Empirical Methods in Finance - Exercise week 2

WG 08 Subgroup2
Furia Edoardo, Kiks Stan, Kim Jinhyun

Q1. DATA PREPARATION (Dataset -1)

```
* keep only forecasts issued before the announcement of the actual:
keep if statpers<ead
* keep only forecast that is closest to announcement date of the actual:
bysort ticker ead: egen minstatpers=min(statpers)
keep if statpers==minstatpers
drop minstatpers
duplicates report ticker fpe

save dataset-1, replace
```

Here we are interested in forecasts issued only before the announcement. So we have deleted forecasts after announcement. Additionally we have kept forecasts that are closest to the actual.

```
* merge in the IBES-CRSP linking table available at WRDS-->Linking Suite by WRDS--> IBES CRSP Link
clear
use dataset-1
joinby ticker using Linktable_IBES_CRSP, unmatched(master)
tab _merge
drop if merge_!=3
drop_merge
save dataset-1_Linked, replace
```

Now we have merged IBES-CRSP to the dataset-1 by using the ticker.

Q1. DATA PREPARATION (Dataset -2)

```
* merge in stock price one week prior to announcement date

//5 trading days prior (-7 days)
g date = ead-7
duplicates report permno date
duplicates drop permno date, force
joinby permno date using CRSP-Stocks1, unmatched(master) update
tab _merge
drop _merge
count if PRC!=.
```

First, we merged stock price 5 trading day prior data to dataset-1. To do so, we used 5 prior trading date.

```
//6 trading days prior
replace date=date-1 if PRC == .
duplicates report permno date
duplicates drop permno date, force
joinby permno date using CRSP-Stocks1, unmatched(master) update
tab _merge
drop _merge
count if PRC!=.
```

Since 5 trading day prior stock prices were not fully merged, we used 6 trading day prior stock price.

Q1. DATA PREPARATION (Dataset - 2)

```
//7 trading days prior
replace date=date-1 if PRC ==.
duplicates report permno date
duplicates drop permno date, force
joinby permno date using CRSP-Stocks1, unmatched(master) update
tab _merge
drop _merge
rename PRC prc
count if prc!=.

* check variable prc:
count id prc!=.
drop if prc==.

* apply filters according to DellaVigna/Pollet,
*  Hirshleifer et al. and Livnat/Mendenhall
    drop if prc<5
    drop if value>prc
    drop if medest>prc
    drop if prc*shrout<5000
    drop date
    drop ret
    drop shrcd
    drop permco
    drop shrout
    drop if prc==.

save dataset-2, replace
```

Because 6 trading day prior stock prices were not fully merged, we have used 7 trading day prior stock price.

We have applied filters according to the given papers. (DellaVigna/Pollet, Hirshleifer et al. and Linvnat/Mendenhall)

Q1. DATA PREPARATION (Dataset - 3)

```
* calculate earnings surprise
    clear
    use dataset-2
    g Sur=(value-medest)/prc
    drop if Sur==.
    drop prc
```

Compute the earnings surprise as (actual - forecast)/price.

```
* define 11 quantiles of earnings surprises separately by quarter
    * install xtileJ from Judson Caskey's web page
    g quarter = qofd(ead)
    xtileJ SurQ7to11 = Sur if Sur>0, nquantiles(5) by(quarter)
    xtileJ SurQ1to5 = Sur if Sur<0, nquantiles(5) by(quarter)
    //Get the quantile for the whole data instead of separate data(<0 or >0)
    gen SurQ = SurQ7to11 + 6
    replace SurQ = SurQ1to5 if SurQ ==.
    replace SurQ = 6 if SurQ ==.
    drop SurQ1to5 SurQ7to11

save dataset3, replace
```

Then we have divided the sample into 11 bins sorted in ascending order of the earnings surprise. Here the middle bin contains all observations with exactly zero surprise.

