

# Randomized Experiments

Campbell Miller

Due by Noon on Friday, November 12th

**Directions** Write your solutions in Rmarkdown and submit your solutions as a rendered HTML or PDF file. Only (sufficiently well done) HTML or PDF submissions will be counted as complete. Install packages in your R console and only load them in your markdown file. Installing packages in your markdown file will create an error.

## 1. Bias in Resume Screening

In class, we briefly discussed the challenge of using experiments to ask important questions where the “treatment” of interest is difficult to randomize. For example, is there gender or racial discrimination in hiring? How could we run an experiment to test this given that we cannot randomize an individuals gender or race?

In “Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination” Bertrand and Mullainathan (2004) solve this challenge by randomizing resume characteristics rather than individuals. Specifically, they generated fake resumes and randomly varied whether the name was identifiable as belonging to a man or a woman or a black or white person. They measure gender and racial discrimination by looking at the treatment effect of gender and race on call back rates.

The data is in the file **BertrandMullainathan2004.csv**. The treatment variables are defined by the sex and race variables. The main outcome of interest is the variable call which equals 1 if someone received a call back.

- a. Is there evidence of racial discrimination in resume screening in the data? (This answer should be based on an estimated causal effect.)

```
setwd("C:/Users/campb/OneDrive/econometrics/hw6")
data <- read.csv("C:/Users/campb/OneDrive/econometrics/hw6/BertrandMullainathan2004.csv")
library(broom)
library(stargazer)
```

```
##
```

```
## Please cite as:
```

```
## Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
## R package version 5.2.3. https://CRAN.R-project.org/package=stargazer
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5    v purrr  0.3.4
## v tibble  3.1.6    v dplyr  1.0.8
## v tidyr   1.2.0    v stringr 1.4.0
## v readr   2.1.2    v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
#data <- data %>%
# mutate(race = factor(race))

reg1 <- lm(call ~ race - 1, data = data)

reg1
```

```
##
## Call:
## lm(formula = call ~ race - 1, data = data)
##
## Coefficients:
##   raceb   racew
## 0.06448  0.09651
```

```
#Yes, white people received 3.2% more callbacks than black people
```

b. Is the amount of racial bias the same for men and women?

```
data <- data %>%
  mutate(black = race == "b")
data <- data %>%
  mutate(women = sex == "f")
data <- data %>%
  mutate(white = race == "wh")
data <- data %>%
  mutate(men = sex == "m")

data <- data %>%
  mutate(bwomen = black * women)
#data <- data %>%
# mutate(whwomen = white * women)
data <- data %>%
  mutate(bmen = black * men)
#data <- data %>%
# mutate(whmen = white * men)

reg3 <- lm(call ~ bwomen + bmen, data = data)

reg3
```

```
##
## Call:
## lm(formula = call ~ bwomen + bmen, data = data)
##
## Coefficients:
## (Intercept)      bwomen      bmen
##    0.09651    -0.03023    -0.03822
```

```
# Black women got just under a 1% higher callback rate
```

- c. The data includes a number of other resume characteristics. These are the baseline covariates in the data. These include the variables: education, ofjobs, yearsexp, honors, volunteer, military, empholes, workschool, email, computerskills, and specificskills. Pick 5 of these characteristics and discuss whether they are balanced across race categories.

```
reg4 <- lm(call ~ race + yearsexp - 1, data = data)
reg4
```

```
##
## Call:
## lm(formula = call ~ race + yearsexp - 1, data = data)
##
## Coefficients:
##      raceb      racew  yearsexp
## 0.038598  0.070543  0.003305
```

```
#Years of experience has a 3% higher callback rate for white,
#close to the uncontrolled difference in black vs white we saw.
```

```
reg5 <- lm(call ~ race + military - 1, data = data)
reg5
```

```
##
## Call:
## lm(formula = call ~ race + military - 1, data = data)
##
## Coefficients:
##      raceb      racew  military
## 0.06631    0.09818   -0.01805
```

```
#Military has a 3.2 % increase in callback rate for white #applicants.
```

```
reg6 <- lm(call ~ race + computerskills -1, data = data)
reg6
```

```
##
## Call:
## lm(formula = call ~ race + computerskills - 1, data = data)
##
## Coefficients:
##      raceb      racew  computerskills
##    0.08042    0.11199    -0.01915
```

*#computer skills has around 3% higher call back rate for white*

```
reg7 <- lm(call ~ race + specialskills - 1, data = data)
reg7
```

```
##
## Call:
## lm(formula = call ~ race + specialskills - 1, data = data)
##
## Coefficients:
##          raceb          racew  specialskills
##          0.04346          0.07530          0.06422
```

*#There was a 3.2% higher call back rate for white applicants when  
# controlled for specialskills*

```
reg8 <- lm(call ~ race + workinschool - 1, data = data)
reg8
```

```
##
## Call:
## lm(formula = call ~ race + workinschool - 1, data = data)
##
## Coefficients:
##          raceb          racew  workinschool
##          0.07300          0.10499         -0.01519
```

