MGMT 237E:
Empirical Methods in Finance

Homework 6

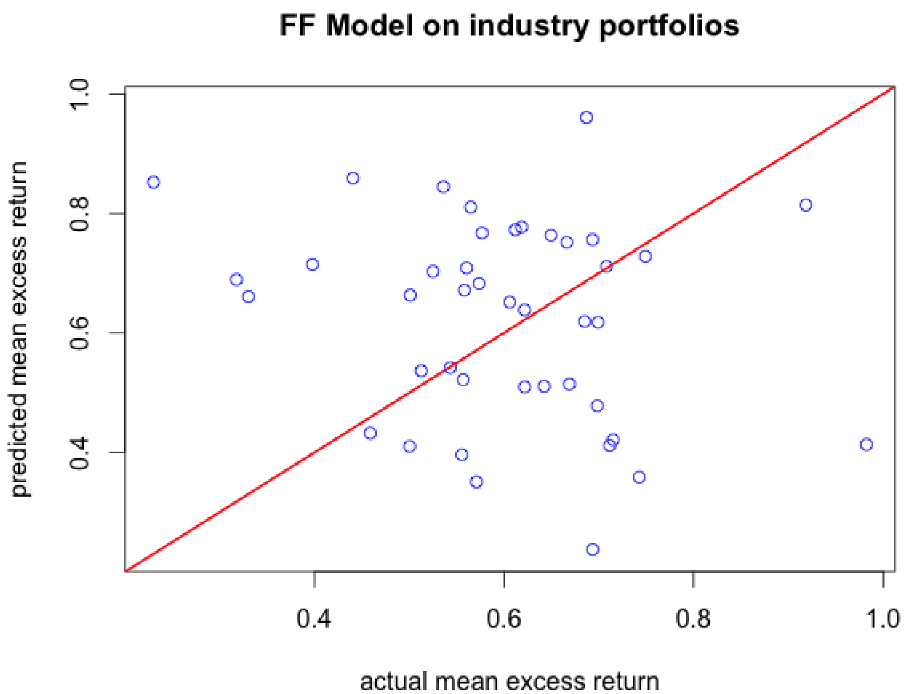
Yi-Chi Chan Zhaofang Shi
Ahswin Kumar Ashok Kumar George Bonebright

March 8, 2016

1.a. We can use time series regression on excess returns to test the Fama-French model on the industry portfolios. To do this, we will run a time series regression to estimate the values of β_i.

We use the formula: $R^e_{it} = \alpha_i + \beta^m_i R^{e,m}_t + \beta^{smb}_i R^{smb}_t + \beta^{hml}_i R^{hml}_t + \epsilon_{it}$ as our regression model. There is no need to estimate risk prices because the factors are traded assests, therefore: $\hat{\lambda} = (\lambda^m, \lambda^{smb}, \lambda^{hml})$ where $\hat{\lambda}^j = \frac{1}{T}\sum_{t=1}^{T} \mathbf{R}^j_t for j = m, smb, hml$. The coefficients $(\beta^m_i, \beta^{smb}_i, \beta^{hml}_i)$ measure how the asset's return R^e_{it} covaries with the risk factors. The regression intercepts are the pricing errors and the should be zero under the null hypothesis.

We plot the predicted mean excess returns $\hat{\beta}'_i\hat{\lambda}$ against the realized mean excess returns $\bar{R}^e_{it} = \frac{1}{T}\sum_{t=1}^{T} R^e_{it}$.



