HEC MONTRÉAL

Cross-Sectional Momentum Return and Crash Risk

par

Jingjing Zhang

Geneviève Gauthier
HEC Montréal
Directrice de recherche

Sciences de la gestion
(Spécialisation M.Sc. Financial Engineering)

Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences
(M. Sc.)

December 2020

Cross-Sectional Momentum Return and Crash Risk

Abstract

Although momentum is robust and persistent for its high return and Sharpe ratio, it is negatively skewed with excess kurtosis and occasional crashes. These crashes happen with high ex-ante volatility and market rally following a downturn. A dynamic strategy significantly improves the performance of benchmark momentum and reduces crash risk, by adding the optimal weight on momentum excess return based on estimating its conditional expected return and variance with a scaling factor chosen such that the annualized volatility of in-sample dynamic return is the same as that of in-sample momentum excess return. Besides, using US stock data and option data, this thesis documents a dynamic option information strategy that includes the stock price-jump information offered by implied volatility difference between out-of-the-money options and at-the-money options in enhancing the momentum performance. Our result shows that the implied volatility difference in level and the implied volatility relative difference in percentage have a certain prediction power on future returns. In addition, this strategy brings about a certain improvement on benchmark momentum with appropriate stop trading signals and helps to reduce the depth of drop during out-of-sample crisis to a certain extent.

keywords: Momentum; crashes; unexpected jumps; implied volatility of option.

Résumé

Bien que le momentum soit robuste et persistant pour son rendement élevé et son ratio Sharpe, il est négativement biaisé avec un kurtosis excessif et des crashes occasionnels. Ces crashes se produisent avec une forte volatilité ex ante et un rebond du marché après une récession. Une stratégie dynamique améliore considérablement les performances du benchmark du momentum et la réduction du risque de crash, en ajoutant le poids optimal sur le rendement excédentaire du momentum basé sur l'estimation de son rendement espéré conditionnel et de sa variance avec un facteur d'échelle choisi de telle sorte que la volatilité annualisée du rendement dynamique dans l'échantillon soit la même que celle du rendement excédentaire du momentum dans l'échantillon. En outre, en utilisant les données des actions américaines et les données des options, cette thèse documente une stratégie d'information dynamique des options qui comprend les informations sur le saut de prix des actions offertes par la différence de volatilité implicite entre les options hors-de-monnaie et les options à-la-monnaie dans le but d'améliorer la performance dynamique. Notre résultat montre que la différence de niveau de volatilité implicite et la différence relative en pourcentage de volatilité implicite ont un certain pouvoir de prédiction sur les rendements futurs. En outre, cette stratégie apporte une certaine amélioration sur le benchmark du momentum avec des signaux d'arrêt de négociation appropriés et contribue à réduire dans une certaine mesure la profondeur de la baisse pendant la crise hors échantillon.

Mots-clés: Momentum; crashes; sauts inattendus; la volatilité implicite d'option.

Acknowledgement

First and foremost, I would like to express my deepest gratitude and appreciation to my supervisor Professor Geneviève Gauthier for her consistent support and guidance during my master study and research process. Her patience, enthusiasm, encouragement, broad knowledge and professional attitude help me a lot for my research and writing this thesis. Without her wise guidance, persistent help and advice, this thesis could not have reached its goal, especially under the current situation we work and study remotely.

Besides my supervisor, I would like to thank the rest of my thesis evaluation committee for their time, review and examination of this thesis.

I wish to offer my sincere thanks to Mr. Mohamed Jabir of the LACED at the HEC for his technical support and his help and guidance on using the WRDS database and data collection. He patiently answered my questions related to collect data and provided technical support on my coding and programming.

Fin-ML CREATE program is also truly appreciated. Without their high-quality machine learning training and funding, this thesis would have not been the same.

In addition, I would like to thank the following schoolmates and friends for their help and their encouragement: Mengna Wang, Rémi Galarneau-Vincent, Samuel Léveillé, Oussama Charifi, Peike Zhou, Qi Guo, Chendi Ni, Shen Li, Jiayu Li and Zehua Huang. They kept me going on, and, without their input, this work would not have been possible.

Last but not least, I would like to thank my parents, Xuhong Zhang and Wenjun Zhang, for their constant care and support throughout my life.

1 Introduction

Momentum strategy is widely applied owing to its robustness and persistence on strong return and high Sharpe ratio. It bets on the autocorrelation of equity returns by buying recent winner stocks and selling recent loser stocks; that is, past return has strong impact on future returns. The study of momentum strategy in this paper is on cross-sectional momentum, and the momentum portfolio is cap-weighted. Compared to time-series momentum, which assigns stocks into winner portfolio and loser portfolio on the basis of their absolute performance over the formation period, cross-sectional momentum assigns stocks based on their relative performance (Jegadeesh & Titman, 1993). Therefore, cross-sectional momentum doesn't have a timing element in stock selection since the number of stocks that are assigned into each portfolio will not be affected by market performance. In contrast, time-series momentum assigns more stocks into the winner portfolio in the market with strong performance and assigns more stocks into the loser portfolio in the market with weak performance (Bird, Gao and Yeung, 2017). Compared to equal-weighted momentum portfolio, cap-weighted portfolio allocates different capital to stocks according to their capitalization and alters the weights on stocks as their prices and numbers of shares outstanding changes, which introduces higher liquidity and higher Sharpe ratio in the portfolio.

Momentum strategy has an average excess return that is twice higher than that of market and a higher Sharpe ratio. The stock selection is based on the cumulative returns of each stock in the formation period (from 12 months to 1 month before). The past winners are stocks with the top 10% past cumulative return in the formation period, while the past losers are stocks with the bottom 10%. In this paper, the empirical research on sample data from 1987 to 2019 shows the high returns and robustness of momentum strategy. However, it suffers from negative skewness, excess kurtosis and occasional crashes, which is consistent with the results of Mahdi Heidar (2015). The momentum return drops by 45% in the two months ending 2002 and 73% in last two months of 2009. By investigating ex-ante bear market indicators, annualized market volatility, and annualized momentum volatility, we find that these crashes happen when the market rebounds and the ex-ante volatility is high after overall market downturn. This is consistent with the result of Stivers and Sun (2010), Heidari (2015), and Daniel and Moskowitz (2016). More precisely, when the market rebounds from previous large downturns, past loser stocks that declined dramatically previously achieve higher returns than past winner stocks; this results in negative return in momentum strategy. During the crash of 2009, the return of past loser portfolio increased by 163% from March to May, while the return of the past winner portfolio only rose by 8%.

To improve the performance of the momentum portfolio, especially during crashes, much academic research addresses the potential predictability of momentum return and crashes from three directions: time-varying betas (Kothari and Shanken (1992), Grundy and Martin (2001)); market illiquidity (Avramov, Cheng and Hameed (2016)); and market volatility and momentum volatility (Barroso and Santa-Clara (2015) and Daniel and Moskowitz (2016)). Daniel and Moskowitz (2016) apply a dynamic strategy which maximizes the unconditional Sharpe ratio of the momentum portfolio to improve the performance of static momentum portfolio. The benchmark momentum portfolio is composed of a position in winner-minus-loser portfolio (WML) and a position in bank account, while the dynamic strategy proposed by Daniel and Moskowitz (2016) adjusts the weights on the WML portfolio dynamically and generates a Sharpe ratio that is more than double. Based on their idea, the optimal weight in this paper is proportional to the conditional mean of momentum excess return and inversely proportional to the conditional variance, with a scale factor α that makes the annualized volatility of in-sample dynamic strategy return equal to that of in-sample momentum excess return. From the result, the dynamic strategy significantly outperforms the benchmark momentum strategy. We also find that the weights are high in bull markets and relatively low in bear markets, and they are negative when the market rallies fast after stagnation; in other words, the dynamic strategy longs past loser and shorts past winner when the market experiences a rapid rebound.

In addition, capturing unexpected jumps in stock prices should be beneficial to predicting future return. In behavioral theory, Jiang and Yao (2013) and Daniel, Hirshleifer, and Subrahmanyam (1998) advocate that unexpected shocks are important components in forecasting future stock returns. We consider including jump information in stock prices as a predictor of momentum return in the near future. Theoretically, the option market contains the information about future underlying asset performance. Compared to long-maturity options, short-term options have a more intimate link between the pricing of options close to maturity and the state of the underlying asset return process (Andersen, Fusari and Todorov, 2018). The short-term deep out-of-the-money (OTM) options reflect jump risk since their price should be close to zero unless market participants anticipate that there will be jumps in the near future. Short-term at-the-money (ATM) options reflect current underlying stock volatility since their price trends are close to the underlying price trend. As the level of underlying and option prices changed a lot from 1996 to 2019, we consider the option implied volatility difference in level (IVD, the difference between implied volatility of short-term OTM and ATM options) and the implied volatility relative difference in percentage (RIVD, which equals IVD divided by implied volatility of ATM options) as the proxy of jump information to distinguish between jump and high volatility of underlying asset.

We consider Generalized Linear Models on extreme values of IVD and RIVD, and Nonlinear Machine Learning Models, Random Forest and Extreme Gradient Boosting, to test the predictive power of option IVD and RIVD on future return. The result shows that the IVD and RIVD data have a certain prediction effect on future returns and the effects are different across stocks. We constructed a dynamic strategy based on option market information that stops trading stocks in the past winner portfolio when their put option RIVD exceeds their threshold x, and stops trading stocks in the past loser portfolio when their call option RIVD exceeds their threshold y. By observing in-sample option data, the RIVD data is used to set unequal values to thresholds x and y that are also different across stocks. We find that the performance of benchmark momentum portfolio is improved by applying this strategy. However, it does not help to enhance the achievement of dynamic momentum portfolio. Applying dynamic thresholds x and y that are updated every six months further improves the performance of benchmark momentum a bit. In addition, by constructing this option information strategy on the subsample S&P500 momentum portfolio, we find that it helps to reduce the depth of fall during out-of-sample crash to a certain extent.

The rest of the thesis is organized as follows: Section 2 reviews some of the most relevant research and literature; Section 3 describes stock data, option data and data construction methodology; Section 4 sets out the portfolio construction, momentum performance analysis and comparison; Section 5 presents the process and technique of optimal dynamic strategy based on the idea of Daniel and Moskowitz (2016), the approach to estimate conditional excess return and conditional variance, and the results and analysis of dynamic weighting strategy; Section 6 introduces a dynamic strategy based on option market information (We explain the methodology of this strategy, and apply Generalized Linear Models Nonlinear Machine Learning Models to test the predictive effect of option IVD and RIVD data on future return; we also investigate option data to set the stop trading signals, and evaluate the performance of this strategy); Section 7 contains the conclusion and discussion of this thesis, followed by the references and the appendices.

2 Literature Review

Jegadeesh and Titman (1993) first find that past winners outperform past losers in the US stock market. The theoretical basis behind momentum mainly falls into three directions: behavioral biases, risk premium, and trading friction. Clifford S. Asness and al. (2014) elaborated the first two directions. In behavioral theory, momentum is a phenomenon where investors underreact or overreact to information and price changes for a variety of reasons (ex. investors chase return and tend to sell winners fast and hold onto losers longer.) In risk premium theory, momentum premium compensates economic risk with two different views: some risk-based models repute that economic risks have different effects on the investment and growth rates of different firms, thus affecting long-term cash flows and dividends which generate momentum patterns (Sagi and Seasholes, 2007); others deem that there is a shared economic risk indicated by consistent value, momentum return premia, and correlation structure across markets and asset classes (Asness, Moskowitz, and Pedersen, 2013). Avramov, Cheng and Hameed (2013) provide the trading-based explanation that momentum profitability crucially depends on market liquidity, that is, high liquidity results in strong momentum effect.

To forecast the future momentum return, time-varying beta is one of the highly debated predictors. Kothari and Shanken (1992) find that the returns of portfolios with stock selection based on past performance are significantly exposed to time-varying systematic factors. Applying this result to momentum portfolio, Grundy and Martin (2001) repute that momentum crashes are closely related to time variation in betas. That is, when the stock market return exceeds Treasury bills return, the stocks that are selected to past winner portfolio will have a beta greater than one and the past loser portfolio will contain stocks with beta less than one. This, in turn, will generate a positive beta for momentum. However, following a down market, the momentum portfolio is seemly to long stocks of low-beta firms (past winners) and shorts stocks of high-beta firms (past loser), which generates negative betas and causes momentum crash. Daniel and Moskowitz (2016) verify Grundy and Martin's idea empirically to find a strong bias in their hedging strategy based on forward-looking betas and to show that a hedging strategy based on ex-ante betas does not improve the momentum performance. Apart from using time variation in betas as a predictor, Avramov, Cheng and Hameed (2014) conclude that there is a robust negative relation between momentum profits and aggregate market illiquidity. Barroso and Santa-Clara (2015) exhibit improvement on momentum performance with scaling the WML by its realized volatility in the previous six months.

Based on the fact that momentum volatility is itself predictable, Daniel and Moskowitz (2016) design a dynamic strategy that adds a dynamic optimal weight on momentum excess return to maximize the unconditional Sharpe ratio of WML. They prove that the optimal weight at time t should be proportional to the conditional expected return of momentum excess return and inversely proportional to the conditional variance, with a coefficient λ. Daniel and Moskowitz (2016) illustrate that "λ is chosen so that the in-sample annualized volatility of the strategy is 19%, the same as that of the CRSP value-weighted index over the full sample" (p.233). However, it's not possible to implement this value by using annualized market volatility (CRSP value-weighted index in Daniel and Moskowitz (2016)) over the full sample, which includes future information, to calculate scale factor. We find that a coefficient α, which makes the annualized volatility of in-sample dynamic strategy return equal to that of in-sample momentum return, significantly enhances the momentum performance.

