

SPL: Fama French Factor Model

Xin Tan, Daria Skidnova, Jeff Giddens

2018/6/9

Structure

1. Model
2. Data
3. Results

1. Model

► CAPM $R_i - R_F = \beta \cdot (R_M - R_F)$

► Fama and French (1993) 3 Factors

$$R_i - R_F = \beta_M \cdot (R_M - R_F) + \beta_S \cdot SMB + \beta_V \cdot HML$$

► Fama and French (2015) 5 Factors

$$R_i - R_F = \beta_M \cdot (R_M - R_F) + \beta_S \cdot SMB + \beta_V \cdot HML + \beta_P \cdot RMW + \beta_I \cdot CMA$$

2. Data

- ▶ Model factors including the estimated market return and the risk-free rate from Kenneth French's data library: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
- ▶ S&P 500 stock prices downloaded with BatchGetSymbols library.

```
library(BatchGetSymbols)
Companies <- GetSP500Stocks()
Stocks<- BatchGetSymbols(tickers = Companies$tickers,
                        first.date = "2010-01-01",
                        last.date = "2017-12-31")
```

2. Data: S&P 500

- ▶ `GetSP500Stocks()` returns company info like name and sector.
- ▶ Downloaded stock data contains two dataframes:
 - df.control** descriptive info. e.g. whether the download for the ticker is successful.
 - df.tickers** downloaded price data, stacked in rows.
- ▶ `BatchGetSymbols()` function **keeps** stocks with 75% or more price points compared to the benchmark.
We took *3M Company's* dates as they are complete.

```
good.tickers <- Stocks$df.control$  
  ticker[Stocks$df.control$threshold.decision=="KEEP"]  
  
SP500.data<-data.frame(date = Stocks$df.tickers$  
  ref.date[Stocks$df.tickers$ticker=="MMM"])
```


2. Data: from Daily Price to Monthly Return

`quantmod::monthlyReturn()` requires non-NA daily prices in `xts` format.

```
Stock.Prices.Daily <-  
  xts(Stock.Prices.Daily[,-1],  
      order.by = as.POSIXct(Stock.Prices.Daily$date))  
  
Stock.Prices.Daily <-  
  Stock.Prices.Daily[!is.na(Stock.Prices.Daily)]  
  
Stock.Prices.Monthly <- monthlyReturn(Stock.Prices.Daily)
```

How to chose?

- ▶ **Remove stocks with NA** ensures remaining stocks have same number of observations.
- ▶ **Remove NAs in each series** results in a larger sample size.

Price data with NAs in the middle might result in strange monthly returns.
(“BHY” Brighthouse Financial Inc. removed for 2015-2017 runs)

2. Data: Linear Regression

► Linear Regression in R

```
y <- lm(rirf ~ rmrf + smb + hml);  
round(summary(y)$coefficients, digits = 4)
```

##	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	-0.3818	0.1075	-3.5504	4e-04
## rmrf	1.0349	0.0264	39.2255	0e+00
## smb	1.3985	0.0393	35.6070	0e+00
## hml	-0.2979	0.0440	-6.7676	0e+00

- We can read out the regression summary from the coefficients matrix.
- Similarly R^2 from `summary(y)$r.squared` and standard errors from `summary(y)$sigma`.

3. Results

Overview:

- ▶ Replicating Fama and French (1993)
- ▶ Regression over S&P500 stocks
- ▶ 3 Factors vs. 5 Factors
- ▶ Stock selection?

3.1 Results: Replicating the 3 Factors Model

<i>b</i>	LOW	2	3	4	HIGH
SMALL	1.03	0.97	0.94	0.89	0.95
2	1.10	1.02	0.96	0.97	1.07
3	1.10	1.02	0.97	0.97	1.06
4	1.06	1.07	1.04	1.03	1.15
BIG	0.96	1.02	0.96	1.01	1.03

<i>s</i>	LOW	2	3	4	HIGH
SMALL	1.40	1.27	1.16	1.10	1.19
2	1.00	0.94	0.83	0.71	0.85
3	0.70	0.63	0.54	0.45	0.65
4	0.30	0.27	0.25	0.22	0.36
BIG	(0.20)	(0.19)	(0.27)	(0.19)	(0.04)