Q1. DATA PREPARATION (Dataset - 4)

```
*****************************************************************
* expand dataset to [-35/+35] days
*****************************************************************
    clear
    use dataset-3
    duplicates report ticker fpe
    expand 71
    sort ticker fpe
    by ticker fpe: g period = [_n]
    replace period=36-period if period>=36
    sort ticker fpe period

* fill in calendar dates
    g date = ead+period
    format date %td


* merge in crsp market return
    joinby date using (--> downloaded return crsp data)
    tab _merge
    drop if _merge!= 3
    drop_merge
```

Here dataset was expanded from -35 to 35 days

New variables date was generated by using announcement date and period, and then merged with the CRSP value-weighted index (NYSE/AMEX/NASDAQ/ARCA)

Merge in the index returns for -35 to 35 days.

Q1. DATA PREPARATION (Dataset - 4)

```
*************************************************************************
* define event time such that date zero is the day of the announcement day
*************************************************************************
    drop if vwretd==.
    g period=0 if date==ead
    forvalues i=-35/35 {
        replace period= `i' id date==anndats+`i'
        }
    sort ticker fpe ead date


 * drop observations outside of tradingdays [-20,+20]
    drop if !inrange(period,-20,+20)
    tab period
    bysort ticker fpe: egen numperiods=count(period)
    tab numperiods
    drop if numperiods<21
    drop numperiods
    sort ticker fpe date
    tab period

* merge in stock returns
    joinby permno date using (downloaded crsp_stocks data)
    tab _merge
    drop _merge


* drop forcasts observations with incomplete return data on [-20,+20]
    bysort ticker fpe: egen numobs=count(ret)
    drop if numobs<41
    drop numobs
    drop shrcd permco prc shrout

save dataset-4, replace
```

Event time has been defined as date 0 if the day is the announcement date.

Restrict the sample to [-20,+20] trading days around the announcement date.

CRSP stock data was again merged with the master data.

Trading days that had insufficient return data were dropped

Q1. DATA PREPARATION (Dataset - 5)

```
* code market-adusted CARs around t=0
    clear
    use dataset-4
    sort ticker fpe period
    g AR  = ret-vwretd
    * normalize CAR to zero in t=-1:
    g CAR = 0 if period==-1
    * cumulate forward:
    replace CAR = CAR[_n-1]+AR[_n-1] if CAR==. & period>-1 & CAR[_n-1]!=. & ticker==ticker[_n-1] & fpe==fpe[_n-1]
    * cumulate backward:
    forvalue n=1/19{
        replace CAR = CAR[_n+1]-AR[_n+1] if CAR==. & period<-1 & CAR[_n+1]!=. & ticker==ticker[_n+1] & fpe==fpe[_n+1]
    }

    sort ticker fpe period

* express CAR in percent
    replace CAR  = CAR*100


* define CAR[-1,+1] for each forecast. This should be the same number for each ticker-fpe combination
    bysort ticker fpe:g CAR3days= CAR if period==2
    egen CAR3day = max(CAR3days), by (ticker fpe)
    drop CAR3days
    drop AR


save dataset-5, replace
```

Finally, we have defined 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.)

In such way, we can generate the CAR3day as in DellaVigna and Pollet

Q2. Descriptive statistics

```
* Panel A
clear
use dataset-5
rename anndats ead
duplicates drop permno fpe ead, force
g dow=dow(ead)
tab dow
```

dow	Freq.	Percent	Cum.
1	19,371	11.37	11.37
2	39,831	23.38	34.74
3	44,738	26.26	61.00
4	56,181	32.97	93.97
5	10,273	6.03	100.00
Total	170,394	100.00	

As we can see from this panel, the earning announcements are usually concentrated during Tuesday, Wednesday and Thursday. While, during Monday and Friday, they are significantly less. This is true especially with regard to Friday, for reasons that we will see later in the presentation.