Jiang and Yao (2013) advocated that investors may have biased expectations about firms' future values that relate to firms' characteristics, especially in turbulent markets. The biases will result in stocks mispricing, which is subsequently reversed by updated news and new information. Jumps, which are large discontinuous changes in stock prices, are the proxy of significant information shocks. Daniel, Hirshleifer, and Subrahmanyam (1998) also illustrated that investors are overconfident in their private expectations of future stock return based on current information. When new information appears, investors will react asymmetrically to confirming news versus disconfirming news. To capture the jumps, Andersen, Fusari, and Todorov (2016) find that weekly options provide information to acquire or lay off exposure to diffusive and jump price risks since they are highly sensitive to the temporary shifts in underlying stock price. More precisely, when the time-to-maturity is short, the expected volatility and jump intensity will not change much over the remaining life of the option. Therefore, the short-term deep OTM options behave as a pure jump process, and the short-term ATM options reflect mainly diffusion risk, i.e. current underlying stock volatility.

3 Data Description

In this section, the stock data and option data that are used in the emprical analysis are illustrated and described.

3.1 Stock Data

The monthly data and daily data of stock price, return with dividend and number of shares outstanding are obtained from the Center for Research in Security Prices (CRSP) database on Wharton Research Data Services (WRDS).[1] The time period covered is from Jan. 1986 to Dec. 2019. All common shares (CRSP Share Code 10 and 11)[2] listed on the NYSE, AMEX and Nasdaq (CRSP Exchange Code 1, 2 and 3)[3] are included, 19397 stocks in total. All the prices are close price and are adjusted to stock events such as mergers, cash dividends, stock dividends or splits, total liquidations, and delisting by using adjusted factor offered on CRSP, in order to make an equivalent basis for prices before and after distribution. The number of shares outstanding were also adjusted correspondingly.

Let S_t be the adjusted post-dividend close price at the end of month t. The monthly simple return for stock i is

$$R_{i,t} = \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}} = \frac{S_{i,t}}{S_{i,t-1}} - 1. \tag{1}$$

Figure 1 reports the monthly return data statistics of 10%, 25%, median, 75% and 90% across individual stocks over time. As shown in the graph, the extreme values occur in turbulent market in 2001 and the end of 2008.

3.2 Option Data

The daily option implied volatility, delta and time to maturity of stocks are obtained from the Option Metrics database on Wharton Research Data Services (WRDS). The time period is from Jan. 1996 to Dec. 2019, 6042 days in total. The underlying assets for these options are the stocks in the previous stock dataset. To select the short-term options, let's assume that the option days to expiration is between 3 days and 16 days and the option volume ≥ 10. Delta measures the sensitivity of option price to underlying asset price changes.[4] The

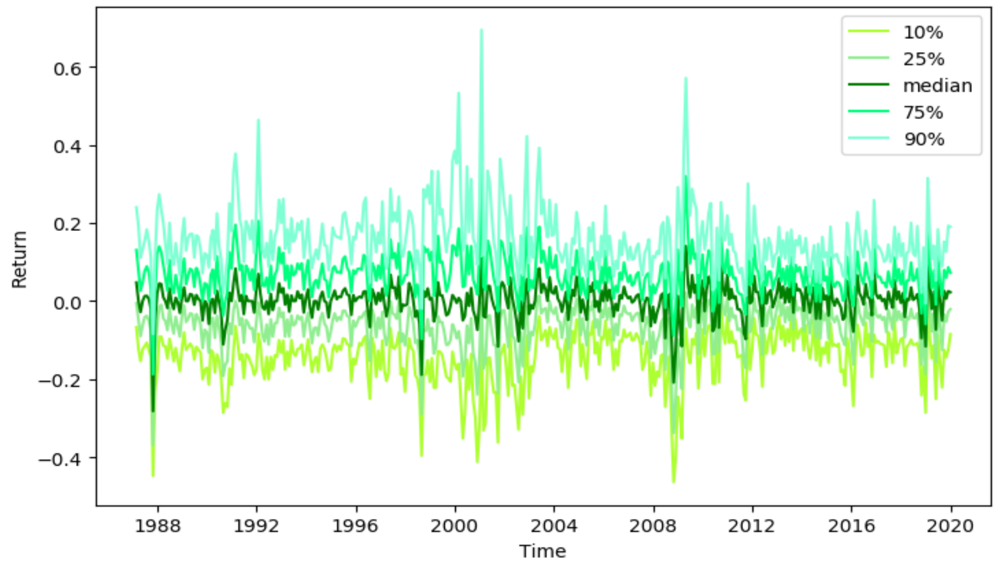
[1] https://wrds-www.wharton.upenn.edu/.

[2] Share Code (SHCD): A two-digit code describing the type of shares traded. Share Code 10 and 11 represent ordinary common shares securities which have not been further defined or need not be further defined.

[3] Exchange Code (EXCHCD) is a code indicating the exchange on which a security is listed. Exchange Code 1, 2 and 3 represent for securities listed on New York Stock Exchange (NYSE), American Stock Exchange (AMEX) and The Nasdaq Stock Market (Nasdaq) respectively.

[4] Delta is calculated by: delta $= \frac{\partial g}{\partial S}$, where g denotes the option price that uses a proprietary pricing algorithm based on the industry-standard Cox-Ross-Rubinstein (CRR) binomial tree model, with the continuously-compounded interest rate r and implied volatility σ that is illustrated in section 6.1.

Figure 1: Stock Data Statistics

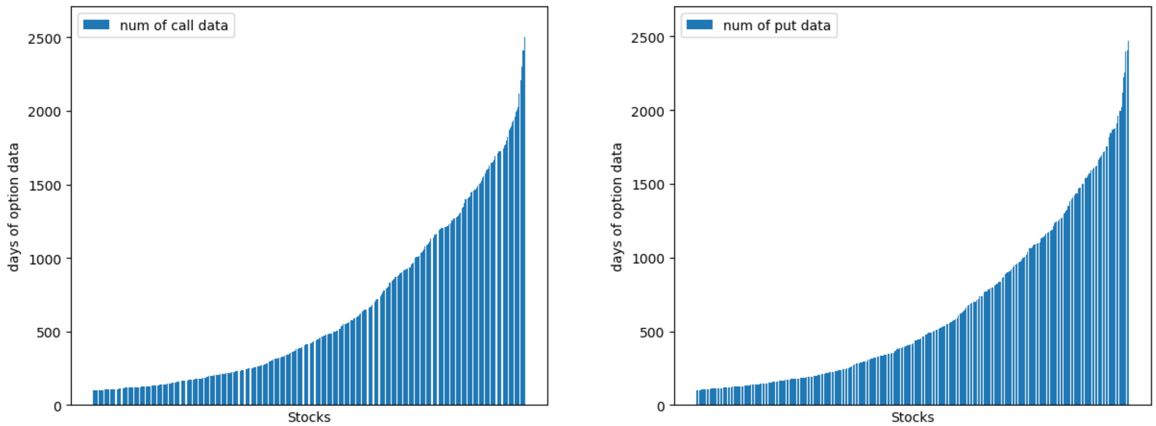


This figure presents the monthly return statistics across stocks. For each month t, the 10%, 25%, median, 75% and 90% of monthly returns of all the stocks with valid data are shown in different colors. The total number of stocks are different across months and are listed in Figure 3.

delta values are positive for call options and negative for put options. Therefore, we use delta values as the criteria to select out-of-the-money (OTM) options and at-the-money (ATM) options; that is, ATM options have absolute deltas around 0.5 and deep OTM options have absolute deltas between 0 and 0.25. For stocks that have many options, the option with highest volume and shortest time-to-maturity is selected.

Under these assumptions and data availability, we collect the short-term deep OTM call option impled volatility (IV) data of 5061 stocks, the short-term ATM call option IV data of 4913 stocks, the short-term deep OTM put option IV data of 4904 stocks, and the short-term ATM put option IV data of 4644 stocks. The implied volatility difference in level (IVD) of call options and put options are given by equation (16) and (17), and the implied volatility relative difference in percentage (RIVD) of call options and put options are given by equation (19) and (20). There are 3631 stocks that have IVD and RIVD of call options, and 828 stocks among them have more than 100 days data. For put options, there are 3314 stocks that have IVD and RIVD of put options, and 739 stocks among them have more than 100 days data. Panel 1 and Panel 2 in Figure 2 report the total numbers of call option data and put option data across stocks. 70% of stocks have less than 1000 call and put option data. The stock of Google Inc. (stock ticker: GOOGL) has the most call (2579 days) and put (2572 days) option data.

Figure 2: Numbers of Option Data



PANEL 1 PANEL 2

Each bar in Panel 1 represents the total number of call option data of a stock, while each bar in Panel 2 presents that of put option data of a stock. In both Panel 1 and Panel 2, the x-axis from left to right are the stocks with least option data to most option data.

4 Momentum Strategy

In this section, we present the momentum portfolio construction and its performance analysis over time period from Feb. 1987 to Dec. 2019.

For each investing month t, all the stocks with valid data[5] are ranked based on their cumulative returns of the formation period from month $t-12$ to month $t-2$.

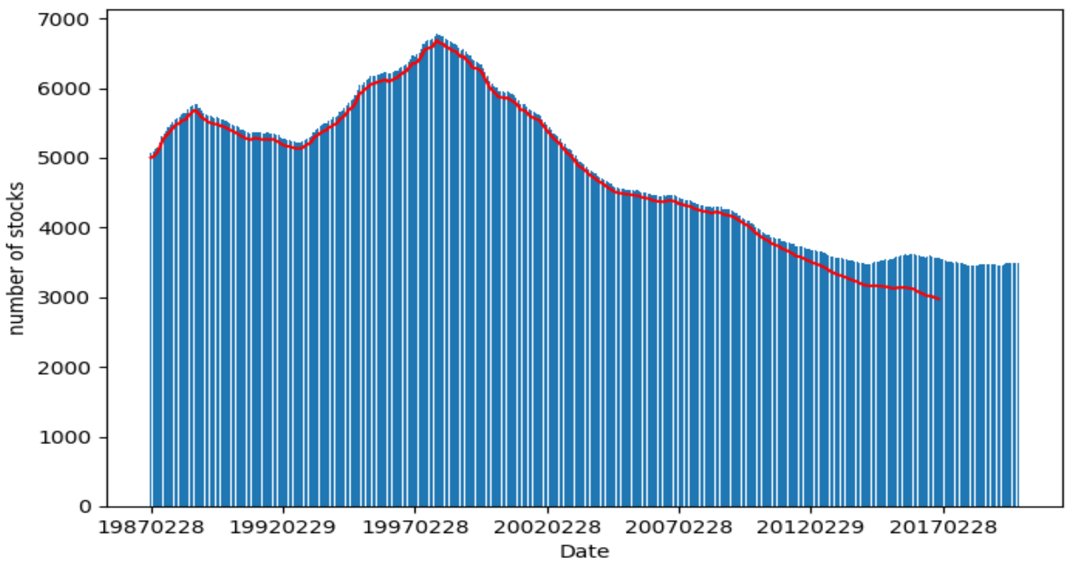
$$C_{i,t-12,t-2} = \frac{S_{i,t-2}}{S_{i,t-12}} - 1 = \prod_{t-11}^{t-2} \frac{S_{i,u}}{S_{i,u-1}} - 1 = \prod_{t-11}^{t-2} (R_{i,u} + 1) - 1. \tag{2}$$

The one-month gap between formation period and the investing month (i.e. month t) is for eliminating the effect of market inefficiency and stock price "overreaction" over the short time intervals (Lehmann, 1990). All the stocks with valid data are then assigned into one of the 10 decile portfolios according to their ranks. Each one of these 10 decile portfolios are cap-weighted. Assume that these 10 decile cap-weighted portfolios keep the same stocks among this month and rebalance at the end of month t with no transaction cost. Compared to Daniel and Moskowitz (2016) (abbreviated as DM (2016) hereafter), there are two differences in stock selection when constructing 10 decile portfolios: firstly, DM (2016) requires that there are at least 8 out of 11 monthly returns for a stock, while this is not a requirement in

[5] It requires that the following data of a stock is not missing: the stock price at month $t-13$ (the formation period is from month $t-12$ to month $t-2$, we need $t-13$ stock price to calculate the return at month $t-12$), the stock return at month $t-2$ and the capitalization at month $t-1$.

our stock selection; secondly, the stock data period of DM (2016) is from Jan. 1927 to Mar. 2013, 1035 months in total; while the stock data period in this paper is from Jan. 1986 to Dec. 2019, 408 months in total. Figure 3 compares the total number of stocks in our 10 decile portfolios (the bar charts) and in DM's 10 decile portfolios (the red curve).

Figure 3: Numbers of Total Stocks Over Time



This figure presents the total amount of stocks with valid data in 10 decile portfolios at each month t. The bar charts shows total number of stocks with valid data in our dataset, and the red curve shows that in the dataset of Daniel and Moskowitz (2016).

Among the 10 decile portfolios, the past winner (portfolio 10) contains stocks with the top 10% past cumulative return, while the past loser (portfolio 1) contains stocks with the bottom 10%. The monthly returns of a cap weighted portfolio is

$$\begin{aligned} R_t^{(p)} &= \frac{V_t^{(p)} - V_{t-1}^{(p)}}{V_{t-1}^{(p)}} = \frac{\sum_{i \in \mathcal{P}_t} n_{i,t-1} S_{i,t} - \sum_{i \in \mathcal{P}_t} n_{i,t-1} S_{i,t-1}}{V_{t-1}^{(p)}} \\ &= \sum_{i \in \mathcal{P}_t} \frac{n_{i,t-1} S_{i,t-1}}{V_{t-1}^{(p)}} \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}} = \sum_{i \in \mathcal{P}_t} w_{i,t-1} R_{i,t}, \end{aligned} \tag{3}$$

where $n_{i,t-1}$ denotes the number of shares outstanding of stock i holding during time period $t-1$ to t, the weight $w_{i,t-1} = \frac{n_{i,t-1} S_{i,t-1}}{V_{t-1}^{(p)}}$ is the proportion of the portfolio invested in stock i, and $\mathcal{P}_t$ is the set of stocks that belong to the portfolio at month t.

