The Information Effects of Salary Disclosure on Application Rates During Job Postings

Robert Zhang, Shu Wang, Zheman Zhong, Sheng Zhou, Guang Yang

03/15/2023

Introduction

In labor economics, researchers use the "*job search and matching theory*" framework to improve the underlying process of job matching efficiency, application rates, and diversity of applicant pools (Yashiv, 2007). Since finding the ideal match is important for both firms and candidates, how information is displayed on job posting platforms like LinkedIn and Google Jobs could impact overall application rate (Gee, 2017). Past research has suggested adding information regarding other people's actions (e.g. how many people clicked a button to start an application) will increase the likelihood of individual job application rate by about 1.9% - 3.6% (Gee, 2017). A recent phenomenon in online job postings is the disclosure of salary range to narrow the gap of asymmetric information between employers and potential candidates. **Our research goal is to assess the effects of pertinent salary information on overall application rates.** We find this experiment interesting since it is unclear whether disclosing salaries will create a positive (pay range exceeds expectation) or negative (pay range below expectation) signal to candidates. Given the ambiguous nature of this experiment's outcome, we decide to use Qualtrics survey experiment to observe the impact of salary disclosure on overall job application rates.

Numerous studies have been done in the demand-side of the job market to analyze how applicant's salary history shapes wage offers or how employers have negative inferences about non-disclosing candidates (Agan, Cowgill, Gee, 2021). However, our research will focus on the supply-side of the market by studying how different information of job postings impact candidates' application rate. We hope our experiment's outcome can help policymakers and employers create a more efficient labor market to find the ideal match effectively.

Research Question: Does disclosing salary range on job postings increase candidates' overall application rate?

Hypothesis: The existence of salary information on job postings increases candidates' overall application rate for that particular job.

Data Analysis & Insights

Data Importing

```
library(data.table)
library(ggplot2)
library(tidyverse)
library(pwr)
```

```
library(modelsummary)
library(readr)
library(tidyr)

#original data entering & cleaning
data <- fread('BA830 Experiment.csv')
```

Modifying Dataset

```
#creating dummy variable: Treatment=1, Control=0;
#treat_1=1 if treated, salary_1=1 if choose salary info job posting
data$treat_1 <- ifelse(data$`Treatment 1` != '',1,0)
data$salary_1 <- ifelse(data$`Treatment 1` == 'Job A (Top)'
                        | data$`Control 1` == 'Job A (Top)',1,0)
data$treat_2 <- ifelse(data$`Treatment 2` != '',1,0)
data$salary_2 <- ifelse(data$`Treatment 2` == 'Job A (Top)'
                        | data$`Control 2` == 'Job A (Top)',1,0)
data$treat_3 <- ifelse(data$`Treatment 3` != '',1,0)
data$salary_3 <- ifelse(data$`Treatment 3` == 'Job A (Top)'
                        | data$`Control 3` == 'Job A (Top)',1,0)
data$treat_4 <- ifelse(data$`Treatment 4` != '',1,0)
data$salary_4 <- ifelse(data$`Treatment 4` == 'Job A (Top)'
                        | data$`Control 4` == 'Job A (Top)',1,0)
data$treat_5 <- ifelse(data$`Treatment 5` != '',1,0)
data$salary_5 <- ifelse(data$`Treatment 5` == 'Job A (Top)'
                        | data$`Control 5` == 'Job A (Top)',1,0)
data$treat_6 <- ifelse(data$`Treatment 6` != '',1,0)
data$salary_6 <- ifelse(data$`Treatment 6` == 'Job A (Top)'
                        | data$`Control 6` == 'Job A (Top)',1,0)
data$treat_7 <- ifelse(data$`Treatment 7` != '',1,0)
data$salary_7 <- ifelse(data$`Treatment 7` == 'Job A (Top)'
                        | data$`Control 7` == 'Job A (Top)',1,0)
data$treat_8 <- ifelse(data$`Treatment 8` != '',1,0)
data$salary_8 <- ifelse(data$`Treatment 8` == 'Job A (Top)'
                        | data$`Control 8` == 'Job A (Top)',1,0)
data$treat_9 <- ifelse(data$`Treatment 9` != '',1,0)
data$salary_9 <- ifelse(data$`Treatment 9` == 'Job A (Top)'
                        | data$`Control 7` == 'Job A (Top)',1,0)
data$treat_10 <- ifelse(data$`Treatment 10` != '',1,0)
data$salary_10 <- ifelse(data$`Treatment 10` == 'Job A (Top)'
                         | data$`Control 10` == 'Job A (Top)',1,0)
```

