Job Credentials in the Labor Market: Pre-Analysis Plan

Lisa Abraham

May 20, 2018

Abstract

Women are underrepresented in certain jobs relative to men, which may be partly due to the composition of the applicant pool. There is also anecdotal evidence that women are less likely than men to apply for jobs for which they do not meet all the posted credentials, perhaps leading to gender gaps in job applications and subsequent hiring. We run a randomized control trial (RCT) with a company to test this in the field. In our RCT, we examine whether deleting optional credentials and reframing language in the company's job postings could encourage more women to apply and progress through the application process. This document describes the pre-analysis plan for our RCT.

1 Introduction

An important component of gender differences in the labor market stems from inequities in representation. There is evidence that females are underrepresented in certain sectors of the U.S. economy and in higher level positions. While some of this may be attributable to different preferences for task content and culture, another potential explanation focuses on pipeline challenges, particularly at the point of job application.

Though there is research examining factors that influence job application decisions (e.g., advertised amenities), as well as research on gender differences in confidence, there is less work investigating how gender differences in beliefs about meeting all the listed credentials in a job application interacts with a female's decision to apply for the job, particularly a traditionally "masculine" or highly technical job. The best evidence is anecdotal in nature: Sheryl Sandberg's *Lean In* highlights an internal Hewlett-Packard report which found that males apply for open jobs when they meet 60% of the credentials, whereas females only apply if they meet 100% of the listed credentials. While this descriptive anecdote has been referenced in numerous media sources, there are no rigorous economic analyses confirming this fact or analyzing potential solutions. Our RCT examines whether deleting optional credentials and reframing language in job postings could encourage more women to apply and progress through the application process.

2 Experimental Design

2.1 Sample

We have partnered with Uber to implement our RCT. As of May 2018, Uber advertised approximately 800 job postings on their career website (https://www.uber.com/careers), comprising both technical and non-technical roles. Individuals can apply to Uber by going directly to the career website, or by linking to the career website from a job posting on a 3rd party website (e.g., LinkedIn, Glassdoor, Indeed). Our sample consists of individuals who *directly* view a job posting on Uber's career website. We exclude individuals that link to the career website from a 3rd party site,[1] as well as referrals and external hires (given the different nature of the referral / external hiring process). We consider the first three weeks of the RCT a pilot; we plan on including data from the pilot in the analyses, unless we there is a problem with the execution of the experiment.

2.2 Design

The experiment randomizes individuals who directly[2] visit the Uber career website. We identify individuals by their IP address and randomly sort these individuals[3] into one of two groups: (1) a Control group where individuals see the original version of the job posting, or (2) a Treatment group, where individuals see a version of the job posting which deletes optional credentials, deletes adjectives describing skills, and reframes vague credentials. Treatment or Control status is maintained across all jobs an individual views. Through the use of cookies, we are also able to serve the same treatment to an individual if they return to the website at a later date.

Upon seeing the job posting on the Uber career website, individuals decide whether or not to apply for the position. For those who decide to apply, the first step is to click on the "Apply

[1] We are unable to include individuals that link from a 3rd party website in our RCT due to technical constraints.

[2] We can identify and eliminate individuals who link to the career website from a 3rd party website.

[3] Randomization is at the IP address level, but for ease of exposition we will refer to randomization as being at the "individual" level.

Now" button, after which the individual is given several options for applying: (1) by signing in (if they already have an account on the career website); (2) by logging into their LinkedIn profile (so that data from the individual's LinkedIn profile is transmitted to the Company); (3) by submitting their resume; (4) by manually by filling out responses to a series of questions.

We are able to track the treatment status of the applicants in our experiment at all stages of the hiring process (i.e., application, initial phone screen, technical phone screen, on-site interview, and offer). If an individual applies to more than one job posting over the duration of the experiment, we are able to examine the portfolio of their job application choices.

2.3 Treatment

Our RCT only alters language in the section of the job posting listing required credentials and desired credentials. In some job postings, desired credentials are interspersed in the required list (e.g., "Python experience is a plus"), while other job postings break out the desired credentials in a separate section. The required credentials are labeled "What You'll Need," (or a variant of this) and the desired credentials, when broken out, are labeled "Bonus Points" (or a variant of this).

Treatment varies significantly depending on the job posting. While some job postings have a lot of optional credentials (e.g., "PhD preferred"), adjectives describing skills (e.g., "Excellent" before "coding skills"), and vague credentials (e.g., "think like your enemy"), others have less scope to be edited. Given the heterogeneity in treatment, for every job posting included in our experiment, we document the original version of the credentials and the treated version of the credentials. We also indicate how many adjectives were deleted, how many vague credentials were reframed, how many optional credentials were deleted, and whether the optional credentials were interspersed in the required credentials or broken out in a separate section. There are different ways that these edits can be classified,[4] and the strength of the treatment across jobs could also be a function of the rest of the text in the job posting, so our classification represents an initial attempt at quantifying the treatment. We plan to have workers on Mechanical Turk rate the job postings on various dimensions after the RCT concludes as this might provide a more accurate

[4]In our initial classification, *Adjectives Deleted* represent the number of deleted adjectives (at the word level) which are not included in a vague credential or optional credential, *Vague Credentials Reframed* represent the number of vague credentials that are reframed (at the phrase level), and *Optional Credentials Deleted* represent the number of optional credentials that are deleted (at the bullet point level).

measure than the quantitative metrics. We also preserve the text of the Treatment and Control versions for alternative classification and finer textual analysis.

