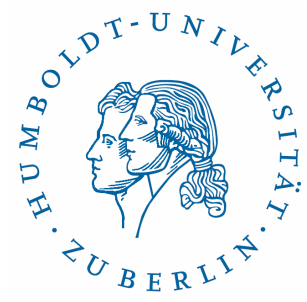
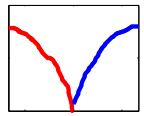
Replication of the Fama French 3 Factor Model and Relevance Today

Xin Tan
Daria Skidnova
Jeff Giddens

Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. – Center for Applied Statistics
and Economics
Humboldt–Universität zu Berlin
http://lvb.wiwi.hu-berlin.de
http://case.hu-berlin.de

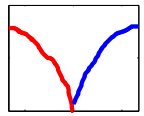


Structure I



Structure II

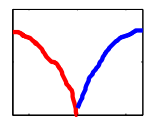
Varying Samples
Correlation of The factors and Top 20 and Bottom 20 returns



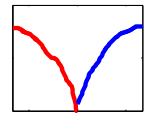
What is the Fama French model?

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

Where $R_I - R_F$ is an excess return of the stock or portfolio; SMB is a small minus big (in terms of capitalization); HML is high minus low (in terms of book-to-value ratio).

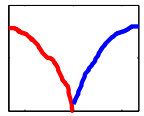


Purpose of the Study



Purpose of the Study

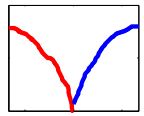
Replicate the model on 25 value-weighted portfolios from July 1963 to December 1991 and check it on updated data from 1980 till 2015 for portfolios of companies within SP 500.



Theory and Design

Our Data:

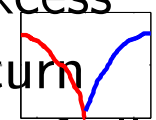
- 3 Factors 1926.07.01 to 2018.03.29 as daily / weekly / monthly data
- 25 Portfolios (5x5) formed on Size and Book-to-Market 1926.07 to 2018.03 corresponding to the Fama and French (1993) 3-factor setup



Data

	Mkt-RF	SMB	HML	RF
192607	2.96	-2.30	-2.87	0.22
192608	2.64	-1.40	4.19	0.25
192609	0.36	-1.32	0.01	0.23
192610	-3.24	0.04	0.51	0.32
192611	2.53	-0.20	-0.35	0.31
192612	2.62	-0.04	-0.02	0.28
192701	-0.06	-0.56	4.83	0.25
192702	4.18	-0.10	3.17	0.26

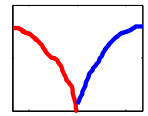
Where SMB (Small minus Big) is the average return on the three small portfolios minus the average return on the three big portfolios; HML is the average return on the two value portfolios minus the average return on the two growth portfolios; Mkt-RF is the excess return on the market shows the difference between market return and free-risk rate. The first component is value-weight return of all



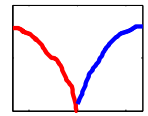
Data

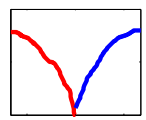
	Book Equity to Market Equity				
Capitalization (Market Equity)	Small ME; Low B/M	2	3	4	Small ME; High B/M
	2			2	
	3			3	
	4			4	
	Big ME; Low B/M	2	3	4	Big ME; High B/M

The 25 portfolios are formed at the end of June of each year. They are the intersection of five portfolios constructed on size, i.e. ME, and five portfolios formed on the book-to-market equity ratio, i.e. BE/ME.



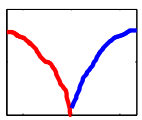
Replication of the Three-Factor Model

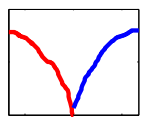
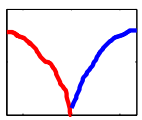


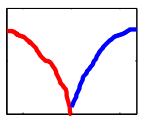
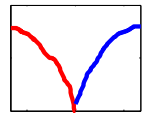
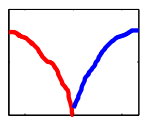
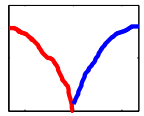
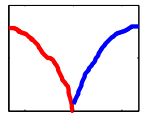
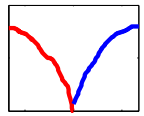


