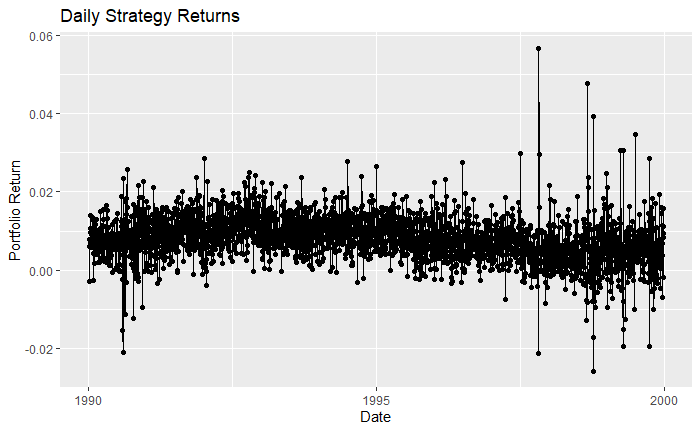
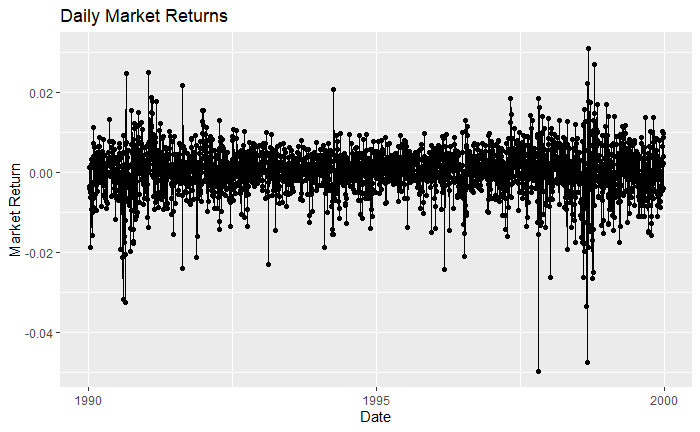
Project B: Algorithmic Trading Strategy Simulation by Ce Luo

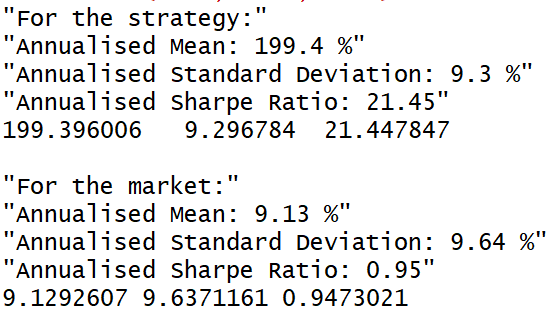
1.

(a). Strategy and market returns plot:



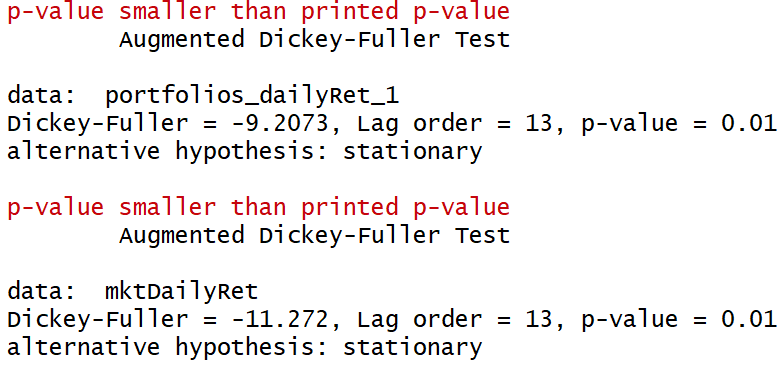


(b) Return Summary:

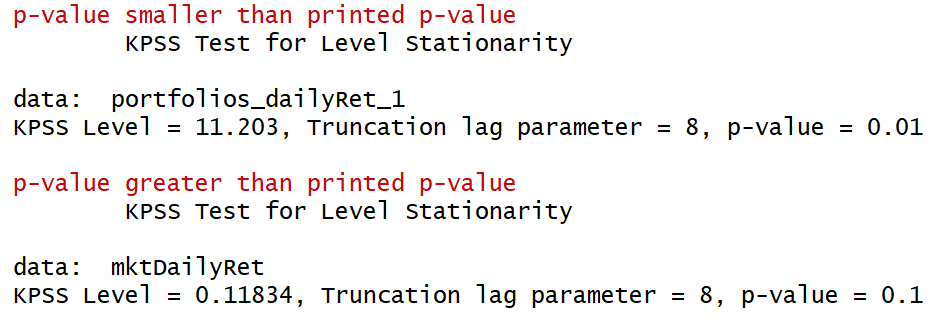


(c)

It’s hard to tell from the return plots by eyeball testing whether the two series are stationary. There seems to be some minor shift of strategy returns and there are some persistent peaks for both series.

However, the ADF test for unit root/stationarity strongly reject that the series are non-stationary. 

It is a little surprising to see the statistical conclusion that the daily return distributions haven’t changed over 10 years for both cases. Thus, to make the art of return analysis a little more scientific, KPSS test is performed to crosscheck the conclusions.

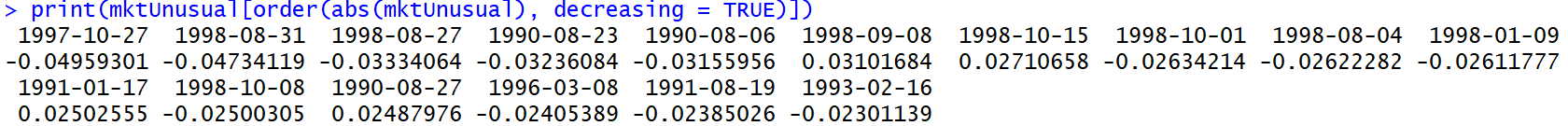


Indeed, we have the contradictory results where KPSS test rejects the stationarity null hypothesis for both series. It is possible that ADF and KPSS give opposite conclusions, since ADF is testing the difference of a series for unit root, while KPSS is testing trend stationarity around a mean. It is suggested that heteroskedasticity could be causing the divergence of test results here, and we may draw the conclusion that the return series are not stationary, but are not unit root non-stationary.

Reference: stats.stackexchange.com/questions/30569/what-is-the-difference-between-a-stationary-test-and-a-unit-root-test/235916#235916

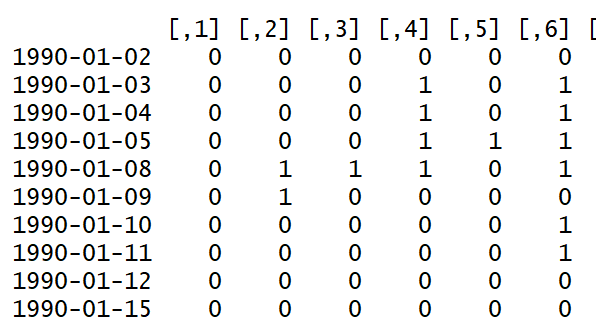
(d)

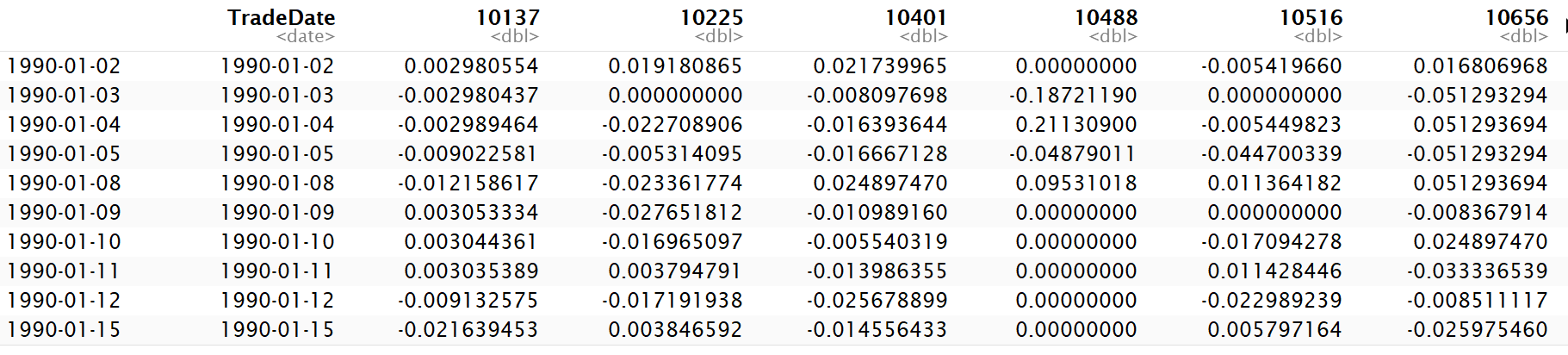
To define unusual market event and stock outliers, I decide to use the more robust statistics such as median and mean absolute deviation (MAD). The scale constant of mad() function in R is 1.4826, which will make it an unbiased estimator of the standard deviation of a normal distribution. Sometimes 2.5 MAD is used to detect outliers, but I decide to be more aggressive here by adopting the threshold of 5 MAD with financial market data which are believed to have heavy tails. Therefore, market daily returns that are 5 MAD beyond the median are selected out as being “unusual,” while the stock returns that are 5 MAD beyond the daily stock median of the day are selected out as “outliers.”