<i>h</i>	LOW	2	3	4	HIGH
SMALL	(0.30)	0.08	0.27	0.38	0.62
2	(0.48)	0.03	0.23	0.47	0.70
3	(0.43)	0.04	0.31	0.50	0.71
4	(0.44)	0.03	0.30	0.56	0.74
BIG	(0.44)	(0.02)	0.20	0.56	0.76

<i>t(b)</i>	LOW	2	3	4	HIGH
SMALL	39.23	50.60	58.42	57.99	57.76
2	53.20	58.56	59.98	62.77	63.25
3	59.68	56.81	53.35	58.93	51.14
4	57.16	52.61	50.34	51.30	46.30
BIG	57.20	56.98	42.80	55.04	37.70

<i>t(s)</i>	LOW	2	3	4	HIGH
SMALL	35.61	44.82	48.65	48.10	48.63
2	32.62	36.36	34.80	30.78	33.82
3	25.53	23.41	20.04	18.46	21.03
4	10.92	8.75	8.06	7.49	9.64
BIG	(8.10)	(7.08)	(7.99)	(6.91)	(1.05)

<i>t(h)</i>	LOW	2	3	4	HIGH
SMALL	(6.77)	2.43	9.92	14.93	22.43
2	(13.93)	0.88	8.73	18.32	24.74
3	(14.04)	1.39	10.27	18.28	20.34
4	(14.24)	0.79	8.77	16.68	17.79
BIG	(15.96)	(0.68)	5.25	18.41	16.65

3.1 Results: Replicating the 3 Factors Model

Results from Fama and French (1993):

Table 6

Regressions of excess stock and bond returns (in percent) on the excess market return ($RM - RF$) and the mimicking returns for the size (SMB) and book-to-market equity (HML) factors: July 1963 to December 1991, 342 months.^a

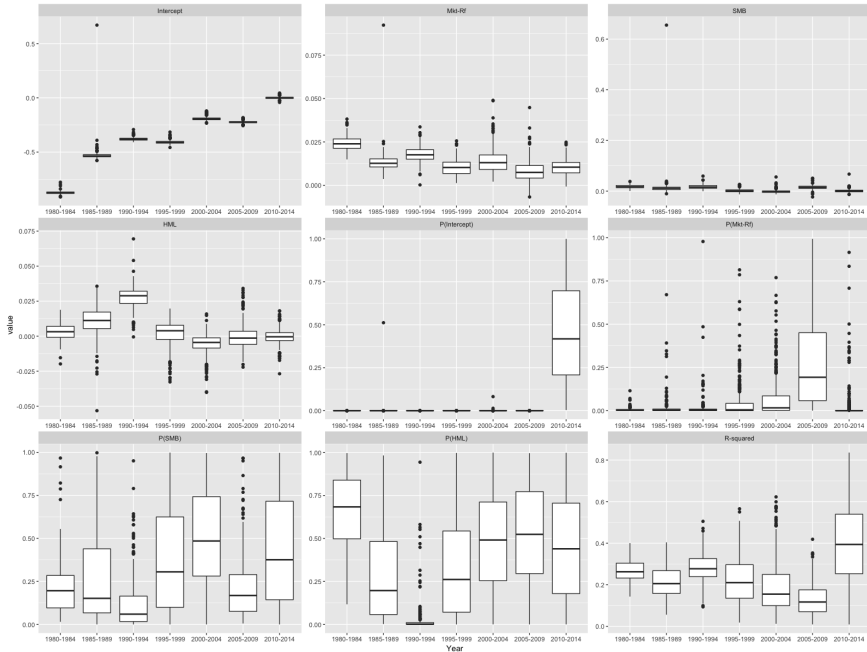
$$R(t) - RF(t) = a + b[RM(t) - RF(t)] + sSMB(t) + hHML(t) + e(t)$$

Dependent variable: Excess returns on 25 stock portfolios formed on size and book-to-market equity