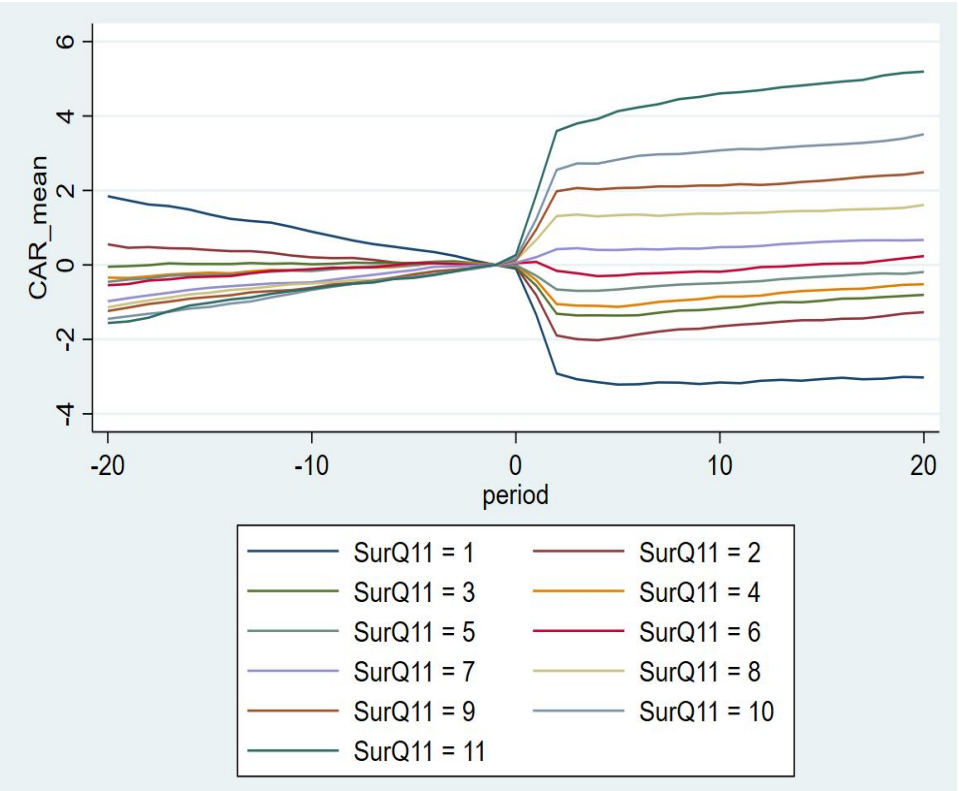
Q2. Descriptive statistics

SurQ11	obs Friday	Other D..	Total
1	-.04653883 1188	-.03664961 14884	-.03738059 16072
2	-.00806464 1101	-.00789086 14935	-.00790279 16036
3	-.00370577 1067	-.00371147 14989	-.00371109 16056
4	-.00178171 1010	-.00173425 15034	-.00173724 16044
5	-.00058898 971	-.00059391 15130	-.00059362 16101
6	0 617	0 11495	0 12112
7	.00047071 832	.00047212 14784	.00047204 15616
8	.00116465 830	.0011551 14758	.00115561 15588
9	.00225159 846	.00217019 14735	.00217461 15581
10	.00431719 853	.00415334 14731	.00416231 15584
11	.01898823 958	.01650121 14646	.0166539 15604
Total	-.00441511 10273	-.00246794 160121	-.00258533 170394

```
* Panel C
g obs = "Other Days"
replace obs="Friday" if dow == 5
tab obs SurQ11 , sum(Sur) mean obs
```

- Descriptive statistics for total observations, observations on Friday and observations on other days.
- The table shows the distribution of observations sorted in quantiles: from 1 to 5 for negative earning surprises, 6 is for earning surprises equal to zero; and from 7 to 11 for positive earning surprises.
- Similar to the paper there is no real difference in earnings surprise for most Qs, however for the two lowest Q there is about 10% difference.

