

NBER WORKING PAPER SERIES

WHY THE REFERENTIAL TREATMENT:
EVIDENCE FROM FIELD EXPERIMENTS ON REFERRALS

Amanda Pallais
Emily Glassberg Sands

Working Paper 21357
<http://www.nber.org/papers/w21357>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2015

We would like to thank David Autor, Felipe Barrera-Osorio, Patrick Bayer, 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, John List, Ben Schoefer, Sarah Turner, Marty West, seminar participants at Berkeley, Booth, Brookings Institution, Columbia, Duke, Harvard, Kellogg, and NBER Summer Institute Labor Studies, the New York Federal Reserve, Princeton, RAND, University of British Columbia, University of Chicago, and Wharton, as well as Jesse Shapiro and four anonymous referees 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@coursera.org. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Amanda Pallais and Emily Glassberg Sands. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Why the Referential Treatment: Evidence from Field Experiments on Referrals
Amanda Pallais and Emily Glassberg Sands
NBER Working Paper No. 21357
July 2015
JEL No. C93,J24,J63,M51

ABSTRACT

Referred workers are more likely than non-referred workers to be hired, all else equal. In three field experiments in an online labor market, we examine why. We find that referrals contain positive information about worker performance and persistence that is not contained in workers' observable characteristics. We also find that referrals performed particularly well when working directly with their referrers. However, we do not find evidence that referrals exert more effort because they believe their performance will affect their relationship with their referrer or their referrer's position at the firm.

Amanda Pallais
Department of Economics
Harvard University
Littauer Center
Cambridge, MA 02138
and NBER
apallais@fas.harvard.edu

Emily Glassberg Sands
Coursera, Inc.
381 East Evelyn Avenue
Mountain View, CA 94041
emily@coursera.org

1 Introduction

A large empirical literature has shown that the majority of jobs are found through informal contacts, firms are more likely to hire applicants referred by current employees than non-referred applicants, and some firms even give bonuses to employees for successful referrals.¹ Yet the literature remains divided on *why* firms draw so heavily on referred applicants. Referrals may provide (positive) information about worker quality, or being referred may induce a worker to work harder or more productively; alternatively, firms may hire referrals for nepotistic reasons or to decrease recruiting costs.² This paper analyzes a set of field experiments in an online labor market to answer two open questions about referrals: first, do referrals contain information about worker productivity? Second, do referred workers work harder or more effectively because they are referred?

Answering the first of these questions with observational data is difficult because we only observe the productivity of workers who are hired. If referrals provide information about worker quality and firms (rationally) incorporate this information into their hiring decisions, hired referred workers may not perform better than hired non-referred workers, even though the referral provides positive information about worker productivity.

Our experiments circumvent differential selection of referred and non-referred workers into employment. By working in an online marketplace (oDesk), we were able to hire workers directly, allowing us to compare the performance of referred and non-referred *applicants*, not just the workers a given firm chose to hire. oDesk has over 2.5 million workers (Horton, 2013a) and 35 million hours billed in 2012 (oDesk Corporation, 2013). The experiments took place between January and June 2013. We ran three experiments: the peer influence experiment, the team experiment, and the selection experiment. To recruit our samples for the peer influence and team experiments, we first hired experienced workers, asked them to complete a short task unrelated to the experimental tasks, and solicited referrals from those who complied. We then invited referred workers and a random sample of non-referred workers to apply, and hired all applicants who met our basic wage criteria. These two experiments were designed primarily to answer whether referred workers perform better *because* they are referred: either because (1) they work harder because they think their performance will affect their referrers' position at the firm or their relationship with their referrers (peer influence)³

¹See, e.g., Bewley (1999); Ioannides and Datcher Loury (2004); Granovetter (1995); Topa (2011); Fernandez and Weinberg (1997); Peterson et al. (2000); Brown et al. (2012); and Burks et al. (2015).

²See, e.g., Montgomery (1991); Simon and Warner (1992); Heath (2015); Kugler (2003); and Wang (2013).

³Peer influence leads referrals to work harder in Kugler's (2003) model because referrals face a psychic cost of exerting less effort than their referrers, while Heath's (2015) and Dhillon et al.'s (2012) models suggests that referred workers work hard because if they perform poorly the firm will punish their referrers. This is similar to microfinance group lending wherein a worker's peers may pressure the worker to repay the loan

or (2) they perform better when working directly with their referrers (team production).⁴ Four months later, we conducted the selection experiment, designed to test whether referrals perform better than non-referred workers even without on-the-job interactions with their referrers. We made job offers from a new firm to all referred and non-referred workers (but not referrers) in the peer influence experiment.

We find that referrals do reveal positive information about worker quality independent of on-the-job interactions with referrers. In the selection experiment, referred workers exhibited substantially higher performance and lower turnover than did non-referred workers even at a firm to which they had not been referred and at which their referrers did not work. Little to none of the information contained in the referral was otherwise observable to the employer through workers' resumes.

The peer influence experiment provides additional evidence that referrals contain information about worker quality. In this experiment, referred and non-referred workers tested an airline flight website by answering questions about the site every other day over 12 days. 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. Referrals in the peer influence experiment were randomized into two treatments. The non-monitoring treatment was designed to minimize peer influence. Referrals in this treatment were told their referrers would never know their performance and (after referring) referrers were told they not be judged on the performance of their referrals. As in the selection experiment, referred workers in this treatment performed better and had less turnover than non-referred workers, and these differences could not have been predicted from workers' observable characteristics. We also use data from this experiment to simulate a realistic hiring process and to show that we could have obtained misleading results if we had only compared the performance of applicants employers chose to hire.

The monitoring treatment of the peer influence experiment was designed to maximize peer influence. Each referrer in this treatment received an update on her referral's performance after each day of work. We implied to each referrer that her referral's performance and willingness to continue working for us would affect whether the referrer was promoted. Yet, we do not find that monitored referrals performed significantly better or had less turnover than non-monitored referrals.

(e.g., Bryan et al., 2014).

⁴While team production has not been emphasized as an explanation for hiring referrals in the economics literature, general research on team production implies it may be an important benefit of referrals. For example, Bandiera et al.'s (2013) model finds that when working in teams with their friends, workers are less likely to free-ride; 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) finds that Civil War soldiers were less likely to desert when more of their unit was from their own birthplace.

The team experiment, however, does suggest that working directly with her referrer makes a referral more productive. Here, 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 exactly three words long). We asked the partners to use the chat box provided on the site to discuss the task and then to each submit the same slogan, which should have satisfied both criteria. Workers completed three such PSA tasks, each with a different partner. Importantly, each referral completed one task with her referrer and one task with another randomly-chosen referrer. Referred workers performed substantially better when paired with their own referrers.

An important caveat to our findings is that while the job tasks performed on oDesk are representative of important parts of the US economy, employer-employee relationships on oDesk are typically much shorter than those in offline labor markets. We discuss implications of these differences for the interpretation of our findings in Section 5 of the paper.⁵

We see our results as reconciling the seemingly inconsistent findings from papers comparing the performance of referred and non-referred workers. Among call center workers, Castilla (2005) finds that referred workers perform better than non-referred workers, while among bank tellers, Blau (1990) finds that referred workers perform worse. Studying nine firms in three different industries, Burks et al. (2015) finds that referred workers perform similarly to non-referred workers on most metrics, though they have less turnover. We show that referrals contain information about worker quality, but that if employers utilize that information in the hiring process, hired referred workers could perform better than, worse than, or the same as non-referred workers.⁶

Other papers directly test predictions of models where referrals contain information about worker quality. Using firm data, Brown et al. (2012) finds results consistent with these models: referred applicants are more likely to be hired and hired referrals have lower turnover and higher initial wages, though the wage advantage decreases over time. Dustmann et al. (2011) finds similar results using matched employer-employee data and ethnic-minority groups to proxy for referrals. Inconsistent with these models, however, Pistaferri (1999) and Bentolila et al. (2010) find that workers who find jobs through informal networks earn lower wages. Our paper adds to this literature by directly analyzing worker performance and by

⁵There are a number of recent papers that use oDesk to learn about general features of labor markets. See, for example, Horton (2013b), Ghani et al. (2014), Horton (2014), Lyons (2014), Pallais (2014), Stanton and Thomas (2014), and Gilchrist et al. (2015).

⁶Burks et al. (2015) shows that referred applicants are more likely to be offered jobs, all else equal. This is consistent with a model in which referrals contain information about worker quality, but also with other models like nepotism. The paper also finds that referred workers are more likely to accept job offers.

constructing a setting (the selection experiment) in which referrals’ superior performance can’t result from on-the-job interactions with referrers. In an experiment, Beaman and Magruder (2012) find that, when told they will be paid on their referrals’ performance, employees refer higher-performing workers. Our paper builds on this by showing that referred applicants perform better than non-referred applicants. Finally, Heath (2015) finds that referrers’ and referrals’ wage changes are highly correlated, consistent with a peer influence mechanism.⁷

There is also closely related research that uses the oDesk platform. Stanton and Thomas (2014) carefully analyzes oDesk agencies, formal groups of oDesk workers often formed through offline connections. Agency-affiliated workers pay a fraction of their earnings to their agency and, in return, their agency affiliation is listed on their resume. The paper finds that employers view agency-affiliation as a signal that inexperienced workers are productive. Among inexperienced workers, employers are more likely to hire agency-affiliates than unaffiliated workers and they pay affiliates higher wages. Once workers have accumulated other signals of productivity (in particular employer feedback scores), the importance of this signal declines. A related paper, Horton (2013b), finds that oDesk employers value recommendations of whom to hire. Employers who randomly received recommendations about workers from oDesk itself were both more likely to hire these workers and more likely to hire anyone for their jobs. Yet, workers hired as a result of these recommendations were not more successful than other hired workers.

The remainder of the paper proceeds as follows. Section 2 describes the marketplace and our experimental designs. Section 3 analyzes whether referrals contain information about worker quality, Section 4 examines whether referrals perform better because they were referred, and Section 5 discusses external validity. Section 6 concludes, discussing how these results could inform strategies to improve unconnected workers’ labor market outcomes.

2 Experimental Context and Recruitment Design

2.1 Online Labor Market

oDesk is an online labor market where employers, mostly from the United States, hire independent contractors from all over the world for jobs that can be completed remotely. The jobs

⁷A few papers suggest firms prefer referrals for reasons other than improved productivity. Consistent with firms hiring workers’ children as a favor to existing workers, Kramarz and Skans (2014) finds parents’ wage growth drops dramatically exactly when one of their children is hired. Wang (2013) also finds evidence of nepotism in referrals. Holzer (1987) and Burks et al. (2015) find that hiring referred workers lowers recruiting costs.

range from those that require significant skills such as computer programming or software development to less skill-intensive tasks such as data entry, internet research, or administrative support. Unlike Amazon’s Mechanical Turk, another online marketplace commonly used in economics research, oDesk employers have complete discretion in whom they hire and they have real relationships with hired workers.

Employers post job listings describing their jobs and any required worker characteristics. They consider applicants’ resumes when deciding whom to hire. (Figure 1 shows a sample oDesk resume from a worker not in the experiment.) These resumes contain information about workers’ skills and qualifications as well as their past experience. The resumes list previous oDesk jobs, educational degrees, skills tests that workers have passed, and a one-to-five feedback score from previous employers. Employers can also choose to interview workers remotely before deciding whom to hire, though many employers do not.

Most jobs on oDesk, including all the jobs in this experiment, are hourly jobs (Pallais, 2014). In these jobs, workers propose an hourly wage when they apply. Workers are then paid their set hourly wage for all hours worked, regardless of the output, though the employer can end the job and fire the worker at any time. Workers also post a desired hourly wage at the top of their resumes, which firms can observe.

During the employment relationship, workers and employers communicate through the oDesk messaging system. They also use non-oDesk methods such as email and Skype. oDesk allows employers to monitor workers’ progress, similar to the monitoring that would be possible in an in-person environment. Workers log into an oDesk application that shows employers when they are working. This application provides information about workers’ keystroke volume and shows screen shots of the workers’ computers, taken six times per hour.

