The More You Know: The Effect of Information on Job Application Rates In A Large Field Experiment*

Laura K. Gee[1]

[1]Tufts University, Economics (laura.gee@tufts.edu)

July 9, 2014

(Keywords: field experiment, labor search, big data, risk aversion, ambiguity aversion, herding)

(JEL: C93, D01, D83, J21, J22)

Abstract

This paper presents the results from a field experiment which varies the amount of information seen by two million job seekers when viewing 100,000 job postings on a large online job posting website. The information seen is the true number of people who previously started an application. I find that showing this information increases the likelihood a person will start/finish an application by 2-5%, representing a potential increase of thousands of applications per day. Beyond increasing applications, the treatment also changes the makeup of the applicant pool by increasing the number of women who apply. Firms in industries like high tech and finance that are highly represented on this job posting website, may be particularly interested in this low cost, light touch intervention to increase the number of female applicants.

*A huge thanks to Joe Adler and Jim Baer at LinkedIn. Additionally this work would not have been possible without the assistance of Mathieu Bastian,Joi Deaser, Vitaly Gordon, Abishek Gupta, Kuo-Ning Huang,Sachit Kamat, Navneet Kapur, Shreya Oswal and others at LinkedIn for sharing their knowledge and expertise. Thank you to Kevin Morsony, Product Counsel at LinkedIn, for reviewing this paper. This research was approved by the Tufts IRB (1308011).

1 Introduction

There are two sides to the labor market: firms which demand labor and job seekers who supply labor. Job seekers ultimately want to find the job that is the best fit for them. While firms want to maximize the number of quality applicants who apply for the position, while possibly targeting certain types of applicants (e.g. to increase diversity in their workforce).[1] This paper uses a large scale field experiment of almost 2 million job seekers viewing 100,000 real job postings to test how the addition of information to a job posting effects the likelihood a job seeker will apply to that posting and the final makeup of the applicant pool.

Job seekers may be affected by the extra information because it helps them to weigh the costs of application against the benefits of a possible job offer. The costs can quite high with most people estimating that it takes over an hour to finish an application.[2] The benefits of applying are often modeled as a known probability of a job offer with a known utility level. However, in reality people may not know much about the position when deciding to apply. For example, people may not know the exact probability they will get called for an interview, enjoy the actual position, or that the compensation package will meet their needs. Yet, most previous work assumes that job search is performed under known risks rather than unknown risks (Galenianos and Kircher, 2009; Mortensen, 1970; Das and Tsitsiklis, 2010; Chade and Smith, 2006; Weitzman, 1979; Kohn and Shavaell, 1974; Telser, 1973; Nachman, 1972; Stigler, 1961). This paper begins to bridge the gap between the assumptions of these theories and the reality of the job application process.

Intuitively if one increases the amount of information on the job posting, this may reduce the amount of ambiguity in the job application decision. The specific piece of information in this experiment is the number of people who previously began an application at the time a job seeker looks at the job posting online on the website LinkedIn. This piece of information might increase, decrease or leave unaffected the likelihood of application. Some job seekers may try to avoid congestion when there is high number of started applications while others may herd toward more popular postings (Anderson and Holt, 1997). Yet others, regardless of the number seen, may simply prefer having more information to less because they dislike ambiguity (Halevy

[1] For example in May 2014 Google announced that only 30% of its workforce is female, and only 17% of its "tech" workforce is female. Google also acknowledged that they would like to increase diversity in their workforce. See http://www.forbes.com/sites/jaymcgregor/2014/05/29/2-of-google-employees-are-black-and-just-30-are-women/

[2] See the online Appendix for survey results available at http://laurakgee.weebly.com/index.html

and Feltkamp, 2005; Ellsberg, 1961).

I find there is no strong pattern of either congestion effects or herding, and that is likely because each person has their own interpretation of the number shown (10 seems high to some people but low to others; see online Appendix). However, regardless of the number seen the addition of that information increases the likelihood of starting or finishing an application by 2%-5%. However, that increase is largely driven by female job seekers being induced to apply which is in-line with previous findings that women are more risk and ambiguity averse than men (Eckel and Grossman, 2008; Croson and Gneezy, 2009). For example, showing this information results in an almost 10% increase in the likelihood a female applicant will finish an application, while the effect is not measurable for males. Thus, this paper offers a low cost, light touch intervention to increase the number of female applicants in industries like high tech and finance that are both highly represented on LinkedIn and have higher male participation rates.

2 Field Experiment

The experiment took place on the professional social networking website LinkedIn in March 2012. LinkedIn was launched in 2003. In January 2014 the website had 259 million members from over 200 countries worldwide.[3] LinkedIn is well known for its professional social networking functionality, but it also acts as a job posting website. This paper will concentrate on the job posting functionality of LinkedIn.

Although the population on LinkedIn is not a representative sample of the total worldwide labor force, it is a particulary important population to study when considering gender differences in the labor force. The largest industries on LinkedIn are "High Tech" and "Finance".[4] Industries like this tend to be have lower levels of female labor force participation. For example only 32.5% of US professionals in STEM (Science Technology Engineering and Mathematics) are female.[5]

When a member searches for job postings on LinkedIn, the member would begin on the Jobs landing page

[3]See http://press.linkedin.com/about. Worldwide there are about 3.5 billion people in the labor force (https://www.cia.gov/library/publications/the-world-factbook/rankorder/2095rank.html). So, only about 7% of the worldwide labor force has a LinkedIn account.

[4]http://www.linkedinppc.com/target-by-industry-company-category/

[5]http://dpeaflcio.org/programs-publications/issue-fact-sheets/women-in-stem/

as pictured in Figure 1 and would be shown some pre-selected job postings.[6] At this point the member can click on one of the posting listed, or can enter a term into the search bar which will return results like those shown in Figure 2. After clicking on a posting a member will see a full page description of the posting, and this is where the treatment and control group differ in their job search experience.

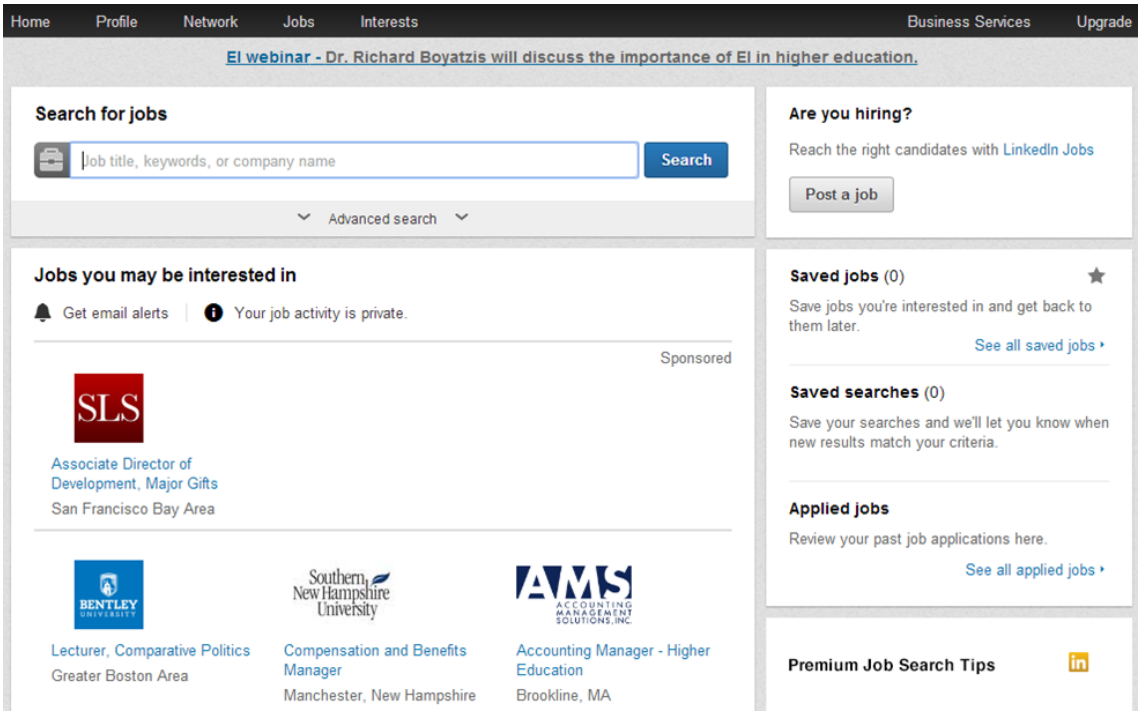


Figure 1: Jobs Landing Page

Note: This figure shows the jobs landing page a LinkedIn user might see when she logs on to the website.