The selected market return days are as follows:

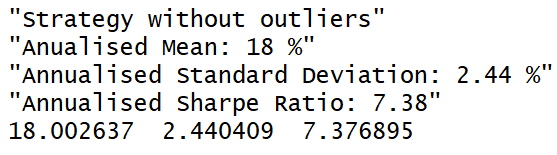
The return that deviates the most from the median happened at Oct. 27, 1997; it is the global market crash originating from Asia markets, especially the Hong Kong market. The second “unusual” market day is Aug. 31, 1998. It seems to be interpreted as a drastic reaction to the Russian political uncertainty and its financial crisis started earlier in August.

For the way I define stock “outliers,” there are many every day. A quick eyeball check shows that stock 10488 was an outlier for four consecutive days at the beginning, as well as stock 10656. However, the identical entry for 10656 indicates that there might be some data issue.



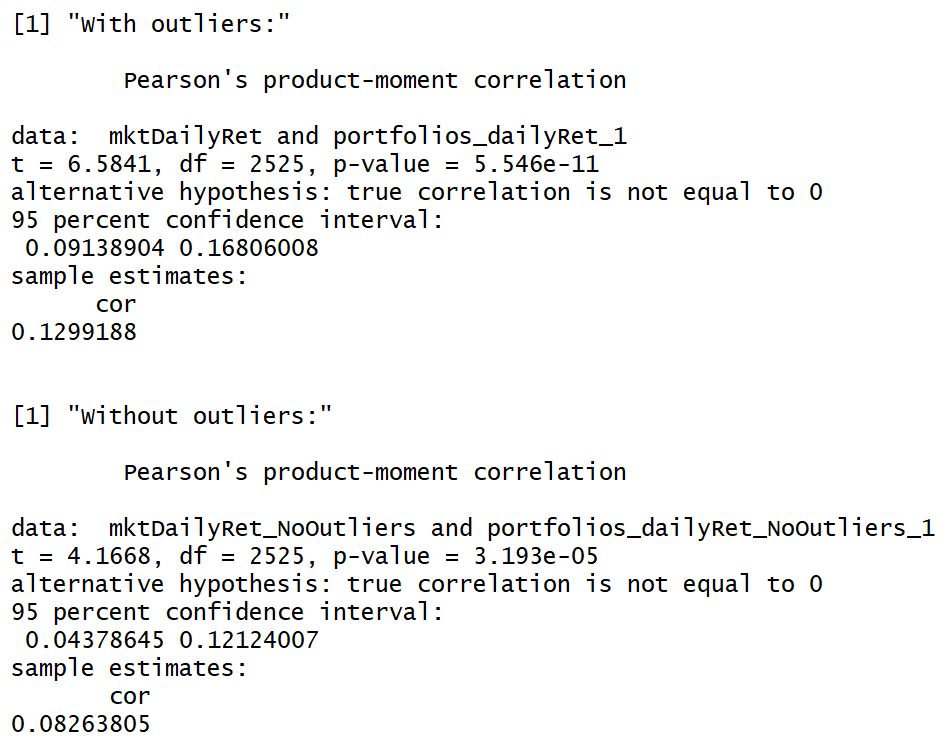


Setting the “outlier” return to be the median of returns without them, the strategy return turns out to be as follows:



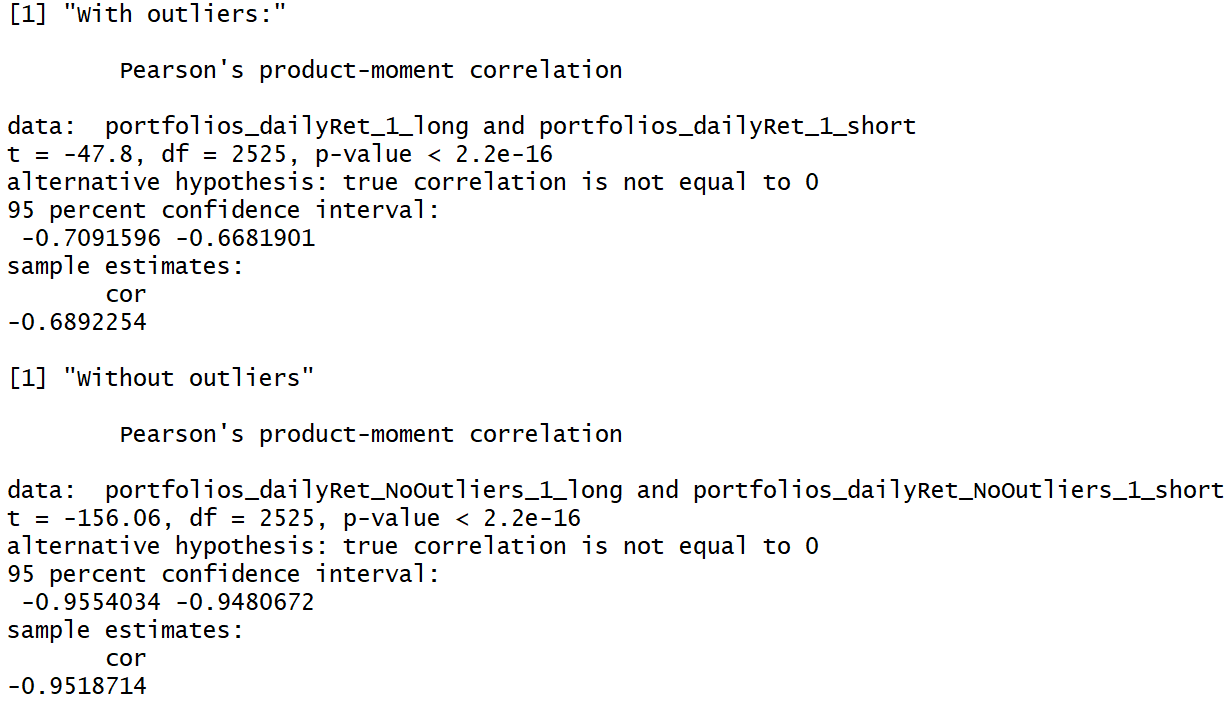
The annualized mean and Sharpe ratio dropped from about 200% to 18% and 21.45 to 7.38, respectively. Clearly the “outliers” have significant effect on the overall strategy returns.

(e)



Whether with or without the stock outliers, based on the given confidence intervals, the correlation between the strategy return and the market return is significantly positive at 5% level, though numerically small in either case. Therefore, the strategy is just dollar neutral, but is probably not far from being market neutral.

(f)



Regarding the correlation between the long and short portfolio, it turns out to be quite different with or without the outliers. Without the outliers, the long and short portfolios are very negatively correlated, which is consistent with the logic of the strategy. However, with the stock outliers present, the correlation is not that negative; the strategy nonetheless earns a much higher annualized return in this case as we have seen. Therefore, it seems that the returns of outliers which are contributing largely to the raw strategy returns actually tend to disagree with the logic of the contrarian strategy.

(g)

According to the above comparisons between the statistics with or without the outliers, it is critical to confirm whether outlier returns can be reliably reproduced with real trading. Even ignoring the transaction costs and market impact, there could be some issue with the data leading to the presence of outliers as suggested in part (d). If those returns cannot be realized in real trading, we see that the return and Sharpe ratio will drop significantly.