Data Cleaning

Preliminary screening of dataset was conducted in Excel by eliminating null and uncompleted responses; the filtered dataset file is named as `data_use_t.csv`, which is imported here.

```
#import cleaned data
data_cleaned <- fread('data_use_t.csv')
```

```r
#rename column
data_cleaned = rename(data_cleaned, treatment = treat_1, outcome = salary_1)
knitr::kable(head(data_cleaned))
```

V1	Duration (in seconds)	Education	Gender	Age	Work	treatment	outcome
1	132	Masters	Female	23-27	No	1	1
2	795	Masters	Female	23-27	No	0	0
3	127	Masters	Male	23-27	No	0	1
4	101	Masters	Male	18-22	Yes	0	1
6	103	Masters	Female	23-27	Yes	1	1
7	199	Masters	Male	23-27	No	0	1

Data Dictionary

Key Variables	Meaning
Duration	Total response time, measured in seconds
Education	4 types: Bachelor, Masters, PhD, and Others
Gender	Male, Female, Non-binary
Age	3 ranges: 18-22, 23-27, >27
Work	Yes and No depends on work experience
treatment	=1 if received treatment question
outcome	=1 if choose jobs with salary

Research Methodology

Procedure

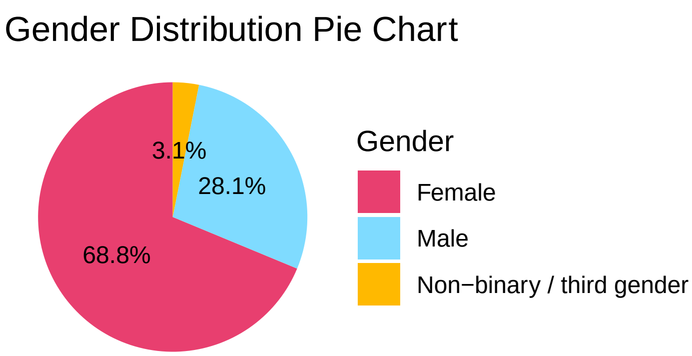
Regarding the survey design for our experiment, we first decided on the types of job listings we wanted to include in our survey. As we are all students in the Master of Business Analytics program at Boston University and our survey respondents are mostly MSBA students, we chose analyst-related jobs such as data analyst, business analyst, marketing analyst, etc., which are more relevant to our survey respondents. We also included five different industries - tech, consulting, banking, marketing, finance, and healthcare - to reduce industry bias within our survey respondents. Our goal was to have 60 respondents, each of whom would answer a total of 10 treatment and control questions, followed by questions regarding their educational level, gender, age, and full-time work experience.

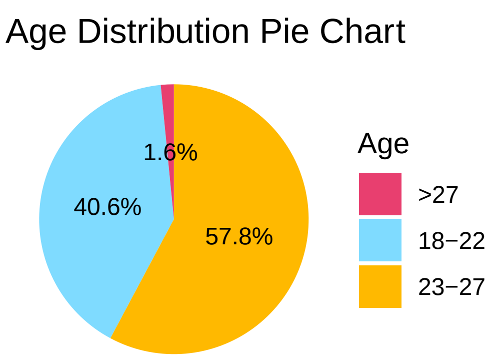
Essentially, we had 10 sets of questions in our survey design. In each set of questions, we had two similar job listings from two companies with similar reputation in their respective industries. We created two versions for each question, one is the treatment version and the other is the control version. Within the treatment version questions, only one of the job listings was shown with the salary range, while both of the job listings in the control version questions were shown without the salary range. We used the Qualtrics platform to create our survey and utilized the randomization function in Qualtrics to evenly display either treatment version or control version in each set of questions. As a result, each respondent answered 10 questions in total, and in each question, they would be assigned with either treatment or control version. Respondents were asked to indicate which job they would most likely apply to. Finally, we performed t-test and regression analysis to see whether disclosing the salary range on job listings would affect candidates' application rates.

