Crossectional asset pricing post CAPM-APT: Fama - French

November 24, 2021

Contents

1 The Fama French debate

1.1 Introduction

We now look at the *Beta is dead* debate.

Big hububb, claims of "CAPM is dead" flying everywhere

Can view this as a "case" in how empirical research in finance can have major repercussions, it can affect the whole finance profession

1.2 Background: Fama on efficient markets

Reason for why the big uproar connected to efficient markets debate, need to say something about that here, and Fama's role in the efficient markets literature.

Fama has been on the forefront of the efficient markets debate since the 70's.

Market efficiency: Asset Prices reflect all available information

His classical survey Fama (1970) is widely cited, its main conclusion is that markets are efficient. His book Fama (1976) has similar conclusions.

Main point made in that survey:

> We can only test whether information is *properly* reflected in prices in the context of a pricing model that defines the meaning of *properly*.

(Fama [1991], pg 1576.)

Because of its preeminence as a model of asset prices, the CAPM became the model that most had in mind when talking about a model that "properly" describes asset prices.

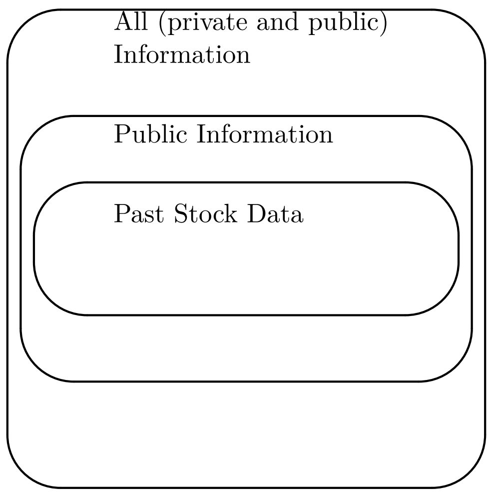
This has had some unfortunate consequences for the current debate.

Reminder of the setup in Fama (1970).

Introduced 3 categories of efficiency

1. Weak form tests
   - Can past prices be used to predict public information
2. Semi strong form tests
   - Do prices reflect public information
3. Strong form efficiency
   - Do prices reflect private info.

The efficiency definition depend on the information available: Think about the information sets



In terms of the original setup, CAPM used as a model of "expected returns" in semi–strong tests (event studies). Note that tests of the CAPM does not really fit into these categories.

2 Fama (1991), Efficient Markets, the sequel

In Fama (1991) he then came back to the efficient markets debate, concluding that markets were on the main efficient. This paper serves as a useful starting point. Let us summarize its main points.

The three categories above were changed, to better fit what had happened in the last 20 years.

1. Tests for return predictabiliy
   - time series predictability (will return to)

- cross-sectional predictability (tests of asset pricing models)

2. Event studies
3. Tests for private information
   - Insider trading
   - Security analysts
   - Portfolio managers

What does he say about return predictability in the context of an asset pricing model?
Summarize results

- Early evidence: Positive relation returns & beta
- Roll critique: Without the market portfolio, can not claim to have tested CAPM
- Anomalies: Other variables beside beta important in explaining returns.
  - Size (Banz (1981))
  - Seasonality
  - E/P ratios
  - Leverage

  See Hawawini and Keim (1995) for international summary.

- Bottom Line: (Asset pricing models) ((Fama, 1991, pg 1593))

  > The SLB model also passes the test of practical usefulness. Before it became a standard part of MBA investments courses, market professionals had only a vague understanding of risk and diversification. ...The SLB model gave a summary measure of risk, market β, interpreted as market sensitivity, that rang mental bells. Indeed, in spite of the evidence against the SLB model, market professionals (and academics) still think about risk in terms of market β.

  Fama [1991], pg 1593.

This was the state of the efficient markets (and the CAPM debate), according to Fama, in 1991.
However, at the same time Fama and French (1992) is lurking...

3 The Fama and French (1992) (FF) paper.

What does FF do in the paper?
Use the Fama and MacBeth (1973) empirical methods, on

- newer data
- portfolios also split according to other criteria (size, B/M)

Results, summary:

- β does a poor job in explaining cross–section of asset returns.
- Size and B/M does a much better job.

FF goes on in a series of papers (Fama and French (1993), Fama and French (1995), Fama and French (1996)) to make similar points.

4 Reactions to FF

The FF paper produced big headlines in newspapers, practical journals. Beta is dead!

Most of the finance profession smells a good fight, and rises to meet the challenge.

Some of the fronts at which attacks are made

- Theory is bad (Berk (1995))
- Data is bad (Kothari, Shanken, and Sloan (1995))
- Econometrics is bad (Kim (1995))
- Unconditional CAPM is dead, lets hear it for the conditional CAPM (Jagannathan and Wang (1996))
- Hooray for the NEW finance (Haugen (1995))

5 Size, the Berk (1995) paper

One of the most interesting (and simple) papers commenting on the size effect is Berk (1995). It is no less than a *theoretical* argument for why we actually observe both a size effect and a B/M effect.

It is unfortunately not cited enough among the major players in the debate.

Berk claim: *Expect* to see a positive relation between size as measured by market value, and return.