4.1 The Market Return

The market return is the cap-weighted return of all listed firms with valid data in CRSP:

$$R_t^{\mathrm{Mkt}} = \frac{V_t^{\mathrm{Mkt}} - V_{t-1}^{\mathrm{Mkt}}}{V_{t-1}^{\mathrm{Mkt}}} = \sum_{i \in \mathcal{M}_t} \frac{n_{i,t-1} S_{i,t-1}}{V_{t-1}^{\mathrm{Mkt}}} R_{i,t}, \tag{4}$$

where $\mathcal{M}_t$ is the set of all stocks available at month t.

And the cumulative market returns over the last T months is

$$C_{t-T,t}^{\mathrm{Mkt}} = \frac{V_t^{\mathrm{Mkt}}}{V_{t-T}^{\mathrm{Mkt}}} - 1 = \prod\nolimits_{u=t-T+1}^{t} \frac{V_u^{\mathrm{Mkt}}}{V_{u-1}^{\mathrm{Mkt}}} - 1 = \prod\nolimits_{u=t-T+1}^{t} R_u^{\mathrm{Mkt}} - 1. \tag{5}$$

4.2 The Risk-free Rate

As DM (2016) do, the risk-free rate series $R^{(f)}$ is the one-month Treasury bill rate. The monthly risk-free rate data is obtained from Ken French's data library.[6].

4.3 The Momentum Strategy

At the end of month $t-1$, the value of the momentum investment strategy is V_{t-1}^{Mom}. As the calculation methods of the strategy value at the end of month t are different when V_{t-1}^{Mom} is positive and negative, they are discussed separately in section 4.3.1 and 4.3.2.

4.3.1 If $V_{t-1}^{\mathbf{Mom}} > 0$

At the beginning of month t, there is a long position in the winners' portfolio for an amount of V_{t-1}^{Mom}, a short position of $-V_{t-1}^{\mathrm{Mom}}$ in the losers' portfolio and an investment of V_{t-1}^{Mom} in the risk-free asset for a total investment of V_{t-1}^{Mom}. Therefore, the strategy is self-financing.

At the end of the month, the strategy value is

$$V_t^{\mathrm{Mom}} = V_{t-1}^{\mathrm{Mom}} \left(1 + R_t^{(10)}\right) - V_{t-1}^{\mathrm{Mom}} \left(1 + R_t^{(1)}\right) + V_{t-1}^{\mathrm{Mom}} \left(1 + R_t^{(f)}\right).$$

The monthly return of this long-short strategy is

$$R_t^{\mathrm{Mom}} = \frac{V_t^{\mathrm{Mom}} - V_{t-1}^{\mathrm{Mom}}}{V_{t-1}^{\mathrm{Mom}}} = R_t^{(10)} - R_t^{(1)} + R_t^{(f)}. \tag{6}$$

[6] $https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/$

4.3.2 If $V_{t-1}^{\text{Mom}} \leq 0$

At the beginning of month t, there is a long position in the winners' portfolio for an amount of $-V_{t-1}^{\text{Mom}}$, short position of V_{t-1}^{Mom} in the losers' portfolio, and a loan of V_{t-1}^{Mom} in the risk-free asset for a total investment of V_{t-1}^{Mom}. Therefore, the strategy is self-financing.

At the end of the month, the strategy value is

$$V_t^{\text{Mom}} = -V_{t-1}^{\text{Mom}} \left(1 + R_t^{(10)}\right) + V_{t-1}^{\text{Mom}} \left(1 + R_t^{(1)}\right) + V_{t-1}^{\text{Mom}} \left(1 + R_t^{(f)}\right).$$

4.3.3 General Case

$$V_t^{\text{Mom}} = \left|V_{t-1}^{\text{Mom}}\right| \left(1 + R_t^{(10)}\right) - \left|V_{t-1}^{\text{Mom}}\right| \left(1 + R_t^{(1)}\right) + V_{t-1}^{\text{Mom}} \left(1 + R_t^{(f)}\right).$$

4.3.4 Cumulative Returns

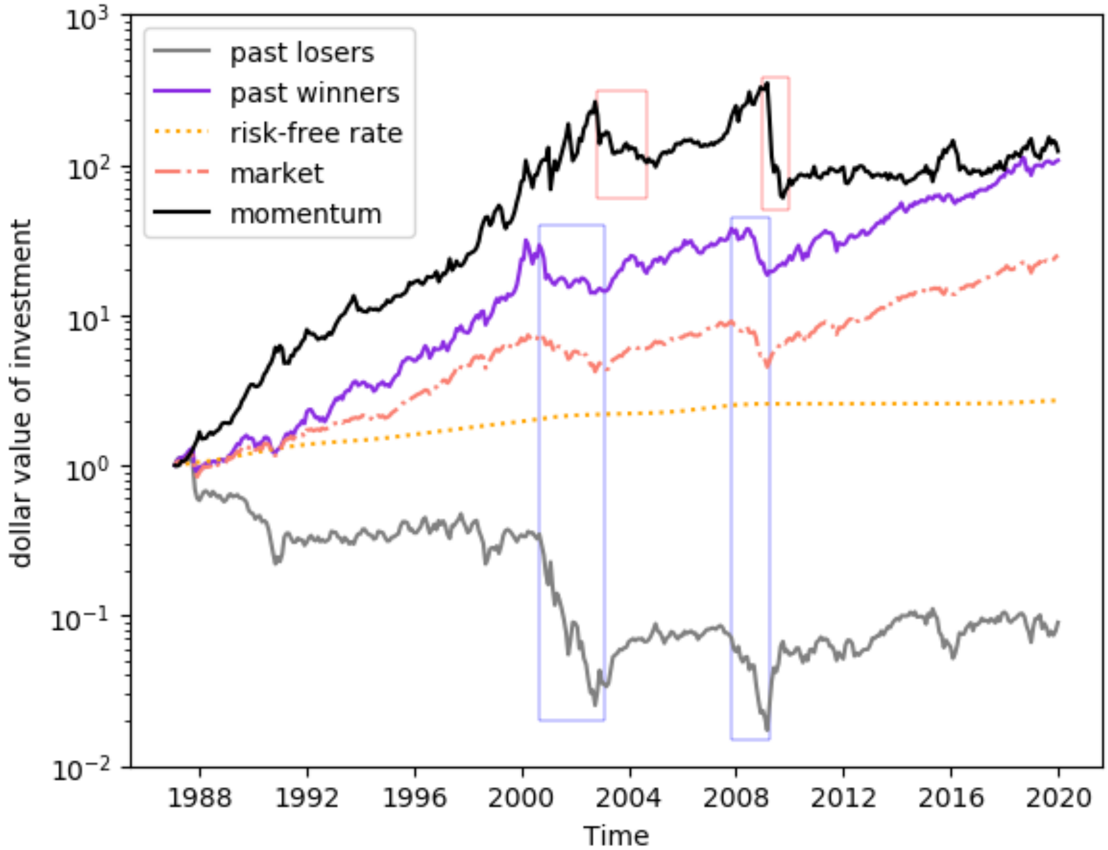
The cumulative returns over T months is

$$C_T^{\text{Mom}} = \frac{V_T^{\text{Mom}}}{V_0^{\text{Mom}}} - 1.$$

To illustrate the performance of momentum strategy, Figure 4 compares the cumulative returns of these five portfolios: past loser, past winner, risk-free rate, market, and momentum from Feb. 1987 to Dec. 2019. We observe that the past winner portfolio not only outperform the past loser portfolio significantly over the whole time period, but also outperform the market, and past winner and past loser almost drop at the same time when market goes down. The momentum portfolio has a large return most of time, but there has been a slight drop during 2003-2004 and a significant drop during 2008-2009. The decline periods of the momentum portfolio are slightly delayed compared to those of past winner and past loser, since momentum crashes happen mostly when the market rebounds. Table 1 presents the statistical analysis of past loser return, past winner return, momentum return and market return. From the results, we can see that the average momentum excess return is twice as high as the average market excess return. The momentum returns are more volatile with higher Sharpe ratio, and are negatively skewed with excess kurtosis, which is consistent with the analysis of Heidari (2015) and DM (2016). Figure 5 reports the daily return of momentum strategy. As it shows, the momentum return fluctuates more dramatically around the same time period when the cumulative momentum return drops.

From further analysis of these two crash periods, which both occur when the market rebounds after a overall downturn, Figure 6 zooms in the two periods framed by vertical bars in Figure 5. The stock market declined rapidly in 2001 after a decade-long bull market where many stocks were overvalued. After the 'September 11 attacks', the market recovered from the bottom but took a sharp downturn again starting from March 2002, triggered by accounting and corporate scandals and fears of deflation. Equity markets rebounded at the

Figure 4: Cumulative Return of Portfolios



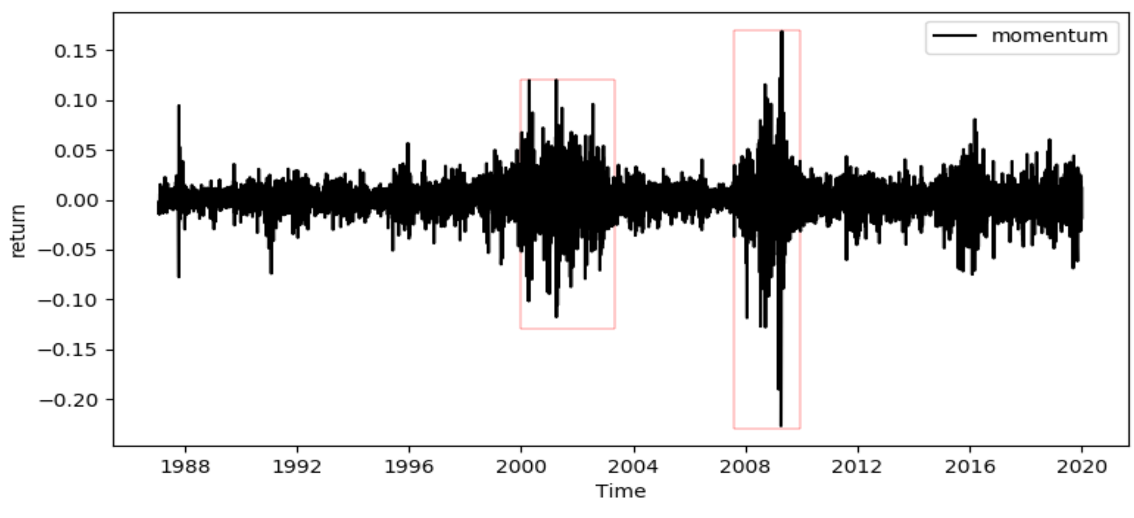
Assume that 1$ was invested on Feb. 1987. This figure compares the cumulative returns of past loser portfolio (the grey curve), past winner portfolio (the purple curve), market portfolio(the salmon curve), risk-free asset (the orange curve) and momentum strategy (the black curve). On Dec. 2019, the total return of these five portfolios are 0.09$, 107.08$, 24.30$, 2.70$ and 123.50$ respectively. The vertical bars present the drop period of these portfolios.

Table 1: Statistics of Portfolio Returns

This table presents the statistical analysis summary of past loser return, past winner return, momentum return and market return. The statistics include average excess return $\overline{r - r_f}$, return standard deviation σ, annualized Sharpe ratio (SR), skewness of monthly and daily return (skew(m) and skew(d)), and kurtosis of monthly and daily return (kurt(m) and kurt(d)).

Statistics	past loser	past winner	momentum	market
$\overline{r - r_f}$	-0.0031	0.0116	0.0147	0.0066
σ	0.1057	0.0664	0.0949	0.0428
SR	-0.0293	0.1747	0.1549	0.1542
skew(m)	0.5310	-0.1965	-1.1211	-0.9736
skew(d)	0.2931	-0.4286	-0.7768	-0.5722
kurt(m)	2.8714	3.0386	4.7940	2.8807
kurt(d)	12.8030	9.7878	12.2719	16.1291

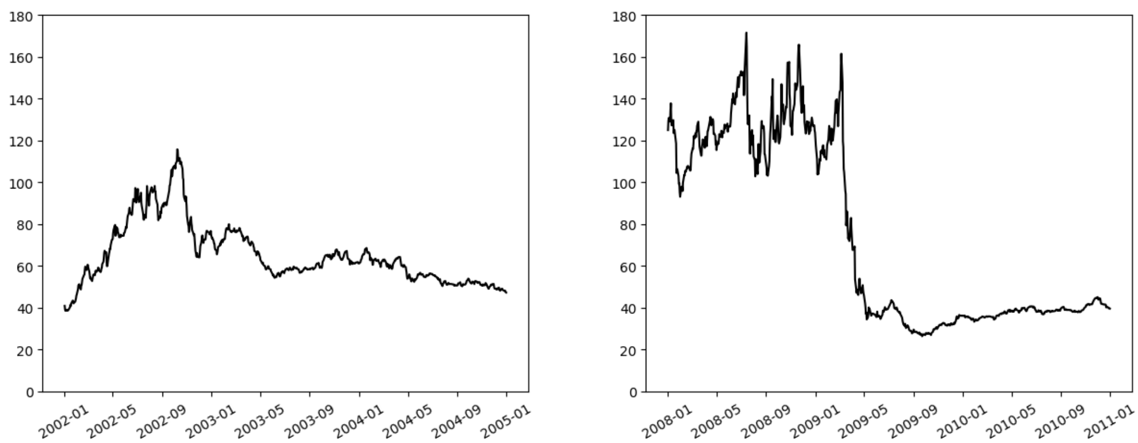
Figure 5: Daily Momentum Return



This graph reports the daily momentum return from Feb. 1987 to Dec. 2019. The two high-volatility periods in momentum returns are 2001-2003 and 2008-2010 when the stock market declines dramatically and revives.

end of 2002 and the momentum strategy dropped by 45% over a two-month period. During 2007-2008, the depreciation of U.S. subprime mortgage market triggered an international banking crisis and evolved into a global financial crisis. On September 29th, 2008, the Dow Jones Industrial Average fell 777.68 points and equity markets crashed. In Panel 2 of Figure 6, the momentum strategy dropped by 73.92% over a two-month period in 2009. These performance features are consistent with the existing literatures (Heidari (2015) and DM (2016)).

