Slim timeframe momentum investing with statistical augmentation

IE471 Term Project

20213581 정재우

1. Introduction

Momentum investing is one of many portfolio management strategy with its advantage in its simple but effective approach. Without a need to predict the future, this strategy simply buys a portfolio when it starts rising, and sells it when it starts to drop. Therefore, it will follow the market, but with a larger magnitude of raise during upward trend and smaller magnitude of drop during downward trend. While such intuitive approach has been actively used throughout the 20^{th} century until today[i], it has mainly been utilized in a larger timescale, from couple of months to even decades. Figure 1 shows an example of momentum investing in a timescale of decades. Smaller sampling timescales of days or weeks have been less preferred due to its large degree of volatility. Hefty volatility is mostly less preferred when managing large assets, even with a considerable expected return. In this work, momentum investing strategy in a slim timescale (in a matter of days) is developed in a way that minimizes risk by incorporating a statistical analysis on several features of stock behavior. In doing so, we have attempted to exploit the daily returns in a manner that minimizes the risk. Our strategy has achieved a sharpe ratio of xx, which is considerably much larger than the markets' ratio, -0.026 and 0.095 for KOSPI and KOSDAQ, respectively.

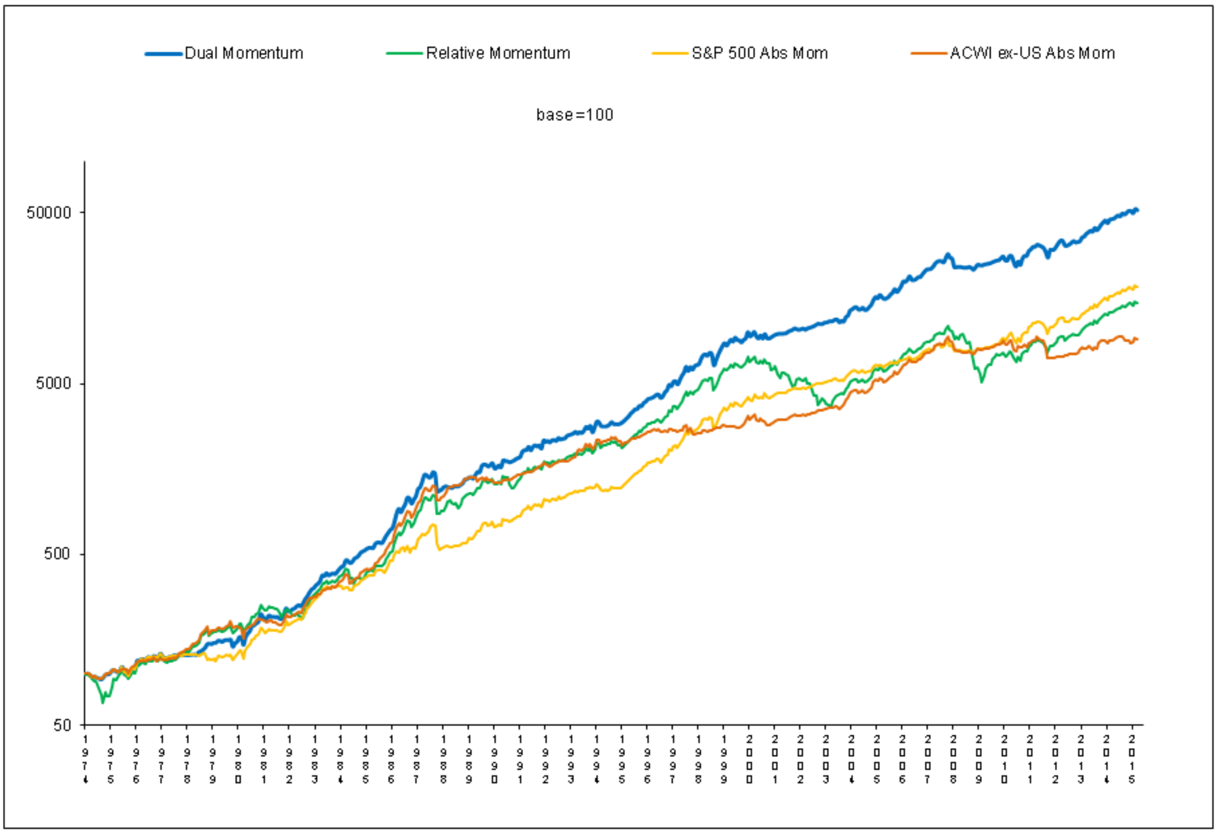


Figure 1. An example of conventional momentum investing strategy, with a timescale in years.

