In-Class Exercise 7

For this in-class exercise, work with your group of 2-3 people, to answer the following questions. These questions are not necessarily easy and sometimes they will not have a clear "correct" answer. The goal is to get you thinking about the material we've learned. Some of these questions may require you to discuss and debate with your group members to come up with an answer or can cover topics that we have not yet covered in class.

Be prepared to share your answers with the class and add to the discussion.

After class submit your a do-file with your answers in comments to Moodle for grading. You will be graded as a group on your submission. Only one group member needs to submit the assignment, but make sure add all group member names.

For this, you will use data from the paper:

"Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination" by Marianne Bertrand and Sendhil Mullainathan

Download the data here: ResumeNames

In this paper, researchers sent fictional resumes to help-wanted ads in Boston and Chicago. They randomly assigned names to the resumes: some were given "White-sounding" names (e.g., Emily, Greg) and others "African-American-sounding" names (e.g., Lakisha, Jamal).

Because the names were randomized, any difference in callback rates can be causally attributed to the name itself (and the race it signals).

The variables you'll be using are:

1. `call` : Binary. 1 if the applicant got a callback, 0 otherwise.
2. `ethnicity` : "cauc" (Caucasian) or "afam" (African American).
3. `quality` : "low" or "high" (quality of the resume credentials).
4. `experience`: Years of experience listed.

Question 1

Read through the brief description of the experiment here.

Think about the design of this experiment. Why is randomization important here? What potential confounding variables does randomization help to eliminate? Can you think of potential issues with this design? What assumptions must hold for us to interpret the results causally?

> ! Solution

Importance of Randomization: Randomization ensures that the only systematic difference between the resumes is the name (and thus the perceived race). This allows us to isolate the effect of racial discrimination on callback rates.

Confounding Variables Eliminated: - Applicant qualifications (education, experience) - Resume quality - Job type and industry (since resumes were sent to a variety of ads) - Geographic location (Boston vs. Chicago) **Potential Issues:** - Employers might infer other characteristics (e.g., socioeconomic status) from names. - The sample may not be representative of the entire job market. - The experiment only captures initial callbacks, not actual hiring decisions.
Assumptions for Causal Interpretation: - Randomization was properly implemented. - No spillover effects (employers do not share information about applicants). - The resumes are perceived similarly except for the name.

Question 2

Create a binary variable `is_black` where 1 = "afam" and 0 = "cauc".

Run the regression:

$$call_i = \beta_0 + \beta_1 is\_black_i + u_i$$

Interpret the coefficient on `is_black`. Is it statistically significant?

> ! Solution

```
import delimited "ResumeNames.dta", clear
gen is_black = (ethnicity == "afam")
reg call is_black
```

Results: - **Coefficient on `is_black`:** ≈ -0.032 (or −3.2 percentage points) - **Use Stata to find t-stat or p-value:** The t-statistic is large (~ -4.1) and the p-value is effectively 0.000.

Interpretation: Resumes with "African-American-sounding" names are **3.2 percentage points less likely** to receive a callback compared to "White-sounding" names (baseline: ~9.7% vs. ~6.5%). This result is **statistically significant** at the 1% level.

Question 3

One major concern in regressions is that it is correlated with the error term (e.g., lower wages might be due to lower experience, not discrimination).

Run a regression of `experience` on `is_black`.

If the experiment was randomized correctly, what should the coefficient on `is_black` be? Explain.

! Solution

```
reg experience is_black
```

Results: - Coefficient on `is_black`: ≈ -0.027 - P-value: ≈ 0.854 (Not significant)

Explanation: If randomization is working correctly, the assignment of names ("treatment") should be uncorrelated with other observed characteristics like experience. The coefficient should be **statistically indistinguishable from zero**. Since the p-value is huge (0.854), we fail to reject the null hypothesis that the difference is zero. This confirms **successful randomization**—black and white applicants have practically identical average experience levels in this sample.

Question 4

There is a theory that "higher quality" resumes overcome discrimination. Subset the data to only white and black ethnicity and look at the effect of callbacks for high-quality and low quality resumes. What do you find? What does this suggest about how high quality resumes affect black and white applicants differently?

! Solution

```
* For White Applicants
reg call quality if is_black == 0
* For Black Applicants
reg call quality if is_black == 1
```

Results: - **Whites:** High quality resumes increase callbacks by ≈ 2.3 percentage points ($p < 0.05$). (Base ~8.5% → ~10.8%) - **Blacks:** High quality resumes increase callbacks by only ≈ 0.5 percentage points ($p > 0.05$, not significant). (Base ~6.2% → ~6.7%)

Interpretation: This contradicts the theory that "credentials overcome discrimination." African American applicants see almost **no return** to having a higher quality resume, whereas White applicants see a significant boost. The gap *widens* for high-quality candidates.

Question 5

Read this passage from the original paper (p. 1010):

> *What do these results imply for existing models of discrimination? Economic theories of discrimination can be classified into two main categories: taste-based and statistical discrimination models. Both sets of models can obviously "explain" our average racial gap in callbacks. But can these models explain our other findings? More specifically, we discuss the relevance of these*

*models with a focus on two of the facts that have been uncovered in this paper: (i) the lower returns to credentials for African-Americans; (ii) the relative uniformity of the race gap across **occupations**, job requirements and, to a lesser extent, employer characteristics and **industries.***

Run the regressions you ran from Question 2 separately for each occupation (`wanted`). What do you find? Can you relate your findings back to the passage above?

Do the same for different industries (`industry`). What do you find?

> ! Hint
>
> The point here is not to make you run lots of very similar regressions. Instead, think about how you can condense all this information so I can read it and understand it.

> ! Solution
>
> **Occupation Analysis:** Running the regression by `wanted` category shows a consistent negative coefficient for `is_black` across almost all job types.
>
> - **Significant Discrimination:** Secretary (~−3.8 pp).
> - **Consistent Negatives:** Supervisor (-4.3 pp), Office Support (-3.1 pp), Retail (-2.7 pp), Manager (-2.1 pp).
> - *Note: Smaller sample sizes in sub-groups reduce statistical significance (higher p-values), but the direction is almost uniformly negative.*
>
> **Industry Analysis:** The pattern holds across industries as well.
>
> - **Significant Discrimination:** Finance/Insurance/Real Estate (-5.8 pp!), Business/Services (-4.1 pp), Trade (-3.5 pp).
> - **Consistent Negatives:** Manufacturing, Health/Education.
> - **Exception:** Transport/Communication shows a (rare) positive coefficient (+2.7 pp), but it is not statistically significant ($p \approx 0.63$).
>
> **Relation to Passage:** The results confirm the authors' statement about the **"relative uniformity of the race gap"**. We don't see one specific industry or job type driving the result; rather, the penalty for a "Black-sounding" name appears to be a systemic constant across the labor market, regardless of the specific occupation or sector.

Question 6

This paper was written in 2004 before the widespread use of algorithms in hiring processes and AI. Do you think the findings of this paper would be different if the same experiment were conducted today? Why or why not?

! Solution

Answers will vary. Some points to consider:

- **Algorithmic Bias:** Modern hiring algorithms may inadvertently perpetuate or even amplify existing biases if they are trained on historical data that reflects human prejudices.
- **Anonymized Applications:** Some companies now use blind recruitment processes (removing names) to reduce bias, which could mitigate the effects observed in the original study.
- **Increased Awareness:** Greater societal awareness of racial discrimination might lead to more cautious behavior by employers, potentially reducing bias.
- **New Forms of Bias:** However, new forms of bias could emerge, such as biases based on other demographic factors or even algorithmic decision-making processes that are not transparent.