Before describing the details of the treatment, one must also know there are two types of job postings on LinkedIn which I will call "interior" and "exterior" postings as pictured in Figure 3. For interior postings LinkedIn collects the finished application and forwards it to the company, so for interior job postings I can observe if a member starts and also finishes an application.[7] Exterior postings link a job seeker to an external website, so I can only observe if a user starts an exterior job application.

The two main outcome variables are the dummy variables "Start Application" and "Finish Application".

[6] Jobs are generally selected by LinkedIn based on the information the member has listed on member's profile like education, industry, and previous employment.

[7] I have the timestamp of when a job seeker clicks "Apply" and also the timestamp for when they submit an application. If a person submits an application within one day of viewing the posting, then this is coded as a finished application. This restriction is likely to bias the number of total finished applications downward since some people may take more than an day to finish an application or may come back at a later date to finish the application. However, I have no reason to believe this bias will be different across the control and treatment.

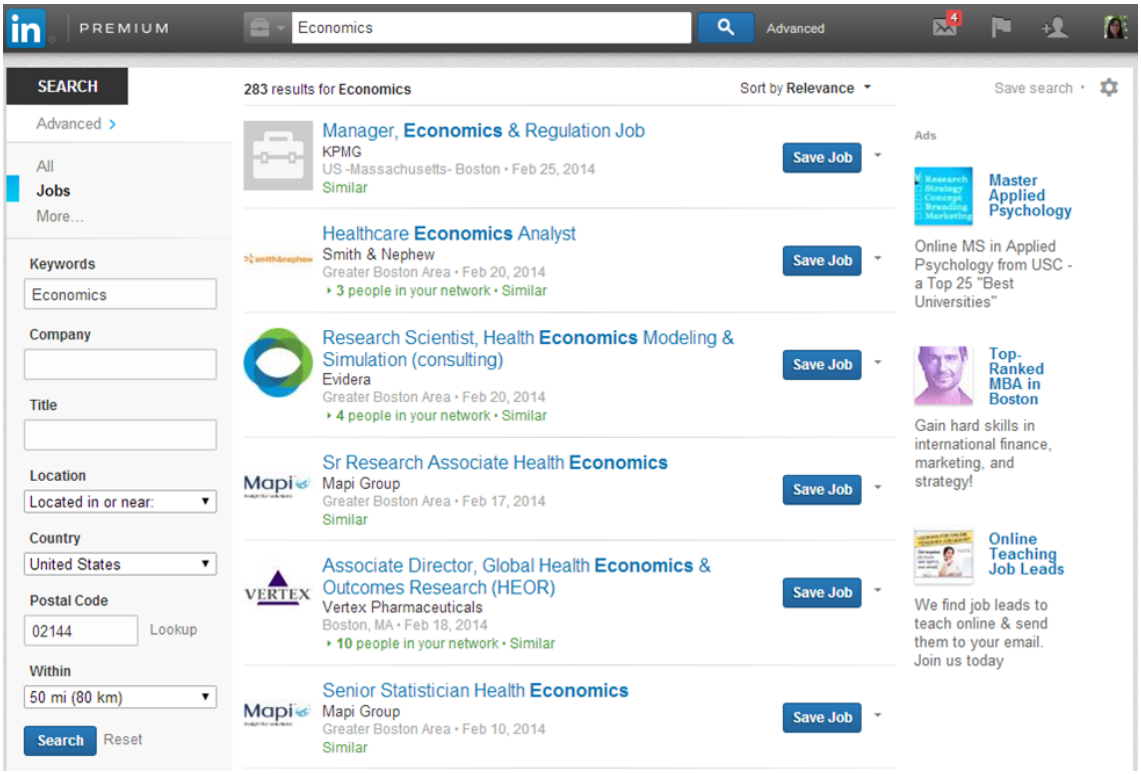
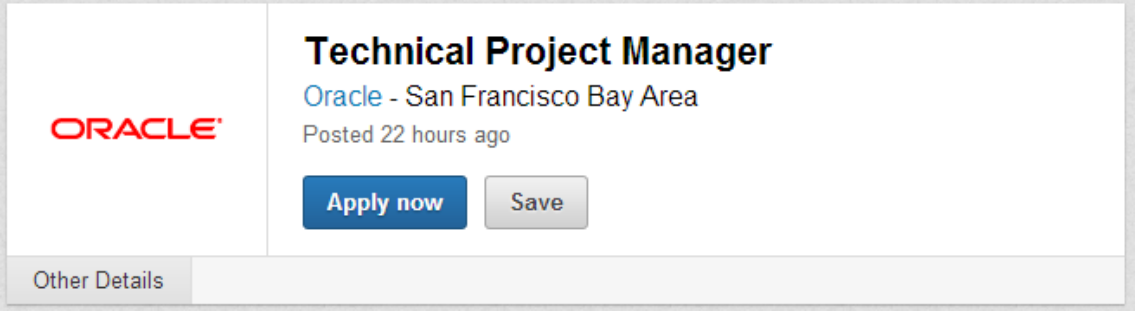


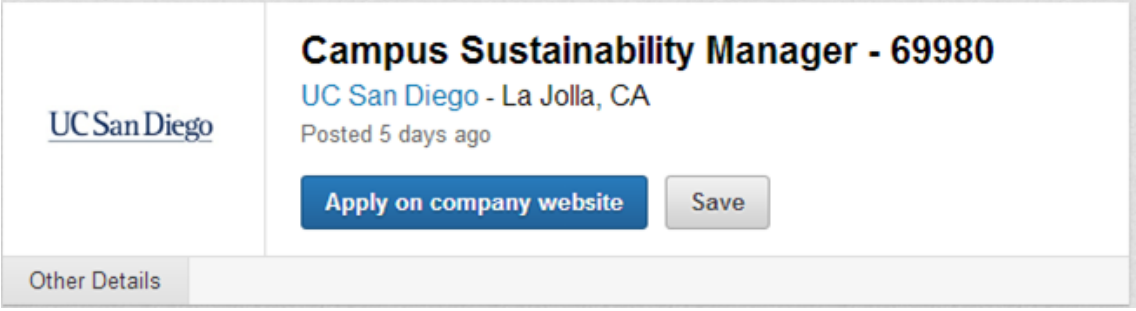
Figure 2: Job Search Landing Page

Note: This figure shows the results from a search for the term "Economics".

For exterior postings I can only tell if someone took the very basic steps of clicking on the "Apply" button. It is impossible for me to determine if that job seeker went on to spend time crafting a cover letter, or instead simply clicked the button by accident. So one can view the Start Application outcome as a noisy measure of interest in the position. Whereas for interior postings I can measure the outcome Finish Application, which is a more accurate measure of investment in the job application.



