Job-Seekers Send Too Many Applications: Experimental Evidence and a Partial Solution

John J. Horton
MIT Sloan & NBER

Shoshana Vasserman
Stanford GSB*

February 5, 2021

Abstract

As job-seekers internalize neither the full benefits or costs of their application decisions, job openings do not necessarily obtain the socially efficient number of applications. Using a field experiment conducted in an online labor market, we find that some job openings receive far too many applications, but that a simple intervention can improve the situation. A treated group of job openings faced a soft cap on applicant counts. However, employers could easily opt out by literally clicking a single button. This tiny imposed cost on the demand side had large effects on the supply side, reducing the number of applicants to treated jobs by 11%—with even larger reductions in jobs where additional applicants were likely to be inframarginal. This reduction in applicant counts had no discernible effect on the probability a hire was made, or in the quality of the subsequent match. This kind of intervention is easy to implement by any online marketplace or job board and has attractive properties, saving job-seekers effort while still allowing employers with high marginal returns to more applicants to get them.

*Email: john.joseph.horton@gmail.com. Thanks to Adam Ozimek, Apostolos Filippas, Dan Walton, Philipp Kircher, and Ada Yerkes Horton for helpful comments and suggestions. Latest draft available at http://www.john-joseph-horton.com/papers/autopause.pdf. COUHS information available at http://www.john-joseph-horton.com/papers/couhs.pdf

1 Introduction

A social planner wants the marginal benefit of using some resource to equal the marginal cost. In the context of the labor market matching process, that valuable resource is the job-seeker's time. Clearly, effort is needed to form matches, but as job-seekers internalize neither the full benefits nor the costs of their application decisions, there is no economic reason to think jobs obtain the socially efficient number of applications in a decentralized market. And to the extent digitization of the search and matching process has dramatically lowered the cost of sending an additional application, we might suspect that there are frequently excess applications.

In this paper, we describe an experiment conducted in an online labor market that influenced the size of applicant pools faced by employers.[1] This was done by imposing a soft cap on the number of applicants that a job opening could receive, as well as limiting the duration of the window of time during which applications could be received: when a job opening received 50 applicants—or when 120 hours (5 days) had passed—no more applicants could apply unless the employer explicitly asked for more applicants. The intent of the intervention was to prevent job-seekers from applying to jobs where their application was likely to either be ignored or simply displace some other applicant, without preventing employers with high marginal returns to more applicants from obtaining them.

[1]We use the terms "employer," "worker" and "hire" to be consistent with the labor literature and not as a comment on the nature of the relationships created on the platform.

We find that the treatment caused a substantial reduction in application counts—about 4 fewer applicants applied on average, or an 11% reduction. However, reductions were largest for jobs that otherwise would have received large numbers of applicants—the quantile treatment effect at the 95th percentile is a reduction of 20 applicants.

Despite the reductions in applicant counts, the treatment did not reduce the probability a hire was made. About 41% of job openings were filled in both the treatment and the control group.[2] Firms denied the right "tail" of 50+ applicants or late-arriving applicants simply hired from their other applicants, with no discernible ill effect. It is not the case that later applicants are adversely selected and thus simply irrelevant—in the control, later-arriving applicants were still in the consideration set of employers.

There is no evidence that better or worse matches were made in the treatment group, as measured by the feedback given by the employer at the end of the contract or in hours-worked. If anything, employer satisfaction rose slightly in the treatment.

The lack of effects on hiring or match quality is seemingly surprising, but likely reflects the fact that price competition among workers "prices in" vertical differences among workers, leaving firms close to indifferent over applicants, as in Romer (1992). Because of this indifference, substitution among applicants is not very costly to employers.

[2] This fill rate is actually quite similar to the fill rate reported by Indeed from 2015. `http://press.indeed.com/wp-content/uploads/2015/01/Time-to-fill-jobs-in-the-US.pdf`.

Our claim is not that the applicant count does not matter—clearly going to 5 or 1 or even 0 would matter a great deal. Our claim instead is that for a substantial number of employers, the marginal benefit to more applicants seems to be less than the *de minimus* cost of pushing a single button. When search costs are already low, marginal applicants might simply not be worth very much, if anything. As it is, only about 7% of employers requested more applicants by pushing the button.

The treatment intervention likely saved job-seekers substantial time—more so than the percentage changes in job post applicant counts would seemingly imply. To see why the treatment has out-sized effects on job seekers, note that although relatively few job openings were affected by the 50 applicant cap (about 10%), these job openings are disproportionately important to job-seekers, as they attracted 43% of applications. This difference simply reflects the fact that a randomly selected application is more likely to be sent to a job with a high applicant count.[3]

Switching our data from job posts to applications to those job posts and using a within-worker analysis, we find that a worker applying to a treated job opening had a 17% increase in their probability of being hired. This increase in application success might raise or lower overall application intensity in equilibrium (Shimer, 2004)—job applications have become more valuable to the worker, but fewer need to be sent to secure a job, on average. However, regardless of the effect on application intensity or the application cost, the

[3]This is a version of the friendship paradox (Feld, 1991).

intervention would still improve worker welfare relative to the *status quo* via a simple envelope theorem argument.

To illustrate the wastefulness of decentralized job search implied by our results, consider the following simple example. Suppose an application costs the job-seeker c and the expected social surplus of a job post is $V(A)$, where A is the number of applicants. And suppose the job-seeker gets a fraction $\theta > 0$ of the job surplus if she is hired. If job-seekers think they are equally likely to be selected from among the applicants, they will apply until $\theta V(A)/A = c$, but a social planner would like $V'(A) = c$. Note that in the decentralized equilibrium, $\theta V(A) = Ac$.[4] That is, the entire hired worker pay-off is consumed by application costs. If θ is sufficiently small or $V'(A)$ is sufficiently large, the platform/social planner might of course welcome more applications—echoing the economic intuition of Hosios (1990). But the soft cap design of the experiment suggests $V'(50) \approx 0$, as most employers did not bother lifting the cap to obtain more applicants. And so while job-seekers might still want to play the congestion game and keep applying past 50, no social planner would want this game to continue.

This is the first experiment we are aware of where the number of applications to a job opening was experimentally reduced. The key contribution of the paper is to use this experimental variation to show that many job applicants are inframarginal in the decentralized labor market equilibrium.

[4]We are assuming none of the wastefulness is due to "ball and urn" matching frictions caused by workers being unable to condition on applicant counts (Gee, 2019; Bhole et al., 2021).

