Assignment 2: Empirical research in accounting
Deadline: Sunday May 17, 2020, 23:59

Below are the question for Assignment 2 of the course *Research laboratory: Business Economics*. Please submit your group solutions in a PDF file. For questions 2 and 3, which requires analyses in Stata, also make sure to write down all of your Stata commands in a "do-file" and separately upload this do-file with the solutions document. Also make sure to include comments in the do-file in order for the reader to understand the steps taken (but do not include a log file of all Stata output). Note: all uploaded documents are automatically checked for plagiarism. Duplications of other groups' code and solutions will be identified (even when the wording has been changed) and all these cases will be forwarded to the examinations board.

You should be able to answer these question based on the material presented to you in the lectures and tutorials of block 5, as well as the Stata commands explained in Chapters 2–3 of the "*Introductory Guide to Using Stata in Empirical Financial Accounting Research*" documented that is posted on Canvas.

QUESTION 1: READING AND UNDERSTANDING AN EMPIRICAL PAPER

Carefully read the paper titled "The information content of annual earnings announcements and mandatory adoption of IFRS" by Landsman, Maydew, and Thornock (2012), which was published in the *Journal of Accounting and Economics*.[1] Answer the following questions:

a) What is the main research question?
b) Why is it important to have an answer to this question? (do not use more than 100 words)
c) What are the theoretical constructs examined in this study?
d) If the study is interested in examining a causal relation,
   - what is the X construct?
   - what is the Y construct?
e) How are the theoretical constructs operationalized (measured)?

In addition to answering these questions, provide "Libby boxes" for the paper. You may ignore the role of control variables in these Libby boxes.

QUESTION 2: PANEL DATA – EFFECTS OF CHANGES IN REGULATION

The Stata file "**shockdata.dta**", which is posted on Canvas, is a dataset that contains panel data for a large set of stock-listed companies from 15 Continental European countries over the years 2003–2017. The dataset was constructed from the Compustat Global Fundamentals Annual and Compustat Global Security Daily databases, which were obtained from WRDS. Next, the dataset was modified to reflect the effects of a *hypothetical* regulation change called "Regulation Random". This hypothetical regulation change is designed to have taken place in some, but not all, of the 15 countries at the start of the year 2010. Assume that the primary reason for this regulation change was to make companies operate more efficiently and become more profitable.[2] We can therefore test the following hypothesis:

HYPOTHESIS 1 (null): Regulation Random did not change the profitability of companies.

[1] Link: https://doi.org/10.1016/j.jacceco.2011.04.002.
[2] The regulation and its effects are *hypothetical* in the sense that we created the effects in the data using a simulation analysis. These data can therefore not be used to make actual inferences about the performance of companies in these countries.

The dataset consists of the following variables:

- *companyid*: unique identifier (anonymized) for each company;
- *countryid*: unique id for each country;
- *country*: name of the country;
- *year*: reporting year for which company variables are measured;
- *profit*: measure of the company's profitability for the year (randomly created);
- *companysizegroup*: a categorical variable capturing the relative size of the company (group 1 reflects the smallest companies, while group 5 reflects the largest companies);
- *regulation*: dummy variable set equal to 1 for companies in countries that were subject to Regulation Random, and 0 for all other companies.

Required:

a) Inspect the data to determine which countries were affected by Regulation Random.
b) Perform an OLS regression with *profit* as the dependent variable and *regulation* as the explanatory variable. Make sure to let Stata compute regression standard errors that are robust to heteroscedasticity and clustering at the company-level. Present the results from this regression and explain what we can conclude from the results. Does this analysis help you to test HYPOTHESIS 1? Why, or why not?
c) Create a graph with two lines that display the time patterns of the average (i.e., mean) profitability of companies for (a) the group of countries subject to the regulation change, and (b) the group of countries not subject to the regulation change.[3] Explain what we can infer from this graph with respect to HYPOTHESIS 1.
d) Based on the dataset, complete the following table by filling in the average of the profitability variable for each group, and by computing the difference-in-difference in average profitability. What does the result tell you about the effect of the regulation?

	Pre	Post	Difference
Regulation=1	…	…	…
Regulation=0	…	…	…
Difference	…	…	…

e) Perform an OLS regression with *profit* as the dependent variable and an interaction between the variables *regulation* and *post*, where the *post* variable is a dummy variable that is set equal to one for the years 2010–2017, and 0 for the years 2003–2009. Again let Stata calculate standard errors that are robust to heteroscedasticity and clustering at the company-level. Present the results from this regression and explain what we can conclude from the results. Does this analysis help you to test HYPOTHESIS 1? Why, or why not?
f) What is the name of the type of test you executed in questions d and e)?
g) Explain what the benefit is of using such a test, instead of simply looking at the changes in profitability for the companies that are subject to the regulation change.