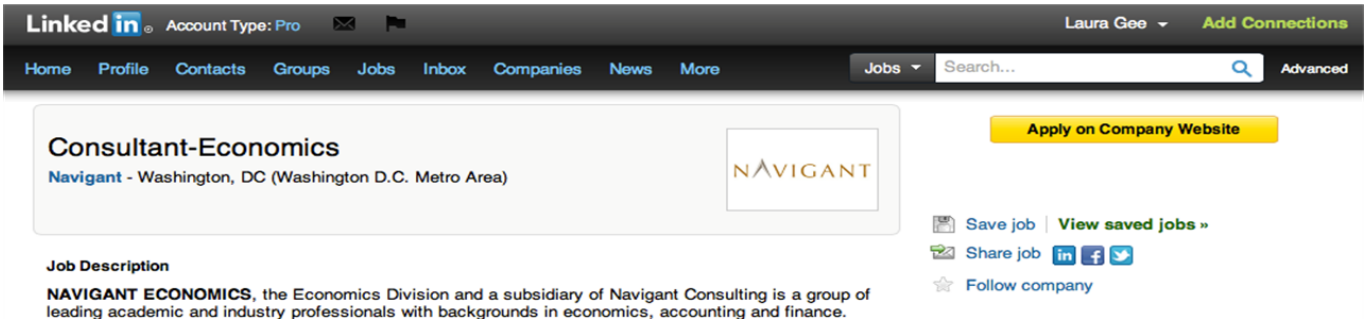
(a) Interior Job Posting



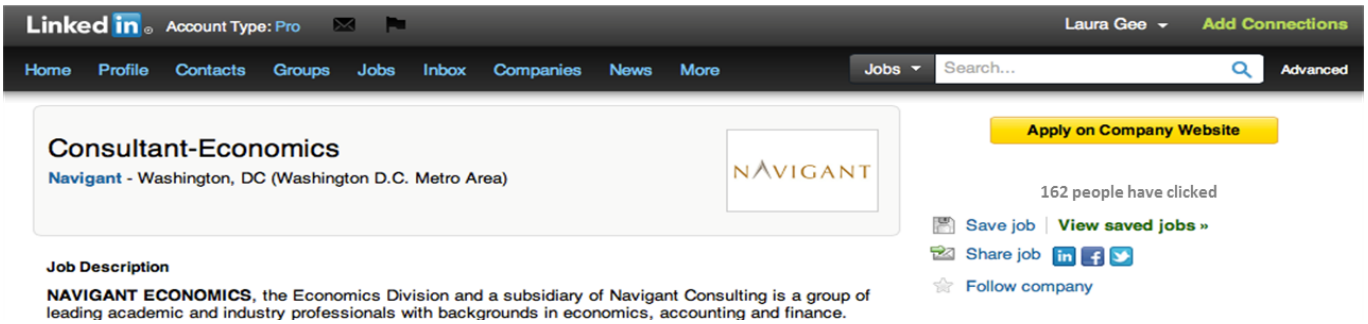
(b) Exterior Job Posting

Figure 3: Types of Job Postings on LinkedIn

Note: This figure shows an example of the two types of job postings on LinkedIn. Panel A shows an interior posting, which means that LinkedIn collects applications for the third party (Oracle) so we can see if a person both begins and finishes an application. Panel B shows an exterior posting, which means that a person is directed to an external website to begin an application so we can only observe is someone begins the application and we cannot observe if she finishes the application. These screenshots were taken in February 2013, which is why they differ very slightly from the formatting seen in the example of the treatment vs. control in Figure 4.



(a) Control



(b) Treatment

Figure 4: Job Posting As Seen In Control and Treatment

Note: This figure shows the way a job posting would be seen by those in the control (Panel A) and the treatment (Panel B). The difference is those in the treatment see that "162 people have clicked" on this job posting to begin an application on the exterior website. Apart from this difference the job posting is displayed identically to those in the control and treatment.

The randomization took place at the member level, so a member in the control would see no information for all the postings he visited during the 16 days of the experiment. On the other hand, a member in the treatment looking at the same job postings would see the number of job seekers who had previously started an application by clicking on the "Apply" button as pictured in Figure 4.[8] Clicking on this button is the first step in starting an application.

This is a unique experiment because I observe how two people looking at the exact same posting change their behavior based on information about how others felt about the attractiveness of the job posting. Namely the number of other people who have already started an application. Additionally, because the information is exogenously assigned I can rule out the possibility that those who seek out more information are already more likely to apply for a position.

One-fourth of the registered members of LinkedIn were randomly assigned to the treatment, so they saw

[8]For an exterior job posting the button reads "Apply on Company Website" while for an interior job the button simply reads "Apply Now".

the true number of job seekers who previously clicked on "Apply" at the time of viewing. The remaining three-fourths were assigned to the control. The experiment only affected registered members who viewed a job posting on LinkedIn during a 16 day period in March 2012.[9]

I restrict the analysis to the first posting a member viewed during the experiment because outcomes may be path-dependent. For example, imagine a person looks at two job postings in total. If she saw 15 applicants on the second posting, that information (the number 15) may have a different effect on her actions if the job posting she saw first displayed 10 previous applicants versus 100 previous applicants (10 then 15 vs. 100 then 15). With these restrictions, the sample includes about two million registered members from 232 countries. There are about 570,000 job seekers in the treatment and 1.4 million job seekers in the control. During the experiment these job seekers viewed a total of about 100,000 job postings from 21 thousand companies. On average each job posting was viewed 20 times and each company had about 4 jobs posted during the experiment.[10]

2.1 Summary Statistics and Balance

The summary statistics for the subjects in the experiment are detailed in Table 1. Subjects in the experiment are about 63% male and 37% female for the 90% where gender is identified.[11] The average age is 36, average year became a LinkedIn member is 2009, about 42% are from the US, with 314-315 links on LinkedIn in Spring 2013, and 3.4 links at the company of the job posting they viewed at the time of viewing (March 2012).[12] The subjects are very well educated with only 2% listing an Associates degree, 52% listing a Bachelors, and 46%

[9] I exclude members who were included in a pilot study for 2 weeks before the main experiment. I also exclude members who visited a posting with 0 previous applicants since these viewers saw no information in either the control or the treatment.

[10] The minimum number of views during the 16 day period was 1 and the maximum was 2,458 with 9 being the median number of views. The minimum number of job postings from a company was 1 and the maximum was 2,088 with the median number of postings from a company being 1. Only 66 companies have 100 or more job postings up during the experiment, and the results are similar if I exclude postings from these companies in the analysis (results available from author by request). Postings viewed by members in the control and treatment both started with an average of 47 previous applicants at the beginning of the experiment.

[11] Members do not actually provide age or gender, so these are imputed from the year the person graduated from college or high school and their name (e.g. Laura in the US is coded female, and Miroslav is coded male in Slovakia). A large portion of the analysis will concentrate on heterogenous treatment effects by sex, so a balance table by sex is provided in the online Appendix. All observable variables are similar across the control and treatment for both men and women.

[12] A "link" is a connection between two LinkedIn members, and it must be approved by both members to exist. For example, a person may ask to be "linked" to a co-worker, and then that co-worker would have to approve that link before it was made on the website. LinkedIn keeps records of the number of connections at the company at the time of viewing, but they do not keep systematic records of the total number of links at the time of viewing.

listing a post-Bachelors degree as their highest education. Overall the randomization worked well, since both the control and treatment are similar on observable variables. There are statistically significant differences between the control and treatment for three observable variables (US, BA listed, Post-BA listed) but the magnitude of these differences is extremely small. Subjects in the experiment came from 232 countries. Looking at Figure 5 one can see that proportion of subjects in the treatment and control is very similar for the most common countries in the sample (a listing of the number of subjects by country is available from the author upon request).

One may fear that those in the treatment may systematically view different postings than those in the control. However Table 2 illustrates that the attributes of the postings are similar across treatment and control.