Most workers state that they are available to work full-time (30+ hours per week), though others are available part-time or only a few hours per week.⁸ In general, oDesk workers are relatively young and well-educated, and, among the lower-wage segment employed in these experiments, disproportionately likely to be female. 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 current workers and workers can recommend people they know to their employers. oDesk also has agencies, formal groups of oDesk workers often formed through offline connections (Stanton and Thomas, 2014).

⁸This statistic is from personal correspondence with John Horton and is based on calculations using oDesk administrative data.

2.2 Hiring our Experimental Samples

We hired workers for the peer influence and team experiments in the same way. (The sample for the selection experiment was a subset of the peer influence experiment sample.) 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 resumes, (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. We eliminated workers with ratings below four because we wanted only 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. We only included workers from the Philippines because we wanted all workers in the team experiment to be able to communicate easily and be in the same time zone and the Philippines is the most common country of residence for low-wage oDesk workers.⁹ 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.¹⁰

Original hires were asked 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." We did not provide workers with financial incentives for referring.¹¹ On each referral form we included questions about how well the referrer knew her referral, how often they interacted (remotely 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 or collaborated with referrals working in the same room, we eliminated from our

⁹That the Philippines is the most common country of residence for low-wage oDesk workers comes from one of the authors' calculations using oDesk administrative data.

¹⁰We chose a \$3 wage cutoff to minimize the cost of the experiment, while ensuring a sufficient sample size and a sample that was representative of the low-wage segment on oDesk. We initially contacted workers with wages of up to \$5 as many workers are willing to work for wages below those listed on their resumes (Pallais, 2014). For logistical reasons, we needed to hire workers at the same time. Because oDesk workers tend to remove their job applications if they do not hear back quickly, we gave workers 48 hours to apply. Prior experience suggested that 48 hours would maximize the size of the applicant pool.

¹¹Appendix Table 1 describes the characteristics of workers whom we asked to refer. It shows that workers who referred someone look somewhat more qualified than those who did not.

sample any referral who ever worked in the same room as her referrer.

We 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.¹² We again gave workers 48 hours to apply. Referred workers were much more likely to apply to our job: 68 percent of referred workers applied versus only six percent non-referred workers. We then hired all referred and non-referred workers who applied at an hourly wage of \$3 or less.¹³ We did not tell original hires or their referrals anything about how they would be treated before the referral was made and the referred worker applied for the job. For example, original hires and referrals in both treatments of the peer influence experiment had the exact same information up until the time the referral was hired.

This recruiting process, used for both the peer influence 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). Figure 2 depicts this recruitment process. 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 any performance results. In the selection experiment, we made job offers to all referred and non-referred workers from the peer influence experiment; no referrers were included. Figure 3 shows the recruitment of the selection experiment sample.

2.3 Peer Influence Experiment Design

The peer influence experiment was designed primarily to determine whether referrals work harder as a result of being referred because they think their performance and persistence will affect either their referrer's position at the firm or their relationship with their referrer. It also allows us to analyze whether referrals contain information about worker quality.

Panel A of Appendix Table 2 describes the characteristics of the referred and non-referred

¹²We eliminated from the pool of both referred and non-referred workers any workers who had already been invited as a potential referrer. We also eliminated from the team experiment anyone who had been invited in the peer influence experiment. As a result, referred and non-referred workers in the team experiment look worse on observables than do referred and non-referred workers in the peer influence experiment.

¹³We 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 referrers to choose referrals with high worker-firm match quality.

workers in the peer influence experiment. Referred workers had, on average, been on oDesk for about 18 months and almost three-quarters had prior oDesk employment. They averaged over 9 previous jobs and \$1,382 in prior oDesk earnings. Non-referred workers had been on oDesk slightly (insignificantly) longer, but were less than half as likely to have previously been hired. Referred workers also had higher feedback scores from prior employers and were more likely to have passed oDesk tests. Despite being seemingly more experienced than non-referred workers, referred workers posted wages on their resumes that were 15 percent *lower* than those posted by non-referred workers, and they proposed significantly lower wages to our jobs. Recall that referred workers were also much more likely to apply to our job. This suggests referrals may reduce recruiting costs by providing a way to identify workers with good resumes who are interested in the job.

We designed our task in this experiment 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 (Holzer, 1999; Regenstein et al., 1999; and Pallais, 2014). 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; Burks et al., 2015; Dustmann et al., 2011).

All referred and non-referred workers in the experiment completed the same task. We told them they would be doing testing for an airline flight website, and asked that they visit the site every other 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. Appendix Figure 1 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 11 am 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 11 am 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 2 shows an example performance report.

Referrers were randomized into the monitoring and non-monitoring treatments. Each referred worker was assigned the same treatment as her referrer. Appendix Table 3 shows that the randomization produced balanced samples between the treatment groups within

both the referrer and referral samples. Out of 26 comparisons between the two treatments groups, only one difference is significant at the 10 percent level.¹⁴

The monitoring treatment was designed to facilitate monitoring of the referred worker by her referrer while the non-monitoring treatment was designed to minimize peer influence. Referred workers in the monitoring treatment were told that their daily performance statistics would be sent to their referrer as well as the manager. Referred workers in the non-monitoring treatment meantime, were explicitly told that their referrer would never see their performance statistics, only the manager would. The difference in performance and persistence between referred workers in these treatments is due to peer influence. The difference in performance between referred workers in the non-monitoring treatment and non-referred workers sheds light on whether referrals contain information about worker quality. However, even referred workers in the non-monitoring treatment may have worked harder because they felt grateful for having been referred or faced informal pressures from their referrers.

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 sent them about their referrals. Thus, we asked them to answer questions on a website every other day over the same twelve-day period and we assigned them a soft deadline of 2 pm Philippine Time for submitting. We did not, however, want the referrers to garner insights from their own task with which they could potentially help their referrals, so we had them work on a site that had a different login method, was focused on consumer products rather than flights, and asked a different set of questions.

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 the monitoring treatment that whether they were promoted would depend on their referrals' performance.¹⁵ Referrers in the non-monitoring treatment were also informed of the management position, 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 the monitoring treatment to her referrer. We also sent the 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

¹⁴While there are 28 comparisons in the table, by construction, there is no variation in prior experience or in having a feedback score among referrers.

¹⁵All 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 the monitoring treatment about the position, we also said that they would receive daily performance updates on their referrals "because we care about workers' performance." To make sure we were as truthful as possible, we hired some of these workers for management positions after the experiment.

continue on the same project. We use this as an (inverse) measure of worker turnover. Each referred and non-referred worker was told that the manager would receive an update on whether she accepted our offer to re-apply. Referred workers in the monitoring treatment were told this update would also go to their referrers while referred workers in the non-monitoring treatment were explicitly told their referrers would not see this information. To strengthen the treatment, when we invited referrers in the monitoring treatment 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 the non-monitoring treatment to apply for the management position as well, but made no mention at all of their referrals. This experimental design is summarized in Panel A of Figure 4.

2.4 Selection Experiment Design

The selection experiment was designed explicitly to determine whether referrals contain information about worker quality. Four months after the peer influence 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. We created a firm with a different name, location, job posting, and writing style from that of the peer influence experiment. 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 peer influence experiment and sent three reminders to accept to workers who had not yet responded. None of the referrers was contacted by this firm. Panel B of Appendix Table 2 describes the characteristics of the referred and non-referred workers who accepted our offer. (These characteristics were measured at the time we first contacted them for the peer influence experiment.)

Similar to the peer influence experiment, workers who accepted the job offers were given a task that measured individual diligence over time. Workers 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 prior knowledge of Twitter and explained where to find the relevant information. Most of each day's task involved reporting on the Twitter activity of the artist from the day before. Although we asked workers to complete the task on the correct day, we also accepted retroactive submissions and automatically recorded the time of submissions. Appendix Figure 3 displays the questionnaire. After the last assigned day of work, we again invited workers to a continuation of the task and recorded whether they re-applied.

2.5 Team Experiment Design

The team experiment was designed to determine whether directly working with their referrers leads referrals to perform better (team production).¹⁶ The task involved brainstorming and we encouraged teamwork. Each worker was paired with three successive partners and asked to come up with a slogan for each of three different public service announcements. We chose this task because there are many jobs on oDesk that ask low-skill workers to come up with advertisements, including jobs that specifically ask workers to create slogans. 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 her partner and to come up with a single slogan that both partners would submit through our online form. Appendix Figure 4 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 different 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 drivers. The stated justification was that there was a lot of information to process and that by giving the partners different information, each partner 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 the 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 told them only that the sheets would contain information, not that they would contain

¹⁶Panel C of Appendix Table 2 shows the characteristics of referred and non-referred workers in the team experiment. As in the peer influence and selection experiments, referred workers were more likely than non-referred workers to have previously been hired and had higher feedback scores from prior employers, but proposed significantly lower wages to our jobs.

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.¹⁷

For comparison with the peer influence and selection experiments, we also collected measures of individual diligence. We monitored whether each worker logged in to the site and whether she submitted work. We also asked each worker an "individual question," the answer to which was in her own information sheet (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 criterion from their own information sheet in their slogan regardless of whom they were partnered with.

In the experiment, each referrer completed the 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. Panel B of Figure 4 gives an example of these three team types. 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.

Comparing the performance of referred workers in Type A and B teams provides the value of team production: how much better a referred worker performs when working with her own referrer than with someone else's referrer. Comparing the performance of workers in Type B and C teams shows the difference between referred and non-referred workers when both work with partners they don't know.

Because we thought worker performance might be correlated not just between partners, but also among partners' partners, we placed workers into blocking groups. Each of the 47 blocking groups contained six referrers, their six referred workers, and two non-referred workers. By definition, every worker in the blocking group only partnered with others in the same blocking group. In all analyses of the team experiment, we cluster standard errors by

¹⁷Because we wanted to measure how effectively workers worked with their partners, we strongly encouraged each worker to complete each PSA. Unlike in the peer influence 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.

blocking group.¹⁸ The placement into blocking groups was random, except that a referrer and her referral were always in the same group.¹⁹ Within a blocking group, the 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.

3 Referrals and Information about Worker Quality

We now examine whether referrals provide information about worker quality. First, we compare the performance and turnover of referred and non-referred workers in the selection experiment. Then, we compare non-monitored referred workers with non-referred workers in the peer influence experiment.

3.1 Selection Experiment

The selection experiment shows that referrals do contain information about worker quality: 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.

Table 1 compares the outcomes of the referred and non-referred workers in the selection experiment. First, we consider workers' likelihood of accepting a job. Panel A includes no controls. Consistent with the idea that hiring referred workers decreases recruiting costs, even among workers contacted for the selection experiment – who had previously participated in an experiment – referred workers were more likely to accept our job offer. While 51 percent of non-referred workers accepted, 68 percent of referred workers did. To determine how much of the information contained in the referral would have been observable to employers through workers' resumes, Panels B and C of Table 1 add control variables to the regressions in Panel

¹⁸We do find evidence of learning from partners, supporting our decision to cluster by blocking group. We show in Appendix Table 4 that a team performed better when one of its members had previously been in a Type A team, controlling for the current team type and the task number. Since the task order was random, this may suggest that when workers are in successful pairings, they learn how to do the task successfully and use that knowledge in subsequent tasks.

¹⁹As in the peer influence experiment, we hired all referred and non-referred workers who met the selection criteria. However, only one randomly-selected referral from each referrer and only 94 non-referred workers were included in this experiment.

A. Panel B adds our main covariates: what we call first-order controls.²⁰ Panel C adds the squares of each of the (non-binary) covariates and the interaction of each pair of covariates (our second-order controls) to the regressions. The table shows that the 17 percentage point difference in job acceptance is almost entirely explained by observable characteristics (in particular, prior oDesk experience and prior earnings in the marketplace), leaving only an (insignificant) 4.6 percentage point difference in acceptance rates once we add the first- and second-order controls.