Batch Regression

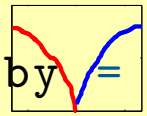
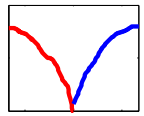
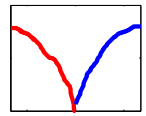
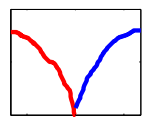
Using multivariate regression. To run regression on each portfolio and saves the results in a list.

```
```{r}
# batch regressing 25 portfolios
results <- list()
# Data starts from the 2nd col of P25
for (i in 1:(ncol(P25) - 1)) {
  rirf <- unlist (P25[, i + 1]) - rf
     y <- lm (rirf ~ rmrf + smb + hml)
  results[[i]] <- summary (y)
}
```
```



Formatting the Results (Betas and t-value)

```
```{r}
betas <- vector()
std.errors <- vector()
t.values <- vector()
R.squareds <- vector()
# save all betas
for(i in 1:(ncol(P25)-1)) {
  betas <- cbind(betas,results[[i]]$coefficients
    [,1])
  std.errors <- cbind(std.errors,results[[i]]$sigma)
  t.values <- cbind(t.values, results[[i]]
    $coefficients[,3])
  R.squareds <- cbind(R.squareds, results[[i]]$adj.r
    .squared)
}
```



```

Formatting the Results (Betas and t-value)

```
# resize the output to 5x5 format like Fama French paper
resize <- function(x) {
  df = data.frame (matrix(x, nrow=5, byrow = TRUE))
  colnames(df) = c("Low", "2", "3", "4", "High")
  rownames(df) = c("Small", "2", "3", "4", "Big")
  return(df)
}
```




Formatting the Results (Betas and t-value)

Resize the results to keep the same display as the Table 6 in the paper

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

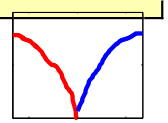
# resize betas
market.beta <- resize(betas[2, ])
SMB.beta <- resize(betas[3, ])
HML.beta <- resize(betas[4, ])
```




b	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

t(b)	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

s	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)

t(s)	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)

h	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

t(h)	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

Figure 1: Results of the Fama French 3 Factor Model



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

Size quintile	Book-to-market equity (BE/ME) quintiles									
	Low	2	3	4	High	Low	2	3	4	High
	b					$t(b)$				
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
	s					$t(s)$				
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
	h					$t(h)$				
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

Figure 2: Results of the Fama French 3 Factor Model in Fama and French 1993b





Retrieving SP500 Data

Download the SP500 stocks from Yahoo finance

```
companies<-GetSP500Stocks()  # get S&P 500 stocks
  with sectors information
head(Companies, n=5)
```

tickers	company	SEC.filings	GICS.Sector	GICS.Sub.Industry
MMM	3M Company	reports	Industrials	Industrial Conglomerates
ABT	Abbott Laboratories	reports	Health Care	Health Care Equipment
ABBV	AbbVie Inc.	reports	Health Care	Pharmaceuticals
ABMD	ABIOMED Inc	reports	Health Care	Health Care Equipment
ACN	Accenture plc	reports	Information Technology	IT Consulting & Other Services

Figure 3: example of the data collected



Retrieving SP500 Data

```
stocks<-BatchGetSymbols(tickers = companies$tickers,
                        first.date = "2010-01-01",
ast.date = "2017-12-31")
```

stocks is a list that contains 2 dataframes:

- df.control contains descriptive information like whether the download for the ticker is successful.
- df.tickers contains the downloaded price data. Each row is the price data for one ticker at one date, hence we need to process the data into a format easier to work with.



Retrieving SP500 Data

Loop over the good data: read one ticker at a time and merge into SP500.data