Table 1: Member Level Summary Statistics

Variable	**Mean (All)**	**N (All)**	**Mean (Control)**	**N (Control)**	**Mean (Treatment)**	**N (Treatment)**	**Min.**	**Max.**	**t-test for diff.**
male	0.635	1,798,968	0.635	1,283,988	0.635	514,980	0	1	0.35
female	0.365	1,798,968	0.365	1,283,988	0.365	514,980	0	1	0.35
gender known	0.900	1,999,964	0.900	1,427,286	0.899	572,678	0	1	0.75
age	35.833	1,578,996	35.840	1,127,672	35.817	451,324	18	94	1.24
year membership	2008.888	1,999,964	2008.888	1,427,286	2008.886	572,678	2003	2012	0.59
US	0.417	1,999,964	0.418	1,427,286	0.416	572,678	0	1	2.54
total links	314.481	1,981,917	314.198	1,414,216	315.189	567,701	0	40,500	1.21
links at company	3.453	1,999,964	3.450	1,427,286	3.461	572,678	0	17,442	0.12
Assoc. listed	0.018	907,675	0.018	650,912	0.018	256,763	0	1	0.23
BA listed	0.518	907,675	0.518	650,912	0.520	256,763	0	1	1.87
Post BA listed	0.462	907,675	0.462	650,912	0.460	256,763	0	1	1.92
Viewed Exterior	0.705	1,999,964	0.705	1,427,286	0.705	572,678	0	1	0.94

Table 2: Posting Level Summary Statistics

Variable	**Mean (All)**	**N (All)**	**Mean (Control)**	**N (Control)**	**Mean (Treatment)**	**N (Treatment)**	**Min.**	**Max.**	**t-test**
start prev. apps	47.07	1,999,964	47.05	1,427,286	47.12	572,678	1	3320	0.29
unixtime 1st seen	1332.79	1,999,964	1332.79	1,427,286	1332.79	572,678	1332	1334	1.31
views per posting	79.99	1,999,964	79.97	1,427,286	80.04	572,678	1	2458	0.34
firm total postings	117.27	1,999,964	117.30	1,427,286	117.20	572,678	1	2088	0.19

Notes: These are not weighted by the number of times the job posting is viewed since the randomization takes place at the member (not the posting) level. The summary statistics weighted by 1 over the number of times a posting was seen are available in the online Appendix.

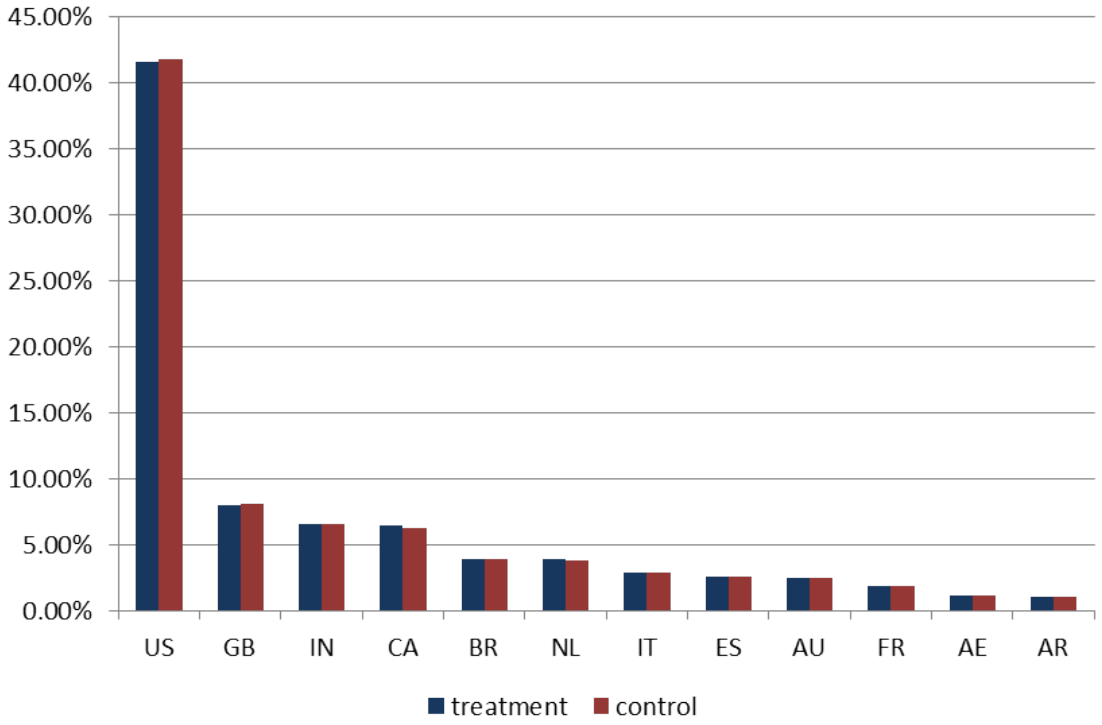


Figure 5: Proportion of Users By Country

Note: The proportion of subjects from each of the 232 coutries is very similar across the control and "See Number" treatment. Key: US United States, GB Great Britain, IN India, CA Canada, BR Brazil, NL Netherlands, IT Italy, ES Spain, AU Australia, FR France, AE United Arab Emirates, AR Argentina.

3 Results

Why would showing the number of people who previously started the application affect the application start or finish rate? Three candidate mechanisms come to mind (see survey evidence in online Appendix). The first is that job seekers are ambiguity or risk averse, so showing them more information increases the likelihood they will apply over a range of numbers shown. The second is that job seekers avoid congested job postings where there are a higher number of applicants. The third is that job seekers herd toward more popular job postings. If job seekers are ambiguity or risk averse, then the treatment should have an overall positive effect on application rates regardless of the actual number shown. If job seekers are instead avoiding congestion or herding, then showing them more information will have a differential effect as the number shown changes.

I begin by presenting the results for the overall treatment effect. I then proceed to test for congestion or herding by exploring the size of the treatment effect by the number of previously started applications shown. Last, I show differential treatment effects by the type of job being applied to.

3.1 Overall Treatment Effect

Each of the 1,999,964 observations is a LinkedIn member who viewed a job posting during the experiment. The outcome variables are whether a person started an application and whether a person finished an application. As already explained I can observe starting an application for both exterior and interior job postings, while I can only observe finishing for interior job postings. So one can think of the outcome variables over two groups: those who saw an exterior posting (N=1,410,384), and those who saw an interior posting (N=589,580). The data only include a user's first job posting viewed during the experiment, so a user is either in the exterior posting group or the interior posting group, but not both.

When a job seeker decides to apply to a job postings this may be a decision with unknown risks on a number of dimensions. The job seeker may not know the probability of an offer, probability the position is a good fit, probability of liking the corporate culture, and so on. Ambiguity aversion describes a preference for known risks (decisions with risk) over unknown risks (decisions with ambiguity). So, for example, an ambiguity averse job seeker might prefer to apply to a job posting with a known 50% chance of an offer, rather than a posting where the odds are unknown. This pattern of decisions can be explained by a number

of models including max-min expected utility or bundled risky decision making (see the online Appendix for a short discussion of Ellsberg (1961); Gilboa and Schmeidler (1989); Halevy and Feltkamp (2005); Halevy (2007)). This experiment decreases the ambiguity about the number of other potential applicants, so showing this information should change the behavior of ambiguity averse job seekers. Previous work has found that women are more ambiguity and risk averse (Moore and Eckel, 2003; Eckel and Grossman, 2008; Croson and Gneezy, 2009) than men. So if ambiguity aversion is driving the results, the treatment should have a larger effect on female job seekers.