Figure 6: Daily Cumulative Momentum Return during Two Crash Periods



PANEL 1 PANEL 2

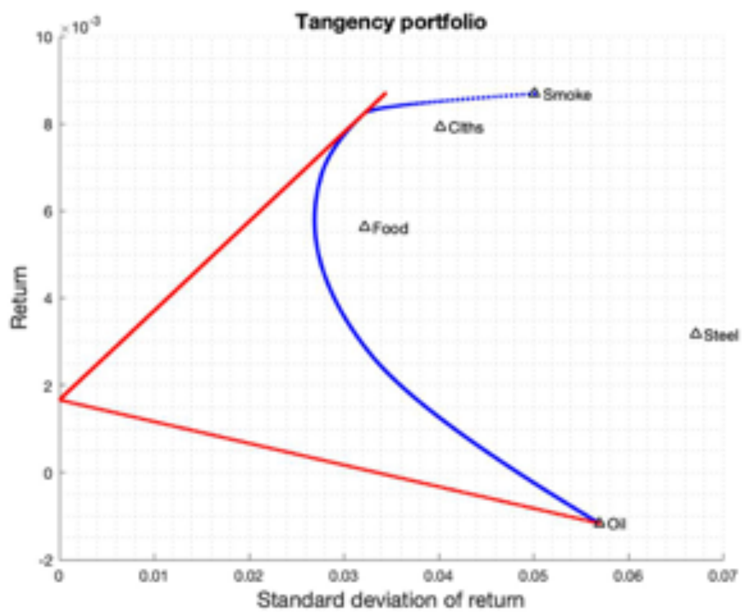
Zooming in the two periods where momentum strategy crashes, we observe that the daily cumulative momentum return declines rapidly in a short time. In the left graph, the cumulative return of momentum portfolio decreases from 115.77$ on October 8th, 2002 to 63.99$ on December 2nd, 2002. In the right graph, it drops from 161.51$ on March 6th, 2009 to 34.29$ on May 8th, 2009.

5 Dynamic Weighting of the Momentum Portfolio

Based on DM's idea, we apply an optimal weight on the momentum excess return and construct a dynamic strategy that maximizes the Sharpe ratio and improves the momentum performance.

The Sharpe ratio measures the risk premium earned per unit of risk. In general, the risk premium is measured by average excess return while risk is measured by standard deviation of return. The mean-variance optimization (Markowitz, 1952) presents the efficient frontier, and each point on the locus represents the optimal portfolio with the best expected return for a given level of risk, which is measured by variance. Simultaneously, we see the optimal portfolio with lowest risk for a given level of expected return. Under the assumption that an investor can invest in a risky asset and a risk-free asset, the optimal portfolio of the risky asset is the tangency portfolio that has the maximum Sharpe ratio, which locates at the point where a line drawn from the reference rate of return on the zero-risk axis is tangential to the efficient frontier (Tobin, 1958 and Ross A. Maller and al., 2010). For example, in Figure 7, the x-axis is risk and the y-axis is return, the navy blue curve is efficient frontier of the portfolio composed of 5 industries, and the red line that is tangential to the efficient frontier is the optimal portfolio with maximum Sharpe ratio. Therefore, maximizing Sharpe ratio adjusts the future expected performance for the excess risk in the dynamic strategy.

Figure 7: Tangency Portfolio



We consider the monthly excess returns of the momentum investment strategy:

$$R_t^{\text{MomEx}} = R_t^{\text{Mom}} - R_t^{(f)} = R_t^{(10)} - R_t^{(1)}. \tag{7}$$

The excess return is the same as the profit and loss of winner-minus-loser portfolio (WML). That is, at the beginning of month t, there is an investment of 1\$ in the winner portfolio and a short position of 1\$ in the loser portfolio. The initial investment is zero. At the end of the month, the profit or loss is

$$P\&L^{WML} = \left(1 + R_t^{(10)}\right) - \left(1 + R_t^{(1)}\right) = R_t^{(10)} - R_t^{(1)}. \tag{8}$$

The conditional expected mean μ_t and variance σ_t^2 of the monthly excess return are

$$\mu_t = \mathrm{E}_t^{\mathbb{P}}\left[R_{t+1}^{\text{MomEx}}\right] \text{ and } \sigma_t^2 = \mathrm{Var}_t^{\mathbb{P}}\left[R_{t+1}^{\text{MomEx}}\right]. \tag{9}$$

5.1 Estimation of μ_t and σ_t

Based on DM's idea, the initial regression is

$$R_{t+1}^{\text{MomEx}} = \gamma_0 + \gamma_1 \mathbb{I}_{B,t} + \gamma_2 s_{\mathsf{MKT},t}^{\text{annualized}} + \gamma_3 \mathbb{I}_{B,t} s_{\mathsf{MKT},t}^{\text{annualized}} + \epsilon_t, \tag{10}$$

where the ex-ante bear market indicator $\mathbb{I}_{B.t}$ worth 1 if the cumulative value-weighted market returns is negative in the past 24 months prior to $t+1$, that is

$$\mathbb{I}_{B,t} = \mathbb{I}_{C_{t-23,t}^{\text{Mkt}}<0}. \tag{11}$$

To make the estimation results of Equation (10) consistent with those in DM (2016), here we use $s_{\mathsf{MKT},t}^{\text{annualized}}$, the annualized sample standard deviation of the daily market returns over the last 126 days prior to $t+1$, that is

$$s_{\mathsf{MKT},t}^{\text{annualized}} = \sqrt{252} s_{\mathsf{MKT},t}^{\text{daily}}.$$

The parameters in equation (10) are estimated using monthly momentum excess return on an initial in-sample period of 228 months, from Jan. 1996 to Dec. 2014. Column (5) in Table 2 reports the estimation results of this equation, where the dependent variable is the excess return of momentum investment strategy and the independent variables are the bear market indicator, annualized historical market volatility over the 126 days prior to $t+1$, and the interaction between the bear market indicator and historical market volatility. Other columns in Table 2 present the estimation results when only one or two independent variables in equation (10) are included.

There are compatibility and similarity points between the estimation results in Table 2 and those in DM (2016).[7] From the results, we see the parameters of the bear market indicator variable and annualized historical market volatility variable are not statistically significant when including the interaction term of these two variables. Therefore, instead of applying regression (10), a regression that only includes the interaction term is applied to estimate conditional expected excess return, that is

$$\widehat{\mu}_t = 0.0232 - 0.5590 * \mathbb{I}_{B,t} s_{\mathsf{MKT},t}. \tag{12}$$

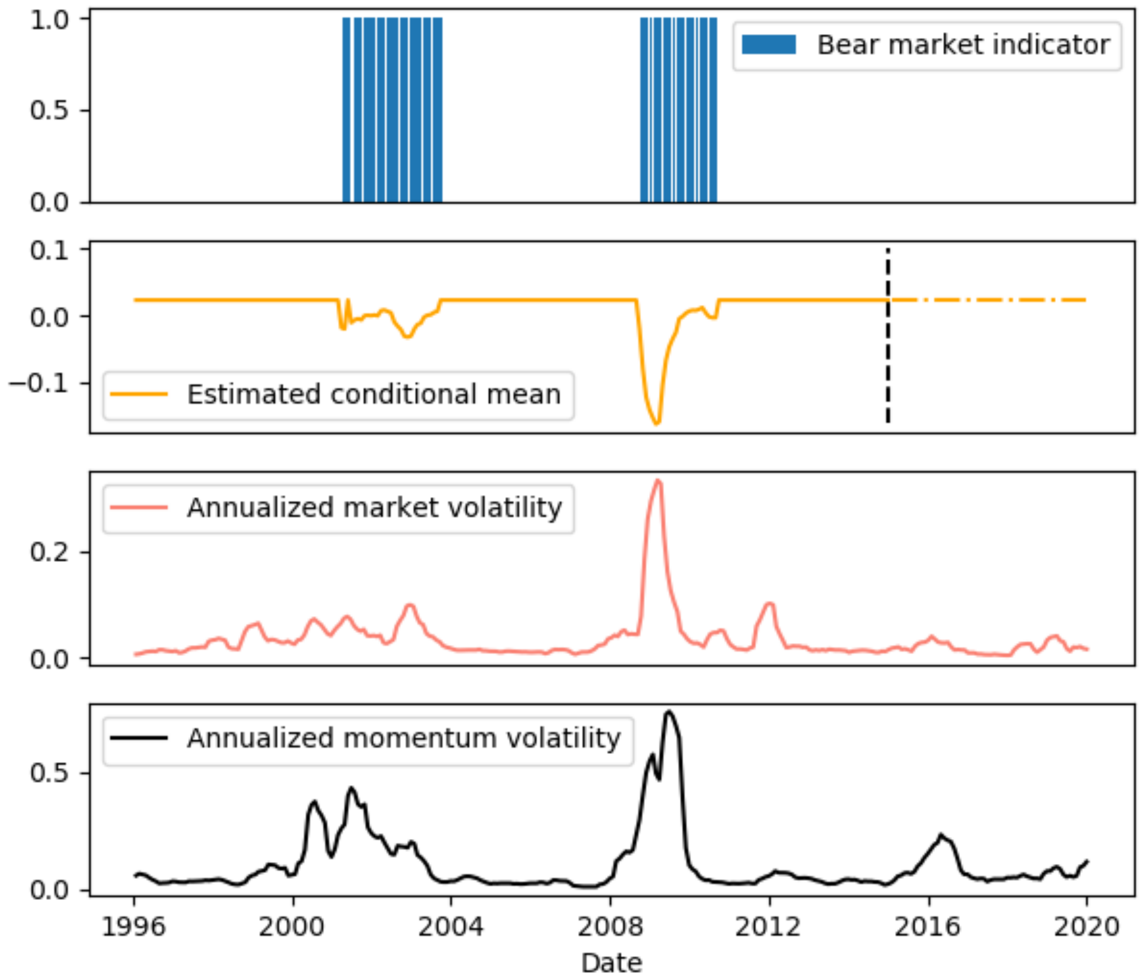
To estimate the conditional variance, we use $s_{\mathsf{Mom},t}^{\text{annualized}}$, the historical annualized momentum volatility of the daily momentum excess return over the past 126 days before month t as the proxy of conditional volatility σ_t, [8] that is

$$\widehat{\sigma}_t = s_{\mathsf{Mom},t}^{\text{annualized}} = \sqrt{252} s_{\mathsf{Mom},t}^{\text{daily}}. \tag{13}$$

[7]The estimation result of DM (2016) is shown in Table A1 of Appendix A.1.

[8]Beside the historical variance, we applied 3 other ways to fit the time-varying conditional variance and get around this, including a GARCH framework, GARCH(1,1) model and GJR-GARCH(1,1,1) model. The model detail, estimation result and analysis of these methods are presented in Appendix A.2.

Figure 8: Bear Market Indicator, Estimated Conditional Mean, Annualized Market and Momentum Volatility



This graph reports the bear market indicator $\mathbb{I}_{B.t}$, the estimated conditional expected return $\widehat{\mu}_t$, the annualized rolling-window historical market volatility over the past 126 days $s^{\text{annualized}}_{\mathsf{MKT},t}$, and the annualized rolling-window historical momentum volatility over the past 126 days $s^{\text{annualized}}_{\mathsf{Mom},t}$. The parameters to calculate the conditional expected return $\widehat{\mu}_t$ are estimated based on the in-sample monthly data from Jan. 1996 to Dec. 2014.

Table 2: OLS Regression Summary

This table presents the summary of OLS regressions (10) where the dependent variable is the excess return of momentum investment strategy and the independent variable are the bear market indicator, annualized historical market volatility over the 126 days prior to $t+1$, and the interaction between bear market indicator and historical market volatility. The parameters are estimated using an initial in-sample period of 228 months, from Jan 1996 to Dec 2014. The t-statistics results are shown in parentheses and the parameters with star means that it is significantly different than zero at the 95% confidence level.

Variable	Coefficient	Estimated Parameters				
		(1)	(2)	(3)	(4)	(5)
1	γ_0	0.0188*	0.0343*	0.0340*	0.0232*	0.0114
		(2.260)	(3.782)	(3.702)	(3.091)	(0.856)
$\mathbb{I}_{B,t}$	γ_1	-0.0284		0.0058		0.0425
		(-1.648)		(0.299)		(1.706)
$s^{\text{daily}}_{\mathsf{MKT},t}$	γ_2		-0.5560^*	-0.5807^*		0.2800
			(-3.902)	(-3.521)		(0.689)
$\mathbb{I}_{B,t}s^{\text{daily}}_{\mathsf{MKT},t}$	γ_3				-0.5590^*	-1.0267^*
					(-4.233)	(-2.314)
R^2		0.012	0.063	0.063	0.073	0.085

Figure 8 presents the bear market indicator $\mathbb{I}_{B.t}$, the estimated conditional expected return $\widehat{\mu}_t$, the annualized rolling-window historical market volatility over the past 126 days $s^{\text{annualized}}_{\mathsf{MKT},t}$, and the annualized rolling-window historical momentum volatility over the past 126 days $s^{\text{annualized}}_{\mathsf{Mom},t}$. We find that when market volatility is high and the bear market indicator equals to one (in 2003 and 2009), the estimated conditional expected return is negative. In other words, market rebounds from bottom are coupled with high market volatility and momentum volatility, which is consistent with the result of Stivers and Sun (2010) and Heidari (2015). Furthermore, the fluctuation range of momentum volatility is greater than that of market volatility.

5.2 The Weights

We can trade in two assets: a risk-free asset and the WML portfolio. At the end of month t, the value of the dynamic strategy is denoted V_t^{Dyn}.

At the beginning of month $t+1$, the agent invest a proportion ω_t of V_t^{Dyn} in the WML portfolio and the rest $1-\omega_t$ in the risk-free asset. Consequently, this investment strategy is self-financing.