2. Methods

2.A. Datasets

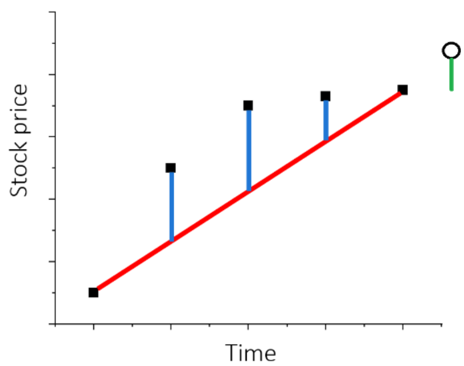
We have chosen KOSPI and KOSDAQ as the markets from which portfolios are optimized. Statistical analysis on stocks have been conducted on close price data from January 1st to April 31st of 2022 when the overall market has shown a declining trend. The proposed strategy is tested for various dates with different market behaviors. Only a period of months has been analyzed due to limitations in computation capability. Vanilla momentum investing model, which our proposed strategy is based on, is analyzed on 3 different period of times with different market behavior: inclining, declining, and stationary behaviors. Details are discussed in the results / discussion section. Sharpe ratio is calculated with risk free interest rate of zero.

2.B. Vanilla momentum investing model

Vanilla model minimizes the volatility of a portfolio from a given pool of stocks. Every stock in the market is iterated for a certain duration to determine its criterion, which we have set as $\frac{\ln(Return)}{Standard\ deviation^{0.2}}$. Return is calculated as the average log value of daily ratio compared to the previous day. Then, return is divided by standard deviation to the power of 0.2 to filter extremely volatile stocks. The power of 0.2 component normalizes the standard deviation value to become closer to 1, therefore slightly reducing the overall significance of division by standard deviation. The value of 0.2 was chosen as an intuitive measure. 10 stocks with the highest criterion value are chosen as the stock pool, from which the portfolio is optimized to minimize the volatility using cvxpy library. Portfolio is optimized based on performance of 5 days, and the portfolio is maintained in duration between 1~3 days.

2.C. Statistical analysis

Although the vanilla model performs well considering its naïve approach, statistical analysis has been applied for further stabilization of the model. For the entire stock data from January 1st to April 31st of 2022, two features have been compared in regard to their future performance: net return during sampling duration and net deviation of stock behavior from the net return slope. Sampling window has been chosen as 5 days to account for a week of stock's behavior. Figure 2 illustrates how the features are calculated. Net return is calculated as the natural log of ratio between last day performance and first day performance. Net deviation is calculated as the sum of deviation of middle dates, normalized by the day 1 price. Next day's return is calculated as the log return between next day and 5th day of window. 237,680 datapoints were acquired by sampling with a moving window for each stock.



237,680 datapoints are averaged into 48 datapoints with equal ranges of next day return. Then, the

Figure 2. Illustration of extracted features, net return as red and net deviation as blue, and next day return as green.

distribution of two features is plotted in terms of next day return. Particular feature range of interest in regard of next day return is selected in reference to the distribution, where linearity is preferred for decrease in volatility. Instead of iterating all stocks in the market, only stocks that exhibit behavior that fits in the feature range are selected for the stock pool for portfolio optimization.

3. Results & discussion

3.A. Vanilla model performance

Figure 3-5 shows the performance of a vanilla model during different market trends. The number of days denote how long the optimized portfolio is held. For a 1 day vanilla model, the portfolio is bought at a close price and sold at a close price of the following day. For this work, we assume buying and selling at close prices for simplicity.