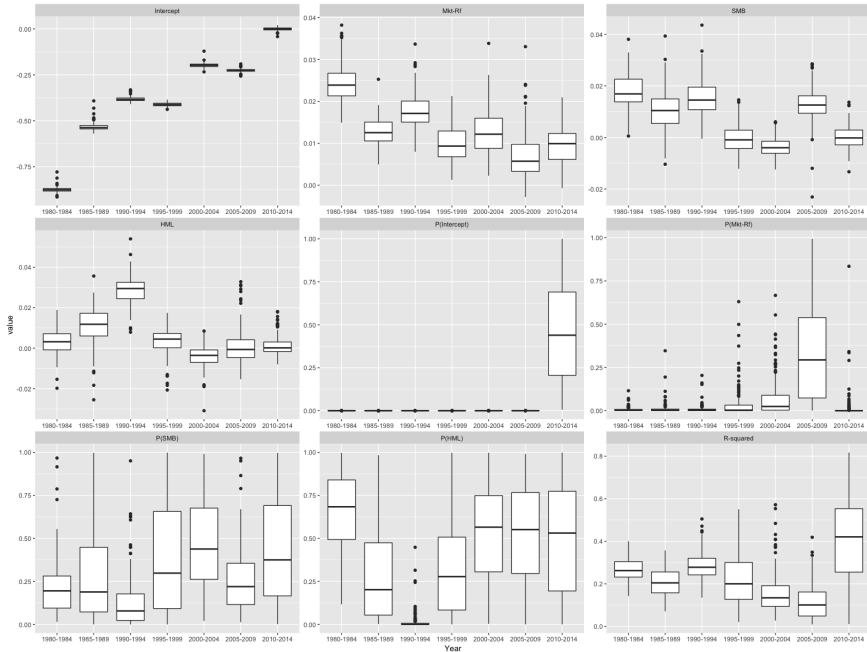
Book-to-market equity (BE/ME) quintiles

Size quintile	Low	2	3	4	High	Low	2	3	4	High
<i>b</i>						<i>t(b)</i>				
Small	1.04	1.02	0.95	0.91	0.96	39.37	51.80	60.44	59.73	57.89
2	1.11	1.06	1.00	0.97	1.09	52.49	61.18	55.88	61.54	65.52
3	1.12	1.02	0.98	0.97	1.09	56.88	53.17	50.78	54.38	52.52
4	1.07	1.08	1.04	1.05	1.18	53.94	53.51	51.21	47.09	46.10
Big	0.96	1.02	0.98	0.99	1.06	60.93	56.76	46.57	53.87	38.61
<i>s</i>						<i>t(s)</i>				
Small	1.46	1.26	1.19	1.17	1.23	37.92	44.11	52.03	52.85	50.97
2	1.00	0.98	0.88	0.73	0.89	32.73	38.79	34.03	31.66	36.78
3	0.76	0.65	0.60	0.48	0.66	26.40	23.39	21.23	18.62	21.91
4	0.37	0.33	0.29	0.24	0.41	12.73	11.11	9.81	7.38	11.01
Big	-0.17	-0.12	-0.23	-0.17	-0.05	-7.18	-4.51	-7.58	-6.27	-1.18
<i>h</i>						<i>t(h)</i>				
Small	-0.29	0.08	0.26	0.40	0.62	-6.47	2.35	9.66	15.53	22.24
2	-0.52	0.01	0.26	0.46	0.70	-14.57	0.41	8.56	17.24	24.80
3	-0.38	0.00	0.32	0.51	0.68	-11.26	0.05	9.75	16.88	19.39
4	-0.42	0.04	0.30	0.56	0.74	-12.51	1.04	8.83	14.84	17.09
Big	-0.46	0.00	0.21	0.57	0.76	-17.03	0.09	5.80	18.34	16.24

