Rauli Susmel

Financial Econometrics FINA 4397

Homework 2 (Due September 24)

Instructions: Send your solved homework, along with the code, to my TA, Yousaf, Hammad. His email address is: hyousaf@CougarNet.UH.EDU.

1.1 (Tests of Hypothesis). Download the Shiller dataset (Shiller data.csv) from my homepage. Or just cut-and-paste the following line:

Sh da <- read.csv("http://www.bauer.uh.edu/rsusmel/4397/Shiller data.csv", head=TRUE,sep=",")

You have stock prices (P), dividends (D), earning (E), consumer prices (CPI) and long interest rates (Long i). Regress log stock returns, r_i , against log earning changes, *earn*_i, inflation rate (in log changes), *Inf*_i, and interest rates, *int*_i (need to subtract one observation): 1

$$r_i = \beta_0 + \beta_1 \ earn_i + \beta_2 \ Inf_i + \beta_3 \ int_i + \varepsilon_i$$

- a. Report the regression
- b. Interpret the \mathbb{R}^2 .
- c. Interpret the estimated coefficient β_1 .
- d. Test with a goodness of fit test H₀: $\beta_1 = \beta_2 = \beta_3 = 0$.
- e. Test with an F-test H₀: $\beta_1 = \beta_3 = 0$.
- f. Test with a Wald test H₀: $\beta_2 = 0.5$ and $\beta_3 = -0.1$

1.2 (Bootstrapping). Bootstrap the t-statistics in the above regression, with B = 1,000. a. Report the mean and the bias in your estimation for each parameter.

b. Build a 95% C.I. for β_2 .

1.3 (Data Problems) Check for outliers and multicollinearity in the above regression

a. Plot standardized residuals and Cook's D statistic

b. Does the regression suffer from outliers? Report the proportion of leverage observations and standardized residuals that are significant according to the "Rules of Thumb" presented in class c. Does the regression suffer from multicollinearity problems? Report VIF and Condition Index numbers.

1.4 (Theory Review)

a. What does it mean that an estimator is unbiased? Consistent? Would you ever consider an inconsistent estimator?

b. Suppose you suspect the unobservable error terms (ε) in a regression does not follow a Normal distribution. Describe how would you test that ε is not normally distributed (state the Null Hypothesis and the test used).

c. Under what circumstances you would use a bootstrap to compute SE for a regression?

d. What are the consequences for the CLM that the errors are not normally distributed?

e. What does it mean that a regression suffer from multicollinearity? What is the possible effect of multicollinearity on a regression? Can you fix it?