Next, we consider the performance and persistence of workers who accepted the job offer. Measures of performance and persistence are regressed on a dummy for being a non-referred worker (the base group is referred workers). We consider three measures of performance: (1) an indicator for submitting the day’s work, (2) an indicator for submitting it on time, and (3) the fraction of questions answered correctly (accuracy). Unanswered questions are marked as incorrect. We also consider whether workers applied for a continuation of the task as a measure of persistence.

The table shows that referred workers submitted work on 76 percent of days and the vast majority of these submissions were made on time. However, non-referred workers were 11 percentage points less likely both to submit work and to submit the work on time. While 82 percent of referred workers re-applied for a continuation of the task, non-referred workers were 20 percentage points less likely to do so. However, despite the fact that these coefficients are large and significant, the non-referred dummy explains only a small share of variation in the outcome measures: just over 1 percent in the case of submission and on-time submission and approximately 5 percent in the case of persistence.

Panel A of Figure 5 shows performance over the course of the experiment by worker type. Submission rates of referred workers were consistently higher than those of non-referred workers. Both types of workers became less diligent over time, but diligence fell off much more for non-referred workers. Thus, the performance gap between referred and non-referred workers grew over the course of the job. Panel A of Appendix Figure 5 shows that the other performance measures (on-time submission and accuracy) follow similar trends.

Workers’ resume characteristics are predictive of their performance and persistence: the proportion of variation explained increases to a quarter (for submission) and a third (for re-application) when the first- and second-order controls are added. However, adding covariates

²⁰These are 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 passed oDesk tests, an indicator for having a portfolio, 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.

does not change the coefficient on the referral dummy at all. This suggests that while the referral mostly contained observable information about workers’ willingness to accept the job, most of the information contained in the referral about workers’ performance and persistence was not otherwise observable through the workers’ resumes. Panel A of Appendix Table 5 displays the coefficients on the first-order controls from Panel B of Table 1. (Coefficients on the second-order controls are harder to interpret.) Unsurprisingly, the coefficients suggest that prior oDesk experience, more degrees, and passing oDesk tests – variables on which referred workers look better than non-referred workers – are positively related to performance and persistence (though these coefficients are typically not significant). However, all else equal, the two characteristics that explain the most variation in performance are (1) having been on oDesk longer and (2) not being in an agency. Referred workers look worse on both these metrics.

3.2 Peer Influence Experiment

Next, we compare the performance and turnover of non-monitored referred workers and non-referred workers in the peer influence experiment. The results are very similar to those of the selection experiment. The main difference is that in the peer influence experiment, we also compare the performance of monitored and non-monitored referred workers. We discuss this comparison in Section 4.

Each column of Table 2 presents the results of regressing an outcome on an indicator for being a monitored referred worker and an indicator for being a non-referred worker. (The omitted group is non-monitored referred workers.) We use the same performance and persistence metrics as in the selection experiment – submission, on-time submission, accuracy, and re-application.²¹ Panel A includes no controls, Panel B includes first-order controls, and Panel C includes first- and second-order controls.

Referred workers performed better than non-referred workers. Non-monitored referred workers were 13 percentage points more likely to submit, 8 percentage points more likely to submit on time, and 23 percentage points more likely to re-apply for the job than were non-referred workers. Panel B of Figure 5 shows that, as in the selection experiment, the submission gap between referred and non-referred workers grew with time.²²

Observable characteristics from workers’ resumes explain a lot of the variation in out-

²¹Two of the three performance metrics we consider 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 11 am. Workers were also told that the manager would see whether they answered all questions, but we exclude this metric from our analysis since 99.8 percent of submissions were complete.

²²Panel B of Appendix Figure 5 shows that the other performance measures (on-time submission and accuracy) follow similar trends.

comes, but they do not diminish the predictive power of the referral. The proportion of variance in performance explained increases from approximately 2 percent to 15 to 20 percent when all the controls are added, but the coefficient on the non-referred dummy remains constant. Panel B of Appendix Table 3 shows that the coefficients on the first-order controls are similar to the coefficients on these controls in the selection experiment regressions.

These results suggest that referrals contain information about worker performance that is not present in workers' resumes. In addition to using workers' resumes, firms could gain information about worker quality through interviews or a job test, both of which are costly.²³ While we do not know what information firms would gain through interviews, we can approximate the information that might be gained from a job test using workers' initial job performance. Panel D shows that the referral still has predictive power for worker performance on the last day of the contract, conditional on worker performance on all prior days. Panel D replicates Panel C, limiting the observations to the last day of the contract. Regressions in the first three columns now additionally control for the worker's performance (on the same metric as measured by the dependent variable) on each of the first five days. All differences in performance between referred and non-referred workers remain large and significant.

The referral also provides information about worker persistence at the firm above and beyond the information provided by the worker's performance throughout the full contract. The final column of Panel D adds controls for each of our performance measures (submission, on-time submission, and accuracy) on each of the six days. Even controlling for all our performance measures on all days, referred workers were 15 percentage points more likely than non-referred workers to want to continue with the firm.²⁴

The results suggest that referrals provide important information about worker quality. Even when referred workers were not monitored by their referrers, they performed much better than non-referred workers and were more eager to continue with the firm. This information was not present on workers' resumes or in their performance on the majority of the contract.

3.3 Heterogeneity by Referral Type

The above analysis suggests that referrals contain information. Here, we find that some referrals contain more information than others. In particular, referrals made by high-performing

²³Even when firms undertake interviews, firms have considerable uncertainty about worker productivity when hiring (e.g., Autor and Scarborough, 2008).

²⁴Unreported coefficients in the final column of Panel D show that workers who performed better were more likely to want to continue with the firm.

referrers and referrals of workers with strong ties to their referrers are particularly informative.

Using data from the peer influence experiment, the first column of Table 3 shows that a referrer's performance is a strong predictor of her referral's performance. This is not just a result of the referrer and her referral facing common shocks. The referrer's performance in the peer influence experiment is a strong predictor of her referral's performance in the selection experiment, four months later (Column 2, Table 3).

Some of this can be accounted for by observable characteristics. Appendix Table 6 shows that workers with better observable characteristics refer workers who also have better observables. Controlling for the referred worker's observable characteristics in the regression of referral performance on referrer performance reduces the point estimate on referrer performance. Nonetheless, the referrer's performance remains an important predictor of her referral's performance. This suggests that higher performers refer workers who perform better than would even be expected based on their observable characteristics. It also suggests that not all referred workers are predicted to outperform non-referred workers. In both the selection and peer influence experiments, these results suggest that referrals from the worst-performing 20 percent of referrers are predicted to under-perform non-referred workers and, in fact, they do.

We turn now to the relationship between referrers and their referrals. Appendix Table 7 shows the distributions of the three relationship variables we have from referrers' reports at the time of the referral. Referrers tended to refer workers they were close to. 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. Just under half of referred workers knew 20 or more people in common with their referrers.

Because the relationship variables are positively correlated and predict performance in the same way, we build an index of relationship strength and for parsimony focus here on the resulting estimates.²⁵ In the final columns of Table 3, we regress referral performance in the different experiments on this index. In each experiment, the coefficients suggest that referrals with stronger ties to their referrers performed better. These coefficients actually increase

²⁵In 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 day, and knowing at least thirty people in common. Our relationship index is defined as the standardized sum of these three binary variables. We exclude the five referred workers whose referrers did not answer all the relationship questions at the time of the referral.

slightly when we add controls for worker characteristics (Panel B). That is because referrals with stronger ties to their referrers look worse on paper: they have lower earnings, have been on oDesk for less time, and have fewer educational degrees. Conditional on observable characteristics, a referred worker with a one standard deviation stronger relationship with her referrer was approximately four percentage points more likely to submit work in the peer influence experiment, five percentage points more likely to submit the same slogan as her partner in the team experiment, and (an insignificant) three percentage points more likely to submit work in the selection experiment.

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

3.4 Potential Bias from Employers' Hiring Decisions

To test whether referrals provide information about the expected performance of job applicants, we hired all applicants who met our basic hiring criteria. Here, 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.

Using data from the peer influence experiment, we first simulate which workers employers would hire if they observed only the characteristics on workers' resumes; we then simulate whom employers would hire if they additionally observed which workers had been referred. For comparison, we also show the characteristics of workers hired if employers only observed workers' referral status and no other characteristics. We assume that employers want to maximize the fraction of workers who submit a response on a given day and that they know the relationship between demographics and referral status, and performance.²⁶ Employers predict each applicant's performance using the information they observe and then hire the half of the applicant pool with the best predicted performance.

Table 4 shows the results of the simulations. Results in the first row simulate hiring under the assumption that employers only see workers' resumes, not who was referred. The second row simulates hiring under the assumption that employers only see workers' referral status, so they hire a random sample of referred workers. Finally, the third row simulates hiring under the assumption that employers observe workers' resume characteristics and who was referred.

²⁶In 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.

If employers observed only workers' resume characteristics, a higher fraction of referred (58 percent) than non-referred (39 percent) workers would be hired (Panel A). However, if employers also observed who was referred, the fraction of referred applicants that would be hired jumps to 79 percent; meantime, only nine percent of non-referred applicants would be hired. If employers observed only referral status, they would hire 85% of referred workers and no non-referred workers.

Panel B displays the summary measure of the hired workers' observable characteristics. It shows that when employers observe workers' resumes as well as who was referred, hired non-referred workers are positively selected on observables relative to hired referred workers.

Panel C shows the actual submission rates of the workers hired in each scenario. Compared to hiring at random, both (1) hiring using only observable characteristics and (2) hiring using only referral status substantially improve the performance of hired workers. (Hiring using these strategies relative to hiring at random improves the performance of hired workers by seven and six percentage points, respectively.) Observing both referral status and observable characteristics brings slightly larger gains in performance than using either in isolation. These results suggest that referrals might provide a way for firms to reduce recruiting costs. Given that much of the gain from using workers' characteristics in hiring could be obtained from using referral status alone, if collecting information on workers' characteristics is costly, employers might choose to forgo collecting these characteristics in favor of using referrals.

The table also shows that if employers did not observe who was referred, hired referred workers would be substantially (13 percentage points) more likely to actually submit work than non-referred workers. However, this difference would be only three percentage points (and statistically indistinguishable from zero) if employers also observed who was referred. This suggests that if we had only observed the performance of hired workers and did not observe all the characteristics employers used in their hiring decisions, we might have mistakenly concluded that referrals contained little to no information about worker performance.

4 Effect of On-the-Job Interactions with Referrers

We now consider whether being referred actually makes referred workers more productive. First, we consider whether referrals work harder because they believe their performance will affect their relationship with their referrer or their referrer's position at the firm (peer influence). Second, we consider whether referrals perform better when working directly with their referrers (team production).

4.1 Peer Influence

The peer influence experiment shows that peer monitoring does not have a detectable effect on performance.

Anecdotal evidence suggests that referred workers in the monitoring treatment were, in fact, monitored by their referrers. Many referrers in this treatment 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 or had not completed it by the soft deadline, and assured us they would encourage their referrals to do better on subsequent days. Yet, Table 2 shows that while the coefficients indicate that monitored referred workers performed better than non-monitored referred workers, these differences are much smaller than the differences between non-monitored referred workers and non-referred workers and are never statistically significant.²⁷ The negative (though again insignificant) coefficient on the monitored referred worker dummy in the final column suggests that monitored referred workers were, if anything, slightly less likely to be interested in continuing with the firm, perhaps because they disliked being monitored.

Panel B of Figure 5 sheds some light on how the performance of monitored and non-monitored referred workers evolved over time. On the first day of work, before any performance reports had been sent, 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 monitored referred workers, though the differences between monitored and non-monitored referred workers on those days is not significant. By day six, however, monitored referred workers were no more likely than their non-monitored counterparts to submit work.

Overall, 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.

4.2 Team Production

The team experiment shows that referred workers perform better when working directly with their referrers. In particular, referred workers performed much better when working with their own referrer than with a randomly-selected referrer they did not know.