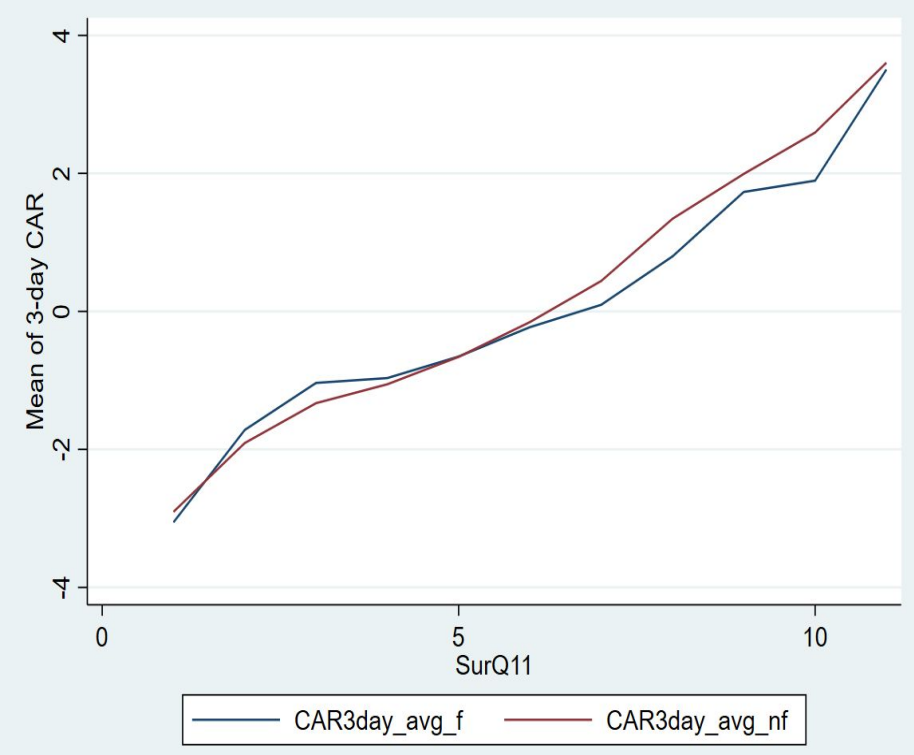
Q3: Illustration of the market reaction to surprises




```
clear
use dataset-5
by period SurQ11, sort: egen CAR_mean = mean(CAR)
duplicates drop period SurQ11, force
xtset period
xtline CAR_mean, overlay t(period) i(SurQ11)
```

- Anticipating the earning announcements, CAR is already changing before day “0”.
- Following the earning announcements, we see that the higher is the quantile, the more positive the reaction to surprises is (and vice versa for bottom quantiles, i.e. more negative)

Q4. Illustration of immediate response, Fridays vs non-Fridays




```
clear
use dataset-5
duplicates drop ticker fpe, force
rename anndats ead
g Fri = dow(ead)==5
* generate average 3-day CAR for each surprise quantile and separately for Fri and non-Fri:
by SurQ11 Fri, sort : egen CAR3day_avg = mean(CAR3day)
duplicates drop SurQ11 Fri, force
g CAR3day_avg_f = CAR3day_avg if Fri == 1
g CAR3day_avg_nf = CAR3day_avg if Fri == 0
line CAR3day_avg_f CAR3day_avg_nf SurQ11, ytitle("Mean of 3-day CAR")
```

- The graphs show the different response to friday announcements compared to non-friday ones.
- As predicted, in the bottom quantiles, the response for fridays is less negative than for non-fridays.
- In the top quintiles, the response for fridays is less positive that for non-fridays