3.2 Regression over S&P 500 stocks (All)

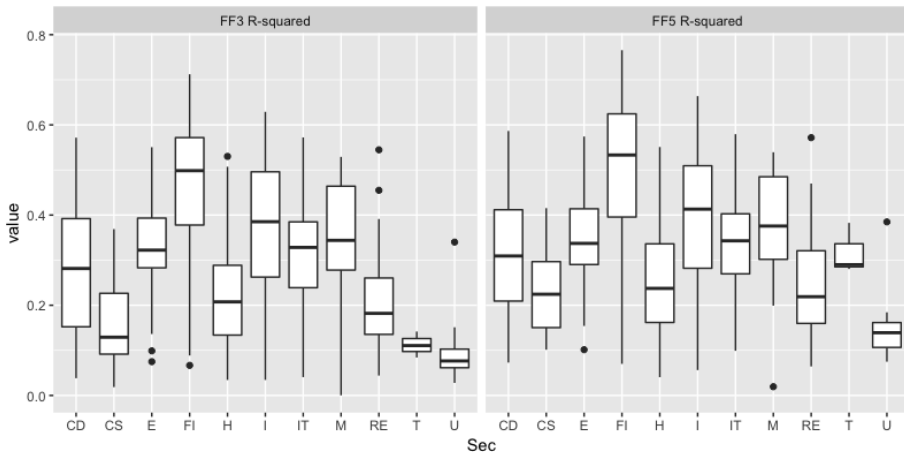


3.2 Regression over S&P 500 stocks (Survivors)



3.3 3 Factors vs. 5 Factors

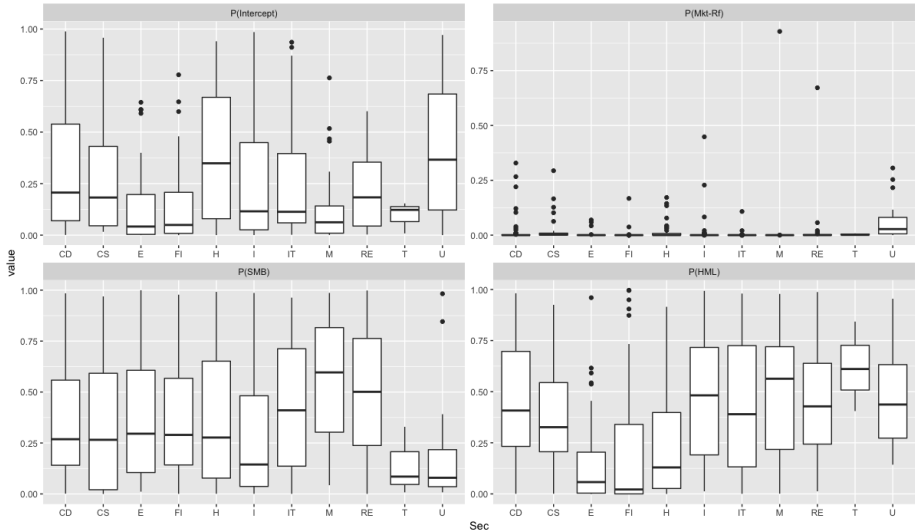
► R^2 Comparison



CD: Consumer Discretionary, CS: Consumer Staples, E: Energy, FI: Financials, H: Health Care, I: Industrials, IT: Information Technology, M: Materials, RE: Real Estate, T: Telecommunication Services, U: Utilities

3.3 3 Factors vs. 5 Factors

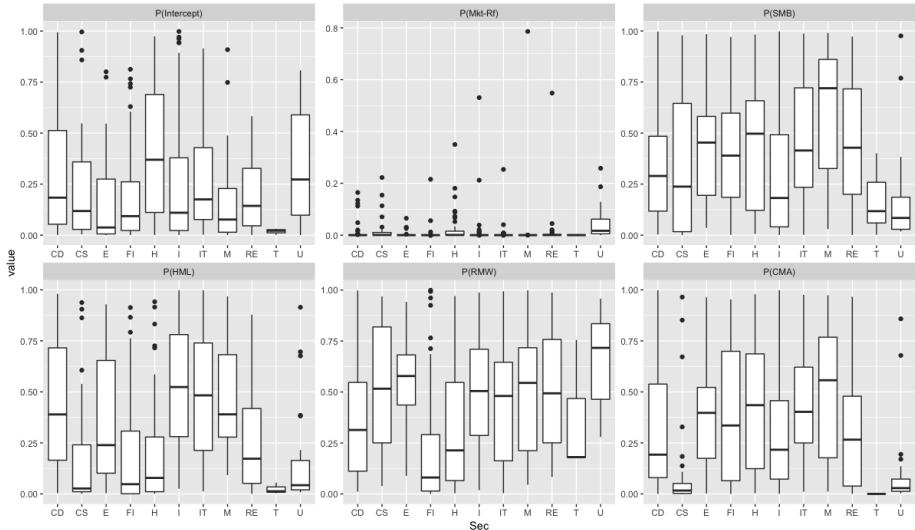
► P-Values: 3 Factors



CD: Consumer Discretionary, CS: Consumer Staples, E: Energy, FI: Financials, H: Health Care, I: Industrials, IT: Information Technology, M: Materials, RE: Real Estate, T: Telecommunication Services, U: Utilities

