Buyer Uncertainty about Seller Capacity: Causes, Consequences, and a Partial Solution*

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

July 12, 2017

Abstract

Employers in an online labor market often pursue workers with little capacity to take on more work. The pursuit of low-capacity workers is consequential, as these workers are more likely to reject employer inquires, causing a reduction in the probability a job opening is ultimately filled. In an attempt to shift more employer attention to workers with greater capacity, the market-designing platform introduced a new signaling feature into the market. It was effective, in that when a worker signaled having high capacity, he or she received more invitations from employers, rejected a smaller fraction of those invitations, quoted a lower price to do the work, and was more likely to be hired. A back-of-the-envelope calculation suggests the signaling feature alone could increase market surplus by as much as 6%, both by increasing the number of matches formed and by helping to allocate projects to workers with lower costs.

1 Introduction

In some markets, buyers propose transactions to sellers, and if the pursued seller has no capacity—or at least no capacity at a price the buyer is willing to pay—the buyers' efforts are wasted. An unsuccessful pursuit may be consequential to whether a "match" is ever made, depending on how hard it is for buyers to find substitute sellers. A potential remedy to this problem might be for a market designer to encourage buyers to pursue higher capacity sellers. The market designer might do this by providing buyers with timely, accurate, and fine-grained information about the capacity of individual sellers. This kind of informational intervention would be hard to implement in conventional markets—which may explain why this kind of information is rare—but would be relatively easy to implement

*This paper appeared as an extended abstract at the 2014 ACM Conference on Economics and Computation. Thanks to Eduardo Azevedo, Peter Coles, Panos Ipeirotis, Joe Golden, Ramesh Johari, Amanda Pallais, Richard Zeckhauser, and especially Andrey Fradkin for helpful comments and suggestions. Thanks to the participants at the market design seminar at the Wharton School and to the Economics of Digitization NBER Summer Institute participants for helpful comments and suggestion. I am sincerely grateful to many team members from the platform used in this study for their encouragement, ideas and assistance in this project.

in online marketplaces, which possess information systems to collect, process, and present market information.[1]

This paper explores the phenomenon of buyers pursuing sellers in the context of an online labor market. It also reports the results of an informational intervention designed to improve the recruiting and matching process. In this market, employers (i.e., buyers) posting job openings frequently pursue workers (i.e., sellers) by "inviting" them to apply to their job openings, just as employers recruit candidates in conventional labor markets.[2] A worker can "accept" the invitation and apply to the associated job opening, or, for any number of reasons, reject it. The most commonly cited reason for rejecting an invitation is insufficient capacity. These rejections appear to be consequential, as the associated job opening is much less likely to be filled, even though spurned employers can invite other workers, or hire from among the non-recruited "organic" applicants their job openings receive.

Despite worker claims about capacity being important—and the correlation between rejections and job openings going unfilled—it is far from clear whether there is actually a problem. Perhaps the negative relationship between a rejection and a match being formed is not causal, and instead reflects heterogeneity in the desirability of the associated job opening. Even if capacity does matter, an employer might be well-aware that his or her preferred worker has little capacity, but simply has a strong preference for that worker relative to the next best option. For the platform to successfully intervene, (1) there should be a causal relationship between recruiting success and match formation, and (2) employers must also find it useful to condition on platform-provided capacity information. Exploring these issues is the focus of this paper.

There are three main empirical portions to the paper: (1) a panel analysis exploring the relationship between worker capacity and invitation acceptance, (2) an instrumental variables analysis of how consequential rejections are to match formation, and (3) an analysis a platform intervention designed to reduce employer uncertainty about worker capacity.

To begin, I establish that busier workers receive more invitations, but are also less likely to accept an employer's invitation. However, as busier workers are likely to be better workers, the attributes of the worker are likely confounds—busier workers could also be pickier workers. To address this concern, I construct a panel dataset and then run regressions that include both worker- and time-specific fixed effects. Using this "within worker" approach, I find that when a worker gets more invitations, his or her acceptance rate goes down. As such, from an employer's perspective, the probability that a recruiting invitation to a particular worker is accepted varies based on time-varying factors that the employer likely cannot observe or condition upon, such as the number of invitations from other employers. If employers could condition on additional information about capacity, they might make different—and perhaps better—decisions about which workers to recruit.

Turning to the question of whether rejections are actually consequential, I first show that rejections are highly correlated with the associated job opening going unfilled. This relationship persists even with the inclusion of both worker- and employer-specific fixed effects, ruling out some of the more obvious selection stories that could explain the relationship. To

[1]In online contexts where platforms can convey capacity information, they do: Airbnb now shows up-to-date host calendars; ManuscriptCentral (the platform many scholarly journals use to handle the peer-review process) shows a reviewer's outstanding reviews to the assigning editor; Facebook shows individual relationship statuses; Uber matches passengers to drivers that currently have no passengers (or will have no passengers momentarily). Thanks to an anonymous reviewer for some of these other examples of platform responses to capacity uncertainty.

[2]I use the words "employer" and "worker" to be consistent with the extant literature rather than as a comment on the legal nature of the contractual relationships created on these platforms.

establish that this relationship is in fact causal, I conduct an instrumental variables analysis that takes advantage of some unique institutional features of the marketplace. I find that a rejection causes a reduction in the probability that the opening is filled. This causal relationship suggests that reducing rejections could increase the quantity of matches formed in the marketplace. A rejection also causes employers to interview and hire more non-recruited applicants; there is some evidence that spurned employers also engage in more recruiting. However, these employer adaptations are not sufficient to offset the overall negative effect of a rejection on whether a job opening is filled.

With the goal of reducing the number of rejected invitations, the platform created new features to give employers more information about the capacity of workers. The most interesting—and most consequential—intervention was the introduction of a signaling mechanism that allowed workers to publicly state their current capacity. Exploiting within-worker changes in the capacity signal being sent, I find that when workers signal they have more capacity, they (1) get more invitations from employers, (2) are more likely to accept an invitation by applying, (3) quote a lower hourly rate to perform the work, and (4) are more likely to be hired.

Interestingly, the bidding and hiring results imply that some workers with little capacity shade up their bids rather than simply reject an invitation outright. This suggests that employer uncertainty about worker capacity not only leads to job openings going unfilled, but also raises costs. In short, capacity uncertainty causes both a quantity and a price effect in the market. If the price effect is due to workers completing projects when they have higher costs, a back-of-the-envelope calculation suggests that employer uncertainty reduces surplus by as much as 3%. If we combine the causal estimates of the effects of a rejection on match formation with the estimates from the signaling feature introduction on acceptance rates, the loss in surplus from unfilled openings could be an additional 3%. Of course, these estimates rely on some strong assumptions, which I will discuss.

The main contributions of this paper are documenting an important market failure, quantifying its importance, and exploring potential remedies. Given the generality of the conditions that cause the failure—supply constrained sellers with imperfectly observable capacity—the problem is likely to be commonplace. This paper shows the problem is remediable in part, at least in markets where substantial market design powers exist, as in the case of computer-mediated markets. This paper adds to the growing literature on the design and functioning of online marketplaces.

2 Related work

There is a large literature on online marketplaces, online IT service marketplaces and a burgeoning literature specifically on online labor markets. Much of this work has, naturally, focused on information, as the defining characteristics of these online markets are that trading partners are usually strangers and goods cannot be directly inspected (Resnick et al., 2006). Dimoka et al. (2012) make an important conceptual distinction about whether the lack of information is about the product or the seller. In labor markets—or markets for services more generally—the two concepts are not so clearly distinct: a "bad" seller can by chance offer a great product, but a "bad" worker or service provider almost by definition produces bad work. The tight coupling of the seller and the "product"—and all the complications this coupling creates—is why the study of IT service markets is conceptually distinct from online markets generally.

The literature on IT service markets focuses largely on the determinants of match formation as mediated by either bidding (as in the case of procurement auctions) or as mediated by marketplace reputations. For example, Snir and Hitt (2003) explore entry into the reverse auctions run by buyers and identify a market failure: excess bidding, as would-be sellers do not internalize the costs of bid evaluation. Yoganarasimhan (2013) studies IT firms bidding for projects and explores how the dynamic nature of job-filling could lead to erroneous inferences about seller reputations if analyzed as a static estimation problem. In both of these examples, the focus is on sellers pursuing buyers, in which case seller capacity constraints are not directly relevant, as only sellers with capacity bother to pursue buyers. Further, even if the analysis did focus on buyers pursuing sellers, if the sellers are actual firms rather than individual workers, capacity constraints are presumably less important.

In existing online labor markets, there is a mix of sellers that are individual workers and sellers that are true firms. There is also a mix of buyers proposing fixed-price and hourly projects, which lead both to different bidding dynamics and very different work relationships (see Bajari and Tadelis (2001) on this distinction). There is also a mix of buyers pursuing sellers and sellers pursuing buyers. Further complicating things, both kinds of activity often occur simultaneously for the same job opening, with buyers both recruiting workers to apply to their jobs and evaluating unsolicited "organic" applications.

Given the diversity of the methods parties use to form matches in online labor markets, most of the research has focused on some particular aspect of the market to answer a research question rather than offer a general theory of online labor markets. For example, Pallais (2014) shows via a field experiment that past worker experience in online labor markets is an excellent predictor of being hired for subsequent work on the platform. Stanton and Thomas (2016) use data from an online labor market to show that agencies (which act as quasi-firms) help workers find jobs and break into the marketplace. Agrawal et al. (2016) investigate what factors matter to employers in making selections from an applicant pool and present some evidence of statistical discrimination. Hong and Pavlou (2015) provide a detailed look at how differences in time-zone, language and cultural factors affect prices in online labor markets. This paper fits this pattern of focusing on some aspect of the market, but it also takes a more design-based view.

As online marketplaces can be readily changed through changes in software, they are remarkably amenable to being designed. To date, research taking a "market design" view on platforms has largely focused on price structure and levels, often in the context of competition between platforms, e.g., Armstrong (2006), Rochet and Tirole (2003), Rochet and Tirole (2006) and Parker and Van Alstyne (2005). See Hong et al. (2015) for a comparison of open and sealed bid auctions in online labor markets.

Although price structure and levels are undoubtedly important, these decisions are generally made once or a small number of times, before the platform has received "traction." In contrast, the firm continually makes decisions that are consequential to the formation of matches, such as through algorithmic recommendations (Horton, 2017a), or the ranking of search results and the generation of choice sets (Halaburda and Piskorski, 2010; Casadesus-Masanell and Halaburda, 2014).

There is a growing recognition in the literature of the importance of the platform in shaping who matches with whom, even in decentralized settings. Tadelis and Zettelmeyer (2015) show that information disclosures by the platform can raise revenue at both the low end *and* high end of a used car market, mainly by helping buyers and workers sort, thereby thickening the market. Motivated by this finding, Lewis and Wang (2013) develop a model of the platform deciding how much to invest in search technology. Horton and Johari (2013) show that buyers will readily reveal their "type" with regards to quality preferences in anticipation of the workers sorting, even though this revelation allows workers to charge them a substantial premium. Arnosti et al. (2014) explore "stockout" and congestion in a dynamic matching market and find that application quotas can improve market efficiency.

Also in this "design" vein, Allon et al. (2012) present a theoretical model of the platform's choice about facilitating communication among platform participants, and the effects their decision has on efficiency. Goes and Lin (2012) examine the effects of a platform introducing paid certifications and, later, costly certifications. While the goal is partly to test a theory about information revelation, the paper also speaks to the platform's decision-making regarding signals that might reduce information asymmetries. This current paper also considers the platform's role in creating new signaling opportunities.

The most closely related paper to this one is Fradkin (2014), in which he shows in the context of Airbnb that over 70% of buyer inquiries do not lead to a match. Decomposing the reasons, Fradkin shows that screening by hosts explains half of the rejections, but the other half is explained by guests pursuing unavailable property listings—essentially the same market problem explored in this paper. Given the qualitative similarity in results despite such different settings, it seems likely that the phenomenon explored in this paper is commonplace in matching markets where buyers pursue capacity-constrained sellers whose at-that-moment capacity is not common knowledge.

