

University of California, Los Angeles
Department of Statistics

Statistics C183/C283

Instructor: Nicolas Christou

Accuracy of historical betas

Forecasting betas with accuracy is important because they affect the inputs for the portfolio analysis problem. The variance covariance matrix is based on the value of beta for each stock. Different techniques have been proposed for better estimation of betas.

a. Unadjusted betas:

These are the betas obtained by the regression of the returns of the stock on the returns of the market.

$$R_{it} = \alpha + \beta R_{Mt} + \epsilon_{it}.$$

b. Blume's technique (1975):

The betas are adjusted as follows. Let us assume that we want to forecast betas for the period 2015-18. Then we will need two five-year periods, 2009-13 and 2005-09. First, we calculate the betas for all stocks of interest for period 2005-09. We then calculate the betas for the same stocks for period 2009-13. And then we run the regression of the betas in 2009-13 on the betas in 2005-09 to get the equation

$$\hat{\beta}_{i2} = \hat{\beta}_0 + \hat{\gamma}\hat{\beta}_{i1}.$$

Assume now that we want to forecast the beta of a stock in 2007-11. Then we find its beta in the period 2002-06 and substitute it in the equation above. For example, if the equation that connects the betas in the two historical periods is

$$\hat{\beta}_{i2} = 0.3 + 0.6\hat{\beta}_{i1}.$$

and the beta for a stock in period 2 is equal to 2, then the forecasted beta for this stock for 2014-18 will be 1.5.

c. Vasicek's technique (1973):

Let $\bar{\beta}_1$ be the average beta for the sample of stocks in the historical period (2009-13), and $\sigma_{\bar{\beta}_1}^2$ be the variance of the betas for the sample of these stocks. Also, let β_{i1} be the beta of stock i in the historical period and $\sigma_{\beta_{i1}}^2$ be the variance of β_{i1} (this is obtained from the regression of the stock on the market). Then the forecasted β_i for 2014-18 will be a weighted average of β_i and $\bar{\beta}_1$ in 2002-06.

$$\beta_{i2} = \frac{\sigma_{\beta_{i1}}^2}{\sigma_{\bar{\beta}_1}^2 + \sigma_{\beta_{i1}}^2} \bar{\beta}_1 + \frac{\sigma_{\bar{\beta}_1}^2}{\sigma_{\bar{\beta}_1}^2 + \sigma_{\beta_{i1}}^2} \beta_{i1}.$$

Comparison:

Elton, Gruber, and Ulrich (1978) compared the following 4 models in terms of their ability to predict the correlation matrix of n stocks:

- a. The historical correlation matrix.
- b. The correlation matrix based on the unadjusted betas.
- c. The correlation matrix based on the Blume's technique.
- d. The correlation matrix based on the Vasicek's technique.

They found that the historical matrix perform the poorest predictions. Therefore, the single-index model not only reduces the amount of data input, but also produces better estimates of the variance-covariance matrix. However, the comparison of the three beta techniques was more difficult, but in some cases the Blume technique was the winner, while in some other the Vasicek's technique was the winner.

Example:

Data were collected for four stocks IBM, BOEING, J&J, CATERPILLAR in two periods, February 1980 - December 1984 and February 1985 - December 1989.

```
#Access the data for the two periods.
#Period 1: February 1980 - December 1984
a1 <- read.table("http://www.stat.ucla.edu/~nchristo/statistics_c183_c283/
  stocks5_period1.txt", header=TRUE)

#Period 2: February 1985 - December 1989
a2 <- read.table("http://www.stat.ucla.edu/~nchristo/statistics_c183_c283/
  stocks5_period2.txt", header=TRUE)

#Initialize the vectors and matrices.
#Period 1:
beta1 <- rep(0,4)
var_beta1 <- rep(0,4)
beta_adj1 <- rep(0,4)
#Drop the first column (date):
a1 <- a1[,2:6]

#Perform regression of each stock on the index and record beta and the
#variance of beta in period 1:
for(i in 1:4){
  q <- lm(data=a1[,i], formula=a1[,i] ~ a1[,5])
  beta1[i] <- q$coefficients[2]
  var_beta1[i] <- vcov(q)[2,2]
}

#Vasicek's method:
for(i in 1:4){
  beta_adj1[i] <- var_beta1[i]*mean(beta1)/(var(beta1)+var_beta1[i]) +
  var(beta1)*beta1[i]/(var(beta1)+var_beta1[i])
}

#Compute betas for period 2:
beta2 <- rep(0,4)
#Drop the first column (date):
a2 <- a2[,2:6]

for(i in 1:4){
  q <- lm(data=a2[,i], formula=a2[,i] ~ a2[,5])
  beta2[i] <- q$coefficients[2]
}
```

```

#Compare:
betas <- as.data.frame(cbind(beta1, beta2, beta_adj1))
> betas
      beta1      beta2 beta_adj1
1 0.7508617 0.8146947 0.7965522
2 1.4539573 1.0752770 1.2618966
3 0.8096664 0.8482616 0.9221883
4 1.1980350 1.0867649 1.1568188

#Blume's method: Forecast of betas in period 3 from February 1990 to
#December 1994:
blume <- lm(betas$beta2 ~ betas$beta1)
beta3 <- blume$coefficients[1]+blume$coefficients[2]*beta2
> as.data.frame(beta3)
      beta3
1 0.8588805
2 0.9652936
3 0.8725881
4 0.9699849

The regression equation for Blume's method:

$$\hat{\beta}_{i2} = 0.5261865 + 0.4083665\beta_{i1}$$


#Use Vasicek's method to forecast betas in period February 1990 to
#December 1994. Perform regression of each stock on the index and
#record beta and the variance of beta in period 2:
beta_adj2 <- rep(0,4)
var_beta2 <- rep(0,4)

for(i in 1:4){
q2 <- lm(data=a2, formula=a2[,i] ~ a2[,5])
beta2[i] <- q2$coefficients[2]
var_beta2[i] <- vcov(q2)[2,2]
}

#Vasicek's method:
for(i in 1:4){
beta_adj2[i] <- var_beta2[i]*mean(beta2)/(var(beta2)+var_beta2[i]) +
var(beta2)*beta2[i]/(var(beta2)+var_beta2[i])
}

#Compare the forecast betas using the two methods for period 3:
> beta3      #Blume's method.
[1] 0.8588805 0.9652936 0.8725881 0.9699849
> beta_adj2  #Vasicek's method.
[1] 0.8737402 0.9903623 0.9216130 1.0060186

```