At the end of the month $t+1$, the value of this investment strategy is

$$V_{t+1}^{\text{Dyn}} = \omega_t V_t^{\text{Dyn}} \left(1 + R_{t+1}^{\text{Mom}}\right) + \left(1-\omega_t\right) V_t^{\text{Dyn}} \left(1 + R_{t+1}^{(f)}\right).$$

The dynamic strategy excess return is

$$\begin{aligned}
R_{t+1}^{\text{DynEx}} &= R_{t+1}^{\text{Dyn}} - R_{t+1}^{(f)} \\
&= \frac{V_{t+1}^{\text{Dyn}} - V_t^{\text{Dyn}}}{V_t^{\text{Dyn}}} - R_{t+1}^{(f)} \\
&= \omega_t \left(1 + R_{t+1}^{\text{Mom}}\right) + (1-\omega_t)\left(1 + R_{t+1}^{(f)}\right) - 1 - R_{t+1}^{(f)} \\
&= \omega_t \left(R_{t+1}^{\text{Mom}} - R_{t+1}^{(f)}\right) \\
&= \omega_t R_{t+1}^{\text{MomEx}}. \qquad (14)
\end{aligned}$$

Therefore, the conditional expected return and variance of dynamic strategy are

$$\text{E}_t^{\mathbb{P}}\left[R_{t+1}^{\text{DynEx}}\right] = \omega_t \mu_t \text{ and } \text{Var}_t^{\mathbb{P}}\left[R_{t+1}^{\text{DynEx}}\right] = \omega_t^2 \sigma_t^2.$$

DM (2016) sets the weights proportional to μ_t/σ_t^2, that is [9]

$$\omega_t = \frac{1}{2\lambda}\frac{\mu_t}{\sigma_t^2},$$

where "λ is chosen so that the in-sample annualized volatility of the strategy is 19%, the same as that of the CRSP value-weighted index (their market return) over the full sample" (DM, 2016, p.233). However, using the annualized volatility of full-sample market return at time t includes future information, which may cause bias in the achievement of this dynamic strategy. To improve this, we set that

$$\omega_t = \alpha \frac{\mu_t}{\sigma_t^2}, \qquad (15)$$

where the scaling factor α is chosen so that the in-sample variance of dynamic strategy is the same as the in-sample variance of momentum excess return. In this thesis, the in-sample monthly momentum time series is from Jan. 1996 to Dec. 2014, and the out-of-sample data is from Jan. 2015 to Dec 2019.

To calculate the scaling factor α, let

$$\overline{R}^{\text{MomEx}} = \frac{1}{T}\sum_{t=1}^{T}\left(R_t^{\text{MomEx}}\right) \text{ and } \sigma_{\text{MomEx}}^2 = \frac{1}{T}\sum_{t=1}^{T}\left(R_t^{\text{MomEx}} - \overline{R}^{\text{MomEx}}\right)^2$$

be the sample average and variance of the monthly excess return time series of the momentum strategy. T is the total number of months of in-sample monthly momentum time series, which equals to 228 in our case. The scaling factor α in equation (15) is chosen so that the sample variance of dynamic strategy

$$\sigma_{\text{DynEx}}^2 = \frac{1}{T}\sum_{t=1}^{T}\left(R_t^{\text{DynEx}} - \overline{R}^{\text{DynEx}}\right)^2,$$

[9] The proof of this equation is in Appendix C of DM (2016).

is the same as the sample variance of momentum excess return σ^2_{MomEx}. More precisely,

$$\sigma^2_{\text{DynEx}} = \frac{1}{T}\sum_{t=1}^{T}\left(\omega_t R_t^{\text{MomEx}} - \frac{1}{T}\sum_{u=1}^{T}\omega_u R_u^{\text{MomEx}}\right)^2 = \alpha^2 \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\mu_t}{\sigma_t^2}R_t^{\text{MomEx}} - \frac{1}{T}\sum_{u=1}^{T}\frac{\mu_u}{\sigma_u^2}R_u^{\text{MomEx}}\right)^2,$$

that is,

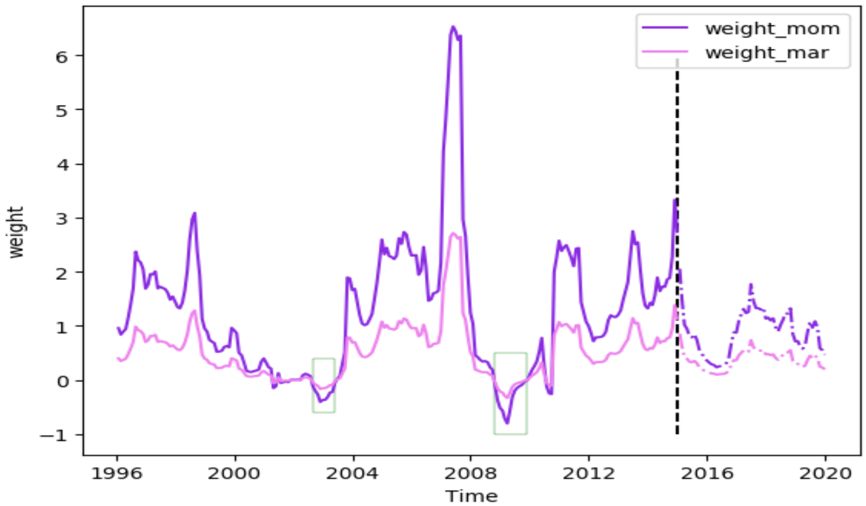
$$\alpha = \frac{\sigma_{\text{MomEx}}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\mu_t}{\sigma_t^2}R_t^{\text{MomEx}} - \frac{1}{T}\sum_{u=1}^{T}\frac{\mu_u}{\sigma_u^2}R_u^{\text{MomEx}}\right)^2}}.$$

To calculate the dynamic weight, as illustrated in section 4.1, the conditional excess momentum return is estimated by equation (12) and the conditional variance is estimated by the historical annualized variance of momentum excess returns over the past 126 days, which is presented in equation (13). Figure 9 compares the dynamic weights between choosing scaling factor α so that the in-sample variance of dynamic strategy σ^2_{DynEx} equals to that of momentum excess return σ^2_{MomEx} (the purple curve) and equals to that of market return σ^2_{MKT} (the pink curve). In our dataset, the annualized in-sample momentum volatility is 38.17%, and the annualized in-sample market volatility is 15.74%. For these two curves, the weights are both high in bull market (ex. in 1999 and 2002-2007) and low in bear market (in 2002-2004 and 2007-2009). Besides, the purple curve has higher weights when market performs well and has lower weights when the market crashes. (The dynamic weight ω_t in the purple curve is calculated by choosing scaling factor α to ensure that the in-sample variance of dynamic strategy σ^2_{DynEx} equals to that of momentum excess return σ^2_{MomEx}.) In particular, the dynamic weights are negative from Jan. 2003 to May 2003 and from Jan. 2009 to Oct. 2009, where the market rallies from previously financial crisis. At the same time, the bear market indicator equals to one, the estimated conditional expected return is negative and the market volatility and momentum volatility are high.

To illustrate the performance of dynamic strategy, Panel 1 in Figure 10 assumes that 1$ is invested in Jan. 1996 and compares the values of dynamic strategy with the sample variance being equal to σ^2_{MomEx}, dynamic strategy with the sample variance being equal to σ^2_{MKT}, and benchmark momentum strategy. From the results, the dynamic strategy with sample variance equaling to σ^2_{MomEx} has the best performance among the three curves, and it outperforms the benchmark momentum strategy and significantly improves the performance during the crash period from 2008 to 2010. In Panel 2, we assume that the 1$ investment is on Jan. 2015, the start time of out-of-sample dataset, we observe the same result as in Panel 1 that the dynamic strategy (red curve) performs best and beats the benchmark momentum strategy.

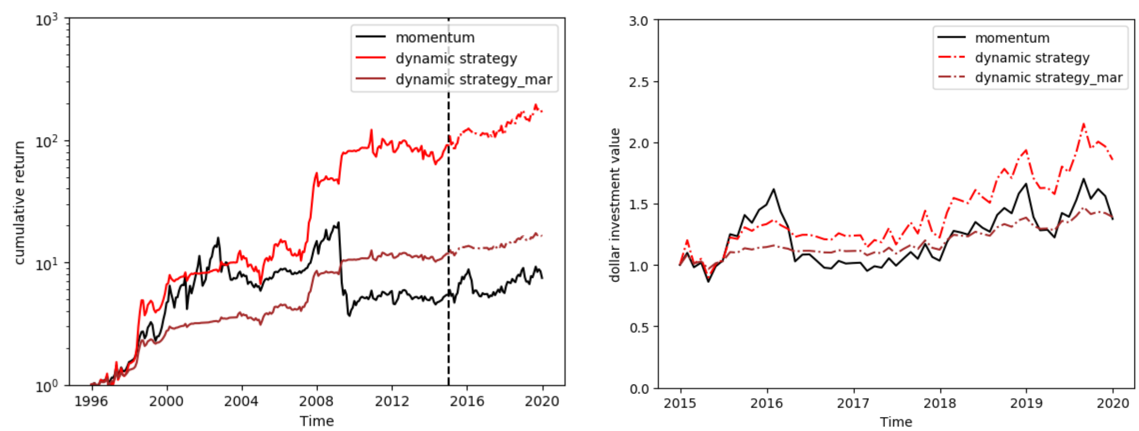
By investigating the optimal weights, benchmark momentum return, and dynamic strategy return, we reveal the explanation for the achievement of this dynamic strategy. As illustrated before, when the formation period is within the market downturn, stocks with great drops are assigned into past loser portfolio and those with less drops are assigned into

Figure 9: Dynamic Weight



The purple curve is the dynamic weight ω_t by chosing scaling factor α so that the in-sample variance of dynamic strategy σ^2_{DynEx} equals to that of momentum excess return σ^2_{MomEx}, and the violet curve is the dynamic weight ω_t by chosing scaling factor α so that the in-sample variance of dynamic strategy σ^2_{DynEx} equals to that of market return σ^2_{MKT}.

Figure 10: The Value of Dynamic Strategy



PANEL 1 PANEL 2

The black curve is the benchmark momentum strategy, the red curve is the dynamic strategy with sample variance equals to σ^2_{MomEx}, and the brown curve is the dynamic strategy with sample variance equals to σ^2_{MKT}. The in-sample period is from Jan. 1996 to Dec. 2014. Panel 1 assumes that 1$ is invested in Jan. 1996, while Panel 2 assumes that the 1$ investment is on Jan. 2015, the start of out-of-sample dataset.

past winner portfolio. However, past losers that have declined sharply previously rise more than past winners when the market experiences a fast rally, which temporarily results in a negative momentum return. We find that the weights are negative when the market rebounds and momentum crashes happen. In other words, the dynamic strategy reduces the momentum crash risk and enhances the return during crashes by longing past losers and shorting past winners.

6 Dynamic Strategy Based on the Option Market Information

Expected return and unexpected shocks are important components in forecasting future stock returns. We consider a dynamic strategy that includes jumps in stock prices, which is measured by short-term option implied volatility difference in level (IVD) and implied volatility relative difference in percentage (RIVD), to improve the perfomance of momentum portfolio.

6.1 Strategy on a Single Stock

As illustrated before, the momentum strategy has long position in past winner stocks and short position in past loser stocks. We first concentrate on a single stock to demonstrate this option information strategy.

6.1.1 Long Position on a Stock

In the case where it is a long position on the stock, the investors does not expect downward jumps. The option market contains information about the underlying asset behavior in the future. In particular, a short-term deep out-of-the-money (OTM) put option price must be close to zero unless the market participants anticipate a negative jump in the near future. Indeed, it takes more than a large volatility for the stock price to go from S_t to the exercise price K before the time-to-maturity τ when K is much smaller that S_t.

Because the level of the underlying price has changed a lot in the past 20 years, it is difficult to construct a trading strategy based on option price level. For that reason, we consider the Cox-Ross-Rubinstein (CRR) binomial tree implied volatility (IV)[10] for American

[10] Let σ_{CRRIV} denotes the CRR implied volatility. More precisely, the time from time t to maturity T is divided into N sub-periods, each period is $h = \frac{T-t}{N}$. We assume that the price of the underlying stock follows a binomial distribution, which is represented using a tree,

$$S_{t+h}^{\text{up}} = S_t \exp(\sigma\sqrt{h}),$$
$$S_{t+h}^{\text{down}} = S_t \exp(-\sigma\sqrt{h})$$

where S_t is the underlying stock price at time t, σ is the volatility. The CRR option price is obtained by using backward induction under risk-neutral measure. The CRR put price at time t is dependent on its price options. A large put option price corresponds to a large implied volatility. To distinguish between a high volatility period and a risk of negative jump, we consider the difference between the IV of two short-term put options: a OTM and an ATM ones[11]:

$$\Delta^{\text{Put}} = \sigma_{\text{CRRIV}}^{\text{Put,OTM}} - \sigma_{\text{CRRIV}}^{\text{Put,ATM}}. \tag{16}$$

1. At beginning of month t, we buy n_{t-1} shares of the stock.[12]

2. During month t, we monitor the daily evolution of the IV difference $\Delta_{j,t}^{\text{Put}}$, with θ^{Put} denoting the first day where $\Delta_{j,t}^{\text{Put}}$ is above a given threshold x.

   (a) If $\max \Delta_t^{\text{Put}} = \max_j \Delta_{j,t}^{\text{Put}}$ is below x, we keep that stock for the whole month.