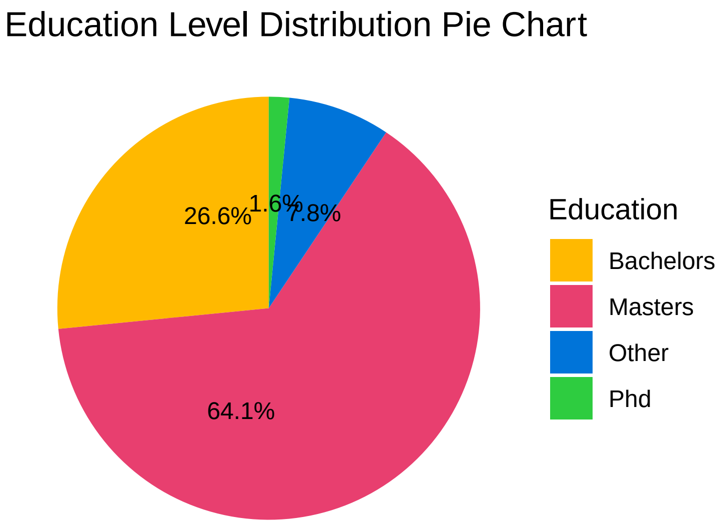
Participants

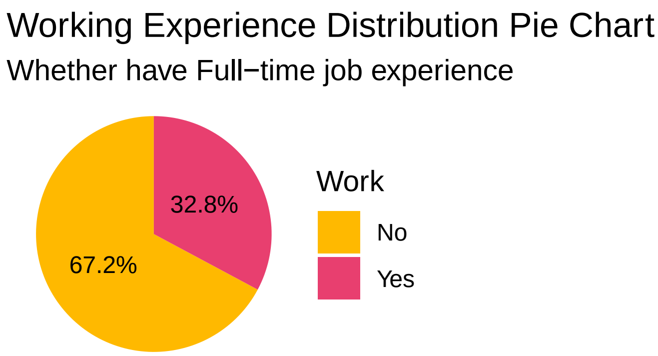
Since our research goal is to assess the effects of pertinent salary information on overall application rates and analyst-relevant jobs in different industries are included when designing the survey, we decided to target all demographics participants whose major is related to business analytics, data analytics and other relevant majors. We mainly recruit participants by directly sending the survey to our friends and classmates. Besides, we posted our survey on some social networking platforms such as Wechat, Instagram and so on in order to reach more people. All participants were required to fulfill a Qualtrics survey that started from treatment and control questions shown randomly and ended with demographic information questions including gender, age, education level and full-time working experience. Finally, we got 70 responses in total, among which 64 are valid responses, so the data includes 640 valid observations when considering the experiment in question-level.

Following pie charts show the distribution of the demographic variables, including gender, age, education level and working experience.









Randomization

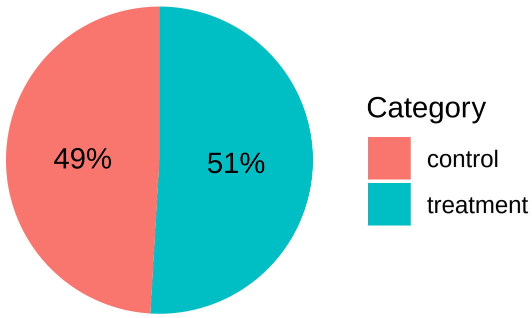
To test the hypothesis, We conduct a **question-level A/B testing** where we randomly assign job postings to each respondent:

Treatment - job postings disclose salary range,

Control - job postings don't disclose salary range;

When conducting our survey experiment on Qualtrics, we utilized the *Question Randomization* feature to assign treatment and control questions to participants. Each question contains unique job postings and participants are randomly assigned to either treatment or control on the question level. The result is approximately an even split between both groups (49% control and 51% treatment).

```
##    treatment   n
## 1:         0 314
## 2:         1 326
```