We first consider the effect of team type on measures that do not rely on teamwork, but may be indicative of individual diligence. These are indicators for logging in to our site

²⁷Using seemingly unrelated regression, we calculate the variance-covariance matrix between the coefficients in these three performance regressions and test the hypothesis that all three monitored referred coefficients are equal to zero. We are unable to reject this hypothesis.

to see the given PSA task, submitting work, correctly answering the question about their own individual reading, and including the criteria from their own information sheets in their slogans.²⁸

In Panel A of Table 5, each measure of individual diligence 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 coefficients on the Type A dummy indicate how much better referred workers perform when paired with their own referrer than with someone else’s referrer; the coefficients on the Type C dummy indicate 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. First- and second-order controls for both partners’ observable characteristics are included throughout.

On average, referred workers performed well on these diligence measures. Similar to our previous results, non-referred workers were less diligent than referred workers, even when neither group was working with a partner they previously knew.

Referred workers were five percentage points more likely to submit work and to correctly answer the question about their own reading when they were paired with their own referrer than when paired with someone else’s referrer. Given that these are measures of diligence more than teamwork, this could suggest that referred workers exerted more effort when working with their referrer. This may be because, in this case, their performance affected their referrers’ output. Alternatively, it could result from peer influence if working together made it easier for referrers to monitor their referrals.

Panel B compares team performance by team type. Observations are again at the partner-pair level. Referred workers 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 than when paired with referrers they did not know; of the Type A teams that both submitted responses, only six 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, while Type A teams were more than twice as likely to do so.²⁹ Appendix A shows that in addi-

²⁸If a worker did not answer the question about her reading, she is marked as not answering it correctly. Similarly, if she did not submit a slogan, she is marked as not including her own criterion in the slogan.

²⁹One hypothesis is that firms could replicate the benefit of team production that comes from referrals by creating teams of workers with similar observable characteristics. However, we do not find evidence that

tion to performing better, Type A teams enjoyed their task more, spent more time on the task, and communicated more. They performed better even conditional on time spent and communication.

5 External Validity

Completing these experiments in an online labor market provides two major benefits. First, it allows us to observe the performance and persistence of workers without the filter of firms' hiring decisions. Second, it allows us to vary parameters of the jobs to cleanly identify why referred workers perform better and have less turnover than non-referred workers. As with any field experiment we might run, however, the results of this experiment come from one particular labor market, in this case an online labor market.

The types of tasks in our experiments are not uncommon in offline labor markets. Autor, Levy, and Murnane (2003) classifies tasks into five categories, now prevalent in the skills literature: expert thinking, complex communication, routine cognitive tasks, routine manual tasks, and non-routine manual tasks. Our selection and peer influence experiments center on *routine cognitive* tasks like basic computations and data entry.

Routine cognitive tasks are prevalent in offline labor markets, especially among workers with a high school degree or some college. Autor, Levy, and Murnane (2003) defines a composite measure of routine cognitive tasks which they map to Census occupations using O*Net data. They find that occupations in office and administrative support are particularly heavy in routine cognitive tasks; examples include cashiers, customer service representatives, and tellers.

We think the principal difference between oDesk and offline labor markets is the incentives workers face. Because oDesk jobs are typically shorter than offline jobs, oDesk workers are often less tied to any particular employer than are workers in other labor markets. Prior to our experiment, the average job taken by the referrers in our sample paid \$237 and lasted 81 working hours. If oDesk workers are less concerned about their reputations with their employers than are most workers in offline labor markets, this could lead referrals to contain less information about worker quality and on-the-job interactions with referrers to be less effective in improving worker performance.

teams in which partners had similar characteristics perform better. 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. 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.

Referrers in our experiments were not provided compensation to provide referrals or to provide high-quality referrals. They may have received a social benefit or felt a warm glow from helping a friend find employment (e.g., Beaman et al., 2013). But, their incentive to make high-quality referrals was implicit: by making high-quality referrals they could improve their relationship with our firm. We did try to provide some incentives for workers to care about their relationship with our firm by implying that if they performed well, they could have a long-term relationship with us. Nonetheless, this working relationship was still far less long-sighted than working relationships in most labor markets. The fact that referrals still contained positive information about worker quality despite referrers' relatively weak incentives to refer high-quality workers suggests that referrals are likely to contain positive (and perhaps even more positive) information about worker quality in other labor markets.

Relatedly, if referrers are less concerned about their reputations with their employers on oDesk, they may exert less pressure on their referrals to perform well, weakening the effect of peer influence. Since we were aware that peer influence might not be as strong a motivator on oDesk as in other labor markets, we aimed explicitly to maximize the effect of peer influence in the monitoring treatment of the peer influence experiment. That we find very limited effects of peer influence, then, suggests peer influence is likely not an important mechanism in this context. Nonetheless, peer influence may still be important in other labor markets, especially in labor markets in which referrers care more about ongoing relationships with firms.

6 Conclusion

The use of social connections is ubiquitous in the labor market. More than half of jobs are found through informal connections and firms are more likely to hire referred than non-referred applicants, all else equal. This suggests that workers without social connections may be disadvantaged in the labor market (e.g., Calvo-Armengol and Jackson, 2004; Montgomery, 1991). This paper examines why firms prefer to hire referred workers: do referrals allow firms to hire more productive workers because they signal worker quality or because being referred actually makes workers more productive?

Understanding why firms prefer to hire referred workers can inform potential policy responses that may help unconnected workers. For example, if referrals provide information about worker quality, then providing unconnected workers with other ways to signal their abilities may improve their labor market outcomes (as in Pallais, 2014). On the other hand, if team production actually causes referrals to be more productive, information approaches may not help unconnected workers. Nepotism may also be harder to eliminate.

We find strong evidence that on-the-job interactions between referred workers and their referrers lead referrals to perform better. While we do not find evidence of peer influence, our results suggest that team production is an important benefit of referrals. However, we also find strong evidence that referrals contain information about worker performance and turnover. In our context, referrals contain information about general productivity. In other contexts, referrals might also signal that a worker is a particularly good match for a given firm or job. While this explanation is precluded in our experiments because referrers did not have information about the job they were referring for, it could be important in other settings.

From our experiments, we learn that referrals made by high-performers and referrals of workers with strong ties to their referrers were particularly informative. Yet, we do not know why: that is, whether referrers actively choose referrals they know will perform well (as in Beaman and Magruder, 2012) or whether these results obtain simply because productive workers have productive friends (as in Montgomery, 1991). Understanding why referrals contain so much information about worker quality is an important question for future research.

References

- AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): “The skill content of recent technological change: An empirical exploration,” *The Quarterly Journal of Economics*, pp. 1279–1333.
- AUTOR, D. H., AND D. SCARBOROUGH (2008): “Does job testing harm minority workers? Evidence from retail establishments,” *The Quarterly Journal of Economics*, pp. 219–277.
- BANDIERA, O., I. BARANKAY, AND I. RASUL (2005): “Social preferences and the response to incentives: Evidence from personnel data,” *Quarterly Journal of Economics*, 120(3), 917–962.
- (2012): “Team incentives: Evidence from a firm level experiment,” *CEPR Discussion Paper DP8776*.
- BEAMAN, L., N. KELEHER, AND J. MAGRUDER (2013): “Do job networks disadvantage women?: Evidence from a recruitment experiment in Malawi,” *Working Paper*.
- BEAMAN, L., AND J. MAGRUDER (2012): “Who gets the job referral? Evidence from a social networks experiment,” *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,” *FRB of New York Staff Report 568*.
- BRYAN, G. T., D. KARLAN, AND J. ZINMAN (2014): “Referrals: Peer screening and enforcement in a consumer credit field experiment,” *American Economic Journal: Microeconomics*, forthcoming.
- BURKS, S., B. COWGILL, M. HOFFMAN, AND M. HOUSMAN (2015): “The value of hiring through referrals,” *Quarterly Journal of Economics*, 130(2), 805–839.

- CALVO-ARMENGOL, A., AND M. O. JACKSON (2004): “The effects of social networks on employment and inequality,” *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,” *Quarterly Journal of Economics*, 118(2), 519–548.
- DHILLON, A., V. IVERSEN, AND G. TORSVIK (2012): “Employee referral, social proximity, and worker discipline,” *Working Paper*.
- DUSTMANN, C., A. GLITZ, AND U. SCHÖNBERG (2011): “Referral-based job search networks,” *IZA Discussion Paper 5777*.
- FERNANDEZ, R. M., AND N. WEINBERG (1997): “Sifting and sorting: Personal contacts and hiring in a retail bank,” *American Sociological Review*, 62(6), 883–902.
- GHANI, E., W. R. KERR, AND C. STANTON (2014): “Diasporas and outsourcing: Evidence from oDesk and India,” *Management Science*, 60(7), 1677–1697.
- GILCHRIST, D., M. LUCA, AND D. MALHOTRA (2015): “When $3+1>4$: Gift structure and reciprocity in the field,” *Management Science*, *forthcoming*.
- GRANOVETTER, M. (1995): *Getting a Job: A Study of Contacts and Careers*. University of Chicago Press, Chicago.
- HEATH, R. (2015): “Why do firms hire using referrals?: Evidence from Bangladeshi garment factories,” *Working Paper*.
- HOLZER, H. J. (1987): “Hiring procedures in the firm: Their economic determinants and outcomes,” *NBER Working Paper 2185*.
- HOLZER, H. J. (1999): “Will employers hire welfare recipients? Recent survey evidence from Michigan,” *Journal of Policy Analysis and Management*, 18(3), 449–472.
- HORTON, J. J. (2013a): “Computer-mediated matchmaking: Facilitating employer search and screening,” *Working Paper*.
- HORTON, J. J. (2013b): “The effects of subsidizing employer search,” *Working Paper*.
- (2014): “Price floors and employer preferences: Evidence from a minimum wage experiment,” *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.
- KRAMARZ, F., AND O. N. SKANS (2014): “When strong ties are strong: Networks and youth labour market entry,” *Review of Economic Studies*, 81(3), 1164–1200.
- KUGLER, A. D. (2003): “Employee referrals and efficiency wages,” *Labour Economics*, 10(5), 531–556.
- LYONS, E. (2014): “Team production in international labor markets: Experimental evidence from the field,” *Working Paper*.
- MONTGOMERY, J. D. (1991): “Social networks and labor-market outcomes: Toward an economic analysis,” *American Economic Review*, 81(5), 1408–1418.
- ODESK CORPORATION (2013): *Slide Deck with Statistics on the oDesk Marketplace*.
- PALLAIS, A. (2014): “Inefficient hiring in entry-level labor markets,” *American Economic Review*, 104(11), 3565–3599.
- 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.
- REGENSTEIN, M., J. A. MEYER, AND J. D. HICKS (1999): “Job prospects for welfare recipients: Employers speak out,” *Urban Institute Series A, No. 25*.
- 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.
- STANTON, C., AND C. THOMAS (2014): “Landing the first job: The value of intermediaries in online hiring,” *CEP Discussion Paper 1316*.
- TOPA, G. (2011): “Labor markets and referrals,” *Handbook of Social Economics*, pp. 1193–1221.
- WANG, S.-Y. (2013): “Marriage networks, nepotism, and labor market outcomes in China,” *American Economic Journal: Applied Economics*, 5(3), 91–112.

