popular type of job and the type used throughout our experiments. In these jobs, workers propose an hourly wage when they apply. Employers choose which workers to hire. Hired workers track the time they are working and oDesk monitors that they are actually working during these periods by taking screenshots and analyzing keystroke volume. Workers are then paid their set hourly wage for the time worked regardless of output quality.[9]

Many workers have friends and relatives who also work on oDesk. Though there is at present no explicit referral mechanism on oDesk, employers can solicit referrals from their

[8] This worker was not in our experiment.

[9] In contrast, in fixed wage jobs workers and employers agree on a price for the entire job and employers have discretion at the job's end over how much to actually pay.

current workers and workers can recommend people they know to their employers.

2.2 Recruiting Workers

We recruited workers for the individual and team experiments in the same way. We first invited a random sample of oDesk workers who (1) were from the Philippines, (2) listed an hourly wage of $5 or less on their resume, (3) had earned $50 or more on oDesk, and (4) had an average job feedback score of four or higher to apply to our job.[10] We told these workers very little about the task, only that we were hiring "for a variety of ongoing administrative support tasks of varying durations" and that we were looking for "diligent and highly-motivated individuals who are competent in the English language and interested in an ongoing relationship with our firm." We also told them that the position came with the possibility of promotion to managerial roles. We gave workers 48 hours to apply and then hired all workers who applied at an hourly wage of $3 or less.

We told these original hires that they would need to visit our website to initialize the job. The initialization step was intended to give workers some connection to our firm and to weed out the least responsive workers. (We fired the 5 percent of workers who did not initialize.) We then asked the workers who initialized to refer up to three other oDesk workers who were "highly-qualified" and whom they thought would "do a good job and be interested in an ongoing relationship with our firm." On each referral form we included questions about how well the referrer knew her referral, how often they interacted (remotely and/or or in person), and how many people they knew in common. We also asked if they ever worked in the same room. Since referrers might have more easily monitored and/or collaborated with referrals working in the same room, we eliminated from our sample any referral who ever worked in the same room as her referrer.

We then invited to our job all referred workers who listed an hourly wage of $5 or less; all workers who were referred were located in the Philippines. We simultaneously invited to our job a random sample of oDesk workers from the Philippines with hourly wages of $5 or less.[11] We again gave workers 48 hours to apply and then hired all referred and non-referred

[10] We only included workers from the Philippines because we wanted all workers in the team task to be able to communicate easily and the Philippines is the most common country of residence for low-wage oDesk workers doing these types of jobs. We eliminated workers with ratings below four because we only wanted referrals from workers we would actually hire; because most oDesk ratings are very positive, only 16 percent of workers who met our other criteria had ratings below four.

[11] We eliminated from the pool of both referred invites and non-referred invites any workers who had already been invited as a potential referrer. We also eliminated from the team experiment anyone who had been invited to the individual experiment. As a result, referred and non-referred workers invited to the team experiment look worse on observables than do referred and non-referred workers invited to the individual experiment.

workers who applied at an hourly wage of $3 or less.[12]

This recruiting process, used for both the individual and team experiments, produced an experimental sample with three types of workers: referred workers, non-referred workers, and "referrers" (i.e., workers who made a successful referral). Workers who did not refer anyone or who referred a worker we did not hire performed a different, shorter task and are not included in the performance results. In the supplemental experiment, we made job offers to all referred and non-referred workers from the individual experiment; no referrers were included.

2.3 Descriptive Statistics

Panel A of Table 1 describes the characteristics of three groups of workers: (1) workers who referred someone (regardless of whether their referral met our referral criteria or applied for the position), (2) workers who were asked to refer, but referred no one, and (3) included referrers (i.e., workers who made a successful referral). Because of the criteria we had for potential referrers, all three groups are relatively experienced. On average, they have been on oDesk for just under two years, held about 12 jobs, and earned almost $3,000.

Workers who referred someone look somewhat more qualified than those who did not. They tended to be more experienced on oDesk, with 22 percent higher earnings and 14 percent more previous jobs. Workers who referred also tended to post slightly higher wages on their resumes, though they requested about the same wages for our jobs.

Panel B of Table 1 describes the characteristics of three other groups of workers: (1) all referred workers, regardless of whether they met our criteria, (2) included referred workers (i.e., referred workers who met our criteria and applied at a wage of $3 or less) and (3) included non-referred workers (i.e., non-referred workers who were invited and applied at a wage of $3 of less).[13] All three groups look substantially less experienced than the experienced workers in Panel A.[14]

[12] We designed the recruitment process so that when referrers were submitting their referrals, they had no information about our actual tasks. The initialization step, for example, was unrelated to the tasks themselves. From their own invitation to apply and from our request for referrals, referrers did know that we were hiring "for a variety of ongoing administrative support tasks of varying durations" and that we were looking for "diligent and highly-qualified individuals who are competent in the English language and interested in an ongoing relationship with our firm." However, all referred and non-referred workers saw this same description on our job posting. Since referred workers had no private information about the job before referring, in our context there is no scope for selection on match quality.

[13] While we hired all referred workers who met our criteria and applied at a wage of $3 or less, only one (randomly-selected) referred worker per referrer was actually included in the team experiment. Thus, there are some referred workers who applied and met our hiring criteria but are not considered "included referred workers."

[14] After all three experiments, we were required by the Harvard IRB to inform all participants about the study and give them the opportunity to remove their data from our study. We received many responses

Included referred and non-referred workers have very different characteristics. While 64 percent of included referred workers had prior experience, only 28 percent of included non-referred workers did. Referred workers' prior earnings were more than double those of non-referred workers and, conditional on having a feedback score, their feedback scores were significantly higher. Despite being seemingly more experienced and qualified than non-referred workers, referred workers posted wages on their resumes that were over 20 percent *lower* than those posted by non-referred workers, and they proposed significantly lower wages to our job. Referred workers were also much more responsive to our invitation to apply: among those we invited, 68 percent of referred workers applied versus only 6 percent non-referred workers.[15] This suggests referrals are a way to identify workers with good resumes who are interested in the job.

3 Individual Experiment

3.1 Design: Identifying Selection and Peer Influence

The task for the individual experiment was designed to measure referred and non-referred workers' diligence when working alone on a project. We designed our task to emphasize diligence because showing up to work and completing tasks in a timely manner are key determinants of success for low-skilled workers, both in more general labor markets and on oDesk.[16] We also designed the task to measure worker turnover since decreased turnover is emphasized in the literature as a benefit of hiring referrals (e.g., Brown et al., 2012; Holzer, 1987; Datcher, 1983). The treatments were designed to determine the extent to which observed differences in workers' performance and turnover were driven by selection relative to peer influence.

All referred and non-referred workers completed the same task. We told them they would be doing testing for an airline flights website, and asked that they visit the site every other

from participants who were excited to participate. However, one worker who was referred but had been excluded from participating in any of the experiments requested to have his data removed and we have done so. Removing this worker's data only affected the "All Referred Workers" column in Table 1, Panel B and the "Excluded Referred Workers" column in Appendix Table 6.

[15] Despite looking less qualified than referred workers, non-referred workers were positively selected among workers in the marketplace; non-referred workers who applied for our job were much more qualified than the invited non-referred workers who did not.

[16] For example, on oDesk, Pallais (2013) finds that employers care more about whether a worker completed a data entry task by the deadline than the worker's accuracy. In more general labor markets, firms respond to absenteeism by having other employees work overtime, reassigning workers from other jobs, and/or hiring temporary workers. These adjustments are all costly and often require manager time. Moreover, the replacement workers may not be as productive as the absent workers (e.g., Herrmann and Rockoff, 2010).

day for twelve days (six visits total), answering the questions on the site each day. For each worker on each day, the site displayed a table with a randomly-generated set of ten flights. Each flight was identified by a flight number and included a departure and arrival city, price, and number of available seats. Just below the flights table were six fill-in-the-blank questions (e.g., the flight number of the cheapest flight). The questions were the same each day, but the correct answers changed with the set of flights shown. Figure 2 displays a sample flights table followed by the questionnaire.

We told all referred and non-referred workers to complete the task on the assigned day and asked, but did not require, that they complete each day's task by 11am Philippine Time. We also informed all referred and non-referred workers that we would send performance updates to a manager after each working day reporting (1) whether they submitted a response on the assigned day, (2) whether they submitted a response by 11am on that day, (3) whether they answered all the questions, and (4) the percentage of working days they had met each of these three performance criteria. Appendix Figure 1 shows an example performance report.

Referrers were randomized to Treatments 1 and 2. Each referred worker was assigned the same treatment as her referrer. Treatment 1 was designed to facilitate monitoring of the referred worker by her referrer while Treatment 2 was designed to minimize peer influence on the referred worker. Referred workers in Treatment 1 were told that their daily performance statistics would be sent to their referrer as well as the manager. Referred workers in Treatment 2, meantime, were explicitly told that their referrer would never see their performance statistics, only the manager would.

Referrers worked on a different task. We wanted to employ them for the duration of their referrals' contracts and we wanted them to understand the performance metrics we would be sending them about their referrals. Thus, we asked them to answer questions on a website every other day over the same time twelve day period and we assigned them a soft deadline for submitting on each day. We did not, however, want the referrers to garner insights from their own task with which they could potentially help their referrals. Thus, the referrers visited a site that had a different login method, was focused on consumer products rather than flights, and asked a different set of questions. Referrers also had a different soft deadline (2pm Philippine Time).

To strengthen the treatment, we told all referrers before work began that they were being considered for a higher-paying management position. We implied to referrers in Treatment 1 that their own employment prospects with the firm were dependent on their referrals' performance.[17] Referrers in Treatment 2 were also informed of the management position,

[17] All referrers were told that the management position would require being able to identify "high-ability workers interested in an ongoing relationship with our firm." When we told referrers in Treatment 1 about

but were assured that they would be "judged on their own merits" and that the performance of their referral would in no way influence the promotion decision. As promised, we sent the performance statistics of each referred worker in Treatment 1 to her respective referrer. We also sent referred and non-referred workers' statistics to a manager we hired.

At the end of the task, we invited all referred and non-referred workers to re-apply to continue on the same project. We use this as an (inverse) measure of worker turnover. Re-application updates by worker type and treatment mirrored performance updates. Each referred and non-referred worker was told that the manager would receive an update on whether they accepted our offer to re-apply. Referred workers in Treatment 1 were told this update would also go to their referrers while referred workers in Treatment 2 were explicitly told their referrers would not see this information. To strengthen the treatment, when we invited referrers in Treatment 1 to apply for the management position, we told them that we had just invited their referrals to continue on with their task and hoped their referrals would accept the invitation. We invited referrers in Treatment 2 to apply for the management position as well, but made no mention at all of their referrals.

3.2 Performance and Persistence by Worker Type and Treatment

Table 2 shows that the randomization produced balanced samples between Treatment 1 and Treatment 2 within both the referrer and referred worker samples. Out of 26 comparisons between the two treatments groups, only one difference is significant at the 10 percent level.[18]

Table 3 shows how monitored referred workers (Treatment 1), non-monitored referred workers (Treatment 2), and non-referred workers compare on three measures of performance and our measure of persistence with the firm (the inverse of turnover). Each column presents the results of regressing an outcome on an indicator for being a referred worker in Treatment 1 (a referred worker monitored by her referrer) and an indicator for being a non-referred worker. The omitted group is referred workers in Treatment 2 (non-monitored referred workers). The first three metrics are the performance metrics. Two are metrics the workers were told the manager would see daily: an indicator for submitting any response on a given day and an indicator for submitting the response by 11am.[19] The third performance metric is the fraction of questions answered correctly. (Non-responses are marked as incorrect.) Each observation is a worker-day and standard errors are clustered at the worker level. The

the position, we also said that they would receive daily performance updates on their referrals "because we care about workers' performance."

[18] While there are 28 comparisons in the table, by construction, there is no variation in prior experience nor in having a feedback score among referrers.

[19] Workers were told the manager would also see whether they answered all questions, but since 99.8 percent of submissions were complete, for parsimony we omit this metric from our tables.

dependent variable in the final column is an indicator for whether the workers responded to our invitation to continue working at our firm after the twelve days of initial work.

Referred workers consistently outperformed non-referred workers even when the referred workers were not monitored. Non-monitored (Treatment 2) referred workers submitted responses on just over three-quarters of assigned days; non-referred workers were 13 percentage points less likely to submit. And while those same referred workers submitted on-time 56 percent of assigned days, the corresponding statistic for non-referred workers was only 48 percent. Referred workers were also much more likely to want to continue with our firm. While almost all referred workers in Treatment 2 (95 percent) wanted to continue working with us, only 73 percent of non-referred workers expressed a desire to stay at the firm.

Across the three performance metrics, the coefficients on the Treatment 1 dummy suggest that peer influence may lead referred workers to perform better still. Anecdotal evidence suggests that referred workers in Treatment 1 were, in fact, monitored by their referrers. Many Treatment 1 referrers replied to our daily performance reports and indicated a strong interest in their referrals' performance. They often apologized when their referrals had not completed the task on the preceding day and/or had not completed it by the soft deadline, and assured us they would encourage their referrals to do better on subsequent days. Nonetheless, all of the performance differences between referred workers in Treatment 1 and referred workers in Treatment 2 appear smaller than the differences between the referred workers in Treatment 2 and non-referred workers, and none is significant. The negative (though again insignificant) coefficient on the Treatment 1 dummy in the final column suggests that referred workers in Treatment 1 were, if anything, slightly less likely to be interested in continuing with the firm, perhaps because they disliked being monitored.

The above results are from regressions with no controls (Panel A). Although referred workers have more positive observable characteristics than non-referred workers, Panel B shows that the referral still has substantial predictive power even conditional on these characteristics.[20] Throughout the paper, we use the covariates listed in footnote 20 as our main controls. However, in appendix tables, we show that the our results are robust to adding the squares of each of the (non-binary) covariates and the interaction of each pair of covariates to the regressions. (We do not use the results with the second order controls as our main specifications because estimating coefficients on 88 control variables increases our standard

[20] The observable characteristics included in the regressions are as follows: an indicator for having any oDesk experience, total oDesk earnings, the number of previous oDesk assignments, oDesk feedback score, an indicator for not having a feedback score, the wage listed on the worker's resume, the number of days since joining oDesk, an indicator for having a portfolio, the number of oDesk tests passed, the self-reported English skill level, an indicator for not reporting an English skill level, an indicator for being affiliated with an agency of oDesk workers, and the number of degrees listed on the resume.

errors to make some results lose significance.) Panel A of Appendix Table 1 shows when these second order controls are added to the regressions in Table 3, two of the coefficients on the non-referred dummy increase and two decrease, though none changes significantly. The explanatory power of the regressions also increase. Unsurprisingly given random assignment, adding control variables does not affect the estimated differences between monitored and non-monitored referred workers.