   (b) If not, then at day θ^{Put}, we sell the n_{t-1} shares of the stock for a total value of $n_{t-1} S_{\theta^{\text{Put}}}^{\text{daily}}$ and reinvest that money in the risk free rate to obtain

   $$m_\theta = \frac{n_{t-1} S_{\theta^{\text{Put}}}^{\text{daily}}}{M_{\theta^{\text{Put}}}^{\text{daily}}}$$

   shares of the risk-free asset where $M_{\theta^{\text{Put}}}^{\text{daily}}$ is the value of the risk-free asset at day θ^{Put} within month t.[13]

3. At the end on month t, the value of this strategy is

$$\begin{aligned}
& n_{t-1} S_t \mathbb{I}_{\max \Delta_t^{\text{Put}} < x} + m_\theta M_t \mathbb{I}_{\max \Delta_t^{\text{Put}} \geq x} \\
= \; & n_{t-1} S_t \mathbb{I}_{\max \Delta_t^{\text{Put}} < x} + \frac{n_{t-1} S_{\theta^{\text{Put}}}^{\text{daily}}}{M_{\theta^{\text{Put}}}^{\text{daily}}} M_{\theta^{\text{Put}}}^{\text{daily}} \left(1 + R_t^{(f)} \left(1 - \frac{\theta^{\text{Put}}}{D}\right)\right) \mathbb{I}_{\max \Delta_t^{\text{Put}} \geq x} \\
= \; & n_{t-1} S_t \mathbb{I}_{\max \Delta_t^{\text{Put}} < x} + n_{t-1} S_{\theta^{\text{Put}}}^{\text{daily}} \left(1 + R_t^{(f)} \left(1 - \frac{\theta^{\text{Put}}}{D}\right)\right) \mathbb{I}_{\max \Delta_t^{\text{Put}} \geq x}.
\end{aligned}$$

---

at time $t+h$,

$$P^{\text{CRR}}(\sigma)_t = \max\left\{\exp(-rh)\left[q P_{t+h}^{\text{up}} + (1-q) P_{t+h}^{\text{down}}\right], S_t - K\right\}$$

where K is the exercise price, r denotes the risk-free rate, P_{t+h}^{up} and P_{t+h}^{down} are the prices of the put option at time $t+h$ when the stock price moves "up" or "down", and q is the risk neutral probability, given by

$$q = \frac{\exp(rh) - \exp(-\sigma\sqrt{h})}{\exp(\sigma\sqrt{h}) - \exp(-\sigma\sqrt{h})}$$

The implied volatility is the the value of σ that should be used in the CRR tree model to match the observed price P, that is,

$$P^{\text{CRR}}(\sigma_{\text{CRRIV}}) = P.$$

[11] The option prices and time-to-maturities vary every day and the methodology is described in the data section.

[12] n_{t-1} is defined the same as in Section 2, it denotes the number of shares outstanding of the stock holding during time period $t-1$ to t.

[13] The daily series of risk-free rate is obtained from the monthly series by converting the risk-free rate at the beginning of each month to a daily rate and we assume that this daily rate is valid throughout the month (Daniel and Moskowitz, 2016).

where D denotes the total number of days within month t.

6.1.2 Short Position on a Stock

In the case where it is a short position on the stock, the investors are worried about upward jumps. In that case, the short term OTM call option price will be a good signal. Indeed, the price of such a call option must be very small unless the market participants anticipate and positive jump in the underlying asset. The trading signal is based on the IV variation between tow short-term call options: an OTM and a ATM ones:

$$\Delta^{\text{Call}} = \sigma^{\text{Call,OTM}}_{\text{CRRIV}} - \sigma^{\text{Call,ATM}}_{\text{CRRIV}}. \tag{17}$$

1. At beginning of month t, we short sale n_{t-1} shares of the stock.
2. During month t, we monitor the daily evolution of the IV difference $\Delta^{\text{Call}}_{j,t}$, with θ^{Call} denoting the first day where $\Delta^{\text{call}}_{j,t}$ is above a given threshold y.
   (a) If $\max \Delta^{\text{Call}}_t = \max_j \Delta^{\text{Call}}_{j,t}$ is below y, we maintain the short sale.
   (b) If not, then at day θ^{Call}, we buy back the n_{t-1} shares of the stock for a total value of $n_{t-1}S^{\text{daily}}_{\theta^{\text{Call}}}$ by borrowing that amount in the risk free rate to obtain
   $$-m_\theta = -\frac{n_{t-1}S^{\text{daily}}_{\theta^{\text{Call}}}}{M^{\text{daily}}_{\theta^{\text{Call}}}}$$
   shares of the risk-free asset where $M^{\text{daily}}_{\theta^{\text{Call}}}$ is the value of the risk-free asset at day θ^{Call} assuming there are 30 days within a month.
3. At the end on month t, the value of this strategy is

$$\begin{aligned}
& -n_{t-1}S_t\mathbb{I}_{\max \Delta^{\text{Call}}_t < y} - m_\theta M_t \mathbb{I}_{\max \Delta^{\text{Call}}_t \geq y} \\
= \; & -n_{t-1}S_t\mathbb{I}_{\max \Delta^{\text{Call}}_t < y} - \frac{n_{t-1}S^{\text{daily}}_{\theta^{\text{Call}}}}{M^{\text{daily}}_{\theta^{\text{Call}}}} M^{\text{daily}}_{\theta^{\text{Call}}} \left(1 + R^{(f)}_t \left(1 - \frac{\theta^{\text{Call}}}{D}\right)\right) \mathbb{I}_{\max \Delta^{\text{Call}}_t \geq y} \\
= \; & -n_{t-1}S_t\mathbb{I}_{\max \Delta^{\text{Call}}_t < y} - n_{t-1}S^{\text{daily}}_{\theta^{\text{Call}}} \left(1 + R^{(f)}_t \left(1 - \frac{\theta^{\text{Call}}}{D}\right)\right) \mathbb{I}_{\max \Delta^{\text{Call}}_t \geq y}.
\end{aligned}$$

6.2 Strategy on Momentum Portfolio

Let OPW and OPL denote the dynamic strategy with option market information on past winner portfolio and past loser portfolio respectively, and OP denotes this strategy on momentum portfolio. At the beginning of month t,

$$\begin{aligned}
V^{\text{OPW}}_{t-1} &= \sum_{i \in \mathcal{P}10_t} n_{i,t-1}S_{i,t-1} \\
V^{\text{OPL}}_{t-1} &= \sum_{j \in \mathcal{P}1_t} n_{j,t-1}S_{j,t-1},
\end{aligned}$$

6.2.1 Strategy on Past Winner Portfolio

At the end of the month, the value of the past winner portfolio in this strategy is

$$V_t^{\text{OPW}} = \sum_{i\in\mathcal{P}10_t}\left(n_{i,t-1}S_{i,t}\mathbb{I}_{\max\Delta_{i,t}^{\text{Put}}<x} + n_{i,t-1}S_{i,\theta_i^{\text{Put}}}^{\text{daily}}\left(1+R_t^{(f)}\left(1-\frac{\theta_i^{\text{Put}}}{D}\right)\right)\mathbb{I}_{\max\Delta_{i,t}^{\text{Put}}\geq x}\right).$$

Therefore, the return of the past winner portfolio in this strategy is

$$\begin{aligned} R_t^{\text{OPW}} &= \frac{V_t^{\text{OPW}} - V_{t-1}^{\text{OPW}}}{V_{t-1}^{\text{OPW}}} \\ &= \sum_{i\in\mathcal{P}10_t}\left(w_{i,t-1}R_{i,t}\mathbb{I}_{\max\Delta_{i,t}^{\text{Put}}<x} + w_{i,t-1}R_{i,\theta_i^{\text{Put}}}\mathbb{I}_{\max\Delta_{i,t}^{\text{Put}}\geq x}\right), \end{aligned}$$

where $w_{i,t-1} = \frac{n_{i,t-1}S_{i,t-1}}{V_{t-1}^{(p)}}$ is defined the same as in Equation (3), $R_{i,t} = \frac{S_{i,t}-S_{i,t-1}}{S_{i,t-1}}$ is the monthly return of stock i and $R_{i,\theta_i^{\text{Put}}}$ is the return between the beginning of the month t and the stop trading time θ_i^{Put} of stock i during month t, which is

$$R_{i,\theta_i^{\text{Put}}} = \frac{S_{i,\theta_i^{\text{Put}}}^{\text{daily}}\left(1+R_t^{(f)}\left(1-\frac{\theta_i^{\text{Put}}}{D}\right)\right) - S_{i,t-1}}{S_{i,t-1}}$$

6.2.2 Strategy on Past Loser Portfolio

Similarly, the value of the past loser portfolio in this strategy at the end of month t is

$$V_t^{\text{OPL}} = \sum_{j\in\mathcal{P}1_t}\left(n_{j,t-1}S_{j,t}\mathbb{I}_{\max\Delta_{j,t}^{\text{Call}}<y} + n_{j,t-1}S_{j,\theta_j^{\text{Call}}}^{\text{daily}}\left(1+R_t^{(f)}\left(1-\frac{\theta_j^{\text{Call}}}{D}\right)\right)\mathbb{I}_{\max\Delta_{j,t}^{\text{Call}}\geq y}\right),$$

and the return of the past loser portfolio in this strategy is

$$\begin{aligned} R_t^{\text{OPL}} &= \frac{V_t^{\text{OPL}} - V_{t-1}^{\text{OPL}}}{V_{t-1}^{\text{OPL}}} \\ &= \sum_{j\in\mathcal{P}1_t}\left(w_{j,t-1}R_{j,t}\mathbb{I}_{\max\Delta_{j,t}^{\text{Call}}<y} + w_{j,t-1}R_{j,\theta_j^{\text{Call}}}\mathbb{I}_{\max\Delta_{j,t}^{\text{Call}}\geq y}\right), \end{aligned}$$

where $R_{j,\theta_j^{\text{Call}}}$ is the return between the beginning of the month t and the stop trading time θ_j^{Call} of stock j during month t,

$$R_{j,\theta_j^{\text{Call}}} = \frac{S_{j,\theta_j^{\text{Call}}}^{\text{daily}}\left(1+R_t^{(f)}\left(1-\frac{\theta_j^{\text{Call}}}{D}\right)\right) - S_{j,t-1}}{S_{j,t-1}}$$

6.2.3 Strategy on Momentum Portfolio

Assume that at the end of month $t-1$, the value of the option information strategy is V_{t-1}^{OP}, the value of this strategy at the end of month t is

$$V_t^{\mathrm{OP}} = \left|V_{t-1}^{\mathrm{OP}}\right| \left(1 + R_t^{\mathrm{OPW}}\right) - \left|V_{t-1}^{\mathrm{OP}}\right| \left(1 + R_t^{\mathrm{OPL}}\right) + V_{t-1}^{\mathrm{OP}} \left(1 + R_t^{(f)}\right).$$

Therefore, the excess returns of this option market information strategy is

$$R_t^{\mathrm{OPEx}} = R_t^{\mathrm{OP}} - R_t^{(f)} = R_t^{\mathrm{OPW}} - R_t^{\mathrm{OPL}}. \tag{18}$$

For the dynamic strategy with option market information, at the end of month t, the value of this strategy is denoted as V_t^{DynOP}. At the beginning of month $t+1$, the agent invests a proportion ω_t of V_t^{DynOP} in the option information strategy and the rest in the risk-free assets. As we can see, the strategy is self-financing.

At the end of month $t+1$, the value of this dynamic strategy with option information is

$$V_{t+1}^{\mathrm{DynOP}} = \omega_t V_t^{\mathrm{DynOP}} \left(1 + R_{t+1}^{\mathrm{OPEx}}\right) + \left(1 - \omega_t\right) V_t^{\mathrm{DynOP}} \left(1 + R_{t+1}^{(f)}\right).$$

6.2.4 Cumulative Returns

The cumulative returns of this dynamic strategy with option market information over T months is

$$C_T^{\mathrm{OP}} = \frac{V_T^{\mathrm{OP}}}{V_0^{\mathrm{OP}}} - 1$$

6.3 The Prediction Effect of Option IV Difference

Before constructing the strategy empirically, we consider linear models and non-linear models to test the predictive power of short-term option implied volatility difference in level (IVD) and implied volatility relative difference in percentage (RIVD) on future returns. The IVDs between short-term OTM options and ATM options are given by equation (16) and (17) for put and call respectively, and the RIVDs are

$$\delta_R^{\mathrm{Put}} = \frac{\sigma_{\mathrm{CRRIV}}^{\mathrm{Put,OTM}} - \sigma_{\mathrm{CRRIV}}^{\mathrm{Put,ATM}}}{\sigma_{\mathrm{CRRIV}}^{\mathrm{Put,ATM}}} \tag{19}$$

$$\delta_R^{\mathrm{Call}} = \frac{\sigma_{\mathrm{CRRIV}}^{\mathrm{Call,OTM}} - \sigma_{\mathrm{CRRIV}}^{\mathrm{Call,ATM}}}{\sigma_{\mathrm{CRRIV}}^{\mathrm{Call,ATM}}} \tag{20}$$

6.3.1 Generalized Linear Models

Based on the strategy, Generalized Linear models are applied on single stocks to test the forecast ability of the extreme values of IVD and RIVD at time t on stock return at time $t+1$. The dataset is splitted into training set (from Jan. 1996 to Dec. 2014, 4784 data) and test set (from Jan. 2015 to Dec. 2019, 1258 data) and we conduct the analysis to the training set. For each stock, we first sort the IVD and RIVD data of call options and put options from small to large respectively. Under the assumption that the stock has more than 100 IVD and RIVD data, the highest 30% are extracted as extreme values to proxy the jumps.