We also illustrate the crowd-out effect of other applicants in a particularly direct way, compared to the literature (Lalive et al., 2015). Our crowd-out results call into further question the equilibrium justification for job search assistance (Crépon et al., 2013; Marinescu, 2017).[5]

Paired with our contribution to the literature is a practical—albeit partial—solution that could be implemented by any computer-mediated labor matching marketplace. Market design interventions that save workers time or direct their applications to relatively under-subscribed openings could offer substantial welfare gains, even setting aside any employer benefits from more efficiently directed applications. The US non-institutional population *on average* spends about 15 hours a year on job search activities, which is about $75B per year in time value at the median US wage.[6]

The rest of the paper is organized as follow. Section 2 describes the experimental context. Section 3 explains the design and discusses internal and external validity. Section 4 presents the results. Section 5 concludes.

[5] Though there is evidence that more targeted recruiting assistance can be helpful without much crowd-out (Horton, 2017, 2019) and that interventions that have job-seekers consider a wider range of options could be beneficial, as in Belot et al. (2019). The lack of an *increase* in hiring in the treatment is evidence against the "choice overload" hypothesis (Iyengar and Lepper, 2000), which itself has been called into question (Scheibehenne et al., 2010).

[6] Using data from 2013 and assuming 252 working days per year—see `https://www.bls.gov/tus/current/work.htm`.

2 Empirical context

Our setting is a large online labor market. In this market, employers post job openings to which workers can typically apply without restriction. The kinds of offered work include tasks that can be done remotely, including programming, graphic design, data entry, translation, writing and so on. Jobs can differ substantially in scope, with some formed matches lasting for years, while others lasting a day or two as a simple project is completed. See Horton et al. (2017) for roughly contemporaneous details on the distribution of kinds of work, contract structure, and patterns of trade in an online labor market.

Employers can solicit applications by recruiting workers, or workers can just apply to openings they find. The majority of applications on the platform come from workers finding job openings through various search tools and then submitting an application. Applying workers submit a wage bid (for hourly contracts) or a fixed amount (for fixed price jobs). When applying, the worker can observe the number of applicants that have already applied. Employers then screen applicants and potentially make a hire or make multiple hires—though hiring a single worker is by far the most common choice, conditional upon hiring anyone.

Applicants arrive very quickly. The reason for this speed is that workers have an incentive to apply as quickly as possible, all else equal, as they do not know exactly when the employer will start making a decision. Fast applications also seem to be the case in conventional markets when application behavior is observed (see van Ours and Ridder (1992)).

There is a burgeoning literature that uses online labor markets as a domain for research. Pallais (2013) shows via a field experiment that past on-platform worker experience is an excellent predictor of being hired for future job openings. Stanton and Thomas (2016) shows that agencies (which act as quasi-firms) help workers find jobs and break into the marketplace. Agrawal et al. (2013) investigate what factors matter to firms in making selections from an applicant pool and present some evidence of statistical discrimination, which can be ameliorated by better information. Horton (2017) explores the effects of making algorithmic recommendations to would-be employers. Barach and Horton (2020) reports the results of an experiment in which employers lost access to wage history when making hiring decisions.

Although our setting offers a rich, detailed look at hiring, there are limitations. A downside of our context is that it is one marketplace. However, when applications are observable in conventional markets, the success probability also appears to be quite low and is similar to what we observe (Skandalis and Marinescu, 2018). Although our context is unique, the basic economic problem—workers not internalizing the externalities of search intensity—is commonplace, and there is emerging evidence that the precise context matters less than we might imagine for generalization (DellaVigna and Pope, 2019). Furthermore, the job search that occurs on online job boards presently is quite similar to our setting, even if the resulting jobs are different (Marinescu and Wolthoff, 2020).

3 Design of the experiment

How the experiment worked was simple: once either a job opening had 50 applicants or 120 hours (5 days) had elapsed since posting, the job was made "private" and no further would-be applicants could apply. The employer was notified of this change when it happened in the interface and via email. Employers could, at any time, revert the change from public to private by pushing a single button. Appendix A.1 shows the interfaces where these notices were presented to employers.

Randomization was at the level of the employer and the data are consistent with successful randomization. A total of 45,742 jobs openings posted by employers were assigned, covering job openings posted between 2013-11-04 and 2014-02-14. The software used by the platform to randomize employers to treatment cells has been used successfully in many experiments. There were 23,075 job posts in the treatment and 22,667 in the control.[7] The experimental sample was itself randomly drawn from all job openings being posted on the platform. We do not report the exact fraction, but it was less than 1% of all job openings posted in the market, which reduces concerns about cross-group interference.

[7]The p-value for a χ^2 test is 0.056, which is slightly concerning, but daily counts of allocated jobs show has no obvious imbalance and a table of pre-randomization job attributes shows excellent balance, suggesting the low p-value from the χ^2 test is simply due to sampling variation.

After being assigned to a cell, any subsequent job openings by that employer received the same treatment assignment. However, we only use the first job opening in our analysis, as subsequent job openings could have been affected by the experience in the first opening.

4 Results

Job posts were allocated to the experiment over time, and so we can begin by plotting daily statistics by experimental group, which we do in Figure 1. We then explore each outcome in more depth, as well as consider match outcomes. We then shift our lens to take a job-seeker perspective, exploring how the treatment affected their experiences and decision-making.