Referrals provide the firm with information about worker quality; firms might also get information about a worker's quality by hiring the worker for some trial duration and observing her performance directly. Longer trials almost certainly provide better information, but at a direct wage cost to the firm. In Panel C, we ask whether the referral still has predictive power for worker performance on the last day of the contract, conditional on worker performance on all prior days. In the first three columns, we regress a worker's performance on the last day of the contract on a dummy for being a monitored referred worker, a dummy for being a non-referred worker, and our control variables. We also control for the worker's performance (on the same metric as measured by the dependent variable) on each of the first five days. Because a worker's performance is highly correlated across days, controlling for her performance on the first five days increases the predictive power of the regressions. But all differences in performance between referred and non-referred workers remain large and significant when considering only the last day of work and controlling for performance in the first five. In fact, the measured performance differences between referred and non-referred workers are slightly, though not significantly, larger when controlling for prior performance.

The referral also provides information about worker persistence at the firm above and beyond the information provided by performance throughout the full contract. In the final column of Panel C, we regress a dummy for wanting to continue with our firm on indicators for being a monitored referred worker and a non-referred worker, controlling for each of our performance measures (submission, on-time submission, and accuracy) on each of the six days, and include again controls for workers' observable characteristics. Even controlling for all our performance measures on all days, referred workers were 18 percentage points more likely than non-referred workers to want to continue on with the firm.[21] Panel B of Appendix Table 1 shows that these results are robust to adding the second order controls.

Figure 3 shows worker performance over the course of the experiment by worker type and treatment. Submission rates of referred workers are consistently higher than those of non-referred workers. Both types of workers become less diligent over time, but diligence falls off much more for non-referred workers. Thus, the performance gap between referred

[21] Unreported coefficients in the final column of Panel C show that workers who performed better were more likely to want to continue.

and non-referred workers grows with time.

The performance gap between Treatment 1 and Treatment 2 referred workers, meantime, is less stark (Figure 3). On the first day of work, before any performance reports had been sent out, monitored and non-monitored referred workers performed equivalently. The graph suggests that peer influence may have stemmed the drop-off in performance in days two, three, and four among Treatment 1 referred workers, though the differences between monitored and non-monitored referred workers on these days is not significant. By day six, however, monitored referred workers were no more likely than their non-monitored counterparts to submit.

Taken together, the results of the individual experiment suggest that selection is important. Even when referred workers are not monitored by their referrer, they perform much better than non-referred workers and are more eager to continue on with the firm. The referral, moreover, contains information that is not present on a worker's resume or in her performance on the majority of her contract. In contrast, we do not find robust evidence in favor of peer influence, though we cannot rule out the presence of peer influence, particularly at the beginning of the contract.

4 Supplemental Experiment

4.1 Design: Isolating Selection

Even though all referrers in Treatment 2 of the individual experiment were assured that they would be judged only on their own merits, and all referred workers in this treatment were assured that their referrers would not see their performance statistics, these referred workers may still have been influenced by the presence of their referrers at the firm. They may have, for example, felt grateful for having been referred or faced informal pressures from their referrers, either of which could affect their performance or persistence in the individual experiment. The supplemental experiment was designed eliminate any such potential influences.

In the supplemental experiment, we measured the performance and persistence of referred and non-referred workers in a job to which the "referred workers" had not been referred. Four months after the individual experiment, we created a firm with a different name, location, job posting, and writing style from that of the individual experiment. None of the referrers was contacted by this firm. To minimize selection, we sought to hire the maximum possible number of referred and non-referred workers. We made direct job offers to all referred and non-referred workers from the individual experiment and sent to workers who had not yet

responded three reminders to accept.

Workers who accepted were given a task that, like the task of the individual experiment, measured individual diligence over time. They were asked to visit the Twitter pages of three successful musicians and to answer a ten-question survey about those accounts every day for five consecutive days (Monday through Friday). We assured workers they needed no knowledge of Twitter and explained where to find the important information. Most of each day's task involved reporting on the Twitter activity of the artist from the day before. For example, on Monday the worker was to answer questions about the artists' Twitter activity on Sunday. Although we asked workers to complete the task on the correct day, we also accepted retroactive submissions and automatically recorded the time of submissions. Figure 4 displays the site and questionnaire. After the last assigned day of work, we again invited workers to a continuation of the task and recorded whether they re-applied.

4.2 Pure Selection Effects

The majority (61 percent) of workers from the individual experiment accepted our offer and so were hired for the supplemental experiment; referred workers were significantly more likely to accept than non-referred workers. However, regardless of whether we include all referred and non-referred workers in the analysis (counting as not submitting work those who did not accept our employment offer) or instead analyze performance conditional on accepting our job offer, our key results remain unchanged: even working at a job for which they were not referred at a firm with which their referrers were not affiliated, referred workers outperformed non-referred workers and had less turnover.

The three performance metrics in the supplemental experiment mirror those of the individual experiment: an indicator for submitting a response for a given day, an indicator for submitting that day's response on the correct day (analogous to the soft deadline of the individual experiment in that it was requested, but not required), and the fraction of questions answered correctly. Panel A of Table 4 shows that unconditional on accepting our employment offer, referred workers were 18 percentage points more likely to submit a response and to submit it on the correct day; they were also eight percentage points more accurate than non-referred workers. In addition to performing better, referred workers were 24 percentage points more likely than non-referred workers to apply for a continuation of the task.

Panel B shows the coefficients from the same regressions with the addition of controls for worker characteristics. Controlling for observables reduces the magnitude of the coefficients, but referred workers were still 9 percentage points more likely to submit any work and to

submit on time, and 12 percentage points more likely to indicate they wanted to continue the job at the end of the contract.

Next, we compare the performance of referred and non-referred workers who accepted our job offer. Appendix Table 2 provides suggestive evidence of, if anything, differentially positive selection of non-referred workers into accepting. In this table, we regress an indicator for accepting our job offer in the supplemental experiment on an indicator for being non-referred, a metric of performance or persistence in the individual experiment, and the interaction of that metric and the non-referred dummy. The base group is all referred workers. Each column uses a different performance or persistence metric from the individual experiment. The large standard errors render many of the results statistically insignificant, but the coefficients suggest that referred workers who performed well in the individual experiment were more likely to accept the job at the new firm. The coefficient on the non-referred indicator is always negative and large, suggesting that among workers who performed poorly, non-referred workers were much less likely to apply. Finally, the interaction between performance and being non-referred is large and positive. This is suggestive of positive selection of non-referred workers relative to referred workers.[22]

Panel C of Table 4 shows the results of estimating the same regressions as in Panel B, now limited to workers who accepted our job offer. In this sample, referred workers were 10 percentage points more likely to submit and 13 percentage points more likely to re-apply. However, given the results in Appendix Table 2, we might expect the conditional results to slightly underestimate the true performance and turnover differences between referred and non-referred workers. Appendix Table 3 shows that the results in Panels B and C of Table 4 are robust to the addition of the second order controls.

The magnitudes of the estimates from the supplemental experiment that condition on accepting our job offer are similar to those from the individual experiment, suggesting that the tasks were roughly comparable. Referred workers performed about as well here as did the non-monitored (Treatment 2) referred workers in the individual experiment. They submitted responses 76 percent of working days and submitted on the correct day (the analog to submitting on-time) 74 percent of the time. The performance gap between non-referred workers and referred workers is also similar to that of the individual experiment and, as in the individual experiment, the gap between referred and non-referred workers widened over the course the contract (Figure 5).

[22] The final column investigates re-application at the end of the individual experiment. Unsurprisingly, it shows that wanting to continue the job at the end of the individual experiment is positively correlated with accepting the job at the new firm. Conditional on re-applying at the end of the individual main experiment, there are not large differences between referred and non-referred workers' likelihood of accepting this new job.

5 Team Experiment

5.1 Design: Identifying Team Production

The team experiment was designed to determine whether referred workers outperform non-referred workers in a task involving teamwork and, if so, to determine how much of the performance difference between referred and non-referred workers is due to the fact that referred workers may perform particularly well when working with their referrers than other similarly-talented workers (team production).

The task involved brainstorming and we encouraged teamwork. Each worker was paired with three successive partners and asked to come up with slogans for three different public service announcements (PSAs).[23] The first PSA was to encourage highway drivers to wear seat belts, the second was to encourage children to practice good dental hygiene, and the third was to encourage college students to get the flu vaccine. For each PSA, we asked the worker to use the chat box we provided on our site to communicate with their partner and to come up with a single slogan that they would each submit through our online form. Figure 6 gives an example of what workers saw when they logged in to the team task site.

Though a worker could complete the task without her partner, the task was designed so that the best output necessitated teamwork. Each partner received a separate sheet with information relevant to the PSA. For the first PSA, for example, one partner received information on seat belts' efficacy, while the other received information about highway driving. The justification was that there was a lot of information to process and that by giving each partner different information, they would only have to read half as much. We told workers we wanted them to work with a partner to come up with their slogan because brainstorming is often more effective in teams.

Each information sheet contained a specific criterion we wanted the slogan to meet as well as a reason for that criterion. In the first round, for example, we told one partner that we wanted the slogan to be only three words long (so as not to distract drivers) and we told the other that we wanted the slogan to be in all capital letters (so drivers would be more responsive to it). In the second round, we told one partner to use an emoticon in their slogan (to make dental hygiene seem more upbeat) and the other to use the name of a real or fictitious person (since kids may respond to role models). In the third, we told each partner we wanted one of four specific words included in the PSA; one partner's word choices emphasized that getting the flu shot would be quick, the other partner's word choices emphasized that flu shots are effective. When giving workers their information sheets, we

[23] As with the prior tasks, we chose this task because there are many jobs on oDesk that ask low-skill workers with coming up with advertisements, for example in the form of flyers, posters, and/or slogans.

told them only that the sheets would contain information, not particular criteria for the slogans.

When workers submitted their slogans, we asked them also to answer a "team question:" a multiple choice question about the slogan. Each of the three PSA assignments had a different team question (what color sign the PSA should be printed on, what type of lettering the slogan should be written in, and where the PSA should be placed). This question had no correct answer, but partners were instructed to give the same answer.[24]

For comparison with the individual and supplemental experiments, we also collected measures of individual diligence. We monitored whether each worker logged into the site and whether she submitted work. We also asked each worker an "individual question," the answer to which was in her own reading (e.g., the fraction of highway drivers who wear seatbelts). Because workers were instructed that they should complete the task even if they could not make contact with their partner, workers should have logged in, submitted work, answered their individual question correctly, and used the criteria from their own information sheets in their slogan regardless of whom they were partnered with.

In the experiment, each referrer completed three different PSA tasks as part of three different types of teams: (1) a Type A team, in which she was paired with her own referral, (2) a Type B team, in which she was paired with someone else's referral, and (3) a Type C team, in which she was paired with a non-referred worker. Each referred worker worked with her own referrer when her referrer was in a Type A team and with someone else's referrer when her referrer was in a Type B team. (When her referrer was in a Type C team, she worked with another referred worker in the same position; results from this treatment are not presented.) Non-referred workers worked with referrers for all three rounds; that is, they were always in Type C teams.

Because we thought worker performance may well be correlated not just between partners, but also among partners' partners, we placed workers into blocking groups. By definition, every worker in the blocking group only ever partnered with others in the same blocking group. In all analyses of the team experiment, we cluster standard errors by blocking group.

Each of the 47 blocking groups contained six referrers, their six referred workers, and two non-referred workers. The placement into blocking groups was random, except that a referrer and her referral were always in the same group.[25] Within a blocking group, the

[24] Because we wanted to measure how effectively workers worked with their partners, we strongly encouraged each worker to complete each PSA. Unlike in the individual experiment, in which we sent workers no reminders about the task, in the team experiment we sent two reminders about each PSA to each worker who had not already submitted work.

[25] As in the individual experiment, we hired all referred workers who met the selection criteria, but only one randomly-selected referral from each referrer was included in this experiment. We also hired all non-referred workers who applied, but randomly selected 94 for inclusion in the blocking groups; the rest are not included

ordering of the type of team workers participated in was random. And, within team type, when relevant, workers' assigned partners were also random.

In addition to measuring worker performance, we collected a proxy for worker enjoyment of the partnered task and willingness to continue working with each partner. After the worker submitted her last slogan, we asked, "In case we have more tasks like this in the future, which if any of the partners that you've worked with would you be interested in working with again?" Workers could select all, none, or a subset of their partners.

5.2 Performance by Team Type

Table 5 compares referred and non-referred worker performance across team types on measures that do not rely on teamwork, but may be indicative of individual diligence. These are indicators for logging in to our site to see the given PSA task, submitting the day's work, correctly answering the question about their own individual reading, and including the criteria from their own information sheets in their slogans.[26]

Each outcome is regressed on an indicator for being in a Type A team (a referred worker paired with her own referrer) and an indicator for being in a Type C team (a non-referred worker paired with a referrer). The omitted group contains workers in Type B teams (referred workers paired with someone else's referrer). Thus, the coefficient on the Type A dummy indicates how much better referred workers perform when paired with their own referrer than with someone else's referrer; the Type C dummy indicates how much worse non-referred workers perform than referred workers when both are paired with someone else's referrer.

Each observation is a partner pair, but in these diligence measures, we consider only referred and non-referred workers. Referrers' performance does not vary significantly across team types.

On average, referred workers performed well on these diligence measures. When paired with someone else's referrer, referred workers logged in 88 percent of the time, submitted work 84 percent of the time, and correctly answered their own question 76 percent of the time. Less than half (44 percent), however, included the criteria from their own information sheet in their slogan (Panel A).[27] Non-referred workers, meantime, were substantially less diligent than referred workers, even when neither group was working with a partner they

in the results.

[26] If they did not answer the question about their reading, they are marked as not answering it correctly. Similarly, if they did not submit a slogan, they are marked as not including their own criteria in the slogan.

