Estimating Beta for a Portfolio of Stocks

Geoffery Mullings

Section 1: Preparing the data

**require**(tseries)

## Loading required package: tseries

**require**(zoo)

## Loading required package: zoo  
##   
## Attaching package: 'zoo'  
##   
## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

mytick <- **c**('ATI','TSN','GME','LB','PAYX')  
  
data <- **get.hist.quote**(mytick[1], quote="AdjClose", start="1989-12-01", compress="m", quiet=TRUE, retclass="zoo")  
  
data <- **zooreg**(data, start=**as.yearmon**(**start**(data)), frequency=12)  
  
for (i in 2:**length**(mytick))  
 {  
 temp <- **get.hist.quote**(mytick[i], quote="AdjClose", start="1989-12-01", compression="m", quiet=TRUE, retclass="zoo")  
 temp <- **zooreg**(temp, start=**as.yearmon**(**start**(temp)), frequency=12)  
 data <- **cbind**(data, temp)  
 }  
**colnames**(data) <- mytick  
**head**(data)

## ATI TSN GME LB PAYX  
## Dec 1989 NA 6.80 NA 3.43 NA  
## Jan 1990 NA 6.70 NA 3.23 NA  
## Feb 1990 NA 7.56 NA 3.64 NA  
## Mar 1990 NA 7.76 NA 4.05 0.40  
## Apr 1990 NA 8.07 NA 3.89 0.39  
## May 1990 NA 9.27 NA 4.94 0.38

**tail**(data)

## ATI TSN GME LB PAYX  
## Sep 2014 36.89 39.27 40.89 64.89 43.48  
## Oct 2014 32.67 40.25 42.44 69.87 46.56  
## Nov 2014 33.50 42.34 37.81 78.72 47.03  
## Dec 2014 34.77 40.09 33.80 84.22 45.80  
## Jan 2015 28.53 39.04 35.25 82.35 45.26  
## Feb 2015 32.92 41.34 37.53 90.03 49.23

Section 2: Retrieving the market factor.

SP <- **get.hist.quote**("^GSPC", start='1989-12-01', quote="AdjClose", compression="m", quiet=TRUE)  
SP <- **zooreg**(SP, start=**as.yearmon**(**start**(SP)), frequency=12)  
**names**(SP) <- "SP500"  
data <- **cbind**(data, SP)

Section 3: Logarithmic returns for the stocks and the index

ret <- 100 \* **diff**(**log**(data))  
**head**(ret)

## ATI TSN GME LB PAYX SP500  
## Jan 1990 NA -1.481509 NA -6.007812 NA -7.1299678  
## Feb 1990 NA 12.076366 NA 11.950154 NA 0.8502706  
## Mar 1990 NA 2.611114 NA 10.673320 NA 2.3965543  
## Apr 1990 NA 3.917115 NA -4.030772 -2.531781 -2.7255168  
## May 1990 NA 13.862990 NA 23.895617 -2.597549 8.8000911  
## Jun 1990 NA -9.261694 NA -3.501903 2.597549 -0.8926024

**tail**(ret)

## ATI TSN GME LB PAYX SP500  
## Sep 2014 -12.806105 3.366449 -2.392285 4.781976 5.946097 -1.5635458  
## Oct 2014 -12.148329 2.464914 3.720578 7.394284 6.844074 2.2936394  
## Nov 2014 2.508821 5.062226 -11.551770 11.926088 1.004389 2.4237469  
## Dec 2014 3.720951 -5.460534 -11.211282 6.753517 -2.650160 -0.4197308  
## Jan 2015 -19.779878 -2.654017 4.200473 -2.245397 -1.186045 -3.1532822  
## Feb 2015 14.312421 5.724379 6.267508 8.916449 8.407955 5.0083283

Section 4: Analyzing the data: Calculating the descriptive stats for the assets and the market returns, plotting them, and calculating correlation coefficients.