To see why, consider two firms with identical expected future cashflows $\tilde{c}$. If cashflows are perpetuities, current price is

$$p = \frac{E[\tilde{c}]}{r}$$

If the firm have different risk, the one with the higher risk will have lower price. . . .

The paper formalizes this notion by showing how any mis–estimation of β gives a positive relation between unexplained returns and firms market value.

Go over the thoretical arguments

- I firms
- $\tilde{c}_i$: end of period cashflow
- p_i: market value of firm i.
- $\tilde{r}_i = \log\left(\frac{\tilde{c}_i}{p_i}\right)$
- $C_i = E[\log \tilde{c}_i]$: "true" size.
- $R_i = E[\tilde{r}_i]$: Expected return.
- $L(C_i, R_i)$ cumulative distribution function.

  Assume independence between cash flows and returns: $L(C,R) = G(C)H(R)$.

Claim: $\log p_i$ will predict expected return.

Consider the regression

$$R_i = \alpha + \theta \log p_i + \epsilon_i$$

Recall the definition of R_i

$$R_i = E[\tilde{r}_i] = E\left[\log\left(\frac{\tilde{c}}{p_i}\right)\right] = E[\log c_i - \log p_i] = E[\log c_i] - \log p_i$$

giving

$$\log p_i = C_i - R_i$$

Recall what the coefficient is for a univariate regression

$$y = a + bx + \epsilon$$

$$\hat{b} = \frac{\text{cov}(x, y)}{\text{var}(x)}$$

(See for example Berck and Sydsæter (1995) 32.18)

$$\hat{\theta} = \frac{\text{cov}(R_i, \log p_i)}{\text{var}(\log p_i)} = \frac{\text{cov}(R_i, C_i - R_i)}{\text{var}(\log p_i)} = \frac{\text{cov}(R_i, C_i)}{\text{var}(\log p_i)} - \frac{\text{cov}(R_i, R_i)}{\text{var}(\log p_i)} = 0 - \frac{\text{var}(R_i)}{\text{var}(\log p_i)} < 0$$

Thus, in the regression

$$R_i = \alpha + \theta \log p_i + \epsilon_i$$

$\theta < 0$, there is an inverse relation between returns and size, as measured by market value.

Thus, market value is correlated with return because the market value is dicounted using the same return

5.1 Relation between market value and unexplained part of return

The previous explains why market value and returns are correlated.

Let us go a step further and look at the relation between the *unexplained* part of return and $\log p_i$

Suppose an *asset pricing model* explains returns. A similar analysis will show that the unexplained returns are also related to market value, even though the asset pricing completely explains returns.

Let $\hat{R}_i$ be the returns according to some asset pricing model. Assume $\text{cov}(\hat{R}, C) = 0$.

Run a regression of $\bar{R}_i$.

$$R_i = \omega + \beta \hat{R}_i + \epsilon_i$$

Take residuals, regress on $\log p_i$.

$$\epsilon_i = \eta + \gamma \log p_i + \xi_i$$

Consider the regression coefficient γ

$$\frac{\text{cov}(\epsilon_i, \log p_i)}{\text{var}(\log p_i)}$$

We can show this also to be negative if we can show $\text{cov}(\epsilon_i, \log p_i) < 0$.

$$\text{cov}(\log p_i, \epsilon_i) = \text{cov}(C_i - R_i, \epsilon_i) = \text{cov}(C_i, \epsilon_i) - \text{cov}(R_i, \epsilon_i)$$

$$= 0 - \text{cov}(R_i, R_i - (\omega + \beta \hat{R}_i)) = -\text{var}(R_i) + \text{cov}(R_i, \omega + \beta R_i) = -\text{var}(R_i) < 0$$

The intuituion carries over, there are theoretical reasons for why the pricing error ϵ_i is related to $\log p_i$.

In tests of the CAPM, what may happen, even if the CAPM actually *is* true

- Market portfolio incorrectly specified
- Beta measured with error.

Both these types of errors are likely, and thus the "size effect" may be partly explained by the results in the paper

5.2 Book to market

In addition to the size problem, Berk also points out similar reasons for why book to market is important for explaining returns.

Consider the variable

$$Q = E[\log \frac{c}{p}] = E[\log c - \log p] = C - \log p$$

(Since p is known.)

Claim: Q will predict returns perfectly.

To see why, consider the perpetuity case.

$$p = \frac{c}{r}$$

$$\log p = \log c - \log r$$

$$\log r = \log c - \log p$$

i.e.

$$Q = E[\log r] = E[\log c] - \log p = C - \log p$$

Thus, taking $\log \frac{\tilde{c}}{\tilde{p}}$ will perfectly predict returns.

Now, if book values are correlated with expected future cashflows (c), B/M will be a very good predictor of returns.

6 Data

Kothari, Shanken, Sloan: Another look

Reaciont to FF 92.

An attempt to ask: Are the results in FF a result of the particular way their data is slized?

In particular

- Why use monthly returns? No particular theoretical reason. Check using annual data: Do we get similar results? Claim to find a significant relation between beta and return.
- Selection biases. there may be a survivorship bias in the tests of FF.