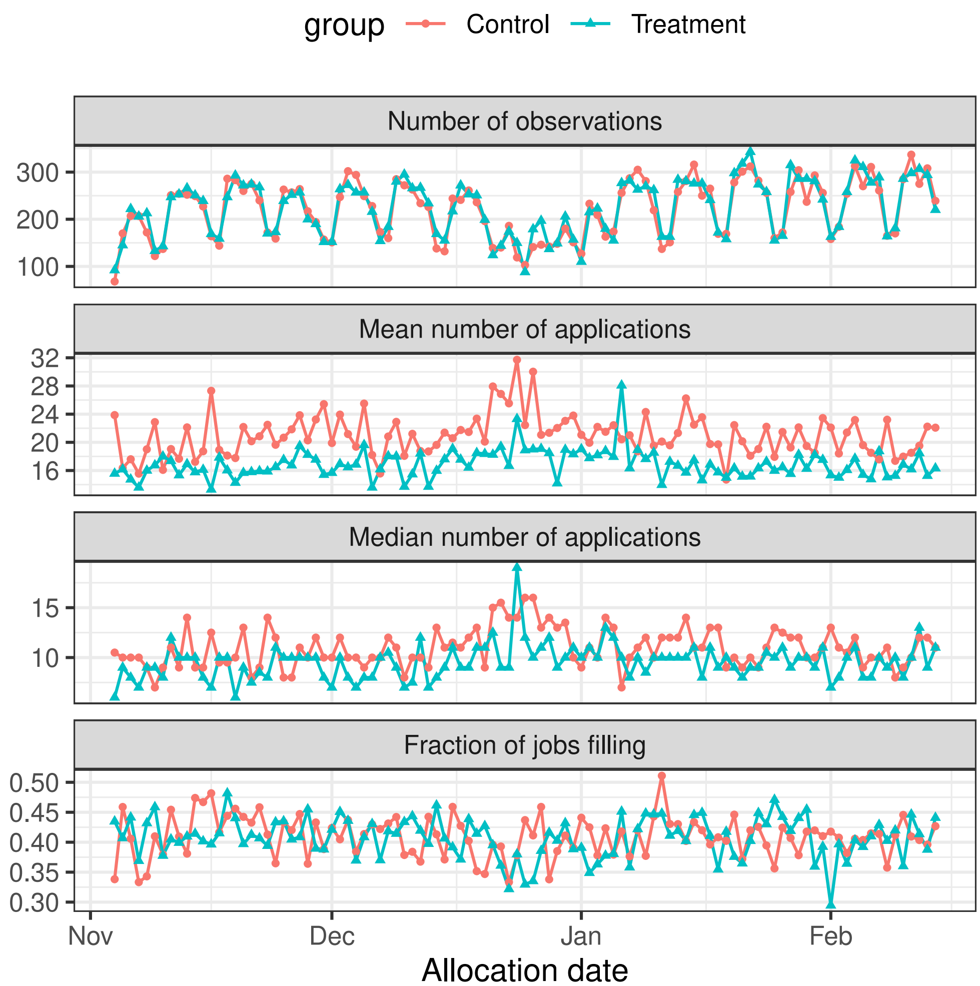
4.1 Experimental outcomes, day by day

The facets of the figure show that the randomization was likely effective, there was a "first stage" of reduced applicant counts, but that the reduction in applicant counts did not reduce match formation.

As expected given random allocation, the top facet of Figure 1 shows the counts of allocated job posts by treatment and control track closely. We also confirm there is no evidence of imbalance by conducting t-tests on pre-randomization attributes, in Appendix A.2.

The treatment reduced the mean number of applications, which we can see in the second facet from the top of Figure 1. The treatment mean is

Figure 1: Group-specific outcomes by allocation date, over time



Notes: This plot shows by-day times series for the two experimental groups. In the experiment, employers posting jobs were randomized to a treatment or a control. Employers in the treatment could not receive additional applicants once they received 50 applicants or 5 days had passed since posting. However, the employer could opt out of this cap by clicking a single button.

always substantially below the control. However, in the facet below that, when we instead plot the *median* number of applications, the difference is smaller, yet still visually evident. This is suggestive that the intervention likely had effects that were not concentrated equally over all jobs, but rather were stronger for jobs that would otherwise receive many applicants.

Previewing one of our main results, the bottom facet for Figure 1 shows there is no obvious evidence of a difference in the probability that a job was filled.[8] We explore these outcomes—and measures of match quality—in the sections that follow.

4.2 Effects of the treatment on applicant pool composition

The treatment had a strong "first stage," lowering applicant counts, with particularly large effects for job posts that would otherwise have received large counts. We visualize the effects of the treatment intervention on applicant pools in Figure 2.

Figure 2a shows the effects of the 120 hour time limit. We plot the kernel density estimate of the relative arrival time of applicants, by treatment and control groups (with some restrictions).[9] We can see that distributions are nearly identical up until 5 days, at which point the treated group shows a

[8] A table of summary statistics for our primary outcomes, by cell are in Appendix A.3.

[9] The sample is restricted to jobs that received 50 or fewer applicants and arrived within the first 10 days. We also remove a small fraction of applications that arrive in less than one minute, so as have a sensible distribution given our log scale. We observe the application arrival times—measured down to the millisecond—relative to when the job was posted.

marked fall-off, consistent with how the treated intervention worked. We can also see how quickly applications typically arrive—both groups exhibit a peak around 20 minutes after posting, with flows declining sharply afterwards.

Figure 2b shows the effects of the 50 applicant soft cap. We plot the kernel density estimates for the application counts for treatment and control. We restrict the domain to less than 200 applicants. As expected, there is a "jump" around 50 applicants in the treatment and no such jump for the control.[10] Prior to 50, there is some slight visual evidence of fewer applicants, but this is better explored with a quantile regression.

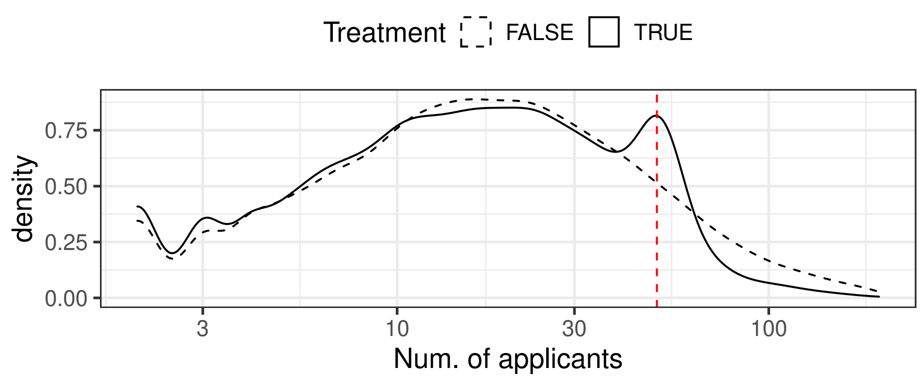
Figure 2c shows precisely where the treatment effects on application counts were concentrated using quantile regressions. The y-axis is log transformed. The x-axis is the associated percentile. Below about the 25th percentile, there is no evidence of an effect. From the 25th to about the 90th percentile, the reduction is about 1 applicant or 2 applicants, but is much larger above the 90th percentile. For comparison, the OLS estimate of the treatment effect is plotted as a horizontal dashed line, which is about 4.

The question we turn to now is how these applicant pool changes affected the probability that an applicant was hired in each job, and what kind of match they formed.

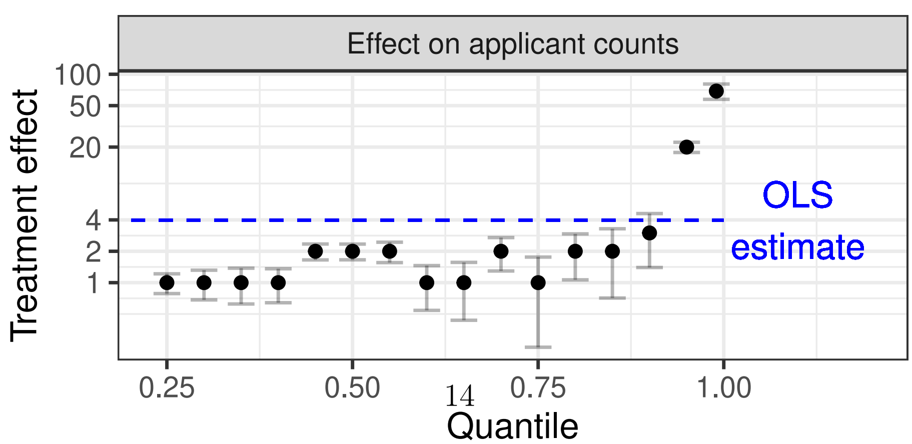
[10] Despite the cut-off of 50 applicants, there is actually excess mass at numbers slightly greater than 50—a fact obscured by the density plot. What causes this is that some applicants withdraw their applications, and withdrawn applicants do not count against the cap.