3 Empirical context

In online labor markets, firms hire workers to perform tasks that can be done remotely, such as computer programming, graphic design, data entry, research, and writing. Markets differ in their scope and focus, but common services provided by platforms include publishing job listings, hosting user profile pages, arbitrating disputes, certifying worker skills, and maintaining reputation systems.

The online labor market used for this analysis is one of the largest, with over $1 billion in lifetime transaction volume. The platform focuses primarily on hourly contracts performed by independent workers, or "freelancers." Hours-worked are measured with a proprietary tracking software that workers install on their computers. The tracking software essentially serves as a digital punch clock that allows for remote monitoring of employees. This monitoring makes hourly contracts, and hence employment-like relationships, possible. This in turn makes the platform's marketplace more like a traditional labor market than project-based online marketplaces where contracts are usually arm's-length and fixed price. With this individual worker focus, the capacity of a worker is more important than in marketplaces where the supply side is composed of large firms.

On the platform used in this study, would-be employers write job descriptions, self-categorize the nature of the work and required skills, and then post the job openings to the platform's website. Workers learn about job openings primarily via electronic searches. Workers submit applications, which include a wage bid (for hourly jobs) or a total project bid (for fixed-price jobs), and a cover letter. In addition to worker-initiated applications, employers can also search worker profiles and invite workers to apply to their job openings; I will discuss this alternative channel of initial connection in more detail below, as recruiting is the central topic of this paper.

The platform used in this study is not the only marketplace for online work, or work more generally. As such, one might worry that every job opening we see on this platform is simultaneously posted on several other online labor market sites *and* in the conventional labor market. To assess this "multi-homing" question, the platform hired a market research firm, which in turn surveyed 6,192 randomly selected employers from the platform. When asked about what they would have done with their most recent project if the platform were not available, only 15% of employers responded that they would have made a local hire. Platform employers report that they are generally deciding among (1) getting the work done online, (2) doing the work themselves, and (3) not having the work done at all. The survey also found that 83% of employers said that they listed their last job opening only on the platform in question, and not on a competitor platform. The survey evidence suggests that online and offline hiring are only very weak substitutes, and that multi-homing of job openings is relatively rare (see Agrawal et al. (2015) for more on these points).

3.1 Employer recruiting

As in conventional labor markets, on the platform employers may choose to actively recruit candidates to apply for their jobs. These recruiting invitations are not job offers, but rather invitations to apply to the employer's already-posted job opening. Recruiting is common on the platform—about half of employers send at least one recruiting invitation.

Employer recruiting on the platform begins with the employer searching for some skill or attribute they are looking for in candidates. Employers search for workers using a search-engine like interface (see Appendix A for screenshots of the interfaces discussed in this section). As such, the ordering of search results is consequential. The search results were historically ordered by a weighted combination of attributes employers care about, such as on-platform experience and feedback scores. However, as it became clear that worker invitation responsiveness mattered to match formation, a measure of "eagerness" was also added to the weighting formula, with workers recently applying to jobs (hence revealing their availability) getting more weight. These early attempts at conveying worker capacity proved, however, to be a less than satisfactory solution, as there were some reports of workers applying to jobs they were not interested in simply to appear more eager and hence appear higher in search results.

Once an employer finds a worker they are interested in, they can click on a button to invite that worker to apply for their job opening. The employer can also "click through" to the worker's full profile before inviting them. The full profile has more information about the worker, such as their full disaggregated work history (rather than just summaries, such as the total hours worked on the platform).

When the employer clicks "contact" they are brought to a new "invite" interface. The "message" text box is pre-populated with a short written request for the worker to apply, which the employer is free to customize. From the worker perspective, invitations appear as messages in an inbox of sorts. For each invitation, the worker can see the date and time the invitation was sent, the title of the associated job, and the employer that sent the invitation, as well as the employer's message. By clicking the title of the job, the worker can learn more about both the job opening and the employer.

When workers turn down a recruiting invitation, they may give a reason from a list of reasons, or write their own. Using all responses from November of 2013, the most common reason selected is "Too busy on other projects" at 48%. The next most common is "Not interested in the project" at 29%. All other reasons get less than 10%. This data—as with all data used in the paper—was shared directly by the platform. If we take these worker responses at face value, worker capacity is important, but there is also evidence of substantial discretion, with invited workers conditioning their response decisions on the attributes of the job opening. If an invitation is accepted, the invited worker applies and the employer eventually evaluates applicants. If the employer ultimately makes a hire, the platform intermediates the relationship.

3.2 What do employers know about worker capacity when they recruit?

Employers want to recruit workers that are likely to accept their recruiting invitations. As such, we would expect employers to try to infer worker capacity. However, in the earlier days of the platform, employers had little to work with: employers could not easily observe how many other recruiting invitations a worker had received, nor the response to those invitations. However, employers did have some imperfect proxies for capacity, in that they could observe the cumulative hours-worked by that worker for each of the worker's projects. From this, the employer could try to estimate how many hours the worker was working per week. However, there was, and still is, substantial individual heterogeneity in hours-worked per week—some workers work full-time, while others work only a few hours per-week—and so it is unlikely that employers could infer much about capacity.[3]

More recently, the platform has introduced a number of features to make worker responsiveness and proxies for capacity available to employers, some of which I will discuss in Section 6. One change not discussed is that employers can now also see the relative difference between their own timezone and that of the worker, in hours, to help them gauge whether they will get a quick response to an invitation.

4 Empirics of employer recruiting and the worker response

In this section, using a large dataset of recruiting invitations, I show that a worker's probability of accepting a recruiting invitation declines with the number of other invitations received during the same time period. Using time variation in how heavily a worker is recruited, I present evidence that the negative relationship is not driven by selection but rather that getting more invitations causes the worker to be pickier about which projects to accept. Although I lack an experiment, I use various approaches to rule out alternative, non-causal explanations. The importance of this finding is that if workers have time-varying capacity, then the platform has some justification for trying to "balance" recruiting invitations across workers.

4.1 Data description

I collected a sample of all recruiting invitations sent by employers from January 1st, 2010 until January 1st, 2013.[4] I then restricted the sample to invitations to openings that were

[3]An emerging stylized fact in computer-mediated labor markets is that when workers are free to choose how many hours to work, there is substantial heterogeneity in the realized number compared to most traditional employment relationships. See Hall and Krueger (2016) on this point in the context of the market for Uber drivers in the US.

[4]Prior to 2010, there are some database-related complexities in tracking who was invited to a job opening that can be avoided by restricting the sample. As the marketplace was considerably smaller then, left-truncating the data is not important in terms of the sample size. The right limit of 2013 was chosen because

public, meaning all workers could see the opening and apply. I also removed any invitations where the recruited worker and employer had interacted previously, such as through a past employment relationship, a previous recruiting invitation, or a completed application by that worker to one of the employer's previous job openings. The motivation for this restriction is to remove any invitation that was, in some sense, pre-arranged.[5] I also eliminated invitations by employers sending 10 or more invitations for their job opening, as these "mass invite" cases are not bona fide recruiting attempts, but are more akin to spam. These restrictions leave 1,247,379 usable recruiting invitations sent to 240,878 distinct workers by 68,255 distinct employers. The gross invitation acceptance rate for the sample is 40.1%.

I collapsed the sample of invitations to create a worker-week panel. I can do this because for each invitation, I observe the precise time it was sent, to the millisecond. The panel has 1,011,380 worker-week observations. For each week, I observe the number of hours worked, the number of invitations received, and the number of invitations accepted. As my focus is on changes in the acceptance rate, the panel only includes worker-week observations in which at least one invitation was received.

Table 1: Summary statistics on weekly invitation panels (n = 1,011,380)

Statistic	Mean	St. Dev.	Min	Max
Number of invites received / week	2.096	6.780	1	1,061
Number of invites accepted / week	0.495	0.631	0	15
Fraction accepted (accepted invites)/(invites)	0.354	0.445	0.000	1.000
Fraction accepted (1 + accepted invites)/(2 + invites)	0.467	0.171	0.023	0.933
Any hours worked that week?	0.210	0.407	0	1
Exactly one invite?	0.695	0.460	0	1

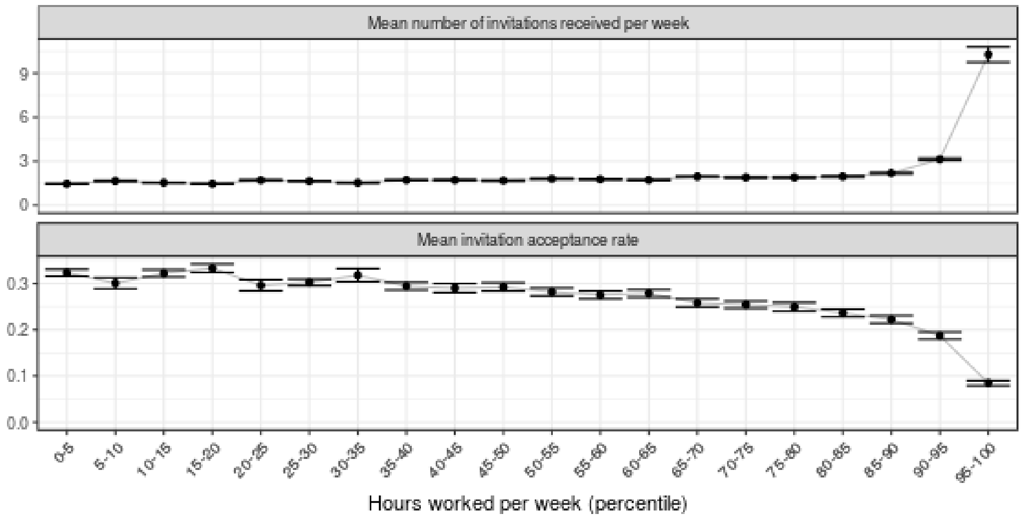
Notes: This table reports summary statistics for an unbalanced weekly panel of workers. The data is from January 1st, 2010 through January 1st, 2013. There are 240,878 distinct workers in the panel. A worker only has an observation for a week if they received at least one invitation that week.

Summary statistics for the panel are reported in Table 1. Note that there is substantial heterogeneity in the weekly number of invitations received by a worker. Although the mean number of invitations is a little more than two, almost 70% of the worker-week observations are weeks in which the worker received exactly one invitation. The maximum number of invitations received is over a thousand, though even triple digit numbers of invitations is extraordinarily rare; the 99.9th percentile is only 73 invitations. Consistent with workers being selective about which invitations they pursue by applying, the maximum number of accepted invitations is only 15. We can also see that during many of the worker-week observations, the worker in question worked no hours on the platform. These workers might be on vacation or simply have no work to do, but some have presumably left the platform. Their inclusion in the sample presumably biases downward the acceptance rate compared to what it would be if the sample only consisted of still-active workers.

the platform made a number of changes in response to earlier versions of this analysis that potentially affected the phenomena being investigated.

[5] Gefen and Carmel (2008) also look at an IT marketplace and find that buyers show a strong preference for firms/workers that they have worked with in the past. This is certainly true on the platform, and it is more evidence towards the importance of information in forming matches: presumably part of the preference for previous contacts is that there is far more information available to both sides.

Figure 1: Distribution and acceptance of recruiting invitations by worker weekly hours-worked bins



Notes: This figure shows the mean number of invitations received per week (top panel) and the mean invitation acceptance rate (bottom panel) for a sample of workers receiving at least one recruiting invitation and working at least one hour. Workers are classified into 20 equally spaced bins based on their hours worked per week. 95% confidence intervals for the means are shown in both panels.