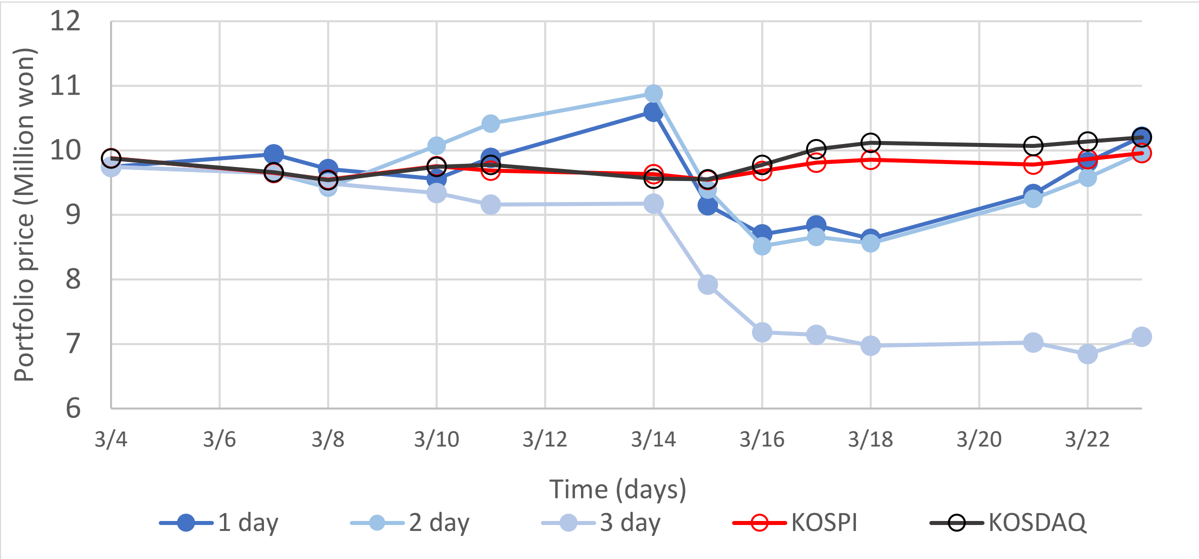


Figure 3. Vanilla model performance over a steady market, denoted by n-days. The number of days represent how long the calculated portfolio has been held.

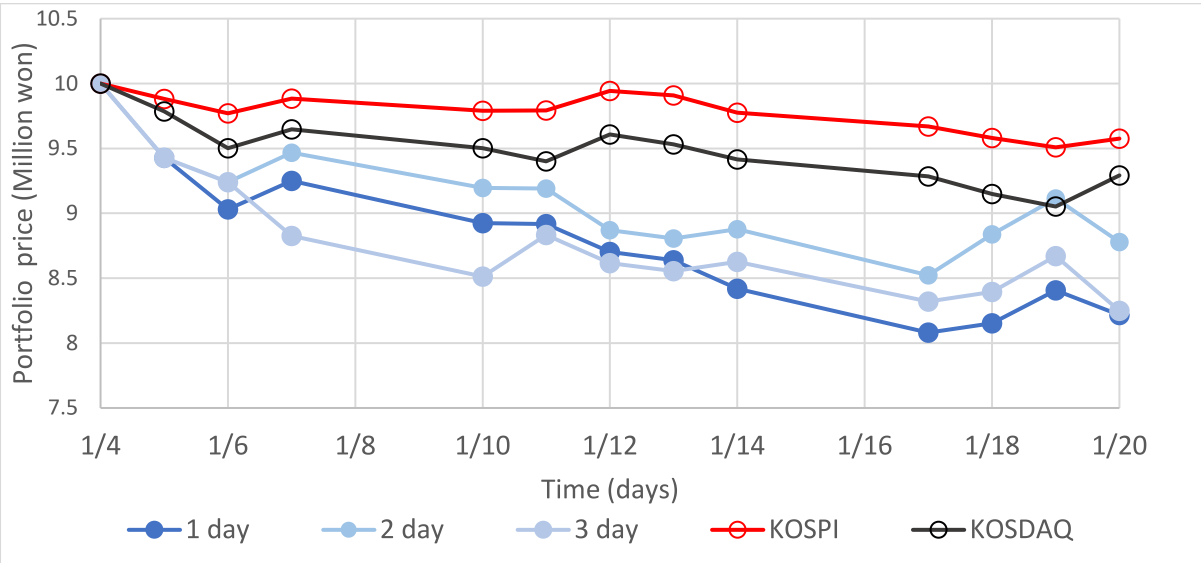


Figure 4. Vanilla model performance over a declining market.