[27] For comparison, referrers on these teams were four percentage points more likely to log in and eight percentage points more likely to use their own criteria in the slogan. It is not surprising that the referrers outperformed their referrals given that we asked for referrals only from workers with strong oDesk records. Moreover, these referrers were (positively) self-selected from the set of workers we asked for referrals.

previously knew. As compared to referred workers in Type B teams, non-referred workers (all on Type C teams) were approximately 30 percentage points less likely to log in to the site, to submit a slogan, and to correctly answer their individual question. They were also 14 percentage points less likely to incorporate their information sheets' criteria into their slogan.

The coefficients on Type A teams show that referred workers were five percentage points more likely to submit their work and to correctly answer their own question when they were paired with their own referrer instead of with someone else's referrer. Given that these are measures of diligence more than teamwork, it could suggest that peer influence may have played a role in the team task. When working together, referrers may have put more pressure on their referrals to be diligent because in this context, their referrals' performance affected their own.

Panel B estimates the same regressions but includes controls for the referred and non-referred worker's own characteristics. Since referred workers in Type A and Type B teams are the same workers, the inclusion of these controls does not impact the coefficient on the Type A team dummy. Although between 15 and 35 percent of the difference between referred and non-referred workers can be accounted for by the observable characteristics (depending on the outcome), the differences between referred and non-referred workers remain large and mostly significant. Even with controls, non-referred workers were 23 percentage points less likely to submit a slogan and 25 percentage points less likely to correctly answer their individual question as compared to referred workers in Type B teams. Panel A of Appendix Table 4 shows that these differences are still significant when the second order controls are added.

Table 6 compares team performance by team type. Observations are again at the partner-pair level. Each team performance metric is regressed on an indicator for being a Type A team and an indicator for being a Type C team; the base group consists of Type B teams. The dependent variable in the first column is an indicator for both partners submitting work. This is provided for reference because if both partners did not submit then they cannot, by definition, have achieved the other performance measures. The dependent variable in the last three column are indicators for (1) whether the partners answered the multiple-choice team question the same way (e.g., what color sign the PSA should be printed on), (2) submitted the same slogan, and (3) submitted the same slogan and that slogan included the criteria given to each partner.

The first column of Panel A shows that 73 percent of Type B teams had both partners submit work. Type A teams were 10 percentage points more likely than Type B teams to have both partners submit. This is driven by two effects. First, referred workers were

five percentage points more likely to submit work when paired with their referrer than when paired with someone else's referrer. Second, partners' submission was more highly correlated in Type A teams than in Type B teams. (More Type A teams than Type B teams also had neither partner submit work.) As compared to Type B teams, Type C teams (teams with a non-referred worker) were 28 percentage points less likely to have both partners submit.

On measures of team performance, teams with a referred worker consistently outperformed those with a non-referred worker, even when the referred worker was working with someone else's referrer. For example, while half of Type B teams answered the team question the same way, Type C teams were 21 percentage points less likely to do so. While referred workers did well relative to non-referred workers even when not working with their referrers, they did particularly well when working with their referrers. Referred workers were, for example, substantially (29 percentage points) more likely to answer the team question the same way when working with their own referrers; of the Type A teams that both submitted responses, only 6 percent failed to submit the same response to the team question.

The results are consistent across team performance metrics. The third column shows similar results for submitting the same slogan. Only about one-third of Type B teams submitted the same slogan. Type C teams were about half as likely to do so; Type A teams were more than twice as likely. And while 14 percent of Type B teams submitted the same slogan with both criteria, the same statistic for Type C teams is 9 percent, and for Type A teams, 24 percent.[28]

Panel B estimates the same regressions with the addition of controls for the referred and non-referred workers' characteristics. Because the same referred workers participated in Type A and Type B teams, adding these controls does not affect the coefficients on the Type A indicator. It does decrease the coefficients on the Type C indicator by about 25 percent, but the team performance differences between Type B and Type C teams remain large and generally significant. Panel B of Appendix Table 4 replicates these results adding the second order controls and shows that these results are the least robust to the addition of the second order controls. While Type C teams are still significantly (16 percentage points) less likely to both submit work, the other differences between Type B and C teams fall and lose their significance.

[28] One potential explanation for why referred workers performed better when working with their referrers is that a referred worker and her referrer were, on average, more similar than a randomly-selected referrer and referred worker. We find no evidence, however, that this drives our results. We create indicators for whether both partners were of the same gender (using workers' names and honorifics), whether they lived in the same city, and whether they had previously worked at the same oDesk firm; we also measure the difference between the partners' wages. Partners in Type A teams look more similar on each of these dimensions than do partners in Type B teams. But none of these similarities positively predicts performance, nor does including measures of them in the regressions affect the estimated effect of working with one's own referrer.

5.3 Enjoyment by Team Type

One potential motivation for hiring referrals is that workers might enjoy working with their friends and, thus, might be willing to spend more time on the job when working with their friends. Because oDesk requires workers to record the time they spend working on oDesk tasks, we can analyze the amount of time workers spent on each of the three PSAs. Panel A of Table 7 shows time spent on the task by team type, first for referrers and then for referred and non-referred workers.

When partnered with someone they did not know, referrers spent the same amount of time (around 37 minutes) on the task regardless of whether their partner was a non-referred worker or someone else's referral. When working with their own referral, however, they spent an average of six extra minutes on the task. The average referred worker on a Type B team, meanwhile, spent 39 minutes on the task. When working with her own referrer, though, she spent significantly (about 14%) more time on the task. Unsurprisingly given that non-referred workers were much less likely to log in or submit a response, non-referred workers (all on Type C teams) spent 16 fewer minutes than did referred workers on Type B teams.

Panel B of Table 7 provides additional insight into how much workers enjoyed their work experience on each type of team. After they had completed all three tasks, workers reported which partners they would be interested in partnering with again; workers could choose as many or a few partners as they wanted. Some workers (about 20 percent) did not answer the question, mostly because they did not complete their third task. But for those who answered, we know whether or not they wanted to work again with each of their three partners. We regress an indicator for wanting to work with the same partner again on a dummy for being in a Type A team and a dummy for being in a Type C team, first for referrers and then for referred and non-referred workers.

While 41 percent of referrers reported wanting to work again with the referred worker they did not know, only 31 percent wanted to work again with their non-referred partner.[29] Meantime, referrers were substantially more likely to want to partner again with their own referral: 96 percent reported wanting to work again with the person they had referred.

Referred workers were also much more likely to say they wanted to work again with their own referrer than with the other referrer with whom they had worked. While 93 percent of referred workers reported wanting to work again with their own referrer, only 48 percent reported wanting to work again with their Type B partner. The coefficient on the Type C team dummy suggests referred and non-referred workers may have been equally likely to

[29] Referrers did not know who, besides their own referrals, had been referred to the firm.

want to work again with referrers they did not previously know, but because non-referred workers were much less likely than referred workers to answer this question, this estimate should be interpreted with caution.

5.4 The Role of Communication

One reason why the different team types may have performed differently is that they may have communicated differently. While each worker always had access to a chat box on the site in which she could chat live with and/or leave messages for her partner, Type A teams may have been advantaged by having other means of communicating. Here we discuss communication differences across team types and, in light of our findings, for robustness we re-estimate performance differences by team type with the inclusion of several additional controls.

Panel A of Table 8 shows how team types differed in their communication methods. We regress each communication outcome on indicators for being in Type A and Type C teams; as before, the base group is Type B teams. Controls for the characteristics of referred and non-referred workers are included throughout. The first column considers chat box use, defined as both partners typing at least one message in the chat box. The second column considers the total number of messages sent by both partners during the task and is limited to teams that used the chat box. Because we directly observe what is written in the chat box, both of these measures are known for all teams and do not rely on worker reports.

The last two columns consider communication outside the chat box, such as on Skype. When workers submitted their slogans for each task, we asked if they had used other forms of communication. We code teams as using other forms of communication if at least one partner reported doing so. The third column addresses selection into answering this question. Here we regress a dummy for whether at least one teammate answered this question on team type. In the final column, we regress an indicator for reporting using other forms of communication on team type. This final specification includes only teams that answered the communication question.

Type A teams seem to communicate the most, both in and out of the chat box. Relative to Type B teams, Type A teams were slightly, though insignificantly, more likely to use the chat box to communicate. When they did use the chat box, Type A teams sent about one-third more messages. The biggest difference between the communication of Type A and Type B teams, however, is in the frequency with which they used other forms of communication. While 38 percent of Type B teams reported using other forms of communication, Type A teams were twice as likely to do so; the magnitude of this difference implies that the difference

itself cannot be driven by the small difference in answering this question.[30]

The only significant communication difference in this panel between Type B and Type C teams was that Type C teams were significantly less likely to use the chat box. This is not surprising since the chat box was on the site and non-referred workers were significantly less likely to ever log in.

We have observed that, relative to Type B teams, Type A teams communicated more both in and out of the chat box. They also spent more time on the task (Table 7). Panel B of Table 8 provides evidence that, even controlling for communication and time spent, Type A teams still outperformed Type B teams. We replicate the main team performance specifications with controls for referred and non-referred workers' characteristics (from Panel B of Table 6), adding controls for using the chat box, the number of messages sent in the chat box, and using other methods of communication as well as controls for the minutes spent by each partner.[31]

Unsurprisingly, more communication and more time spent both lead to better outcomes. For example, teams that sent the median number of messages in the chat box (eight) were 21 percentage points more likely to answer the team question the same way and 15 percentage points more likely to provide the same slogan than teams that did not use the chat box, all else equal. Moreover, either partner spending more time is positively correlated with better outcomes. Teams in which each partner spent an additional five minutes each on the task were, all else equal, three percentage points more likely to have their team question match and two percentage points more likely to submit the same slogan.

Even conditioning on the type of communication used, number of messages sent, and minutes spent by each partner, however, Type A teams remained 14 percentage points more likely to provide the same answer to the team question and 23 percentage points more likely to submit the same slogan than Type B teams. Type C teams, meanwhile, remained substantially less likely to do either (11 percentage points and seven percentage points, respectively).

6 Predictors of Referral Performance

We find across our experiments that having been referred is a powerful, positive predictor of performance: in each of our three experiments, referred workers substantially outperformed

[30] At least one partner answered the question about communication in 95 percent of Type B teams; Type B teams were slightly more likely to answer this question than either of the other team types.

[31] If neither partner answered the question about using other forms of communication, we set the indicator for having reported communication outside the chat box to zero. Thus, this dummy also directly captures the effect of having at least one partner submit work.

their non-referred counterparts. But not all referrals are created equal. In this section we focus on referred workers and identify predictors of their performance. We look first at a referrer's performance as a predictor of the performance of her referral and find that higher performers refer higher-performing workers. We then turn to the relationship between the referrer and her referral. Here we find that workers who have stronger ties with their referrers tend to look worse on observable characteristics, but actually perform better.

6.1 Referrer's Performance

The first column of Table 9 shows that a referrer's performance is a strong predictor of her referral's performance. We regress the referred worker's performance in the individual experiment on her referrer's performance in the same experiment. (We use submission as our performance metric here, but using other performance metrics provides similar results.) A worker referred by someone who never submitted any work is predicted to submit work on only 46 percent of days. Meanwhile, a worker whose referrer submitted work every day is predicted to herself submit work on 88 percent of days.[32]

What this result does not illuminate is *why* the performance of the referring worker is a good predictor of her referral's performance. For example, referrers and referred workers may perform similarly because on any given day they experience common shocks in their ability or desire to complete the job. In what follows, we provide evidence that (1) even absent common shocks, workers tend to refer people who perform like they do, (2) part of this seems to be driven by the positive correlation between a worker's own observables and those of the worker she refers, and (3) even absent common shocks and controlling for the referred worker's observable characteristics, the referrer's performance still predicts her referral's performance.

Because it was executed four months after the individual experiment, the supplemental experiment allows us to disentangle the common shocks hypothesis from others by comparing the performance of referred workers at this later date with that of their referrers from an earlier period. (This assumes that common shocks do not persist for four months.) In the second column of Table 9, we regress the referred worker's performance in the supplemental experiment on her referrer's performance in the individual experiment four months before. We find that, even absent common shocks, the referrer's performance remains a powerful predictor of the referral's performance. For comparison, in the final column we regress the

[32] Since the average non-referred worker submitted 63 percent of days, workers referred by the worst-performing referrers are predicted to under-perform the average non-referred worker. That said, a referrer would have to submit only 40 percent of days for her referral to be predicted to perform as poorly as the average non-referred worker. Less than the bottom quarter of referrers performed this poorly.

referred worker's performance in the supplemental experiment on her *own* performance in the individual experiment. Knowing the performance of a worker's referrer four months prior leads to almost two-thirds as much updating as knowing her own performance four months ago.

As demonstrated in Appendix Table 5, workers with better observable characteristics refer workers who also have better observables. Referrers who have passed more oDesk tests, or who have higher earnings, for example, refer people with similar credentials. This seems partially to explain why referred workers perform like their referrers even absent common shocks. Controlling for the referred worker's observables in the regression in Table 9 reduces the point estimate on the referrer's performance by more than half. Nonetheless, the referrer's performance remains a large and positive (albeit not statistically significant) predictor of her referral's performance.[33] The results in Table 9, then, suggest that higher performers refer workers who not only have more positive observable characteristics, but also perform better than would be expected conditional on these characteristics, even absent common shocks.[34]

6.2 Strength of Referrer-Referral Relationship

We turn now to the relationships between referrers and their referrals. Appendix Table 6 shows the distributions of the three relationship variables we have from referrers' reports at the time of the referral.[35] Referrers tended to refer workers they were close to. Among those included in the experiment, most reported knowing their referrals "extremely well" (six on a scale of one to six), while only one percent said they knew their referral "hardly at all" (one on the same scale). According to referrers, 32 percent of referrals interacted with their referrers more than once a day (in person or remotely) and another 19 percent interacted about once a day; meanwhile, only 7 percent interacted once a month or less. We also asked workers how many other people they knew in common with their referral: 37 percent of referred workers knew 30 or more people in common with their referrer, while the same fraction knew fewer than 10.

We are interested here in how referrer performance varies with the strength of the referrer-

[33] When on-time submission is used as the performance metric instead of submission, this coefficient is significant at the five percent level.

[34] In the team experiment, when referred workers and their referrers were in Type B teams (and so not working together), their performance was unsurprisingly also positively correlated. The magnitude of this correlation is quite similar to the correlation in the individual experiment. We do not, however, focus here on the results from the team experiment because we cannot tease out the role of common shocks in this experiment.