In the first three columns of Table 3 I present the results from a simple regression:

$$A_{ijt} = \beta T_i + \epsilon_{ijt}$$

Each observation is a user i who viewed a job posting j at time t. In Panels A and B.i the dependent variable A_{ijt} takes the value 1 if that user *started* that job application by clicking on the "apply" button. In Panel B.ii the dependent variable A_{ijt} takes the value 1 if that user *finished* that job application by submitting all the requested materials. The dependent variable takes the values zero or one, so a logit model would be appropriate. However, I am most interested in the average probability of applying, and the coefficients from the linear model are more easily interpreted. So, I ignore the special nature of the dependent variable and report the results from the linear model in text.[13] Column 1 shows the results for all LinkedIn users, Column 2 is only for female users, and Column 3 is only for male users. The independent variable T_i takes the value one if a user was assigned to the treatment group which sees the number of previously started applications. All standard errors are clustered at the job posting j level.

Looking at Column 1, I find that the treatment of showing the number of previously started job applications increases the likelihood a user will start and/or finish an application by 0.164 to 0.394 percentage points; representing a proportional increase above the control mean of between 2.037%-4.631%.[14] This 2-5%

[13]The sign and significance of the logit models are the same as those presented in text with the exception that there is movement from 5% to 10% significance when comparing the linear fixed effects model with the logit fixed effects model for the *finishing* application outcome for *all* users. See the online Appendix for details.

[14]I will concentrate my analysis on the unconditional finish rate. Meaning I look at the rate of finishing in general, not conditional on starting. The reason is that starting an application is endogenous so I cannot make any causal statements about the effect of the treatment on the conditional finish rate. However it may be of interest that the conditional finish rate is 31.32% for control (18,175 of 58,002) and 31.86% for the treatment (7,586 of 23,800). The difference is not statistically significantly

increase represents a potential increase of thousands of applications per day.[15]

Previous research has found that women are more ambiguity and risk averse than men, so we may expect differential treatment effects by gender. I compare the results from female users in Column 2, to those of male users in Column 3. I find that the effect of the treatment is always larger for women than for men. For example in Panel B.ii the treatment increases the likelihood a female users will finish an application by 0.338 percentage points, while the coefficient for male users is only 0.098; and furthermore this coefficient is not statistically significant for male users.[16]

When comparing Column 2 to Column 3, we can see that the positive and significant effect of the treatment on starting and finishing applications is largely driven by female LinkedIn users being induced to apply. In fact the coefficient for male users is insignificant in Panel B.i and B.ii, meaning there is no measurable effect of the treatment on male users who view an interior job posting. So that the gain in interior job applications is driven solely by increased female application rates. This is all the more surprising since there are far more men than women in the sample (656,774 female vs. 1,142,194 male), so that this is not driven by differences in the number of observations.

different ($t = 1.49$). This is likely partially due to loss of sample size (N=589,580 for all those who view an internal posting vs. N=81,802 for all those who start an internal application) but may also be driven by selection.

[15] A back of the envelope calculation would be that 344,671 users viewed an exterior job posting on the first day of the experiment. If they had all been in the treatment group we would expect 8.902% (instead of 8.508% in the control) to start an application, which would be an increase of 1,358 started applications.

[16] The male and female coefficients in Panel B.ii are statistically significantly different from each other ($Prob > chi2 = 0.048$ for column 2 vs. 3; $Prob > chi2 = 0.023$).

Table 3: Likelihood of Starting/Finishing An Application

	Simple			With Fixed Effects		
	1	**2**	**3**	**4**	**5**	**6**
	A. Exterior: Likelihood Starting Application					
	All	**Female**	**Male**	**All**	**Female**	**Male**
Control Mean $\overline{A}_{T=0}$	8.508	8.044	8.419	8.605	8.634	8.498
Treatment β	0.394***	0.397***	0.383***	0.372***	0.409***	0.376***
	(0.052)	(0.089)	(0.069)	(0.055)	(0.103)	(0.076)
Adj. R2	0.000	0.000	0.000	0.049	0.045	0.051
Effective N	1,410,384	464,679	806,378	892,474	194,215	415,342
Pct Increase $\frac{\beta}{\overline{A}_{T=0}}$	4.631%	4.935%	4.594%	4.323%	4.737%	4.425%
	B.i Interior: Likelihood Starting Application					
	All	**Female**	**Male**	**All**	**Female**	**Male**
Control Mean $\overline{A}_{T=0}$	13.794	12.860	13.932	13.818	13.581	13.903
Treatment β	0.281**	0.361*	0.228	0.281**	0.372*	0.192
	(0.099)	(0.169)	(0.133)	(0.098)	(0.174)	(0.134)
Adj R2	0.000	0.000	0.000	0.050	0.056	0.050
Effective N	589,580	192,095	335,816	483,206	116,931	241,520
Pct Increase $\frac{\beta}{\overline{A}_{T=0}}$	2.037%	2.807%	1.637%	2.034%	2.739%	1.381%
	B.ii Interior: Likelihood Finishing Application					
	All	**Female**	**Male**	**All**	**Female**	**Male**
Control Mean $\overline{A}_{T=0}$	4.322	3.792	4.584	4.201	3.809	4.454
Treatment β	0.164**	0.338***	0.098	0.155**	0.359***	0.074
	(0.059)	(0.099)	(0.080)	(0.060)	(0.103)	(0.081)
Adj R2	0.000	0.000	0.000	0.015	0.028	0.023
Effective N	589,580	192,095	335,816	483,206	116,931	241,520
Pct Increase $\frac{\beta}{\overline{A}_{T=0}}$	3.795%	8.914%	2.138%	3.690%	9.425%	1.661%

Notes: The dependent variable takes the value 1 if a job seeker started or finished an application. All coefficients are multiplied by 100 for ease of reading results. Columns 1, 2 & 3 are simple models that only use the treatment as the right hand side variable. Columns 4, 5, & 6 include job posting fixed effects, days posted (omitted category 1st day) fixed effects and categorical dummies for the previous number of people who started a job application at time of viewing (omitted category is 1-24, other bins 25-49, 50-74, 75-99, 100-124, 125-149, 150-175, 175+). All standard errors clustered at the job posting level. The coefficients for male vs. female job seekers are only statistically significantly different from each other for panel B.ii ($Prob > chi2 = 0.0591$ for column 2 vs. 3; $Prob > chi2 = 0.0225$ for column 5 vs. 6). Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

The differences described so far may be driven by a number of other factors including selection of the job posting viewed, time of viewing, and actual number displayed to those in the treatment. In Columns 3-6 I explore these more complex relationships using the following model:

$$A_{ijt} = \beta T_i + P_j + D_t + \alpha NumPrevApply_{ijt} + \epsilon_{ijt}$$

The dependent variable A_{ijt} still takes the value 1 if a user decides to start or finish an application after

viewing the posting. The independent variable T_i takes the value one if a user was assigned to the treatment group which sees the number of previously started applications. I include a fixed effect P_j for each job posting j, so that the treatment identifies differences in likelihood of application between two members viewing the exact same posting.[17] This posting fixed effect controls for all time invariant attributes of the posting like: firm, industry, description, pay range and title. Additionally I use a fixed effect D_t for the number of days the posting has been live during the experiment because there are some time trends in the raw data (see online Appendix).

Last $NumPrevApply_{ijt}$ is a set of categorical variables that divide the number of previous applicants into eight bins: (1) 1-25, (2) 25-49, (3) 50-74, (4) 75-99, (5) 100-124, (6) 125-149, (7) 150-174 and (8) 175+. Keep in mind that even if a user is in the control group there is still a true number of previously started applications, but that information is simply not revealed to those in the control group. In the next section I will explore how that number shown vs. not shown varies the effectiveness of the treatment. For now I will flexibly control for the true underlying number with the categorical variable $NumPrevApply_{ijt}$.

Columns 3-6 represent the effect of the treatment while controlling for all the time invariant attributes of the job posting, the number of days the posting has been online, and the true number of previous applicants at the time of viewing. With these controls in place, I find that results are quite similar to those in Column 1-3. The treatment increases the likelihood a user will start or finish an application by 0.155-0.372 percentage points; representing a proportional increase above the control mean of between 2.034%-4.323%.