Estimated Average Treatment Effect (ATE)

```
treat_ratio <- c()
for (k in 1:10){
  treat_ratio[k] = nrow(data%>%filter(data[[2*k+2]] == 'Job A (Top)')) /
    +nrow(data%>%filter(data[[2*k+2]] != ''))}
treat_ratio
```

```
##  [1] 0.8000000 0.7741935 0.6562500 0.7941176 0.7878788 0.7575758 0.5151515
##  [8] 0.6666667 0.4242424 0.4848485
```

```
cont_ratio <- c()
for (k in 1:10){
  cont_ratio[k] = nrow(data%>%filter(data[[2*k+3]] == 'Job A (Top)'))/
    +nrow(data%>%filter(data[[2*k+3]] != ''))}
cont_ratio
```

```
##  [1] 0.5666667 0.7272727 0.4545455 0.7333333 0.6451613 0.9032258 0.3333333
##  [8] 0.6250000 0.2258065 0.2903226
```

```
#calculate estimated ATE
diff_in_ratio = treat_ratio - cont_ratio
#ATE for each question
treat_effect <- data.frame(treat_ratio,cont_ratio,diff_in_ratio)
ATE = mean(treat_effect$diff_in_ratio)
paste("ATE is", round(ATE,3))
```

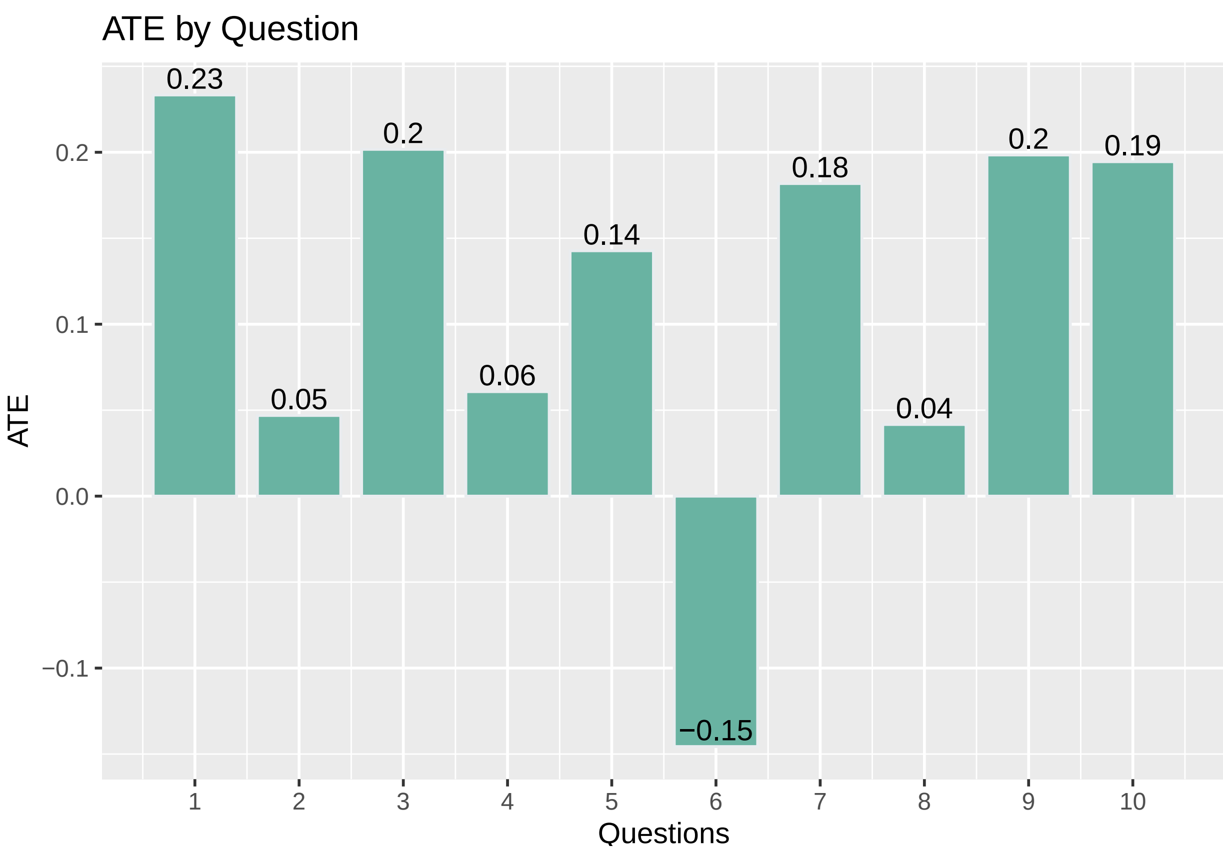
```
## [1] "ATE is 0.116"
```

```
knitr::kable(treat_effect, digits = 2,
             col.names = c("Treatment Ratio", "Control Ratio", "ATE"))
```