Figure 2: Evidence of the effect of the interventions of applicant pools

(a) Distribution of application arrival times, by treatment group, for jobs receiving fewer than 49 applicants



(b) Kernel density estimate of the distribution of applicants per job opening, by treatment and control



(c) Quantile treatment effects

Table 1: Effects of the treatment on number of applications and whether the job opening filled

	Any hires?	Total hires	Any hires after 55?
	(1)	(2)	(3)
Treatment	-0.002	-0.016	-0.007***
	(0.005)	(0.010)	(0.001)
Intercept	0.411***	0.528***	0.020***
	(0.003)	(0.007)	(0.001)
N	45,742	45,742	45,742
R squared	0.00000	0.00006	0.00068

Notes: This table reports the effects of treatment assignment on whether a hire wade in Column (1), the total quantity of hires, in Column (2), and whether a hired applicant was greater than the 55th arrival. In the experiment, employers posting jobs were randomized to a treatment or a control. Employers in the treatment could not receive additional applicants once they received 50 applicants or 5 days had passed since posting. However, the employer could opt out of this cap by clicking a single button. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : * * *$.

4.3 Effects on match formation probability

Despite reducing applicant counts, the treatment had no discernible effect on the probability a treated job was filled. In Table 1, Column (1), the outcome is an indicator for whether any applicant was hired. We can see that the treatment effect is a precise 0. Note that the baseline fill rate in the control group is about 41%.

In addition to not affecting whether a hire was made, the treatment had no discernible effect on total quantity of hires. In Column (2) of Table 1 we regress the total number of hires on the treatment indicator. The point estimate is negative, but not significant and fairly close to zero. This is important, as some employers do make multiple hires for one job opening, and so the treatment could have thwarted this desire if employers with these intentions could not easily opt out. If the experimental intervention was performed again, it would be prudent to ask the employer upfront whether they plan to make multiple hires.

4.4 Later applicants crowd-out earlier applicants

The treatment caused employers to hire from applicants arriving earlier. We can see this directly in Column (3) of Table 1, where the outcome is whether the employer hired anyone who applied after the 55th applicant. "Late" hiring is substantially reduced in the treatment—the treated group was 33% less likely to hire a 55+ applicant than the control.

Note that the decline in late hiring in the treatment is *not* mechanical but instead reveals applicant crowd-out. To illustrate this point, imagine if employers always received at most one acceptable applicant and that that applicant could arrive before or after the cap. With a hard cap, treated employers whose acceptable candidate would have arrived after the cap would simply not hire. With a soft cap, those treated employers would still be able to hire, but we would find no average difference between treatment and control in the probability that an applicant arriving after the cap was hired. We have a soft cap but still find substitution towards applicants arriving earlier, i.e., we have crowd-out.

Table 2: Match outcomes, conditional upon a hire

	Log hired worker wage	Log hours-worked	Feedback on worker
	(1)	(2)	(3)
Treatment	-0.002	0.042	0.011
	(0.014)	(0.044)	(0.012)
Intercept	2.240***	2.987***	4.667***
	(0.010)	(0.031)	(0.009)
N	9,354	7,082	16,330
R squared	0.00000	0.00013	0.00005

Notes: The sample for these regressions are those job openings where a hire was made. In the experiment, employers posting jobs were randomized to a treatment or a control. Employers in the treatment could not receive additional applicants once they received 50 applicants or 5 days had passed since posting. However, the employer could opt out of this cap by clicking a single button. In Column (1), the sample consists of hourly job openings; in Column (2), job openings where at least 1 hour was billed. In Column (3) the sample is all job openings, including fixed price jobs. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

4.5 Effects of the treatment on match quality

There is no evidence that the treatment affected the characteristics of formed matches, including quality. To look for match quality effects, in Table 2, we regress several match outcomes on the treatment indicator. It is important to note that the samples used in Table 2 are selected, in the sense that these are only filled job posts. However, we have no evidence that the treatment changed the composition of filled jobs.

The treatment had no discernible effect the wage of the hired worker. This hourly wage is the outcome of the regression reported in Column (1). The coefficient is close to zero and precisely estimated. There is no evidence the employer was getting less surplus in terms of price.

The treatment had, if anything, a small positive effect on the hours-worked within the match, though hours-worked is an ambiguous metric of match quality. The outcome in Column (2) is the log total hours-worked per hired worker, conditional upon at least one hour (this is why the sample size is smaller). The estimate is positive, but imprecise. If hours-worked did increase, this could be a sign of a better match (the employer wants to buy more hours) or be a sign of a worse match (the hired worker takes more time to complete the task).

Our best metric for match quality is post-contract feedback, and on this measure, there is some slight evidence of a higher feedback in the treatment. The outcome in Column (3) is the average feedback that the employer left for the worker (on a 1 to 5 star scale). We see no large significant difference in feedback by treatment assignment, though the point estimate is positive.[11]

4.6 Which employers wanted more applicants?

Only about 7% on employers pushed the "opt out" button, with some variation by category of work. With so little uptake, it is hard to conclude very much about what kinds of employers had high marginal returns to more applicants. See Appendix A.4 for further analysis.

[11]Numerical feedback is prone to inflation and strategic misreporting, but as (Filippas et al., 2018), star feedback is still highly correlated with measures of reviewer satisfaction. Note that the sample in Column (3) is larger because fixed price jobs also generate feedback.

4.7 Effects of the treatment on job-seekers

With smaller applicant pools but the same probability a job is filled, we should expect that job-seekers applying to treated job openings enjoyed a higher probability of being hired. To measure this effect, we can compare per-application win rates, based on the treatment assignment of the applied-to job opening. Workers did not know the treatment status of the job openings when deciding whether or not to apply. We observe 129,520 distinct job-seekers collectively sending 738,861 applications to job openings assigned to the experiment. The mean number of applications per worker is 5.7, while the median is 2.

As job-seekers typically send many applications, we can include a worker-specific fixed effect to perform a within-worker analysis, obviating concerns about worker selection. The selection we would otherwise be worried about is that the kinds of workers that apply to job openings with many applicants (which are disproportionately found in the control) are different from those applying to jobs in the treatment. We estimate a regression of the form

$$y_{ij} = \beta \cdot \textsc{Trt}_j + \textsc{AppCount}_j + \gamma_i + \epsilon \quad (1)$$

where y_{ij} is some outcome or choice for worker i applying to job post j, $\textsc{Trt}_j$ is the treatment assignment of the applied-to job opening, γ_i is a worker-specific fixed effect and $\textsc{AppCount}_j$ is a fixed effect for the total applicant count for that opening. As workers send different numbers of applications, we weight these regressions by the inverse of the total number of applications sent by the worker.