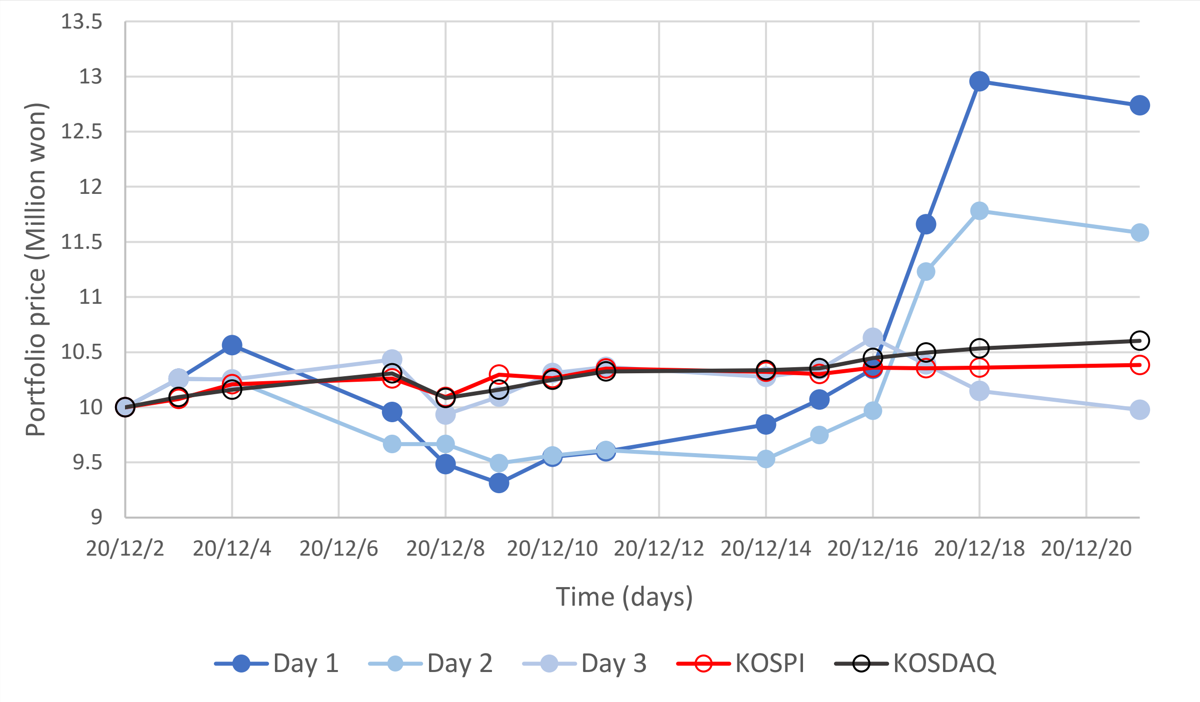


Figure 5. Vanilla model performance over an inclining market.

As shown in figure 5, the vanilla model performs considerably well during an inclining market since it is designed to find the stocks with the most return and optimize the portfolio to minimize volatility. If the market inclines as expected, the portfolio experiences grows much larger compared to the market's growth. However, when the market is declining, vanilla model suffers greater loss since stocks with large returns are likely to suffer large losses, even if it did not decline in the past. When market is stationary, the vanilla model shows analogous performance compared to the market.

3.B. Statistical analysis

One of the main drawbacks of vanilla model is that it is essentially a first order system. The convex minimization approach minimizes volatility based on the covariance matrix of the portfolio. However, this minimization does not account for second order behavior of stocks. While ongoing research attempts to address such limitation by using a separate 2nd order covariance matrix[ii], using a separate matrix increases the computation cost and only minimizes volatility in a mathematical manner based on behavior of individual stocks. Analyzing the wholistic behaviors require a different approach, and this work attempts to do so by extracting features from stock behaviors that represent a second-order derivative behavior. According to figure 2, stocks with convex behavior will result negative deviation value for either positive or negative slope. Vice versa, concave stocks will have positive deviation value for either slope sign. If a stock continues its behavior with same characteristic, the concavity and convexity measure will serve as an indicator of future performance. In doing so, the degree of concavity or convexity is quantified with net return and net deviation. Figures 6-7 quantitatively explores this idea.

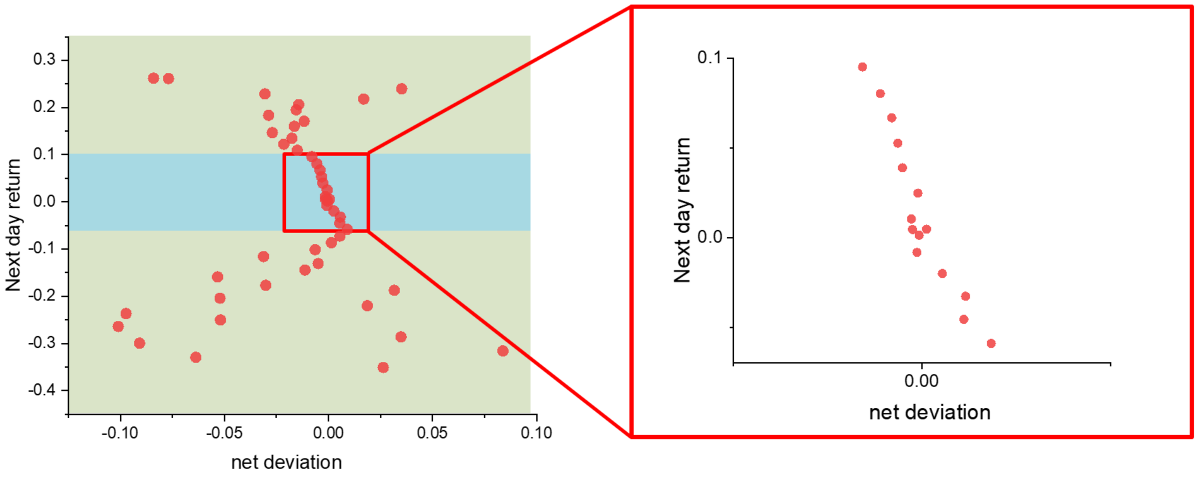


Figure 6. Net deviation vs. next day return shows a linear trend in the range of -0.05 to 0.1 next day return, highlighted in blue. Next day return expresses the return between last day of 5 day sampling window and the subsequent day, and is expressed in a log scale.