The average acceptance rate over all worker-weeks is about 35%. As it will be useful to have an individual estimate of acceptance probability strictly in $(0, 1)$, I also report $(1 + A)/(2+I)$, where I is the number of invitations received and A is the number accepted. This monotonic transformation gives an acceptance fraction that is slightly larger, as expected given that the untransformed acceptance rate is less than 1/2.[6]

4.2 Where do employers send their invitations?

Not apparent from Table 1 is the extent to which employers focus on recruiting the busiest workers. Using the panel data described above, the top facet of Figure 1 shows the mean count of per-week invitations by the number of hours-worked per week (split into 20 evenly spaced percentile bins). The sample is restricted to workers working at least one hour in the week in question. The count of applications received is strongly increasing in hours-worked per week, and invitations are concentrated in the right tail. As such, many invitations go to workers who are already quite busy and thus unlikely to have the capacity to take on more work. The data bears out this intuition, in that the bottom facet of Figure 1 plots the average acceptance rate by the same hours-worked bins, showing that the busiest workers also have the lowest acceptance rates. Acceptance rates go from above 30% for the workers with the fewest hours-worked to below 10% for workers with the most hours-worked.

The pattern in Figure 1 does not imply that employers are mistaken in their recruiting

[6]From a Bayesian perspective, the transformed acceptance rate is the posterior of the acceptance rate when the prior on the acceptance rate is an uninformative uniform distribution.

decisions, as the busiest workers might still be the most attractive workers to pursue, even with their lower acceptance probability. However, if we see that a worker's acceptance probability fluctuates over time, it is likely that employers are not fully taking advantage of these changes, as fully-informed employers would shift invitations towards periods with high acceptance rates from periods with low acceptance rates, at least in a thick market with many alternative workers. In a nutshell, if there is within-worker variation in the acceptance rate that can be explained by within-worker variation in attributes, it undercuts the notion that employers are doing the best they can do already.

One time-varying attribute of potential interest to an employer is the number of invitations the worker has recently received from other employers. If the worker gets lots of invitations in some unit of time, we might suspect that they would be more selective during that time, and vice versa when they get fewer.

In Table 2, I report estimates of the effects of the weekly invitation count on a worker's log transformed acceptance rate. Before presenting within-worker estimates, I first present a cross-sectional regression of the log of the transformed weekly acceptance rate on the weekly invitation count. Column (1) of Table 2 reports a pooled OLS estimate of

$$\log p_{it} = \alpha I_{it} + \textsc{Week}_t + \epsilon_{it}, \tag{1}$$

where p_{it} is the transformed acceptance rate for worker i in week t, I_{it} is that worker's invitation count, and $\textsc{Week}_t$ is a week-specific effect.[7] We can see that there is a strong negative relationship between the number of invitations received and the acceptance probability. In terms of magnitude, a worker going from receiving 10 invitations per week to 20 invitations per week would have an acceptance rate that is about 4.6% lower.[8] Note that as the regression includes week-specific fixed effects, the relationship cannot be explained by changes in the platform or general market conditions.[9]

Now I turn to the within-worker estimates of how invitation counts affect response rates. Column (2) of Table 2 reports an estimate of Equation 1, but with the inclusion of a worker-specific fixed effect. The negative relationship between invitations and the acceptance rate is still highly significant and close in magnitude to the Column (1) estimate. In short, when a worker gets more invitations, he or she is pickier about which invitations to accept.

Despite the persistent negative relationship between invitations received and acceptance probability, even when using only within-worker variation, workers and markets are not static. Even if worker attributes are not changing quickly, perceived worker attributes could be—say through better or worse feedback scores.[10] Demand for certain skills might also be changing over time. To address these concerns, I cannot include worker-week effects—this fixed effect would absorb all the variation in weekly invitation counts. However, I can include worker-month fixed effects, but still use a weekly panel.

Column (3) of Table 2 reports an estimation of Equation 1, but with the inclusion of worker-month fixed effects. There is still a strong negative relationship between the number of invitations received and the acceptance rate of those invitations. The magnitude of the

[7]One empirical concern is that because the count of invitations appears on both the left- and right-hand sides of the regression, the estimates are subject to attenuation bias. However, because of the computer-mediated nature of the platform, the count of invitations and whether or not they are accepted is measured without error, making this concern unfounded.

[8]The percentage change is $\approx (e^{20\hat{\alpha}} - e^{10\hat{\alpha}})/e^{10\hat{\alpha}}$.

[9]The inclusion of these week fixed effects seemingly does little, as the coefficient on invites does not change when they are removed. This analysis is not shown but is available upon request.

[10]This is somewhat unlikely, as workers receiving invitations are already likely to have nearly perfect reputation scores (Horton and Golden, 2015).

Table 2: Association between weekly invitation acceptance rates and the number of recruiting invitations received that week

	Dependent variable:		
	Log transformed invitation acceptance rate:		
	(1)	(2)	(3)
Invitations/week	-0.007^{***}	-0.006^{***}	-0.005^{***}
	(0.001)	(0.001)	(0.001)
Worker FE	No	Yes	No
Week FE	Yes	Yes	Yes
Worker-Month FE	No	No	Yes
Observations	1,011,380	1,011,380	1,011,380
R^2	0.021	0.393	0.808
Adjusted R^2	0.021	0.203	0.228
Residual Std. Error	0.369 (df = 1011169)	0.333 (df = 770292)	0.327 (df = 250939)

Notes: This table reports regressions where the dependent variable is the worker's transformed weekly invitation acceptance probability. See Section 4.1 for a description of the sample. In Column (1), only week fixed effects are included. In Column (2), worker-specific and week-specific fixed-effects are included. In Column (3), worker-month specific and week-specific fixed effects are included. Standard errors are clustered the level of the individual worker. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

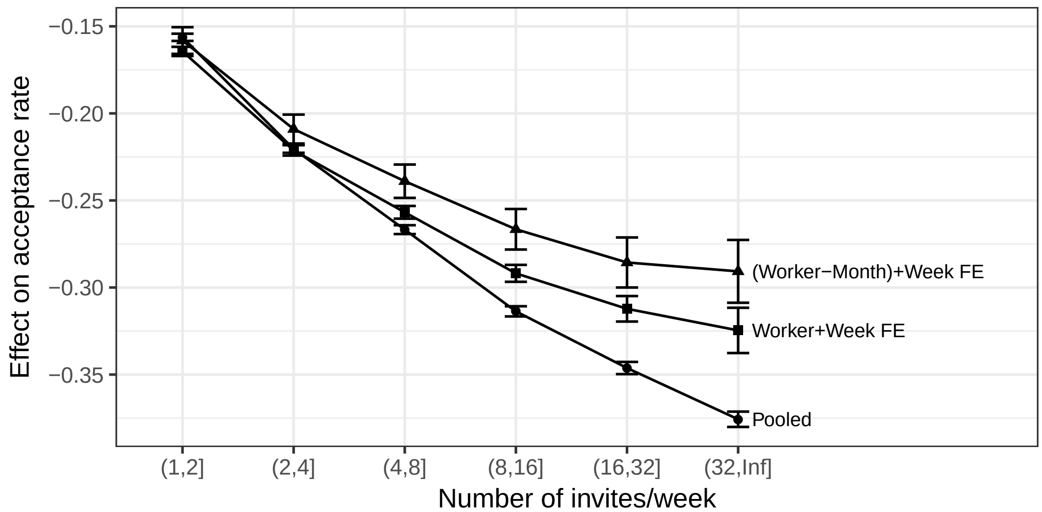
effect is only somewhat closer to zero, compared to the Columns (1) and (2) regression estimates, and the difference in the estimates is far from conventionally significant. To find a spurious relationship between the count of invitations and the acceptance rate with the presence of worker-month fixed effects, demand and perceived worker productivity would have to be (implausibly) fluctuating within the week, but not within the month.

The regression specification in Equation 1 forces all fluctuations in the count of invitations to have the same marginal effect on the response rate in percentage terms. This is an unattractive assumption—a worker getting 1 invitation one week and 2 the next is clearly different from a worker getting 50 one week and 51 the next. An alternative empirical approach that relaxes this linearity assumption is to create a series of indicators for different "bins" of invites per week, and then regress the untransformed acceptance rate on this collection of indicators, while still including worker fixed effects:

$$p_{it} = \sum_k \beta_k \cdot \textsc{InviteBand}_{it}^k + \textsc{Week}_t + \gamma_i + \epsilon_{it}. \tag{2}$$

Figure 2 reports the $\hat{\beta}$ coefficients from Equation 2, using InviteBands split at powers of 2, with workers receiving exactly one invitation as the omitted group. Mirroring the Table 2 analysis, estimates are shown for the regression without worker effects, labeled "Pooled," with worker fixed effects, labeled "Worker+Week FE," and with worker-month fixed effects and week fixed effects, or "(Worker-Month)+Week FE." The point estimates can be interpreted as the effects for a worker moving from receiving one invitation to some other invitation "band." For example, a worker going from 1 invitation to some number between 9 and 16 (the $(8, 16]$ band) would have about a 25-30 percentage point reduction in

Figure 2: Effect of weekly invitation counts on acceptance fraction



Notes: This figure reports point estimates from Equation 2 using different sets of controls. The sample is the same as used in Table 2. The key independent variables are indicators for "bands" of weekly recruiting invitations received. Mirroring the Table 2 analysis, estimates are shown for the a regression without worker effects, labeled "Pooled," with worker fixed effects, labeled "Worker+Week FE," and with worker-month fixed effects and week fixed effects, or "Worker-Month+Week FE." The outcome is the acceptance rate for invitations received, untransformed. For each point estimate, 95% CI are shown, with standard errors clustered at the worker level.

their acceptance rate. We can see that various controls affect the estimates, but the general pattern is clear: with more invitations, acceptance rates fall.

4.3 Discussion of employer recruiting and worker response results

Regardless of the method, the reduction in invitation response rates is fairly small, which is seemingly consistent with workers being flexible about how much work to take on.[11] This finding would seemingly undercut the notion that conditioning on worker capacity is important, as even "busy" workers would readily scale up to meet additional demand. However, there are several counter-arguments to consider. For one, accepting an invitation simply means applying, and so a capacity-constrained worker can respond by applying but with a very high price. Indeed, later in this paper, I show that some workers with low stated capacity still applied when recruited, but they raised their wage bids. Even if a worker is offered a contract, they do not necessarily have to agree to the contract or complete the work—they have a job offer, which is useful, even if not pursued.

The results from the worker panel analysis strongly suggest that workers become pickier when they have many projects to choose from. If projects could be "stored" the average market-wide acceptance rate could be increased by cross-leveling invitations "within" a

[11] For a market that focuses on relatively short-term tasks and has a work-force with many people working part-time, this is perhaps unsurprising. Farronato and Cullen (2015)—looking at TaskRabbit, a market with similar task sizes and relationship durations—also find that workers have very high labor supply elasticities, which can imply a high degree of flexibility in hours-worked.

worker. Although most projects are time-sensitive, making storage infeasible, so long as not all workers are busy at the same time, it also seems likely that cross-leveling "between" workers is desirable. However, the expected benefit from improving the invitation acceptance rate depends on how consequential invitation rejections are to match formation, which is the focus of the next section.

5 Recruiting success and match formation

In this section, I first show that when an employer sends an invitation and it is rejected, the associated job opening is much less likely to be filled. This correlation cannot be interpreted as causal, but I also show that with both worker- and employer-specific fixed effects, the negative relationship persists, undercutting the most obvious selection-driven explanation. To show that the relationship is in fact causal, I conduct an instrumental variables analysis.

5.1 Data used to estimate the effects of rejected invitations on match formation

I constructed a sample of 57,250 recruiting invitations. The time period covered by the sample was January 1st, 2011 through January 1st, 2013, inclusive, which is the same time period used for the panel analysis. The sample is not the universe of all invitations, but rather it is a selected sample that meets several geographic, job-opening, and time-based restrictions. I will discuss these restrictions in detail, but the motivation for all restrictions is to create a sample where the conditions for causal inference can be met. The restrictions will make more sense when I discuss the instrument, but first I will simply describe the sample.

The sample was restricted to invitations sent by US-based employers to workers residing in Bangladesh, India, Philippines, Russia and Ukraine.[12] These countries are the major non-US worker countries on the platform, with the UK excluded. I excluded the UK because the close cultural and linguistic ties make it more likely that US-based employers would consider the local time in the worker's country when sending their invitations. This conditioning would violate the exclusion restriction assumption, which I will discuss later when I explain the instrumental variable. I further restricted the sample to invitations by employers that sent one and only one "early" invitation for their opening, which I define as being sent within the first hour after posting the opening.