[35] We caveat this section by underlining that these relationship characteristics are self-reported by the referrer. Referrers could have reported being close with workers whom they thought would perform particularly well and/or whom they particularly wanted the employer to hire.

referral relationship. Because each relationship variable is consistently a positive predictor of the referral's performance, we build an index of relationship strength and for parsimony focus here on the resulting estimates.[36] We exclude the five referred workers whose referrers did not answer all the relationship questions at the time of the referral.

Panel A of Table 10 shows how characteristics of the referred worker vary with the strength of her relationship with her referrer. Each column shows the results of a different observable characteristic regressed on the relationship index. The reported coefficients each reflect the effect of a one-standard deviation increase in the relationship index on the observable characteristic.

Referred workers who have stronger relationships with their referrer look worse on observables. They have passed fewer oDesk tests, completed fewer assignments, and, conditional on receiving feedback, have received (insignificantly) worse feedback.[37] These results suggest that when referrers refer people with whom they have weaker ties, they refer people who look better on paper.

Panel B investigates how a referral's performance varies with the strength of the referrer-referral relationship. For each experiment, a worker's performance on a given day (or a given PSA in the team experiment) is regressed on the relationship index. For parsimony we present only one outcome per experiment. The first column considers submission in the individual experiment, the second considers submission in the supplemental experiment, and the third considers slogan matching in the team experiment. Within experiments, the magnitudes of the coefficients are similar when performance is defined using the other metrics.

In each experiment, referred workers performed better the stronger their relationship with their referrer. A referred worker with a one standard deviation stronger relationship with her referrer was four percentage points more likely to submit work in the individual experiment and two percentage points more likely to submit work in the supplemental experiment, though the latter point estimate is not statistically significant. In the team experiment a referred worker with a one standard deviation stronger relationship with her referrer was five

[36] In building the index, we first create dummy variables for reportedly knowing the referred worker well (responding more than three on a scale of one to six when asked how well she knew the referred worker), interacting with the referral at least once a week, and knowing at least twenty people in common. Our relationship index is defined as the standardized sum of these three binary variables. The magnitudes of the coefficients are similar if we define the index instead as the standardized average of z-scores for the three raw variables (on scales of 1 to 6 for how well the referrer knew her referral, 1 to 7 for how often they interacted, and 1 to 5 on how many people they knew in common) or of z-scores for three constructed variables (with how well they knew each other on the same scale, but with how often they interacted coded as the estimated number of days per month they interacted and with how many people they knew in common coded as the midpoint in the chosen range).

[37] They are also significantly (5 percentage points) less likely to have received any feedback, probably because they have completed fewer assignments. Point estimates suggest they also look worse on the other observable characteristics we have, but the coefficients are not generally significant.

percentage points (nearly ten percent) more likely to have her slogan match her partner's.

Panel C presents the results of these same regressions with the inclusion of controls for the referred worker's observable characteristics. Given that referred workers with stronger ties to their referrers tend to have worse observable characteristics, it is unsurprising that the coefficients on the relationship index are, on average, larger when the controls are added.[38]

These results are consistent with the idea that when workers refer people they know less well, they refer workers who look better on paper. In contrast, when they refer people they know better, they choose workers who do not look as good on paper, but who perform well in ways that would not be predicted by their observables.

7 Potential Bias from Employers' Hiring Decisions

In each experiment, we hired all applicants who met our basic hiring criteria. A benefit of this approach is that there is not differential selection of referred and non-referred workers into our sample based on an employer's hiring decision. In this section, we use our experimental data to simulate how our comparisons between referred and non-referred workers might have been biased had we only observed the performance of workers an employer chose to hire. Though the results of this exercise are qualitatively similar under different assumptions, our primary aim in this section is not to pin down the particular bias that would be generated by an employer's hiring decisions, but rather simply to demonstrate that such a bias might exist.

To do this, we first simulate which of the workers in the individual experiment an employer would hire under three different hiring scenarios. We then compare how the hired referred and non-referred workers would perform in each scenario. The three scenarios vary in the information an employer sees when choosing whom to hire. They are (1) the employer sees no information about applicants (provided only for reference), (2) the employer sees workers' resumes, but not who was referred, and (3) the employer sees both workers' resumes and who was referred.

In each of the three hiring scenarios, we assume that employers want to maximize the fraction of workers who submit a response on a given day and also that they know the relationship between demographics and referral status and performance.[39] They incorporate

[38] We do not have enough power to test the interaction of the relationship variables and treatment (in the individual experiment) or team type (in the team experiment). However, the coefficients on the interactions are not consistently signed across outcomes within each experiment.

[39] In practice, an employer may prefer to hire a referred worker over a non-referred worker who is predicted to perform slightly better either as a source of compensation to an existing employee or because the referred worker is predicted to persist longer at the firm. For simplicity and clarity, we abstract away from any such considerations here.

what information (if any) they have on applicants to predict that applicant's performance; they then hire the half of the applicant pool with the best predicted performance.

Table 11 shows the results of the simple simulation. Panel A shows the fraction of the referred and non-referred applicants that would be hired under each scenario, while Panels B and C display characteristics of these hired referred and non-referred workers. Panel B provides a summary measure of the hired workers' observable characteristics. This is the predicted probability that a worker will submit a response based on her observable characteristics. For a given worker, it is calculated by regressing the performance of all the other workers (excluding herself) on their demographic characteristics (not including referral status). These coefficients are then use to predict the excluded worker's own performance. Panel C shows the average actual (observed) submission rates of the referred and non-referred workers that would be hired in each scenario.

The first row corresponds to the first hiring scenario, in which employers have no information about applicants and so hire a random 50 percent sample of the applicant pool. Here, as in our experimental results, referred workers look better on observables than do non-referred workers. They are also 15 percentage points more likely to submit work than non-referred workers.

Results in the second row simulate hiring under the assumption that employers see workers' resumes, but not who is referred. In this scenario, employers simply hire the workers with the highest summary measure of observables. A higher fraction of referred (58 percent) than non-referred (39 percent) workers is hired since referred applicants have better observable characteristics. Among those hired, referred and non-referred workers unsurprisingly have similar observable characteristics. But the hired referred workers remain substantially (19 percentage points) more likely to actually submit work.

The final row presents the results of simulating hiring under the assumption that the employer also observes which workers were referred. For each worker, we regress the performance of all other workers (excluding herself) on their observable characteristics and an indicator for being referred. We then predict the excluded worker's performance using the estimated coefficients and assume that employers hire the half of workers with the highest predicted performance. When the referral indicator is incorporated into hiring decisions, the fraction of referred applicants that is hired increases to 77 percent, while only 12 percent of non-referred applicants are hired. The non-referred workers hired are now a select group with very positive observable characteristics: based on their observables, they are predicted to submit much more often than are hired referred workers (86 percent versus 77 percent of the time).

Even when the referral is used in hiring decisions, hired referred workers perform a little

better than non-referred workers, but the difference is much smaller (only five percentage points) and is no longer significant.[40] This suggests that if we had only observed the performance of hired workers and did not observe all the characteristics employers used in the hiring decision, we might have mistakenly concluded that referrals contained little to no information about worker performance.

8 Conclusion

This project presents the results of three field experiments in an online marketplace comparing the performance and turnover of referred and non-referred workers. We find that referrals contain information: referred workers performed better and had less turnover than non-referred workers, even conditional on their observable characteristics and initial job performance. On-the-job interactions between referred workers and their referrers drove some of the superior performance of referred workers: they performed particularly well when working directly with their referrers. However, much of the performance and turnover differential between referred and non-referred workers was driven by selection. Referred workers outperformed and persisted longer than non-referred workers even at a job for which they were not referred for a firm at which their referrers did not work.

These results are somewhat similar to the results in Burks et al. (2013), which also compares the performance of referred and non-referred workers. Burks et al. (2013) uses rich administrative data from firms in three industries (call centers, trucking, and high technology). Consistent with our results, it finds that referred workers have lower turnover than non-referred workers. At odds with our results, however, it finds that referred workers perform better only on rare "high-impact" metrics (e.g., generating patents and avoiding accidents). It concludes that it is selection on match quality that leads referred workers to perform better than non-referred workers on these metrics, not general selection nor on-the-job interactions with referrers.

One potential reason for this difference in findings is that our paper analyzes the performance of all applicants while most of the performance data in Burks et al. (2013) comes from workers the firms chose to hire. If employers can observe who is referred, then they may already take the positive signal contained in a referral into account when making hiring decisions. Consistent with this, on a job test taken by all applicants (not just hires), Burks et al. (2013) does find that referred applicants perform better than non-referred applicants.

[40] In fact, if we assume employers hire the top third or top quarter of the applicant pool (rather than the top half), hired referred workers perform two or three percentage points *worse* than hired non-referred workers when the employer uses referral status in the hiring decisions. As in the main specification, these differences are not significant.

Relatedly, its conclusion that on-the-job interactions with referrers cannot explain referred workers' superior performance comes largely from call center workers and truck drivers; we similarly find that when assigned a predominately individual task, the main benefit of hiring a referred worker is selection. However, on a task in which there are large returns to teamwork, we find that referred workers perform particularly well when working with their referrers. This is consistent with Burks et al.'s (2013) finding that high tech workers with more friends at the firm are more productive.

Overall, we find that selection and team production can explain why referred workers outperform and outlast non-referred workers. Yet we have limited evidence on why the referral contains so much information. One explanation (as in Montgomery, 1991) is that there is simply homophily among friends: productive workers have productive friends. Another explanation is that (as in Beaman and Magruder, 2012) workers have information on which of their friends are particularly productive and may choose these particularly productive workers to refer. Understanding the relative contributions of these two factors is an important question for future research.

References

Bandiera, O., I. Barankay, and I. Rasul (2005): "Social preferences and the response to incentives: Evidence from personnel data," *The Quarterly Journal of Economics*, 120(3), 917–962.

——— (2012): "Team incentives: Evidence from a firm level experiment," *CEPR Discussion Paper No. DP8776*.

Beaman, L., and J. Magruder (2012): "Who gets the job referral? Evidence from a social networks experiment," *The American Economic Review*, 102(7), 3574–3593.

Bentolila, S., C. Michelacci, and J. Suarez (2010): "Social contacts and occupational choice," *Economica*, 77(305), 20–45.

Bewley, T. F. (1999): *Why wages don't fall during a recession*. Harvard University Press.

Blau, G. (1990): "Exploring the mediating mechanisms affecting the relationship of recruitment source to employee performance," *Journal of Vocational Behavior*, 37(3), 303–320.

Brown, M., E. Setren, and G. Topa (2012): "Do informal referrals lead to better matches? Evidence from a firm's employee referral system," *Evidence from a Firm's Employee Referral System (August 1). FRB of New York Staff Report*, (568).

Bryan, G. T., D. Karlan, and J. Zinman (2012): "You can pick your friends, but you need to watch them: Loan screening and enforcement in a referrals field experiment," *NBER Working Paper 17883*.

Burks, S., B. Cowgill, M. Hoffman, and M. Housman (2013): "The Value of Hiring through Referrals," *Available at SSRN 2253738*.

Calvo-Armengol, A., and M. O. Jackson (2004): "The effects of social networks on employment and inequality," *The American Economic Review*, 94(3), 426–454.

Castilla, E. J. (2005): "Social Networks and Employee Performance in a Call Center," *American Journal of Sociology*, 110(5), 1243–1283.

Costa, D. L., and M. E. Kahn (2003): "Cowards and heroes: Group loyalty in the American Civil War," *The Quarterly Journal of Economics*, 118(2), 519–548.

——— (2010): "Health, wartime stress, and unit cohesion: Evidence from Union Army veterans," *Demography*, 47(1), 45–66.

Datcher, L. (1983): "The impact of informal networks on quit behavior," *The Review of Economics and Statistics*, 65(3), 491–495.

Dustmann, C., A. Glitz, and U. Schönberg (2011): "Referral-based job search networks," *IZA Discussion Paper 5777*.

Fernandez, R. M., E. J. Castilla, and P. Moore (2000): "Social capital at work: Networks and employment at a phone center," *American Journal of Sociology*, pp. 1288–1356.

Fernandez, R. M., and N. Weinberg (1997): "Sifting and sorting: Personal contacts and hiring in a retail bank," *American Sociological Review*, pp. 883–902.

Granovetter, M. (1995): *Getting a Job: A Study of Contacts and Careers*. University of Chicago Press, Chicago.

Heath, R. (2013): "Why Do Firms Hire Using Referrals?: Evidence from Bangladeshi Garment Factories," *Working Paper*.

Herrmann, M. A., and J. E. Rockoff (2010): "Worker absence and productivity: Evidence from teaching," *NBER Working Paper 16524*.

Holzer, H. J. (1987): "Hiring procedures in the firm: their economic determinants and outcomes," *NBER Working Paper 2185*.

Horton, J. J. (2013): "Computer-Mediated Matchmaking: Facilitating Employer Search and Screening," *Working Paper*.

Ioannides, Y. M., and L. D. Loury (2004): "Job information networks, neighborhood effects, and inequality," *Journal of Economic Literature*, 42(4), 1056–1093.

Kugler, A. D. (2003): "Employee referrals and efficiency wages," *Labour Economics*, 10(5), 531–556.

Montgomery, J. D. (1991): "Social networks and labor-market outcomes: Toward an economic analysis," *The American Economic Review*, 81(5), 1408–1418.

oDesk Corporation (2013): *Slide Deck with Statistics on the oDesk Marketplace*.

Pallais, A. (2013): "Inefficient Hiring in Entry-Level Labor Markets," *NBER Working Paper*, 18917.

Petersen, T., I. Saporta, and M.-D. L. Seidel (2000): "Offering a Job: Meritocracy and Social Networks," *American Journal of Sociology*, 106(3), 763–816.

Pistaferri, L. (1999): "Informal networks in the Italian labor market," *Giornale degli Economisti e Annali di Economia*, pp. 355–375.

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

Simon, C. J., and J. T. Warner (1992): "Matchmaker, matchmaker: The effect of old boy networks on job match quality, earnings, and tenure," *Journal of Labor Economics*, 10(3), 306–330.

Topa, G. (2011): "Labor markets and referrals," *Handbook of Social Economics*, pp. 1193–1221.