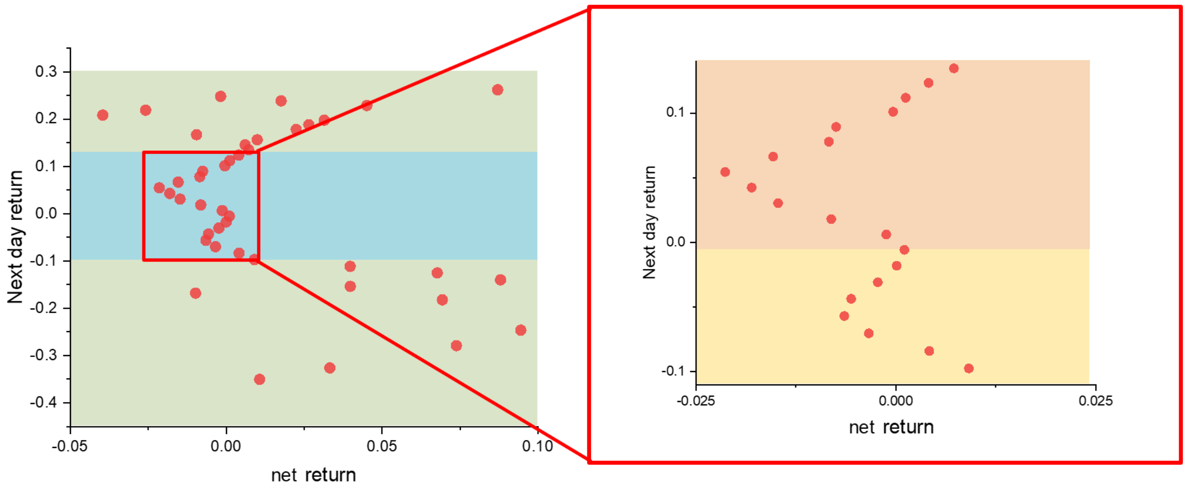


Figure 7. Net return vs. next day return shows a combination of linear trends in the range of – 0.1 to 0.13 next day return, highlighted in blue. Within the blue region, V-shaped linear trend is manifested in both positive and negative regions of next day return, highlighted in orange and yellow. Both returns in x and y axes are expressed in a log scale.

Figure 6 shows a linear trend between net deviation and next day return in the range of -0.05 to 0.1 next day return value. Within the linear range highlighted in blue, as net deviation increases, next day return decreases. Increase in net deviation signifies the stock behavior's transition from a convex to concave stock. Indeed, as the net deviation value becomes positive, next day return accordingly becomes negative. A different trend is found outside of the blue linear region. For both upper and lower non-linear regions highlighted in green, net deviation either increases or decreases as the absolute magnitude of next day return increases. These trends address the random nature of a stock. Stocks with large magnitude of next day return of either sign is likely to be incorporated with large degree of net deviation, regardless of sign of deviation.

Figure 7 also shows a linear trend between net return of 5 day sampling window and next day return. Unlike a simple linear trend shown in figure 6 in the linear region, the linear region show a combination of linear trends for the relationship between the two returns. Namely, for both positive and negative values of next day return, a V-shaped relationship is manifested in the same direction. Both region's V-shaped relationship accounts for a stock's randomness in behavior and limitation of slope as an indicant for stock's future behavior. Net return could be interpreted as the slope, a first-order derivative of stock's behavior. As shown in the orange and yellow regions' V-shaped relationships, slope does not directly have a correlation with next day return as much as net deviation does. The second order derivative represented by net deviation is therefore more valuable, and slope merely serves as a supplementary indicant to interpret a stock's behavior.

To interpret stocks' behavior based on these indicants, the two features are plotted with next day return as Z-axis, shown in figure 8. The region of linear trend in figure 8 is analogous to those of figure 6 and 7, in the range of -0.1 to 0.1 next day return. For the purpose of long position investment, the region of interest is that with maximal next day return values while maintaining a linear relationship with net return and net deviation. This region is marked by the blue lines in figure 8, namely between 0.1 and 0.075 next day return values. Since this region maintains a linear relationship, predicting next day return within this region from net return and net deviation values is statistically possible.

