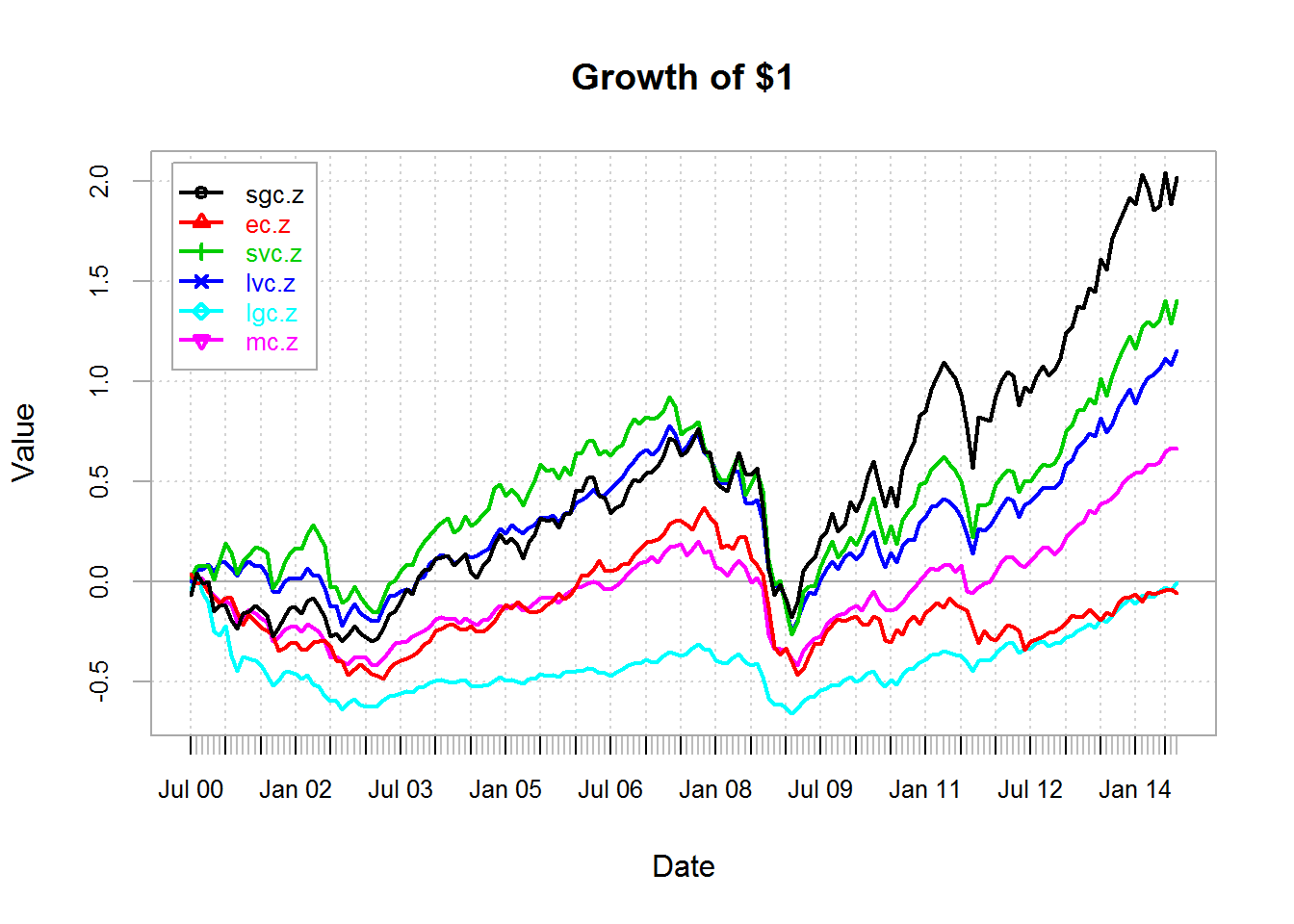
Using R for Portfolio Analysis

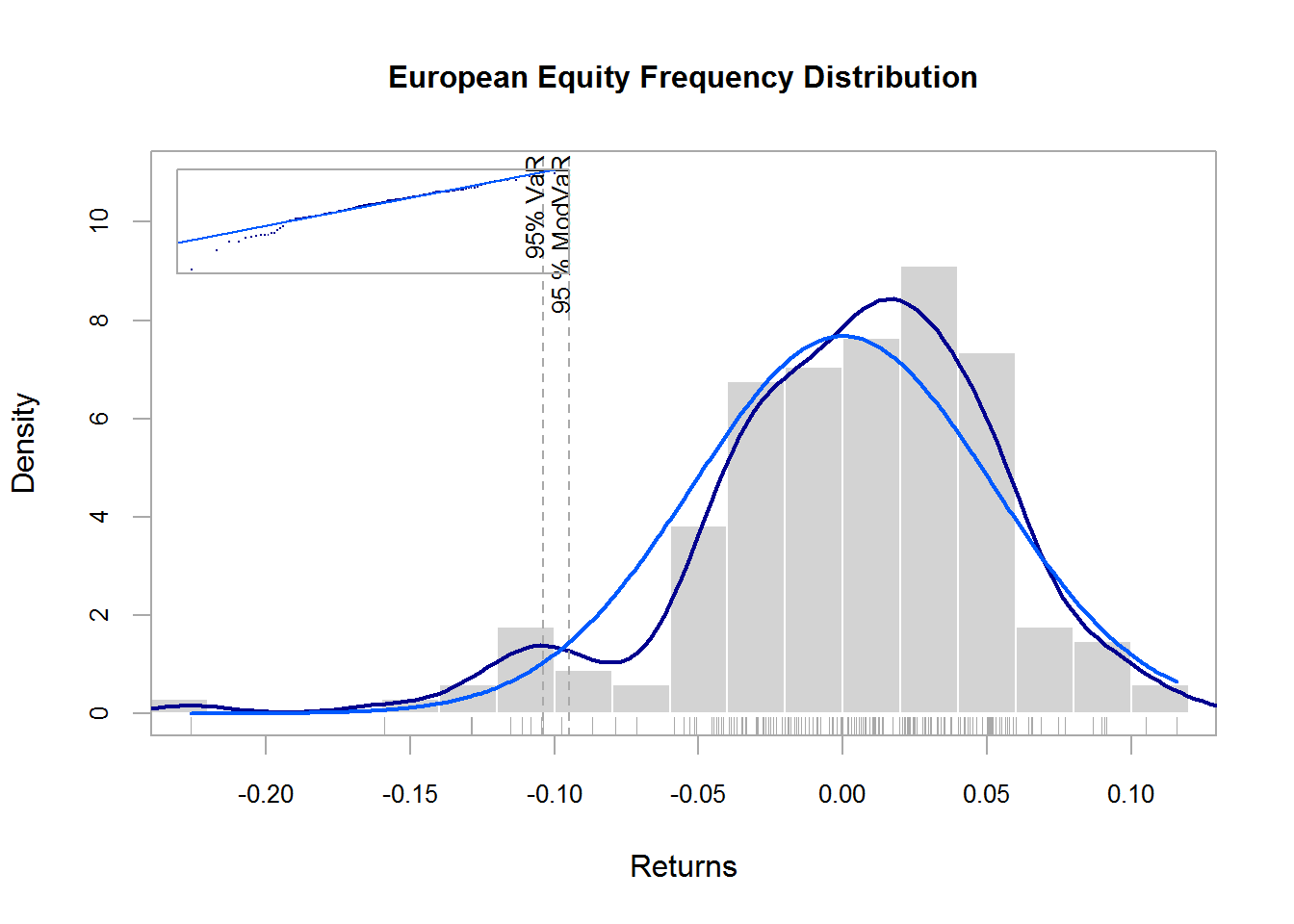
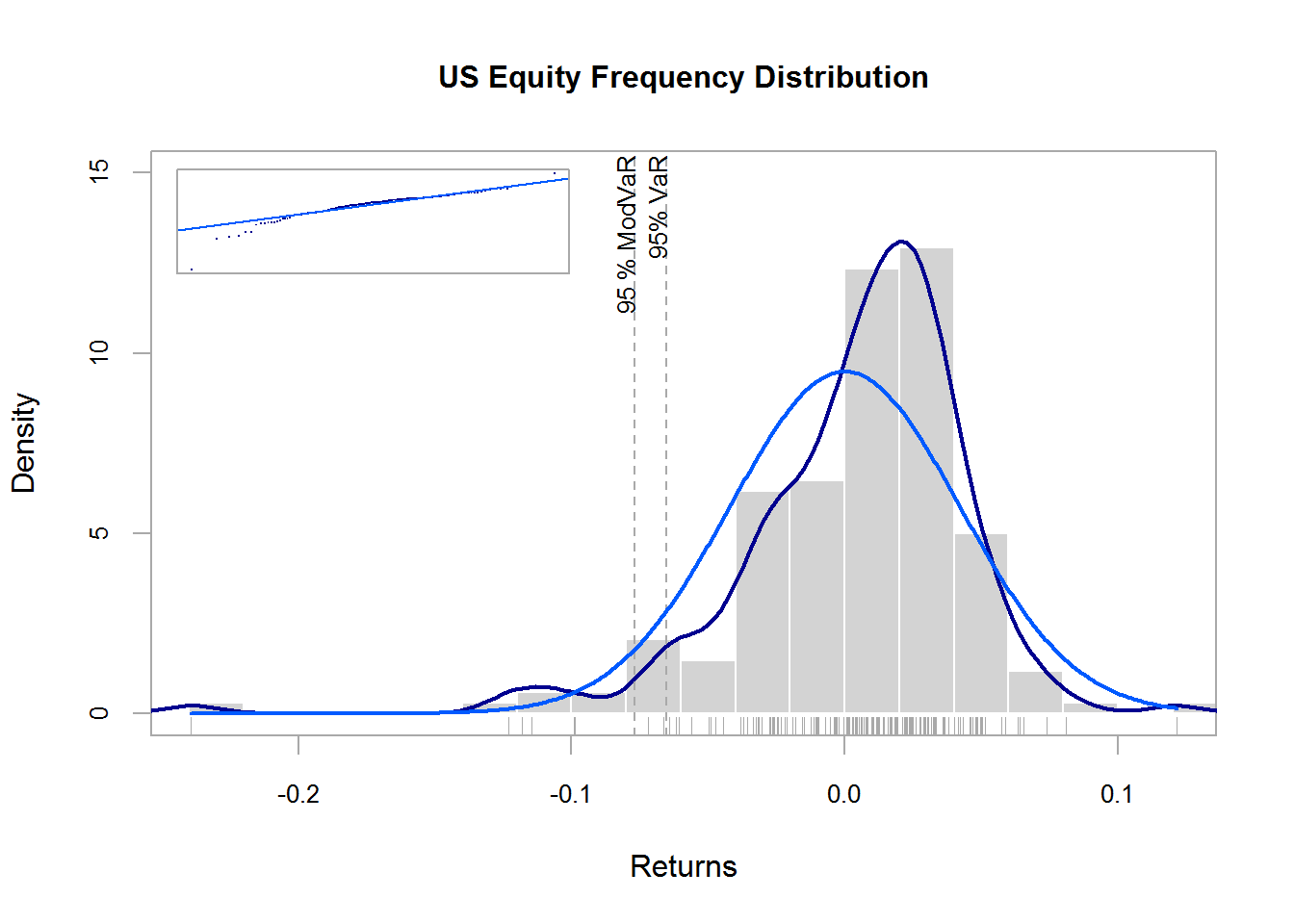
# Descriptive Statistics