Among US employers doing any recruiting at all, the fraction sending a single early invitation is about 30%.[13] This is also the most common number to send, conditional upon sending any. I restrict the sample to job openings with a single early invitation, in part, to simplify the analysis and interpretation. This "early" restriction allows me to look at later compensatory recruiting by employers as a response to having their invitations rejected. However, the main reason to use the single invitation scenario is that employers sending multiple invitations are likely doing so precisely because they are concerned that some or all of their invitations will be rejected—and this probability of rejection might be high precisely because of the capacity uncertainty problem this analysis investigates. As such, the single

[12] I only include jobs that were open to all applicants, both recruited and organic (i.e., "private" jobs for which a worker needed to be invited to apply were excluded).

[13] The figure is with respect to all public job openings posted before 2014.

invitation context offers the best scenario to understand whether rejections are, per se, consequential.

Table 3: Summary statistics for employer recruiting invitations (n = 57,250)

Statistic	Mean	St. Dev.	Min	Max
Job opening filled?	0.446	0.497	0	1
Invitation accepted?	0.545	0.498	0	1
Hour invitation sent (PT)	12.900	7.426	0	23
Value of the instrumental variable	0.040	0.017	0.008	0.077

Notes: This table reports summary statistics for the data set used in the instrumental variables analysis. The sample consists of recruiting invitations by a sample of employers and the ultimate outcome for that associated job post. See Section 5.1 for the full description of the dataset.

The final sample of 57,250 invitations were sent by 40,971 distinct employers to 24,282 distinct workers. Table 3 contains the summary statistics for the dataset. The gross acceptance rate for this sample is about 55%. This is higher than in the panel, which could reflect that employers sending a single invitation are more confident that their invitation will be accepted. Note that the probability that a job opening is filled is less than 50%, meaning that there is plenty of “room” to increase the number of matches. Also note that the average invitation hour (on a 24 hour clock) is 12.90, or about 1pm Pacific Time/4pm Eastern Time, consistent with the sample being restricted to US-based employers that keep customary US business hours. This concentration of recruiting activity during the middle of the US work day will be important when explaining the identification strategy.

My main empirical focus in this section will be on how the acceptance of the single recruiting invitation affects job opening outcomes. Before turning to regression results, we can simply compare hiring outcomes for job openings by whether the worker accepted or not, which I do in Table 4. The top panel shows the outcomes when the recruited worker rejects the invitation. Unsurprisingly, that rejected worker is never hired. About 6% of the time the firm hires a later recruited worker, but the most common action is to hire an non-recruited organic applicant, which occurs about 44% of the time. 52% of the time, the firm hires no one at all. In the next panel, the same results are shown, but for openings where the worker accepts the invitation. Here, the most common outcome is for the accepting early recruited worker to be hired, which occurs nearly 40% of the time. The fraction hiring organic applicants drops substantially, to only about 24%. There is no evidence in an uptick in hiring other recruited applicants. Note that the fraction of employers not hiring drops substantially, going to a bit less than 40%, compared to 52% when the recruited worker rejects the invitation.

5.2 Correlation between invitation acceptance and match formation

Now I consider the effects of the invited worker's acceptance decision in a regression framework. Consider an employer, j, sending a single recruiting invitation to their preferred worker i. Let $a_{ij} = 1$ if the worker accepts the invitation and applies, and $a_{ij} = 0$ if he or she declines or ignores the invitation. Using the sample of invitations, Column (1), Table 5

Table 4: Characteristics of the worker hired by whether the the initially recruited applicant applied (in %s)

Early recruit declined		
	Early recruit hired	0.0
	Late recruit hired	6.0
	Organic applicant hired	44.0
	No hire made	52.0
Early recruit accepted		
	Early recruit hired	38.4
	Late recruit hired	6.0
	Organic applicant hired	24.0
	No hire made	39.4

Notes: This figure reports the composition of the worker hired in the sample, by the acceptance status of the early recruited worker. See 5.1 for a description of the sample. As employers can make multiple hires, percentages do not have to sum to 100%. For these hire measures, I use whether a contract was formed rather than whether any money was spent.

reports an OLS estimate of

$$\mathbf{1}\{Y_j > 0\} = \beta_0 + \beta_1 a_{ij} + \epsilon_{ij}, \tag{3}$$

where Y_j is the amount of money spent by the employer on workers hired to that opening.[14] Standard errors in each regression of Table 5 are clustered at the level of the worker.

We can see in Column (1) that an accepted invitation is strongly associated with the job opening being filled, with an effect of more than 15 percentage points. However, this regression clearly cannot be interpreted causally. On the firm side, firms proposing poorly described, low-value projects might be more likely to be rejected, and less likely to be filled by some other worker; firms that require highly sought-after skills are also more likely to be turned down; firms with unobserved off-line arrangements with the worker always fill their jobs and have a nearly 100% accept rate, and so on. On the worker side, if the "best" workers are more likely to turn down a recruiting invitation, workers observed accepting invitations will be relatively less attractive to employers and their acceptance could have a smaller effect on the probability a job opening is filled. Essentially any job opening, employer or worker characteristic that is correlated with the probability that the job opening will be filled would bias the OLS estimate. Appendix B presents these arguments more formally, showing how bias in any direction is plausible given either worker- or employer-selection.

I can address some of these worker or employer omitted variable concerns by exploiting the matched worker/employer nature of the invitation data. Column (2) of Table 5 shows the results from a regression

$$\mathbf{1}\{Y_j > 0\} = \beta_0 + \beta_1 a_{ij} + \gamma_i + \delta_j + \epsilon_{ij}, \tag{4}$$

where γ_i and δ_j are worker- and employer-specific fixed effects, respectively. With the inclusion of these controls, the association between invitation acceptance and match-formation

[14] I use money spent as an indication of match formation because it is a revealed preference measure, whereas simply stating an intention to work with someone is only a stated preference, and may not necessarily imply an economic relationship was formed.

Table 5: Association between a worker accepting an employer's recruiting invitation and whether that employer fills his or her opening

	Dependent variable:	
	Job opening filled, $\mathbf{1}\{Y_j > 0\}$	
	(1)	(2)
Workers accepts invitation, $a_{ij} = 1$	0.163^{***}	0.210^{***}
	(0.005)	(0.060)
Constant	0.357^{***}	
	(0.004)	
Worker and Firm FEs?	No	Yes
Observations	57,250	57,250
R^2	0.027	0.945
Adjusted R^2	0.027	0.202
Residual Std. Error	0.490 (df = 57248)	0.444 (df = 3914)

Notes: This table reports regressions where the outcome variable is an indicator for whether an employer filled his or her opening. The dataset consists of a sample of recruiting invitations described in Section 5.1. The key independent variable is whether the recruiter worker "accepted" the invitation by applying to the employer's opening. In Column (1), no controls are included, whereas in Column (2) worker- and employer-specific fixed effects are included. Standard errors in each regression are clustered at the level of the individual worker. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : * * *$.

does not disappear—on the contrary, an acceptance is associated with a 20 percentage point increase in match probability, though the estimate is now considerably less precise.

Although the double employer and worker fixed effect approach does nothing to deal with omitted variables at the level of the job opening—or time-changing attributes of the worker and employer—it does rule out selection explanations that depend solely on the identity of the invited worker and employer. In short, it rules out the possibility that workers simply turn down "bad" employers who are unlikely to fill their job openings, or that naive employers invite workers with no capacity. Next I turn to an approach that can deal with job-specific attributes and time-changing worker/employer characteristics that could affect both the probability a job opening is filled and whether the invited worker accepts the invitation. However, before doing so, it is useful to first consider why rejections might be consequential to match formation.

Firms presumably recruit to improve their eventual pool of applicants. Following a recruiting invitation rejection, this applicant pool is "missing" the recruited worker. Without this recruited worker, the firm's payoff from making a hire might be negative, and so they may choose to hire no one at all. Horton (2017a) demonstrates this argument in the opposite direction, showing that an increase in the size of the applicant pool increases the probability a job opening is filled.

Assuming for a moment that the reduction in match formation is *caused* by a rejection, a natural question, however, is why, in a large and presumably thick market, does the rejected employer not simply recruit someone else? Or barring that, why not simply recruit more workers ex ante as a precaution? Recruiting is costly, so it is no puzzle why some firms invite only a single applicant, or even none at all, despite the possibility of rejection. But stopping a search following a rejection cannot be explained solely by costly recruiting—it was also costly to send the initial recruiting invitation. One possibility is that after a rejection, the firm might infer that recruiting another candidate will be costlier. Given that most employers post a job opening and then recruit candidates in a single session, starting another search session after a rejection might have a higher fixed cost than the original recruiting session. After a rejection, the employer might also infer that their project is less desirable, lowering the returns to additional recruiting.

5.3 Instrumental variable construction

To estimate the causal effect of a recruiting acceptance, ideally the platform would, at random, turn some invitation acceptances into rejections. For a number of reasons this approach is infeasible. However, the platform offers a natural experiment that approximates this hypothetical, providing an instrumental variable for a_{ij}. The instrument requires some explanation, but the core idea is that the relative time of day in which a worker receives a recruiting invitation has a strong effect on their probability of accepting that invitation, and yet there is no evidence that employers take this relationship into account when deciding which workers to invite.

As noted earlier, most employers on the platform are from the US. They post job openings and send invitations during customary US business hours (recall Table 3). Workers are distributed around the globe, and so the local time an invitation is received varies, depending on when it is sent and to which worker. Although workers adapt to the US-centric rhythm of the market, this adaptation is only partial, with most workers still keeping their home country work hours, more or less. As such, the variation in the local time of an invitation leads to variation in whether a worker happens to be online when an invitation arrives.

Why does the local time when an invitation is received matter? One might imagine that since the recruited worker will eventually come online and see the invitation, a few hours of difference cannot matter to the decision of whether to apply. However, this is empirically not the case. The reason is interesting, and has an economic basis. Because workers are uncertain about when and how an employer will evaluate candidates, all else equal, it is better to apply sooner rather than later. To see this, first suppose the employer will consider applications "in batch" after some amount of elapsed time. As long as the worker is in the batch, they will be considered. This would seemingly undercut the urgency of applying, but if workers do not know when the batch will be processed, applying earlier increases their chance of being considered. The longer the expected time until processing, the weaker the incentive to apply early, but the incentive does not go away. The incentive to apply early also exists if employers monitor the "flow" of applicants, hiring the first one above their reservation utility—now the worker wants to apply before anyone better than them applies.[15]

Because of the urgency created by worker uncertainty about the employer's evaluation process, job openings get all the applicants they will ever receive quite quickly. The modal applicant in some categories of work comes less than 24 hours after posting (see Horton (2017a) on this point). The upshot of this incentive for speed is that a worker that receives

[15] This same argument explains why it would be difficult to show empirically that early applicants are more likely to be hired. When employers monitor the flow of applicants, hiring the first above their reservation value, it is worthwhile from the worker's perspective to apply sooner rather than later. However, now, the worker hired is always the last one that applied!

an invitation when they happen to be online is more likely to respond immediately and apply; a worker that receives an invitation when they are offline is more likely to not apply at all, as they will only learn about the opportunity some number of hours or days in the future, when the value of applying has diminished. This difference in the probability of applying, caused by time-zone differences, is the identification strategy.