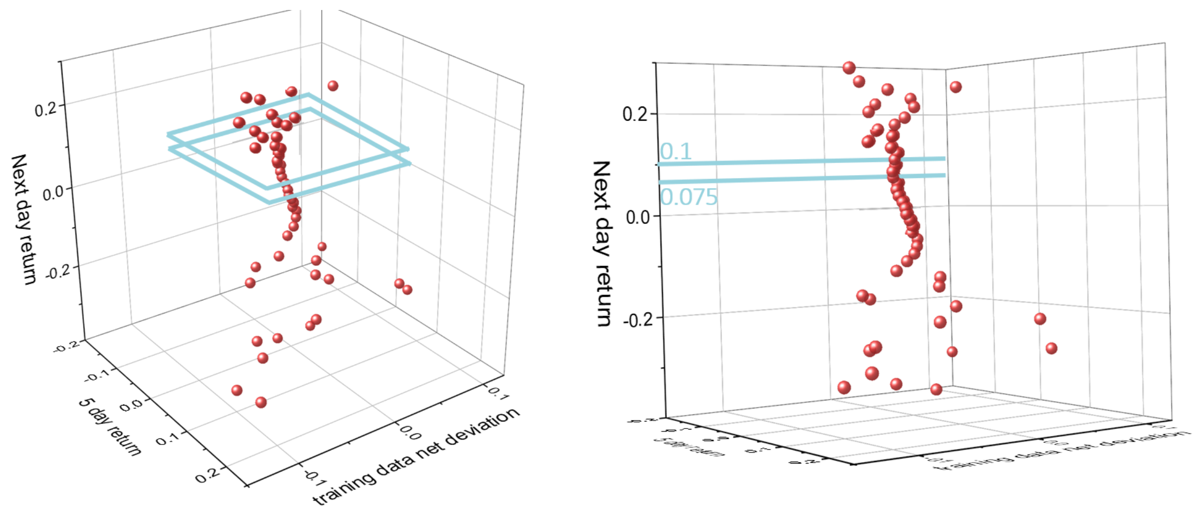


Figure 8. 3-D representation of correlation between net return (5 day return), net deviation, and next day return. Parameters for augmented momentum investment is sampled from data with next day return values between 0.1 and 0.075, depicted by blue lines.

Figure 9 shows the distribution of net deviation and net return for all data with next day return value between 0.075 and 0.1. From the 2-D distribution, net deviation and net return parameters ranges are chosen as parameters for augmented momentum investment model. Specifically, instead of finding assets with highest criterion explained in section 2.B from all stocks with any behavior, we first filter the stocks using net deviation and net return parameters. If we only use stocks from the red sampling area in figure 9, they are likely to result in positive next day return since they are the empirical parameters of stocks that resulted in next day return value of 0.075 ~ 0.1. The red sampling area is determined by mean ± 0.5 standard deviations for each parameter. The average value is acquired as 0.0085 for net return and -0.0097 for net deviation, which models a convex behavior. Standard deviation is acquired as 0.1193 and 0.1327, respectively.

