Instructions for

Tutorial 2

Empirical Methods in Finance
Academic Year 2021/2022

For this tutorial you are asked to replicate, on a different sample period and with some modifications, selected parts of the article by DellaVigna and Pollet, "Investor Inattention and Friday Earnings Announcements", Journal of Finance, Vol. LXIV, No.2, 2009.

1. **Data preparation.** You are provided two of the necessary datasets for this exercise: summary analyst EPS forecasts and actual EPS, both from IBES. In addition, you will need to download daily stock prices/returns and daily index returns from CRSP. The sample period for forecasts and actuals is Jan 2000 to Dec 2016. You are also provided with STATA code that manipulates the data and runs regressions, except that certain lines of the code are blanked and replaced with dots ("..."). These are the parts that you need to fill in in order to complete the exercise. In the code there are five lines in which intermediate files are saved (eg "save dataset-1, replace"). These intermediate files are also provided to you so that you can proceed with the exercise in case you are not able to correctly fill in all of the blanks. The data preparation proceeds as follows:
   1. Merge IBES forecasts and actuals.
   2. Merge in the closing stock price from five trading days (or one calendar week) prior to the earnings announcement. If there is no price available for that day, use trading day -6, then day -7. If there is no match for any of the trading days -5, -6, or -7, leave the observation unmatched.
   3. Compute the earnings surprise as (actual - forecast)/price. Then divide the sample into 11 bins sorted in ascending order of the

earnings surprise where the middle bin contains all observations with exactly zero surprise.
4. Expand the dataset to accommodate 35 trading days around the earnings announcement, and merge in the index returns for those days.
5. Define a new variable that counts event time, i.e. a variable taking the value of zero on the announcement day, +1 on the day following the announcement, -1 on the day preceding the announcement, etc.
6. Merge in stock returns and index returns and restrict the sample to [-20,+20] trading days around the announcement.
7. Construct cumulative abnormal returns (CARs) around the announcement day normalizing the CAR to zero on the trading day just before the announcement, i.e. event day -1. The CAR of stock *i* in period *t* (when *t* is in the post-announcement period) is computed as

$$CAR_i[-1,t] = \sum_{s=0}^{t}(r_{i,s} - r_{mkt,s})$$

8. Construct a variable for the three-day abnormal return around the announcement: $CAR_i[-1,+1] = \sum_{s=-1}^{1}(r_{i,s} - r_{mkt,s})$.
You may, in addition, try the two-day abnormal return as in DellaVigna and Pollet).

2. **Descriptive statistics.** Provide a table of summary statistics, analogous to DellaVigna and Pollet's Table I, Panel A and C (you do not need to produce Panel B). Which numbers or patterns of numbers in this table do you find notable, reassuring or suspicious in the sense that they may support or question the validity of the latter analysis?

3. **Illustration of the market reaction to surprises.** Compute, for each event day and earnings surprise bin, the average CAR. Then plot the average CARs (one line for each bin) against event time. See, e.g. Foster, Olsen and Shevlin, Accounting Review, 1984,

Figure 1, or Bernard and Thomas, Journal of Accounting Research, 1989, for how the figure should approximately look like.

4. **Illustration of immediate response, Fridays vs non-Fridays.** Compute, for each earnings surprise bin and separately for Friday and non-Friday announcements, the average three-day CAR. Produce a graph analogous to DellaVigna and Pollet's Figure 1(a). Because you are working with a different sample, it does not have to look exactly like the figure in the article.