The charts the cumulative growth of $1 invested in the respective asset classes since July 2000, the earliest I could find a complete history for the indices I used which are: the Wilshire 5000, the MSCI EAFE, the Russell 1000 Growth, the Russell 1000 Value, Russell 2000 Growth, Russell 2000 Value to represent US Equity, Developed Market Equity (Europe for short), US Large Growth, US Large Value, US Small Growth and US Small Value, respectively



## Examining the Frequency Distribution

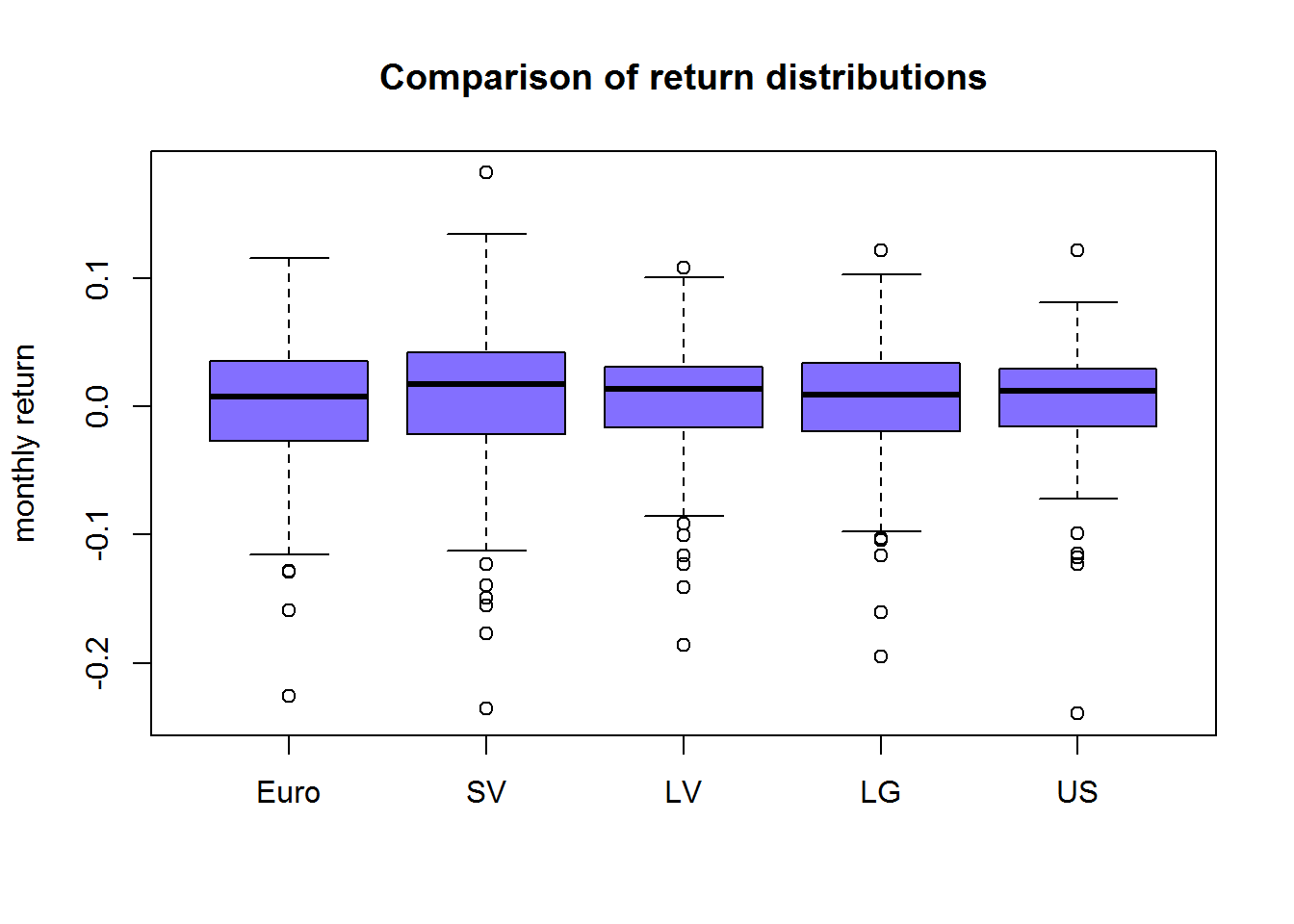
The following two histograms are unique for having the same number of breaks and the same breaks for the bins facilitating eye-ball comparison between the two. The two overlaid histograms represent the normal distribution histogram and the actual data smoothed histogram. It is clear that US equity is more clustered in the middle, which is reinforced by the high excess kurtosis of 6.5 (leptokurtic). European Equity is slightly Platykurtic. Both distributions are only slightly negatively skewed.