1.b. Model testing H_0:pricing errors (alphas) are jointly zero.
We can test the model: $E[R^e_{it}] = \beta_i^{'} E[f_t]$ by running time series regressions: $R^e_{it} = \alpha_i + \beta_i^{'} f_t = \epsilon_{it}$, $t = 1, ..., T$ with iid errors, homscedasticity, and independence of the factors, the test statistic for the pricing errors is given by: $(T - K - N)/N[1 + \bar{f}^{'} \hat{\sum}_f^{-1} \bar{f}^{-1}]\hat{\alpha}^{'} \hat{\sum}_\epsilon^{-1} \hat{\alpha} \sim F_{N,T-N-K}$ where $\hat{\sum}_\epsilon$ denotes the covariance matrix of ϵ, $\hat{\sum}_f$ denotes the covariance matrix of the factors f_t, $\bar{f}$ is the average factor, and $\hat{\alpha}$ are the OLS estimates of α.
In this case, the F-statistic is 0.4184848 which is significant at the 99% confidence level for $43, 626$ degrees of freedom. We reject the null that the pricing errors are jointly zero.

1.d. The variation of industry portfolio returns are more dispersed than the variation in the 25 portfolios sorted by B/M ratios and size. It would be harder to explain the increased variance.

```
# MGMT237E HW6
# Question1

library(lubridate)
library(xts)
library(dplyr)

##Import Data
#Read 48 Industry Data/3 Factor Model Data
ind48<-read.csv("48_Industry_Portfolios.CSV",header=T)
fac3<-read.csv("F-F_Research_Data_Factors.CSV",header=T)
#Rename Date Column
colnames(ind48)[1]<-"Date"
colnames(fac3)[1]<-"Date"
#Reformat Date Column
ind48$Date<-parse_date_time(ind48$Date,"%y%m")
fac3$Date<-parse_date_time(fac3$Date,"%y%m")
ind48$Date<-as.Date(ind48$Date)
fac3$Date<-as.Date(fac3$Date)
#Select data from 1960 to 2015
ind48<-ind48[ind48$Date>=as.Date("1960-01-01"),]
fac3<-fac3[fac3$Date>=as.Date("1960-01-01"),]
fac3<-fac3[fac3$Date<as.Date("2015-12-31"),]

#Redefine NA convention
ind48[ind48==-99.99]=NA
#Remove columns with NA
ind48<-ind48[,colSums(is.na(ind48))==0]

#Excess return matrix: subrtract risk free rate from portfolio returns
xsret<-ind48[,2:length(ind48)]-fac3$RF

# number of periods
T=length(fac3$Date)
# number of portfolios/industries
N=dim(ind48)[2]-1
# number of factors
K=3

# run the time series regression
beta=matrix(0,K+1,N)
predxsret=matrix(0,1,N)


# X is T*4 maxtrix for factors scale overtime
# beta is a 4*N matrix: factor loading for each industry with first row constant
X=cbind(1,fac3[,2:4])
X=as.matrix(X)
for (i in 1:N){
  out=lm(xsret[,i]~fac3$Mkt.RF+fac3$SMB+fac3$HML)
  beta[,i]=out$coefficients
}
beta=as.matrix(beta)
pred=X%*%beta
#regression intercept are the pricing errors.
```

```
alpha=as.matrix(beta[1,])

# difference between actual excess return and predicted excess return
error=xsret-pred

# plot predicted mean excess return vs. realized mean excess return
Min=min(colMeans(xsret))
Max=max(colMeans(xsret))
plot((colMeans(pred)-beta[1,])~colMeans(xsret),xlim=range(Min,Max),ylim=range(Min,Max),xlab="actual mean excess return", ylab="predicted mean excess return", main="FF Model on industry portfolios", col="blue")
fit=lm(colMeans(xsret)~(colMeans(pred)-beta[1,]))
abline(fit$coef,col="red",lwd=2)
summary(fit)

# covariance matrix across industry portfolios
sigma=cov(error)
facmean=as.matrix(colMeans(fac3[,2:4]))
facsigma=cov(fac3[,2:4])
# calculate F-statistic for pricing error, which follow F(N,T-N-K))
Fstat=(T-N-K)/N*(1+t(facmean)%*%facsigma%*%facmean)^(-1)*(t(alpha)%*%chol2inv(chol(sigma))%*%alpha)
pf(Fstat,df1=N,df2=T-N-K)
# Reject the null, pricing error are jointly deviated from zero.
```