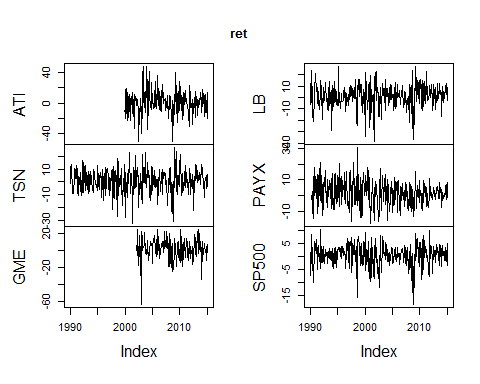
**require**(fBasics)

## Loading required package: fBasics  
## Loading required package: timeDate  
## Loading required package: timeSeries  
##   
## Attaching package: 'timeSeries'  
##   
## The following object is masked from 'package:zoo':  
##   
## time<-  
##   
##   
##   
## Rmetrics Package fBasics  
## Analysing Markets and calculating Basic Statistics  
## Copyright (C) 2005-2014 Rmetrics Association Zurich  
## Educational Software for Financial Engineering and Computational Science  
## Rmetrics is free software and comes with ABSOLUTELY NO WARRANTY.  
## https://www.rmetrics.org --- Mail to: info@rmetrics.org

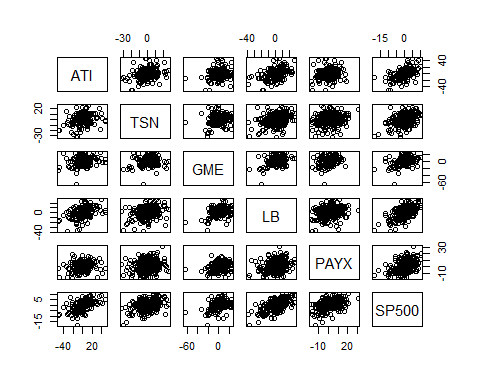
**basicStats**(ret)

## ATI TSN GME LB PAYX  
## nobs 302.000000 302.000000 302.000000 302.000000 302.000000  
## NAs 119.000000 0.000000 146.000000 0.000000 3.000000  
## Minimum -50.582850 -32.437495 -64.102189 -38.813365 -17.831724  
## Maximum 47.757260 26.602058 24.198801 27.284112 30.486041  
## 1. Quartile -7.928835 -4.888740 -4.908240 -4.018786 -3.888885  
## 3. Quartile 9.419565 6.601945 8.755280 8.227696 6.871681  
## Mean 0.339388 0.597652 0.941499 1.081981 1.609630  
## Median 0.340189 1.466547 1.728554 1.804020 1.769958  
## Sum 62.108087 180.490794 146.873803 326.758269 481.279393  
## SE Mean 1.160293 0.536596 0.989158 0.593851 0.458608  
## LCL Mean -1.949967 -0.458302 -1.012472 -0.086645 0.707109  
## UCL Mean 2.628743 1.653605 2.895469 2.250607 2.512151  
## Variance 246.369083 86.956309 152.635676 106.503182 62.886081  
## Stdev 15.696149 9.325037 12.354581 10.320038 7.930074  
## Skewness -0.060566 -0.452142 -1.053220 -0.616875 0.105061  
## Kurtosis 1.280306 1.010546 3.869442 1.347896 0.015104  
## SP500  
## nobs 302.000000  
## NAs 0.000000  
## Minimum -18.563649  
## Maximum 10.578951  
## 1. Quartile -1.795770  
## 3. Quartile 3.364167  
## Mean 0.589694  
## Median 1.101731  
## Sum 178.087704  
## SE Mean 0.245016  
## LCL Mean 0.107532  
## UCL Mean 1.071856  
## Variance 18.129969  
## Stdev 4.257930  
## Skewness -0.803669  
## Kurtosis 1.658987

**plot**(ret)



**plot**(**as.data.frame**(ret))



**cor**(ret, use="pairwise.complete")

## ATI TSN GME LB PAYX SP500  
## ATI 1.0000000 0.2998431 0.2296781 0.4142882 0.1922053 0.5422705  
## TSN 0.2998431 1.0000000 0.1724798 0.3049466 0.1724369 0.3847046  
## GME 0.2296781 0.1724798 1.0000000 0.3796804 0.2833974 0.3834414  
## LB 0.4142882 0.3049466 0.3796804 1.0000000 0.2701079 0.5605976  
## PAYX 0.1922053 0.1724369 0.2833974 0.2701079 1.0000000 0.4568763  
## SP500 0.5422705 0.3847046 0.3834414 0.5605976 0.4568763 1.0000000