Treatment Ratio	Control Ratio	ATE
0.80	0.57	0.23
0.77	0.73	0.05
0.66	0.45	0.20
0.79	0.73	0.06
0.79	0.65	0.14

Treatment Ratio	Control Ratio	ATE
0.76	0.90	-0.15
0.52	0.33	0.18
0.67	0.62	0.04
0.42	0.23	0.20
0.48	0.29	0.19

```r
#visualize ATE by question number, display outlier (Q6)
dodgewidth <- position_dodge(width=0.9)
ggplot(treat_effect, aes(x=c(1:10), y=diff_in_ratio)) +
  geom_bar(stat = "identity", fill="#69b3a2", color="#e9ecef", width = 0.8) +
  labs(x = "Questions", y = "ATE", title = "ATE by Question") +
  scale_x_continuous(breaks=seq(1,10,1)) +
  geom_text(aes(label = round(diff_in_ratio,2), x=c(1:10), y=diff_in_ratio),
            position = position_dodge(width = 0.8), vjust = -0.3)
```



Based on the above table, 6 out of 10 questions (1, 3, 5, 7, 9, 10) show significant differences in percentage of participants choosing Job A in treatment and control groups. This is important because Job A is the one with salary range information. 3 out of 10 questions (2, 4, 8) show slight difference in treatment and control groups. The only outlier is Question 6 where the control group exhibits greater likelihood in choosing Job A compared to the treatment's proportion, creating a -0.15 ATE. This may be caused by the specificity of job posting or our small sample size, although it remains unclear what actually caused such an outlier.

Overall Treatment Effect

As our survey is randomized at question level, it would be reasonable for us to examine the treatment effect on a aggregated level by viewing each response to a question as an observation. In this part, we will be using 640 observations in our analysis.

Treatment Effect by sample-mean test

```
t.test(data_cleaned[treatment ==1,outcome], data_cleaned[treatment==0,outcome])
```

```
##
##  Welch Two Sample t-test
##
## data:  data_cleaned[treatment == 1, outcome] and data_cleaned[treatment == 0, outcome]
## t = 3.7309, df = 631.35, p-value = 0.000208
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.06767795 0.21808658
## sample estimates:
## mean of x mean of y
## 0.6779141 0.5350318
```

Treatment Effect by regression

```
reg_treat <- feols(outcome ~ treatment,data = data_cleaned,se = 'hetero')
etable(reg_treat)
```

```
##                           reg_treat
## Dependent Var.:             outcome
##
## Constant        0.5350*** (0.0282)
## treatment       0.1429*** (0.0383)
## _______________ __________________
## S.E. type       Heteroskedas.-rob.
## Observations                   640
## R2                         0.02140
## Adj. R2                    0.01987
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We can see that the result of treatment is significant at 5% significance level. The average percentage of people choosing Job A is 67% in treatment group vs. 54% in control group. The standard error is 0.03 and 0.04 respectively.

Conditional Average Treatment Effect

We are thinking that treatment effect may differs between experienced workers and junior workers so we conduct a conditional treatment effect on these two groups.