*#There was around 2.5 % higher call back rate for white applicants  
#controlled for work in school*

*#Overall controlling for almost every skill still gets around 3%  
#higher call back white for white applicants. This shows that  
#regardless of skills or experience, equally skilled white  
#applicants were still called back 3% more often.*

- d. How do the estimates from part a change if you control for these baseline characteristics when estimating the impact of race and gender on call back rates?

```
reg9 <- lm(call ~ race + education + ofjobs + yearsexp + honors + volunteer+ military + empholes + workinschool - 1, data = data)
reg9
```

```
##
## Call:
## lm(formula = call ~ race + education + ofjobs + yearsexp + honors +
##   volunteer + military + empholes + workinschool + email +
##   computerskills + specialskills - 1, data = data)
##
## Coefficients:
##          raceb          racew      education          ofjobs          yearsexp
##       -0.0034057       0.0276033       0.0050856      -0.0015447       0.0031482
##          honors          volunteer          military          empholes          workinschool
```

```
##      0.0531384      0.0005037      0.0016530      0.0293771      0.0140156
##      email  computerskills  specialskills
##      0.0123552      -0.0229834      0.0599767
```

*#The raceb coefficient is -.0034 and the racew is .027 so the  
# difference between black and white is still about .03 or 3% which #is what we got in part A.  
#After controlling it is clear to see being black is negative  
#and being white is a positive effect of .027 towards getting  
#called back.*

```
reg10 <- lm(call ~ sex + education + ofjobs + yearsexp + honors + volunteer+ military + empholes + work  
tidy(reg10)
```

```
## # A tibble: 13 x 5
##   term          estimate std.error statistic  p.value
##   <chr>          <dbl>     <dbl>     <dbl>    <dbl>
## 1 sexf          0.0118     0.0250     0.473  6.36e- 1
## 2 sexm          0.0140     0.0255     0.550  5.83e- 1
## 3 education     0.00508    0.00571     0.889  3.74e- 1
## 4 ofjobs       -0.00156    0.00359    -0.433  6.65e- 1
## 5 yearsexp      0.00315    0.000860     3.66  2.55e- 4
## 6 honors        0.0534     0.0183     2.92   3.53e- 3
## 7 volunteer     0.0000292   0.0110     0.00265 9.98e- 1
## 8 military      0.000449    0.0146     0.0307  9.75e- 1
## 9 empholes      0.0298     0.00936     3.19   1.45e- 3
## 10 workinschool  0.0145     0.00984     1.47   1.42e- 1
## 11 email        0.0130     0.0112     1.16   2.46e- 1
## 12 computerskills -0.0238    0.0111    -2.16   3.12e- 2
## 13 specialskills  0.0601     0.00862     6.98   3.40e-12
```

*#It is a bit less clear the effects on gender with scientific  
#notation but it appears that being male is a more positive  
#effect than being female*

- e. Look at the names used in the data for the resumes in each of the 4 gender/race categories. Could potential employers be inferring anything besides race and gender from the names? If so, how does this change your interpretation of the results?

```
table(data$firstname)
```

```
##
##   Aisha Allison Anne Brad Brendan Brett Carrie Darnell
##   180      232    242    63      65      59    168      42
##   Ebony Emily Geoffrey Greg Hakim Jamal Jay Jermaine
##   208      227     59    51      55      61     67     52
##   Jill Kareem Keisha Kenya Kristen Lakisha Latonya Latoya
##   203      64    183    196      213     200     230     226
##   Laurie Leroy Matthew Meredith Neil Rasheed Sarah Tamika
##   195      64     67    187      76      67     193     256
##   Tanisha Todd Tremayne Tyrone
##   207      68     69     75
```

*# I don't think much more than race and gender can be inferred from  
# these names. Maybe religion or something of the like could be  
# if the employer has preconceived notions about what race is what  
# religion and if this could affect the hiring process if the  
# employer was say Christian and wanted to hire more Christians  
# and thought one type of person was more likely to fit that group.  
# But this case has more to do with racial or gender discrimination  
# and how those impact the ideas of who is what religion.*

## 2. Impact of Dogs on Health

On several occasions, we have discussed the correlation between dog ownership and health. Propose a randomized experiment to test this. How might individuals in your experiment feel about your randomization? Could this create problems?

With no constraints on length or money there are many ways this experiment could be done. A basic experiment would be to randomize a large group of people and then randomly tell half of the sample to get a dog and half to not get a dog. You could measure health traits such as weight, ability to complete exercises, blood work, and mental health from the start of the experiment to a few years later. If you wait long enough it could potentially include the time after the dog's death to see if the health measures you are taking change after losing the dog. Then you could compare these statistics for both groups and see if one had a statistically significant increase or decrease in health measures through the experiment. The individuals in this experiment could feel like you telling them to get a dog or not does not matter and not listen to you. This would cause problems because the individuals would likely just continue with their lives and get a dog or not based on what they want to do and not the experiment and this could change the results. Or if people got a dog after you told them to but they hate dogs or are allergic to them, this could negatively impact health and show an incorrect conclusion for people who like dogs or are not allergic.