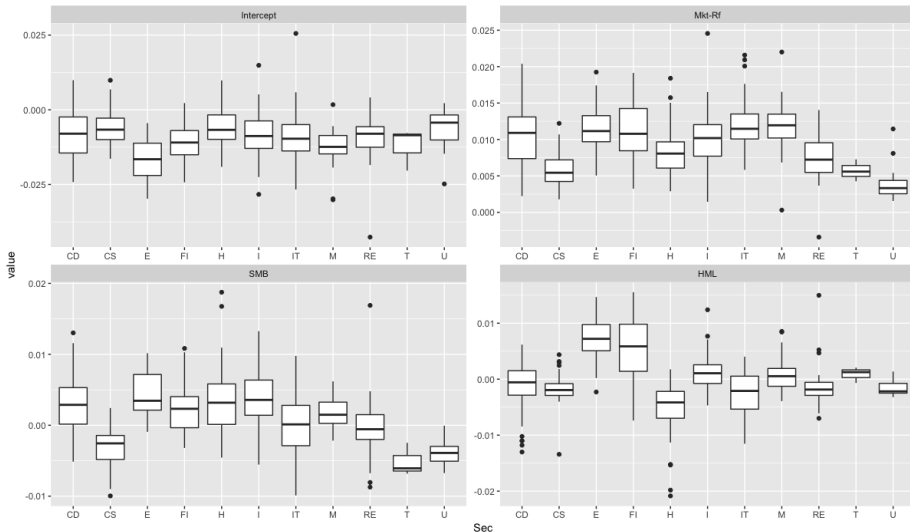
3.3 3 Factors vs. 5 Factors

► P-Values: 5 Factors



3.3 3 Factors vs. 5 Factors

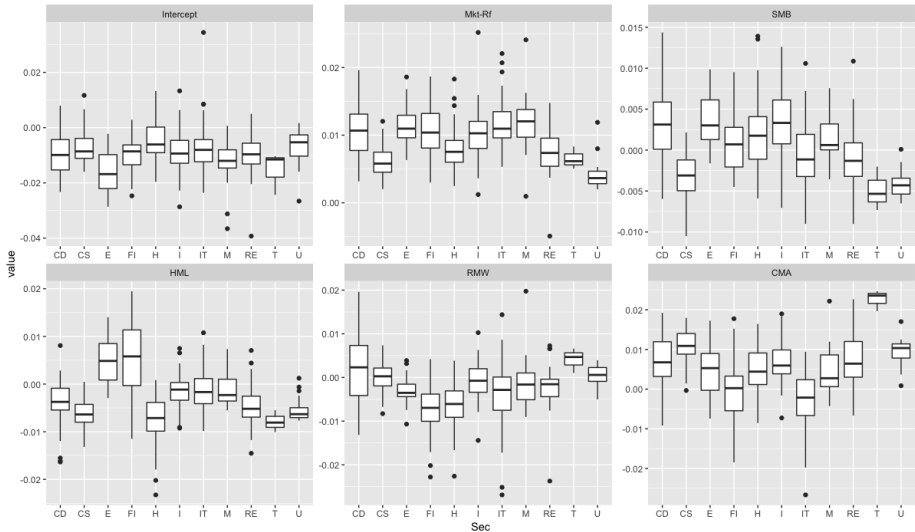
► Regressed Coefficients: 3 Factors



CD: Consumer Discretionary, CS: Consumer Staples, E: Energy, FI: Financials, H: Health Care, I: Industrials, IT: Information Technology, M: Materials, RE: Real Estate, T: Telecommunication Services, U: Utilities

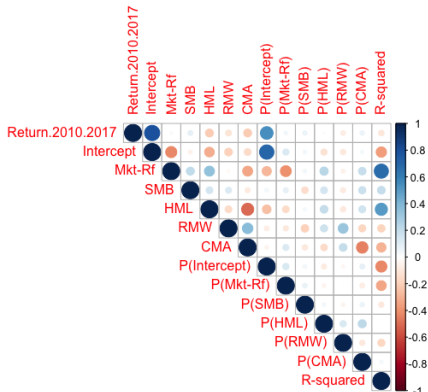
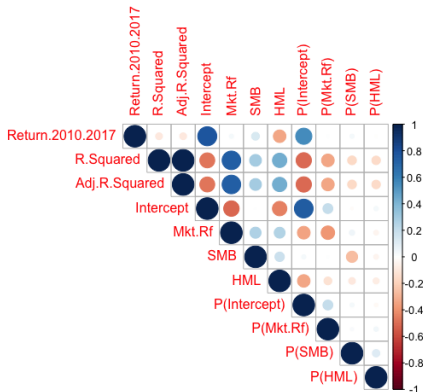
3.3 3 Factors vs. 5 Factors