## US Equity vs European Equity

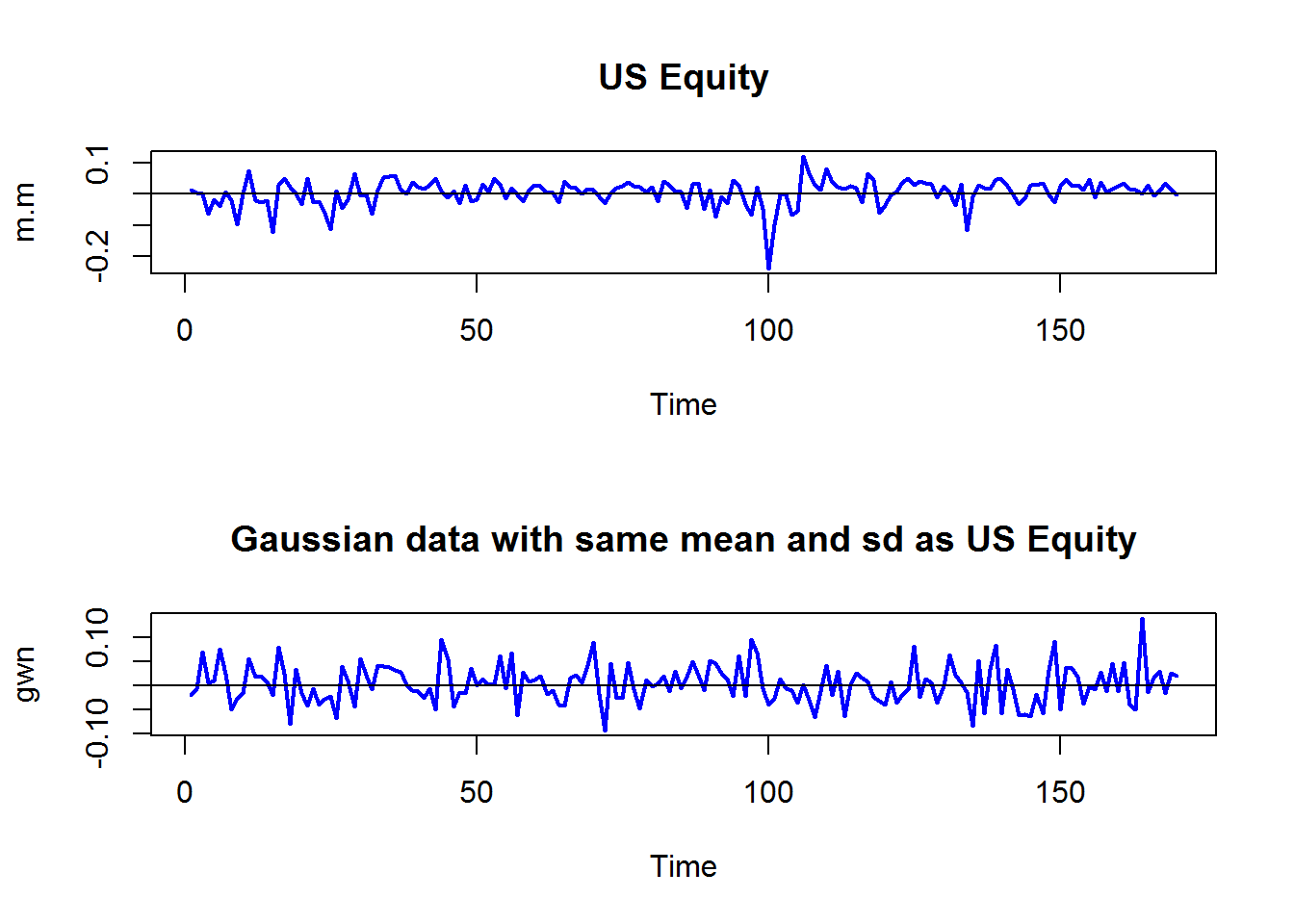
|  |  |  |
| --- | --- | --- |
|  | US Equity | Eurpean Equity |
| median | 0.01196 | 0.007262 |
| Inter Quantile Range | 0.04457 | 0.06186 |
| mean | 0.003923 | 0.001048 |
| var | 0.001768 | 0.002699 |
| sd | 0.04205 | 0.05195 |
| skewness | -1.64 | -0.9227 |
| Excess kurtosis | 6.445 | 1.906 |
| Min | -0.2394 | -0.2261 |
| Max | 0.1215 | 0.1157 |

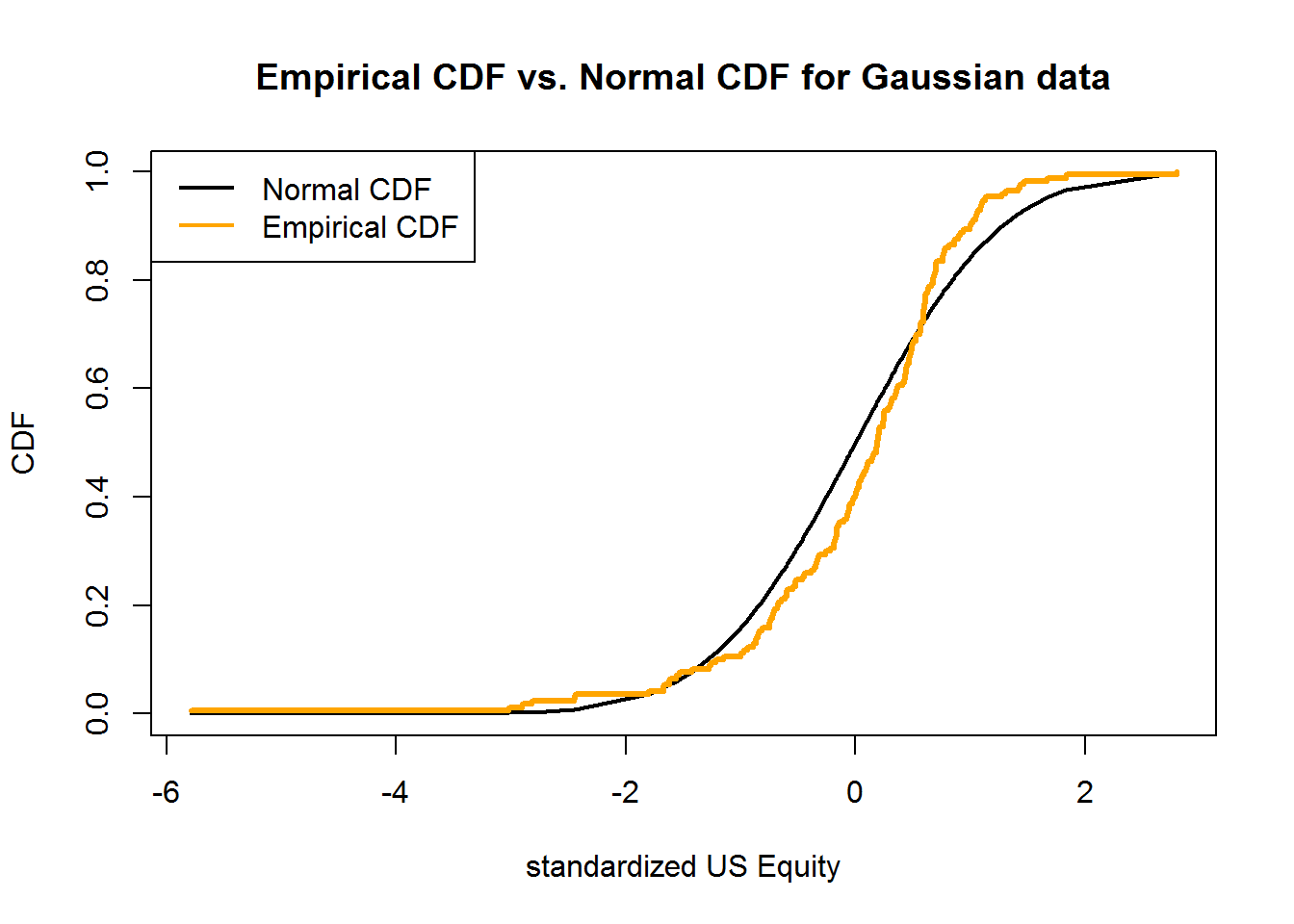
## Box Plots and Scatter Plots



## Normality Eye Ball Exercise

Comparing the Market Return with that of an IDD Gaussian Distribution with the same Mean and SD as US Equity, US data again seems more clustered around the mean with large infrequent outliers.