  Generally, selection biases can occur when the selection of data is done on an *ex post* basis. For example: what if the tests are done only considering stocks that have survived the whole period? May exaggerate returns since all bankrupt firms are excluded. (But stocks may disappear for other reasons. If the result of a merger, may have opposite effect.)

  What is the potential bias in the FF paper? To find B/M, need for the firms to have been on the COMPUSTAT tapes for the whole period. Small firms that "do not make it" will not be backfilled by COMPUSTAT, and will not be on the tape.
- Artifact of test period (Post '62) B/M only from '63 onwards in FF. Look at earlier time periods too, results may be specific to time of measurement.

Overall, KSS points to the FF results influenced by data, but KSS not able to completely reverse FF.

See FF's response to KSS, JF.

7 Econometrics

Kim (1995) looks at how sensitive the results of FF are to the econometric errors–in–variables problem coming from the "rolling beta" estimation procedure.

Extention of Gibbons (1982) and Shanken (1992).

Hairy econometrics, showing how to correct for EIV problem under certain assumptions.

Simulation evidence, FF type estimation is biased.

After corrections, beta is significant. But firm size is important still.

8 Conditional CAPM

Jagannathan and Wang points to the fact that what FF test is the *unconditional* CAPM.

Present theorizing have pointed to a more reasonable hypothesis being a *conditional* version of the CAPM. The conditional CAPM nests the unconditional, but conditional CAPM may hold even if unconditional is false.

Setup

$$E[R_{it}|I_{t-1}] = \gamma_{0t-1} + \gamma_{1t-1}\beta_{it-1}$$

$$\beta_{it-1} = \frac{\text{cov}(R_{it}, R_{mt}|I_{t-1})}{\text{var}(R_{mt}|I_{t-1})}$$

Take expectations over I_{t-1}

$$E[E[R_{it}|I_{t-1}]] = E[\gamma_{0t-1}] + (E[\gamma_{1t-1}]E[\beta_{it-1}] + \text{cov}(\gamma_{1t-1}, \beta_{it-1}))$$

The expression in parenthesis on the right comes from the definition of covariance.

$$E[R_{it}] = \gamma_0 + \gamma_1\bar{\beta}_i + \text{cov}(\gamma_{1t-1}, \beta_{it-1})$$

where

$$\gamma_0 = E[\gamma_{0t-1}]$$

$$\gamma_1 = E[\gamma_{1t-1}]$$

$$\bar{\beta}_i = E[\beta_{1t-1}]$$

In an unconditional sense, the CAPM holds exactly if $\text{cov}(\gamma_{1t-1}, \beta_{it-1}) = 0$, which of course holds of β_{it} is constant.

Decompose further to get an estimable model.

Let

$$\nu_i = \frac{\text{cov}(\gamma_{1t-1}, \beta_{it-1})}{\text{var}(\gamma_{1t-1})}$$

$$\eta_{it-1} = \beta_{it-1} - \bar{\beta}_i - \eta_i(\gamma_{it-1} - \gamma_1)$$

γ_1 is a measure of sensitivity of conditional β to market risk premium.

Write above as

$$\beta_{it-1} = \bar{\beta}_i + \nu_i(\gamma_{it-1} - \gamma_1) + \eta_{it-1}$$

JW claims it can be shown that

$$E[\eta_{it-1}] = 0$$

$$E[\eta_{it-1}\gamma_{1ti-1}] = 0$$

Given that, perform substitution

$$
\begin{aligned}
E[R_{it}] &= \gamma_0 + \gamma_1 \bar{\beta}_i + \text{cov}(\gamma_{1t-1}, \beta_{it-1}) \\
&= \gamma_0 + \gamma_1 \bar{\beta}_i + \text{cov}(\gamma_{1t-1}, \bar{\beta}_i + \eta_i(\gamma_{1t-1} - \gamma_1) + \eta_{it-1}) \\
&= \gamma_0 + \gamma_1 \bar{\beta}_i + \eta_i \text{cov}(\gamma_{1t-1}, \gamma_{1t-1}) \\
&= \gamma_0 + \gamma_1 \bar{\beta}_i + \eta_i \text{var}(\gamma_{1t-1})
\end{aligned}
$$

Now we almost have an estimable unconditional relation. Problem: Have to estimate $\text{var}(\gamma_{1t-1})$ JW put conditions to make it possible to write

$$E[R_{it}] = \alpha_0 + \alpha_1 \beta_i + \alpha_2 \beta_i^{\gamma}$$

The rest is the typical work done to get an estimable relation.

Results: Strong evidence for a conditional CAPM.

9 Summarizing

The FF paper was a launching pad for a large number of reactions to the paper.

10 FF - current usage (2013)

The previous gives some summary of the early reactions to the FF research.

Note that little of it concerns methodological innovations. We are still in a setting with methods based on the original Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973) articles, but we are expanding on the number of "factors" considered in the pricing equation.

The best known such factors are the two Fama and French factors `SMB` and `HML`.

$$E[r_{it}] = \beta_i er_{mt} + b^{SMB} SMB_t + b^{HML} HML_t$$

The two factors `SMB` and `HML` were introduced in Fama and French (1996). For the construction they split data for the US stock market as shown in figure 1.