To construct this instrument, I need an estimate of the probability that a worker from a particular country is online at a particular hour. A perfect indicator that a particular worker is online and using the platform at a given time is that they sent an application at that time; job applications are invariably sent by workers who are awake and using the platform. The top panel of Figure 3 shows the distribution of applications sent in each of the 24 hours of the day, as measured in the Pacific timezone (PT), for the top platform worker countries, namely Bangladesh, India, Philippines, Russia, Ukraine and the United Kingdom. The data for this figure is pooled over the entire history of platform job applications through March 10th, 2015.[16]

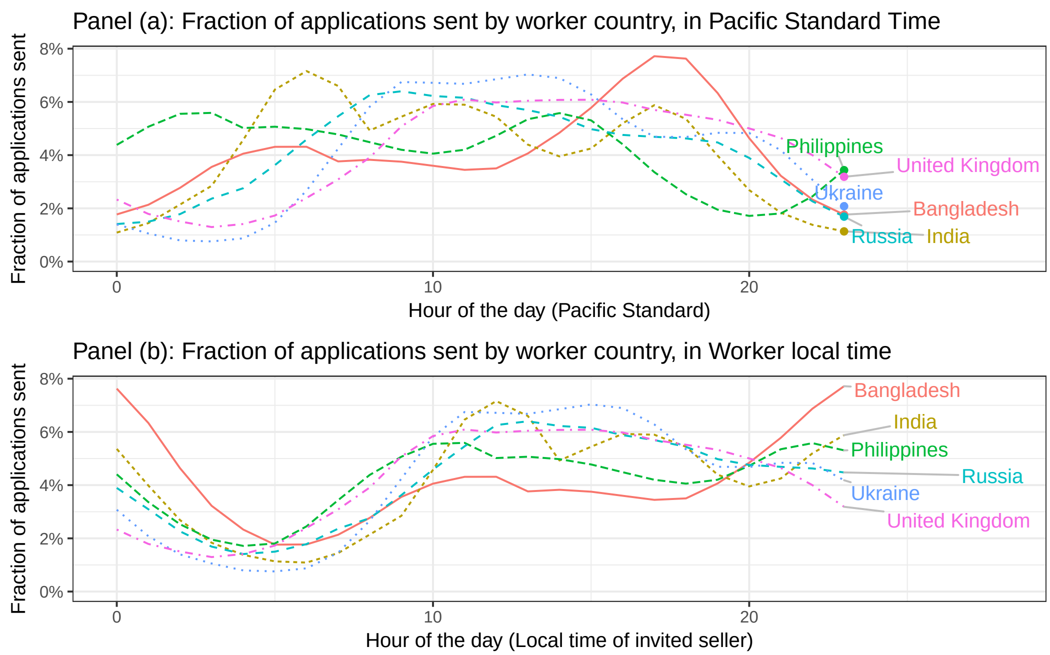
In this top panel, the lines are quite jumbled and interwoven with each other, but when we shift the data so that the fraction of applications is reported in each worker's local time, the pattern becomes quite clear: all workers show the same basic activity cycle, with on-site activity reaching a peak around 2pm or 3pm (14:00 on the 24 hour clock) and then trailing off in the evening, reaching a trough at around 5am. Figure 3 illustrates that despite the global nature of the market, workers "keep" their home country hours. As such, when an invitation is sent to a worker, the probability that worker is online at that moment and using the platform varies, and so his or her probability of applying varies, giving us a first stage for a two stage least squares estimation (2SLS).

To give an example of how the instrument is calculated, suppose an employer in New York City posts a web design job at 2pm local time on March 12th. If she were to invite a worker from Manila, the invitation would arrive at 2am; if the invited worker was in Moscow, it would arrive at 9pm; if Bangalore, then 11pm. If we use Figure 3 as a proxy for the probability of whether a worker from a particular country is online at a particular hour, then in descending order of being online, we have Russia, India and then the Philippines. Of course, we might have a Filipino night owl and a Russian early riser with an 8pm bed time, but on average, we would expect the employer to catch more Russians online than Filipinos.

The precise measure I use for the instrument comes directly from Figure 3: the instrumental variable associated with the acceptance decision for an invitation made by an employer at hour h (24 hour clock) to a worker from country l is the fraction of all job applications by workers from country l sent at hour h (in the employer's time zone). For example, an invitation by our hypothetical NYC-based employer sent at 2pm local time would be received by a worker in the Philippines at 2am; of all applications sent by Filipino workers on the platform, a little bit less than 3% are sent during that 2am-3am period. This 3% value would be the instrument for acceptance for this particular observation. If the Filipino worker had instead received the invitation at 10am local time (which is the peak of

[16]While it may be tempting to reduce the sample of applications to a smaller window covering the period of the IV analysis, this would likely be a bad trade in the bias/variance trade-off. The goal in constructing the instrument is to get the most accurate possible estimate of a worker's propensity to be online and thus respond to an invitation based on the time of day. For some worker/hour bins, the data is sparse and restricting the sample makes the problem worse. Restricting the sample might seemingly reduce bias—suppose the propensity to be online has changed over time—but as we are trying to capture something that is a general feature of workers in that country (i.e., their patterns of work and sleep), I doubt the bias reduction would be great, while the cost in increased sampling variation would be large.

Figure 3: Distribution of job applications sent by hour (0-23) of the day and applying worker country



Notes: Both plots show the fraction of applications sent each hour by worker country. Panel (a) shows the hour in terms of Pacific Standard and Panel (b) is the hour of the day in the modal timezone of the worker country.

activity), the instrument value would be nearly 6%. Note that each worker from the same country has the same instrument value for a given invitation hour.

Following Angrist and Pischke (2008), I formulate the IV estimation problem as a 2SLS regression using the linear probability model. Given that our interest is in marginal effects rather than prediction—and that our sample is already highly selected to have the "right" properties vis-a-vis the IV—the simplicity of using the linear probability model in this case makes it advantageous over an IV probit or logit model, particularly since I will also report a regression with employer-specific fixed effects.

A worker i is invited to a job opening, j. Let $l(i)$ index the worker's country. This invitation occurs at hour $H(i,j)$ in the worker's timezone, but at hour $H(i,j)+\Delta\mathrm{TZ}(l(i),j)$ in PT, where $\Delta\mathrm{TZ}(l(i),j)$ is the offset. Let $z_{H(i,j),l(i)}$ be the value of the instrument, which depends on the local arrival hour and the country of the worker. The two stages are:

$$a_{ij} = \gamma_0 + \gamma_1 z_{H(i,j),l(i)} + \gamma^{\text{Time}}_{H(i,j)+\Delta\mathrm{TZ}(i,j)} + \gamma^{\text{Loc}}_{l(i)} + \xi_{ij} \quad \text{(IV First stage)} \tag{5}$$

$$\mathbf{1}\{Y_{ij} > 0\} = \beta_0 + \beta_1 a_{ij} + \beta^{\text{Time}}_{H(i,j)+\Delta\mathrm{TZ}(i,j)} + \beta^{\text{Loc}}_{l(i)} + \epsilon_{ij} \quad \text{(IV Second stage)}, \tag{6}$$

where β^{Time} and γ^{Time} are fixed effects for the hour that the invitation was sent (in PT) and β^{Loc} and γ^{Loc} are fixed effects for the invited-worker country.

The exclusion restriction assumption is the following: conditional upon the hour of day and the country of the invited worker, the instrument is independent of whether the job opening fills or not. With the inclusion of country- and hour-specific fixed effects, this exclusion restriction assumption would still hold even if employers that recruit at different times differ systematically from each other (at an hour level of granularity), or if employers have worker country preferences that are correlated with the probability of the job being filled. What would violate the exclusion restriction assumption is if some employers send their invitations to workers they think are more likely to respond quickly because they are online, based on the invited worker's country. Other non-country signals of capacity that employers could condition on do not invalidate the instrument. This conditioning-on-the-IV hypothesis is somewhat testable, and I will present evidence below that employers do not learn to invite workers with higher values of the instrument. However, I will first simply present the 2SLS results.

5.4 Instrumental variables estimates of the effects of recruited worker invitation acceptance on match formation

In Column (1), Table 6, the first stage for the instrumental variable regression is shown (Equation 5), with the dependent variable being the invitation acceptance indicator, regressed on the instrument. In this regression—and all regressions in the table—standard errors are clustered at the level of the individual worker. The F-statistic for the first stage regression is 78.99, which implies a very strong instrument (Bound et al., 1995).

Note that although the coefficient on the instrument is highly significant, the effect is not absolutely large: the range of the instrument (from maximum to minimum value) is only about 0.07 and so the first stage coefficient means that from peak to trough, acceptance probabilities only vary by about 5 percentage points. This strong-but-not-large characterization is useful for my purposes, as the larger the effect, the less plausible it is that employers do not learn to condition on the instrument directly, which would violate the exclusion restriction.

In Column (2) of Table 6, the 2SLS estimate shows that a worker acceptance causes an increase in whether any worker was hired at all. Presumably, this effect is largely driven by the firm hiring the invited worker if he or she accepts, but not hiring anyone else if the worker rejects, consistent with Table 4. The 2SLS effect is substantially larger than the OLS estimate, though as expected, the 2SLS estimate is considerably less precise; the 95% CI for the 2SLS estimate overlaps the 95% CI for the OLS estimate.

Although the IV is valid even with job- or employer-specific omitted variables, we can control for the identity of the inviting employer. Column (3) reports the same regression as in Column (2), but with employer-specific fixed effects. The point estimate is similar to the Column (2) estimate, but the estimate is far less precise, which is to be expected given that many of the employers only send one invitation, leaving much less identifying variation.

Table 6: Effect of an invitation acceptance on the probability a job opening is filled

	Dependent variable:		
	Accepted $a_{ij} = 1$	$\mathbf{1}\{Y_j > 0\}$	
	1st Stage	2SLS	2SLS
	(1)	(2)	(3)
Instrument, z_{ij}	1.400^{***}		
	(0.165)		
Invite Accepted, $a_j = 1$		0.673^{***}	0.583
		(0.132)	(0.374)
Constant	0.577^{***}	0.034	
	(0.014)	(0.083)	
Employer FE?	No	No	Yes
Observations	57,250	57,250	57,250
R^2	0.037	-0.218	0.712
Adjusted R^2	0.037	-0.218	-0.014
Residual Std. Error	0.489 (df = 57221)	0.549 (df = 57221)	0.500 (df = 16251)

Notes: This table reports the results of an IV analysis of the effects of a recruited worker accepting an employer's recruiting invitation on the probability that the job opening is filled. Column (1) is the first stage of the 2SLS regression and Column (2) is the 2SLS estimate without employer fixed effects, while Column (3) includes employer fixed effects. Y_j is the amount of money spent against the opening and a_{ij} is an indicator for whether the invited worker i accepted the invitation. The instrument z_{ij} is the fraction of applications from i's country that are sent at the hour when the actual recruiting invitation was sent. This instrument is essentially a country-level proxy for the probability that the invited worker was online when they were invited to apply. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

The 2SLS estimate of the effect of rejections is 0.67, which is quite high. Recall from the OLS estimate, the fill rate following a rejection is about 0.36, whereas following an acceptance it is about $0.36 + 0.16 \approx 0.52$. If we imagine the most extreme case, where the true effect of an acceptance is to bring the fill rate to 100%, then the maximum possible treatment effect is $1.00 - 0.36 \approx 0.64$. The 95% CI for the 2SLS estimate is $[0.413, 0.933]$, and so clearly a substantial portion of the interval is impossible, and furthermore, a 100% fill conditional upon acceptance is improbable. As such, it seems likely that the 2SLS estimate is a high-side estimate, and the true value is closer to the OLS estimate, albeit by some unknown amount. However, as it is imprudent to search for alternative specifications to get point estimates that more closely approximate the researcher's prior, I report the 2SLS estimate with the caveat that it is likely an over-estimate.

The direction of the OLS bias relative to the 2SLS estimate, assuming it is not due to sampling variation, offers some insight into the nature of the recruiting process. If the main source of omitted variables bias was job-specific attributes, we would expect the OLS estimate to overstate the true effect of an acceptance. For example, if invitations to "better" job openings are more likely to be accepted, and these better job openings are inherently more likely to filled, the 2SLS estimate of the effect of an acceptance would be closer to zero (and actually be zero if the OLS result was purely due to selection).

If the main source of omitted variables bias was worker-specific attributes, the relative size of the 2SLS estimate would depend on whether those worker-specific attributes made a hire more or less likely, conditional upon an acceptance. For example, suppose invitations are sent to "better" workers who are less likely to accept because they have relatively less capacity—but that if they do accept, these better workers are more likely to be hired, as the employer feels fortunate to get such a good applicant. In this case, the acceptances induced by the instrument would come from relatively better workers, making it more likely that the job opening will be filled. This scenario is consistent with the Table 6 finding that the 2SLS effect from an acceptance is larger than the OLS estimate. See Appendix B for this argument presented more formally.