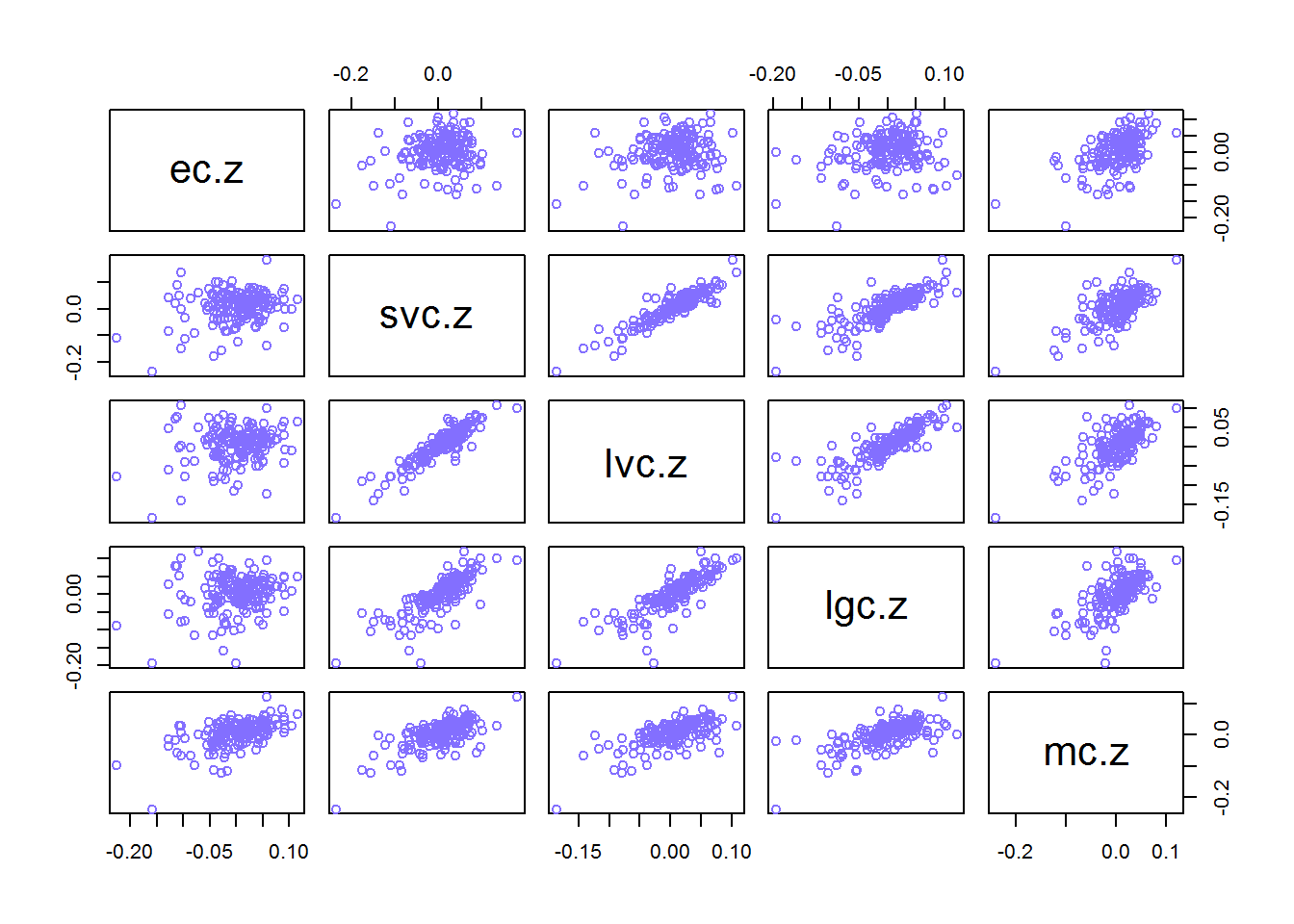




After standardizing the US equity return (forcing the mean to be zero, and the standard deviation 1) to make it comparable to the normal CDF, comparison suggests thicker tails for the market, and a negative skew.

|  |  |  |
| --- | --- | --- |
|  | 1% C.I. | 5% C.I. |
| US Equity | -0.11958 | -0.06528 |
| Normal Quantiles | -0.09389 | -0.06524 |

# correlation and the covariance Matrix



Correlations look stronger within the same capitalization category (i.e. small growth with small value). The MSCI has poor correlation with all, suggesting less integration between US and the Developed Equity in the period of the analysis.

Mean

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Europe | Small Value | Large Value | Large Growth | US |

## 0.001048 0.006850 0.005546 0.001375 0.003923

Variance

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Europe | Small Value | Large Value | Large Growth | US |

## 0.002699 0.003253 0.002001 0.002699 0.001768

Standard Deviation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Europe | Small Value | Large Value | Large Growth | US |

## 0.05195 0.05704 0.04474 0.05195 0.04205

Covariance Matrix

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Europe | Small Value | Large Value | Large Growth | US |
| Europe | 0.0026986 | 0.0006631 | 0.0004645 | 0.0004704 | 0.001199 |
| Small Value | 0.0006631 | 0.0032531 | 0.0023290 | 0.0023509 | 0.001652 |
| Large Value | 0.0004645 | 0.0023290 | 0.0020015 | 0.0019258 | 0.001274 |
| Large Growth | 0.0004704 | 0.0023509 | 0.0019258 | 0.0026992 | 0.001442 |
| US | 0.0011989 | 0.0016523 | 0.0012743 | 0.0014422 | 0.001768 |

Correlation Matrix

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Europe | Small Value | Large Value | Large Growth | US |
| Europe | 1.0000 | 0.2238 | 0.1999 | 0.1743 | 0.5489 |
| Small Value | 0.2238 | 1.0000 | 0.9128 | 0.7934 | 0.6890 |
| Large Value | 0.1999 | 0.9128 | 1.0000 | 0.8286 | 0.6775 |
| Large Growth | 0.1743 | 0.7934 | 0.8286 | 1.0000 | 0.6602 |
| US | 0.5489 | 0.6890 | 0.6775 | 0.6602 | 1.0000 |

### Estimated standard error for mean and \

|  |  |  |  |
| --- | --- | --- | --- |
| ## | muhat.vals | se.muhat | tratios |
| Europe | 0.001048 | 0.003984 | 0.2632 |
| Small Value | 0.006850 | 0.004374 | 1.5660 |
| Large Value | 0.005546 | 0.003431 | 1.6164 |
| Large Growth | 0.001375 | 0.003985 | 0.3451 |
| US | 0.003923 | 0.003225 | 1.2165 |

In some cases the SE of the mean is even bigger than the mean itself! (in Large Growth)

### Compute exact 95% confidence intervals

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean Lower | Mean Upper | Mean Width |
| Europe | -0.006857 | 0.008954 | 0.01581 |
| Small Value | -0.001829 | 0.015530 | 0.01736 |
| Large Value | -0.001262 | 0.012354 | 0.01362 |
| Large Growth | -0.006531 | 0.009282 | 0.01581 |
| US | -0.002476 | 0.010322 | 0.01280 |