Figure 1: oDesk Profile Example

Joanah M.
data entry and virtual assistant

\$3.33 / hr
★★★★★ (4.80)

Contact

Save as Favorite

	Jobs	Stars	Hours
All Time	13	4.80	680
Last 6 months	1	4.55	38

Location: **cavite, Philippines**
10:06 AM (UTC+08)

English Skills: **5 out of 5** (self-assessed)

Last Worked: **May 14, 2013**

Member Since: **August 2, 2008**

Work History and Feedback (13)

Web Research and Data Entry (May 2013)
38 hours @ \$3.33/hr
Earned \$126
★★★★★ Joanah was a great addition to the team, and it was a pleasure working with her. I would definitely hire her again.

data entry (Dec 2012)
4 hours @ \$3.33/hr
Earned \$18
★★★★★ It was a pleasure working with joanah

Website Data entry (Apr 2011 - Nov 2012)
29 hours @ \$2.22/hr
Earned \$64
No feedback given

Figure 2: The Recruitment Process

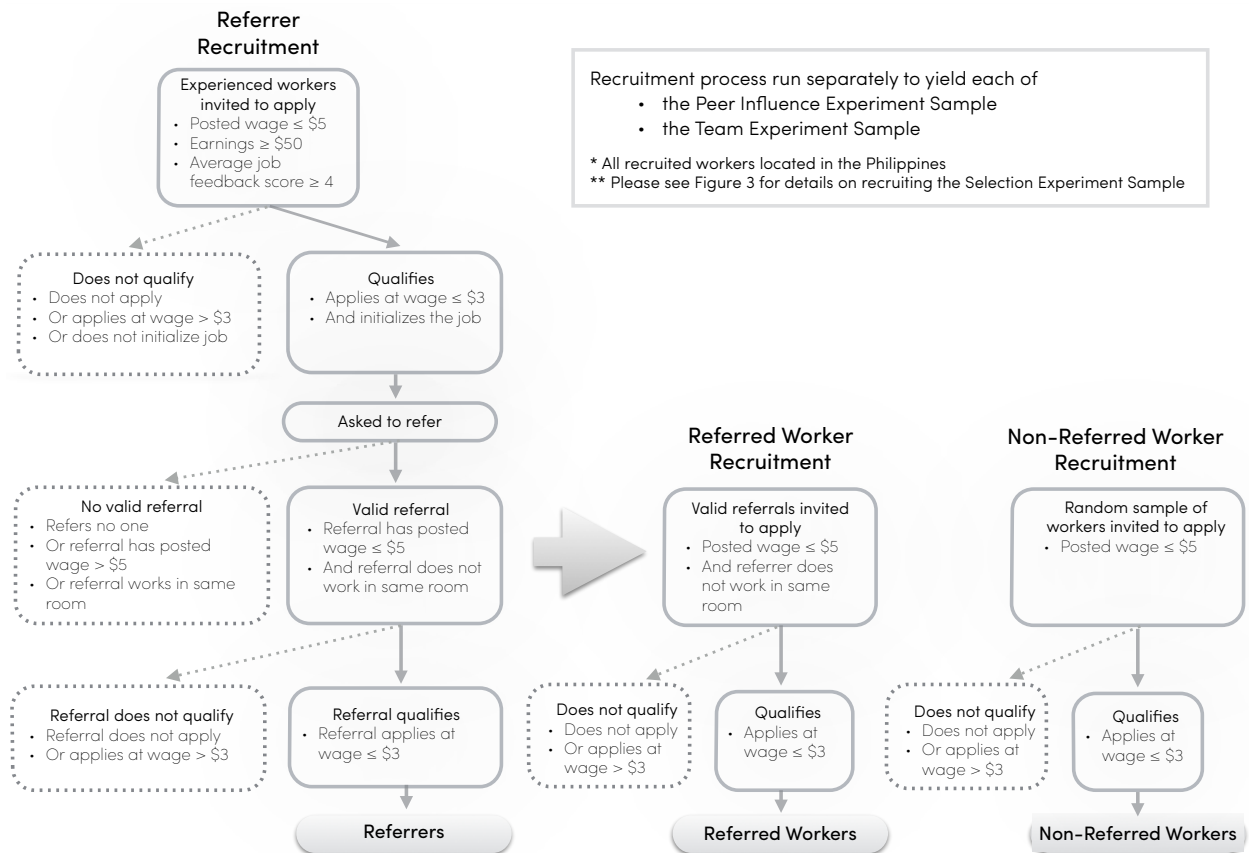


Figure 3: *The Experiment Samples*

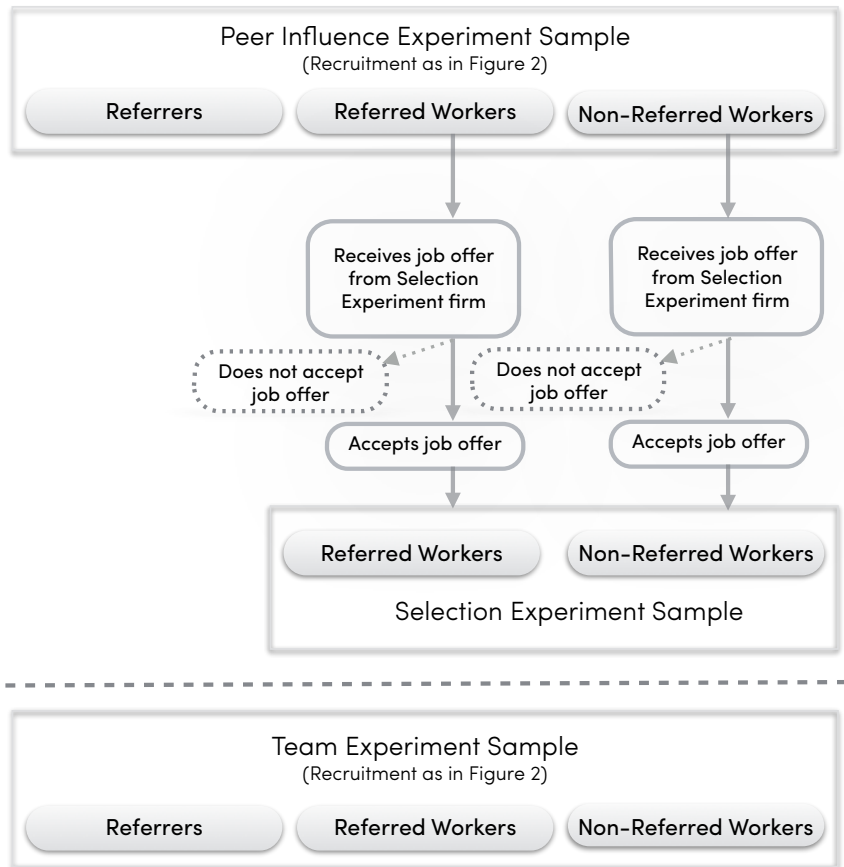


Figure 4: *Treatments in the Peer Influence and Team Experiments*

(Recruitment as in Figure 2)

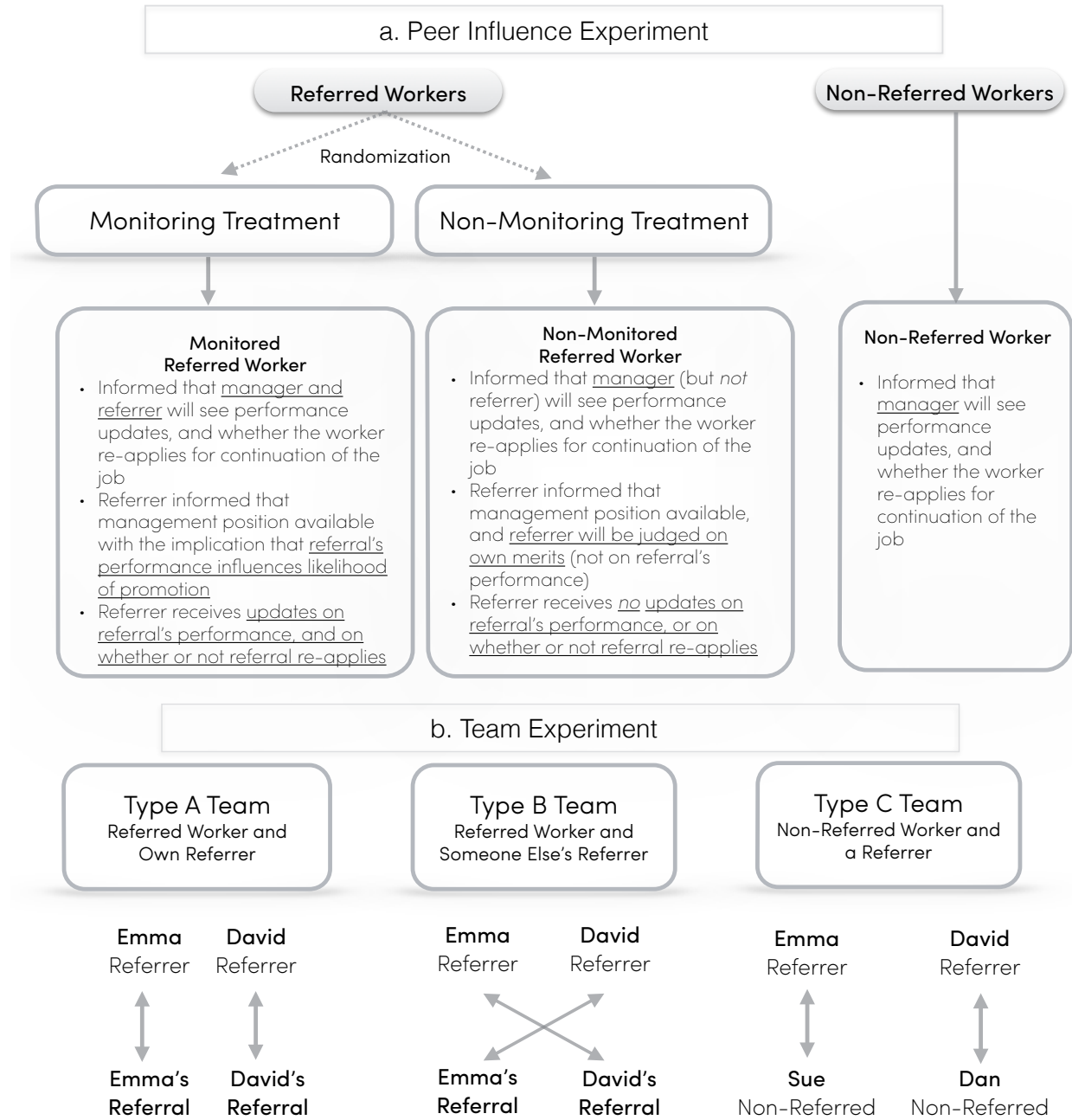
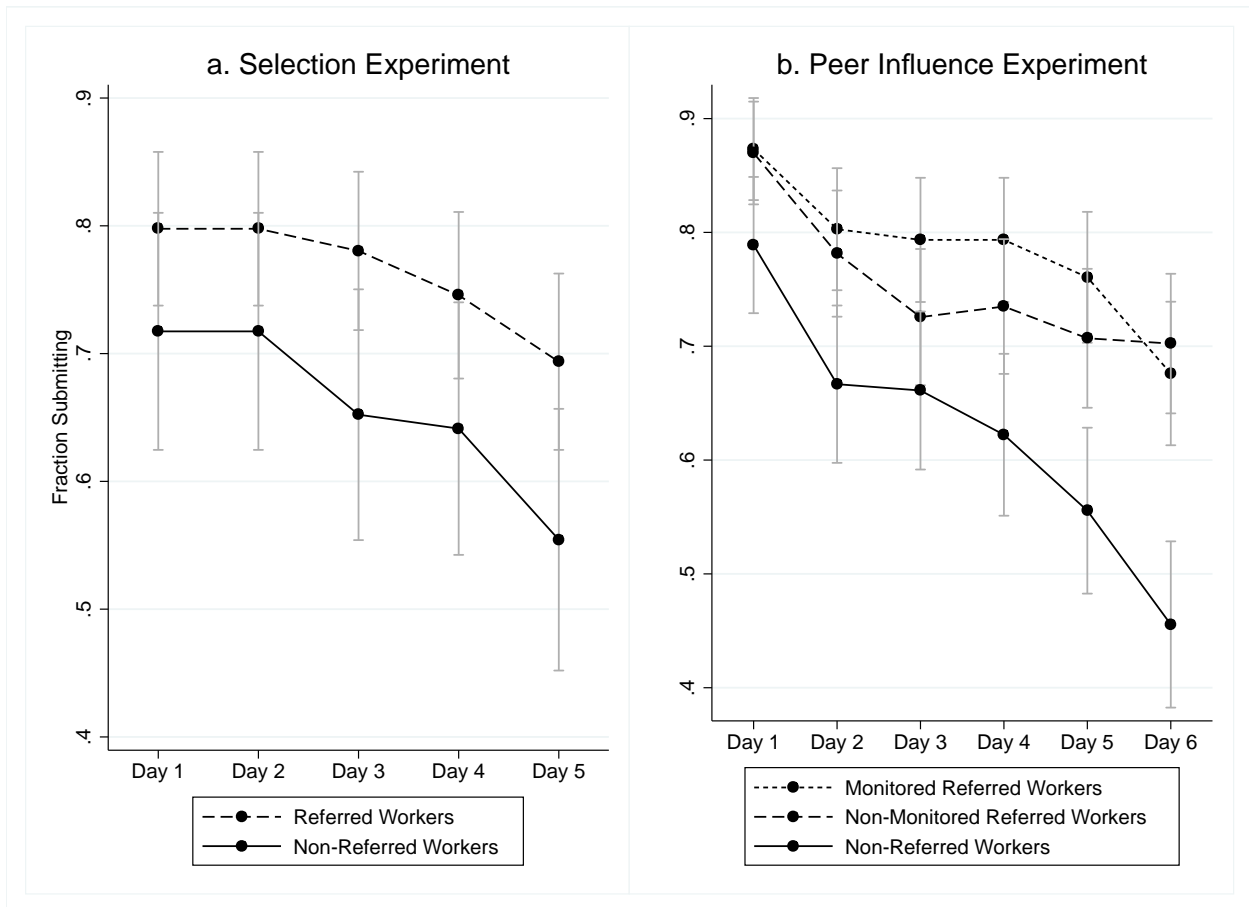


Figure 5: *Submission Rates by Day*



Notes: Error bars denote 95% confidence intervals.