QUESTION 3: EVENT STUDY – STOCK MARKET REACTIONS TO EARNINGS ANNOUNCEMENTS

Every quarter, listed companies announce the results of their performance over the most recently completed 3-month reporting period. This event is known as an "earnings announcement". Earnings announcements typically trigger much attention from the stock market, and stock prices often move substantially after a company has reported good or bad news at the earnings announcement. Accounting researchers have long been interested in understanding how the news in these earnings announcements

[3] You do not have to use Stata to create this graph. You may use Excel, or any other program you prefer, and paste the graph into your solutions document.

is incorporated into stock prices. These researchers have investigated questions such as: (1) is there a significant relation between the news in earnings announcements and stock prices? and (2) how quick and efficient is the stock market in pricing this information? These questions are typically studied using an "event study", in which the researcher examines the relative change in stock price (the "stock return") around the date of the event, and in which s/he assesses the statistical relation of these stock returns with other variables.

In this particular context, the researcher can look at results from a regression similar to the following:

$$RET_{iq}^{[0,1]} = \alpha + \beta SURPRISE_{iq} + \varepsilon_{iq} \quad (1)$$

where $RET_{iq}^{[0,1]}$ is a variable that captures the stock returns around the announcement, from the day of the announcement (day 0) to the day after the announcement (day 1). The subscript i refers to a company, while the subscript q refers to a quarter. $SURPRISE_{iq}$ is a variable that captures the news in the earnings announcement, which is typically defined as:

$$SURPRISE_{iq} = \frac{EPS_{iq} - FEPS_{iq}}{P_{iq}} \quad (2)$$

where EPS_{iq} is the measure of company performance over the period, the earnings per share (EPS), and $FEPS_{iq}$ is a measure of the expected EPS performance of the company. These expectations are often measured using the forecasts made by professional analysts in the period before the earnings announcement. If the company's EPS is higher than expected by analysts, this suggests good news. If the EPS is lower than expected, this suggests bad news. By examining whether stock returns have a significant positive relation with the news in $SURPRISE_{iq}$, we can formally test the following hypothesis:

HYPOTHESIS 2 (alternative): Accounting information contains information that is relevant for the stock market valuation of companies.

If, instead, the test suggests there is *no* significant positive relation between the news in $SURPRISE_{iq}$ and stock returns around the announcement, we cannot conclude that accounting information contains information that is relevant for the stock market valuation of companies. This part of the assignment relies on four datasets. The questions below can be answered by combining these datasets.

The first dataset "**announcements.dta**" is a panel dataset that consists of a large sample of quarterly observations of U.S. stock-listed companies over the period 2001–2018. For each company-quarter, the dataset provides information on the date of the earnings announcement event and the earnings per share (EPS) that was reported by the company for the quarter.

- *compustatcode*: unique company identifier variable;
- *crspcode*: unique company identifier variable;
- *edgarcode*: unique company identifier variable;
- *symbol*: unique company identifier variable;
- *periodenddate*: the final date of the reporting quarter;
- *cyear*: the calendar year in which the reporting quarter ends;
- *cquarter*: the calendar quarter in which the reporting quarter ends;;
- *prccq*: the stock price of the company at the end of the quarter in $;
- *rdq*: the date of the earnings announcement;
- *earnings*: the profitability of the company in the quarter, or earnings per share (EPS) in $.

The second dataset is "**analysts.dta**", which is a panel dataset with the average forecast of EPS made by analysts before each company makes its earnings announcement.

- *symbol*: unique company identifier variable;
- *datadate*: the final date of the reporting quarter;
- *forecast*: average forecast of analysts for the reporting quarter.

The third dataset is "**stockreturns.dta**", which is a panel dataset that contains stock return variables for each company. For each company and trading day, the dataset provides two stock return variables, which are measured over different time windows.

- *crspcode*: unique company identifier variable;
- *anndate*: the date of the earnings announcement (day 0);
- *bhar01*: stock return, measured from *date* (day 0) until the next trading day (day 1).

The fourth dataset is "**big4.dta**", a panel dataset that contains information about whether or not a company's financial statements were audited by a large ("Big 4") audit firm during the year.

- *compustatcode*: unique company identifier variable;
- *year*: the calendar year in which the financial statements were audited;
- *big4auditor*: dummy variable set equal to 1 if the company was audited by a Big 4 auditor (PwC, EY, KPMG, or Deloitte), 0 otherwise.

Required:

a) Use the combined information in these datasets to perform the OLS regression that is specified in equation (1) above. Make sure to let Stata calculate regression standard errors that are robust to heteroscedasticity and clustering at the company-level. Report the regression results. What do the results tell you with respect to HYPOTHESIS 2?
   - Before running the regression, make sure to eliminate from your sample all observations with missing values for one of the relevant variables, and for which the stock price is smaller than $2.50. Also eliminate all observations for which the absolute value of *surprise* is greater than 0.05. Finally, winsorize the stock return variable *bhar01* at the 1st and 99th percentiles.

b) Create an interaction regression in order to test the following hypothesis. Report and interpret the results:

HYPOTHESIS 3 (alternative): Accounting information is less relevant for stock market valuations when companies do *not* have a Big 4 auditor.

c) Create five groups ("quintiles") of observations based on the level of the $SURPRISE_{iq}$ variable. For each of these groups, compute the β coefficient from equation (1). Create a bar (column) chart in which you plot the coefficients for each of the quintiles.[4] What do the differences in coefficients across the groups tell you about the overall relation between earnings news and stock market reactions?
   - Note: you do not have to assess the statistical significance of the coefficients for this question.

[4] You do not have to use Stata to create this graph. You may use Excel, or any other program you prefer, and paste the graph into your solutions document.