► Regressed Coefficients: 5 Factors



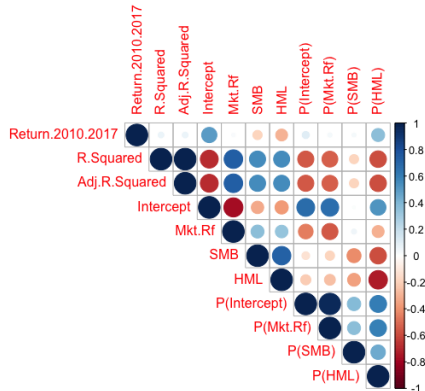
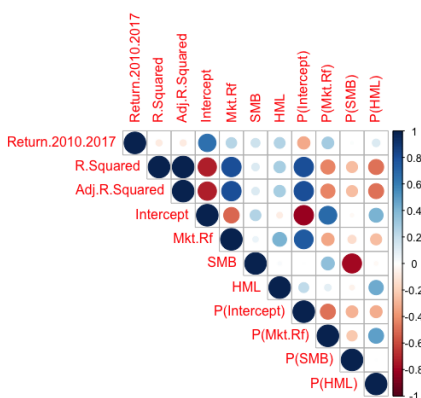
3.4 Stock Selection?

- ▶ Stock returns from Jan. 2010 to Dec. 2017.
Corr. plot of the stock returns and the regressed coefficients.
- ▶ *alpha* or *Intercept* highly correlated to stock returns.
- ▶ Positive corr. to SMB: Small cap premium.
Negative corr. to HML: Market seemed to favor **Growth** stocks (instead of **Value**) in the bullish years.
- ▶ Negative corr. to RMW and CMA.



3.4 Stock Selection: 3 Factor / Top and Bottom 20s

- ▶ Winners have positive correlations to both *SMB* and *HML*, while losers have negative correlations.
- ▶ Small Cap and Value premiums work at the extremes.



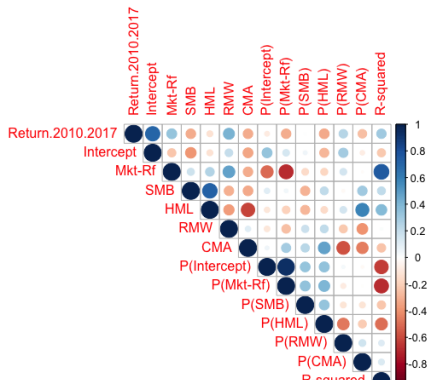
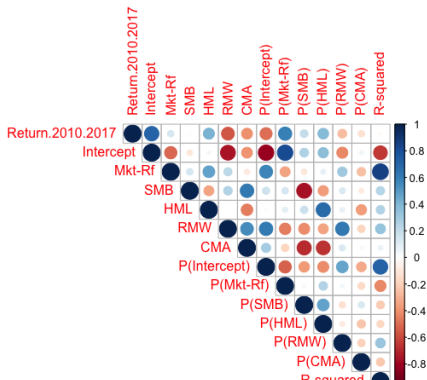
Although the overall winner, *NETFLIX*, is clearly neither *small cap* nor *value*.

3.4 Stock Selection: 5 Factor / Top and Bottom 20s

- **Profitability** factor *RMW*: negatively corr. to returns of winners and positively corr. to that of losers: might not be good indicator for stock selection.

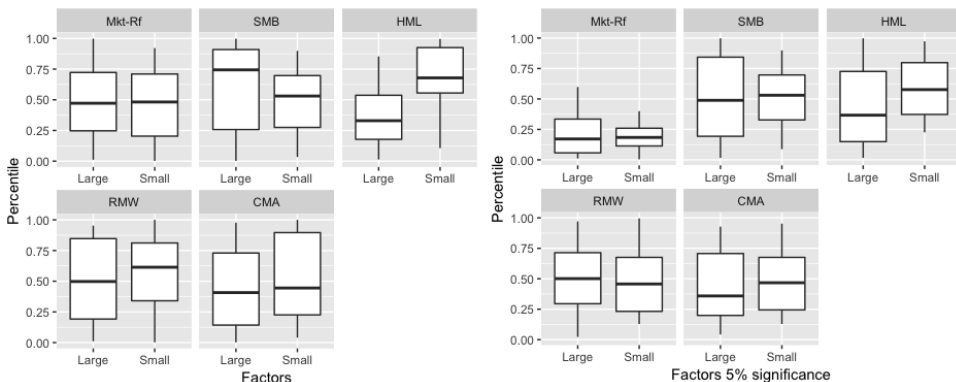
Negative regressed beta for *RMW* for certain sectors: companies might have poor near-term operating performance but very high future potentials, which is *priced-in* to the current stock prices.