```
# Find out the good tickers / 465 stocks are "KEEP"
  for 2010 - 2017
good.tickers <- stocks$df.control$ticker [stocks$df.
  control$threshold.decision == "KEEP"]

for(i in 1:length(good.tickers)) {
X <- data.frame(date = stocks$df.tickers$ref.date[
  stocks$df.tickers$ticker == good.tickers[i]],
                  stocks$df.tickers$price.adjusted
                    [stocks$df.tickers$ticker ==
                    good.tickers[i]])
   colnames(X)[2] <- stocks$df.tickers$ticker[
     stocks$df.tickers$ticker == good.tickers[i]]

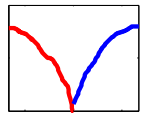
   SP500.data <- merge.data.frame(SP500.data, X, by =
       "date", all.x = TRUE)
}
```




Data Cleaning and Data Transforming

Convert downloaded daily data to monthly price data series into XTS series

```
1 library(quantmond)
2 temp <- xts(SP500.data$AMZN, order.by = as.POSIXct(
    SP500.data$date))
3 temp.monthly <- monthlyReturn(temp)
```




Data Cleaning and Data Transforming

Remove stocks with NAs in the series. We need to remove NAs for using the monthlyReturn() function. Most NAs are due to data not available on the starting date of the series, e.g. the company has not IPO yet. Here we face choices:

1. Remove all columns with NAs, then all remaining stocks could have the regression in the same period, i.e. with the same number of observations. (2010-2017)
2. Dynamically frame the data based on the available non-NA data points, but then some stocks in the regression analysis will have fewer observations. (1980-2015 every 5 years case)





Application of Fama-French three-factor model in SP500

FamaFrench in 2010-2017

```
```{r message=F, warning=F}
library(quantmod)

# Read SP500 daily data and convert date column to
  date format
SP500.data <- read.csv("Data/SP500_price.
  adjusted_2010-2017.csv")
SP500.data$date <- as.Date(SP500.data$date)

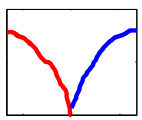
# Select 2010 - 2017 range
Stock.Prices.Daily <- SP500.data[SP500.data$date>="
  2010-01-01" &
    SP500.data$date<="2017-12-31",-1]
```



```

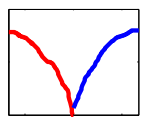
```
11 # Current FF3 till 201803, monthly
12 FF3 <- read.csv("Data/original/FF3.csv")
13 FF <- FF3[FF3$X >= 201001 & FF3$X <= 201712,]
14
15 # Convert series to XTS for using quantmod's
     monthlyReturn function
16 Stock.Prices.Daily <- xts(Stock.Prices.Daily[,-1],
17                          order.by = as.POSIXct(
                              Stock.Prices.Daily$date)
                              )
18
19 # Number of stocks to start with
20 ncol(Stock.Prices.Daily)
21
22 # Remove stocks with NAs in the series, otherwise
     monthly Return will not work properly
23 Stock.Prices.Daily <- Stock.Prices.Daily[, colSums(
     is.na(Stock.Prices.Daily)) == 0]
```

```
24 # Apply monthlyReturn function to each column (it
     seems it converts only one column at a time)
25 Stock.Prices.Monthly <- do.call(cbind, lapply(Stock.
     Prices.Daily, monthlyReturn))
26 # Stock.Prices.Monthly <- na.omit(Stock.Prices.
     Monthly)
27 colnames(Stock.Prices.Monthly) <- colnames(Stock.
     Prices.Daily)
28
29 # Number of stocks left
30 ncol(Stock.Prices.Monthly)
31
32 ```
```