5. **Regression Analysis.** To test whether the immediate market response to earnings surprises announced on a Friday is slower than for those announced on other weekdays, produce a table similar to DellaVigna and Pollet, Table II, Panel A. Use three different specifications for the regressions:

$$CAR[-1/+1]_{t,k} = \beta_0 + \beta_1 Fri_{t,k} + \beta_2 SurQ_{t,k} + \beta_3 Fri_{t,k} \times SurQ_{t,k} + \varepsilon_{t,k} \quad (1)$$

$$CAR[-1/+1]_{t,k} = \beta_0 + \beta_1 Fri_{t,k} + \beta_2 SurQ11_{t,k} + \beta_3 Fri_{t,k} \times SurQ11_{t,k} + \varepsilon_{t,k} \quad (2)$$

$$CAR[-1/+1]_{t,k} = \beta_0 + \beta_1 Fri_{t,k} + \beta_2 SurQTop_{t,k} + \beta_3 Fri_{t,k} \times SurQTop_{t,k} + \varepsilon_{t,k} \quad (3)$$

In the above regression equations, ***Fri*** is a dummy indicating a Friday earnings announcement, ***SurQ*** is the bin number of the earnings surprise, i.e. a number from 1 to 11, ***SurQ11*** is a dummy for bin #11 of earnings surprises, and ***SurQTop*** is a dummy for the two highest bins of earnings surprises, Q10 and Q11. The indices ***t*** and ***k*** stand for the fiscal quarter and the firm, respectively. For regression (2) you need to drop all date in bins 2-10 before running the regression. (Why?) For regression (3) you need to drop bins 3-9. For each regression model, run two regressions, one without additional controls and one with what DellaVigna and Pollet call "standard controls (interacted)". If you do not have all the controls in your dataset, just use those that you can construct from your dataset that you have prepared.

6. **Timing of announcements.** Run the following three regressions:

$$Fri_{t,k} = \beta_0 + \beta_1 1(\text{Sur}_{t,k}<0) + \varepsilon_{t,k},$$
$$Fri_{t,k} = \beta_0 + \beta_1 1(\text{Sur}_{t,k}<q_{25\%}) + \varepsilon_{t,k},$$
$$Fri_{t,k} = \beta_0 + \beta_1 1(\text{Sur}_{t,k}<q_{10\%}) + \varepsilon_{t,k},$$

where $1(\text{Sur}_{t,k}<0)$ is a dummy variable indicating a negative earnings surprise, $1(\text{Sur}_{t,k}< q_{25\%})$ is a dummy variable indicating an earnings surprise below the 25th quantile of surprises, and $1(\text{Sur}_{t,k}< q_{10\%})$ is a dummy variable indicating an earnings surprise below the 10th quantile of surprises. In a fourth regression, include all three dummy variables simultaneously. How do you interpret these results?

7. **"Sticky" forecast errors.** Run regressions of the earnings surprise in quarter *t* on the earnings surprise in quarter *t-1*, *t-2*, *t-3*, and *t-4*:

$$Sur_{t,k} = \beta_0 + \beta_1 \text{Sur}_{t\text{-}1,k} + \varepsilon_{t,k} \quad (1)$$
$$Sur_{t,k} = \beta_0 + \beta_1 \text{Sur}_{t\text{-}2,k} + \varepsilon_{t,k} \quad (2)$$
$$Sur_{t,k} = \beta_0 + \beta_1 \text{Sur}_{t\text{-}3,k} + \varepsilon_{t,k} \quad (3)$$
$$Sur_{t,k} = \beta_0 + \beta_1 \text{Sur}_{t\text{-}4,k} + \varepsilon_{t,k} \quad (4)$$

You do not have do add any particular control variables, but you should think about what kind of fixed effects could be appropriate (and add them to the regression).

How do you interpret the estimate of β_1 in these regressions economically? How much of a surprise is the earnings surprise really? Does β_1 say something about how rationally analysts update their estimates for the next quarter when there was an earnings surprise in the current quarter?

Advanced questions (if you have extra time): What do the regression coefficients in each of the three regression models in Q5 and Q6 measure? Which is the coefficient of (most) interest in each regression? State the economic magnitude of the effects that you estimate in each regression. How should standard errors be computed in these regressions? How do your results compare to those of DellaVigna and Pollet? Do you replicate their findings? You can also try to run the analysis on two different subsamples: the earlier part of the sample that includes data that DellaVigna and Pollet used (Jan 2000 to June 2006) and the later "post-publication"sample (July 2006 to Dec 2016).

For questions 2-7, provide the respective table or figure.

Good luck!