For call options, the four regressions are

$$R_{i,t+1} = \beta_0 + \beta_1 \Delta_{i,t}^{\text{Call}} + \beta_2 R_{i,t} + \epsilon_{t+1} \tag{21}$$

$$R_{i,t+1} = \beta_0 + \beta_1 \delta_{R,i,t}^{\text{Call}} + \beta_2 R_{i,t} + \epsilon_{t+1} \tag{22}$$

$$R_{i,t+1} = \beta_0 + \beta_1 \Delta_{i,t}^{\text{Call}} + \beta_2 \sigma_{\text{BSIV}}^{\text{Call,ATM}} + \beta_3 R_{i,t} + \epsilon_{t+1} \tag{23}$$

$$R_{i,t+1} = \beta_0 + \beta_1 \delta_{R,i,t}^{\text{Call}} + \beta_2 \sigma_{\text{BSIV}}^{\text{Call,ATM}} + \beta_3 R_{i,t} + \epsilon_{t+1} \tag{24}$$

The same regressions for put options are:

$$R_{i,t+1} = \beta_0 + \beta_1 \Delta_{i,t}^{\text{Put}} + \beta_2 R_{i,t} + \epsilon_{t+1} \tag{25}$$

$$R_{i,t+1} = \beta_0 + \beta_1 \delta_{R,i,t}^{\text{Put}} + \beta_2 R_{i,t} + \epsilon_{t+1} \tag{26}$$

$$R_{i,t+1} = \beta_0 + \beta_1 \Delta_{i,t}^{\text{Put}} + \beta_2 \sigma_{\text{BSIV}}^{\text{Put,ATM}} + \beta_3 R_{i,t} + \epsilon_{t+1} \tag{27}$$

$$R_{i,t+1} = \beta_0 + \beta_1 \delta_{R,i,t}^{\text{Put}} + \beta_2 \sigma_{\text{BSIV}}^{\text{Put,ATM}} + \beta_3 R_{i,t} + \epsilon_{t+1} \tag{28}$$

The estimation results of these regressions are listed in the Appendix B.1. From Figure B1 and Figure B2, the prediction power of IVD and RIVD at time t on stock return at time $t+1$ are different from stock to stock. For call options, five stocks with highest R^2, which are higher than 0.5, of equation (21), (22), (23) and (24) are the same: stock of Noble Energy Inc. (NBL), stock of American Tower Corp. (AMT), stock of Reliq Health Technologies Inc. (RHT), stock of L Brands Inc. (LB) and stock of Vitesse Semiconductor Corp. (VTSS). The total number of option data on the training set of these stocks are around 150 days. For put options, five stocks with highest R^2, which are higher than 0.5, of equation (25), (26), (27) and (28) are the same: stock of Sherwin-Williams Co. (SHW), stock of MicroStrategy Inc. (MSTR), stock of Boston Scientific Corp. (BSX), stock of HCA Healthcare Inc. (HCA) and stock of H & R Block Inc. (HRB). And the equations that include IV of ATM options (equation (23), (24), (27) and (28)) have higher R^2 than those that do not include (equation (21), (22), (25) and (26)). However, the R^2 of these models are relatively low for most of the stocks. One of the reasons may be that the relationship between return at day $t+1$ and option IVD and RIVD is non-linear. Therefore, non-linear Machine Learning models are implemented in the next subsections.

6.3.2 Non-linear Machine Learning Models

Instead of implementing the model for each stock, we head to non-linear supervised Machine Learning (ML) regression models to approximate the mapping function that explains the predictive power of all individual options IVD and RIVD to the momentum portfolio return and dynamic momentum portfolio return. Therefore, we have two datasets, the output variable of the first dataset is the return of momentum portfolio at day $t + 1$, and the input variables are the short-term options IVD and RIVD of all the stocks that are selected in momentum portoflio at day t and momentum return at day t. For the second dataset, the output variable is the return of dynamic momentum portfolio at day $t + 1$, and the input variables are the same as the first dataset except that the momentum return at day t is replaced by the dynamic momentum return at day t. The total number of features in these two datasets are both 10177, and there are many missing values in the option information features. The two datasets are both splitted into training set (from Jan. 1996 to Dec. 2014, 4784 data) and test set (from Jan. 2015 to Dec. 2019, 1257 data).

The first Non-linear ML model is Random Forest (RF), which constructs multiple decision trees and combines them to get a more accurate and stable model. RF is a bagged decision tree model by using the Bootstrapping method to select m samples randomly with replacement from the original training set and select n times to generate n subsample sets and decision trees.[14] For a single decision tree model, it is split on multiple features until we conclude. The final predicted value is the average value of all n decision trees. Random is embodied in two aspects: One is to randomly select features, and the other is to randomly select sample, so that each tree in the forest has both similarities and differences. RF model is suitable for high-dimensional, big dataset problems: it is faster to train, and the final predicted value is low bias with moderate variance. In addition, it is less sensitive to missing values.

The second ML model is Extreme gradient boosting (XGBoost). It is an effective gradient boosting decision tree algorithm that has a good performance on various problems. It generates models serially and takes the sum of all models as the output. XGBoost expands the loss function as a second-order Taylor expansion, uses the second derivative information of the loss function to optimize the loss function, and greedily chooses whether to split the node according to whether the loss function is reduced. At the same time, XGBoost adds regularization, learning rate, column sampling, and approximate optimal split point to prevent over-fitting. Certain optimizations have also been made in dealing with missing values. The main difference between RF and XGBoost is that trees are built independently in RF while XGBoost adds a new tree to complement previous built ones (Pan, 2018). The model principle is illustrated in Appendix B.2.

The predicting performance of RF and XGBoost on predicting future momentum return (the first dataset) and predicting future dynamic momentum return (the second dataset) are reported in Appendix B.3. From the result, to forecast future momentum return by RF and XGBoost, the Root Mean Square Error (RMSE) on the test set of the first dataset are 0.05 in RF model and 0.054 in XGBoost model. The RF model outperforms XGBoost model on training set in the first dataset, but XGBoost is better at capturing extreme values. However, both of the RF and XGBoost does not perform well on the test set, with XGBoost performing slightly better than RF. As for predicting future dynamic momentum return, the RMSE on test set iof the second dataset are 0.037 in RF model and 0.042 in XGBoost model. Although the RMSE of XGBoost is slightly higher than RF on test set, The XGBoost is

[14] n is a hyperparameter, and it is set to 100 in the model

better at capturing the return movement, and has better performance than RF on both the training set and test set in the second dataset. After examining the feature importance, the features with highest weight in RF model on two datasets are both the return at day t. While in XGBoost model, the weights on return at day t are not ranked within top 20 features on both two datasets, the features with highest weight are put option IVD of Abbott Laboratories stock (ABT) and call option IVD of Mastercard Inc stock (MA) on the first and second dataset respectively.

From the analysis of Generalized Linear models and ML models, we find that the prediction effect of IVD and RIVD data are different across stocks. The extreme values of IVD and RIVD exhibit good prediction on future return for several stocks with R^2 higher than 0.7 and the statistically significant coefficients, such as the stock of Noble Energy Inc. (NBL). However, for most stocks, the results of Generalized Linear models are not significant statistically. When predicting the future momentum return and dynamic momentum return with all individual option IVD and RIVD data by ML models, the RMSE on the test set is relatively low. We can conclude that the IVD and RIVD data has a certain predictive power on future return, and other factors and predictors should be included to improve the forecast performance.

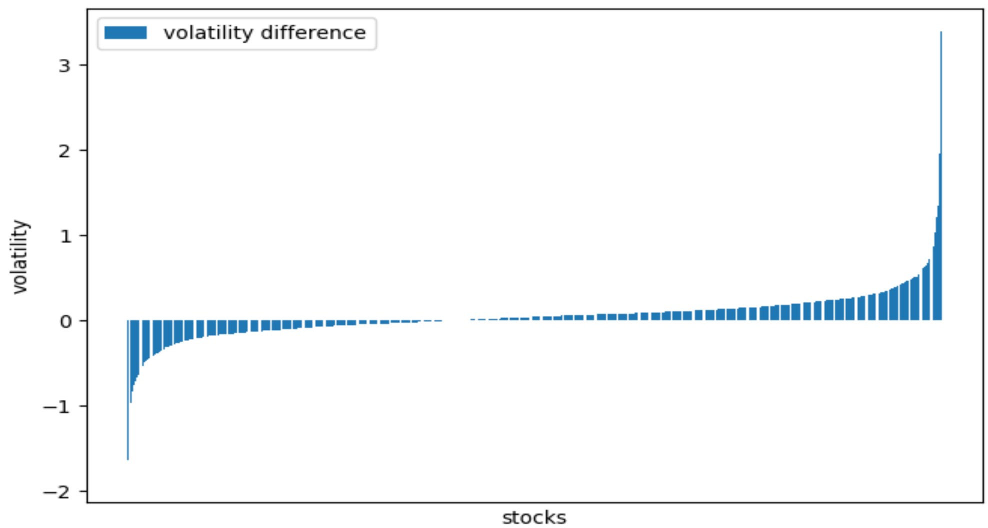
6.4 The Thresholds x and y

One of the key points in this dynamic strategy with option market information is setting the appropriate threshold x and y as stop trading signals. We split the option dataset into training set (from Jan 1996 to Dec 2014, 4784 data) and test set (from Jan 2015 to Dec 2019, 1258 data) By observing in-sample data, the option data varies greatly from stock to stock. By calculating the difference of RIVD standard deviation between put option and call option for each stock, Figure 11 shows that the put option RIVD volatility is higher than that of call option for more than 60% stocks, which illustrates that the IVD and RIVD data of put options fluctuate more dramatically than those of call options. From the in-sample data, the majority call option IVD and RIVD data fluctuate in a small range, with several extreme values that are significantly deviated. While the fluctuation range of put option IVD and RIVD data is larger than that of call option data. Therefore, we consider using RIVD data to set thresholds and choosing different values for x and y of different stocks.

For put option data of each stock, after sorting the RIVD data from small to large, the value at 99% of total number of δ_R^{Put} data is chosen as the threshold of x. For example, stock AAPL has 1359 put option data on the training set. Then the threshold x is set at the 1345th data of RIVD from small to large, which equals to 0.823. In addition, there will be 14 stop points when constructing the strategy on the stock.

After sorting the RIVD call option data from small to large, the value at 96% of total number of δ_R^{Call} data is chosen as the threshold of y. For example, stock GS has 788 call option data on the training set. Then we set the treshold y at the 756th data of sorted RIVD data, which equals to 0.216, and there will be 32 stop points when constructing the strategy on the stock.

Figure 11: RIVD Standard Deviation Difference Between Put and Call



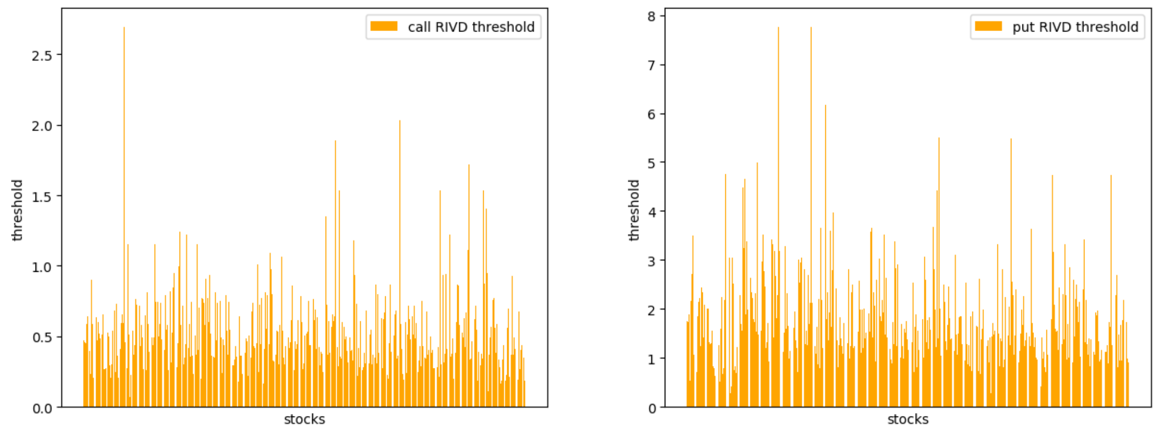
This figure presents the differences of RIVD standard deviation between put options and call options for each stock. The x-axis is every stock, each bar represents the differences of RIVD standard deviation between put options and call options. The RIVD standard deviation of put option is higher than that of call option for more than 60% stocks.

In this case, the thresholds x and y are chosen by taking the difference between stocks, the total numbers of option data and the fluctuation difference between call options and put options into consideration. Figure 12 reports the thresholds x and y across stocks. From the results, the call option stop trading signal, threshold y, is less volatile than the put option signal threshold x. This is consistent with the characteristics of option RIVD data.

With the thresholds x and y that are set on in-sample data (from Jan. 1996 to Dec. 2014), we construct the strategy based on option information. Figure 13 compares the values of benchmark momentum strategy (the black curve), option information strategy on benchmark momentum (the blue curve), dynamic strategy (the red curve), and dynamic strategy with option information (the green curve). In Panel 1, it is assumed that $1 is invested at the beginning of Jan 1996, the option information strategy slightly improves the performance of benchmark momentum strategy by 15.33% (the values at the end of Dec 2019 are $7.44 of benchmark momentum and $8.59 of option information strategy). However, the dynamic strategy with option information does not outperform the dynamic strategy on benchmark. At the same time, the performance of out-of-sample (OOS) data is concerned. In Panel 2, assuming that the 1$ investment is at the beginning of the test set (Jan. 2015), we find that neither the option information strategy nor the dynamic strategy with option information outperform the benchmark momentum and dynamic momentum on the test set.

To break down further this result according to the performance of the strategy on past

Figure 12: Thresholds x and y



PANEL 1 PANEL 2

Panel 1 reports the threshold y as call options stop trading signal across stocks, while Panel 2 reports the threshold x as put options stop trading signal across stocks. The x-axis is every stock and each bar is the option RIVD threshold for a stock in both panels.

Figure 13: Comparison of Different Strategies

PANEL 1 PANEL 2

This graph presents the cumulative return of the benchmark momentum strategy (black curve), dynamic strategy on benchmark (red curve), option information strategy (blue curve) and dynamic strategy on option information (green curve). The thresholds x and y for stop trading are setted on training set (from Jan. 1996 to Dec. 2014). Panel 1 assumes that 1\$ is invested in Jan. 1996, while Panel 2 assumes that the 1\$ investment is on Jan. 2015, the beginning of test set.

winner portfolio and past loser portfolio respectively, which is shown in Figure C1 in Appendix C, we find that the option information strategy exhibit improvement on past losers by stop trading stocks whose call option RIVD exceed their threshold x, while it fails to improve the past winner performance by stop trading stocks whose put option RIVD exceed their threshold y.