As expected, workers are more likely to be hired when applying to a treated job opening. We can see this in Column (1) of Table 3, where the outcome is an indicator for whether the worker was hired. Given the baseline hiring probability, this coefficient implies about a 17% increase in hiring probability on a per-application basis. Note that the baseline probability of being hired is about 3%—which is nearly identical to that rate found by Skandalis and Marinescu (2018).

Workers applying to treated jobs had a lower arrival rank, confirming the treatment affected the worker's application experience. The outcome in Column (2) is the applicant's rank in the applicant pool (i.e., first applicant is 1, second applicant is 2, etc.). Unsurprisingly, this falls as well, as "late" applications are missing from treatment jobs.

There is some evidence that workers applying to treated jobs bid higher, likely reflecting the difference in perceived competition. The outcome in Column (4) is the log wage bid. The treatment raises wage bids by about 1.2%. This is suggestive that aside from simply saving job-seekers time, the treatment could also transfer some surplus from employers to workers by reducing *in situ* competition. However, recall that there was no evidence of a substantial change in hired worker wage at the job level from the treatment, and so it is unclear whether this channel is important in practice.

Table 3: Association between treatment status of applied-to job opening and application outcomes

	Hired	Rank	Log wage bid
	(1)	(2)	(3)
Treatment	0.003***	-13.712***	0.012***
	(0.001)	(0.186)	(0.002)
DV Mean	0.02	33.23	1.89
Worker FE	Y	Y	Y
Worker Cluster SE	Y	Y	Y
N	738,861	738,861	466,592
R squared	0.54924	0.80494	0.95144

Notes: This table reports regressions of applicant-level outcomes on the treatment status of the applied-to job opening. The sample consists of all applications sent to assigned job openings. In the experiment, employers posting jobs were randomized to a treatment or a control. Employers in the treatment could not receive additional applicants once they received 50 applicants or 5 days had passed since posting. However, the employer could opt out of this cap by clicking a single button. Regressions are weighted by the inverse total number of applications sent by each worker. Each regression includes a worker-specific fixed effect. Standard errors are clustered at the level of the individual job applicant. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : ***$.

4.8 Should we have expected match quality effects?

With fewer applicants to choose from, we would expect worse matches. If we imagine each applicant offers some idiosyncratic surplus, fewer choices means a lower expected pay-off. However, there are economic reasons why decreases in match quality would be small—namely if vertical differentiation gets priced into worker wage bids. Horizontal differentiation might still matter, but consider that the amount of variation *within* a self-selecting group of job-seekers responding to a particular job posting might be minimal. As such, price competition would tend to make employers indifferent over applicants in the pool. This in turn would tend to make applicant pool size practically irrelevant and explain why employers would be so willing to substitute among applicants.

It is useful to step back from the labor context to highlight the interplay of pricing, vertical differentiation, and choice. This is perhaps easier with a simple good. Suppose you are at a grocery store and you are selecting an apple to buy from the store's display. If there is a lot of variation in quality—some look great, some are bruised or have soft spots—but all are offered at the same price, having more apples offered would help you form a better "match." You are more likely to get a very good apple if you have more to choose from. And if good apples are very rare, then the marginal benefit from more apples declines slowly. This is the logic of why we might expect worse matches with fewer choices.

But now instead suppose each apple was priced differently and independently to account for its own vertical attributes. The perfect apple has a high price; the bruised apple offers a substantial discount. So long as each apple views itself as not having much market power, having more apples does not help very much for buyer surplus. You, the consumer, are going to get the quality-adjusted market rate apple surplus, no matter which apple you choose. Any additional surplus you get is going to reflect perhaps horizontal preferences or pricing error—both of which would likely be small for a good like apples. This "individually priced apple" notion has strong support in our data.

In Table 4, we report regressions of an indicator for an application leading to a hire (x 1,000) on the log wage bid from the applying worker and an opening-specific fixed effect. The sample consists of applications only to hourly job openings in the control group. The critical difference among specifications is whether a worker-specific fixed effect is included. Although we lack a true experiment in the wage bidding, the residual variation in wage bids is presumably caused in part by "sellers" experimenting, as in Einav et al. (2015).

In cross sections without worker fixed effect, the wage bid and hire probability are slightly *positively* correlated—higher bidding workers are *more* likely to get hired. We can see this in Column (1), where the sign on the wage goes the "wrong" way. But of course, this bid is a choice by the applying job-seeker, and they are not bidding randomly. They know if they are a "good apple" or not and can tailor their bid accordingly.

Table 4: Association between worker application wage bidding and hiring in the control group

	Hires (1/0) x 1000		
	(1)	(2)	(3)
Log wage bid	2.018***	-11.019***	-11.018***
	(0.247)	(1.117)	(1.117)
Applicant arrival rank			-0.002
			(0.008)
Intercept	9.492***		
	(0.526)		
N	262,463	262,463	262,463
DV Mean	13	13	13
Worker FE	N	Y	Y
Job Opening FE	N	Y	Y
Worker Cluster SE	Y	Y	Y
R squared	0.00025	0.75011	0.75011

Notes: The table reports regressions of application-level outcomes—namely whether the applicant was hired. In the experiment, employers posting jobs were randomized to a treatment or a control. Employers in the treatment could not receive additional applicants once they received 50 applicants or 5 days had passed since posting. However, the employer could opt out of this cap by clicking a single button. The regressions are weighted by the inverse of the total number of applications sent by the worker. The sample consists of all applications to all job openings assiged to the control group in the experiment. Standard errors are clusterd at the worker level. Significance indicators: $p \leq 0.05 : *$, $.p \leq 0.01 : **$ and $p \leq .001 : * * *$.

Workers face a strong negative elasticity of being hired with respect to their wage bid. In Column (2), we add a worker-specific fixed effect, exploiting the matched structure of our data, as in Abowd et al. (1999). The effect is enormous relative to the non-fixed effect regression in Column (1).

Rank effects do not matter much. In Column (3), we add the applicant rank to the Column (2) regression. We see essentially the same coefficient on the wage bid in the Column (2) and Column (3) regressions. The small coefficient on rank reiterates the point that rank effects are not very important.

The Table 4 regressions show that workers condition their bids on their perceived productivity, which can rationalize the lack of effects on hiring and match quality. Rather than interpreting the wage as a direct bid into the "auction" for the job, the wage should be thought of as a component of the total bid, which also includes efficiency and quality, as in a scoring auction. This "scoring auction" interpretation of the hiring situation means that competition among workers tends to leave employers indifferent. If some workers is obviously offering a much better deal, they should instead raise their wage bid, at least in a thick market.