When I compare the results from female users in Column 5 to male users in Column 6, I find that the coefficient for female job seekers is always larger than for males. For example in Panel B.ii the treatment increases the likelihood a female users will finish an application by 0.359 percentage points, while the coefficient for male users is only 0.074; and furthermore this coefficient is not statistically significant for male users.[18] Not only are the coefficients different, but the proportional increase for women is quite sizeable. Being in the treatment group increases the likelihood a female job seeker will finish an application (Panel B.ii) by almost 10%.

[17] Only 44 job postings were seen by only a single person during the experiment so the fixed effects have a minimal effect on the effective sample size.

[18] The male and female coefficients in Panel B.ii are statistically significantly different from each other ($Prob > chi2 = 0.023$).

Result 1*: Showing job seekers the number of previously started applications increases the likelihood they will start or finish an application by 2% to 5%; this represents a potential increase in thousands of applications per day. The increase caused by the treatment is similar with or without controls for time invariant attributes of the job posting, the number of days the posting has been online, and the true number of previous applicants at the time of viewing.*

Result 2*: The increase in applications is largely driven by female job seekers being induced to start or finish an application. The size and significance of the coefficient on the treatment is always larger for female vs. male job seekers. For example being in the treatment group increases the likelihood a female job seeker will finish an application by almost 10%, whereas the effect on men is not statistically significantly different from 0.*

3.2 Treatment Effects By Number Shown

Intuitively one may believe that showing the number of previous applicants could have either a positive or negative effect on the likelihood of application, and that the size and sign of the effect may vary by the number shown. On the one hand, if job seekers want to avoid applying to postings with greater competition, we should see the effect of the treatment fall as the number shown rises. Yet on the other hand, if job seekers herd toward more popular postings, we should see the effect of the treatment rise as the number shown rises (see the online Appendix for a short discussion of herding models (Banerjee, 1992; Anderson and Holt, 1997)).

Survey evidence shows that people viewing the exact same number may have two different opinions on whether it signals high or low competition (see online Appendix). For example a person may see the that 26 people have already applied and they may believe that is a good sign they will get an offer and enjoy the position, while another person will believe that 26 people is too many. So it is an empirical question which of these effects is larger, and at what numbers shown are the effects most pronounced.

To test the size of the treatment effect by the number shown I use the following model:

$$A_{ijt} = \beta T_i + \gamma T_i * NumPrevApply_{ijt} + \alpha NumPrevApply_{ijt} + P_j + D_t + \epsilon_{ijt}$$

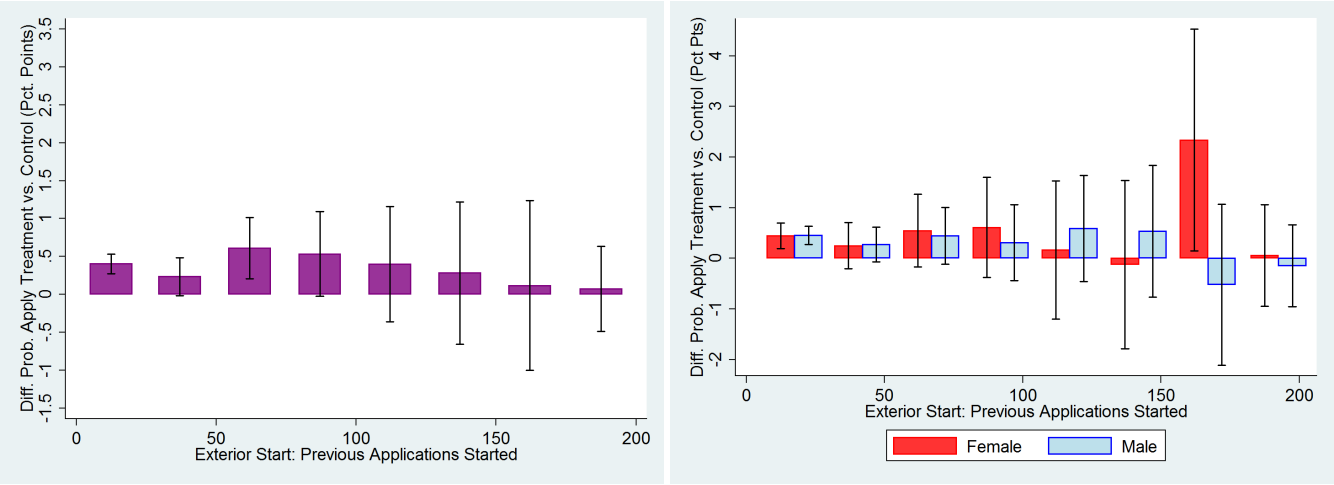
Again the dependent variable A_{ijt} takes the value 1 if a user decides to start or finish an application after viewing the posting. The independent variable T_i takes the value one if a user was assigned to the treatment group which sees the number of previously started applications. The treatment dummy T_i is interacted with a categorical variable for the number of previous applicants $NumPrevApply_{ijt}$. $NumPrevApply$ is a set of dummy variables that divide the number of previous applicants into eight bins (1) 1-24, (2) 25-49, (3) 50-74, (4) 75-99, (5) 100-124, (6) 125-149, (7) 150-174 and (8) 175+ (the omitted category is 1-24, so the coefficient β represents the effect for this category). I also include job posting fixed effects P_j and days posted fixed effects D_t.[19]

Figure 6 graphically represents the results from this model. On the vertical axis of Figure 6 is the percentage point difference in the likelihood of applying between the treatment and the control. On the horizontal axis is the true previous number of applicants that was either shown in the treatment, or not shown in the control. The error bars show the 95% confidence interval around each predicted difference.

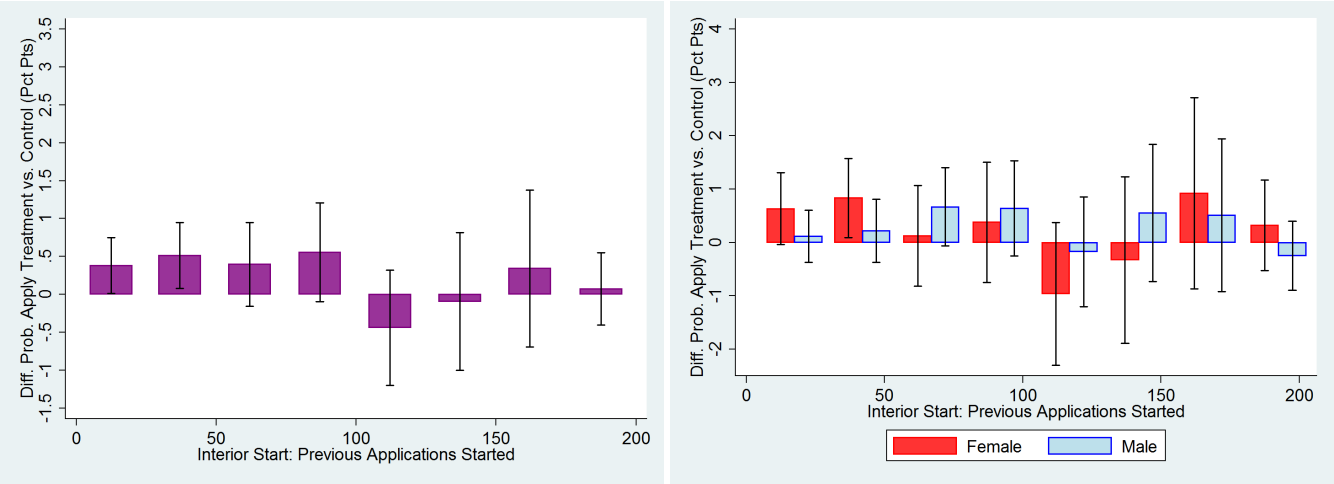
The top left-hand graph in Figure 6 shows the change in the effectiveness of the treatment for all users (male and female) who view an exterior job posting. The first bar states that the treatment increases the likelihood of applying by about 0.5 percentage points above the control group when users see between 1-24 previous applicants. The second bar says that the treatment increases the likelihood of applying by only 0.25 percentage points when users see between 25-49 previous applicants, and that this single point estimate is not statistically significantly different from 0. Looking at the pattern of the bars, there does not seem to be a strong upward or downward trend as the number shown increases. The estimates simply get more noisy as the number increases. This noise may be driven by having fewer observations where users see more than 100 previous applicants, but even concentrating on those bars where the estimates are more precise there isn't a strong upward pattern (herding) or downward pattern (congestion).[20]

[19] Results are similar if I include Days Posted X Number Shown fixed effects; these results are available upon request.