		Book/Market		
		L	H	M
Size	Small	S/L	S/M	S/H
	Big	B/L	B/M	B/H

Figure 1: The construction of the two Fama and French (1996) factors

The pricing factors are then constructed as:

$$\texttt{SMB} = \text{average}(S/L, S/M, S/H) - \text{average}(B/L, B/M, B/H)$$

$$\texttt{HML} = \text{average}(S/H, B/H) - \text{average}(S/L, B/L)$$

Note an important property of these factors: They are zero investment investment strategies, a long position minus a short position.

In estimation settings we need the two factors SMB and HML. These are typically downloaded from Ken French homepage when dealing with US data, or alternatively constructed from the crossection.

Momentum The Carhart factor `PR1YR`

Carhart (1997) introduced an additional factor that accounts for momentum. Figure 2 illustrates this factor construction. Each month the stock return is calculated over the previous eleven months. The returns are ranked, and split into three portfolios: The top 30%, the median 40% and the bottom 30%. The Carhart (1997) factor `PR1YR` is the difference between the average return of the top and the bottom portfolios. The ranking is recalculated every month.

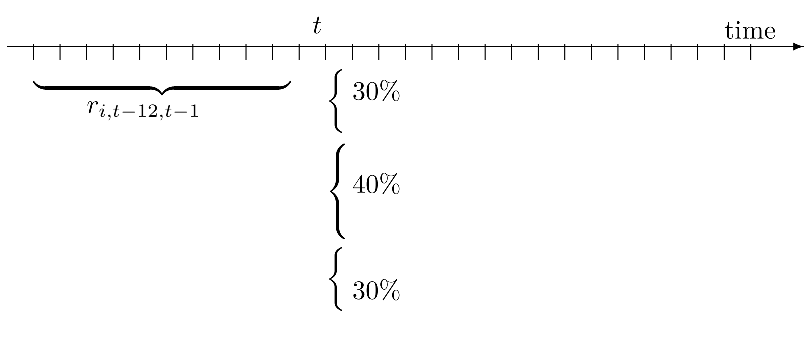


Figure 2: The construction of the Carhart (1997) factor PR1YR

An alternative momentum factor: `UMD` Ken French introduces an alternative momentum factor `UMD`, which he describes as follows:

> *....a momentum factor, constructed from six value-weight portfolios formed using independent sorts on size and prior return of NYSE, AMEX, and NASDAQ stocks. Mom is the average of the returns on two (big and small) high prior return portfolios minus the average of the returns on two low prior return portfolios. The portfolios are constructed monthly. Big means a firm is above the median market cap on the NYSE at the end of the previous month; small firms are below the median NYSE market cap. Prior return is measured from month -12 to - 2. Firms in the low prior return portfolio are below the 30th NYSE percentile. Those in the high portfolio are above the 70th NYSE percentile.* (from Ken French's web site)

Alternative factors Much of the current empirical asset pricing literature concerns the search for alternatives to the original Fama French two factors.

Some examples

- Macroeconomic factors
- Liquidity factors

Each time one suggests a "new" factor one does a similar construction to the FF constuction: Construct a zero investment portfolio of stocks sorted by the given criterion. Does this factor "price" the crossection of asset returns.

10.1 Illustrating the typical current usage

Consider the analysis of Ferreira, Keswani, Miguel, and Ramos (2013), a randomly chosen paper, in which they have a huge crossection of international mutual funds, and want to test for excess performance.

They run the four-factor model regression

$$R_{it} = \alpha_i + \beta_{0i} RM_t + \beta_{1i} SMB_t + \beta_{2i} HML_t + \beta_{3i} MOM_t + \varepsilon_{it}$$

> where R_{it} is the return in US dollars of fund i in excess of the 1 month US Treasury bill in month t, RM_t is the excess return in US dollars on the market, SMB_t (small minus big) is the average return on the small capitalization portfolio minus the average return on the large capitalization portfolio;....

So this formulation for investigating performance is by now standard in current research.

11 Fama and French add factors (again) (2014)

A recent step on the "lets add factors" road is Fama and French (2015). In this article Fama and French add two more variables, profitability and investment, to their three factor model.

The definitions of these variables are (taken from Fama and French (2015), table 8.)

> "In the sort for June of year t, B is book equity at the end of the fiscal year ending in year $t-1$ and M is market cap at the end of December of year $t-1$, adjusted for changes in shares outstanding between the measurement of B and the end of December. Operating profitability, OP, in the sort for June of year t is measured with accounting data for the fiscal year ending in year $t-1$ and is revenues minus cost of goods sold, minus selling, general, and administrative expenses, minus interest expense all divided by book equity. Investment, Inv, is the rate of growth of total assets from the fiscal year ending in year $t-2$ to the fiscal year ending in $t-1$."

12 Examples with US data

Exercise 1.

Collect from Ken French's homepage data on returns on ten industry portfolios (equally weighted) for the period 1926-2012.

1. Estimate the CAPM using the BJS method industry for industry. Do you reject that the constant coefficients are zero?

2. Estimate the three factor model (RMRF plus SMB and HML) using the BJS method industry for industry. Do these provide a better fit?

Solution to Exercise 1.

