Economics 241B Exercise 3
OLS Algebra and Projections Lectures

1. (2017 Prelim) You are asked to determine how the conditional mean of (log) wage, denoted y, depends on the sex of the individual. Let m be an indicator for male (that is, m takes the value 1 if the individual is male) and let f be the indicator for female (that is, f takes the value 1 if the individual is female). For a sample of n individuals, indexed by i, with n_1 males, consider OLS regression of

$$y_i = \beta_1 m_i + \beta_2 f_i + u_i. \tag{1}$$

a) Is $\beta_1 m_i + \beta_2 f_i$ an exact, or approximate, expression for the mean of the male and female wage distributions?

b) Show that $\widehat{\beta}_1$ is the sample mean of male wages.

c) Consider the continuous covariates X (an $n \times k$ matrix) along with y (an $n \times 1$ vector). Describe, in words, the transformations

$$\begin{aligned} y^* &= y - \widehat{\beta}_1 m - \widehat{\beta}_2 f \\ X^* &= X - m\overline{x}_1^{\mathrm{T}} - f\overline{x}_2^{\mathrm{T}}, \end{aligned}$$

where $\overline{x}_1$ and $\overline{x}_2$ are the $k \times 1$ means of the covariates for men and women, respectively.

d) Compare $\widetilde{\alpha}$ from the OLS regression

$$y^* = X^*\widetilde{\alpha} + \widetilde{u}$$

with $\widehat{\alpha}$ from the OLS regression

$$y = \widehat{\beta}_1 m + \widehat{\beta}_2 f + X\widehat{\alpha} + \widehat{u}.$$

2. Computational Exercise

Read through the paper by Charness and Kuhn listed on the syllabus. Write programs in both Matlab and Stata (the results from each program should match) that estimate the model in columns (1), (2) and (3) of Table 3 of Charness and Kuhn. Calculate classic standard errors (the authors report cluster-robust standard errors, so your estimated standard errors will not match those in the table).

Finally, for the models of columns (2) and (3) test the hypothesis that the coefficient on relative wage equals zero and provide the p-value for the estimated test statistic.

The key variables:

e1 - effort of type 1 workers

w1 - wage of type 1 workers

w2 - wage of type 2 workers

For help with Stata programming, consult the reference by Cameron and Trivedi.

For Matlab programming: Matlab files are denoted .m and there are certain naming conventions to follow. To determine these, use Matlab help to search "filenames" and then follow the link to naming m files.

Notation to begin:

% denotes a comment to follow

; denotes the output of the line should not be printed

Commands to study

clc

clear

xlsread

plot

Sample Program

```
%     Program Title
%     Date created
clc;
clear all;                % to ensure a clear workspace before you begin
data=xlsread('dataset');      % to read in data
q=data(1:92,1);           % define quantity
plot(q, 'd');             % plot for outliers
```