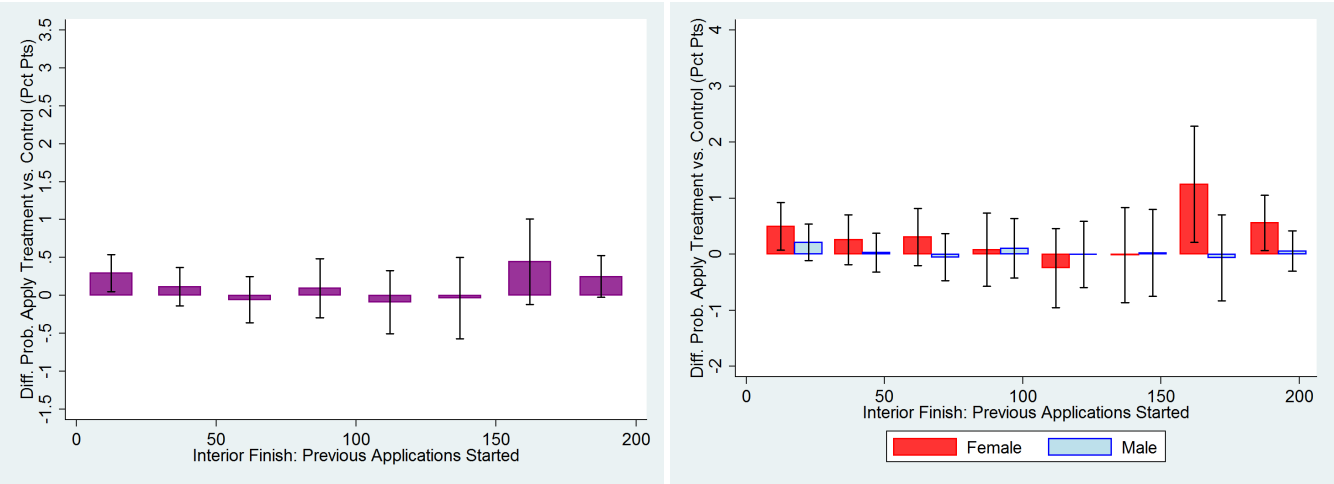
[20] There are actually a large number of observations in each bin. The bin with the fewest observations is those viewing a posting with 150-174 previous applicants, and here $N = 17,025$ for exterior postings and $N = 22,136$ for interior postings.



(a) Exterior: Starting Application



(b) Interior: Starting Application



(c) Interior: Finishing Application

Figure 6: Plots of Coefficients on Treatment Dummy Variable By Number Shown

The top right-hand graph in Figure 6 shows the change in the effectiveness of the treatment for all female vs. male users who view an exterior job posting. Again there doesn't seem to be a strong upward or downward pattern as the number shown increases. A similar lack of a pattern is found in panel (b) for starting an interior job posting and panel (c) for finishing the application for an interior job posting.

If users were avoiding competition we would expect a downward trend as the number shown increases. While if users were herding toward more popular positions we would expect an upward trend as the number shown increases. The fact that we observe neither could mean either that the two balance out, or that we do not have enough information about how each individual interprets the number she sees. Survey evidence indicates that both of these might be at play.

In June 2014 I administered an online survey to better understand what job seekers are thinking when they choose between a job posting with or without the number of previous applicants shown (see online Appendix for details). This was completely hypothetical, so the results are only presented to better understand the results from the field experiment. I found that 50% of respondents used the information to avoid competition, 22% used the information to herd toward more popular jobs, and the remaining 27% simply preferred having more information in-line with being ambiguity or risk averse. While the majority of respondents wanted to use the information to avoid competition, the number that was too much competition varied by person. For example, there were those who felt that 26 previous applicants was too many competitors while others thought this number was lower than expected. So each person had very different views of the same number shown. This means that if herding or competition avoidance was taking place during the field experiment it would be very difficult to detect since I cannot observe the beliefs of the job seekers. This is an important area of future research, but is not something that this paper can address. These findings are summarized in Result 3.

Result 3*: Given that we cannot observe the beliefs of individuals in the field study, from the observed data there is no strong evidence of congestion avoidance or herding toward more popular jobs.*

3.3 Treatment Effect By Job Type

Thus far, we have seen that the treatment of showing the number of previous applicants results in an increase in the likelihood of starting or finishing an application, and that increase is larger for female job seekers than male job seekers. But not all firms will be actively seeking more female applicants, for example about 1% of the jobs seen in the control have only female applicants. So it is important to know if the treatment is simply increasing the number of female applicants to "female jobs".

Let us define a "male job," M_{ijt}, as a job where over 80% of those who start the application in the control are male. M_{ijt} is only defined for those jobs which have at least one person who starts the job in the control, so I restrict the sample to those jobs that have at least one male or female person who starts an application for a job posting in both the treatment and the control.[21] Then I use M_{ijt} as the dependent variable to see the effect of the treatment on those positions that may be most likely to be seeking more female applicants with the following model:

$$M_{ijt} = \beta T_i + P_j + D_t + \alpha NumPrevApply_{ijt} + \epsilon_{ijt}$$

Table 4 reports the results from this model. Column 1 of Table 4 shows that overall the treatment has a positive effect on the likelihood that any person (male or female) will apply to a "male job". However, looking at columns 2 and 3, we see this is largely driven by a large increase in the proportion of female applicants applying to "male jobs". These are most likely the firms that want a larger female applicant pool, so this is further evidence of the effectiveness of the treatment in increasing female applicants in industries which are actively seeking to diversify their workforce.[22]

Result 4*: The treatment increases the number of female applicants to "male jobs".*

[21] This definition is for the outcome variable of *starting* an application. When defining M_{ijt} for *finishing* an application, then it is defined as a job with at least 80% males who finish the application. And in this case I restrict the sample to those jobs with at least one male or female person who finishes the application in the control and the treatment.

[22] The proportional gains are also quite large compared to the control (e.g. a 3.131 percentage point increase off a mean of 0.581 for female users in Panel A), but this is largely driven by the definition of the outcome variable as a job with greater than 80% male applicants in the control.