Table 1. Performance and Persistence in the Selection Experiment
Base Group is All Referred Workers

	Sample: All Referred and Non-Referred Workers	Sample: Referred and Non-Referred Workers Who Accepted Job Offer			
	Accepted Job Offer	Submission	On-Time Submission	Accuracy	Re-Application
		<u>A. No Controls</u>			
Non-Referred	-0.167*** (0.047)	-0.106** (0.046)	-0.107** (0.048)	-0.035 (0.026)	-0.195*** (0.059)
Observations	435	1,325	1,325	1,325	265
R-squared	0.029	0.013	0.012	0.003	0.046
		<u>B. First-Order Controls</u>			
Non-Referred	-0.071 (0.056)	-0.100* (0.057)	-0.098* (0.059)	-0.024 (0.033)	-0.123* (0.071)
Observations	435	1,325	1,325	1,325	265
R-squared	0.125	0.079	0.077	0.048	0.088
		<u>C. First- and Second-Order Controls</u>			
Non-Referred	-0.046 (0.064)	-0.114* (0.064)	-0.108 (0.067)	-0.043 (0.036)	-0.172** (0.086)
Observations	435	1,325	1,325	1,325	265
R-squared	0.268	0.236	0.236	0.186	0.349
Base Group Mean	0.678	0.763	0.735	0.363	0.815

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 a non-referred worker. Panel A includes no controls; Panel B includes the first-order controls for worker characteristics listed in footnote 20; and Panel C also includes second-order controls (the square of each non-binary characteristic in footnote 20 and the interaction of each pair of characteristics in footnote 20). Observations in the first and final columns (*Accepted Job Offer*, *Re-Application*) are workers, while observations in the middle three columns of regressions (*Submission*, *On-Time Submission*, *Accuracy*) are worker-days. Regressions in the first column (*Accepted Job Offer*) include all workers contacted for the selection experiment; regressions in the remaining columns include only workers who accepted the job offer. Standard errors are clustered at the worker level when observations are worker-days, and Huber-White standard errors are presented when observations are workers. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 2. Performance and Persistence in the Peer Influence Experiment
Base Group is Non-Monitored Referred Workers

	Submission	On-Time Submission	Accuracy	Re-Application
<u>A. All Days, No Controls</u>				
Monitored Referred	0.036 (0.042)	0.053 (0.047)	0.034 (0.039)	-0.032 (0.030)
Non-Referred	-0.132*** (0.042)	-0.079* (0.045)	-0.101** (0.039)	-0.225*** (0.038)
Base Group Mean	0.757	0.563	0.640	0.953
Observations	2,610	2,610	2,610	435
R-squared	0.027	0.013	0.020	0.085
<u>B. All Days, First-Order Controls</u>				
Monitored Referred	0.020 (0.042)	0.038 (0.046)	0.014 (0.040)	-0.035 (0.034)
Non-Referred	-0.115** (0.047)	-0.080* (0.048)	-0.095** (0.043)	-0.193*** (0.044)
Base Group Mean	0.757	0.563	0.640	0.953
Observations	2,610	2,610	2,610	435
R-squared	0.075	0.061	0.063	0.156
<u>C. All Days, First- and Second-Order Controls</u>				
Monitored Referred	0.004 (0.041)	0.045 (0.047)	0.002 (0.039)	-0.044 (0.037)
Non-Referred	-0.135*** (0.049)	-0.067 (0.052)	-0.100** (0.046)	-0.196*** (0.052)
Base Group Mean	0.757	0.563	0.640	0.953
Observations	2,610	2,610	2,610	435
R-squared	0.181	0.139	0.163	0.264
<u>D. Last Day Only, First- and Second-Order Controls & Daily Performance Controls</u>				
Monitored Referred	-0.058 (0.048)	0.018 (0.057)	-0.046 (0.042)	-0.053 (0.038)
Non-Referred	-0.172*** (0.053)	-0.102* (0.056)	-0.103** (0.046)	-0.149*** (0.050)
Base Group Mean	0.703	0.500	0.600	0.953
Observations	435	435	435	435
R-squared	0.614	0.506	0.622	0.434

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 a referred worker in the monitoring treatment and an indicator for being a non-referred worker. As in Table 1, Panel A includes no controls; Panel B includes the first-order controls for worker characteristics listed in footnote 20; and Panel C includes first- and second-order controls. Regressions in Panels A, B, and C include observations on all six days of work. Regressions in Panel D are limited to observations on workers' last day of work and include first- and second-order controls, as well as daily performance controls; each of the first three columns includes controls for the worker's performance as measured by the dependent variable on each of the first five days of work. The final column includes controls for each of the three performance measures on each of the six days. Observations in the first three columns of regressions (*Submission*, *On-Time Submission*, *Accuracy*) are worker-days; observations in the final column (*Re-Application*) are workers. Standard errors are clustered at the worker level when observations are worker-days, and Huber-White standard errors are presented when observations are workers. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 3. Heterogeneity in Referral Performance
All Experiments, Dependent Variables Indicate Referred Workers' Performance

	<u>A. No Controls</u>				
	<u>Submission Rate (Peer Influence Experiment)</u>	<u>Submission Rate (Selection Experiment)</u>	<u>Submission (Selection Experiment)</u>	<u>Submission (Peer Influence Experiment)</u>	<u>Same Slogan (Team Experiment)</u>
Referrer's Submission Rate, Peer Influence Experiment	0.421*** (0.066)	0.342*** (0.084)			
Relationship Strength Index			0.015 (0.026)	0.030 (0.022)	0.044** (0.021)
Dependent Variable Mean	0.775	0.763	0.760	0.775	0.520
Observations	255	173	855	1,512	560
R-squared	0.192	0.115	0.001	0.006	0.007
	<u>B. First-Order Controls</u>				
	<u>Submission Rate (Peer Influence Experiment)</u>	<u>Submission Rate (Selection Experiment)</u>	<u>Submission (Selection Experiment)</u>	<u>Submission (Peer Influence Experiment)</u>	<u>Same Slogan (Team Experiment)</u>
Referrer's Submission Rate, Peer Influence Experiment	0.392*** (0.068)	0.194** (0.084)			
Relationship Strength Index			0.027 (0.023)	0.036* (0.021)	0.046** (0.020)
Dependent Variable Mean	0.775	0.763	0.760	0.775	0.520
Observations	255	173	855	1,512	560
R-squared	0.266	0.272	0.186	0.098	0.051

Notes: Each column in each panel reports the results of a separate regression where the dependent variable is indicated by the column. All dependent variables indicate referral performance. In the first two columns, observations are referred workers. In these columns, referred workers' average performance over the course of the indicated experiment is regressed on their referrer's average submission rate in the peer influence experiment. Huber-White standard errors are in parenthesis. In the third and fourth columns, observations are worker-days and standard errors are clustered by worker. In the final column, observations are at the worker-PSA level and standard errors are clustered by blocking group. In each of the three final columns, the dependent variable is regressed on an index for the strength of the referrer-referral relationship. This index is defined in Section 3.3 of the text and has mean zero and standard deviation one. Regressions in Panel A include no controls, while regressions in Panel B include the first-order controls for worker characteristics listed in footnote 20. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 4. Simulated Hiring
Peer Influence Experiment: Assuming Top 50% of Applicants Hired

	<u>A. Fraction Hired</u>		<u>B. Measure of Observables</u>		<u>C. Actual Submission Rate</u>			Difference
	Referred Applicants	Non-Referred Applicants	Hired Referred Workers	Hired Non-Referred Workers	Hired Referred Workers	Hired Non-Referred Workers	All Hired Workers	
Observe Characteristics Only	58%	39%	80%	78%	83%	70%	78.5%	13%**
Observe Referral Status Only	85%	0%	74%	N/A	77%	N/A	77.5%	N/A
Observe Characteristics & Referral Status	79%	9%	77%	83%	79%	76%	79.1%	3%
Applicant Pool Average			74%	68%	77%	63%	71%	15%***

Notes: Each row presents the results of a separate hiring scenario. In each scenario, employers use the available characteristics to predict workers' performance (likelihood of submitting work) and hire the 50% of workers with the highest predicted performance. The first row simulates hiring under the assumption that employers observe only workers' resume characteristics, but not their referral status. To calculate a given worker's predicted performance in this scenario, the performance of all other workers (excluding herself) are regressed on their resume characteristics listed in footnote 20. The estimated coefficients are then used to predict the excluded workers' own performance. Throughout the table, this prediction of performance – based on a worker's observable characteristics alone – is used as the measure of observable characteristics in Panel B. The second row assumes that employers observe only referral status, so they hire a random sample of referred workers such that 50% of the workforce is hired. The third row simulates hiring assuming employers observe workers' resume characteristics and referral status. To calculate a given worker's predicted performance here, the performance of all other workers (excluding herself) are regressed on their resume characteristics and referral status and the resulting coefficients are used to predict the worker's performance. For each scenario, Panel A presents the fraction of referred and non-referred workers hired. Panel B presents the estimated probability, based on their observable characteristics alone, that hired referred and non-referred workers submit work. Panel C presents the actual submission rate of the hired workers. The column labeled Difference provides the difference in average submission rates of the referred and non-referred workers hired under each scenario. **, *** denote the difference is significant at the 5%, and 1% levels, respectively.

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

<u>A. Individual Diligence</u>				
	Logged in	Submitted	Individual Question Correct	Own Criterion in Slogan
Referred Worker When Working with Own Referrer (Type A)	0.018 (0.018)	0.046** (0.018)	0.053* (0.030)	0.004 (0.035)
Non-Referred Worker When Working with Referrer (Type C)	-0.082 (0.053)	-0.129** (0.055)	-0.159*** (0.054)	-0.039 (0.057)
Base Group Mean (Type B)	0.883	0.837	0.755	0.440
Observations	846	846	846	846
R-Squared	0.419	0.381	0.294	0.180
<u>B. Team Performance</u>				
	Both Submitted	Team Question Matches	Same Slogan	Same Slogan & Both Criteria
Referred Worker and Own Referrer Team (Type A)	0.099*** (0.024)	0.287*** (0.030)	0.372*** (0.034)	0.103*** (0.025)
Non-Referred Worker and Referrer Team (Type C)	-0.122** (0.058)	-0.062 (0.054)	-0.023 (0.055)	0.004 (0.036)
Base Group Mean (Type B)	0.730	0.496	0.337	0.142
Observations	846	846	846	846
R-Squared	0.312	0.317	0.313	0.157

Notes: Each column in each panel reports the results of a separate regression of the dependent variable indicated by the column on indicators for being in a Type A team and for being in a Type C team. Observations in Panel A are at the worker-PSA level; only referred and non-referred workers (not referrers) are included. Observations in Panel B are at the team-PSA level. All regressions include the first- and second-order controls for worker characteristics listed in footnote 20. Standard errors are clustered at the blocking group level. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Appendix A: Time Spent, Enjoyment, and Communication in the Team Experiment

One reason that referrals may perform particularly well when working with their referrers is that they exert more effort in these pairings. They may face a higher return to effort because their effort also affects their friends' performance. Alternatively, if they enjoy working with their referrers, they may face lower cost of effort. While we cannot directly measure effort, we can analyze the amount of time workers spent on each of the three PSAs. (oDesk automatically records the time workers spend on the job.) Panel A of Appendix Table 8 shows time worked 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. Referred workers also spent significantly (13 percent) more time on the task when working with their own referrers.¹

Type A teams also communicated more than the other team types. Panel A of Appendix Table 9 shows how the 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 using other forms of communication. 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, the dependent variable is an indicator for reporting using other forms of communication. This final specification includes only teams that answered the communication question.