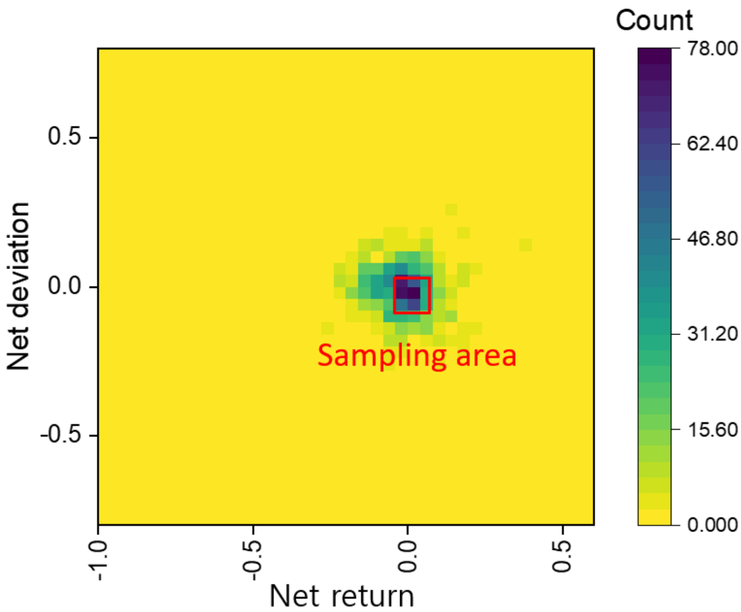


Figure 9. Region for net deviation and net return parameter sampling

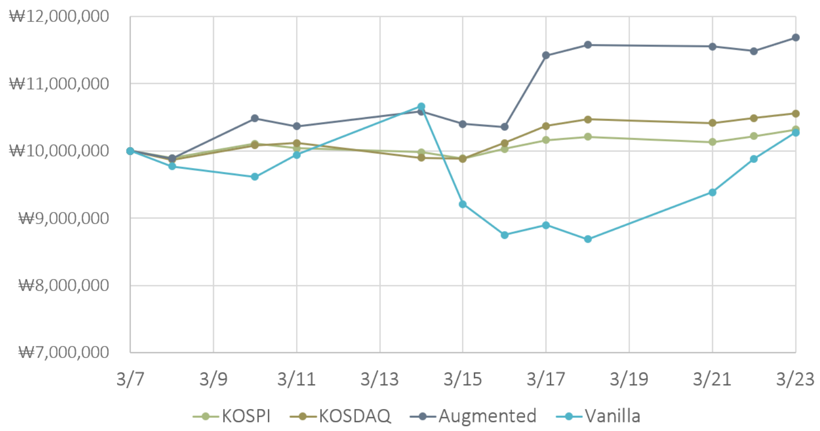
Table 1 shows the average next day return and its standard deviation of filtered stocks from the red sampling area. The average next day return and standard deviation is calculated from the same dataset used for initial calculation of next day return, net return, and net deviation. With such crude filtering, the average return is about 0.002 greater than that of the market, along with doubled standard deviation. We then apply the vanilla momentum investment model to these stocks to construct a portfolio of 10 stocks with maximal return along with minimal volatility. The same criterion of $\frac{\ln(Return)}{Standard\ deviation^{0.2}}$ is used. Selecting maximal criterion from this filtered pool of stocks would provide a much statistically sound portfolio compared to selecting from all stocks, as shown in table 1.

Table 1. Expected return for a single day and its standard deviation. These are sampled from January 1st to April 31st of 2022.

	KOSPI	KOSDAQ	Filtered stocks
Average return	-0.0013	-0.0017	0.0008
Standard deviation	0.0012	0.0167	0.0291

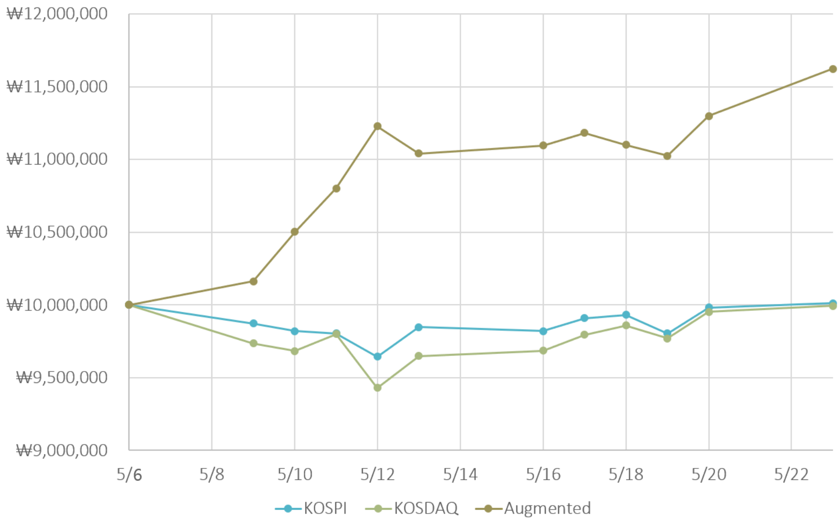
3.C. Augmented momentum investment model performance

Following figures show the augmented model's performance for 4 different timeframes and its respective sharpe ratios for KOSPI, KOSDAQ, and augmented model. Each figure is in the timeframe of 10 ~ 15 working days. As in the vanilla model, the calculated portfolio is held for one day. It is assumed that the portfolio is bought at close price of day n and sold at close price of day n+1, with simultaneous purchase of new portfolio at close price of day n+1. This process is repeated throughout the entire timeframe.