```
experienced <- data_cleaned[Work == 'Yes']
junior <- data_cleaned[Work == 'No']
exp_ate <- experienced[treatment == 1, mean(outcome)] -
```

```
  experienced[treatment == 0, mean(outcome)]
print(exp_ate)
```

```
## [1] 0.1962719
```

```
jun_ate <- junior[treatment == 1, mean(outcome)] -
  junior[treatment == 0, mean(outcome)]
print(jun_ate)
```

```
## [1] 0.1196123
```

Experienced workers have a higher treatment effect compared to non-experienced workers. They might pay more attention to disclosed salary information.

As shown in the response duration chart, most of the responses are completed within 500 seconds, which is around 8.3 minutes (within our expectation).

Randomization Check

```
reg1 <- feols(treatment ~ Education, data = data_cleaned)
reg2 <- feols(treatment ~ Age, data = data_cleaned)
reg3 <- feols(treatment ~ Gender, data = data_cleaned)
reg4 <- feols(treatment ~ Work, data = data_cleaned)
```

```
modelsummary(list("Treatment" = reg1,
                  "Treatment" = reg2,
                  "Treatment" = reg3,
                  "Treatment" = reg4
                  ), type = 'text', stars = T, output = 'markdown', gof_map = c("nobs", "r.squa
```

	Treatment	Treatment	Treatment	Treatment
(Intercept)	0.553***	0.400*	0.520***	0.493***
	(0.038)	(0.158)	(0.024)	(0.024)
EducationMasters	-0.048			
	(0.046)			
EducationOther	-0.133+			
	(0.080)			
EducationPhd	-0.153			
	(0.163)			
Age18-22		0.119		
		(0.161)		
Age23-27		0.105		
		(0.161)		
GenderMale			-0.020	
			(0.044)	
GenderNon-binary / third gender			-0.170	
			(0.114)	
WorkYes				0.050
				(0.042)
Num.Obs.	640	640	640	640

	Treatment	Treatment	Treatment	Treatment
R2	0.005	0.001	0.004	0.002

Note: ^^ + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

None of the covariates that we added shows significant effect at 5% significance level on treatment/control randomization.

Balance Check

```
num_treat <- dim(data_cleaned[treatment==1])[1]
num_obs = nrow(data_cleaned)
proportion_r = 0.5
prop.test(num_treat, num_obs, p = proportion_r)
```

```
##
##  1-sample proportions test with continuity correction
##
## data:  num_treat out of num_obs, null probability proportion_r
## X-squared = 0.18906, df = 1, p-value = 0.6637
## alternative hypothesis: true p is not equal to 0.5
## 95 percent confidence interval:
##  0.4699267 0.5487092
## sample estimates:
##        p
## 0.509375
```

At p-value of 0.66,we fail to reject the null hypothesis that the proportion of observations in treatment group is 50%. Hence the randomization is done properly.

Adding Fixed effects/covariates

```
reg_5 <- feols(outcome ~ treatment + Education,data = data_cleaned,se = 'hetero')
etable(reg_5)
```

```
##                                reg_5
## Dependent Var.:              outcome
##
## Constant          0.4849*** (0.0435)
## treatment         0.1443*** (0.0384)
## EducationMasters    0.0739. (0.0444)
## EducationOther       0.0545 (0.0796)
## EducationPhd        -0.1426 (0.1581)
## ________________ __________________
## S.E. type        Heteroskedas.-rob.
## Observations                     640
## R2                           0.02816
## Adj. R2                      0.02204
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In order to get more accurate regression results, we add one covariate which is shown to be most significant within four features that we considered in the dataset. There's a slight increase in treatment effect (from 0.143 to 0.144) and 'Masters' in Education level is significant at 10% significance level.

Statistical Power

```
ATE = data_cleaned[treatment == 1, mean(outcome)] -
  data_cleaned[treatment == 0, mean(outcome)]
cohen_d <- ATE/(sd(data_cleaned[,outcome]))

pwr.t2n.test(n1 = 326, n2 = 314, d = cohen_d,

sig.level = .05, power = NULL)
```

```
##
##      t test power calculation
##
##              n1 = 326
##              n2 = 314
##               d = 0.29242
##       sig.level = 0.05
##           power = 0.9584225
##     alternative = two.sided
```

The statistical power means we are less likely to make type II error. In other words, we obtain a large chance to detect the true effect of treatment. This is because our sample size is big enough.

