KING'S COLLEGE LONDON
(University of London)

King's Business School

6SSMN961: APPLIED ECONOMETRICS

2019-20

Problem Set 1

1. Consider the following three causal questions:

   Q1: Does going to a high-rank university increase earnings?

   Q2: Does immigration increase house prices?

   Q3: Do children in smaller families have better opportunities?

For each of these questions, answer the following:

   a. What is the outcome variable and what is the treatment?
   b. Define the potential outcomes Y_{0i} and Y_{1i}.
   c. What plausible causal channel(s) runs directly from the treatment to the outcome?
   d. What are possible sources of selection bias in the raw comparison of outcomes by treatment status? Which way would you expect the bias to go and why?

2. For this question we will use a dataset from a randomised experiment conducted by Marianne Bertrand and Sendhil Mullainathan, who sent 4,870 fictitious resumes out to employers in response to job adverts in Boston and Chicago in 2001. The resumes differ in various attributes including the names of the applicants, and different resumes were randomly allocated to job openings. Some of the names are distinctly white sounding and some distinctly black sounding. The researchers collecting these data were interested to learn whether black sounding names obtain fewer callbacks for interviews than white names.

   Download the data set **bm.dta** from course page.

   a. The data set contains two dummy variables (0-1 variables) for female (**female**) and whether the applicant has computer skills (**computerskills**). Tabulate these variables by **black**. Using the command:

      ```
      tab female black, col
      ```

will give you cross-tabulation of female and race, and display the percentages of males and females in each race group.

Do gender and computer skills look balanced across race groups?

b. Do a similar tabulation for **education** and the number of jobs previously held (**ofjobs**). Does education and the number of previous jobs look balanced across race groups?

c. Use the `summarize` command to look at the mean and standard deviation for the variable years of experience (**yearsexp**) separately for black and whites (using the `if` modifier). Does this variable look similar by race?

d. What do you make of the overall results on resume characteristics? Why do we care about whether these variables look similar across the race groups?

e. The variable of interest on the data set is the variable **call**, which indicates a call back for an interview. Calculate the t-statistic for the difference in average call back rates for whites and blacks. Do you find differences in call back rates by race?

f. Run a regression of **call** on **black**. Interpret the results and compare with what you obtained in part (e).

g. Run a regression of **call** on **black**, but now include also as controls **female**, c**omputerskills**, **education**, **ofjobs** and **yearexp.** Interpret the results and compare with what you obtained in part (f).

h. What do you conclude from the results of the Bertrand and Mullainathan experiment?

References:

Bertrand and Mullainathan (2004), "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination", *American Economic Review*, vol. 94, no.4, pages 991-1013