Type A teams communicated the most, both in and out of the chat box. Relative to Type

¹We do not know exactly what the workers were doing during this time. Referred workers could also have been more likely to spend working time off task when working with their referrers.

B teams, Type A teams were slightly, though insignificantly, more likely to use the chat box. When they did use the chat box, Type A teams sent about one-third more messages, though this difference is also not significant. 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 about twice as likely to do so; the magnitude of this difference implies that the difference itself cannot be driven by the small difference in the likelihood of answering this question.²

In general, workers who spent more time on the task and communicated more performed better. For example, teams that sent the median number of messages in the chat box (eight) were 18 percentage points more likely to answer the team question the same way and 14 percentage points more likely to provide the same slogan than were teams that did not use the chat box, all else equal. Teams in which each partner spent an additional five minutes 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, the number of messages sent, and the minutes spent by each partner, however, Type A teams still outperformed Type B teams. We replicate the main team performance specifications (from Panel B of Table 5), adding as controls an indicator for using the chat box, the number of messages sent in the chat box, an indicator for using other methods of communication, and the number of minutes spent by each partner.³ Type A teams remained 14 percentage points more likely to provide the same answer to the team question and 22 percentage points more likely to submit the same slogan than Type B teams.

Panel B of Appendix Table 8 shows that workers also preferred being on Type A teams. After they had completed all three tasks, workers reported which partner(s) they would be interested in partnering with again; workers could choose as many or as few partners as they wanted.⁴ We find that referrers were significantly more likely to want to work again with referred workers they did not know than with non-referred workers.⁵ But, referrers were more than twice as likely to want to partner again with their own referral as with someone else's

²At least one partner answered this question in 95 percent of Type B teams; Type B teams were insignificantly more likely to answer this question than either of the other team types.

³If 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 captures the effect of having at least one partner submit work.

⁴Some workers (about 20 percent) did not answer the question, mostly because they did not complete the third PSA task. But for those who answered, we know whether or not they wanted to work again with each of their three partners.

⁵Referrers did not know who, besides their own referrals, had been referred to the firm.

referral. Similarly, referred workers were substantially more likely to want to work again with their own referrer than with someone else's referrer.

Appendix Figure 1: *Peer Influence Experiment Task Site, Referred and Non-Referred Workers*

The screenshot shows a web browser window with the URL `www.pqanalytics.com/itask_g2/9151`. The page header includes the logo "PQANALYTICS" and a "Sign out" link. The main content is a table with the following data:

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

Below the table, the instructions read: "Please answer the following questions based on the table above." The questions and their corresponding input fields are:

- 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.)

A "submit" button is located at the bottom right of the form area.

Appendix Figure 2: *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 Figure 3: *Selection 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 @MileyCyrus.

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.)

Appendix Figure 4: Team Experiment Task Site

← → ↻ www.pqanalytics.com/team_task/31294

PQANALYTICS Sign out

Welcome, [REDACTED].

Please work with your assigned partner, [REDACTED], to come up with a catchy and informative slogan for a Public Service Announcement (PSA) on seat belt usage to be placed on a highway. One of you has received information on the efficacy of seat belts and the other partner has received information on highway drivers. Based on this information, you and your partner should agree on a single slogan. Both of you need to submit this slogan. Please submit the slogan and your answers to the other questions below by 11:59pm PHT on Friday, April 26. All answers should be submitted on this site; answers received over oDesk message will not be processed. Note that once you have submitted your responses, they cannot be changed.

Please use the chat window on this page to share information and collaborate with your partner. In the event that you are unable to get in contact with your partner, you may complete this task on your own.

Information about Highway Drivers

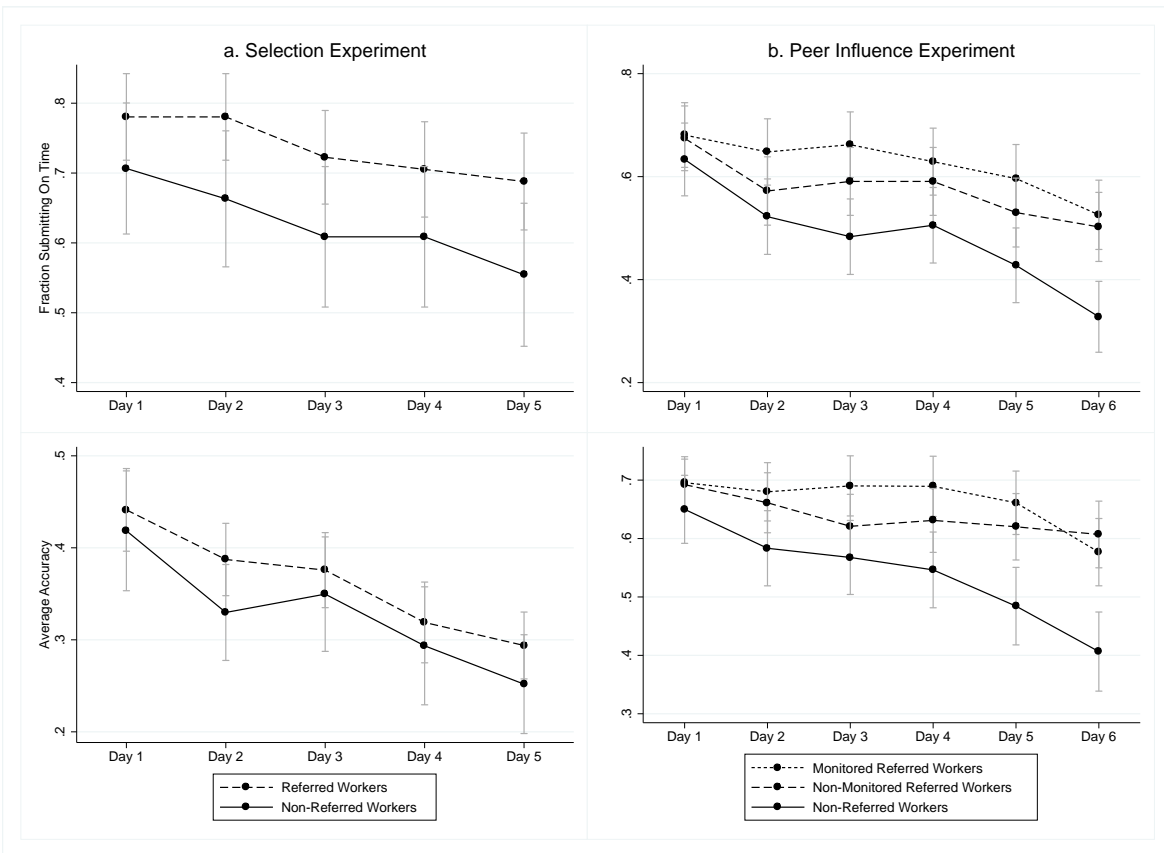
The United States has the world's largest network of highways and these highways are used by millions of Americans every day. There is at least one network in every state and highways interconnect most major cities. Driving on highways is a popular choice for commuters and travelers alike. The I-405 in Los Angeles, California alone sees an estimated 374,000 vehicles per day.

One reason that highways are very popular choices for drivers is that they often permit a high travel speeds. In some parts of some states, such as in rural western Texas, speed limits are as high as 80 mph (129 km/h). These high speeds allow drivers to travel long distances in shorter periods of time. But they also make car crashes on highways very dangerous – and often even fatal.

[REDACTED] 14:43

[REDACTED] 14:56

Appendix Figure 5: *On-Time Submission and Accuracy Rates by Day*



Notes: Error bars denote 95% confidence intervals.

Appendix Table 1. Descriptive Statistics: Workers Asked to Refer
Peer Influence and Team Experiments

	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 Passed Tests	1.00	0.98	***	1.00
Has Portfolio	0.69	0.61	***	0.69
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

Notes: Each statistic in the table presents the mean of the characteristic indicated by the row for the sample indicated by the column. *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. (See Figure 2 for details of the recruitment process.) *English Score* is self-reported English ability on a one-to-five scale, *agency-affiliated* workers pay a fraction of their earnings to report they are part of a given group of oDesk workers (an agency), and a *portfolio* is where a worker posts prior work. *, **, *** denote that the means of the characteristic for *Referred Someone* and *Referred No One* are significantly different at the 10%, 5%, and 1% levels, respectively.

Appendix Table 2. Descriptive Statistics: Referred and Non-Referred Workers

	A. Peer Influence Experiment			B. Selection Experiment			C. Team Experiment		
	Referred Workers	Non-Referred Workers	Difference	Referred Workers	Non-Referred Workers	Difference	Referred Workers	Non-Referred Workers	Difference
Has Prior Experience	0.74	0.34	***	0.82	0.46	***	0.55	0.17	***
Earnings	\$1,382	\$337	***	\$1,807	\$539	***	\$389	\$384	
Number of Previous Jobs	9.22	2.73	***	11.38	4.25	***	2.67	0.76	***
Has Feedback Score	0.63	0.29	***	0.71	0.40	***	0.48	0.16	***
Feedback Score	4.62	4.33	*	4.63	4.32		4.45	4.02	*
Posted Wage	\$2.70	\$3.19	***	\$2.78	\$3.38	***	\$2.46	\$3.48	***
Days Since Joining oDesk	527	573		580	639		401	569	***
Has Passed Tests	0.98	0.77	***	0.99	0.77	***	0.96	0.80	***
Has Portfolio	0.48	0.27	***	0.58	0.32	***	0.43	0.15	***
Has English Score	0.99	0.97	*	0.99	0.95	**	0.99	0.94	**
English Score	4.71	4.61		4.77	4.69		4.66	4.52	*
Agency Affiliated	0.07	0.05		0.08	0.22	**	0.16	0.07	**
Number of Degrees	1.42	1.04	***	1.38	1.16	*	1.28	0.96	***
Proposed Wage	\$2.39	\$2.57	***	\$2.45	\$2.64	***	\$2.30	\$2.63	***
Observations	255	180		173	92		282	94	

Notes: Each statistic in the table corresponds to the mean of the characteristic indicated by the row for the sample indicated by the column. Each observation is a referred or non-referred worker in the peer influence, selection, or team experiment. (See Figure 2 for details on recruitment of referred and non-referred workers into the experimental sample). *English Score* is self-reported English ability on a one-to-five scale, *agency-affiliated* workers pay a fraction of their earnings to report they are part of a given group of oDesk workers (an agency), and a *portfolio* is where a worker posts prior work. Characteristics are recorded at the time we first recruit the workers to the peer influence or team experiments. *, **, *** denote that the means of the characteristic for *Referred Workers* and *Non-Referred Workers* are significantly different at the 10%, 5%, and 1% levels, respectively.

Appendix Table 3. Randomization Assessment
Peer Influence Experiment

	Referrers		Referred Workers	
	Monitoring Treatment	Non-Monitoring Treatment	Monitoring Treatment	Non-Monitoring Treatment
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 Passed Tests	1.00	1.00	0.98	0.99
Has Portfolio	0.64	0.68	0.47	0.50
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 peer influence experiment are included. *English Score* is self-reported English ability on a one-to-five scale, *agency-affiliated* workers pay a fraction of their earnings to report they are part of a given group of oDesk workers (an agency), and a *portfolio* is where a worker posts prior work. * denotes that the Monitoring Treatment and Non-Monitoring Treatment group means are significantly different at the 10% level.