The Standard Error for the mean values are very big when compared to the mean values, indicating very low estimation precision

### Compute estimated standard error for Variance and Standard Deviation

|  |  |  |
| --- | --- | --- |
| ## | Variance | Standard Error |
| Europe | 0.002699 | 0.0002927 |
| Small Value | 0.003253 | 0.0003528 |
| Large Value | 0.002001 | 0.0002171 |
| Large Growth | 0.002699 | 0.0002928 |
| US | 0.001768 | 0.0001918 |

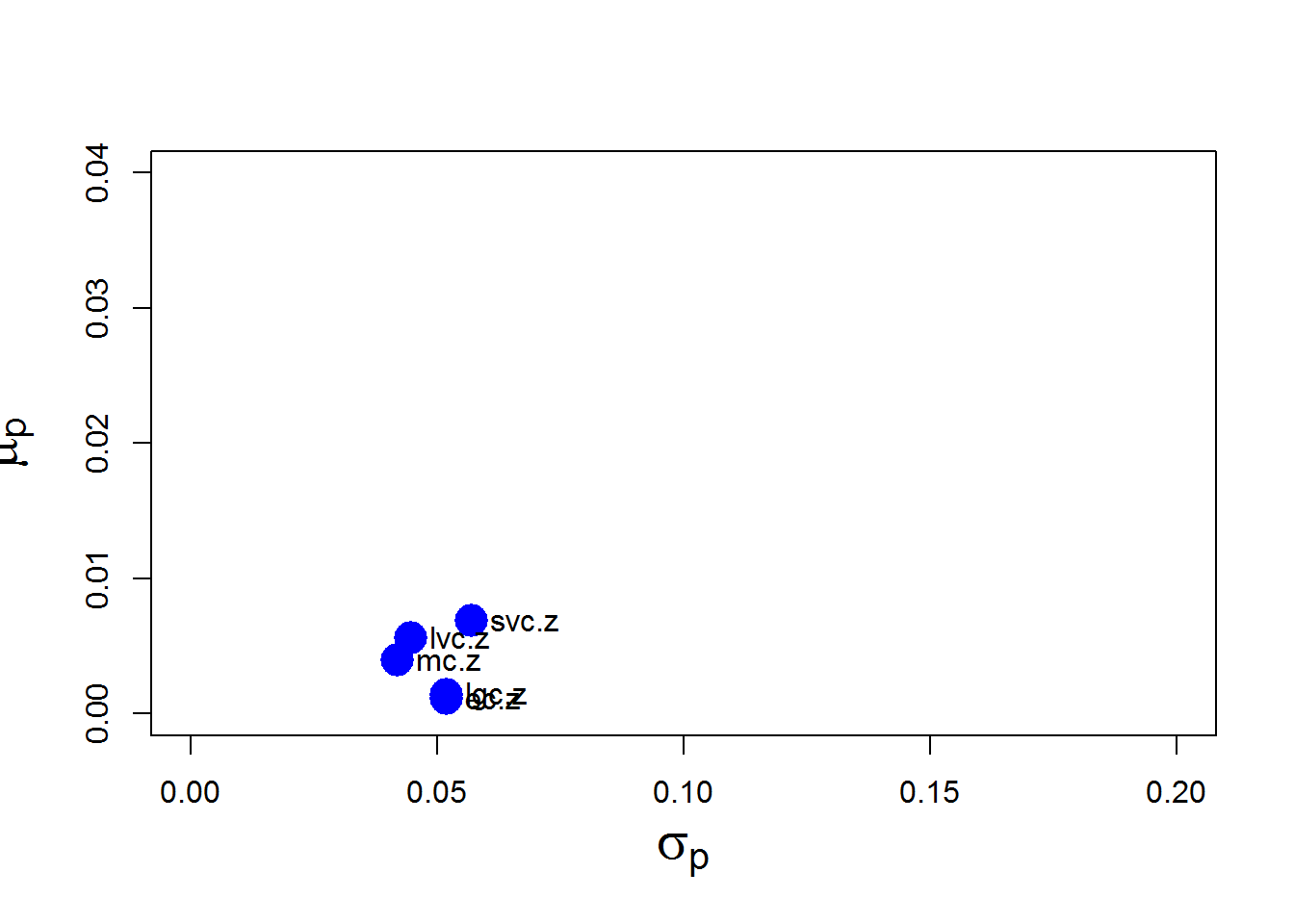
|  |  |  |
| --- | --- | --- |
| ## | Standard Deviation | Standard Error |
| Europe | 0.05195 | 0.002817 |
| Small Value | 0.05704 | 0.003093 |
| Large Value | 0.04474 | 0.002426 |
| Large Growth | 0.05195 | 0.002818 |
| US | 0.04205 | 0.002280 |

### Compute estimated standard error for Correlation

|  |  |  |
| --- | --- | --- |
| ## | Correlation | SE |
| ## sve | 0.2238 | 0.07286 |
| ## lve | 0.1999 | 0.07363 |
| ## lge | 0.1743 | 0.07437 |
| ## me | 0.5489 | 0.05359 |
| ## lvsv | 0.9128 | 0.01280 |
| ## lgsv | 0.7934 | 0.02842 |
| ## msv | 0.6890 | 0.04029 |
| ## lvlg | 0.8286 | 0.02404 |
| ## lvm | 0.6775 | 0.04150 |
| ## lgm | 0.6602 | 0.04327 |

SE for Variance, Standard Deviation and correlation are much lower, indicating higher precision in calculating volatility than in calculating mean returns, which is expected.