````
1
2  ```{r}
3  Results <- list()
4  for(i in 1:ncol(Stock.Prices.Monthly)) {
5    RiRF <- Stock.Prices.Monthly[,i] - FF$RF
6    Regression <- lm(RiRF ~ FF$Mkt.RF + FF$SMB +
       FF$HML)
7    Results[[i]] <- summary(Regression)
8  }
9  ```
````

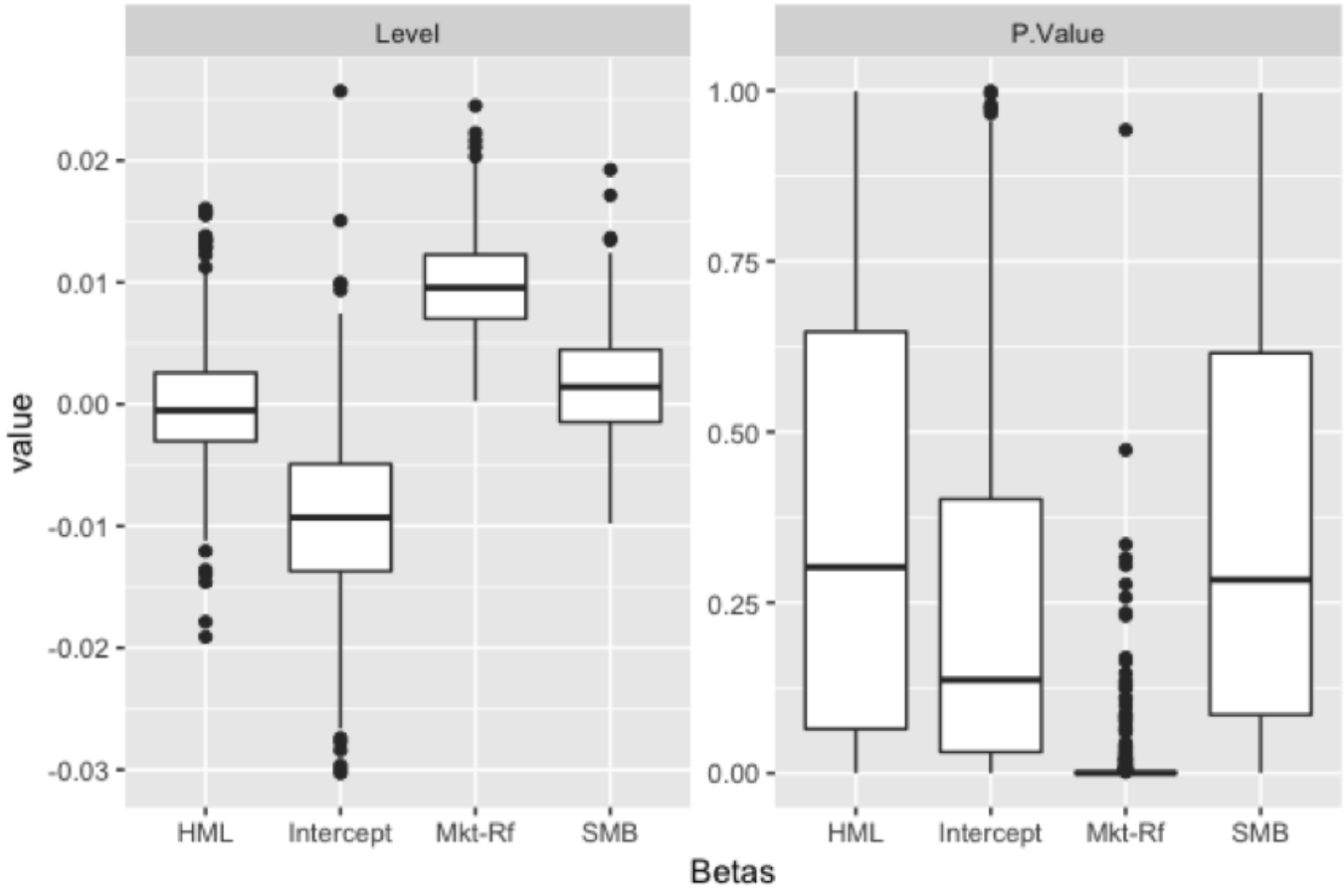


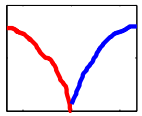
Figure 4: 3 Factor Model Betas

Goodness of Fit

