Lab Session 1

Classroom Environment and Problem Sets

1. If you have questions that are relevant to the current topic or if you are lost, please feel free to ask!
2. For the future session, 15-20 min review PS in previous week, 25-30 min discuss sample codes related to new PS, 20-25 min try new PS by yourself, Q&A (personalized questions)[1].
3. Collaboration and discussion with your neighbors are encouraged when you try PS by yourself.
4. Remember to save your do-file on u-disk or online drive or email to yourself.
5. You must submit your problem sets typed (a do-file with code and a pdf-file with answers to each question), electronic version on Canvas and printed version in Friday's lab session, late problem sets will not be accepted.
6. Grading criteria: full mark is 10, each small question is assigned an equal mark. For instance, if there are 5 questions in the PS, then 2 points for each question. A wrong answer or a right but incomplete answer would end up with no point or 50% of the full points.

Introduction to Stata

1. What is Stata? advantages and disadvantages of Stata

   • Stata is a statistical software for data science.
   • advantages: user-friendly, built-in economic analysis packages, great guides available on web
   • disadvantages: it doesn't have these advantages in other software, R (free, open-source software), Matlab (matrix analysis, multidimensional datasets), Python (free, open-source software, create datasets)

2. Familiarize yourself with the Stata interface

   • command window, results window, review window, variables window
   • do-file editor, Stata help

3. Resources in addition to Google: Stata official website[2], CISER[3]

[1] the time might vary according to the difficulty level of problem sets
[2] https://www.stata.com
[3] https://ciser.cornell.edu

Lab Session 1

PS 1: Capital Asset Pricing Model

1. Question Overview

$$r_j - r_f = \alpha_j + \beta_j(r_m - r_f) + e_j$$

where r_j and r_f are the returns to security j and the risk-free rate, respectively, r_m is the return on the market portfolio, and β_j is the jth security's "beta" value.

2. Dataset Import

```
** ================================================================================
** Filename = MengweiLin_AEM4110_Session1.do
** Author = Mengwei Lin
** Date Modified = 2019/09/06
** Purpose = AEM4110 Lab Session 1: Intro to Stata & Problem Set 1
** ================================================================================

clear //clear memory
set more off //tells Stata not to pause or display the -more- message
cd "/Users/mengweilin/Desktop/AEM4110" //change directory
capture log close //to close any open logs
                  //capture tells Stata to run the command and ignore any errors
log using PS1, text replace //to keep a permanent record of your results
import delimited capm5.csv //import dataset
```

3. Variables Generation and Regression

```
gen x = mkt - riskfree //generate new variable x
gen y_xom = xom - riskfree //generate new variable y_xom
reg y_xom x //regress y_xom on x

mat b_xom = e(b) //save the coefficients in the matrix b_xom
mat list b_xom //display matrix
```

4. Hypothesis Testing and Statistics Compute (Example of Question b. Exxon-Mobil)
   - Null Hypothesis H_0: $\beta_{xom} = 1$
   - Alternative Hypothesis H_1: $\beta_{xom} \neq 1$
   - Test Statistic: $t = \frac{b_{xom}-1}{se(b_{xom})} \sim t_{(N-2)}$
   - P-value (given H_1: $\beta_{xom} \neq 1$): p = Pr(the right of $|t|$) + Pr(the left of $-|t|$)

```
scalar t_b_xom = (_b[x]-1)/_se[x] //compute test statistic
display "the test statistic for beta_xom is:" t_b_xom //show the result
scalar p_b_xom = 2*ttail(178,abs(t_b_xom)) //compute the p-value
display "the p-value for beta_xom is:" p_b_xom //show the result
```

5. Loop

```
foreach a of varlist ford ge xom{
reg y_`a' x
}
//loops allow you to run the same command for several variables at one time
//without having to write separate code for each variable.
```