Figure 1. oDesk Profile Example

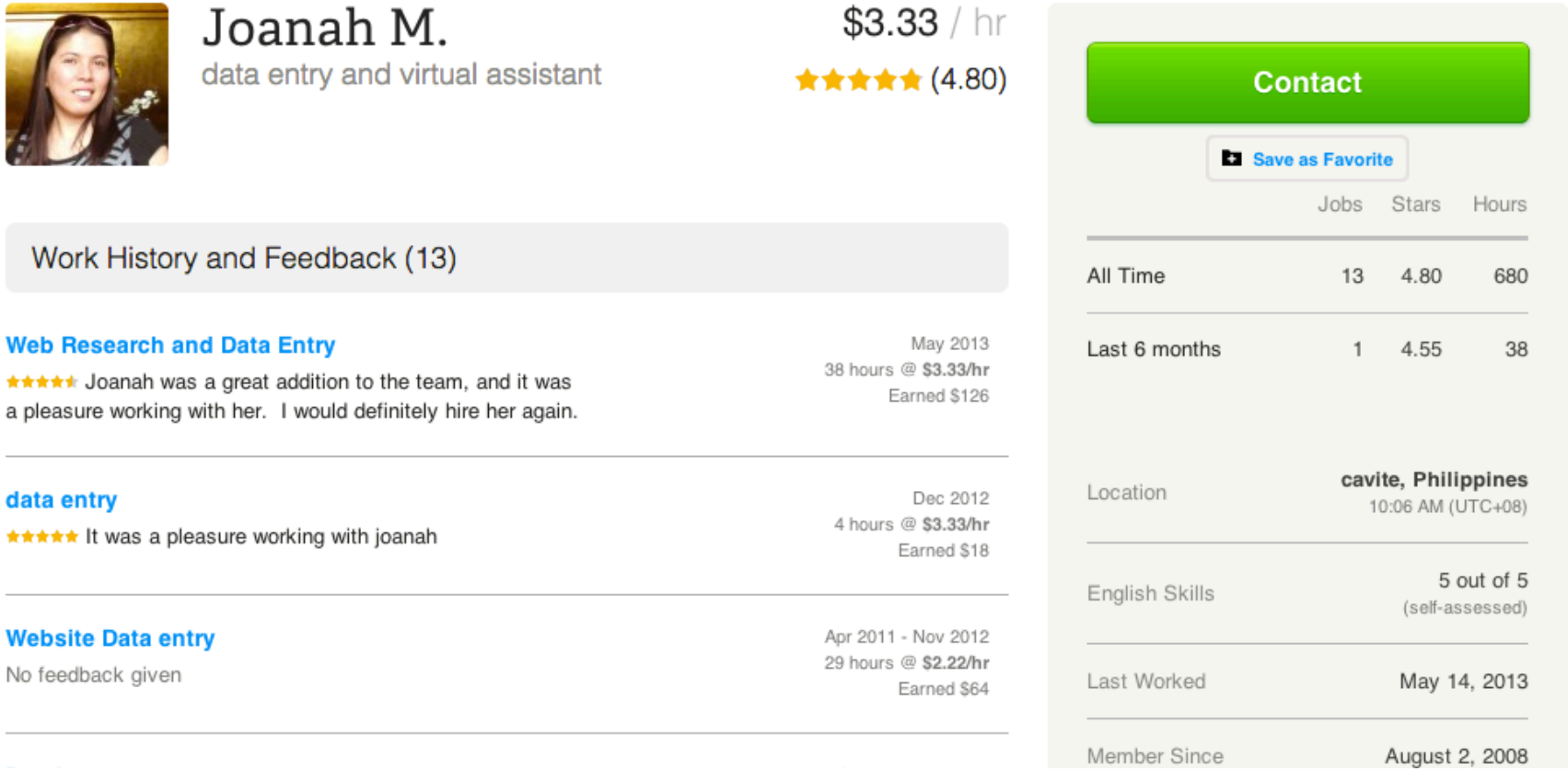
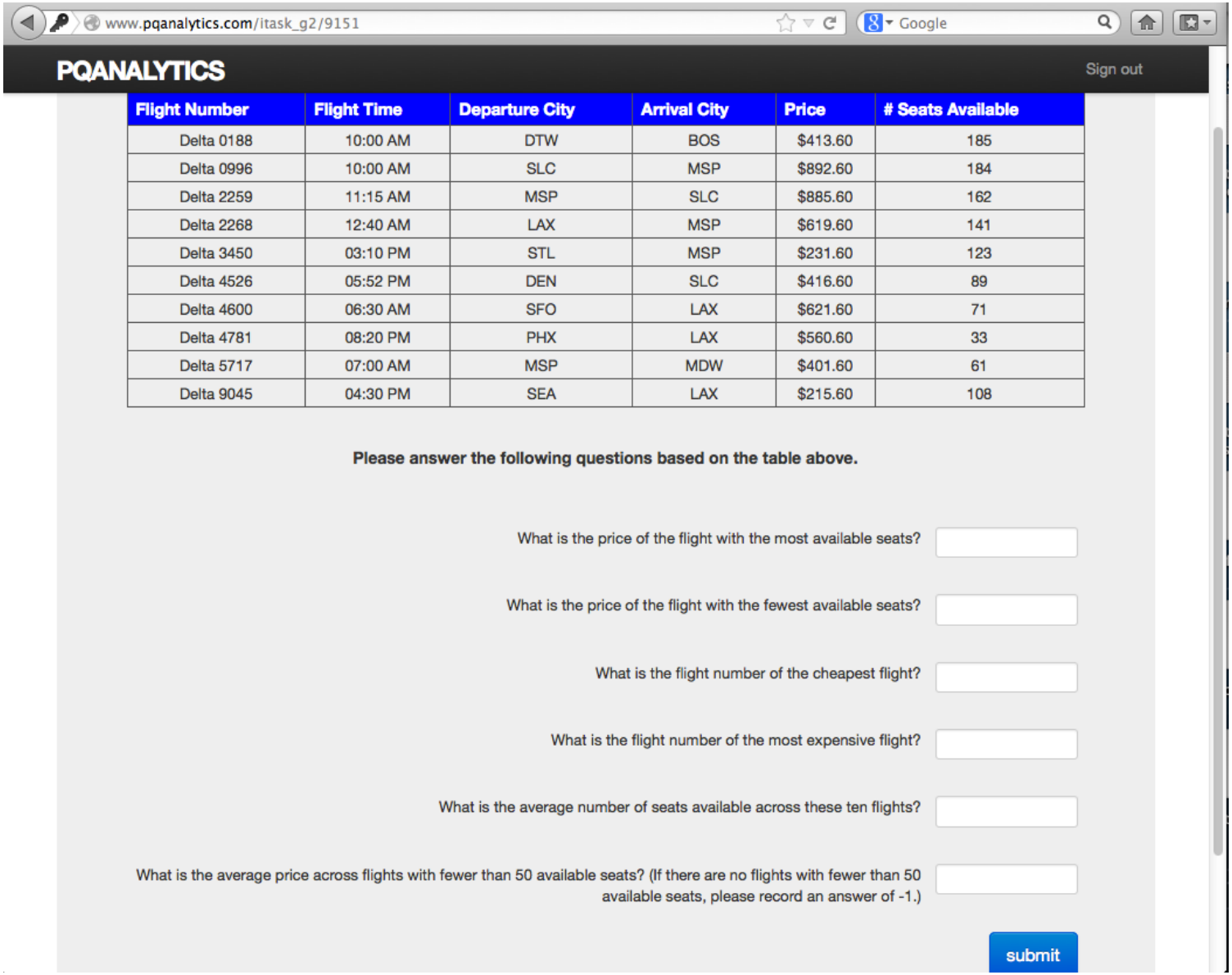


Figure 2. Individual Experiment Task Site for Referred and Non-Referred Workers



www.pqanalytics.com/itask_g2/9151

PQANALYTICS

Sign out

Flight Number	Flight Time	Departure City	Arrival City	Price	# Seats Available
Delta 0188	10:00 AM	DTW	BOS	$413.60	185
Delta 0996	10:00 AM	SLC	MSP	$892.60	184
Delta 2259	11:15 AM	MSP	SLC	$885.60	162
Delta 2268	12:40 AM	LAX	MSP	$619.60	141
Delta 3450	03:10 PM	STL	MSP	$231.60	123
Delta 4526	05:52 PM	DEN	SLC	$416.60	89
Delta 4600	06:30 AM	SFO	LAX	$621.60	71
Delta 4781	08:20 PM	PHX	LAX	$560.60	33
Delta 5717	07:00 AM	MSP	MDW	$401.60	61
Delta 9045	04:30 PM	SEA	LAX	$215.60	108

Please answer the following questions based on the table above.

What is the price of the flight with the most available seats?

What is the price of the flight with the fewest available seats?

What is the flight number of the cheapest flight?

What is the flight number of the most expensive flight?

What is the average number of seats available across these ten flights?

What is the average price across flights with fewer than 50 available seats? (If there are no flights with fewer than 50 available seats, please record an answer of -1.)

submit

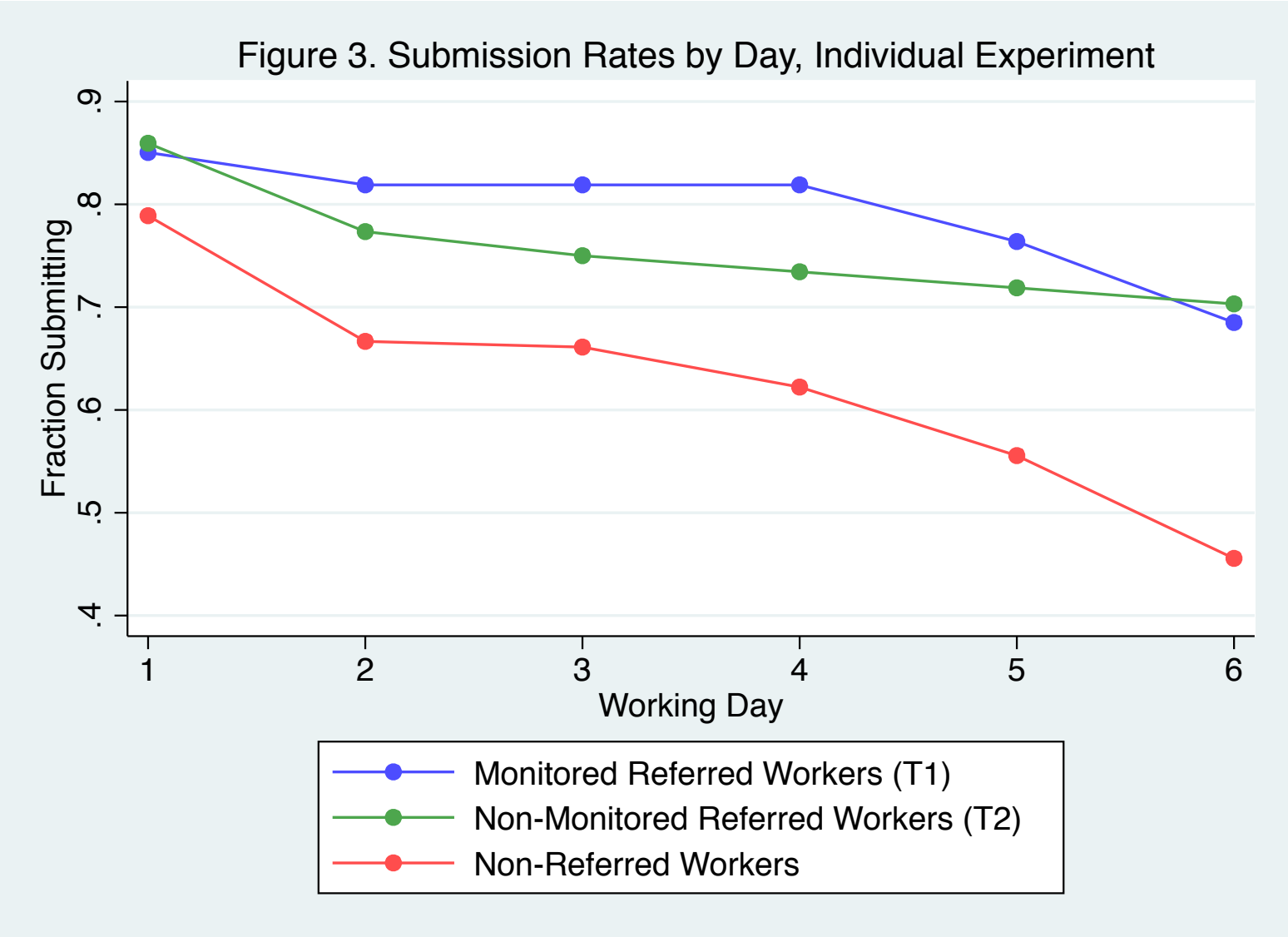


Figure 3. Submission Rates by Day, Individual Experiment
Fraction Submitting
.4
.5
.6
.7
.8
.9
1
2
3
4
5
6
Working Day
Monitored Referred Workers (T1)
Non-Monitored Referred Workers (T2)
Non-Referred Workers

Figure 4. Supplemental Experiment Task Site

We would like you to track the Twitter activity of three artists. Your three artists are (1) Justin Bieber, (2) Taylor Swift, and (3) Miley Cyrus.

Before we begin, what is today's date?

Now, check out how Justin Bieber is trending. The link to his profile is https://twitter.com/justinbieber and for Twitter users, his official handle is @justinbieber.

How many followers does Justin have right now?

How many tweets did Justin himself post yesterday? Please do NOT count retweets.

How many hashtags were there in Justin's own tweets from yesterday? (Recall that "hashtag" refers to the # symbol.)

Next, take a look at Taylor Swift's Twitter profile. The link to her profile is https://twitter.com/taylorswift13 and her official handle is @taylorswift13.

How many followers does Taylor have right now?

How many tweets did Taylor herself post yesterday? Please do NOT count retweets.

How many hashtags were there in Taylor's own tweets from yesterday? (Recall that "hashtag" refers to the # symbol.)

Then look at Miley Cyrus's Twitter profile. The link to her profile is https://twitter.com/MileyCyrus and her official handle is @MileyCyryus.

How many followers does Miley have right now?

How many tweets did Miley herself post yesterday? Please do NOT count retweets.

How many hashtags were there in Miley's own tweets from yesterday? (Recall that "hashtag" refers to the # symbol.)