```
1 ```{r}
2 # sort with R2 from largest to smallest, get top 10
3 kable(head(R2[order(R2$R.Squared, decreasing = T),],
     n=10), digits = 4)
4 ```
```

	Ticker	Name	Sector	R.Squared
388	TROW	T. Rowe Price Group	Financials	0.6806
227	IVZ	Invesco Ltd.	Financials	0.6777
31	AMG	Affiliated Managers Group Inc	Financials	0.6646
283	MS	Morgan Stanley	Financials	0.6334
334	PRU	Prudential Financial	Financials	0.6308
201	HON	Honeywell Int'l Inc.	Industrials	0.6299
270	MET	MetLife Inc.	Financials	0.6279
321	PFG	Principal Financial Group	Financials	0.6262
60	BEN	Franklin Resources	Financials	0.6225
232	JPM	JPMorgan Chase & Co.	Financials	0.6027

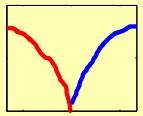
Figure 5: Tell me what to put here!



Trend analysis in 1980-2015

loop over above codes to download data from 1980 - 2015, group every 5 years.

```
1  library(lubridate)
2  List.of.start.date <- seq(as.Date("1980/1/1"), as.
     Date("2016/1/1"), "years")
3  List.of.start.date <- List.of.start.date[year(List.
     of.start.date)%%5==0]
