Empirical Corporate Finance Course
Winter 2014
Paola Sapienza

Problem Set 1-2

Please note: it is a violation of the rule of the class and will be punished accordingly to copy full or part of this exercise, or receive the code from somebody else.

Problem sets 1 and 2 consists in putting your empirical skills and knowledge into practice. After you have completed both you will have replicated some of the basic regressions of a recent empirical paper:

Eberly, Rebelo and Vincent (2009). "Investment and Value: A Neoclassical Benchmark". Working Paper.
You can find the paper here:

http://www.kellogg.northwestern.edu/faculty/sapienza/ftp/Ph.D.Teaching/ECF/Eberlyforproblemset1.pdf

(if you cannot access it, please email In Selwyn).

You need to obtain the data from the same source as the paper, the Compustat Database. For that, you need to create an account and obtain access to the database.

Refer to: www.kellogg.northwestern.edu/rc/data.htm for creating your account. You will have to do the regression analysis and the exercises in STATA.

Problem set 1 (due January 30th)

1) Extracting the data. You should extract the data by using the SAS programming that you have learned in class. You should attach the SAS program used to extract the data at the back of the exercise. PLEASE DO **NOT** USE THE WEB INTERFACE TO EXTRACT THE DATA for the reasons described by Patricia in class. You should extract the data trying to replicate the procedure discussed in the original paper (as much as you can). To do that you should first extract

the data using the legacy data and then using the new Computstat data.

2) The data set:

a. Provide an introduction on how you extract the data (this session should mimic a good data section of a paper – [see for example, A Lobbying Approach to Evaluating the Sarbanes-Oxley Act of 2002, Yael Hochberg, Paola Sapienza, and Annette Vissing-Jørgensen, Journal of Accounting Research, 47(2), pp: 519-583, May 2009]. You should have an accurate description of the sample selection introduced by your method of extracting the data, include the final number of firms, number of firm-year observations, period…

b. Present a table with summary statistics on the variables of your data.

c. Part a and b needs to be done twice because you have two datasets. Comment briefly the differences between the two data, using appropriate statistics to compare the two.

Problem set 2 (due February 6th in class)

3) Estimate the following regressions:

a. The capital-investment ratio on Tobin's Q

b. The capital-investment ratio on Tobin's Q and cash-flow

c. Repeat the two regressions above with a semi-log specification.

d. Repeat the two regressions above with and without fixed effect

e. Present the results of the four regressions in a single table (similar to table 2 in Eberly et al. paper—see attached table here)

f. Briefly, comment on the results of the regressions

4) Since you are using a panel data set, it is likely that the residuals are correlated across firms or across time. This would lead to biased standard errors in your previous OLS estimations.

a. Read the paper:
Mitchell A. Petersen (2009): "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches" (forthcoming in the Review of Financial Studies)

b. Apply the various techniques presented in the paper in order to correct the standard errors in the regressions run in (2).

c. Present the results in a single table and comment briefly on the techniques you used and why you think you obtained different results with different corrections.

Some tips for both problem sets:

- Download only the variables that you will need for your study

- Note that, the description of the variables, with the corresponding numbers of their classification in Compustat, are specified at the appendix of the paper you are going to replicate.
- All the tables of results that you present must be self-contained, i.e. all the variables as well as symbols (like *, or numbers in brackets) should be explained in a text just above the table. The reader of your table should not go back to the text to understand what you have done in the tables. The notes of the table should be enough to understand your results. As an example of well done tables you can refer to:
A Lobbying Approach to Evaluating the Sarbanes-Oxley Act of 2002, Yael Hochberg, Paola Sapienza, and Annette Vissing-Jørgensen, Journal of Accounting Research, 47(2), pp: 519-583, May 2009.
http://www.kellogg.northwestern.edu/faculty/sapienza/htm/HSVJ.pdf