Reading data

```
# make sure that the first date do not change, this hardcodes the first date
library(zoo)
FF1 <- read.table("../data/F-F_Research_Data_Factors_monthly.txt",
                  header=TRUE,skip=3)
FF <- zooreg(FF1[2:5],start=c(1926,7),frequency=12)
RMRF <- FF$Mkt.RF
SMB  <- FF$SMB
HML  <- FF$HML
RF   <- FF$RF
FF10IndusEW <- read.table("../data/10_Industry_Portfolios_monthly_ew.txt",
                          header=TRUE,skip=10)
FF10IndusEW <- zooreg(FF10IndusEW,start=c(1926,7),frequency=12)
```

Running the CAPM

```
eRi <- FF10IndusEW-RF
eRm <- RMRF
data <- merge.zoo(eRi,eRm,all=FALSE)
eRi <- as.matrix(data[,1:10])
eRm <- as.matrix(data[,11])
summary(data)
```

Results

First, look at the data, always safest to look over the sum stats to see if we have the right data.

```
> summary(data)
     Index          NoDur              Durbl             Manuf
 Min.   :1926   Min.   :-28.5300   Min.   :-34.690   Min.   :-32.260
 1st Qu.:1948   1st Qu.: -2.1500   1st Qu.: -3.245   1st Qu.: -2.675
 Median :1970   Median :  0.9300   Median :  0.790   Median :  1.220
 Mean   :1970   Mean   :  0.8662   Mean   :  0.891   Mean   :  1.039
 3rd Qu.:1991   3rd Qu.:  3.7850   3rd Qu.:  4.860   3rd Qu.:  4.590
 Max.   :2013   Max.   : 57.4200   Max.   : 81.250   Max.   : 70.050
     Enrgy             HiTec             Telcm             Shops
 Min.   :-32.570   Min.   :-37.690   Min.   :-27.900   Min.   :-30.2500
 1st Qu.: -3.365   1st Qu.: -3.410   1st Qu.: -2.810   1st Qu.: -2.6050
 Median :  1.120   Median :  1.150   Median :  1.100   Median :  0.9200
 Mean   :  1.150   Mean   :  1.151   Mean   :  0.948   Mean   :  0.9173
 3rd Qu.:  5.180   3rd Qu.:  5.665   3rd Qu.:  4.775   3rd Qu.:  4.3050
 Max.   : 71.530   Max.   : 54.580   Max.   : 52.270   Max.   : 67.8000
      Hlth             Utils              Other               eRm
 Min.   :-33.320   Min.   :-32.0500   Min.   :-30.8000   Min.   :-28.980
```

```
 1st Qu.: -2.780   1st Qu.: -1.6950   1st Qu.: -2.3650   1st Qu.: -2.105
 Median :  1.050   Median :  0.8000   Median :  1.0600   Median :  1.010
 Mean   :  1.077   Mean   :  0.8437   Mean   :  0.9874   Mean   :  0.628
 3rd Qu.:  4.985   3rd Qu.:  3.1350   3rd Qu.:  4.1500   3rd Qu.:  3.655
 Max.   : 42.560   Max.   : 65.5700   Max.   : 77.2400   Max.   : 37.770
```

Note that everything is in percentages, which is obvious from the scale of the data.
Result for industry NoDur (Non Durables)

```
> summary(reg)
Response NoDur :

Call:
lm(formula = NoDur ~ eRm)

Residuals:
     Min       1Q   Median       3Q      Max
-11.6174  -1.8605  -0.2133   1.5024  23.6169

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.19884    0.10209   1.948   0.0517 .
eRm          1.06264    0.01868  56.879   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.263 on 1033 degrees of freedom
Multiple R-squared: 0.758,Adjusted R-squared: 0.7577
F-statistic:  3235 on 1 and 1033 DF,  p-value: < 2.2e-16
```

Summarizing all in one table

	Estimate	Std. Error	t value	Pr(>|t|)	
NoDur	(Intercept)	0.1988	0.1021	1.95	0.0517
	eRm	1.0626	0.0187	56.88	0.0000
Durbl	(Intercept)	-0.0090	0.1427	-0.06	0.9495
	eRm	1.4332	0.0261	54.89	0.0000
Manuf	(Intercept)	0.2169	0.1056	2.05	0.0402
	eRm	1.3092	0.0193	67.75	0.0000
Enrgy	(Intercept)	0.4104	0.1747	2.35	0.0190
	eRm	1.1782	0.0320	36.86	0.0000
HiTec	(Intercept)	0.2548	0.1424	1.79	0.0738
	eRm	1.4267	0.0261	54.77	0.0000
Telcm	(Intercept)	0.2910	0.1351	2.15	0.0315
	eRm	1.0463	0.0247	42.31	0.0000
Shops	(Intercept)	0.1905	0.1231	1.55	0.1221
	eRm	1.1572	0.0225	51.35	0.0000
Hlth	(Intercept)	0.4135	0.1305	3.17	0.0016
	eRm	1.0566	0.0239	44.25	0.0000
Utils	(Intercept)	0.2607	0.1355	1.92	0.0546
	eRm	0.9284	0.0248	37.44	0.0000
Other	(Intercept)	0.2128	0.1315	1.62	0.1058
	eRm	1.2333	0.0241	51.26	0.0000