While ATI & GME have relatively low minimums, ATI's maximum was quite high. For ATI and PAYX, the mean and median are very close, suggesting both securities may have little skew. All securities are at least marginally negatively skewed, except PAYX. Notably ATI's variance was massive, twice that of GME's. GME's is very fat tailed, probably due to its limited data. Yet it is noteworthy that ATI has similarly limited data and not even half of GME's kurtosis.

Elsewhere, in the logarithmic returns plots, significant events like the 2008 crash and the post-Dot-Com Recession show up. And ATI holds the strongest correlation with the market, but all of the assets are well correlated with the S&P 500.

Section 5: Estimating the Linear Regression Model

fit <- **lm**(ret[,1:5] ~ ret[,6])  
 **summary**(fit)

## Response ATI :  
##   
## Call:  
## lm(formula = ATI ~ ret[, 6])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -40.715 -6.625 -1.214 6.652 37.450   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.2442 1.0530 -0.232 0.817   
## ret[, 6] 2.1667 0.2429 8.922 1.27e-15 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.09 on 154 degrees of freedom  
## (146 observations deleted due to missingness)  
## Multiple R-squared: 0.3408, Adjusted R-squared: 0.3365   
## F-statistic: 79.6 on 1 and 154 DF, p-value: 1.266e-15  
##   
##   
## Response TSN :  
##   
## Call:  
## lm(formula = TSN ~ ret[, 6])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -24.2546 -5.2416 0.4866 4.5794 25.4474   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.4680 0.6914 0.677 0.5   
## ret[, 6] 0.8814 0.1594 5.528 1.36e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 8.597 on 154 degrees of freedom  
## (146 observations deleted due to missingness)  
## Multiple R-squared: 0.1656, Adjusted R-squared: 0.1601   
## F-statistic: 30.55 on 1 and 154 DF, p-value: 1.356e-07  
##   
##   
## Response GME :  
##   
## Call:  
## lm(formula = GME ~ ret[, 6])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -57.788 -6.811 -0.147 7.025 24.864   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.4932 0.9206 0.536 0.593   
## ret[, 6] 1.0939 0.2123 5.152 7.78e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 11.45 on 154 degrees of freedom  
## (146 observations deleted due to missingness)  
## Multiple R-squared: 0.147, Adjusted R-squared: 0.1415   
## F-statistic: 26.55 on 1 and 154 DF, p-value: 7.778e-07  
##   
##   
## Response LB :  
##   
## Call:  
## lm(formula = LB ~ ret[, 6])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -17.7741 -5.5911 -0.1615 5.4124 21.4468   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.8898 0.5888 1.511 0.133   
## ret[, 6] 1.3897 0.1358 10.235 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 7.321 on 154 degrees of freedom  
## (146 observations deleted due to missingness)  
## Multiple R-squared: 0.4048, Adjusted R-squared: 0.401   
## F-statistic: 104.8 on 1 and 154 DF, p-value: < 2.2e-16  
##   
##   
## Response PAYX :  
##   
## Call:  
## lm(formula = PAYX ~ ret[, 6])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -16.7939 -3.5276 0.3397 3.1943 13.9126   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.07079 0.42832 0.165 0.869   
## ret[, 6] 0.88344 0.09878 8.944 1.11e-15 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.326 on 154 degrees of freedom  
## (146 observations deleted due to missingness)  
## Multiple R-squared: 0.3419, Adjusted R-squared: 0.3376   
## F-statistic: 79.99 on 1 and 154 DF, p-value: 1.112e-15

ATI slopes dramatically relative to the S&P implying high volatility for that security. All other stocks slope relatively within .3 of the S&P, which is corroborated by the data's statistical significance.