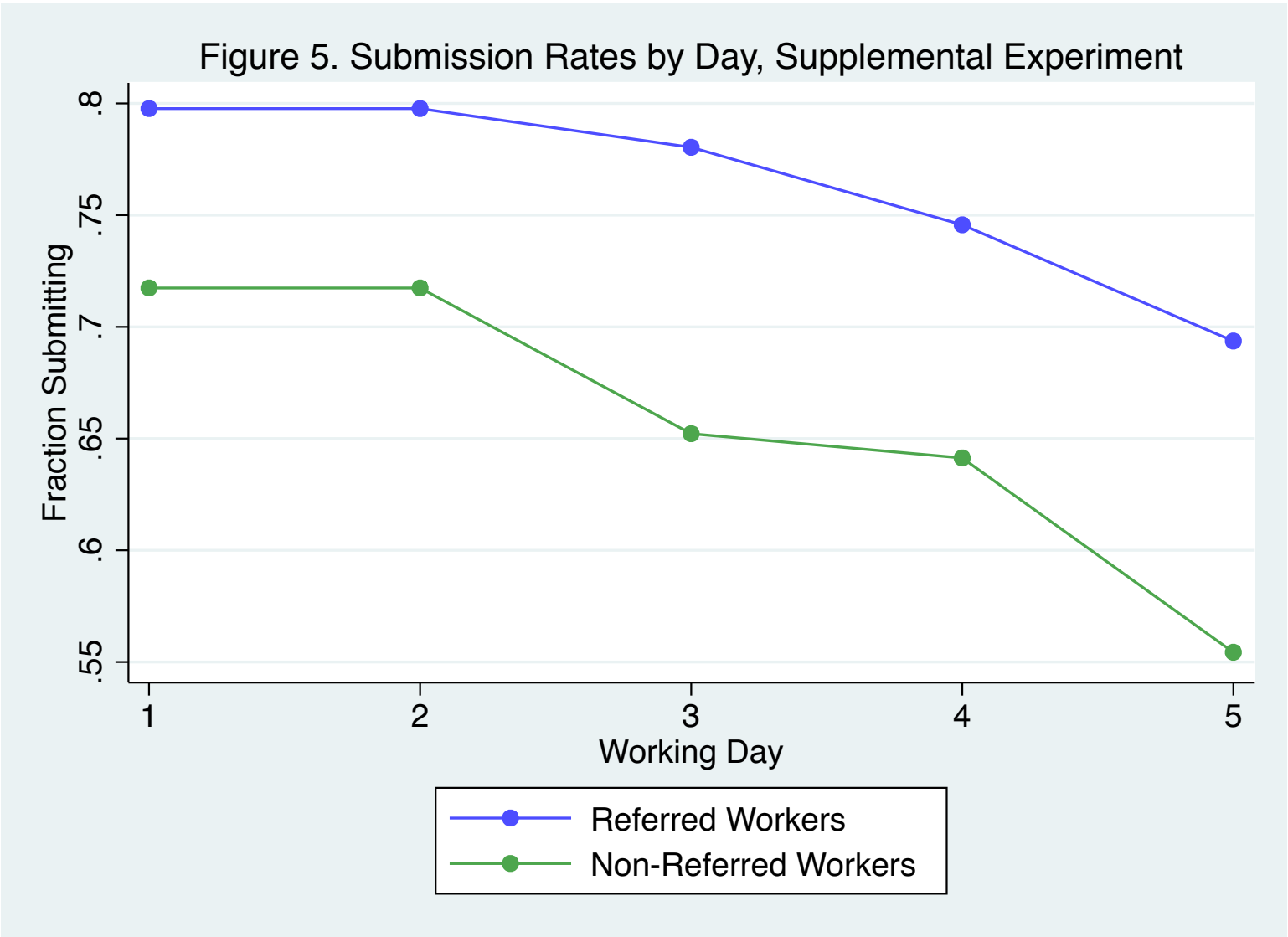


Figure 5. Submission Rates by Day, Supplemental Experiment
Fraction Submitting
.55
.6
.65
.7
.75
.8
1
2
3
4
5
Working Day
Referred Workers
Non-Referred Workers

Figure 6. Team Experiment Task Site

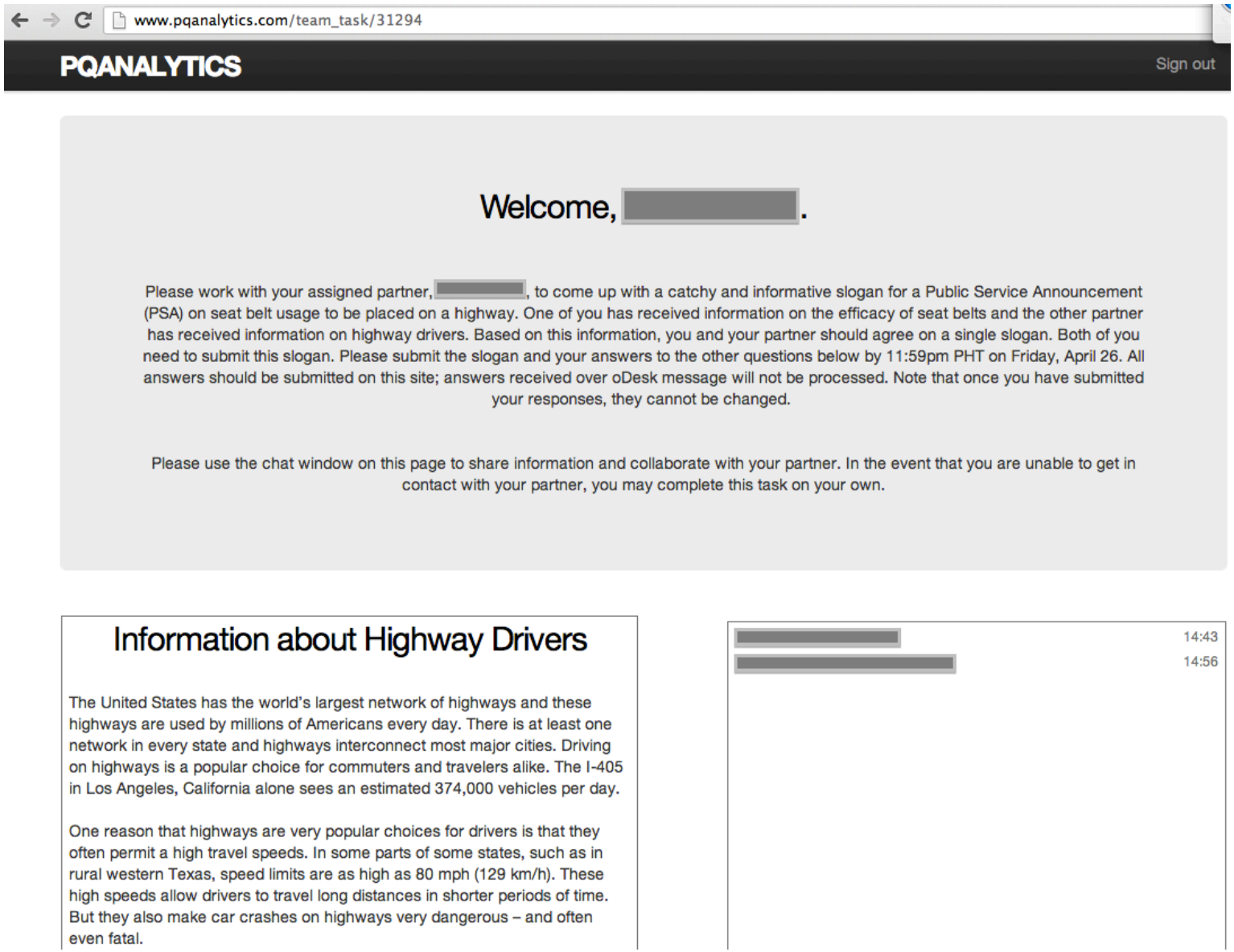


Table 1. Descriptive Statistics
Individual and Team Experiments

A. Workers Asked to Refer

	Referred Someone	Referred No One	Difference	Included Referrers
Has Prior Experience	1.00	1.00		1.00
Earnings	$2,917	$2,397	**	$2,932
Number of Previous Jobs	12.58	11.07	**	12.35
Has Feedback Score	1.00	1.00		1.00
Feedback Score	4.80	4.80		4.81
Posted Wage	2.84	2.77	*	2.85
Days Since Joining oDesk	689	709		666
Has Portfolio	0.69	0.61	**	0.69
Number of Tests Passed	5.80	5.60		5.84
Has English Score	1.00	0.99	**	1.00
English Score	4.79	4.79		4.77
Agency Affiliated	0.25	0.24		0.21
Number of Degrees	1.40	1.35		1.41
Proposed Wage	$2.50	$2.51		$2.51
Observations	1,246	1,867		455

B. Referred and Non-Referred Workers

	All Referred Workers	Included Referred Workers	Included Non-Referred Workers	Difference
Has Prior Experience	0.62	0.64	0.28	**
Earnings	$1,481	$864	$353	**
Number of Previous Jobs	7.29	5.79	2.05	**
Has Feedback Score	0.53	0.55	0.24	**
Feedback Score	4.54	4.55	4.26	**
Posted Wage	2.97	2.58	3.29	**
Days Since Joining oDesk	501	462	572	**
Has Portfolio	0.46	0.48	0.23	**
Number of Tests Passed	4.14	4.49	3.19	**
Has English Score	0.98	0.99	0.96	**
English Score	4.67	4.68	4.58	**
Agency Affiliated	0.18	0.12	0.06	**
Number of Degrees	1.27	1.35	1.01	**
Proposed Wage		$2.34	$2.59	**
Observations	1,854	537	274	

Notes: Each statistic in the table presents the mean of the characteristic indicated by the row for the sample indicated by the column. Both panels use data from both the individual and team experiments. *Referred Someone* denotes workers who referred at least one other worker to our firm, whether or not we hired that worker. *Referred No One* denotes workers who referred no workers to our firm. *Included Referrers* is a subset of *Referred Someone* and includes only those workers whose referral we hired. *All Referred Workers* denotes all workers who were referred, while *Included Referred Workers* is the subset of *All Referred Workers* who applied for our job and whom we hired for the individual or team experiment. *Included Non-Referred Workers* are non-referred workers who applied for our job and whom we hired. *Difference* denotes the significance level of the difference in the mean of the indicated characteristics between *Referred Someone* and *Referred No One* (Panel A) and between *Included Referred Workers* and *Included Non-Referred Workers* (Panel B). *English Score* is self-reported English ability on a one-to-five scale, a *portfolio* is where a worker posts prior work, and *agency-affiliated* workers pay a fraction of their earnings to report they are part of a given group of oDesk workers (an agency). *, ** denote significance at the 10% and 5% levels.

Table 2. Randomization Assessment
Individual Experiment

	Referrers		Referred Workers	
	Treatment 1	Treatment 2	Treatment 1	Treatment 2
Has Prior Experience	1.00	1.00	0.73	0.75
Earnings	$2,919	$2,996	$1,396	$1,379
Number of Previous Jobs	12.78	13.09	8.28	10.14
Has Feedback Score	1.00	1.00	0.62	0.64
Feedback Score	4.80	4.76	4.66	4.59
Posted Wage	2.78	2.85	2.68	2.72
Days Since Joining oDesk	645	676	489	566
Has Portfolio	0.64	0.68	0.47	0.50
Number of Tests Passed	5.78	5.78	4.98	5.31
Has English Score	1.00	1.00	0.98	1.00
English Score	4.84	4.79	4.75	4.66
Agency Affiliated	0.08	0.08	0.05*	0.10*
Number of Degrees	1.50	1.36	1.34	1.51
Proposed Wage	2.53	2.53	2.40	2.37
Observations	86	87	127	128

Notes: Each cell presents the mean of the characteristic indicated by the row for the sample indicated by the column. Only workers in the individual experiment are included. *English Score* is self-reported English ability on a one-to-five scale, a *portfolio* is where a worker posts prior work, and *agency-affiliated* workers pay a fraction of their earnings to report they are part of a given group of oDesk workers (an agency). * denotes the Treatment 1 and Treatment 2 group means are statistically different at the 10% level.

Table 3. Performance and Persistence

Individual Experiment: Base Group is Non-Monitored Referred Workers (Treatment 2)

A. No Controls

	Submission	On-Time Submission	Accuracy	Re-Application
Monitored Referred	0.036	0.053	0.034	-0.032
(Treatment 1)	(0.042)	(0.047)	(0.039)	(0.030)
Non-Referred	-0.132**	-0.079*	-0.101**	-0.225**
	(0.042)	(0.045)	(0.039)	(0.038)
Constant	0.757	0.563	0.640	0.953
	(0.031)	(0.034)	(0.028)	(0.019)
Controls	No	No	No	No
Observations	2,610	2,610	2,610	435
R-squared	0.027	0.013	0.020	0.085

B. With Controls

	Submission	On-Time Submission	Accuracy	Re-Application
Monitored Referred	0.021	0.039	0.015	-0.033
(Treatment 1)	(0.042)	(0.047)	(0.040)	(0.035)
Non-Referred	-0.127**	-0.090*	-0.102**	-0.216**
	(0.046)	(0.048)	(0.042)	(0.043)
Controls	Yes	Yes	Yes	Yes
Observations	2,610	2,610	2,610	435
R-squared	0.078	0.063	0.075	0.130

C. Last Day Only, Controlling for Performance on First Five

	Submission	On-Time Submission	Accuracy	Re-Application
Monitored Referred	-0.055	0.000	-0.051	-0.047
(Treatment 1)	(0.044)	(0.053)	(0.039)	(0.035)
Non-Referred	-0.163**	-0.125**	-0.120**	-0.180**
	(0.048)	(0.050)	(0.043)	(0.041)
	0.040	0.117**	0.036	
Day 1 Performance	(0.058)	(0.048)	(0.056)	
	-0.011	0.022	0.033	
Day 2 Performance	(0.057)	(0.051)	(0.060)	
	0.136*	0.247**	0.170**	
Day 3 Performance	(0.075)	(0.071)	(0.075)	
	0.210**	0.114*	0.206**	
Day 4 Performance	(0.081)	(0.066)	(0.079)	
	0.440**	0.223**	0.396**	
Day 5 Performance	(0.072)	(0.070)	(0.073)	
Controls	Yes	Yes	Yes	Yes
Observations	435	435	435	435
R-squared	0.528	0.405	0.535	0.306

Notes: Each column in each panel presents the results of a separate regression using data from the individual experiment. The dependent variable (indicated by the column) is regressed on an indicator for being a referred worker in Treatment 1 and an indicator for being non-referred. The omitted group is referred workers in Treatment 2. Panels A and B include all six days of observations. Regressions in Panel A include no controls; regressions in Panel B includes the controls for worker characteristics listed in footnote 20. Regressions in Panel C are limited to observations on workers' last day of work. They include the controls in footnote 20 and controls for the worker's prior performance. Each of the first three columns in Panel C includes controls for the worker's performance as measured by the dependent variable on each of the first five days. For example, in the *Submission* column, we control for Day 1 Submission through Day 5 Submission. The final column *(Re-Application)* includes controls for each of the three performance measures on each of the six days. For the first three columns in each panel, observations are worker-days and standard errors are clustered at the worker level. In the final column, observations are workers and Huber-White standard errors are presented. *Submission* is an indicator for whether the worker sent in any response. *On-Time Submission* is an indicator for whether that response was received by the soft deadline (11am). *Accuracy* is the percentage of questions the worker answered correctly; non-responses are counted as incorrect. *Re-Application* is an indicator for whether the worker accepted our invitation for a continuation of the same task. *, ** denote significance at the 10% and 5% levels.