4.9 Characterizing employer decision-making

Our results imply there is substantial heterogeneity among employers in the marginal returns to more applicants. Consider the simplest model of batch hiring—a firm is selecting from a complete pool of applicants and chooses the best one, subject to the candidate exceeding some threshold. If the "scoring auction" framing of the hiring problem is a good description, then we might think of each applicant as having some idiosyncratic probability of exceeding the employer's reservation surplus for hiring. Suppose that search is random and that the firm receives an exogenous number of applicants a, who are all drawn IID from some distribution of worker productivity, which the worker bids, plus some noise. Suppose, further, all firms have some reservation surplus for hiring—they are willing to hire if an applicant is above that value, but otherwise not. Let $\underline{u}$ be the probability that a randomly selected applicant exceeds that hiring reservation surplus.

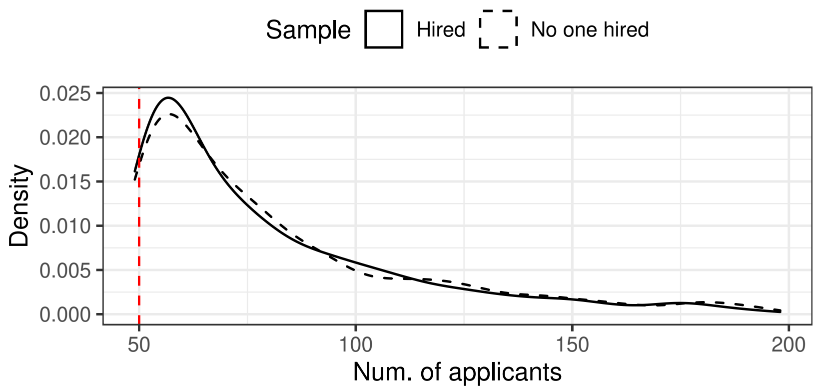
With this collection of a applicants, the probability that the firm would hire anyone is

$$\begin{aligned} Pr(\textsc{Hired}|a) &= 1 - (1 - \underline{u})^a \\ &\approx 1 - \exp(-\underline{u}a). \end{aligned}$$

We estimate these parameters using the control group, finding that the maximum likelihood estimate of $\underline{u}$ is 0.0292, with a standard error 0.0003.

With the fitted model, the treatment should have reduced in the fill rate from the treatment is 26%, which highly highly counterfactual. What is explanation for the divergence? With the size of applicant pools that we observe in the control group and the overall fill rate near 40%, each application must have a low hire probability. But if each applicant has a low hire probability,

Figure 3: Distribution of applications to control job openings, by whether the opening was filled



Notes: This figure plots the kernel density estimate for the log number of applications in the control group, by whether or not the job opening led to a hire.

all else equal, there should be very little crowd-out—it is rare for the firm to get one above-the-bar applicant, never mind two. With little crowd-out, reducing the applicant pool by some percentage amount should have about the same percentage effect on hiring. That we observe no reduction in hiring suggests something is wrong with this simple characterization of the hiring problem.

The problematic assumption is that all employers have the same per-applicant reservation surplus. If instead, employers differ in their value of $\underline{u}$, we can have substantial crowd-out but a far-from-100% fill rate. Some employers are easy to please, get many applicants and hire, but would also still hire with a much smaller pool; other employers are hard to please, get many applicants, do not hire and obviously would not hire with a much smaller pool either.

Job-seekers seemingly *do not* know how to condition on employer "pickiness," exacerbating inefficiency. In Figure 3, we plot the distribution of applications to control openings, by whether or not the job filled. There is no visual evidence that job-seekers try to "avoid" bad jobs, at least in application counts above 49. This is further evidence that more guidance to job-seekers around the likely effect of their applications could be welcome.

5 Conclusion

The key finding of the experiment was that reducing application counts had no discernible effect on whether the employer hired, or the quality of the subsequent match. We argue that price competition among applicants would tend to make employers indifferent over applicants, helping rationalize why substitution was "easy" for the employer.

This paper shows that many applications in a decentralized labor market are likely inframarginal. The value of the marginal proposal for at least some employers was less than the cost of pushing a button. This creates a market design opportunity. Although our results come from a particular marketplace, the issues are commonplace in matching markets and could potentially be addressed by a market-designing platform.

Aside from the pure cost savings from the treatment intervention, presumably some prevented applications could be re-directed to other jobs where they have a higher value. Our context does not allow us to answer the question of whether any of the "saved" application were re-directed to jobs where they would have a higher value. This would be an interesting question for future research.

Another interesting question is what would be the optimal number of applications per job, and whether market design changes could help bring that about. Although we have focused on crowd-out and job-seekers not internalizing costs, job-seekers also do not internalize the full benefits of the applications in creating matches—externalities abound in equilibrium search models (Diamond, 1982; Mortensen and Pissarides, 1994; Hosios, 1990). Beyond crowd-out and crowd-in, there are also congestion concerns that would be interesting to explore empirically (Roth and Xing, 1994; Albrecht et al., 2006).

Our paper illustrates that even small interface and policy changes by platforms can lead to large economic changes. As more of economic life becomes computer-mediated, the opportunity for platforms to control or influence the number of counter-parties a seller interacts with has grown (Varian, 2010; Fradkin, 2015; Halaburda et al., 2017; Li and Netessine, 2019). And although the effects of the Internet on labor markets was initially thought to be limited (Kuhn and Skuterud, 2004; Kroft and Pope, 2014), with the continuing maturation and expansion of the Internet, this characterization might already be changing (Kuhn and Mansour, 2014).

References

Abowd, John M, Francis Kramarz, and David N Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

Agrawal, Ajay K, Nicola Lacetera, and Elizabeth Lyons, "Does information help or hinder job applicants from less developed countries in online markets?," January 2013, (NBER Working Paper 18720).

Albrecht, James, Pieter A Gautier, and Susan Vroman, "Equilibrium directed search with multiple applications," *The Review of Economic Studies*, 2006, *73* (4), 869–891.

Barach, Moshe E. and John J. Horton, "How do employer use compensation history? Evidence from a field experiment," *Journal of Labor Economics*, 2020.

Belot, Michele, Philipp Kircher, and Paul Muller, "Providing advice to jobseekers at low cost: An experimental study on online advice," *The Review of Economic Studies*, 2019, *86* (4), 1411–1447.

Bhole, Monica, Andrey Fradkin, and John Horton, "Dealing with congestion," *Working Paper*, 2021.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora, "Do labor market policies have displacement effects? Evidence from a clustered randomized experiment," *The Quarterly Journal of Economics*, 2013, *128* (2), 531–580.

DellaVigna, Stefano and Devin Pope, "Stability of experimental results: Forecasts and evidence," *Working paper*, 2019.