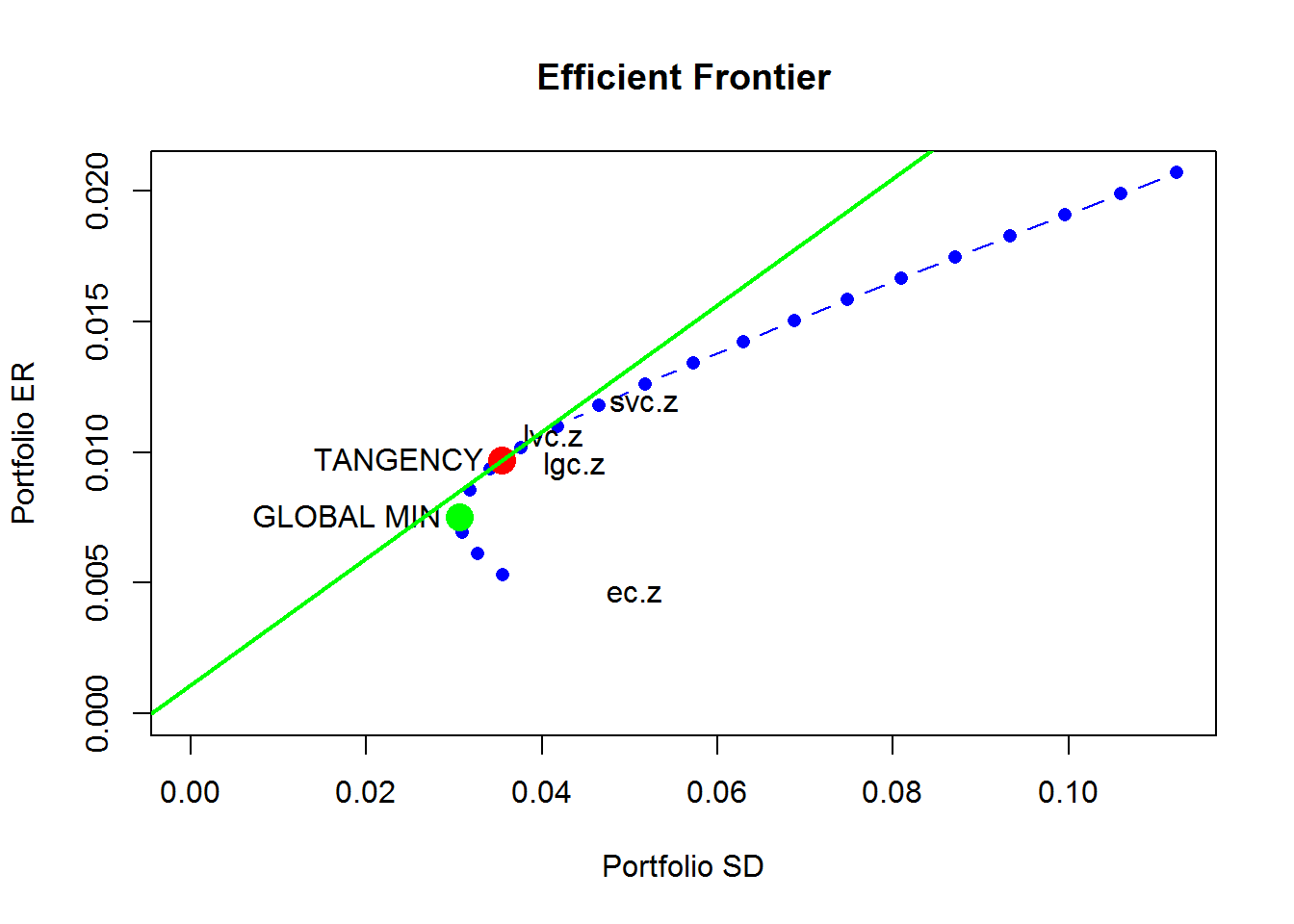
## Show risk return tradeoffs



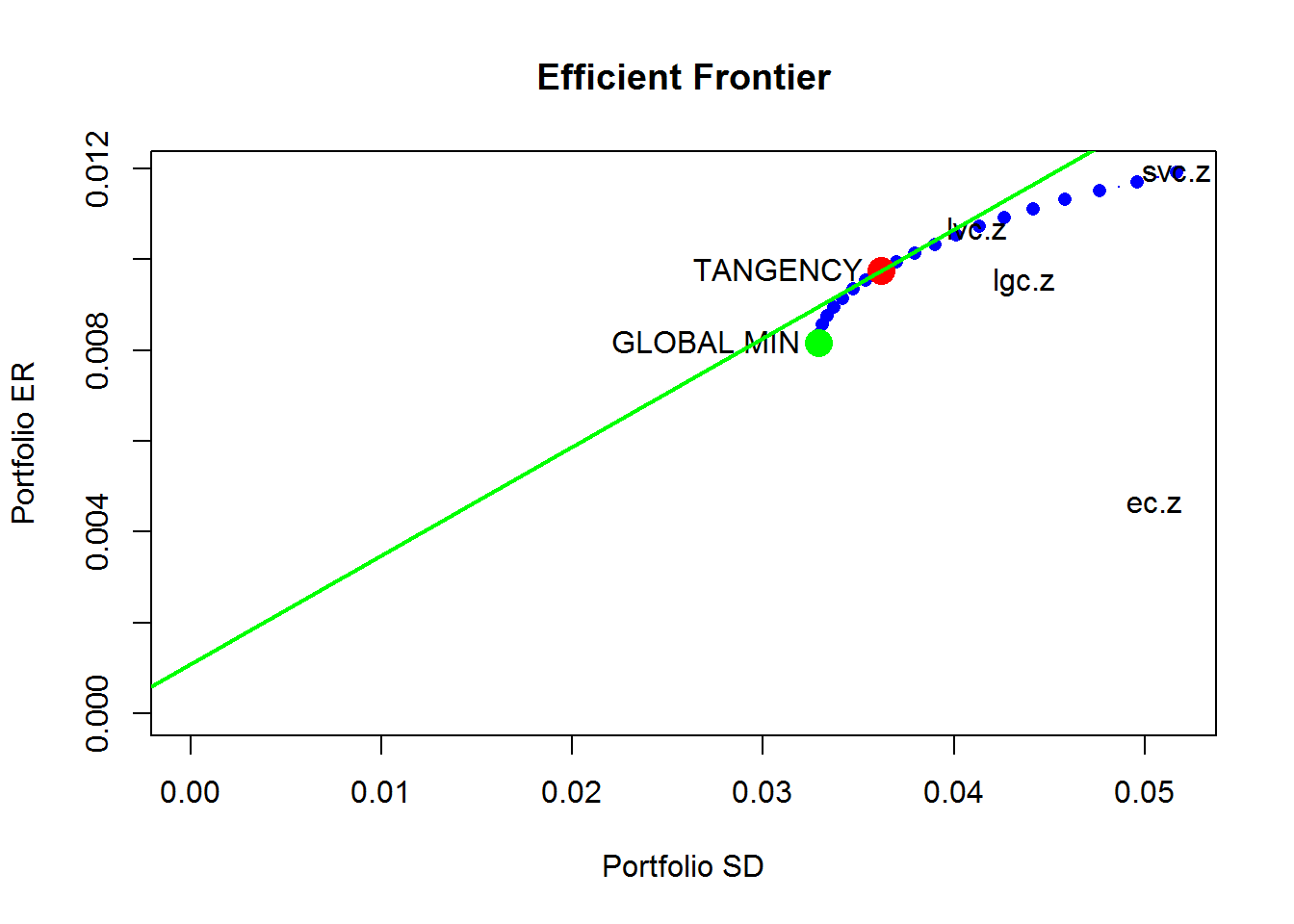
A clear tradeoff between risk and return is unclear for the period 14-year period for which return and volatility are calculated. The Will 5000 index is inferior to all other choices

## Portfolio Construction and Optimization

#### Compute the efficient frontier



### Compute portfolio frontier with no short sales



### Bootstrapping the Asset Returns and Recalculating the Efficient Frontier

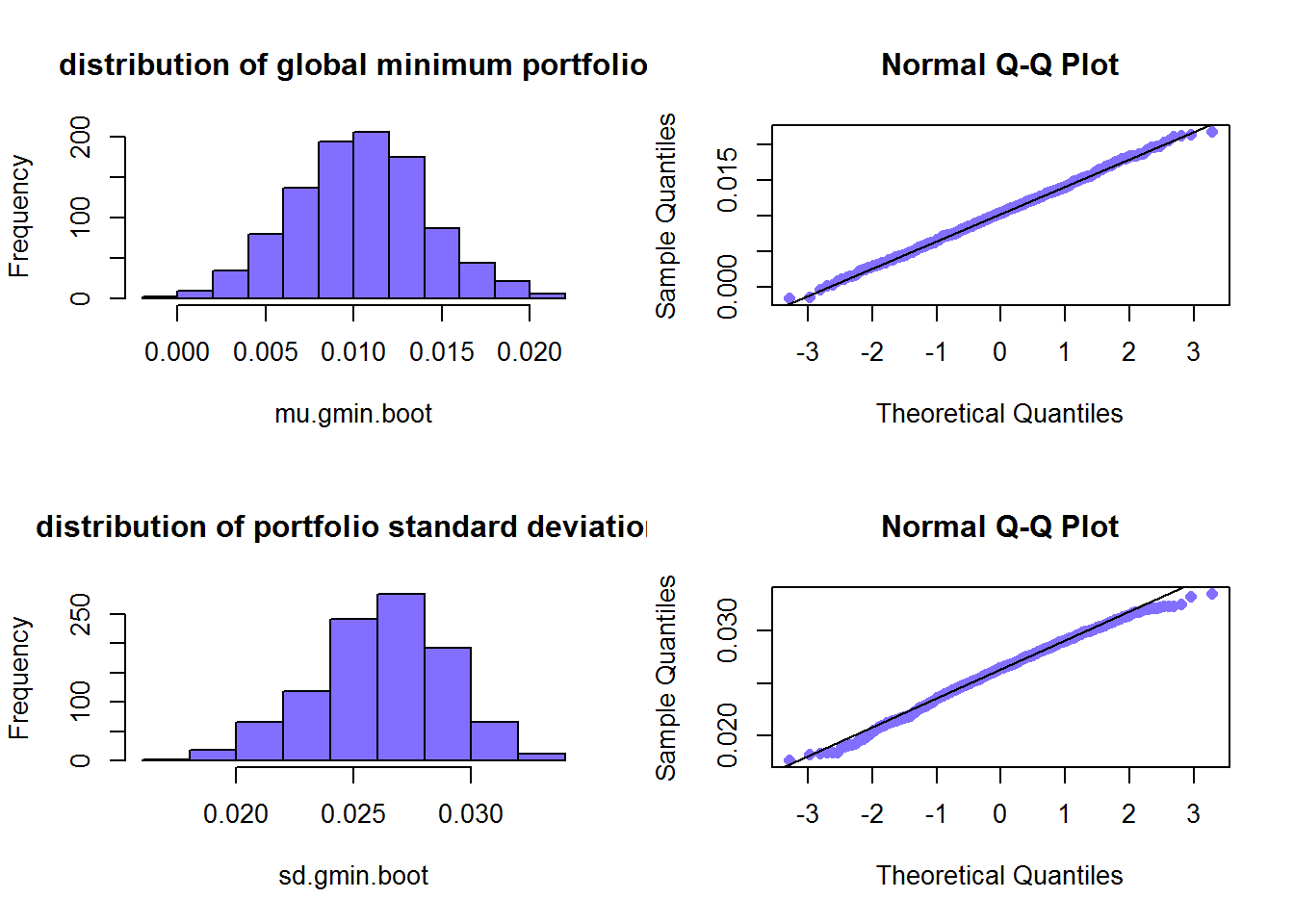
In what follows I resample both the asset returns and their standard deviations, producing a set of 200 resampled *return* *sets* each with the same number of observations as the original data sets. For each of the 200 data sets, R calculates the mean and standard deviations which are shown below.