Table 4. Performance and Persistence in New Firm
Supplemental Experiment: Base Group is All Referred Workers

A. No Controls

	Accepted Job Offer	Submission	On-Time Submission	Accuracy	Re-Application
Non-Referred	-0.167**	-0.182**	-0.178**	-0.079**	-0.236**
	(0.047)	(0.042)	(0.042)	(0.022)	(0.047)
Constant	0.678	0.518	0.499	0.247	0.553
	(0.029)	(0.029)	(0.029)	(0.015)	(0.031)
Controls	No	No	No	No	No
Observations	435	2,175	2,175	2,175	435
R-squared	0.029	0.033	0.031	0.019	0.055

B. With Controls

	Accepted Job Offer	Submission	On-Time Submission	Accuracy	Re-Application
Non-Referred	-0.064	-0.090*	-0.088*	-0.028	-0.121**
	(0.055)	(0.048)	(0.047)	(0.026)	(0.056)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	435	2,175	2,175	2,175	435
R-squared	0.126	0.130	0.132	0.100	0.132

C. Conditional on Accepting Job Offer, With Controls

	Submission	On-Time Submission	Accuracy	Re-Application
Non-Referred	-0.104*	-0.104*	-0.026	-0.131*
	(0.057)	(0.058)	(0.033)	(0.071)
Controls	Yes	Yes	Yes	Yes
Mean of Dependent Variable				
Base Group (Referred Workers)	0.763	0.735	0.363	0.815
Observations	1,325	1,325	1,325	265
R-squared	0.096	0.098	0.063	0.088

Notes: Each column in each panel reports the results of a different regression using data from the supplemental experiment. The dependent variable, indicated by the column, is regressed on an indicator for being a non-referred worker. (The omitted group is referred workers.) Regressions in Panel A include no controls; regressions in Panels B and C include the controls for worker characteristics listed in footnote 20. Panels A and B include all workers to whom we made employment offers; Panel C includes only workers who accepted these offers. Observations in the first and last columns (*Accepted, Re-Application)* are workers while observations the middle three columns of regressions (*Submission, On-Time Submission, Accuracy)* are worker-days. *Accepted Job Offer* is an indicator for whether the worker accepted our invitation to work for the new firm created for the supplemental experiment. *Submission, Accuracy,* and *Re-Application* are defined as in the individual experiment (see the notes to Table 3). *On-Time Submission* is an indicator for whether the worker submitted work before midnight on the assigned day. Huber-White standard errors are presented when observations are workers and standard errors are clustered at the worker level when observations are worker-days. *, ** denote significance at the 10% and 5% levels.

Table 5. Individual Diligence by Team Type: Referred and Non-Referred Workers Only
Team Experiment: Base Group is Referred Workers Paired with Someone Else's Referrer (Type B)

A. No Controls

	Logged in	Submitted	Individual Question Correct	Own Criteria in Slogan
Referred Worker Paired with Own Referrer (Type A)	0.018	0.046**	0.053*	0.004
	(0.017)	(0.017)	(0.028)	(0.033)
Non-Referred Worker Paired with Referrer (Type C)	-0.294**	-0.312**	-0.287**	-0.138**
	(0.051)	(0.048)	(0.049)	(0.052)
Constant	0.883	0.837	0.755	0.440
	(0.016)	(0.018)	(0.023)	(0.028)
Controls	No	No	No	No
Observations	846	846	846	846
R-Squared	0.124	0.134	0.102	0.018

B. With Controls

	Logged in	Submitted	Individual Question Correct	Own Criteria in Slogan
Referred Worker Paired with Own Referrer (Type A)	0.018	0.046**	0.053*	0.004
	(0.017)	(0.017)	(0.029)	(0.034)
Non-Referred Worker Paired with Referrer (Type C)	-0.194**	-0.229**	-0.245**	-0.087
	(0.055)	(0.053)	(0.052)	(0.058)
Controls	Yes	Yes	Yes	Yes
Observations	846	846	846	846
R-Squared	0.188	0.187	0.147	0.066

Notes: Each column in each panel reports the results of a separate regression of the dependent variable indicated by the column on an indicator for being in a Type A team (which consists of a referred worker paired with her own referrer) and an indicator for being in a Type C team (which consists of a non-referred worker paired with a referrer). (The omitted group is Type B teams, which consist of a referred worker paired with someone else's referrer.) Observations are at the worker-PSA level; only referred and non-referred workers (not referrers) are included. Data come from the team experiment. Regressions in Panel A includes no controls; regressions in Panel B control for the worker characteristics in footnote 20. *Logged In* is an indicator for whether the worker logged on to our task site in the allocated window (48 hours). *Submitted* is an indicator for submitting work in that period. *Individual Question Correct* is an indicator for the worker providing an accurate response to the question about her own reading. *Own Criteria in Slogan* is an indicator for the worker submitting a slogan that met the criteria in her own instructions. Standard errors are clustered at the blocking group level. *, ** denote significance at the 10% and 5% levels.

Table 6. Team Performance by Team Type

Team Experiment: Base Group is Referred Workers Paired with Someone Else's Referrer (Type B)

A. No Controls

	Both Submitted	Team Question Matches	Same Slogan	Same Slogan & Both Criteria
Referred Worker Paired with Own Referrer (Type A)	0.099**	0.287**	0.372**	0.103**
	(0.023)	(0.028)	(0.032)	(0.024)
Non-Referred Worker Paired with Referrer (Type C)	-0.280**	-0.206**	-0.142**	-0.053*
	(0.048)	(0.047)	(0.041)	(0.029)
Constant	0.730	0.496	0.337	0.142
	(0.027)	(0.030)	(0.029)	(0.022)
Controls	No	No	No	No
Observations	846	846	846	846
R-Squared	0.117	0.164	0.194	0.031

B. With Controls

	Both Submitted	Team Question Matches	Same Slogan	Same Slogan & Both Criteria
Referred Worker Paired with Own Referrer (Type A)	0.099**	0.287**	0.372**	0.103**
	(0.023)	(0.028)	(0.032)	(0.024)
Non-Referred Worker Paired with Referrer (Type C)	-0.206**	-0.165**	-0.108**	-0.031
	(0.053)	(0.048)	(0.044)	(0.031)
Controls	Yes	Yes	Yes	Yes
Observations	846	846	846	846
R-Squared	0.155	0.193	0.213	0.055

Notes: Each column in each panel reports the results of a separate regression of the dependent variable indicated by the column on an indicator for being in a Type A team (which consists of a referred worker paired with her own referrer) and an indicator for being in a Type C team (which consists of a non-referred worker paired with a referrer). (The omitted group is Type B teams, which consist of a referred worker paired with someone else's referrer.) Observations are at a team-PSA level and data come from the team experiment. Regressions in Panel A include no controls; regressions in Panel B control for the characteristics of referred and non-referred workers listed in footnote 20. *Both Submitted* is an indicator for whether both partners sent in responses in the allocated time window (48 hours). *Team Question Matches* is an indicator for whether both partners chose, as instructed, the same answer to the multiple choice question. *Same Slogan* is an indicator for whether both partners entered the same slogan as instructed. *Same Slogan & Both Criteria* is an indicator for whether both partners submitted the same slogan which met the criterion in each of their information sheets. Standard errors are clustered at the blocking group level. *, ** denote significance at the 10% and 5% levels.

Table 7. Time Spent & Wanting to Partner Again, by Team Type
Team Experiment: Base Group is Referred Workers Paired with Someone Else's Referrer (Type B)

	A. Time Spent (Minutes)		B. Wants to Partner Again	
	Referrers	Referred & Non-Referred Workers	Referrers	Referred & Non-Referred Workers
Referred Worker Paired with Own Referrer (Type A)	5.922**	5.142**	0.556**	0.451**
	(1.752)	(1.559)	(0.030)	(0.033)
Non-Referred Worker Paired with Referrer (Type C)	1.135	-15.532**	-0.100**	0.009
	(1.445)	(3.121)	(0.041)	(0.060)
Constant	37.482	38.723	0.406	0.477
	(1.291)	(1.492)	(0.029)	(0.031)
Controls	No	No	No	No
Observations	846	846	717	612
R-squared	0.009	0.087	0.338	0.211

Notes: Each column in each panel reports the results of a separate regression of the dependent variable indicated by the column on an indicator for being in a Type A team (which consists of a referred worker paired with her own referrer) and an indicator for being in a Type C team (which consists of a non-referred worker paired with a referrer). (The omitted group is Type B teams, which consist of a referred worker paired with someone else's referrer.) No controls are included. Observations are at a worker-PSA level and data come from the team experiment. The dependent variable in Panel A is *Time Spent:* the number of minutes the worker logged on oDesk for the task. The dependent variable in Panel B, *Wants to Partner Again* , is an indicator for whether, after completion of all three tasks, the worker reported wanting to work again with the partner from that particular pairing. The first regression in each panel includes only referrers while the second includes only referred and non-referred workers. Standard errors are clustered at the blocking group level. *, ** denote significance at the 10% and 5% levels.

Table 8. Team Communication and Performance Controlling for Communication and Time Spent
Team Experiment: Base Group is Referred Workers Paired with Someone Else's Referrer (Type B)

A. Communication by Team Type

	Chat Box Use	Total Chat Messages (Conditional on Use)	Answered Communication Question	Reported Outside Communication
Referred Worker Paired with Own Referrer (Type A)	0.025	4.346*	-0.028*	0.376**
	(0.042)	(2.544)	(0.016)	(0.040)
Non-Referred Worker Paired with Referrer (Type C)	-0.090*	0.805	-0.032	-0.043
	(0.047)	(2.706)	(0.027)	(0.038)
Controls	Yes	Yes	Yes	Yes
Mean of Dependent Variable Base Group (Type B)	0.408	13.522	0.947	0.378
Observations	846	307	846	778
R-Squared	0.047	0.062	0.017	0.193

B. Team Performance Controlling for Communication and Time Spent

	Both Submitted	Team Question Matches	Same Slogan	Same Slogan & Both Criteria
Referred Worker Paired with Own Referrer (Type A)	-0.028	0.140**	0.225**	0.025
	(0.026)	(0.031)	(0.037)	(0.029)
Non-Referred Worker Paired with Referrer (Type C)	-0.144**	-0.112**	-0.066*	-0.014
	(0.035)	(0.041)	(0.038)	(0.029)
Used Chat Box	0.223**	0.167**	0.116**	0.020
	(0.028)	(0.037)	(0.041)	(0.041)
Total Chat Messages	0.001*	0.005**	0.006**	0.003*
	(0.001)	(0.001)	(0.001)	(0.002)
Used Outside Communication	0.236**	0.320**	0.336**	0.173**
	(0.030)	(0.033)	(0.042)	(0.031)
Minutes Spent by Referrer	0.003**	0.002**	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.000)
Minutes Spent by Referred or Non-Referred Worker	0.005**	0.003**	0.002**	0.001**
	(0.001)	(0.001)	(0.000)	(0.000)
Controls	Yes	Yes	Yes	Yes
Mean of Dependent Variable Base Group (Type B)	0.730	0.500	0.337	0.142
Observations	846	846	846	846
R-Squared	0.497	0.478	0.467	0.171

Notes: Each column in each panel reports the results of a separate regression of the dependent variable indicated by the column on an indicator for being in a Type A team (which consists of a referred worker paired with her own referrer) and an indicator for being in a Type C team (which consists of a non-referred worker paired with a referrer). (The omitted group is Type B teams, which consist of a referred worker paired with someone else's referrer.) Observations are at a worker-PSA level and data come from the team experiment. *Chat Box Use* is an indicator for whether each partner typed at least one message in the chat box. *Total Chat Messages* is the aggregate number of messages sent between the two partners, and is

conditional on chat box use. *Answered Communication Question* is an indicator for whether at least one partner responded to the question at the end of that task about how the partners had communicated. *Reported Outside Communication* is an indicator for whether either partner reported communicating using methods other than the chat box and is conditional on at least one partner having answered the communication question. Dependent variables in Panel B are as defined in Table 6. Regressions in both panels control for the characteristics of referred and non-referred workers listed in footnote 20. Regressions in Panel B also control for whether the team used the chat box, the number of messages in the chat box, whether they reported using other forms of communication, and (separately) the number of minutes spent by both partners. Standard errors are clustered at the blocking group level. *, ** denote significance at the 10% and 5% levels.

Table 9. Relationship between Referred Worker's Performance and Referrer's Performance
Individual and Supplemental Experiments

	Dependent Variable: Referred Worker's Submission Rate, Individual Experiment	Dependent Variable: Referred Worker's Submission Rate, Supplemental Experiment		
Referrer's Submission Rate,	0.421**	0.246**	0.132	
Individual Experiment	(0.066)	(0.079)	(0.082)	
Individual Experiment				0.409**
Individual Experiment				(0.078)
Constant	0.456	0.331	0.222	0.201
	(0.059)	(0.065)	(0.421)	(0.063)
Controls	No	No	Yes	No
Observations	255	255	255	255
R-squared	0.192	0.034	0.184	0.087

Notes: Each column presents the results of a regression of the dependent variable indicated by the column on the independent variable indicated by the row. Each observation is a referred worker. Data come from the individual and supplemental experiments. Huber-White standard errors are in parenthesis. No controls are included except in the second-to-last column, which includes controls for referred worker characteristics listed in footnote 20. *, ** denote significance at the 10% and 5% levels.