Research Limitations

1.Sample size:

The purpose of this study was to investigate the effect of salary disclosure on job postings and overall application rates. The study was conducted using an online survey, which included questions related to job search behaviors and preferences, as well as demographic information. The survey was distributed to a convenience sample of 64 individuals, primarily consisting of MSBA students and individuals in the immediate vicinity of the researchers. Respondents were asked to rate their likelihood of applying for a job based on various factors, including whether or not the job posting included salary information.Future studies with larger and more diverse sample populations are needed to further investigate the impact of salary disclosure on job search behaviors and preferences.

2. Respondents may have been influenced to choose the top answer option due to its placement

The presence of the treatment effect is a significant limitation of this study. Respondents may have been influenced to choose the top answer option due to its placement, rather than giving it proper consideration. This could lead to biased results and limit the validity of the study's findings. Future studies could consider using a randomized design, where the placement of the answer options is varied for each respondent, to minimize the impact of treatment effects on the study outcomes.

3. Randomization questions by software (*Qualtrics*)

The randomization of questions in a study is an important aspect of the research design as it ensures that participants are assigned to different groups in an unbiased manner, reducing the potential for confounding factors that may affect the study results. While the use of automated randomization software can be helpful in reducing researcher bias, it is possible that the software may not always result in perfect randomization, and the results may be influenced by factors that are not directly related to the study hypothesis. Future studies could consider that it is important for researchers to carefully choose the potential limitations and sources of bias in their study design and to take appropriate steps to minimize their impact.

4. Self-selection bias

Another potential source of bias in the study is self-selection bias. self-selection bias can also impact the internal validity of the study. Participants who choose to participate in the study may be different from those who do not participate in ways that could impact the study results. By carefully considering the study design and recruitment strategies, researchers can minimize the impact of self-selection bias and increase the validity and generalization of their findings.

Conclusion

Based on our analysis on the information effects of disclosing salary information on job postings, we observed an overall ATE of 0.1156 (11.56%), meaning that candidates are 11.56% more likely to apply to jobs with salary information on average. However, as mentioned above, we do encounter limitations in our research, especially in non-compliant individuals who didn't fully complete surveys and limited respondent pool. With more data, we hope our research outcomes can help companies make more informed decision when hiring talented individuals.

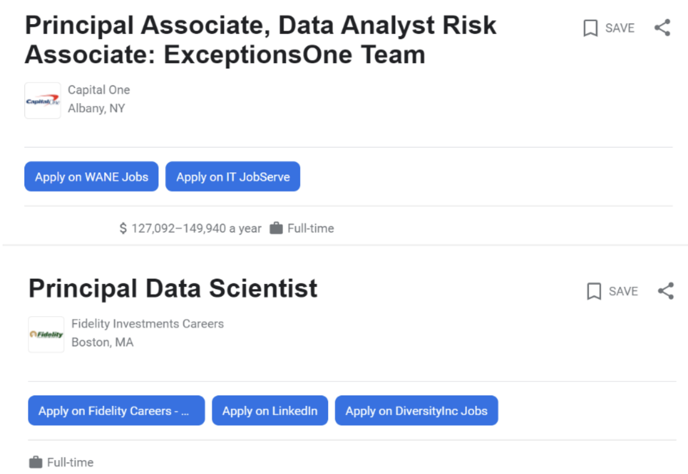
References

1. Impact of salary on job satisfaction. Eurofound. (n.d.). Retrieved March 15, 2023, from https://www.eurofound.europa.eu/publications/article/2013/impact-of-salary-on-job-satisfaction
2. Gee, Laura K. "The More You Know: Information Effects on Job Application Rates in a Large Field Experiment." Management Science 65, no. 5 (2018): 1949–2443. https://gap.hks.harvard.edu/more-you-know-information-effects-job-application-rates-gender-large-field-experiment
3. Wheeler, Laurel, Robert Garlick, Eric Johnson, Patrick Shaw, and Marissa Gargano. 2022. "LinkedIn(to) Job Opportunities: Experimental Evidence from Job Readiness Training." American Economic Journal: Applied Economics, 14 (2): 101-25. https://www.aeaweb.org/articles?id=10.1257/app.20200025
4. Manudeep Bhuller, Domenico Ferraro, Andreas R. Kostøl & Trond C. Vigtel, 2023. "The Internet, Search Frictions and Aggregate Unemployment." National Bureau of Economic Research. https://www.nber.org/papers/w30911