Plot of bootstrapped mean asset returns and volatilities on the left and the original asset mean returns and volatilities on the right.

|  |  |
| --- | --- |
|  |  |

### Exploring the Bootstrap Distribution

Here I calculate the global minimum portfolio mean return and standard deviation for each of the resampled data sets.



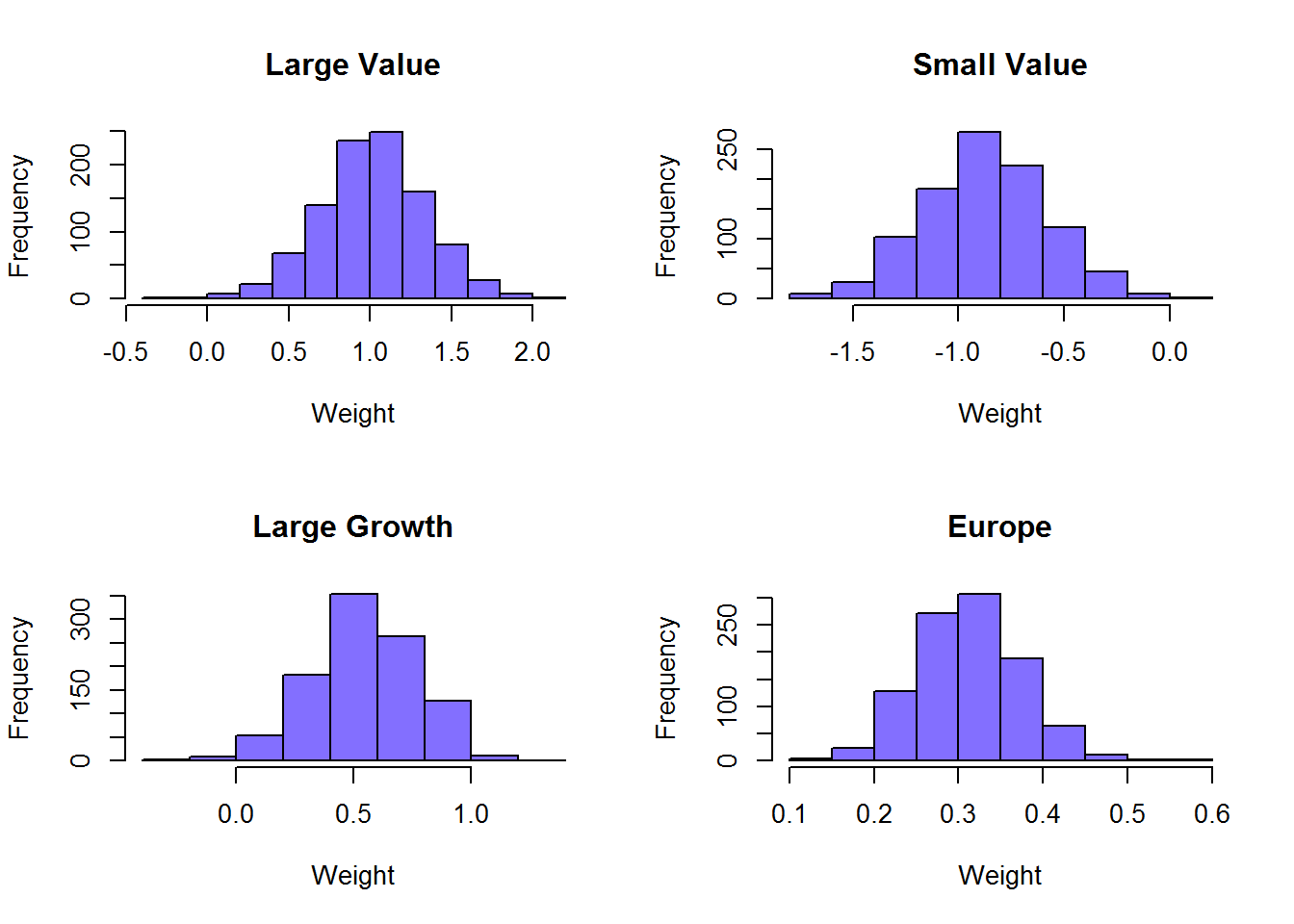
Confidence Intervals for the Global Minimum Portfolio

One of the benefits of bootstrapping is that you can estimate the bias of the original estimates of the Global Minimum Return and Global Minimum Standard Deviation, without the need of making any assumptions about the distribution of the asset returns. This proves to be more helpful the more diverse your asset types are, especially if you are including hedge funds or other option-enhanced assets that have very significant higher moments.