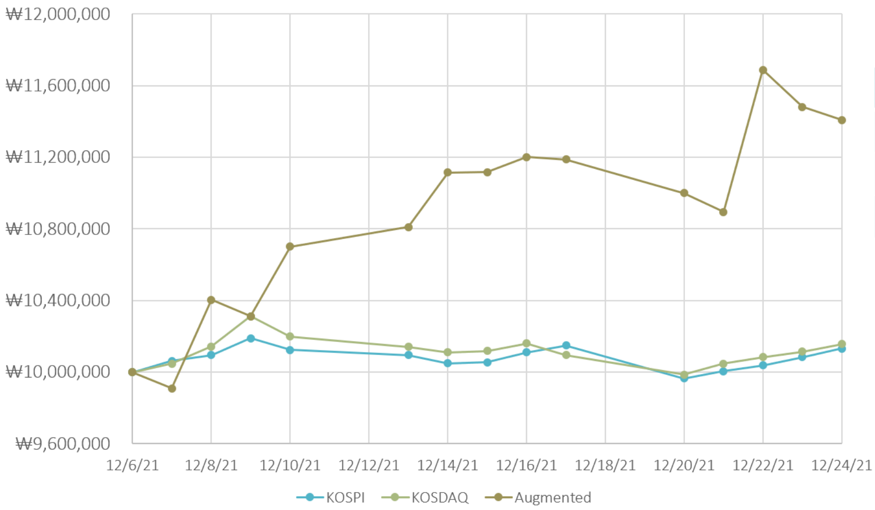
Portfolio	Sharpe ratio
KOSPI	0.248
KOSDAQ	0.331
Vanilla	0.038
Augmented	0.403

Figure 10. Augmented model's performance between 3/7/22 and 3/23/22.



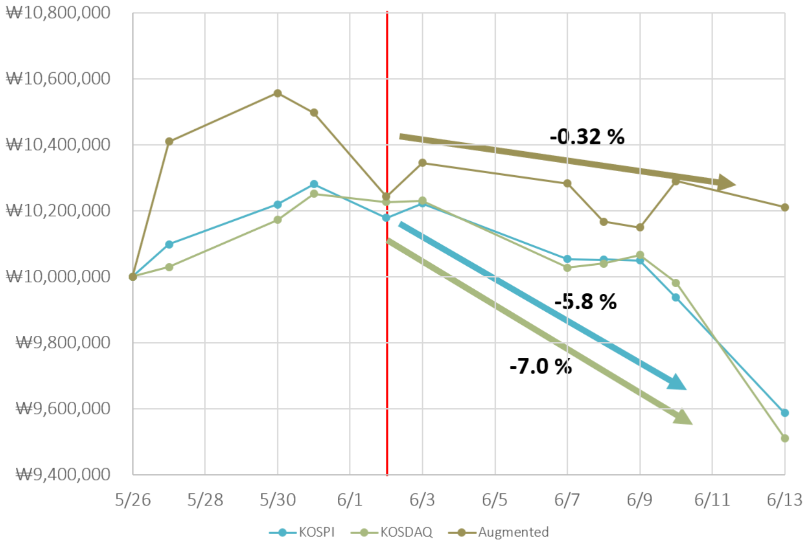
Portfolio	Sharpe ratio
KOSPI	0.008
KOSDAQ	-0.003
Augmented	0.736

Figure 10. Augmented model's performance between 5/6/22 and 5/22/22.



Portfolio	Sharpe ratio
KOSPI	0.133
KOSDAQ	0.143
Augmented	0.352

Figure 10. Augmented model's performance between 12/6/21 and 12/24/21.



Portfolio	Sharpe ratio
KOSPI	-0.290
KOSDAQ	-0.282
Augmented	0.116

Figure 10. Augmented model's performance between 5/26/22 and 6/13/22. Note that the model has also shown competence in hedging.

4. Conclusion

In this work, we propose a statistical analysis-based momentum investing strategy in timeframe of days. The model has been tested for 47 working days in 4 separate timeframes. Sharpe ratio of our model has considerably exceeded the sharpe ratio of market for all timeframes. Despite its supreme performance for all kinds of market behaviors, it is currently distant from practical application due to daily portfolio update. While the algorithm requires refinement to consider transaction fee by reducing unnecessary portfolio updates, we have successfully shown the feasibility and excellent competence of a momentum investment strategy in a daily timeframe with augmentation of statistical analysis. As a future work, the entire pipeline of the proposed methodology will be automated and optimized using machine learning.

5. References

[i] Grundy, Bruce D., and J. Spencer Martin Martin. "Understanding the nature of the risks and the source of the rewards to momentum investing." *The Review of Financial Studies* 14.1 (2001): 29-78.

[ii] Barroso, Lúcia P., Denise A. Botter, and Gauss M. Cordeiro. "Second-order covariance matrix formula for heteroskedastic generalized linear models." *Communications in Statistics-Theory and Methods* 42.9 (2013): 1618-1627.