Here we reject that several of the constants are zero.
Let us look at the same portfolios, but now adding the two Fama French factors SMB and HML.

```
> source("read_industries.R")
> eRi <- FF10IndusEW-RF
```

```
> eRm <- RMRF
> data <- merge.zoo(eRi,eRm,SMB,HML, all=FALSE)
> summary(data)
     Index          NoDur              Durbl             Manuf        
 Min.   :1926   Min.   :-28.5300   Min.   :-34.690   Min.   :-32.260  
 1st Qu.:1948   1st Qu.: -2.1500   1st Qu.: -3.245   1st Qu.: -2.675  
 Median :1970   Median :  0.9300   Median :  0.790   Median :  1.220  
 Mean   :1970   Mean   :  0.8662   Mean   :  0.891   Mean   :  1.039  
 3rd Qu.:1991   3rd Qu.:  3.7850   3rd Qu.:  4.860   3rd Qu.:  4.590  
 Max.   :2013   Max.   : 57.4200   Max.   : 81.250   Max.   : 70.050  
     Enrgy            HiTec             Telcm             Shops         
 Min.   :-32.570   Min.   :-37.690   Min.   :-27.900   Min.   :-30.2500  
 1st Qu.: -3.365   1st Qu.: -3.410   1st Qu.: -2.810   1st Qu.: -2.6050  
 Median :  1.120   Median :  1.150   Median :  1.100   Median :  0.9200  
 Mean   :  1.150   Mean   :  1.151   Mean   :  0.948   Mean   :  0.9173  
 3rd Qu.:  5.180   3rd Qu.:  5.665   3rd Qu.:  4.775   3rd Qu.:  4.3050  
 Max.   : 71.530   Max.   : 54.580   Max.   : 52.270   Max.   : 67.8000  
      Hlth             Utils               Other               eRm         
 Min.   :-33.320   Min.   :-32.0500   Min.   :-30.8000   Min.   :-28.980  
 1st Qu.: -2.780   1st Qu.: -1.6950   1st Qu.: -2.3650   1st Qu.: -2.105  
 Median :  1.050   Median :  0.8000   Median :  1.0600   Median :  1.010  
 Mean   :  1.077   Mean   :  0.8437   Mean   :  0.9874   Mean   :  0.628  
 3rd Qu.:  4.985   3rd Qu.:  3.1350   3rd Qu.:  4.1500   3rd Qu.:  3.655  
 Max.   : 42.560   Max.   : 65.5700   Max.   : 77.2400   Max.   : 37.770  
      SMB                HML         
 Min.   :-16.3900   Min.   :-13.450  
 1st Qu.: -1.5200   1st Qu.: -1.295  
 Median :  0.0500   Median :  0.220  
 Mean   :  0.2352   Mean   :  0.382  
 3rd Qu.:  1.7750   3rd Qu.:  1.745  
 Max.   : 39.0400   Max.   : 35.480  
```

Look at the first industry, nondurables.

```
lm(formula = NoDur ~ eRm + SMB + HML)

Residuals:
     Min       1Q   Median       3Q      Max 
-10.4511  -1.1701  -0.0814   0.9188  13.0919 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.01752    0.06219   0.282    0.778    
eRm          0.86677    0.01226  70.718   <2e-16 ***
SMB          0.72316    0.01974  36.640   <2e-16 ***
HML          0.35155    0.01779  19.756   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.98 on 1031 degrees of freedom
Multiple R-squared: 0.9111,Adjusted R-squared: 0.9108 
F-statistic:  3521 on 3 and 1031 DF,  p-value: < 2.2e-16
```

Summarizing the results

		Estimate	Std. Error	t value	Pr(>|t|)
NoDur	(Intercept)	0.0175	0.0622	0.28	0.7783
	eRm	0.8668	0.0123	70.72	0.0000
	SMB	0.7232	0.0197	36.64	0.0000
	HML	0.3515	0.0178	19.76	0.0000
Durbl	(Intercept)	-0.2364	0.0921	-2.57	0.0104
	eRm	1.1718	0.0181	64.58	0.0000
	SMB	1.0161	0.0292	34.78	0.0000
	HML	0.3995	0.0263	15.17	0.0000
Manuf	(Intercept)	0.0156	0.0588	0.27	0.7907
	eRm	1.0967	0.0116	94.70	0.0000
	SMB	0.7685	0.0186	41.21	0.0000
	HML	0.4034	0.0168	23.99	0.0000
Enrgy	(Intercept)	0.2007	0.1558	1.29	0.1980
	eRm	0.9868	0.0307	32.13	0.0000
	SMB	0.5950	0.0495	12.03	0.0000
	HML	0.4973	0.0446	11.15	0.0000
HiTec	(Intercept)	0.1975	0.1040	1.90	0.0578
	eRm	1.2547	0.0205	61.23	0.0000
	SMB	0.9875	0.0330	29.92	0.0000
	HML	-0.1750	0.0298	-5.88	0.0000
Telcm	(Intercept)	0.2949	0.1234	2.39	0.0170
	eRm	0.9708	0.0243	39.93	0.0000
	SMB	0.5335	0.0392	13.63	0.0000
	HML	-0.2146	0.0353	-6.08	0.0000
Shops	(Intercept)	0.0362	0.0838	0.43	0.6656
	eRm	0.9498	0.0165	57.50	0.0000
	SMB	0.8963	0.0266	33.70	0.0000
	HML	0.1932	0.0240	8.05	0.0000
Hlth	(Intercept)	0.4031	0.1068	3.77	0.0002
	eRm	0.9449	0.0211	44.87	0.0000
	SMB	0.7343	0.0339	21.65	0.0000
	HML	-0.2411	0.0306	-7.89	0.0000
Utils	(Intercept)	0.0968	0.1234	0.78	0.4330
	eRm	0.8238	0.0243	33.88	0.0000
	SMB	0.1542	0.0392	3.94	0.0001
	HML	0.5060	0.0353	14.33	0.0000
Other	(Intercept)	-0.0784	0.0690	-1.14	0.2565
	eRm	0.9658	0.0136	70.99	0.0000
	SMB	0.8372	0.0219	38.21	0.0000
	HML	0.6868	0.0198	34.77	0.0000