Q5. Regression Analysis: Stata codes

```
* Q5: Regression (slow reaction to Friday surprises?)
    clear
    cap erase temp.txt
    use dataset-5
    g month=month(date)
    g year=year(date)
    duplicates drop ticker fpe, force
    rename anndats ead

    * columns 1&2:
    g Fri = dow(ead)==5
    g FrixSurQ11 = Fri*SurQ11
    gen qyear = quarter(date)
    gen monthxyear = month*year
    reghdfe CAR3day Fri SurQ11 FrixSurQ11, absorb (qyear ticker) cluster(qyear ticker)
    outreg2 using q5w2.doc, adjr bdec(2) tdec(2) label
    reghdfe CAR3day Fri SurQ11 FrixSurQ11 monthxyear, absorb (qyear ticker) cluster (qyear ticker)
    outreg2 using q5w2.doc, adjr bdec(2) tdec(2) label append
    drop FrixSurQ11


    * columns 3&4:
    g SurQTop=(SurQ11==11)
    g FrixSurQTop = Fri*SurQTop
    drop if SurQ11 >=2 & SurQ11 <=10
    reghdfe CAR3day Fri SurQTop FrixSurQTop, absorb (qyear ticker) cluster(qyear ticker)
    outreg2 using q5w2.doc, adjr bdec(2) tdec(2) label append
    reghdfe CAR3day Fri SurQTop FrixSurQTop monthxyear, absorb (qyear ticker) cluster(qyear ticker)
    outreg2 using q5w2.doc, adjr bdec(2) tdec(2) label append
    drop SurQTop
    drop FrixSurQTop

    * columns 5&6:
    g SurQTop2=(SurQ11==10 | SurQ11==11)
    g FrixSurQTop2 = Fri*SurQTop2
    drop if SurQ11 >=3 & SurQ11 <=9
    reghdfe CAR3day Fri SurQTop2 FrixSurQTop2, absorb (qyear ticker) cluster(qyear ticker)
    outreg2 using q5w2.doc, adjr bdec(2) tdec(2) label append
    reghdfe CAR3day Fri SurQTop2 FrixSurQTop2 monthxyear, absorb (qyear ticker) cluster(qyear ticker)
    outreg2 using q5w2.doc, adjr bdec(2) tdec(2) label append
    seeout
```

We generated Fri as a dummy equal to 1 in case of friday; 0 otherwise;

For columns 1&2, we ran the regression, with and without controls, using SurQ11 and its interaction with Fri;

For columns 3&4, we had to drop the observations from quantile 2 to 10, in order to interpret properly the dummy variable SurQTop (top quantile compared with reference group, bottom quantile);

We did the equivalent for columns 5&6 (i.e. top two quantiles compared with the bottom two quantiles).

Q5. Regression Analysis: Regression outputs

VARIABLES	(1) CAR3day	(2) CAR3day	(3) CAR3day	(4) CAR3day	(5) CAR3day	(6) CAR3day
Fri	0.15	0.15	-0.14	-0.14	-0.02	-0.02
	(0.168)	(0.169)	(0.448)	(0.448)	(0.198)	(0.198)
SurQ11	0.60***	0.60***				
	(0.014)	(0.014)				
FrixSurQ11	-0.05**	-0.05**				
	(0.023)	(0.024)				
monthxyear		0.000017		0.000046		0.000045
		(0.000042)		(0.000062)		(0.00050)
SurQTop			6.43***	6.43***		
			(0.099)	(0.099)		
FrixSurQTop			0.29	0.30		
			(0.374)	(0.374)		
SurQTop2					5.45***	5.45***
					(0.10)	(0.10)
FrixSurQTop2					-0.24	-0.24
					(-0.28)	(-0.29)
Constant	-3.43***	-3.64***	-2.85***	-3.41***	-2.37***	-2.92***
	(0.076)	(0.46)	(0.039)	(0.75)	(0.039)	(0.59)
Observations	169,905	169,905	30,453	30,453	62,394	62,394
Adjusted R-squared	0.071	0.071	0.125	0.125	0.106	0.106

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

- As it was expectable, the coefficient on SurQTop is the biggest, followed by SurQTop2 and SurQ11
- The coefficient on friday is not stistically significant (as in DellaVigna and Pollet)
- The coefficient on FrixSur is negative and signficant for regression (1) and (2), as we expected; and negative again for (5) and (6), although not statistically significant (when we restrict the sample the observations on fridays are even fewer than before)

6. Timing of announcements: Stata code

```
* Q6: Regression (timing of announcements?)
    clear
    cap erase temp.txt
    use dataset-5
    duplicates drop ticker fpe, force
    rename anndat ead
//generate dummy variables for 25% and 10% quantile
xtile Sur_4=Sur, nq(4)
xtile Sur_10=Sur, nq(10)
g D25=(Sur_4==1)
g D10=(Sur_10==1)
g D0 = (Sur<0)
g Fri = dow(ead)==5

//create regression output
reg Fri D0, cluster(quarter)
outreg2 using Q6, word dec(3)
reg Fri D25, cluster(quarter)
outreg2 using Q6, word dec(3) append
reg Fri D10, cluster(quarter)
outreg2 using Q6, word dec(3) append
reg Fri D0 D25 D10, cluster(quarter)
outreg2 using Q6, word dec(3) append
```

First, we generated three dummy variables for negative surprise earnings (D0), lowest 10th (D10) and 25th (D25) percentiles. We also created a dummy for friday announcements (equal to 1 in case of Friday; 0 otherwise);

Second, we ran four different regressions, friday on D0, friday on D25, friday on D10 and friday on all the three dummies together, clustering for quarter in every case.