On the other hand, the backtesting is done under the assumption that we can trade instantaneously at the market close. It may or may not be realistic to have all the orders filled at the last second, or maybe part of it shortly in the extended hours, depending on how great the technology and market makers work nowadays. More importantly, it is usually not possible to use the short-sell proceeds to make further investment, as they are required to be kept as the margin.

2.

(a) Table based on Rmd output:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| With Outliers | k = 1 | k = 2 | k = 3 | k = 4 | k = 5 |
| Annual Mean | 199.4 % | 19.55 % | 16.93 % | 12.85 % | 9.21% |
| Annual SD | 9.3 % | 7.72 % | 7.14 % | 7.25 % | 7.04 % |
| Sharpe Ratio | 21.45 | 2.53 | 2.37 | 1.77 | 1.31 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Without Outliers | k = 1 | k = 2 | k = 3 | k = 4 | k = 5 |
| Annual Mean | 18 % | 2.99 % | 1.93 % | 1.78 % | 0.75 % |
| Annual SD | 2.44 % | 2.19 % | 2.17 % | 2.1 % | 2.14 % |
| Sharpe Ratio | 7.38 | 1.37 | 0.89 | 0.85 | 0.35 |

The return means and Sharpe ratios appear to decrease with increasing lags, either with or without the outliers. In particular, there is a sharp performance decrease from k = 1 to k = 2, so it indicates that the short-term mean reversion tends to happen right the next trading day. In short, the optimal k appears to be 1 based on the outstanding Sharpe ratio.

3.

I choose to modify the given contrarian strategy with two changes. First, the signal is constructed based on the difference between k-day cumulative (log) return and the market median cumulative return. Second, the weights are proportional to those differences.

The following tables give a summary of return statistics, correlation with market return, and correlation between the long and short sub-portfolios.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| With Outliers | k = 1 | k = 2 | k = 3 | k = 4 | k = 5 |
| Annual Mean | 434.74 % | 356.78 % | 315.05 % | 290.33 % | 271.46 % |
| Annual SD | 22.54 % | 20.17 % | 19.68 % | 19.49 % | 19.44 % |
| Sharpe Ratio | 19.28 | 17.69 | 16.01 | 14.9 | 13.97 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Without Outliers | k = 1 | k = 2 | k = 3 | k = 4 | k = 5 |
| Annual Mean | 21.86 % | 17.07 % | 14.88 % | 11.97 % | 9.61 % |
| Annual SD | 2.6 % | 2.46 % | 2.38 % | 2.34 % | 2.34 % |
| Sharpe Ratio | 8.39 | 6.93 | 6.24 | 5.11 | 4.11 |

Comparing to the previous strategies, similar patterns do emerge here., i.e., the return means and Sharpe ratio decreases with increasing lag. However, taking the cumulative return seems to eliminate the sharp decline of performance; the decreasing means and Sharpe ratios look much smoother. Additionally, the returns seem to be boosted by taking weights proportional to the deviation from the median, and we also get higher Sharpe ratios for the cases without the outliers.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| With Outliers | k = 1 | k = 2 | k = 3 | k = 4 | k = 5 |
| Correlation with Mkt Return | 0.0515  (t= 2.59) | 0.0873  (t=4.40) | 0.0868  (t=4.38) | 0.0846  (t=4.26) | 0.0887  (t = 4.47) |
| Correlation between Long Short Portfolio | -0.2552  (t = -13.26) | -0.3232  (-17.16) | -0.3458  (t = -18.52) | -0.3568  (t = -19.18) | -0.3550  (t = -19.070 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Without Outliers | k = 1 | k = 2 | k = 3 | k = 4 | k = 5 |
| Correlation with Mkt Return | 0.0702  (t=3.54) | 0.1314  (t = 6.66) | 0.0891  (t = 4.49) | 0.0810  (t = 4.08) | 0.0860  (t = 4.33) |
| Correlation between Long Short Portfolio | -0.9460  (t = -146.6) | -0.9512  (t = -154.8) | -0.9532  (t = -158.41) | -0.9548  (t = -161.4) | -0.9547  (t = -161.0) |

Again, similar patterns appear with the correlations: 1) the correlation with market return is low but significant in either case. 2) the correlation between the long short portfolios are not that negative with the outliers, but are very negative after accounting for the outliers.