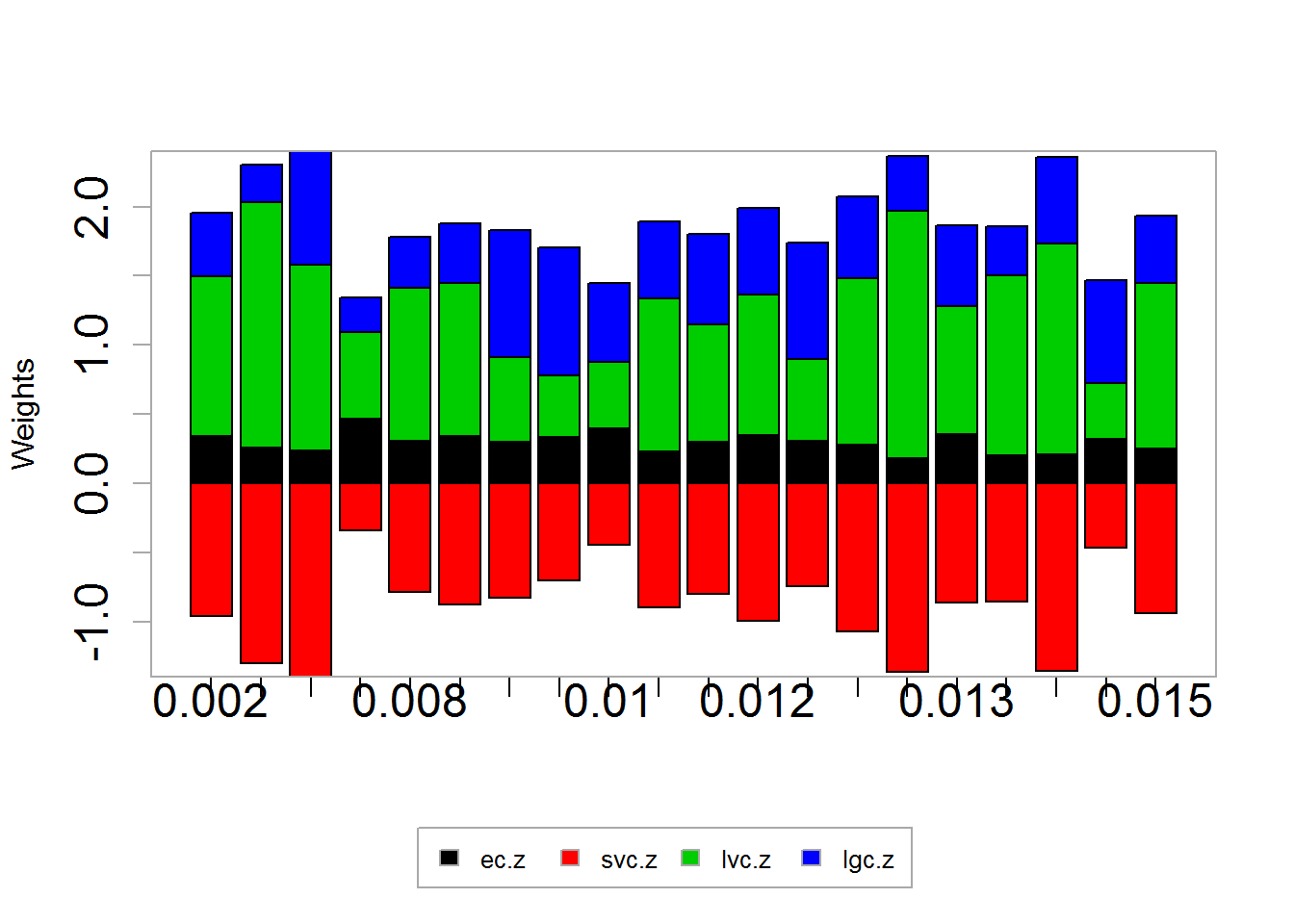
|  |  |  |  |
| --- | --- | --- | --- |
|  | Bias | Standard Error | Confidence Interval (95%) |
| Mean | 0.002853 | 0.00383 | -0.0001451 0.0151745 |
| Standard Error | -0.004387 | 0.002768 | 0.02508 0.03615 |

Asset Weights

|  |  |  |  |
| --- | --- | --- | --- |
|  | Bias | Standard Error | Confidence Interval (95%) |
| Europe | -0.03627 | 0.06216 | 0.2240 0.4726 |
| Small Value | -0.22107 | 0.29354 | -1.23791 -0.06374 |
| Large Value | -0.03985 | 0.32590 | 0.4048 1.7084 |
| Large Growth | 0.29720 | 0.22588 | -0.2058 0.6977 |



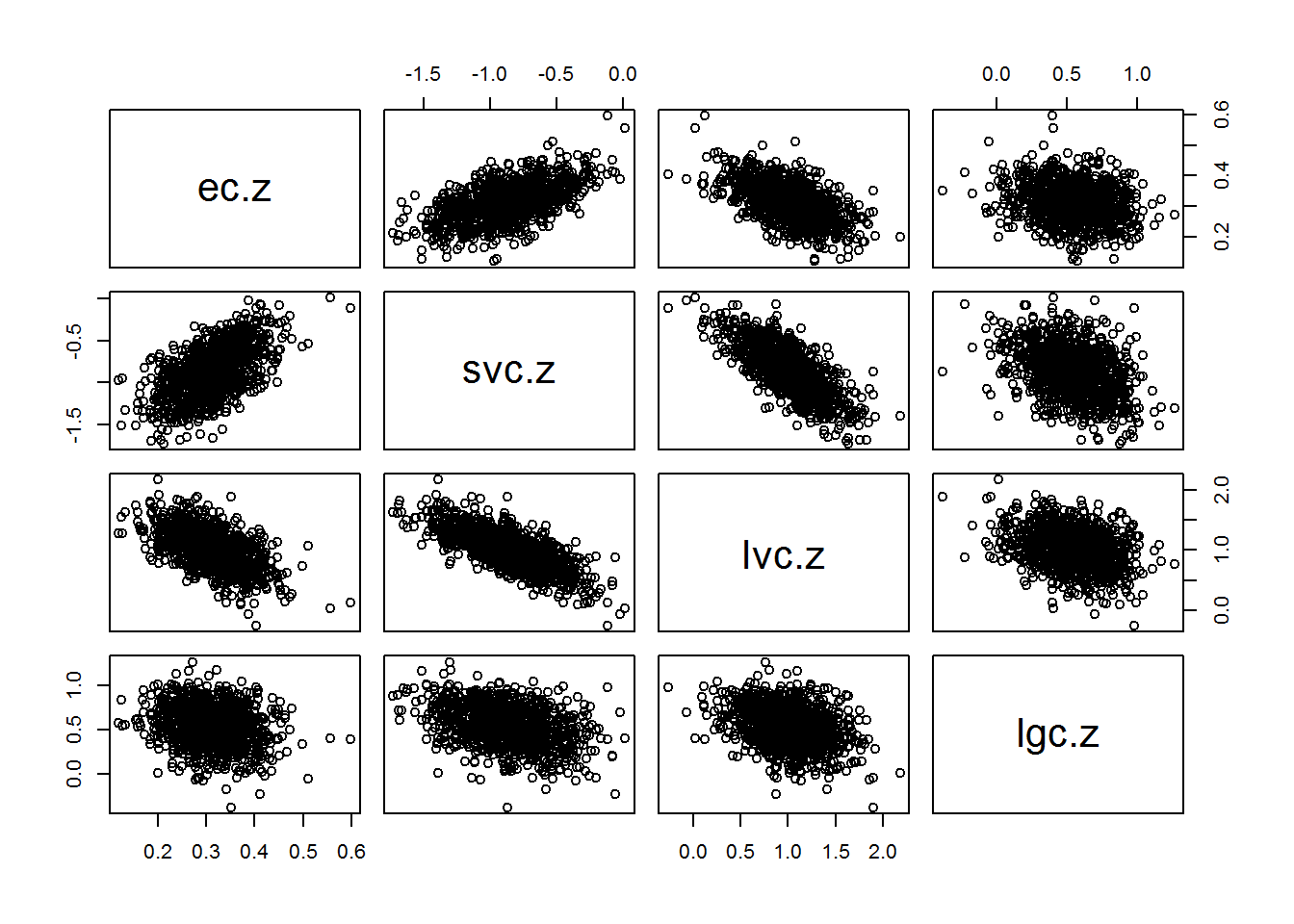
a look at a sample of Portfolios in stacked bar charts



An examination of the correlation between the Global Minimum asset weights, though aren’t comparable to return correlations, can be informative for an asset class allocation strategy.

Bootstrapped Asset Weights Correlation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ec.z | svc.z | lvc.z | lgc.z |
| ## ec.z | 1.0000 | 0.5775 | -0.5572 | -0.2217 |
| ## svc.z | 0.5775 | 1.0000 | -0.7512 | -0.3747 |
| ## lvc.z | -0.5572 | -0.7512 | 1.0000 | -0.3133 |
| ## lgc.z | -0.2217 | -0.3747 | -0.3133 | 1.0000 |



### Bootstrapping the Efficient Frontier

I calculated and charted the efficient frontiers for the first 40 resampled sets. Resampled efficient frontiers can be helpful for understanding the stability of the frontier, measuring its bias, standard error, etc. Also, downside risk measures calculated from resampled efficient frontiers such as VaR can also be more useful than those calculated from a parametric Monte Carlo simulation.

|  |
| --- |
|  |
|  |