6. Timing of announcement: regression output

VARIABLES	(1) Fri	(2) Fri	(3) Fri	(4) Fri
D0	0.012***			0.008***
	(0.001)			(0.002)
D25		0.013***		0.005**
		(0.002)		(0.002)
D10			0.016***	0.008**
			(0.003)	(0.003)
Constant	0.055***	0.057***	0.059***	0.055***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	170,394	170,394	170,394	170,394
R-squared	0.001	0.001	0.000	0.001

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

- As expected, companies that are about to announce earnings that are lower than expected, tend to do it on friday more often.
- This is due to the fact that, as we saw previously in the presentation, earnings announcements on fridays tend to have a less pronounced response.
- Thus, companies in case of negative surprises will have an incentive to announce them on Friday.
- In column 4 we need to interpret the data differently because the quantiles overlap and the coefficients correct for this. This results in lower coefficients.

7. “Sticky” forecast errors: Stata codes

```
ssc install reghdfe
ssc install ftools
ssc install winsor2
clear
cap erase temp.txt
use dataset-5
duplicates drop ticker fpe, force
* winsorize or truncate Sur:
winsor2 Sur,suffix(_win) cuts(0.5 99.5)
* define quarter variable and tsset data:
g qtr = qofd(anndats)
format qtr %tq
egen ID =group(ticker)
sort ID qtr
duplicates drop  ticker qtr,force
tsset ID qtr

* define lags of the earnings surprise:
g SurL1 = l.Sur
g SurL2 = l2.Sur
g SurL3 = l3.Sur
g SurL4 = l4.Sur
```

```
* define lags of the earnings surprise:
g SurL1 = l.Sur
g SurL2 = l2.Sur
g SurL3 = l3.Sur
g SurL4 = l4.Sur

* Regress using company ID and quarterly fixed effects
reghdfe Sur_win SurL1, absorb(ID qtr) cluster(ID qtr)
est store S1
reghdfe Sur_win SurL2, absorb(ID qtr) cluster(ID qtr)
est store S2
reghdfe Sur_win SurL3, absorb(ID qtr) cluster(ID qtr)
est store S3
reghdfe Sur_win SurL4, absorb(ID qtr) cluster(ID qtr)
est store S4
```

7.“Sticky forecast errors”: regression outputs

VARIABLES	(1) Sur	(2) Sur	(3) Sur	(4) Sur
SurL1	0.628*** (22.07)			
SurL2		0.352*** (15.50)		
SurL3			0.254*** (12.16)	
SurL4				0.268*** (12.19)
Constant	-0.002*** (-27.10)	-0.002*** (-30.73)	-0.002*** (-28.99)	-0.002*** (-27.46)
Observations	153,381	145,928	139,669	134,990
R-squared	0.067	0.021	0.011	0.012

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

- Companies that had an earning surprise in previous quarters tend to still have an earning surprise in the current quarter (effect gradually dissipates going backward with the quarters, as we can see from the coefficients).
- This implies that earnings surprises in a defined quarter can be, at least to a certain extent, predictable on the basis of previous (if any) quarters with earning surprises (i.e. they are not real surprises…).
- Moreover, the latter consideration arises doubts about the rationality of analysts (i.e. given this consistency in surprises in following quarters, they probably do not update their own forecast completely rationally)

References

- https://onlinelibrary.wiley.com/doi/full/10.1111/j.1540-6261.2009.01447.x
- https://wrds-www.wharton.upenn.edu/pages/get-data/center-research-security-prices-crsp/