2.4 Slippage

We anticipate that there might be some slippage in assignment to treatment, given the fact that individuals can view an Uber job posting on a 3rd party website, and later find that same job posting directly on the career website; this would result in slippage if the version the individual sees on the 3rd party website does not match the version the individual is randomly assigned on the career website. This could be partly mitigated if the credentials only become salient to individuals when they are deciding whether or not to apply.

Slippage might also occur if individuals access the career website from multiple, different devices (e.g., an ipad and a computer), since randomization is at the IP address level and each device has its own IP address. We believe this is somewhat mitigated by the relative ease of the applying (detailed above), which makes it unlikely that an individual is repeatedly viewing the same posting from different devices over the brief span of time he / she decides whether or not to apply.

There is also the potential for slippage due to the presence of "faulty" job links on Uber's career website. These faulty job links are those which have extra digits in the job requisition ID, representing approximately 10% of all U.S. job postings. When browsing the career website, these links sometimes present a cached version of the last job the individual viewed on the career website, as opposed to the actual job associated with the extra digits. While this does not affect the application process (i.e., individuals apply for the job they view), it might interact with the A/B testing system in a manner that induces some slippage in treatment. We are in the process of confirming whether this is the case.

3 Empirical Strategy

3.1 Individual level Regressions

The main specification is the regression at the individual level. This regression answers the question: is there a difference between the overall fraction of female applicants out of all applicants in the Treatment and Control conditions? The regression is given by:

$$I(Female)_i = \alpha + \beta Treated_i + \epsilon_i$$

where *I(Female)*$_i$ is an indicator for whether person i is female and $Treated_i$ is a dummy variable for treatment status of the individual. β is the percentage point change in the fraction of female applicants between the Treatment and Control groups (before controlling for any other variables). This analysis can be repeated at all other stages of the interview process (i.e., we can examine phone screens, on-site interviews, and offers instead of applicants).

Given the heterogeneity in treatment strength, we also consider a different specification that interacts the "intensity" of the treatment with treatment status. This regression is given by:

$$I(Female)_i = \alpha + \beta Treated_i + \phi Intensity_i + \sigma(Treated * Intensity)_i + \epsilon_i$$

where $Intensity_i$ is a variable denoting the intensity of the treatment.

3.2 Job level Regressions

We can also run a regression at the job-treatment level. This regression answers the question: does the overall fraction of females increase between the Treatment and Control conditions after collapsing to the job-treatment level and accounting for the number of applicants in each job-treatment cell? This regression examines treatment effect heterogeneity and is given by:

$$\%\ Female\ Applicants_{j,t} = \alpha + \beta Treated_{j,t} + \lambda_j + \epsilon_{j,t}\ [aweight = \#applicants_{j,t}]$$

where *% Female Applicants*$_{j,t}$ indicates the fraction of female applicants out of all applicants in job j and treatment condition t, $Treated_{j,t}$ is a dummy variable for treatment status (note that each job has both a treatment and control condition), and λ_j represents job fixed-effects. β is the percentage point change in the fraction of female applicants between the Treatment and Control groups after collapsing to the job-treatment level and weighting by the number of applicants in the job-treatment cell. We also plan to run this reqression without weighting by by the number of applicants in the job-treatment cell; doing so weights all jobs equally. This analysis can also be repeated at all other stages of the interview process (i.e., we can examine phone screens, on-site interviews, and offers instead of applicants).

As with the individual level regression, we also consider a different specification that interacts the "intensity" of the treatment with treatment status. This regression is given by:

$$\% \ Female \ Applicants_{j,t} = \alpha + \beta Treated_{j,t} + \lambda_j + \phi Intensity_{j,t} + \sigma(Treated * Intensity)_{j,t}$$
$$+\epsilon_{j,t} \ [aweight = \#applicants_{j,t}]$$

where $Intensity_{j,t}$ is a variable denoting the intensity of the treatment.

3.3 Other Analyses

3.3.1 Total Number of Applicants

While the primary focus on the experiment is on gender differences, we can also examine whether more individuals apply and reach subsequent stages of the application process via the Treatment condition versus the Control condition; differences between the Treatment and Control condition (irrespective of gender composition) could be indicative of potential inefficiencies in screening.

3.3.2 Conversion Rates

We can examine conversion rates from application to subsequent stages of the interview process (e.g., general phone screen, technical phone screen, on-site interview, and offer). This analysis addresses how the quality of applicants in the Treatment and Control conditions differs (at the aggregate level and by gender).

3.3.3 Portfolio of Job Applications

If an individual applies to more than one job over the duration of the RCT, we can examine the portfolio of jobs he/she applies to. This analysis has the potential to inform our understanding of how listed job credentials interact with job search.

4 Variables of Interest

4.1 Outcome Variables

To execute our analyses, we examine the following outcomes:

1. Gender of applicants
2. Gender of initial phone screens
3. Gender of technical phone screens
4. Gender of on-site interviews

5. Gender of offers
6. Number of applicants
7. Number of initial phone screens
8. Number of technical phone screens
9. Number of on-site interviews
10. Number of offers
11. Data on applicant quality (parsed from candidate resumes / application information)
12. Portfolio of jobs an individual applies to