13 Examples with Norwegian data

Exercise 2.

Running the Black et al. (1972) regression

$$er_{it} = \alpha_i + \beta_i er_{mt} + e_{it}$$

on a set of 10 size-based portfolios on the Oslo Stock Exchange, we find that we on an equation by equation basis reject the null hypothesis that $\alpha_i = 0$ for many of the portfolios. An alternative model is the Fama French model

$$E[r_i] - r_f = E[r_m - r_f]\beta_i + b_i^{smb} SMB + b_i^{hml} HML_t$$

where SMB and HML are "zero investment" portfolios designed to represent size and book-to-market "factors." Using domestic versions of the Fama French factors, consider the regression

$$er_{it} = \alpha_i + \beta_i er_{mt} + +b_i^{smb} SMB_t + b_i^{hml} HML_t e_{it}$$

Run these regressions on 10 (ew) size sorted portfolios at the OSE. Test $\alpha_i = 0$ on a portfolio by portfolio basis. Use an equally weighted market index, and returns data 1980-2012.

Solution to Exercise 2.

Reading the data and running the regressions

```
library(zoo)
library(xtable)
Rets <- read.zoo("../../data/equity_size_portfolios_monthly_ew.txt",
                   header=TRUE,sep=";",format="%Y%m%d")
Rf <- read.zoo("../../data/NIBOR_monthly.txt",
               format="%Y%m%d",header=TRUE,sep=";")
eR  <- Rets - lag(Rf,-1)
Rm <- read.zoo("../../data/market_portfolios_monthly.txt",
               format="%Y%m%d",header=TRUE,sep=";")
eRmew <- Rm$EW - lag(Rf,-1)
FF <- read.zoo("../../data/pricing_factors_monthly.txt",
               header=TRUE,sep=";",format="%Y%m%d")
data  <- merge(eR,eRmew,FF$SMB,FF$HML,all=FALSE)
er  <- as.matrix(data[,1:10])
erm <-as.matrix(data[,11])
SMB <- as.matrix(data[,12])
HML <- as.matrix(data[,13])
reg=lm(er~erm + SMB + HML )
```

Let us now look at the results

For the first portfolio (smallest stocks)

```
Response X1..small.size. :

Residuals:
     Min       1Q   Median       3Q      Max
-0.14418 -0.03050 -0.00379  0.02299  0.33764

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.009332   0.002781   3.356 0.000872 ***
erm         0.762926   0.047877  15.935  < 2e-16 ***
SMB         0.244951   0.058359   4.197 3.38e-05 ***
HML         0.147748   0.051301   2.880 0.004206 **
```

```
Residual standard error: 0.0519 on 374 degrees of freedom
  (6 observations deleted due to missingness)
Multiple R-squared: 0.4203,Adjusted R-squared: 0.4156
F-statistic: 90.38 on 3 and 374 DF,  p-value: < 2.2e-16
```

And for portfolio 2

```
Response X2 :

Residuals:
      Min        1Q    Median        3Q       Max
-0.180823 -0.021229 -0.002246  0.021534  0.155906

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.002393   0.002289   1.046   0.2964
erm         0.928658   0.039405  23.567  < 2e-16 ***
SMB         0.363066   0.048032   7.559 3.17e-13 ***
HML         0.077116   0.042223   1.826   0.0686 .

Residual standard error: 0.04272 on 374 degrees of freedom
  (6 observations deleted due to missingness)
Multiple R-squared: 0.6086,Adjusted R-squared: 0.6054
F-statistic: 193.8 on 3 and 374 DF,  p-value: < 2.2e-16
```

Summarizing the results in a table:

		Estimate	Std. Error	t value	Pr(>|t|)
1(small)	(Intercept)	0.0093	0.0028	3.36	0.0009
	erm	0.7629	0.0479	15.94	0.0000
	SMB	0.2450	0.0584	4.20	0.0000
	HML	0.1477	0.0513	2.88	0.0042
2	(Intercept)	0.0024	0.0023	1.05	0.2964
	erm	0.9287	0.0394	23.57	0.0000
	SMB	0.3631	0.0480	7.56	0.0000
	HML	0.0771	0.0422	1.83	0.0686
3	(Intercept)	-0.0018	0.0020	-0.92	0.3595
	erm	0.9847	0.0338	29.10	0.0000
	SMB	0.2370	0.0412	5.75	0.0000
	HML	-0.0211	0.0363	-0.58	0.5608
4	(Intercept)	-0.0040	0.0019	-2.05	0.0413
	erm	1.0636	0.0335	31.75	0.0000
	SMB	0.3090	0.0408	7.57	0.0000
	HML	-0.0430	0.0359	-1.20	0.2321
5	(Intercept)	0.0012	0.0018	0.68	0.4939
	erm	0.9733	0.0312	31.15	0.0000
	SMB	0.2034	0.0381	5.34	0.0000
	HML	-0.0784	0.0335	-2.34	0.0198
6	(Intercept)	-0.0004	0.0019	-0.21	0.8330
	erm	0.9516	0.0322	29.51	0.0000
	SMB	-0.0539	0.0393	-1.37	0.1714
	HML	0.0544	0.0346	1.57	0.1165
7	(Intercept)	-0.0024	0.0019	-1.25	0.2103
	erm	1.0884	0.0326	33.40	0.0000
	SMB	-0.1652	0.0397	-4.16	0.0000
	HML	0.0748	0.0349	2.14	0.0329
8	(Intercept)	-0.0010	0.0019	-0.55	0.5796
	erm	1.0466	0.0324	32.33	0.0000
	SMB	-0.2380	0.0395	-6.03	0.0000
	HML	-0.0079	0.0347	-0.23	0.8207
9	(Intercept)	-0.0041	0.0019	-2.16	0.0313
	erm	1.2009	0.0323	37.21	0.0000
	SMB	-0.3073	0.0393	-7.81	0.0000
	HML	-0.0144	0.0346	-0.42	0.6768
10(large)	(Intercept)	0.0026	0.0025	1.08	0.2805
	erm	0.9585	0.0422	22.71	0.0000
	SMB	-0.6657	0.0514	-12.94	0.0000
	HML	-0.1656	0.0452	-3.66	0.0003

There are much less rejections of $\alpha_i = 0$, and in particular SMB is significant in most of the regressions.

References

Rolf W Banz. The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9:3–18, 1981.

Per Berck and Knut Sydsæter. *Matematisk Formelsamling for økonomer*. Universitetsforlaget, 2 edition, 1995.

Jonathan Berk. A critique of size related anomalies. *Review of Financial Studies*, 8:275–86, 1995.

Fisher Black, Michael Jensen, and Myron Scholes. The capital asset pricing model, some empirical tests. In Michael C Jensen, editor, *Studies in the theory of capital markets*. Preager, 1972.

Mark M Carhart. On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82, March 1997.

Eugene F Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25:383–417, 1970.

Eugene F Fama. *Foundations of Finance*. Basic Books, 1976.

Eugene F Fama. Efficient capital markets: II. *Journal of Finance*, 46(5):1575–1618, December 1991.

Eugene F Fama and Kenneth R French. The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–466, June 1992.

Eugene F Fama and Kenneth R French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33:3–56, 1993.

Eugene F Fama and Kenneth R French. Size and book-to-market factors in earnings and returns. *Journal of Finance*, 50(1): 131–56, March 1995.

Eugene F Fama and Kenneth R French. Multifactor explanations of asset pricing anomialies. *Journal of Finance*, 51(1):55–85, 1996.

Eugene F. Fama and Kenneth R. French. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1 – 22, 2015.

Eugene F Fama and J MacBeth. Risk, return and equilibrium, empirical tests. *Journal of Political Economy*, 81:607–636, 1973.

Miguel A. Ferreira, Aneel Keswani, António F. Miguel, and Sofia B. Ramos. The determinants of mutual fund performance: A cross-country study. *Review of Finance*, 17(2):483–525, 2013. doi: 10.1093/rof/rfs013. URL `http://rof.oxfordjournals.org/content/17/2/483.abstract`.

Michael R Gibbons. Multivariate tests of financial models, a new approach. *Journal of Financial Economics*, 10:3–27, March 1982.

Robert A Haugen. *The New Finance. The Case Against Efficient Markets*. Prentice–Hall, 1995.

Gabriel Hawawini and Donald B Keim. On the predictability of common stock returns: World wide evidence. In R A Jarrow, V Maksimovic, and W T Ziemba, editors, *Finance*, volume 9 of *Handbooks in Operations Research and Management Science*, chapter 17, pages 497–54. North–Holland, 1995.

Ravi Jagannathan and Zhenyu Wang. The conditional capm and the cross–section of expected returns. *Journal of Finance*, 51 (1):3–54, 1996.

Dongcheol Kim. The errors in variables problem in the cross-section of expected returns. *Journal of Finance*, 50(5):1605–34, December 1995.

S P Kothari, Jay Shanken, and Richard G Sloan. Another look at the cross-section of expected returns. *Journal of Finance*, 50(1):185–224, March 1995.

Jay Shanken. On the estimation of beta-pricing models. *Review of Financial Studies*, 5(1):1–34, 1992.