5.5 Effects of invitation acceptances on other outcomes

In addition to whether a hire is made, I can also examine whether the employer interviewed any non-invited applicants, hired some other worker, or engaged in compensatory recruiting following a rejection or acceptance decision. Detecting changes in these other outcomes potentially increases our confidence in the validity of the instrumental variables approach.

Column (1) of Table 7 reports a 2SLS regression where the outcome is whether the employer interviewed any non-recruited, organic applicants. We can see that when an invited worker accepts, the employer is far less likely to interview any organic applicants. This is consistent with the recruited worker being the employer's first choice and organic applicants serving as less preferred alternatives. Further making this point, in Column (2) the outcome is whether the employer hired a non-recruited applicant. As expected, the effect is negative, and highly significant. In Column (3), the outcome is whether the employer sent any "late" recruiting invitations i.e., those sent an hour or more after the posting of the job opening. As the sign on the coefficient is negative, there is some evidence of compensatory recruiting. However, the effect is imprecisely estimated and not conventionally significant. Adding employer fixed effects to this analysis increases the standard errors, but the point estimates are broadly similar—see Appendix C for this analysis.

The results from Table 7 show that when rejected, employers adjust in comprehensible ways: they rely on organic applicants and (perhaps) recruit again. However, a rejection still lowers hiring overall, as we saw in Table 6, and so the employers adjustments' are not fully effective in making up for the initial rejection.

5.6 Testing the exclusion restriction

As the 2SLS model is exactly identified, I have no statistical test for the validity of the instrument. However, I can test whether employers learn to invite more-likely-to-be-awake

Table 7: Effects of an invitation being accepted on organic candidate interviewing and hiring, as well as compensatory recruiting

	Dependent variable:		
	Non-recruit interviewed? 2SLS	Non-recruit hired? 2SLS	"Late" recruiting?"
	(1)	(2)	(3)
Invite Accepted, $a_i = 1$	-0.595^{***}	-0.239^{**}	-0.135
	(0.133)	(0.105)	(0.095)
Constant	0.857^{***}	0.385^{***}	0.256^{***}
	(0.084)	(0.066)	(0.060)
Observations	57,250	57,250	57,250
Adjusted R^2	-0.209	0.014	-0.030
Residual Std. Error (df = 57221)	0.550	0.430	0.394

Notes: This table reports the results of 2SLS regressions of the effects of a recruited worker accepting an employer's recruiting invitation on the probability the employer interviews any non-recruited organic applicants, in Column (1), or hires such a worker, in Column (2). In Column (3) the outcome is whether the employer engages in "late" recruiting, defined as recruiting 24 hours after the initial invitations, All regressions are 2SLS regressions using the "fraction of applications" instrument. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : * * *$.

workers over time, as they post more job openings. If employers are learning, then the exclusion restriction would be invalid. However, a caveat is that even if there is no evidence of learning, it does not prove the exclusion restriction is met—perhaps all employers already "know" what to do, however implausible that may be, given how small the effect is. That caveat aside, Column (1) of Table 8 reports a regression of the instrument on the log of $\textsc{PriorInvites}_{jt}$, the employer's count of previous invitations by that employer in the sample at time t:

$$z_{jt} = \beta \cdot \log(\textsc{PriorInvites}_{jt}) + \delta_j + h_{jt} + \epsilon_{jt}, \tag{7}$$

where δ_j is an employer-specific fixed effect, h_{jt} is a fixed effect for the hour the invitation was sent, and z_{jt} is the instrument value. The sample is the same sample as used in the instrumental variables analysis.

We can see from the Column (1) regression that the coefficient on the prior invitations measure is essentially a precisely estimated zero, with the "wrong" sign, implying there is no evidence that employers learn to seek out workers with higher values of the instrument. In Column (2), I report the same regression as in Column (1), but instead use the count of prior invitations, while in Column (3) I add a squared prior invitations term. In both regressions, the same result holds—there is no evidence employers learn to invite "awake" workers. To reiterate, these regressions do not prove the exclusion restriction holds, but they are suggestive that it holds.

6 Platform remedies to the worker capacity problem

In response to earlier versions of this analysis, the platform took steps to give recruiting employers more information about worker capacity. The most important step was the

Table 8: Association between employer prior recruiting experience and the instrument value for the invited worker

	Dependent variable:		
	Value of instrument for invited worker:		
	(1)	(2)	(3)
Log PriorInvites	−0.00004		
	(0.0003)		
PriorInvites		−0.00002	−0.00002
		(0.0001)	(0.0001)
PriorInvites^2			−0.00000
			(0.00000)
Employer Effect	Yes	Yes	Yes
Hour-of-day Effect	Yes	Yes	Yes
Observations	57,250	57,250	57,250
R^2	0.850	0.850	0.850
Adjusted R^2	0.471	0.471	0.471
Residual Std. Error	0.012 (df = 16255)	0.012 (df = 16255)	0.012 (df = 16254)

Notes: This table reports regressions where the dependent variable is the value of the invited worker's instrument. All regressions include both hour-of-day and employer fixed effects. The key independent variable is the count of invitations the employer made prior to the current invitation, and various transformations of this variable. Standard errors are clustered the level of the individual employer. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

creation of a new signaling mechanism that allowed workers to publicly declare their capacity on their profiles. The hope for the feature was that workers would change their signal as their actual capacity waxed and waned, and that employers in turn would condition on this signal when deciding which workers to invite.

The capacity signaling feature allowed workers to put one of three messages on their profiles about their "availability": (1) "More than 30 hrs/week," (2) "Less than 30 hrs/week," (3) "As Needed - Open to Offers." The settings were presented in order in the interface to imply that the third option would signal the least amount of capacity. However, as it is somewhat ambiguous as to the relative implied capacity of (2) versus (3), I turn the signal data into a binary indicator for signaling (1), the highest capacity, which I label "Full Time." For conciseness, I refer to the other two signals as "Part Time" and "As Needed," respectively. Workers were free to change the setting whenever they wished, but were prompted to do so after events the platform inferred were likely to indicate a change in capacity.

I use changes in worker availability status settings to obtain within-worker estimates of the effects of the displayed signal on recruiting and subsequent outcomes. I first examine whether workers changing their signal to "Full Time" receive more invitations. Then I examine how workers receiving invitations when in a "Full Time" respond with respect to acceptance and wage bidding. Finally, I examine whether those workers that apply while in a "Full Time" status are more likely to be hired. For each of these outcomes, I use a variety of different regression specifications to explore possible alternative explanations for the results.

Before discussing the results, it is useful to consider the interpretation of any effects that are seemingly due to the "Full Time" signal. The first outcome of interest is the count of invitations a worker receives in some period. This outcome is largely determined by employer behavior, as employers condition on the signal. From an employer's perspective, the change in a worker's signal can reasonably be thought of as exogenous changes in a worker's attributes.[17] In contrast to the count of invitations received by a worker, other outcomes are determined by the individual worker, such as their invitation acceptance and their wage bids. Changes in these worker behavior outcomes are clearly not "caused" by the signal, at least not directly. If we imagine an experiment in which the platform randomly altered a worker's signal, we would not expect workers to change their behavior. In contrast, when workers endogenously change their signal themselves, we might expect their behavior to change because their signal choice was not random, but rather was a response to a change in some aspect of the worker' situation, such as his or her capacity to take on more work, outside options, preferences and so on. In a sense, any effects of the "Full Time" signal on worker behavior should be interpreted as a consequence of some other change that is being truthfully reported.

6.1 Data construction and description

The data used for the analysis of the signal is constructed from the records of workers changing their signals. From these changes, I construct worker "windows" of time in which his or her signal was unchanged. I use data from April 24, 2014 to May 27, 2016. I only use windows that are at least a day in length, as shorter windows are likely to be workers experimenting with the feature. Only 49,378 workers actually changed their signal at least once during the period covered by the data. A much larger number of workers used the

[17]As far as I am aware, the platform did not use the availability settings made by workers to determine ranking results in search.

signal only once, setting it and then not changing it during the period covered by the data. The 49,378 workers with changes are responsible for a total of 158,006 windows. The most common status window type is “Full Time,” which makes up 45% of all windows, whereas only 33% and 22% of windows are “Part Time” and “As Needed,” respectively.[18] Of workers only setting their signal once, a much larger fraction set their signal to “Full Time.”

6.2 Effects of the worker capacity signal on employer behavior

The first question of interest is whether employers conditioned their recruiting decisions on the signal observed. Using worker windows as the unit of observation, I regress a measure of employer recruiting intensity on the “Full Time” indicator. The outcome measure is $\text{InvitesNormalized}_{iw}$, which is the count of invitations received by worker i in their window w, divided by the duration of the window (measured in quarters of a year). The regression is

$$\text{InvitesNormalized}_{iw} = \beta \cdot \text{FullTime}_{iw} + \gamma_i + \text{Week}_{t(i,w)} + \epsilon_{iw}, \tag{8}$$

where FullTime_{iw} is the “Full Time” indicator, γ_i is a workers-specific fixed effect, and $\text{Week}_{t(w,j)}$ is a fixed effect for the week the window began.

The far left panel of Figure 4, labeled labeled “Normalized Invites (per qtr),” reports coefficients on “Full Time” for several regressions based on Equation 8. Four estimates are reported, using regressions with: (1) with neither worker- nor time-specific fixed effects, labeled “No FE,” (2) week fixed effects only, labeled “Week,” (3) both week and worker fixed effects, labeled “Worker+Week,” and corresponding to Equation 8, and (4) worker-quarter specific fixed effects and week fixed effects, labeled “(Worker-Qtr)+Week,”

In the “Worker+Week” regression, we can see that a worker switching to the “Full Time” status received about 2 more invitations per quarter. This coefficient is identified only from workers changing their status, eliminating concerns that certain kinds of workers likely to give the “Full Time” signal also receive more invitations. With the inclusion of worker-quarter fixed effects, in “(Worker-Qtr)+Week,” the coefficient on “Full Time” actually increases, suggesting that non-signal related changing worker attributes are not driving the increased employer interest. However, this estimate is potentially problematic, in the sense that they can absorb some of the useful variation in signal-setting.[19] The “Week” and “No FE” regressions yield somewhat smaller “Full Time” effects relative to the “Worker+Week” estimate, which is consistent with less sought-after workers more frequently being in a “Full Time” status than more sought-after workers.

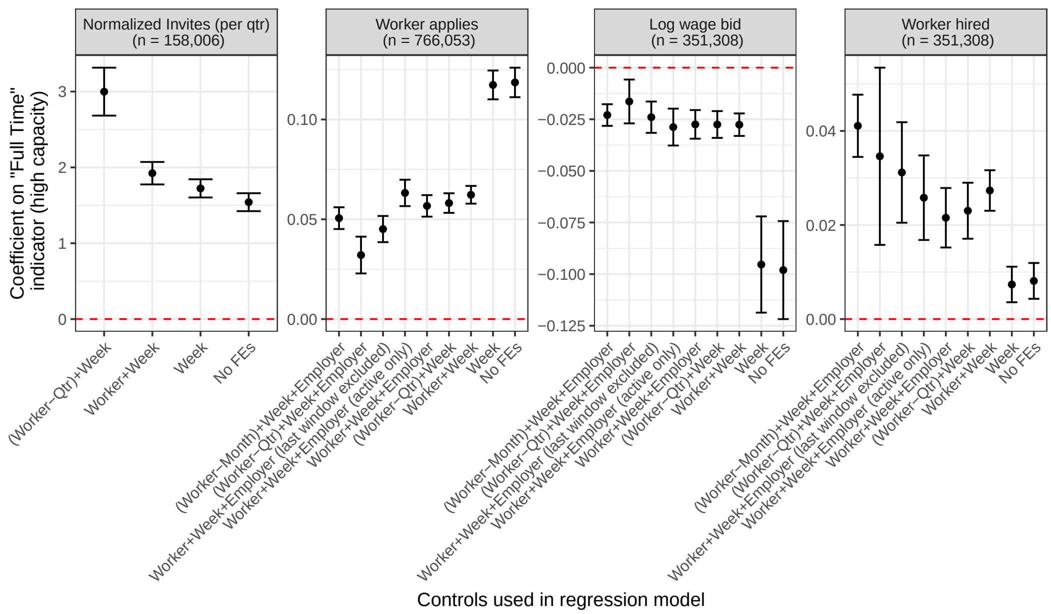
6.3 Worker response to recruiting invitations, by capacity signal status

I now turn to the question of how workers responded to invitations. For this analysis, I switch to using individual invitations as the unit of analysis, with invitations labeled with

[18]The feature was introduced earlier, but the platform made several changes to the language of the status settings and the user interface for making status changes. After April 24, 2014 the design stabilized and was available to 100% of workers.

