# STOR 455 Homework 6

20 points - Due Friday 3/25 5:00pm

**Are Emily and Greg More Employable Than Lakisha and Jamal?**

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review, 94*(4), pp. 991-1013.

*Abstract*

We perform a field experiment to measure racial discrimination in the labor market. We respond with fictitious resumes to help-wanted ads in Boston and Chicago newspapers. To manipulate perception of race, each resume is randomly assigned either a very African American sounding name or a very White sounding name. The results show significant discrimination against African-American names: White names receive 50 percent more callbacks for interviews. We also find that race affects the benefits of a better resume. For White names, a higher quality resume elicits 30 percent more callbacks whereas for African Americans, it elicits a far smaller increase. Applicants living in better neighborhoods receive more callbacks but, interestingly, this effect does not differ by race. The amount of discrimination is uniform across occupations and industries. Federal contractors and employers who list "Equal Opportunity Employer" in their ad discriminate as much as other employers. We find little evidence that our results are driven by employers inferring something other than race, such as social class, from the names. These results suggest that racial discrimination is still a prominent feature of the labor market.

| Variables | Descriptions |
|---|---|
| *call* | Was the applicant called back? (1 = yes; 0 = no) |
| *ethnicity* | indicating ethnicity (i.e., "Caucasian-sounding" vs. "African-American sounding" first name) |
| *sex* | indicating sex |
| *quality* | Indicating quality of resume. |
| *experience* | Number of years of work experience on the resume |
| *equal* | Is the employer EOE (equal opportunity employment)? |

Use the *ResumeNames455* found at the address below:

https://raw.githubusercontent.com/JA-McLean/STOR455/master/data/ResumeNames455.csv

1) Construct a logistic model to predict if the job applicant was called back using *experience* as the predictor variable.

2) Plot the raw data and the logistic curve on the same axes.

3) For an applicant with 6 years of experience, what does your model predict is the probability of this applicant getting called back?

4) Construct an empirical logit plot and comment on the linearity of the data.

5) Use the model from question #1 to perform a hypothesis test to determine if there is significant evidence of a relationship between *call* and *experience*. Cite your hypotheses, p-value, and conclusion in context.

6) Construct a confidence interval for the odds ratio for your model and include a sentence interpreting the interval in the context.

7) For each 2-year increase in *experience*, how does your model predict the odds will change for the applicant getting called back?

8) Construct subsets of the data for each category of *ethnicity* and construct logistic models to predict if the job applicant was called back using *experience* as the predictor variable for each of these subsets. Then plot the raw data and the logistic curves on the same axes. Comment on differences between the curves and what this means in the context of the data.

9) Construct subsets of the data for each category of *sex* and construct logistic models to predict if the job applicant was called back using *experience* as the predictor variable for each of these subsets. Then plot the raw data and the logistic curves on the same axes. Comment on differences between the curves and what this means in the context of the data.

In homework #7 we will continue with this data to investigate how the other variables impact an applicant's chances of being called back using multiple logistic regression models.