2

Factor Models

## Data selection

library(quantmod)

stocks <- stockSymbols()

str(stocks)

stocks[1:5, c(1, 3:4, ncol(stocks))]

library(Quandl)

LIBOR <- Quandl('FED/RILSPDEPM01\_N\_B',

start\_date = '2010-06-01', end\_date = '2014-06-01')

d <- read.table("data.csv", header = TRUE, sep = ";")

d[1:7, c(1:5, (ncol(d) - 6):ncol(d))]

d <- d[, colSums(is.na(d)) == 0]

d <- d[, c(T, colMins(d[, 2:ncol(d)]) > 0)]

## Estimation of APT with principal component analysis

p <- d[, 3:ncol(d)]

r <- log(p[2:nrow(p), ] / p[1:(nrow(p) - 1), ])

r <- r[, runif(nrow(r)) < 0.1]

pca <- princomp(r)

plot(pca$sdev)

factanal(r, 5)

## Estimation of the Fama-French model

d2 <- d[, 2:ncol(d)]

d2 <- log(tail(d1, -1)/head(d1, -1))

d <- cbind(d[2:nrow(d), 1], d2)

d <- merge(LIBOR, d, by = 1)

print(d[1:5, 1:5])]

d$LIBOR <- d$LIBOR / 36000

d[1:5, c(1,(ncol(d) - 3):ncol(d))]

> stocks = stocks[stocks$Symbol %in% colnames(d),]

stocks$BookToMarketRatio <-

stocks$BookValuePerShare / stocks$LastSale

str(stocks)

avg\_size <- mean(stocks$MarketCap)

BIG <- as.character(stocks$Symbol[stocks$MarketCap > avg\_size])

SMALL <- as.character(stocks[stocks$MarketCap < avg\_size,1])

d$SMB <- rowMeans(d[,colnames(d) %in% SMALL]) –

rowMeans(d[,colnames(d) %in% BIG])

avg\_btm <- mean(stocks$BookToMarketRatio)

HIGH <- as.character(

stocks[stocks$BookToMarketRatio > avg\_btm, 1])

LOW <- as.character(

stocks[stocks$BookToMarketRatio < avg\_btm, 1])

d$HML <- rowMeans(d[, colnames(d) %in% HIGH]) –

rowMeans(d[, colnames(d) %in% LOW])

d$Market <- d$SP500 - d$LIBOR

d$C <- d$C - d$LIBOR

model <- glm( formula = "C ~ Market + SMB + HML" , data = d)

estimation <- model$coefficients[1]+

model$coefficients[2] \* d$Market +

model$coefficients[3]\*d$SMB +

model$coefficients[4]\*d$HML

plot(estimation, d$C, xlab = "estimated risk-premium",

ylab = "observed riks premium",

main = "Fama-French model for Citigroup")

lines(c(-1, 1), c(-1, 1), col = "red")

d$EXEL <- d$EXEL – d$LIBOR

model2 <- glm( formula = "EXEL~Market+SMB+HML" , data = d)

summary(model2)

estimation2 <- model2$coefficients[1] +

model2$coefficients[2] \* d$Market +

model2$coefficients[3] \* d$SMB + model2$coefficients[4] \* d$HML

plot(estimation2, d$EXEL, xlab = "estimated risk-premium",

ylab = "observed riks premium",

main = "Fama-French model for EXEL")

lines(c(-1, 1), c(-1, 1), col = "red")