[19]As an extreme illustration of the problem, if workers set their signal weekly, the inclusion of worker-week fixed effects would mechanically lead to the conclusion that the signal had no effect. For this reason, using larger time periods is attractive (so that there is signal variation with in the period), though this has the limitation that these effects are less successful in dealing with the problem they are designed to confront (namely time-varying worker characteristics).

Figure 4: Effects of a worker sending a high capacity signal, or "Full Time," signal



Notes: This figure shows a collection of regression estimates of the effects of a worker setting a high capacity signal to the marketplace. In each panel, the reported value is the coefficient on the "Full Time" status indicator. A 95% CI is shown, with standard errors clustered at the worker level. The various specifications are explained in the text.

the signal the worker was giving when receiving that invitation. For these outcomes, it is possible to include a wider and richer collection of regression specifications. In each panel of Figure 4, I report, from right to left: (1) "No FE," (2) "Week,"which adds a week-specific effect to specification 1, (3) "Worker+Week,"which adds a worker-specific effect to specification 2, (4)"(Worker-Qtr)+Week," which replaces the worker-specific effect in specification 3 with worker-quarter specific effects, (5) "Worker+Week+Employer," which adds an employer-specific effect to specification 3, (6) "Worker+Week+Employer (active only)," which is the same specification as 5, but with the sample restricted to observations occurring during windows in which the worker sent at least one organic application, (7) "Worker+Week+Employer (last window excluded)," which is the same specification as 5, but the sample excludes the last window for each worker, (8) "(Worker-Qtr)+Week+Employer," which adds an employer-specific effect to specification 4, and (9) "(Worker-Month)+Week+Employer," which adds a worker-month specific effect and an employer specific effect to specification 2.

These different specifications address various hypotheses about why the "Full Time" signal might be correlated with various outcomes. As with the invitations received outcome, specifications that include worker-time specific effects are intended to address the possibility of time-varying worker attributes that would not be captured by a single worker-specific effect. Specifications that have a restricted sample are designed to address the possibility that workers leaving the platform or otherwise mechanically unavailable might leave their signal in a "Full Time" status, leading to a biased estimate. Specifications that include employer-specific effects are designed to address the possibility that different kinds of employers might seek out certain kinds of workers.

In the "Worker applies" panel of Figure 4 regressions, results are reported where the outcome is whether the worker actually applied, with the sample restricted to invitations where the worker responded. All regression specifications show a strong positive effect, meaning that workers signaling a "Full Time" status were more likely to apply in response to an invitation. This positive signaling effect holds in the most credible estimates, which are those that include worker-specific effects, namely "Worker+Week" and "Worker+Week+Employer." Replacing the worker-specific fixed effect with a worker-week and worker-quarter effects somewhat lowers the point estimates. However, the effects are still far away from zero.

The "No FE" and "Week" estimates do not differ much from each other, suggesting the particular calendar time is not highly relevant in the response decision. The "No FE" and "Week" regressions show considerably larger effects than the "Worker+Week" regression; the "within" estimates that include an employer-specific effect are about half the size of the two "between" estimates. This is consistent with the most selective workers being less likely to use the "Full Time" signal, which highlights the importance of the within-worker estimation.

Including an employer-specific effect, "Worker+Week+Employer" has almost no effect on the point estimates relative to "Worker+Week." This is not evidence that employers do not matter, but rather simply that whatever attributes workers care about do not appear to vary systematically with however the employers condition their recruiting on the signal. Including only active periods or excluding the last period does little to the point estimate.

When the worker applies, they submit an hourly wage bid. The panel labeled "Log wage bid" of Figure 4 reports regressions where the outcome is the log wage bid, with the sample restricted to invitations where the invited worker applied. In all regressions, workers propose a lower hourly rate when they have given the "Full Time" signal. The "Worker+Week" estimate implies workers bid about 2.5% lower. This "discount" is smaller than what is found in the "No FE" and "Week" regressions. Adding employer-specific fixed effects does not change the point estimate. Including only active periods or excluding the last period does nothing to the point estimate. Adding worker-specific time effects does not change the point estimate. In short, there is strong evidence that workers bid lower when they are sending the "Full Time" signal.

6.4 Employer response to applications, by capacity signal status

We now return back to an outcome that is at least partially out of the worker's control, which is whether or not they were ultimately hired. The rightmost panel of Figure 4, labeled "Worker hired," reports regressions where the outcome is an indicator for whether the applying worker was hired. The sample is restricted to workers that applied in response to the invitation. In the "Worker+Week" regression, we can see that the applying worker is about 3 percentage points more likely to be hired. Including only active periods or excluding the last period does nothing to the point estimate. Adding worker-specific time effects makes the point estimate somewhat larger. Although it is temping to control for the wage bid in this regression, this would be a problematic regression because the wage bid should be thought of as an outcome.[20] Given how much lower workers bid when in a "Full Time" status, it seems likely that the lower wage bid is largely driving the hiring result.

6.5 Discussion of capacity signaling feature

A key finding from the availability signaling analysis is that employers condition their invitation on the signal, sending more invitations to workers signaling more capacity. This conditioning is understandable when we consider how workers respond to invitations contingent upon their signal: when signaling "Full Time" they not only are more likely to apply, they also bid less. This provides the rationale for why employers are willing to treat the signal as informative. In essence, the platform engineered an informative signaling equilibrium.

The capacity signaling feature—despite having zero marginal cost to the platform—had several welcome effects. It shifted invitations to workers who were more likely to accept the invitation and, at that moment, seemed to have had lower costs. Given the results from the IV analysis, it is likely that these increased acceptances by workers increased the total fraction of jobs filled, meaning the signaling feature had both a price effect and a quantity effect.

That these improvements were possible strongly suggests that employers were not fully informed about worker capacity before the feature was introduced. All the evidence suggests workers used the feature to signal truthfully, as their behavior is consistent with what they publicly signaled. However, an important caveat—but also an implied opportunity—is that most workers did not change their status once it was set. It is unclear whether this was because they truly had no change in their capacity, or because they were reluctant to reveal their true capacity.

The signaling feature makes clear that high capacity workers tend to be in relatively low demand. We can see this with the often stark differences between the within-worker

[20] Even if employer hiring was only affected through the channel of changed wage bids, including that bid as a regressor would not guarantee that all of the effect would be loaded on that bid—it is possible that the "Full Time" indicator could still be non-zero and highly significant if the hiring process is not correctly specified.

estimates and the between-worker estimates for some outcomes, such as wage bidding and hiring. An implication of this pattern is that if employers with more information are encouraged to pursue higher capacity workers, they are also more likely to pursue relatively less experienced workers, perhaps with good equity consequences for the marketplaces, ala Pallais (2014).

The signaling feature results allow us to make some ballpark calculations about the likely efficiency effects of capacity uncertainty. In terms of price, in a competitive market, the nearly 3% discount offered by workers when they have more capacity implies that their costs are 3% lower, which means at least a 3% higher social surplus—it could be even more as hours-worked is endogenous with respect to the offered wage (Horton, 2017b). However, an important caveat is that some work is already done by workers with lots of capacity just by chance, so the realized improvement would be lower than 3%, even if all work could be done by workers with high capacity. Furthermore, to the extent employers have elastic demand, they will "naturally" avoid these relatively higher cost workers, partly mitigating any price effects. In terms of quantity, if we take the approximately 6% increase in invitation acceptance rates by high capacity workers times the increase in fill rate from an acceptance from the IV analysis, we get about a 3.5% increase in the quantity of transactions. If these marginal job openings generate as much surplus as the average job opening, then this 3.5% figure would be the increase. However, an important caveat is that many invitations already go to high capacity workers, and furthermore, the IV estimates comes from a sample of employers sending a single invitation, making it likely an over-estimate of the effects of an acceptance. Although for employers sending multiple invitations capacity uncertainty is less likely to lead to unfilled openings, the added cost of sending multiple invitations is also another cost of capacity uncertainty that is not quantified by this analysis but could be important.

7 Conclusion

The central conclusion of the paper is that employer uncertainty about worker capacity matters. Uncertainty affects the market in ways an efficiency-minded market designer would likely view as adverse—it prevents matches from being formed, and when they still are formed, raises costs. However, the paper shows that even relatively simple market design interventions by the platform can at least partially remedy the problem.

Although the intervention discussed here was effective, it is not a fully satisfactory solution. For one, workers had fairly weak incentives to use the signaling mechanism, and uptake was concentrated among a relatively small number workers, albeit those most likely to receive invitations. If the platform wanted to take a stronger hand, it could do more than simply provide signals. It would be possible, for example, to prevent the worker from receiving new invitations until all older invitations have been processed.[21] Or workers could have some non-monetary visibility or search-result prominence "budget" they could allocate dynamically, creating an opportunity cost to claiming full capacity.

If the platform was open to using money, it could have workers pay to appear higher in search results. Only high capacity workers would pay, and the platform could collect additional revenue. However, this method places a high burden on workers to manage their advertising "campaign," and it could potentially undermine the overall level of trust in the platform's actions if organic and paid results were not strongly delineated.

[21] Thanks to Randall Lewis for this suggestion.

An alternative approach that puts less burden on workers would be for the platform to fit predictive models of capacity, and then adjust search rankings in accordance with the platform's objectives. Somewhat ironically, the common industrial practice of fairly static "best-to-worst" search rankings would tend to exacerbate the problems identified in this paper, as the most frequently shown worker is also likely to be capacity constrained (by the very fact that they appear so often). This machine learning approach places a high burden on the platform, and to the extent workers know their own capacity, a worker-revealing mechanism is probably superior. However, different approaches could be complementary, with the worker-provided signal being just one input into the platform's decision about how "visibility" on the platform is allocated to workers.

References

Agrawal, Ajay K, John Horton, Nico Lacetera, and Elizabeth Lyons, "Digitization and the contract labor market: A research agenda," in Avi Goldfarb, Shane M. Greenstein, and Catherine E. Tucker, eds., *Economic Analays of the Digital Economy*, University of Chicago Press, 2015.

_ , Nicola Lacetera, and Elizabeth Lyons, "Does standardized information in online markets disproportionately benefit job applicants from less developed countries?," *Journal of International Economics*, 2016, *203*, 1–12.

Allon, Gad, Achal Bassamboo, and Eren B. Çil, "Large-scale service marketplaces: The role of the moderating firm," *Management Science*, 2012, *58* (10), 1854–1872.

Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly harmless econometrics: An empiricist's companion*, Princeton University Press, 2008.

Armstrong, Mark, "Competition in two-sided markets," *The RAND Journal of Economics*, 2006, *37* (3), 668–691.

Arnosti, Nick, Johari Johari, and Yash Kanoria, "Managing congestion in decentralized matching markets," *Proceedings of the ACM Conference on Economics and Computation*, 2014.

Bajari, Patrick and Steven Tadelis, "Incentives versus transaction costs: A theory of procurement contracts," *The RAND Journal of Economics*, 2001, pp. 387–407.

Bound, John, David A Jaeger, and Regina M Baker, "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak," *Journal of the American Statistical Association*, 1995, *90* (430), 443–450.

Casadesus-Masanell, Ramon and Hanna Halaburda, "When does a platform create value by limiting choice?," *Journal of Economics & Management Strategy*, 2014, *23* (2), 464–482.

Dimoka, Angelika, Yili Hong, and Paul A Pavlou, "On product uncertainty in online markets: Theory and evidence," *MIS Quarterly*, 2012, *36* (2), 395–426.