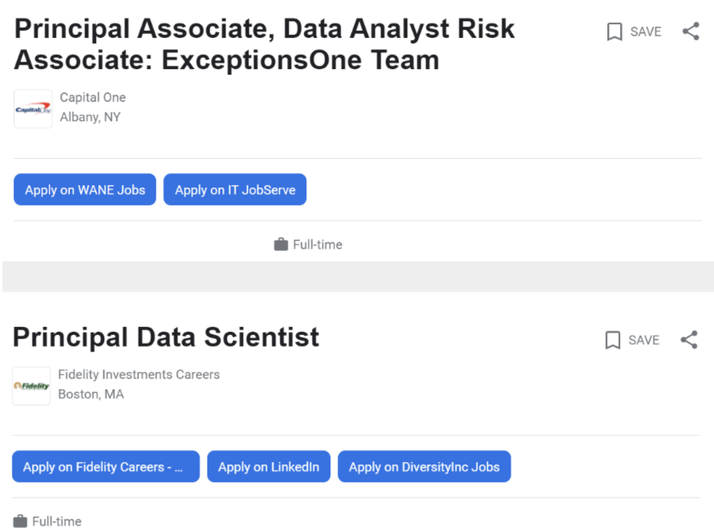
Appendix

Exhibit 1: Survey Design

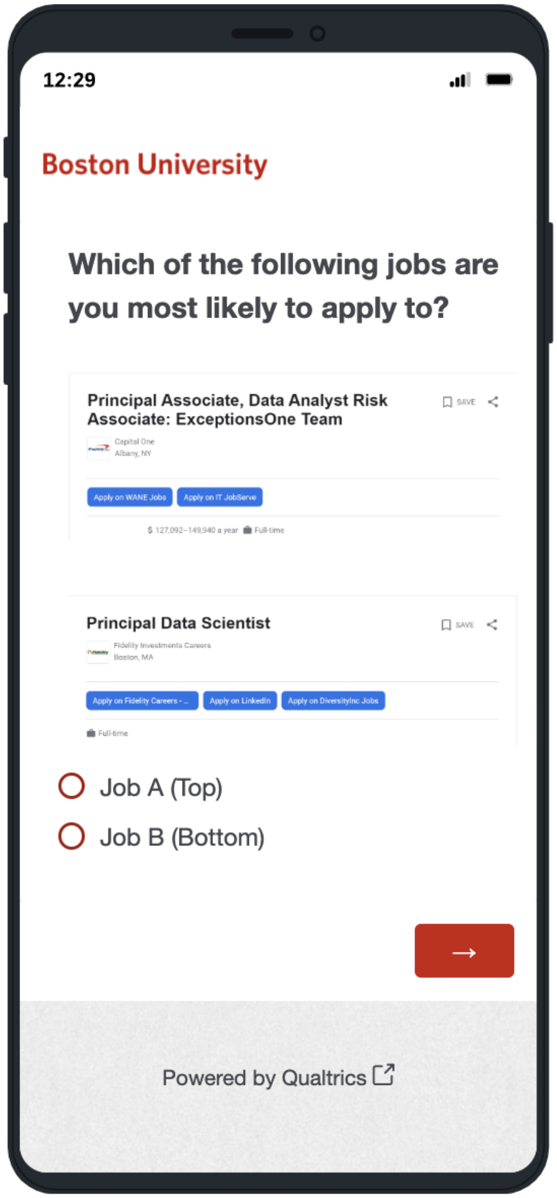
Treatment Posting (no salary information)



Control Posting (no salary information)



Mobile Survey Interface (Qualtrics)


12:29
Boston University
Which of the following jobs are you most likely to apply to?
Principal Associate, Data Analyst Risk Associate: ExceptionsOne Team
SAVE
Capital One
Albany, NY
Apply on WANE Jobs
Apply on IT JobServe
$ 127,092–149,940 a year
Full-time
Principal Data Scientist
SAVE
Fidelity Investments Careers
Boston, MA
Apply on Fidelity Careers - ...
Apply on LinkedIn
Apply on DiversityInc Jobs
Full-time
Job A (Top)
Job B (Bottom)
→
Powered by Qualtrics