- **Investment** factor *CMA* is negatively correlated to the returns of both winners and losers: both winners and losers invest aggressively?



3.4 Stock Selection: 5 Factor / Top and Bottom Factors

- ▶ Calculate return percentile for each stock (best return ~ 1 , worst ~ 0).
- ▶ Filter out largest and smallest estimated coefficient for each factor.
- ▶ With and without considering the significance level of 5%.



- ▶ Not considering significance, market seems to favor **growth** stocks.
- ▶ Most of the stocks with lowest estimated coefficient for *HML* belongs to *Health Care* or *Consumer Discretionary*.

Thank You!

Below are backup slides

Melt and Plotting

- ▶ Regression results in dataframe Results.
- ▶ `reshape::melt()` prepares the data for plots.
- ▶ Each Fama French factor will have 1 sub-plot comparing different sectors

```
df.melt <- melt(  
  Results[,c("Intercept", "Mkt-Rf", "SMB", "HML", "RMW", "CMA", "Sec")],  
  "Sec"  
)
```

- ▶ Boxplot with the melted data

```
ggplot(df.melt, aes(x=Sec, y=value))  
+ geom_boxplot()  
+ facet_wrap(~ variable, scales='free')
```


Melt and Plotting

- Illustration of `reshape::melt()`:
Data is reshaped into one dimensional vector with group / subgroups in additional columns.

```
df.melt <- melt(  
  Results[,c("Intercept", "Mkt-Rf", "SMB", "HML", "RMW", "CMA", "Sec")],  
  "Sec"  
)
```

source data

	Intercept	Mkt-Rf	SMB	HML	RMW	CMA	Sec
1	-0.011	0.010	0.002	-0.002	0.003	0.004	I
2	-0.012	0.010	-0.003	-0.007	0.001	0.003	H
3	-0.008	0.011	-0.004	-0.003	-0.005	0.000	IT
4	-0.001	0.010	-0.008	-0.009	-0.010	0.000	IT
5	-0.005	0.012	0.010	-0.009	-0.014	0.010	I

df.melt

	Sec	variable	value
1	I	Intercept	-0.011
2	H	Intercept	-0.012
3	IT	Intercept	-0.008
...			
464	I	Mkt-Rf	0.010
465	H	Mkt-Rf	0.010
466	IT	Mkt-Rf	0.011
...			
2776	CD	CMA	0.001
2777	H	CMA	0.012
2778	FI	CMA	-0.006

Helper functions

- ▶ Resizing the regression output from an 1 dimensional vector to the same format as in the papers with rows for *SMB* and columns for *HML*.

```
resize <- function(x)
{
  return(
    matrix(x, nrow=5, byrow = TRUE,
           dimnames = list(
             c("SMALL", "2", "3", "4", "BIG"),
             c("LOW", "2", "3", "4", "HIGH")
           )
    )
  )
}
```

Descriptions copied from K.French's Website

Description of the factors: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_5_factors_2x3.html

- ▶ SMB (Small Minus Big) is the average return on the nine small stock portfolios minus the average return on the nine big stock portfolios.
- ▶ HML (High Minus Low) is the average return on the two value portfolios minus the average return on the two growth portfolios.
- ▶ RMW (Robust Minus Weak) is the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios.
- ▶ MA (Conservative Minus Aggressive) is the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios.

Descriptions copied from K.French's Website (cont.)

- ▶ $R_m - R_f$, the excess return on the market, value-weight return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ that have a CRSP share code of 10 or 11 at the beginning of month t , good shares and price data at the beginning of t , and good return data for t minus the **one-month Treasury bill rate** (from Ibbotson Associates).

References

Fama, Eugene F., and Kenneth R. French. 1993. "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics* 33 (1): 3–56. doi:10.1016/0304-405X(93)90023-5.

———. 2015. "A five-factor asset pricing model." *Journal of Financial Economics* 116 (1). Elsevier: 1–22. doi:10.1016/j.jfineco.2014.10.010.