Table 10. Relationship Strength, Observable Characteristics, and Performance
Individual, Supplemental, and Team Experiments: Referred Workers

A. Observable Characteristics

	Tests Passed	Number of Assignments	Feedback Score
Relationship Strength Index	-0.462**	-1.456**	-0.038
	(0.138)	(0.579)	(0.030)
Constant	4.447	5.774	4.545
	(0.125)	(0.593)	(0.041)
Controls	No	No	No
Observations	532	532	293
R-squared	0.025	0.011	0.003

B. Performance, No Controls

	Submission (Individual Experiment)	Submission Experiment)	Same Slogan (Team Experiment)
Relationship Strength Index	0.041**	0.018	0.053**
	(0.020)	(0.027)	(0.020)
Constant	0.780	0.518	0.514
	(0.021)	(0.029)	(0.022)
Controls	No	No	No
Observations	1,512	1,260	560
R-squared	0.012	0.003	0.009

C. Performance, With Controls

	Submission (Individual Experiment)	Submission Experiment)	Same Slogan (Team Experiment)
Relationship Strength Index	0.047**	0.037	0.051**
	(0.020)	(0.026)	(0.022)
Controls	Yes	Yes	Yes
Observations	1,512	1,260	560
R-squared	0.106	0.155	0.048

Notes: Each column in each panel reports the results of a separate regression of the dependent variable indicated by the column on an index for the strength of the referrer-referred worker relationship (defined in Section 6 of the text). All regressions in the table include only referred workers. Regressions in Panel A include referred workers in both the individual and team experiments. Observations are at the worker level. No controls are included; Huber-White standard errors are in parentheses. *Tested Passed* is the number of oDesk tests a worker passed, *Number of Assignments* is the number of prior oDesk assignments the worker had, and *Feedback Score* is a one-to-five rating given by employers on previous assignments. Regressions in Panel B include no controls, while regressions in Panel C include the controls for worker characteristics listed in footnote 20. The first two columns of Panels B and C include workers from only the individual and supplemental experiments, respectively. In these columns, outcomes are observed at the worker-day level and standard errors are clustered by worker. The final column of Panels B and C includes only workers from the team experiment; outcomes are observed at the worker-PSA level and standard errors are clustered by blocking group. *, ** denote significance at the 10% and 5% levels.

Table 11. Simulated Hiring

Individual Experiment: Assuming Top 50% of Applicants Hired

	A. Fraction Hired		B. Measure of Observables		C. Actual Submission Rate		
	Referred Applicants	Non-Referred Applicants	Hired Referred Workers	Hired Non-Referred Workers	Hired Referred Workers	Hired Non-Referred Workers	Difference
No Information (Applicant Pool Average)	50%	50%	73%	68%	77%	63%	15%**
Observe Characteristics Only	58%	39%	80%	79%	82%	64%	19%**
Observe Characteristics & Referral Status	77%	12%	77%	86%	81%	75%	5%

Notes: Each row presents the results of simulating hiring under a different hiring scenario. The first row simulates hiring under the assumption that employers hire at random. The second row simulates hiring if employers observed only workers' resume characteristics, but not their referral status. The third row simulates hiring assuming employers observed workers' resume characteristics and referral status. Details of the simulation are provided in Section 7 of the text. Panel A presents the fraction of referred and non-referred workers that would have been hired under each scenario. Panel B presents a summary measure of the observables of the referred and non-referred workers who would have been hired under each scenario. (This is the average predicted probability of submission, based on the observable characteristics listed in footnote 20. It is described further in the text.) Panel C presents the actual submission rate of the referred and non-referred workers who would have been hired under each scenario. The column labeled *Difference* provides the difference in average submission rates between the referred and non-referred workers who would have been hired under each scenario. *, ** denote that this difference is significant at the 10% and 5% levels.

Appendix Figure 1. Performance Report Example

Me
Feb 01

Dear [redacted],

We are writing to update you, as promised, on the performance of a colleague you referred.

Yesterday,
[redacted] did not submit their work.
[redacted] did not submit their work by 11AM PHT.
[redacted] did not answer all of the questions.

Cumulative Performance :
* Percentage of days work was submitted : 80%
* Percentage of days work was submitted by 11AM PHT : 60%
* Percentage of days all questions were answered : 80%

Sincerely,
The Hiring Team

Appendix Table 1. Performance and Persistence, Including Second Order Controls
Individual Experiment: Base Group is Non-Monitored Referred Workers (Treatment 2)

A. All Days

	Submission	On-Time Submission	Accuracy	Re-Application
Monitored Referred	0.012	0.043	0.015	-0.027
(Treatment 1)	(0.042)	(0.046)	(0.040)	(0.039)
Non-Referred	-0.143**	-0.080	-0.109**	-0.184**
	(0.050)	(0.052)	(0.046)	(0.054)
Controls	Yes	Yes	Yes	Yes
Second Order Controls	Yes	Yes	Yes	Yes
Observations	2,610	2,610	2,610	435
R-squared	0.190	0.160	0.180	0.270

B. Last Day Only, Controlling for Performance on First Five

	Submission	On-Time Submission	Accuracy	Re-Application
Monitored Referred	-0.045	0.023	-0.032	-0.041
(Treatment 1)	(0.049)	(0.057)	(0.043)	(0.039)
Non-Referred	-0.161**	-0.104*	-0.090*	-0.125**
	(0.054)	(0.057)	(0.046)	(0.050)
Day 1 Performance	0.003	0.114*	0.023	
	(0.074)	(0.061)	(0.067)	
Day 2 Performance	0.031	0.030	0.079	
	(0.062)	(0.059)	(0.064)	
Day 3 Performance	0.063	0.208**	0.141	
	(0.084)	(0.078)	(0.086)	
Day 4 Performance	0.222**	0.140*	0.171*	
	(0.097)	(0.075)	(0.090)	
Day 5 Performance	0.460**	0.223**	0.401**	
	(0.083)	(0.079)	(0.085)	
Controls	Yes	Yes	Yes	Yes
Second Order Controls	Yes	Yes	Yes	Yes
Observations	435	435	435	435
R-squared	0.617	0.510	0.625	0.452

Notes: Panels A and B replicate Panels B and C, respectively, of Table 3 with additional control variables. In addition to the characteristics listed in footnote 20, these regressions all include *Second Order Controls:* the square of each non-binary characteristic in footnote 20 and the interaction of each pair of characteristics in footnote 20 (an additional 88 variables). *, ** denote significance at the 10% and 5% levels.

Appendix Table 2. Selection into Accepting Job Offer in the Supplemental Experiment
Data from the Individual and Supplemental Experiments

	Dependent Variable: Accepted Job Offer in Supplemental Experiment			
	Submission	On-Time Submission	Accuracy	Re-Application
Performance in Individual Experiment × Non-Referred	0.131	0.133	0.219	0.008
	(0.129)	(0.123)	(0.135)	(0.143)
Performance in Individual Experiment	0.164*	0.113	0.159	0.217*
	(0.096)	(0.081)	(0.099)	(0.119)
Non-Referred	-0.122	-0.114	-0.165*	-0.026
	(0.106)	(0.086)	(0.098)	(0.136)
Controls	Yes	Yes	Yes	Yes
Observations	435	435	435	435
R-squared	0.154	0.144	0.161	0.149

Notes: Each column presents the results of a separate regression of an indicator for a worker accepting our job offer in the supplemental experiment on a measure of their performance in the individual experiment, an indicator for being a non-referred worker, and the interaction of these two indicators. Each column uses a different performance measure indicated by the column heading. These metrics are defined in the notes of Table 3. Observations are workers; only referred and non-referred workers are included. Each regression contains controls for the individual characteristics listed in footnote 20 . *, ** denote significance at the 10% and 5% levels.

Appendix Table 3. Performance and Persistence in New Firm, Including Second Order Controls
Supplemental Experiment: Base Group is All Referred Workers

A. All Workers

	Accepted Job Offer	Submission	On-Time Submission	Accuracy	Re-Application
Non-Referred	-0.042	-0.073	-0.073	-0.020	-0.097
	(0.064)	(0.055)	(0.055)	(0.029)	(0.064)
Controls	Yes	Yes	Yes	Yes	Yes
Second Order Controls	Yes	Yes	Yes	Yes	Yes
Observations	435	2,175	2,175	2,175	435
R-squared	0.268	0.242	0.239	0.201	0.292

B. Conditional on Accepting Job Offer

	Submission	On-Time Submission	Accuracy	Re-Application
Non-Referred	-0.119*	-0.118*	-0.047	-0.198**
	(0.067)	(0.070)	(0.038)	(0.090)
Controls	Yes	Yes	Yes	Yes
Second Order Controls	Yes	Yes	Yes	Yes
Mean of Dependent Variable				
Base Group (Referred Workers)	0.763	0.735	0.363	0.815
Observations	1,325	1,325	1,325	265
R-squared	0.247	0.25	0.192	0.358

Notes: Panels A and B replicate Panels B and C, respectively, of Table 4 with additional control variables. In addition to the characteristics listed in footnote 20, these regressions all include *Second Order Controls:* the square of each non-binary characteristic in footnote 20 and the interaction of each pair of characteristics in footnote 20 (an additional 88 variables). *, ** denote significance at the 10% and 5% levels.

Appendix Table 4. Performance in Team Experiment, Including Second Order Controls
Team Experiment: Base Group is Referred Workers Paired with Someone Else's Referrer (Type B)

	A. Individual Diligence by Team Type (Referred and Non-Referred Workers Only)			
	Logged in	Submitted	Individual Question Correct	Own Criteria in Slogan
Referred Worker Paired with Own Referrer (Type A)	0.018	0.046**	0.053*	0.004
	(0.018)	(0.018)	(0.030)	(0.035)
Non-Referred Worker Paired with Referrer (Type C)	-0.131**	-0.176**	-0.193**	-0.044
	(0.058)	(0.056)	(0.057)	(0.051)
Controls	Yes	Yes	Yes	Yes
Second Order Controls	Yes	Yes	Yes	Yes
Observations	846	846	846	846
R-Squared	0.420	0.388	0.309	0.213

	B. Team Performance by Team Type			
	Both Submitted	Team Question Matches	Same Slogan	Same Slogan & Both Criteria
Referred Worker Paired with Own Referrer (Type A)	0.099**	0.287**	0.372**	0.103**
	(0.024)	(0.030)	(0.034)	(0.025)
Non-Referred Worker Paired with Referrer (Type C)	-0.162**	-0.075	-0.032	0.026
	(0.061)	(0.054)	(0.054)	(0.031)
Controls	Yes	Yes	Yes	Yes
Second Order Controls	Yes	Yes	Yes	Yes
Observations	846	846	846	846
R-Squared	0.314	0.327	0.311	0.170

Notes: Panels A and B replicate Panel B of Table 5 and Panel B of Table 6, respectively, with additional control variables. In addition to the characteristics listed in footnote 20, these regressions all include *Second Order Controls:* the square of each non-binary characteristic in footnote 20 and the interaction of each pair of characteristics in footnote 20 (an additional 88 variables). *, ** denote significance at the 10% and 5% levels.

Appendix Table 5. Relationship Between Referrer Characteristics and Referred Worker Characteristics
Individual and Team Experiments

	Dependent Variable: Referred Worker Characteristic										
	Earnings	Number of Previous Jobs	Feedback Score	Posted Wage	Days Since Joining oDesk	Has Portfolio	Number of Tests Passed	English Score	Agency Affiliated	Number of Degrees	Proposed Wage
Referrer Characteristic	0.060**	0.095**	0.204	0.231**	0.264**	0.272**	0.157**	0.285**	0.317**	0.075*	0.367**
	(0.022)	(0.043)	(0.171)	(0.051)	(0.045)	(0.045)	(0.037)	(0.050)	(0.033)	(0.041)	(0.047)
Constant	688.6	4.637	3.562	1.928	289.95	0.296	3.580	3.319	0.060	1.238	1.425
	(110.6)	(0.790)	(0.822)	(0.152)	(34.01)	(0.037)	(0.253)	(0.243)	(0.014)	(0.070)	(0.120)
Observations	537	537	296	537	537	537	537	533	537	537	537
R-Squared	0.014	0.009	0.005	0.036	0.060	0.064	0.032	0.057	0.146	0.006	0.103

Notes: Each column reports the results of regressing the value of an observable characteristic for a referred worker on the value of the same characteristic for her referrer. Each column corresponds to a different characteristic, indicated by the column header. All 537 hired referred workers from the individual and team experiments are included, although *Feedback Score* and *English Score* are missing for some workers. *English Score* is self-reported English ability on a one-to-five scale, a *portfolio* is where a worker posts prior work, and *agency-affiliated* workers pay a fraction of their earnings to report they are part of a given group of oDesk workers (an agency). Standard errors are clustered by referrer. *, ** denote significance at the 10% and 5% levels.

Appendix Table 6. Characteristics of the Referrer-Referred Worker Relationship
Individual and Team Experiments

	Included Referred Workers	Excluded Referred Workers
How Well Referrer Knows Referral		
1 (Hardly at all)	1%	2%
2	2%	2%
3	5%	3%
4	14%	9%
5	20%	19%
6 (Extremely Well)	57%	65%
Observations	535	1,314
Frequency of Interaction		
Less than Once a Month	2%	4%
About Once a Month	5%	4%
Less than Weekly, More than Monthly	8%	4%
About Once a Week	13%	9%
Less than Daily, More than Weekly	21%	14%
About Once a Day	19%	16%
More than Once a Day	32%	47%
Observations	533	1,311
Number of People Known in Common		
0 to 4	21%	18%
5 to 9	16%	16%
10 to 19	16%	18%
20 to 29	11%	10%
30 or more	37%	39%
Observations	535	1,314
Sometimes Work in Same Room	0%	44%
Observations	537	1,317

Notes: This table presents the distributions of referrers' responses to questions about their relationships with their referrals for two different samples, indicated by the column headings. *Included Referred Workers* are referred workers we hired in either the individual or team experiment. *Excluded Referred Workers* are workers who were referred to us, but whom we did not hire in either experiment.