4
5  Download.Stat <- data.frame(Data = List.of.start.
     date)
6  temp <- vector()
7
8  for(i in 1:(length(List.of.start.date)-1))
9  {
10   start.date <- as.Date(List.of.start.date[i])
11   end.date <- as.Date(List.of.start.date[i+1])-1
12
13  ?
```



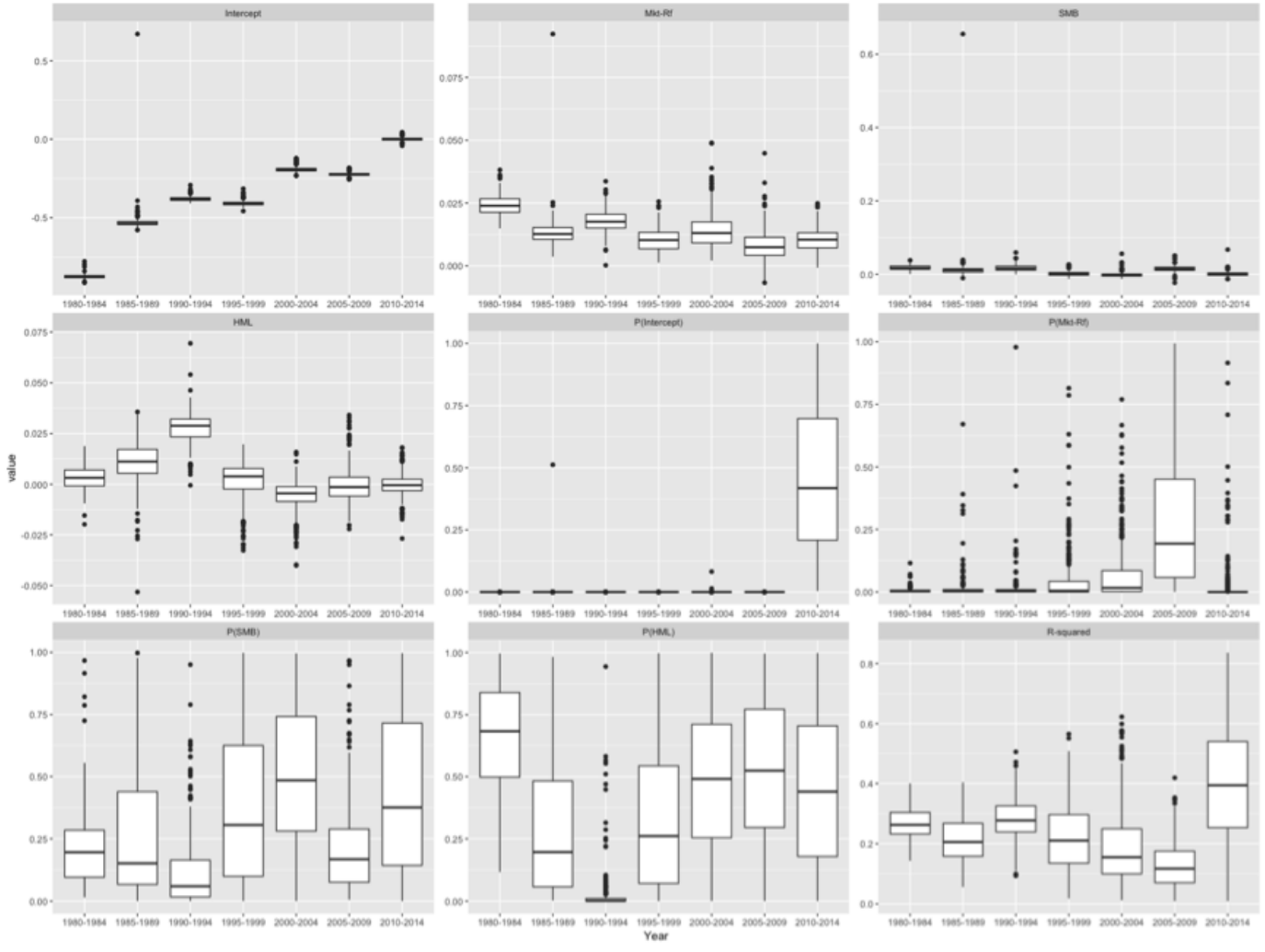
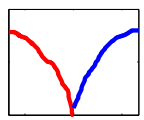


Figure 6: Trend analysis in 1980-2015

Time Period	Number of Stocks
1980-1984	170
1985-1989	229
1990-1994	271
1995-1999	345
2000-2004	394
2005-2009	432