Diamond, Peter A, "Aggregate demand management in search equilibrium," *Journal of Political Economy*, 1982, *90* (5), 881–894.

Einav, Liran, Theresa Kuchler, Jonathan Levin, and Neel Sundaresan, "Assessing sale strategies in online markets using matched listings," *American Economic Journal: Microeconomics*, 2015, *7* (2), 215–47.

Feld, Scott L, "Why your friends have more friends than you do," *American Journal of Sociology*, 1991, *96* (6), 1464–1477.

Filippas, Apostolos, John J. Horton, and Joseph Golden, "Reputation inflation," in "Proceedings of the 2018 ACM Conference on Economics and Computation" 2018, pp. 483–484.

Fradkin, Andrey, "Search frictions and the design of online marketplaces," *Working paper*, 2015.

Gee, Laura K, "The more you know: Information effects on job application rates in a large field experiment," *Management Science*, 2019, *65* (5), 2077–2094.

Halaburda, Hanna, Mikołaj Jan Piskorski, and Pınar Yıldırım, "Competing by restricting choice: The case of matching platforms," *Management Science*, 2017, *64* (8), 3574–3594.

Horton, John J., "The Effects of Algorithmic Labor Market Recommendations: Evidence from a Field Experiment," *Journal of Labor Economics*, 2017, *35* (2), 345–385.

_, "Buyer uncertainty about seller capacity: Causes, consequences, and a partial solution," *Management Science*, 2019.

_, William R Kerr, and Christopher Stanton, "Digital labor markets and global talent flows," in "High-skilled migration to the United States and its economic consequences," University of Chicago Press, 2017, pp. 71–108.

Hosios, Arthur J, "On the efficiency of matching and related models of search and unemployment," *The Review of Economic Studies*, 1990, *57* (2), 279–298.

Iyengar, Sheena S and Mark R Lepper, "When choice is demotivating: Can one desire too much of a good thing?," *Journal of Personality and Social Psychology*, 2000, *79* (6), 995.

Kroft, Kory and Devin G Pope, "Does online search crowd out traditional search and improve matching efficiency? Evidence from Craigslist," *Journal of Labor Economics*, 2014, *32* (2), 259–303.

Kuhn, Peter and Hani Mansour, "Is Internet job search still ineffective?," *The Economic Journal*, 2014, *124* (581), 1213–1233.

— **and Mikal Skuterud**, "Internet job search and unemployment durations," *The American Economic Review*, 2004, *94* (1), 218–232.

Lalive, Rafael, Camille Landais, and Josef Zweimüller, "Market Externalities of Large Unemployment Insurance Extension Programs," *American Economic Review*, December 2015, *105* (12), 3564–96.

Li, Jun and Serguei Netessine, "Higher market thickness reduces matching rate in online platforms: Evidence from a quasiexperiment," *Management Science*, 2019.

Marinescu, Ioana, "The general equilibrium impacts of unemployment insurance: Evidence from a large online job board," *Journal of Public Economics*, 2017, *150*, 14–29.

— **and Ronald Wolthoff**, "Opening the black box of the matching function: The power of words," *Journal of Labor Economics*, 2020, *38* (2), 535–568.

Mortensen, Dale T and Christopher A Pissarides, "Job creation and job destruction in the theory of unemployment," *The review of economic studies*, 1994, *61* (3), 397–415.

Pallais, Amanda, "Inefficient hiring in entry-level labor markets," *American Economic Review*, 2013.

Romer, David, "Why do firms prefer more able workers," *Department of economics, UC Berkeley*, 1992.

Roth, Alvin E and Xiaolin Xing, "Jumping the gun: Imperfections and institutions related to the timing of market transactions," *The American Economic Review*, 1994, pp. 992–1044.

Scheibehenne, Benjamin, Rainer Greifeneder, and Peter M Todd, "Can there ever be too many options? A meta-analytic review of choice overload," *Journal of Consumer Research*, 2010, *37* (3), 409–425.

Shimer, Robert, "Search intensity," Technical Report 2004.

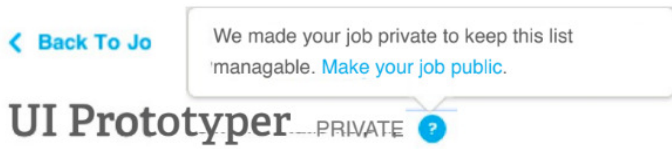
Skandalis, Daphné and Ioana Elena Marinescu, "Unemployment Insurance and Job Search Behavior," *Available at SSRN 3303367*, 2018.

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.

van Ours, Jan and Geert Ridder, "Vacancies and the recruitment of new employees," *Journal of Labor Economics*, 1992, *10* (2), 138–155.

Varian, Hal R, "Computer mediated transactions," *American Economic Review*, 2010, *100* (2), 1–10.

Figure 4: Notices to employers about their job opening



(a) Notice to treated employers that their job was made private

(b) Button to make a public job private

(c) Button to make a private job public

Notes: These are screenshots of the interfaces shown to employers.

A Appendix

A.1 Interfaces

Figure 4 shows the on-app notice that employers received that their job was made private. The employer could request "re-opening" the job with a single button click both on the notice. Both treated and control workers could make their job private at any time—the interface is in Figure 4b. They could also make a private job public at any time—the interface is in Figure 4c. Neither workers nor employers knew about the treatment *ex ante* and workers never learned about the treatment.

A.2 Internal validity

To assess balance, in the top panel of Table 5, labeled "Pre-Treatment" we report means for several pre-experiment job characteristics, by group. In the first row, we have the vertical preference score, which is elicited from the employer when they post a job. It ranges from 1 (an employer interested in low cost and willing to work with relatively inexperienced workers) to 3 (an employer interested in the most experienced workers and willing to pay high wages). We also have the employer's estimate of how long they think the job will take (in weeks). Another pre-experiment outcome is whether the employer included an attachment in the job post (typically images or a more detailed specification). Below that, we have an indicator for whether the job opening was hourly. Across all pre-treatment measures, we obtain excellent balance, consistent with successful randomization.

We can see this excess pooling in the "Post-Treatment" panel of Table 5, where we report the fraction of job openings receiving various numbers of applications. As expected, there is a much larger share in the treatment group receiving exactly 50 applicants. But we can also see a pooling at exactly 51 applicants. We see excess mass all the way up to 55 applications. What explains this "near 50" pooling is that applicants can and do withdraw their applications, which raised the effective cap. Despite this withdrawal, we include these applications in our datasets. This means that even a treated employer hitting the soft cap could still obtain more than 50 applicants in total without requesting more applicants.