Table 4: Likelihood Apply to "Male" Job

	With Fixed Effects		
	1	**2**	**3**
A. Exterior: Likelihood Start App			
	All	**Female**	**Male**
Control Mean $\overline{M}_{T=0}$	4.611	0.581	7.105
Treatment β	2.078***	3.131***	1.310***
	(0.063)	(0.096)	(0.092)
Adj. R2	0.122	0.148	0.122
N	494,006	188,877	305,129
B.i Interior: Likelihood Start App			
	All	**Female**	**Male**
Control Mean $\overline{M}_{T=0}$	5.292	1.178	7.660
Treatment β	0.971***	1.789***	0.464***
	(0.073)	(0.092)	(0.106)
cons	6.180***	1.535***	8.777***
	(0.153)	(0.124)	(0.222)
Adj. R2	0.135	0.127	0.130
N	424,406	155,150	269,256
B.ii Interior: Likelihood Finish App			
	All	**Female**	**Male**
Control Mean $\overline{M}_{T=0}$	2.200	0.327	3.264
Treatment β	0.895***	1.504***	0.532***
	(0.064)	(0.099)	(0.093)
Adj. R2	0.046	0.054	0.044
N	226,968	82,281	144,687

Notes: The dependent variable takes the value 1 if a job seeker started or finished an application to a "male" job. A position is a "male" job if over 80% of the applicants in the control group are male. All coefficients are multiplied by 100 for ease of reading results. Includes job posting fixed effects, days posted (omitted category 1st day) fixed effects and categorical dummies for the previous number of people who started a job application at time of viewing (omitted category is 1-24, other bins 25-49, 50-74, 75-99, 100-124, 125-149, 150-175, 175+). All standard errors clustered at the job posting level. Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

4 Conclusion

Job search is costly for both the job seeker and the firm. Firms hope to maximize the number of applicants who meet their search criteria. Job seekers want to apply to any job posting where the expected marginal benefits are greater than the marginal costs. This paper uses a large scale field experiment with about two million real world job seekers to show that increasing the amount of information on a job posting causes more people to apply and specifically increases the number of female applicants to the jobs which may most need more female applicants. Since there are many campaigns to bring more women into the labor force, this paper illustrates a low cost, light touch intervention that increases the number of female applicants.

I find that showing the number of previous applicants on the job posting increases the liklihood of application by 2-5%. Since millions of job seekers view job posting each week on websites like LinkedIn the actual increase in the number of applications would be thousands per day. This paper only looked at the first job posting a person viewed during the experiment, so these estimates can be seen as a lower bound on the total increase given most job seekers view many postings during a job search. However, path dependence in this type of intervention is an important area for further research. Additionally the long term effects, if any, on unemployment duration and subsequent job tenure are another area I plan to address in future research.

It is interesting that showing the number of previous applicants has a strong positive effect on applications rates, because intuitively seeing this number might either decrease or increase the likelihood a person would apply. If a person sees 100 previous applicants is that a signal of a good quality position or of too much competition for the position (or both)? The pattern of behavior in this experiment is not consistent with job seekers herding toward more popular jobs or avoiding congestion. However, the overall positive treatment effect can be explained by models of ambiguity or risk aversion. Furthermore, consistent with findings that women tend to be more risk and ambiguity averse than men, much of the increase in applications is driven by an increase in female applicants.

References

Anderson, Lisa R. and Charles A. Holt, "Information Cascades in the Laboratory," *The American Economic Review*, 1997, *87* (5), pp. 847–862.

Babcock, Linda and George Loewenstein, "Explaining Bargaining Impasse: The Role of Self-Serving Biases," *The Journal of Economic Perspectives*, 1997, *11* (1), pp. 109–126.

Banerjee, Abhijit V., "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, 1992, *107* (3), 797–817.

Bradley, Gifford W., "Self-serving biases in the attribution process: A reexamination of the fact or fiction question," *Journal of Personality and Social Psychology*, 1978, *36* (1).

Braunstein, Yale M. and Andrew Schotter, "Labor Market Search: An Experimental Study," *Economic Inquiry*, 1982, *20* (1), 133–144.

Chade, Hector and Lones Smith, "Simultaneous Search," *Econometrica*, 2006, *74* (5), 1293–1307.

Chow, Clare Chua and RakeshK. Sarin, "Known, Unknown, and Unknowable Uncertainties," *Theory and Decision*, 2002, *52* (2), 127–138.

Croson, Rachel and Uri Gneezy, "Gender differences in preferences," *Journal of Economic literature*, 2009, pp. 448–474.

Das, Sanmay and John N. Tsitsiklis, "When is it important to know youve been rejected? A search problem with probabilistic appearance of offers," *Journal of Economic Behavior and Organization*, 2010, *74* (12), 104 – 122.

Eckel, Catherine C. and Philip J. Grossman, "Chapter 113 Men, Women and Risk Aversion: Experimental Evidence," in Charles R. Plott and Vernon L. Smith, eds., , Vol. 1 of *Handbook of Experimental Economics Results*, Elsevier, 2008, pp. 1061 – 1073.

Eil, David and Justin M. Rao, "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself," *American Economic Journal: Microeconomics*, 2011, *3* (2), 114–38.

Ellsberg, Daniel, "Risk, Ambiguity, and the Savage Axioms," *The Quarterly Journal of Economics*, 1961, *75* (4), pp. 643–669.

Fernandez, Roberto M. and Collette Friedrich, "Gender Sorting at the Application Interface," *Industrial Relations: A Journal of Economy and Society*, 2011, *50* (4), 591–609.

Galenianos, Manolis and Philipp Kircher, "Directed search with multiple job applications," *Journal of Economic Theory*, 2009, *144* (2), 445 – 471.

Gaucher, Danielle, Justin Friesen, and Aaron C. Kay, "Evidence That Gendered Wording in Job Advertisements Exists and Sustains Gender Inequality," *Journal of Personality and Social Psychology*, 2011, *1* (101).

Gilboa, Itzhak and David Schmeidler, "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 1989, *18* (2), 141 – 153.

Halevy, Yoram, "Ellsberg Revisited: An Experimental Study," *Econometrica*, 2007, *75* (2), 503–536.

_ **and Vincent Feltkamp**, "A Bayesian Approach to Uncertainty Aversion," *The Review of Economic Studies*, 2005, *72* (2), 449–466.

Jianakopolos, Nancy A. and Alexandra Bernasek, "ARE WOMEN MORE RISK AVERSE?," *Economic Inquiry*, 1998, *36* (4), 620–630.

Kohn, M. and S. Shavaell, "The Theory of Search," *Journal of Economic Theory*, 1974.

Moore, Evan and Catherine Eckel, "Measuring Ambiguity Aversion," *Working Paper*, 2003.

Mortensen, Dale T., "Job Search, the Duration of Unemployment, and the Phillips Curve," *The American Economic Review*, 1970, *60* (5), pp. 847–862.

Nachman, David, "On Risk Aversion and Optimal Stopping," *Working Paper*, 1972.

Oberholzer-Gee, Felix, "Nonemployment stigma as rational herding: A field experiment," *Journal of Economic Behavior & Organization*, 2008, *65* (1), 30 – 40.

Orazem, Peter F. and J. Peter Mattila, "Male-Female Supply to State Government Jobs and Comparable Worth," *Journal of Labor Economics*, 1998, *16* (1), pp. 95–121.

Petit, Pascale, "The effects of age and family constraints on gender hiring discrimination: A field experiment in the French financial sector," *Labour Economics*, 2007, *14* (3), 371 – 391.

Schotter, Andrew and Yale M. Braunstein, "Economic Search: An Experimental Study," *Economic Inquiry*, 1981, *19* (1), 1–25.

Small, Deborah A, Michele Gelfand, Linda Babcock, and Hilary Gettman, "Who goes to the bargaining table? The influence of gender and framing on the initiation of negotiation.," *Journal of personality and social psychology*, 2007, *93* (4), 600.

Stigler, George J., "The Economics of Information," *Journal of Political Economy*, 1961, *69* (3), pp. 213–225.

Telser, L. G., "Searching for the Lowest Price," *The American Economic Review*, 1973, *63* (2), pp. 40–49.

van Ours, Jan and Geert Ridder, "Vacancies and the Recruitment of New Employees," *Journal of Labor Economics*, 1992, *10* (2), pp. 138–155.

Weitzman, Martin L., "Optimal Search for the Best Alternative," *Econometrica*, 1979, *47* (3), pp. 641–654.