Appendix Table 4. Effect of Previously Having Been on a Type A Team on Team Performance
 Team Experiment: Base Group is Referred Workers Paired with Someone Else's Referrer (Type B Teams)

	Team Question			Same Slogan & Both Criteria
	Both Submitted	Matches	Same Slogan	
At Least One Partner Has Previously Been on Type A Team	0.105* 0.056	0.115** 0.050	0.055 0.058	0.030 0.046
Referred Worker and Own Referrer Team (Type A)	0.169*** 0.041	0.364*** 0.041	0.409*** 0.054	0.123*** 0.046
Non-Referred Worker and Referrer Team (Type C)	-0.107* 0.058	-0.045 0.054	-0.014 0.056	0.008 0.038
Base Group Mean (Type B)	0.730	0.496	0.337	0.142
Observations	846	846	846	846
R-Squared	0.316	0.321	0.314	0.169

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 whether at least one partner has previously been on a Type A team (i.e., has partnered with her referrer or referral). The regressions also include indicators for being on a Type A team and for being on a Type C team and dummies for this being the second or third task completed. Observations are at the team-PSA level. All regressions include the first- and second-order controls for worker characteristics listed in footnote 20. Standard errors are clustered at the blocking group level. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Appendix Table 5. Performance and Persistence with First-Order Controls: Selection and Peer Influence Experiments
 Base Group is all Referred Workers (Panel A) and Non-Monitored Referred Workers (Panel B)

	<u>A. Selection Experiment</u>					<u>B. Peer Influence Experiment</u>			
	Accepted Job Offer	Submission	On-Time Submission	Accuracy	Re-Application	Submission	On-Time Submission	Accuracy	Re-Application
Monitored Referred						0.020 (0.042)	0.038 (0.046)	0.014 (0.040)	-0.035 (0.034)
Non-Referred	-0.071 (0.056)	-0.100* (0.057)	-0.098* (0.059)	-0.024 (0.033)	-0.123* (0.071)	-0.115** (0.047)	-0.080* (0.048)	-0.095** (0.043)	-0.193*** (0.044)
Has Prior Experience	0.163* (0.093)	0.072 (0.079)	0.064 (0.087)	0.049 (0.049)	0.194** (0.091)	0.006 (0.063)	0.001 (0.066)	0.009 (0.059)	0.048 (0.049)
Earnings/1000	0.214** (0.084)	0.058 (0.076)	0.051 (0.077)	0.031 (0.046)	-0.026 (0.128)	0.009 (0.083)	-0.016 (0.083)	0.071 (0.084)	-0.003 (0.073)
Days Since Joining oDesk/100	0.005 (0.006)	0.009* (0.005)	0.013** (0.005)	0.006* (0.003)	-0.006 (0.008)	-0.005 (0.005)	-0.001 (0.005)	-0.003 (0.004)	0.001 (0.005)
Previous Jobs/10	0.011 (0.013)	-0.014 (0.013)	-0.014 (0.014)	-0.006 (0.007)	-0.001 (0.015)	0.013 (0.009)	0.013 (0.012)	0.000 (0.008)	0.003 (0.006)
Feedback Score	0.002 (0.040)	0.003 (0.037)	0.027 (0.043)	0.000 (0.022)	0.021 (0.048)	-0.021 (0.027)	0.035 (0.036)	-0.022 (0.027)	-0.064*** (0.021)
No Feedback Score	-0.051 (0.203)	0.050 (0.181)	0.163 (0.210)	0.025 (0.110)	0.213 (0.235)	-0.078 (0.134)	0.221 (0.176)	-0.099 (0.131)	-0.231** (0.094)
Posted Wage	0.002 (0.019)	0.005 (0.019)	0.008 (0.020)	-0.001 (0.011)	0.011 (0.025)	0.008 (0.014)	0.002 (0.014)	0.013 (0.014)	0.008 (0.014)
English Score	0.048 (0.037)	0.006 (0.049)	-0.004 (0.050)	-0.010 (0.030)	0.088 (0.062)	-0.028 (0.025)	-0.020 (0.029)	-0.010 (0.024)	-0.006 (0.029)
No English Score	0.480** (0.228)	-0.153 (0.297)	-0.217 (0.300)	-0.118 (0.171)	0.343 (0.351)	-0.139 (0.174)	-0.220 (0.185)	-0.066 (0.167)	0.215 (0.173)
Agency Affiliated	-0.080 (0.092)	-0.393*** (0.117)	-0.357*** (0.115)	-0.196*** (0.056)	0.063 (0.082)	-0.278*** (0.083)	-0.215*** (0.076)	-0.234*** (0.071)	-0.065 (0.075)
Has Passed Tests	-0.032 (0.093)	0.111 (0.104)	0.124 (0.104)	0.056 (0.052)	0.096 (0.137)	0.200*** (0.069)	0.179*** (0.068)	0.167*** (0.063)	0.291*** (0.084)
Has Portfolio	-0.006 (0.054)	-0.035 (0.048)	-0.025 (0.050)	-0.010 (0.028)	0.022 (0.062)	-0.075* (0.041)	-0.125*** (0.043)	-0.097** (0.038)	-0.055 (0.039)
Number of Degrees	-0.009 (0.026)	0.005 (0.026)	0.004 (0.026)	0.009 (0.014)	0.001 (0.032)	0.023 (0.018)	0.042** (0.019)	0.019 (0.017)	0.016 (0.019)
Constant	0.322 (0.287)	0.525 (0.290)	0.385 (0.313)	0.288 (0.175)	0.003 (0.404)	0.806 (0.197)	0.319 (0.233)	0.642 (0.186)	0.914 (0.177)
Observations	435	1,325	0.385	1,325	265	2,610	2,610	2,610	435
R-squared	0.125	0.079	(0.313)	0.048	0.088	0.075	0.061	0.063	0.156

Notes: Panel A replicates Panel B of Table 1, listing out the coefficients on each of the control variables; Panel B replicates Panel B of Table 2, again listing out the coefficients on each of the control variables. See the Table 1 and Table 2 notes for more details. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Appendix Table 6. Relationship Between Referrer Characteristics and Referred Worker Characteristics
Peer Influence and Team Experiments

		Dependent Variable: Referred Worker Characteristic									
	Earnings	Days Since Joining oDesk	Number of Previous Jobs	Feedback Score	Posted Wage	English Score	Agency Affiliated	Has Passed Tests	Has Portfolio	Number of Degrees	Proposed Wage
Referrer Characteristic	0.060* (0.034)	0.264*** (0.049)	0.095* (0.052)	0.204 (0.180)	0.231*** (0.049)	0.285*** (0.068)	0.317*** (0.050)	0.002 (0.002)	0.272*** (0.044)	0.075* (0.044)	0.367*** (0.052)
Constant	688.6 (113.3)	289.95 (33.337)	4.637 (0.685)	3.562 (0.867)	1.928 (0.146)	3.319 (0.333)	0.060 (0.011)	0.959 (0.015)	0.296 (0.035)	1.238 (0.069)	1.425 (0.134)
Observations	537	537	537	296	537	533	537	537	537	537	537
R-Squared	0.014	0.060	0.009	0.005	0.036	0.057	0.146	0.003	0.064	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 peer influence 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, *agency-affiliated* workers pay a fraction of their earnings to report they are part of a given group of oDesk workers (an agency), and a *portfolio* is where a worker posts prior work. Huber-White standard errors are in parenthesis. *, *** denote significance at the 10% and 1% levels, respectively.

Appendix Table 7. Characteristics of the Referrer-Referred Worker Relationship
Peer Influence and Team Experiments

How Well Referrer Knows Referral	
1 (Hardly at all)	1%
2	2%
3	5%
4	14%
5	20%
6 (Extremely Well)	57%
Observations	535
Frequency of Interaction	
Less than Once a Month	2%
About Once a Month	5%
Less than Weekly, More than Monthly	8%
About Once a Week	13%
Less than Daily, More than Weekly	21%
About Once a Day	19%
More than Once a Day	32%
Observations	533
Number of People Known in Common	
0 to 4	21%
5 to 9	16%
10 to 19	16%
20 to 29	11%
30 or more	37%
Observations	535

Notes: This table presents the distributions of referrers' responses to questions about their relationships with their referrals. Each observation is a referred worker hired in the Peer Influence Experiment or the Team Experiment.

Appendix Table 8. Time Spent and Task Enjoyment, by Team Type

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

	A. Time Spent (Minutes)		B. Wants to Partner Again	
	Referrers	Referred & Non- Referred Workers	Referrers	Referred & Non- Referred Workers
Referred Worker and Own Referrer Team (Type A)	5.922*** (1.811)	5.142*** (1.644)	0.556*** (0.031)	0.451*** (0.035)
Non-Referred Worker and Referrer Team (Type C)	1.135 (1.494)	-2.993 (3.985)	-0.100** (0.042)	0.102 (0.081)
Base Group Mean (Type B)	37.482	38.723	0.406	0.477
Observations	846	846	717	612
R-squared	0.183	0.301	0.394	0.308

Notes: Each column in each panel reports the results of a separate regression of the dependent variable on indicators for being in a Type A team and for being in a Type C team. In Panel A, the dependent variable is the number of minutes the worker spent on the task while, in Panel B, the dependent variable is an indicator for the worker reporting wanting to work again with her partner from that task. Observations are at the worker-PSA level. The first regression in each panel includes only referrers while the second includes only referred and non-referred workers. All regressions include the first- and second-order controls for worker characteristics listed in footnote 20. Standard errors are clustered at the blocking group level. *, *** denote significance at the 10%, and 1% levels, respectively.

Appendix Table 9. 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 Teams)

<u>A. Communication by Team Type</u>				
	Chat Box Use	Total Chat Messages (Conditional on Use)	Answered Communication Question	Reported Outside Communication
Referred Worker and Own Referrer Team (Type A)	0.025 (0.044)	5.098 (3.178)	-0.028 (0.017)	0.367*** (0.042)
Non-Referred Worker and Referrer Team (Type C)	-0.052 (0.050)	1.608 (3.996)	-0.003 (0.032)	0.038 (0.053)
Base Group Mean (Type B)	0.408	13.522	0.947	0.378
Observations	846	307	846	778
R-Squared	0.139	0.250	0.102	0.287
<u>B. Team Performance, Controlling for Communication and Time Spent</u>				
	Both Submitted	Team Question Matches	Same Slogan	Same Slogan & Both Criteria
Referred Worker and Own Referrer Team (Type A)	-0.024 (0.025)	0.141*** (0.031)	0.217*** (0.038)	0.021 (0.031)
Non-Referred Worker and Referrer Team (Type C)	-0.119*** (0.036)	-0.063 (0.043)	-0.030 (0.044)	-0.003 (0.030)
Used Chat Box	0.204*** (0.033)	0.151*** (0.040)	0.087* (0.046)	0.003 (0.047)
Total Chat Messages	0.000 (0.001)	0.004*** (0.001)	0.006*** (0.001)	0.003 (0.002)
Used Outside Communication	0.224*** (0.028)	0.307*** (0.032)	0.350*** (0.044)	0.177*** (0.034)
Minutes Spent by Referrer	0.003*** (0.001)	0.002*** (0.001)	0.002*** (0.001)	0.002*** (0.000)
Minutes Spent by Referred or Non-Referred Worker	0.004*** (0.001)	0.003*** (0.001)	0.002*** (0.001)	0.001*** (0.000)
Base Group Mean (Type B)	0.730	0.500	0.337	0.142
Observations	846	846	846	846
R-Squared	0.576	0.544	0.522	0.256

Notes: Each column in each panel reports the results of a separate regression of the dependent variable indicated by the column on indicators for being in a Type A and Type C teams. Observations are at the worker-PSA level. *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 the 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. Regressions in both panels control for the first- and second-order 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 chat messages sent, whether either partner 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 5%, and 1% levels, respectively.