2010-2014	459

Figure 7: Varying Samples: 168 Stocks idk what else



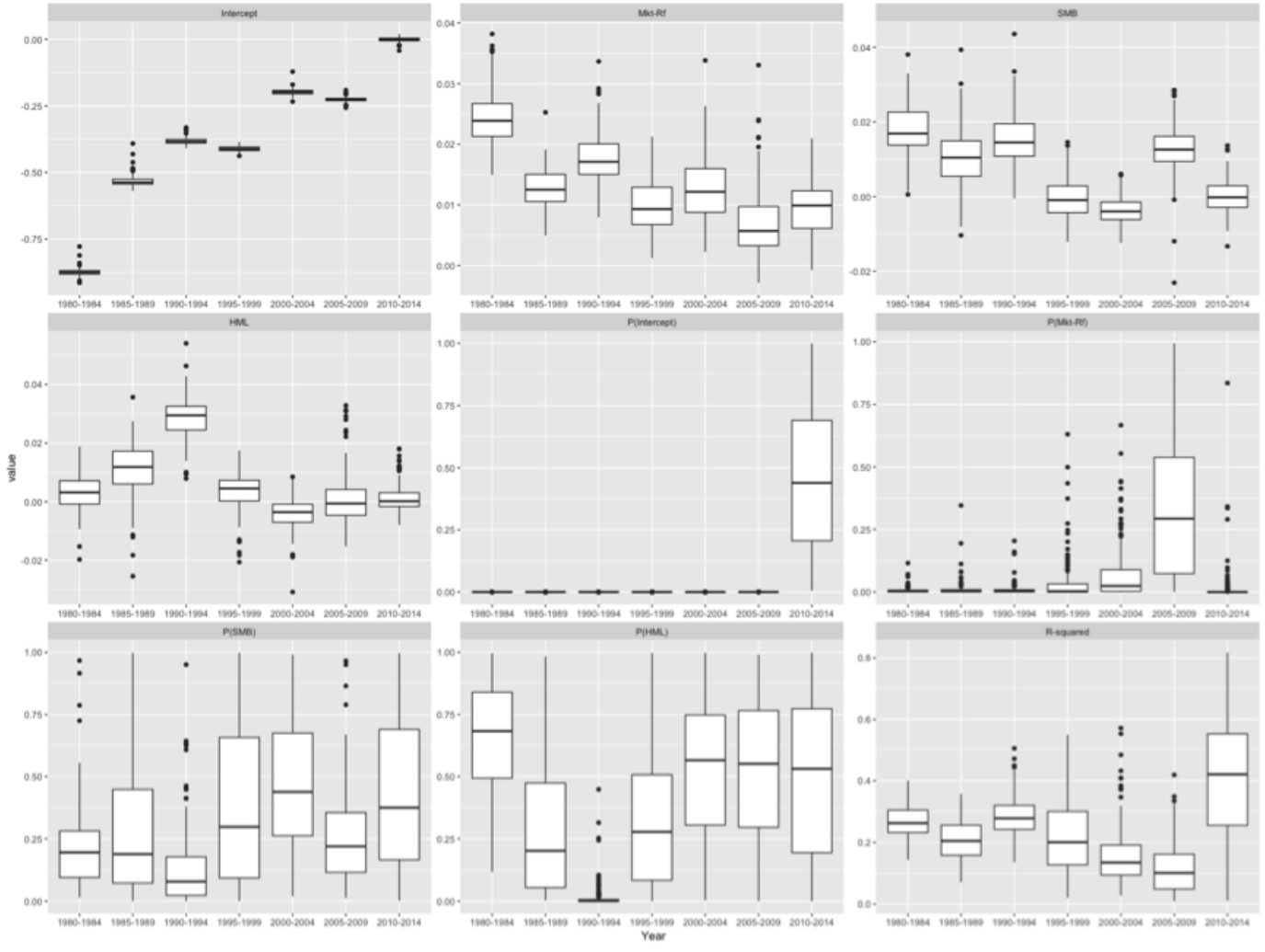
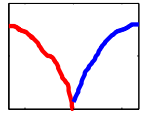


Figure 8: 168 Stocks idk what else



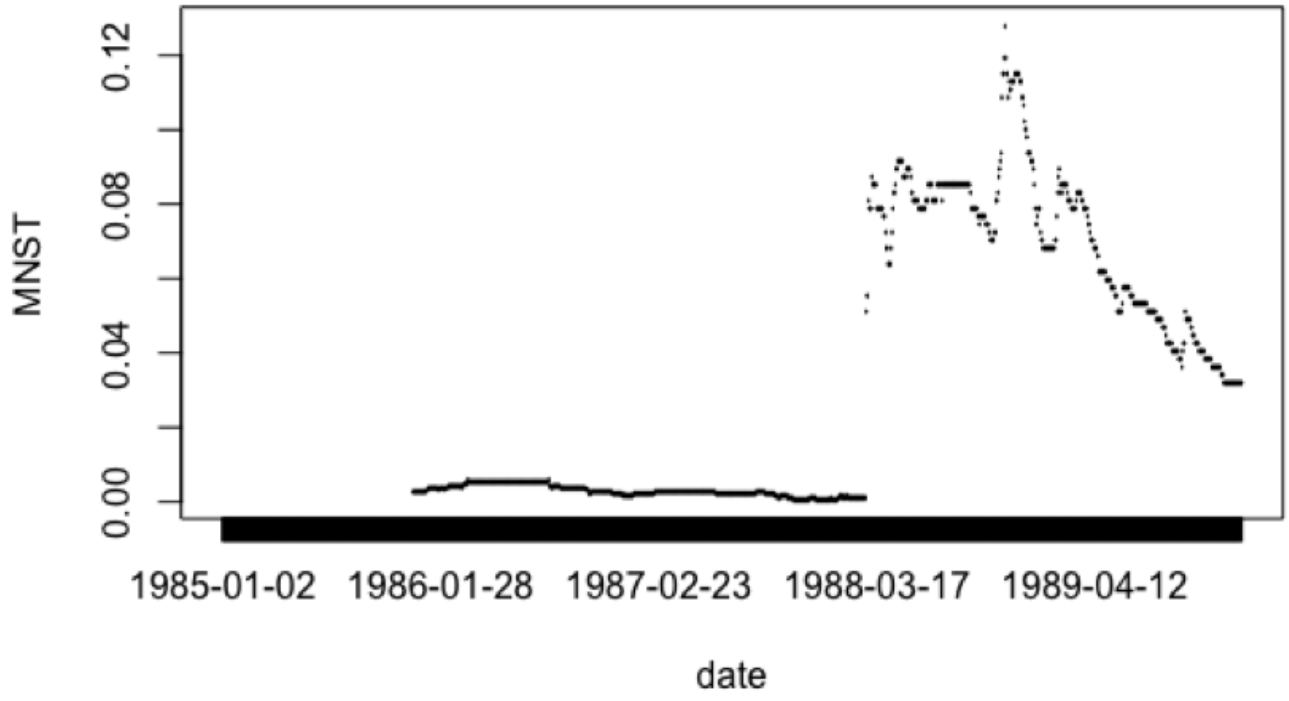
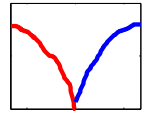


Figure 9: 168 Stocks idk what else



Correlation of The factors and Top 20 and Bottom 20 returns

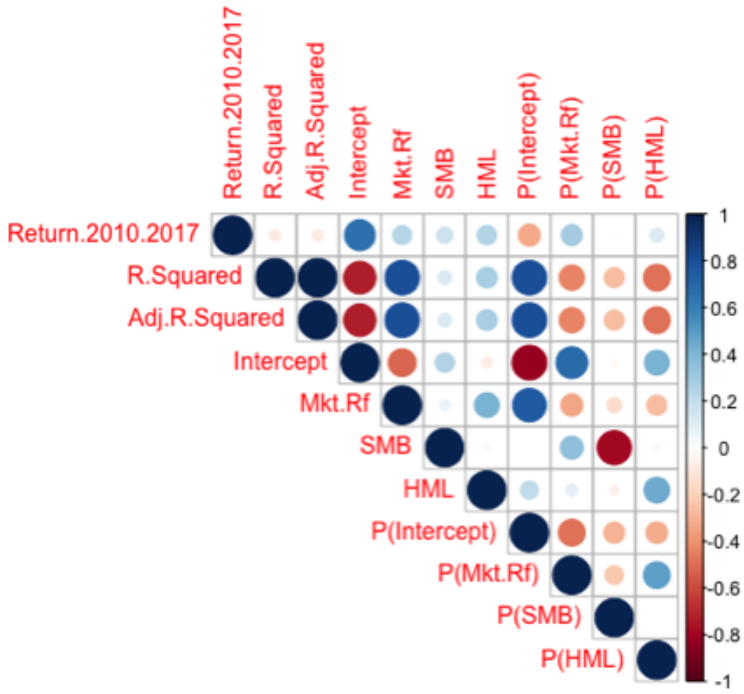


Figure 10: left side stuff

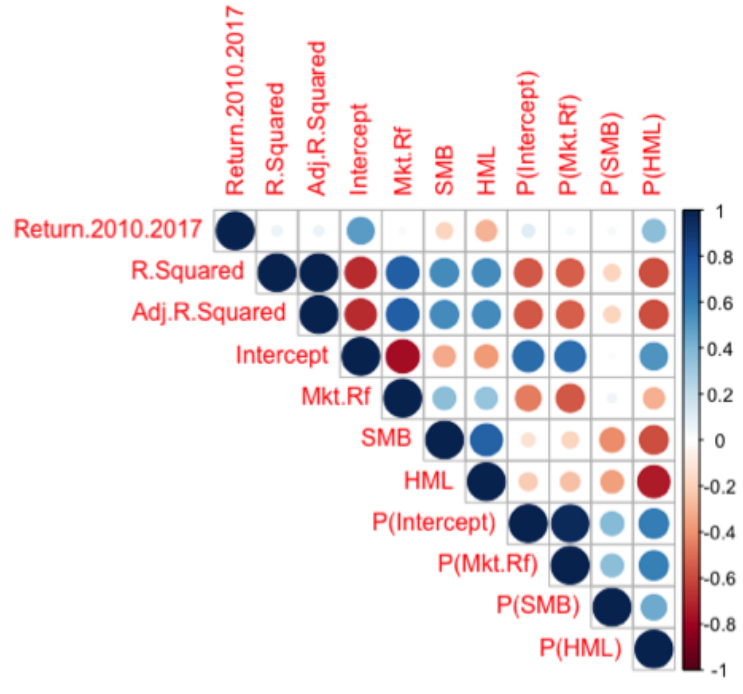


Figure 11: right side stuff