Table 5: Job opening attributes and outcomes for treatment group

	Means		Difference		
	Treatment	Control	Diff.	SE	t-stat
Pre-Treatment					
Vertical preference score	0.894	0.879	0.015	0.010	1.486
Duration weeks	8.435	8.465	−0.030	0.160	−0.187
Has attachment	0.109	0.108	0.001	0.003	0.463
Hourly job?	0.521	0.517	0.005	0.005	0.994
Post-Treatment					
Number of applications	16.849	20.821	−3.972	0.323	−12.315
Number of apps equal to 49	0.003	0.003	−0.000	0.001	−0.782
Number of apps equal to 50	0.018	0.003	0.016	0.001	16.566
Number of apps equal to 51	0.011	0.003	0.008	0.001	10.413
Number of apps equal to 52	0.008	0.003	0.005	0.001	7.032
Number of apps equal to 53	0.006	0.003	0.003	0.001	5.166
Number of apps equal to 54	0.005	0.003	0.001	0.001	2.453
Number of apps equal to 55	0.004	0.003	0.001	0.001	2.236
Number of apps equal to 56	0.003	0.003	0.000	0.001	0.589
Number of apps equal to 57	0.003	0.003	0.000	0.001	0.327
Number of apps greater than 57	0.035	0.077	−0.042	0.002	−19.545

Notes: Table of pre- and post-randomization job level attributes. There are 23,075 job openings in the treatment and 22,667 in the control. In the experiment, employers posting jobs were randomized to a treatment or a control. Employers in the treatment could not receive additional applicants once they received 50 applicants or 5 days had passed since posting. However, the employer could opt out of this cap by clicking a single button.

Table 6: Summary statistics for the experimental sample of job openings

	N	Min	25th	Mean	Median	75th	Max	StDev
Number of apps								
Control	22,667	0.00	2.00	20.82	11.00	13.00	1,536.00	37.31
Treatment	23,075	0.00	1.00	16.85	9.00	11.00	3,194.00	31.37
Any hires								
Control	22,667	0.00	0.00	0.41	0.00	0.00	1.00	0.49
Treatment	23,075	0.00	0.00	0.41	0.00	0.00	1.00	0.49
Total hires								
Control	22,667	0.00	0.00	0.53	0.00	0.00	75.00	1.22
Treatment	23,075	0.00	0.00	0.51	0.00	0.00	33.00	0.92
Average wage bid								
Control	10,277	0.01	6.15	12.27	10.30	11.11	96.89	8.64
Treatment	10,459	0.01	6.06	12.17	10.22	11.06	83.33	8.58
Average wage hired								
Control	4,660	0.01	5.77	11.74	9.93	10.64	96.89	8.42
Treatment	4,694	0.01	5.83	11.72	10.00	10.75	80.00	8.19

Notes: Opening level outcomes by treatment and control group.

A.3 Summary statistics

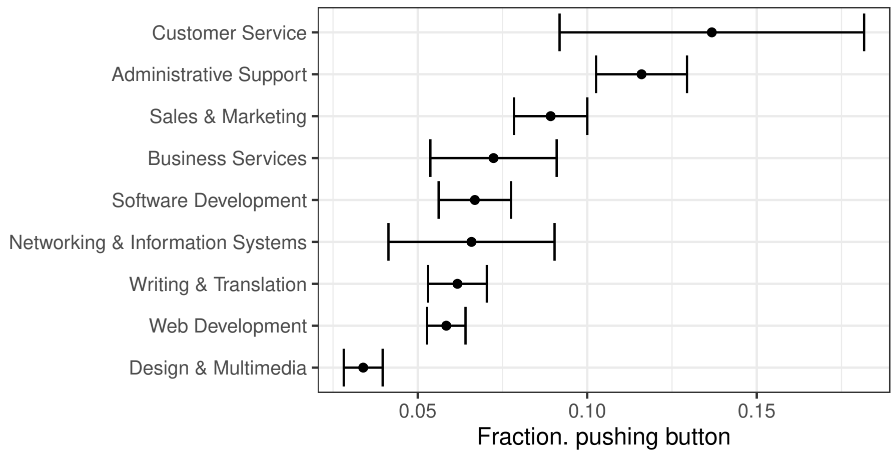
In addition to our regression approaches, many of the main results can be observed simply by comparing means of outcomes. Table 6 reports summary statistics for the job opening sample, by treatment assignment.

A.4 Button-pushing

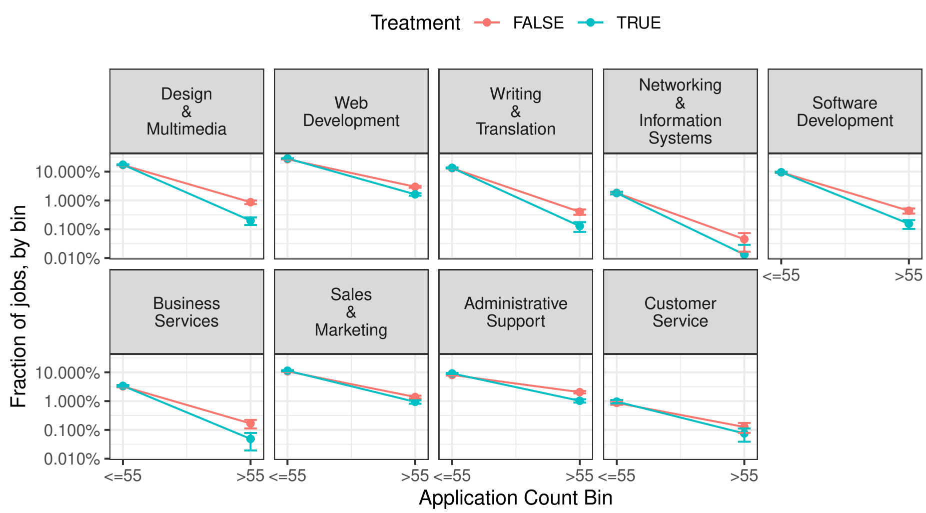
In Figure 5a, we plot the imputed fraction of employers pushing the button to receive more applications. Some categories have a low volume and so the estimated fraction is imprecise. Generally, we see more button-pushing in relatively low-skilled categories. One likely explanation is that these are also the categories with relatively large application counts, and so the employer was more likely to hit the 50 cap. This point illustrates an advantage of the soft cap design rather than the platform trying to decide *a priori* which kinds of jobs have a higher marginal return to more applicants.

Figure 5b shows the fraction of jobs with more than 55 or fewer than 55 applicants, by category and treatment status. The y-axis is on a log scale, as the fraction of openings exceeding the cap is quite small in some categories, and some categories are quite small. We can see that, for example, "Design & Multimedia" has large reductions in the 55+ group, as this category tends to attract many applicants and few would-be employers pushed the button. "Customer Service" shows little evidence of a decline in the 55+ group, but we can also see that this a small category, with less than 1% of jobs posted.

Figure 5: The kinds of jobs that required more applications



(a) Implied fraction of employers pushing the button



(b) Percentage of jobs with 55+ applicants, by category and group