Farronato, Chiara and Zoë Cullen, "Outsourcing tasks online: Matching supply and demand on peer-to-peer Internet platforms," *Working Paper*, 2015.

Fradkin, Andrey, "Search frictions and the design of online marketplaces," *Working Paper*, 2014.

Gefen, David and Erran Carmel, "Is the world really flat? A look at offshoring at an online programming marketplace," *MIS Quarterly*, 2008, pp. 367–384.

Goes, Paulo and Mingfeng Lin, "Does information really 'unravel'? Understanding factors that motivate sellers to seek third-party certifications in an online labor market," *Working Paper*, 2012.

Halaburda, Hanna and Mikolaj Piskorski, "When should a platform give people fewer choices and charge more for them?," *Antitrust Chronicle*, 2010, *7.*

Hall, Jonathan V and Alan B Krueger, "An analysis of the labor market for Uber's driver-partners in the United States," Technical Report, National Bureau of Economic Research 2016.

Hong, Yili and Paul A Pavlou, "Is the world truly "flat"? Empirical evidence from online labor markets," *Working Paper*, 2015.

_ , Chong Wang, and Paul A Pavlou, "Comparing open and sealed bid auctions: Evidence from online labor markets," *Information Systems Research*, 2015, *27* (1), 49–69.

Horton, John, "The effects of algorithmic labor market recommendations: Evidence from a field experiment," *Journal of Labor Economics*, 2017, *35* (2).

_ , "Price floors and employer preferences: Evidence from a minimum wage experiment," *Working paper*, 2017.

_ **and Joseph Golden**, "Reputation inflation: Evidence from an online marketplace," *Working Paper*, 2015.

_ **and Ramesh Johari**, "A what quality and what price?: Eliciting buyer preferences as a market design problem," *Working Paper*, 2013.

Lewis, Gregory and Albert Wang, "Who benefits from improved search in platform markets?," *Working Paper*, 2013.

Pallais, Amanda, "Inefficient hiring in entry-level labor markets," *American Economic Review*, 2014, *104* (11), 3565–3599.

Parker, Geoffery G. and Marshall W. Van Alstyne, "Two-sided network effects: A theory of information product design," *Management Science*, 2005, pp. 1494–1504.

Resnick, P., R. Zeckhauser, J. Swanson, and K. Lockwood, "The value of reputation on eBay: A controlled experiment," *Experimental Economics*, 2006, *9* (2), 79–101.

Rochet, Jean-Charles and Jean Tirole, "Platform competition in two-sided markets," *Journal of the European Economic Association*, 2003, *1* (4).

– **and** – , "Two-sided markets: a progress report," *The RAND Journal of Economics*, 2006, *37* (3), 645–667.

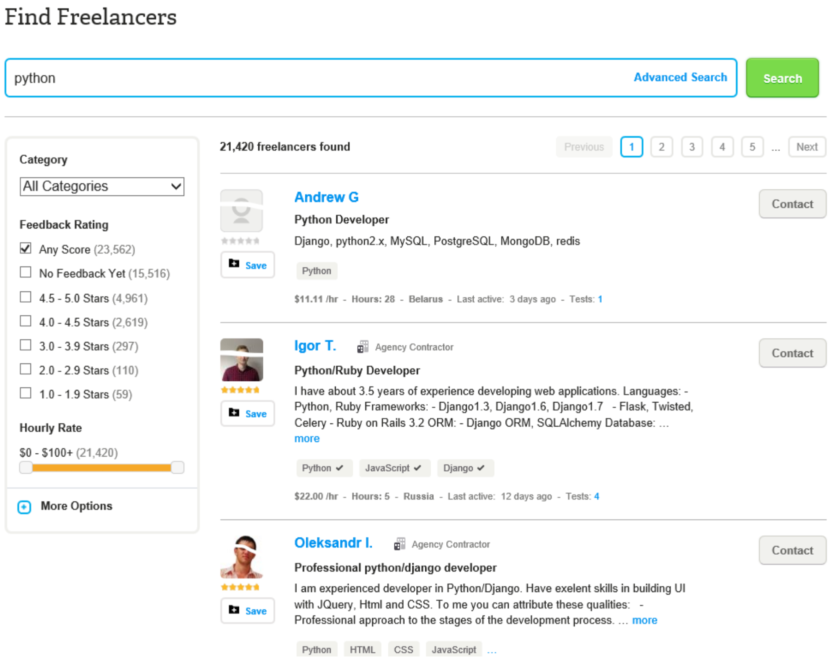
Snir, Eli M and Lorin M Hitt, "Costly bidding in online markets for IT services," *Management Science*, 2003, *49* (11), 1504–1520.

Stanton, Christopher T. and Catherine Thomas, "Landing the first job: The value of intermediaries in online hiring," *The Review of Economic Studies*, 2016, *83* (2), 810–854.

Tadelis, Steven and Florian Zettelmeyer, "Information disclosure as a matching mechanism: Theory and evidence from a field experiment," *The American Economic Review*, 2015, *105* (2), 886–905.

Yoganarasimhan, Hema, "The value of reputation in an online freelance marketplace," *Marketing Science*, 2013, *32* (6), 860–891.

Figure 5: Search results interface presented to an employer



Notes: This figure is a screenshot of the platform search interface take on February 9th, 2015 for the query "python," a popular computer programming language.

A Interfaces

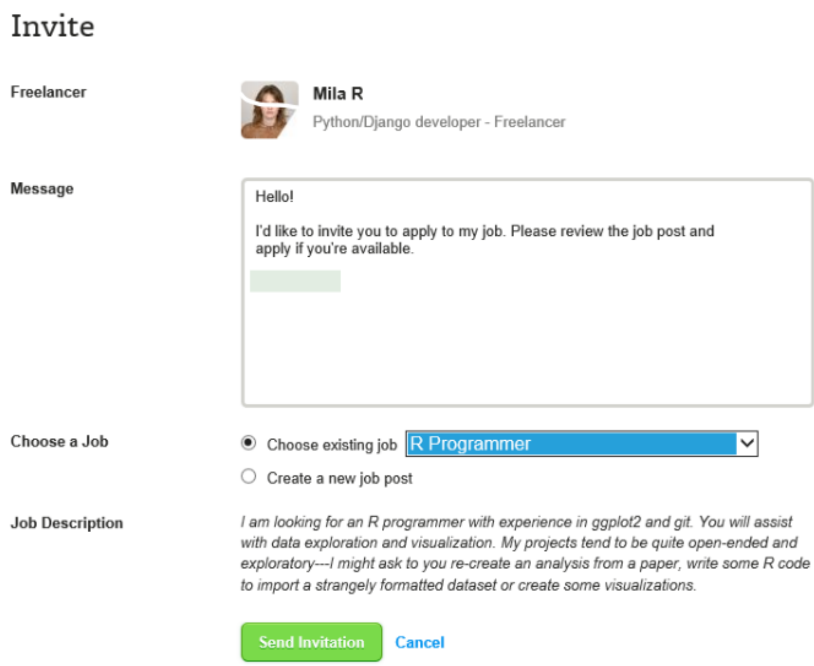
Figure 5 shows the worker search results displayed for a query of "python," a popular programming language. Figure 6 shows the interface employer use when inviting a worker to apply for a job opening. Figure 7 shows the interface shown to workers receiving a recruiting invitation from an employer. To "accept" the invitation, they go to the same job application interface as organic applicants.

B Marginal effects of acceptances

What direction should the bias be in the OLS regressions? The answer depends on the nature of the omitted variables. For example, suppose "better" workers, if they accept, are more likely to cause a job to be filled. And assume further that those better workers are less likely to accept an invitation. In this case, the OLS estimate of the effects of acceptance on hiring will be biased downward. In contrast, suppose workers do not differ in ability but jobs differ in quality, with "better" jobs being more likely to be filled. If workers condition their decision on job quality, the OLS estimate will be be biased upward.

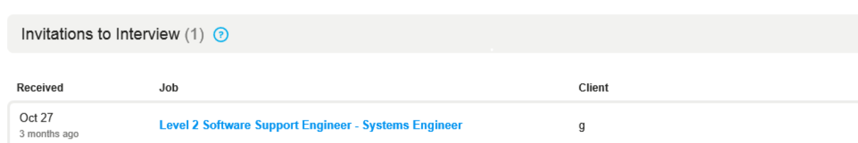
To see this selection argument more formally, let us consider the worker case. Suppose workers have ability a, and the marginal effect of an invitation acceptance of job filling probability is $\beta(a)$. Assume that for all a, $\beta(a) \geq 0$. Furthermore, assume that marginal effects are increasing in worker ability, i.e., $\beta'(a) > 0$.

Figure 6: Employer interface for contacting a worker with an invitation to apply to a job opening



Notes: This figure shows the view for an employer inviting a worker to a job opening.

Figure 7: Recruiting invitation "inbox" on the platform presented to workers



Notes: This figure shows the invitation "inbox" presented to a worker. Note that it shows the date the invitation was received, the title of the associated job and the client [employer] that sent the invitation. By clicking the title of the job, the worker can learn more about both the job opening and the employer.

A worker's probability of accepting an invitation is $\Pr(\text{Accept}|a)$. We can define the probability of observing in our data a worker with ability a accepting an invitation as

$$p_{obs}(a) = \frac{\Pr(\text{Accept}|a)}{\int_0^1 \Pr(\text{Accept}|a)da}.$$

Let $P_{obs}(a)$ be the cdf of p_{obs}. The expected marginal effect of an acceptance on the hire probability is thus

$$\begin{aligned}
\mathbf{E}[\beta|p_{obs}] &= \int_0^1 \beta(a)p_{obs}(a)da \\
&= \beta(a)P_{obs}(a)\big|_0^1 - \int_0^1 P_{obs}(a)\beta'(a)da \\
&= (\beta(1) - \beta(0)) - \int_0^1 P_{obs}(a)\beta'(a)da, \qquad (9)
\end{aligned}$$

with integration-by-parts. Now consider the marginal effect we would observe if all workers, regardless of ability, were equally likely to be accepted. The worker selection probability (i.e., probability they show up in the data) is $p_{act}(a) = 1$, and $P_{act}(a) = a$. From Equation 9, we can see the critical role played by the cdf of the selection vis-a-vis ability. If workers with higher ability are more likely to reject an invitation, for all a, $P_{act}(a) < P_{obs}(a)$ (i.e., P_{act} first order stochastically dominates P_{obs}), then $\mathbf{E}[\beta|p_{act}] > \mathbf{E}[\beta|p_{obs}]$. The average effect of an invitation acceptance, when all workers are sampled equally, is greater than the average effect when those accepting are adversely selected. In the data, we see that "better" workers are less likely to accept, and so we would expect that the OLS estimate to be biased downward relative to the 2SLS estimate. If the reverse were true, the bias would be in the other direction. A similar logic would hold for job quality—if the higher quality job openings were more likely to be accepted, then OLS estimate over-state the effect of an acceptance on job filling probability.

C Auxiliary Analyses

Table 9 re-capitulates the results of Table 7, but with the inclusion of employer-specific fixed effects.

Table 9: Effects of an invitation being accepted on organic candidate interviewing and hiring, as well as compensatory recruiting

	Dependent variable:		
	Non-recruit interviewed? 2SLS	Non-recruit hired? 2SLS	"Late" recruiting?"
	(1)	(2)	(3)
Invite Accepted, $a_i = 1$	-0.829^{*}	-0.255	-0.339
	(0.432)	(0.306)	(0.316)
Observations	57,250	57,250	57,250
Adjusted R^2	-0.272	0.095	-0.152
Residual Std. Error (df = 16251)	0.564	0.412	0.416

Notes: This table reports the results of 2SLS regressions of the effects of a recruited worker accepting an employer's recruiting invitation on the probability the employer interviews any non-recruited organic applicants, in Column (1), or hires such a worker, in Column (2). In Column (3) the outcome is whether the employer engages in "late" recruiting, defined as recruiting 24 hours after the initial invitations, All regressions are 2SLS regressions using the "fraction of applications" instrument. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.