Based on the result, two methods are considered to test the stop trading signals.[15] In the first method, we take the underlying returns that correspond to the option IV jumps into consideration. More precisely, in the case where a long position on the stock, the stop strading signals are: the put option RIVD data at day t exceeds the threshold x and the underlying stock return at day t is lower than the daily risk-free rate at that day; On the opposite, in the case where a short position on the stock, the stop trading signals are: the call option RIVD data at day t exceeds the threshold y and the underlying stock return at day t is higher than the daily risk-free rate at that day. In the second method, the dynamic thresholds are considered to update the thresholds every six months, as more information are known as time goes by.

From Figure D1 in Appendix D.1, the option information strategy (dynamic option information strategy) that uses option RIVD thresholds and underlying return together as the signal to stop trading, the chocolate curve (the plum curve), does not outperform the option information strategy (dynamic option information strategy) with static thresholds, the navy blue curve (the green curve). One of the reasons may be that, the option information strategy stops trading the past winner stocks (past loser stocks) from day t that has option signal within the month to the end of that month, therefore, adding underlying returns before and at day t as another creteria to stop trading is not enough to show the near future returns. More precisely, in the strategy with both option information and underlying return as signals, when the option RIVD data at day t exceeds the static thresholds but the underlying stock return at day t is high for past winners (low for past losers), which is not a stop trading signal, the stock return from day t to the end of that month may lower than the return of risk-free asset over the same time period, and we may miss a stop point.

The second method, applying dynamic thresholds that are updated every six months enhances the performance of the option information strategy, but the performance improvement is minimal.[16] From the result, compared to static thresholds, dynamic thresholds improve 5.41% performance on benchmark momentum and 5.98% on dynamic momentum.

6.5 Subsample Test on S&P500

As the static thresholds x and y are set on the in-sample data, the achievement of the option information strategy over the crisis period within the test set needs to be tested. However, there is no dramatic crash of the benchmark momentum portfolio from Jan. 2015 to Dec.

[15] The performance of the option information strategy with these two methods are presented in Appendix D.1 and D.2 respectively.

[16] The performance comparison between static threshold and dynamic threshold is shown in Figure D2 in Appendix D.2.

2019.[17] We consider a momentum porfolio on S&P500 as a subsample benchmark to test the performance of option information strategy during crisis within test set.

The same method as in section 4 is applied to build the momentum portfolio on S&P500. There are around 50 stocks in the S&P500 past winner portfolio and the S&P500 past loser portfolio for each month. As less stocks are invested in, the return of S&P500 momentum portfolio is more volatile than that of the previous benchmark momentum portfolio. The dynamic S&P500 momentum return is given by adding the optimal weight ω_t that is calculated in section 5.2 on S&P500 momentum excess return and investing the rest weight $(1 - \omega_t)$ into risk-free asset.

Figure 14 compares the cumulative return value in dollar of benchmark S&P500 momentum (the black curve), dynamic S&P500 momentum (the red curve), option information strategy on benchmark S&P500 momentum (the navy blue curve), and option information strategy on dynamic S&P500 momentum (the green curve). Panel 1 assumes that 1$ is invested at the beginning of the whole dataset (Jan. 1996), while Panel 2 assumes that 1$ is invested at the beginning of the test set (Jan. 2015). From the result, we find that the option information strategy improves the performance of benchmark S&P500 momentum both over full-sample and over the test dataset, while it does not help to increase the cumulative return of dynamic S&P500 momentum, neither on whole dataset nor on the test dataset. Besides, we observe that there is a crash in benchmark S&P500 momentum from the middle of 2019 to the end of 2019, which is the economic stagnation caused by the recent shutdown and quarantine because of the COVID-19. The option information strategy helps to improve the performance during this crisis by reducing the extent of drop.

7 Concluding Remarks

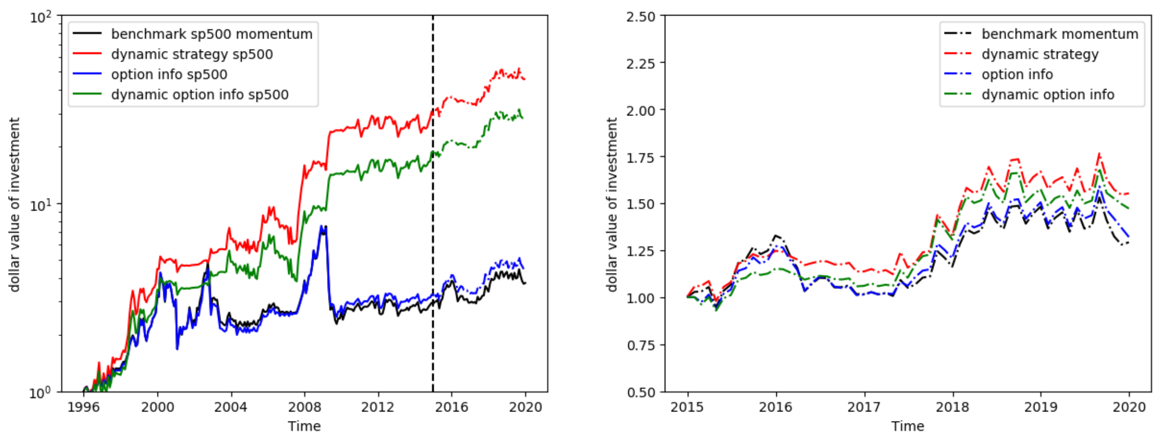
Momentum enjoys a strong performance in a normal market, with infrequent and persistent anomalies in a panic market. This paper investigates the characteristics and predictability of these crashes, conducts a dynamic strategy proposed by DM (2016) and an option information strategy that takes price jumps into consideration to improve the performance of momentum, especially during crashes.

This thesis contributes to the literatures in several ways. Firstly, we collect researches on momentum, investigate momentum crashes by analyzing its performance, its volatility, market volatility and bear market indicators. Ultimately, we find that these crashes happen when the market experiences a rally after a downturn, coupled with a high ex-ante volatility. This is consistent with the existing literatures (Stivers and Sun(2010), Heidari (2015), and DM (2016)).

Secondly, we construct a dynamic strategy based on DM's idea. Instead of using full-sample market volatility to choose a scaling factor that includes future information to estimate conditional volatility, our optimal weight is propotional to conditional mean of excess

[17] The 2015–2016 stock market selloff that happens in the middle of 2016 does not result in a distinct crash in our benchmark portfolio.

Figure 14: Comparison of Different Strategies on S&P500



PANEL 1 PANEL 2

This graph presents the cumulative return of the benchmark S&P500 momentum (the black curve), dynamic S&P500 momentum (the red curve), option information strategy on benchmark S&P500 momentum (the navy blue curve), and option information strategy on dynamic S&P500 momentum (the green curve). The option information thresholds x and y for stop trading are setted on training set (from Jan. 1996 to Dec. 2014). Panel 1 assumes that 1$ is invested in Jan. 1996, while Panel 2 assumes that the 1$ investment is on Jan. 2015, the beginning of test set.

return which is estimated by a regression on the interaction between bear market indicators and the annualized historical market volatility over past 126 days (equation (12)). In addition, the optimal weight is inversely proportional to the conditional variance, which is estimated by annualized realized excess return volatility, with a scaling factor that makes the in-sample dynamic return variance equals to the in-sample momentum variance. By analyzing the optimal weight, we find that dynamic strategy reduces the crash risk by dynamically adding negative weight to benchmark momentum when market rebounds fast after a downturn.

Thirdly, we propose an option information strategy that includes price jumps which are measured by option implied volatility difference in level and relative difference in percentage between OTM options and ATM options. We find that the option IVD and RIVD data have certain predictive powers on future return and this strategy slightly improves the momentum performance, especially in past loser portfolio by stop trading loser stocks whose call option RIVD exceed their threshold x. Besides, it helps to reduce the extent of drop during out-of-sample crisis. However, the improvement is not significant, and applying this strategy to dynamic momentum return does not enhance its performance.

Using the implied volatility difference between short-term OTM options and ATM options to capture the fleeting price jumps is an interesting research direction on understanding the idiosyncratic and systematic risk and constructing trading strategy to enhance return. Finding the proper thresholds as signals is a difficult task and one of the key points in the option information strategy, which would be a potential topic for further research. In addition, instead of reinvesting the money earned by stop trading the stocks that fall sharply in short-term into risk-free asset, investing this amount in other assets which are more profitable is a subject for further research and may help to further improve the performance.

References

[1] B. D. Grundy and J. S. Martin, *Understanding the Nature of the Risks and the Source of the Rewards to Momentum Investing*, The Review of Financial Studies, Vol. 14, No. 1 (Spring, 2001), pp. 29-78.

[2] B. S. Paye, A. Timmermann, *Instability of return prediction models*, Journal of Empirical Finance 13 (2006) 274–315.

[3] Bingyue Pan, *Application of XGBoost algorithm in hourly PM2.5 concentration prediction*, IOP Conf. Series: Earth and Environmental Science 113 (2018) 012127.

[4] D. Avramov, Si Cheng and A. Hameed, *Time-Varying Liquidity and Momentum Profits*, Journal of Financial and Quantitative Analysis, Vol. 51, No. 6, 2016.

[5] C. Stivers and L. Sun, *Cross-Sectional Return Dispersion and Time Variation in Value and Momentum Premiums*, The Journal of Financial and Quantitative Analysis, Vol. 45, No. 4 (AUGUST 2010), pp. 987-1014.

[6] C. S. Asness, A. Frazzini, R. Israel and T. J. Moskowitz, *Fact, Fiction and Momentum Investing*, Journal of Portfolio Management, Fall 2014 (40th Anniversary Issue).

[7] C. S. Asness, T. J. Moskowitz and L. Pedersen, *Value and Momentum Everywhere*, Journal of Finance, 2013, vol. 68, issue 3, 929-985.

[8] D. Amaya, P. Christoffersen, K. Jacobs, A. Vasquez, *Does realized skewness predict the cross-section of equity returns?*, Journal of Financial Economics, Volume 118, Issue 1, October 2015, Pages 135-167.

[9] Doron Avramov, *Stock return predictability and model uncertainty*, Journal of Financial Economics 64 (2002) 423–458.

[10] E. Jondeaua, M. Rockinger *Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements*, Journal of Economic Dynamics & Control 27 (2003) 1699–1737.

[11] George J. Jiang and Tong Yao, *Stock Price Jumps and Cross-Sectional Return Predictability*, Journal of Financial and Quantitative Analysis, 2013, vol. 48, issue 5, 1519-1544.

[12] Harry Markowitz, *Portfolio Selection*, The Journal of Finance, Vol. 7, No. 1. (Mar., 1952), pp. 77-91.

[13] Jacob S. Sagi and Mark S. Seasholes, *Firm-specific attributes and the cross-section of momentum*, Journal of Financial Economics, 2007, vol. 84, issue 2, 389-434.

[14] J. Bégin, C. Dorion, and G. Gauthier, *Idiosyncratic Jump Risk Matters: Evidence from Equity Returns and Options*, The Review of Financial Studies, vol. 33, no 1, 2019, p. 155-211.

[15] J. Tobin, *Liquidity Preference as Behavior Towards Risk*, The Review of Economic Studies, Vol. 25, No. 2 (Feb., 1958), pp. 65-86.

[16] K. Daniel and T. J. Moskowitz, *Momentum Crashes*, Journal of Financial Economics, 122 (2016) 221–247.

[17] K. Daniel, D. Hirshleifer and A. Subrahmanyam, *Investor Psychology and Security Market Under- and Overreactions*, The Journal of Finance, Vol. 53, No. 6 (Dec., 1998), pp. 1839-1885.

[18] Mahdi Heidari, *Momentum Crash Management*, Stockholm School of Economics, March 14, 2015.

[19] M. J. Cooper, R. C. Gutierrez JR., and A. Hameed, *Market States and Momentum*, The Journal of Finance, Vol. Lix, NO. 3, June 2004.

[20] N. Jegadeesh and S. Titman, *Momentum has its moments*, The Journal of Finance, Vol. 48, No. 1. (Mar., 1993), pp. 65-91.

[21] P. Barroso and P. Santa-Clara, *Returns to Buying Winners and Selling Losers:Implications for Stock Market Efficiency*, Journal of Financial Economics, 116 (2015) 111–120.

[22] R. Bird, X. Gao and D. Yeung, *Time-series and cross-sectional momentum strategies under alternative implementation strategies*, Australian Journal of Management, 2017, vol. 42, issue 2, 230-251.

[23] Ross A. Maller, Robert B. Durand, and H.Jafarpour, *Optimal ortfolio choice using the maximum Sharpe ratio*, The Journal of Risk; London Vol. 12, Iss. 4, (Summer 2010): 49-73.

[24] S.P. Kothari and Jay Shanken, *Stock return variation and expected dividends: A time-series and cross-sectional analysis*, Journal of Financial Economics, 31 (1992) 177-210. North-Holland.

[25] T. G. Andersen, N. Fusari, V. Todorov, *Short-Term Market Risks Implied by Weekly Options*, CREATES Research Papers 2018-08, Department of Economics and Business Economics, Aarhus University..

A Estimation Result and GARCH Models

A.1 Estimation Result in DM (2016)

Table A1 reports the estimation results of equation (10) in Daniel and Moskowitz (2016) (Table 5 on p.232).

Table A1: OLS Regression Summary in DM (2016)

The parameters are estimated using monthly data over the period Jul. 1927 to Mar. 2013. The t-statistics results are shown in parentheses.

Variable	Coefficient	Estimated Parameters				
		(1)	(2)	(3)	(4)	(5)
1	γ_0	0.01955	0.02428	0.02500	0.01973	0.02129
		(6.6)	(7.5)	(7.7)	(7.1)	(5.8)
$\mathbb{I}_{B,t}$	γ_1	-0.02626		-0.01281		0.023
		(-3.8)		(-1.6)		(0.0)
$s^{\text{daily}}_{\mathsf{MKT},t}$	γ_2		-0.330	-0.275		-0.088
			(-5.1)	(-3.8)		(0.8)
$\mathbb{I}_{B,t}s^{\text{daily}}_{\mathsf{MKT},t}$	γ_3				-0.397	-0.323
					(-5.7)	(-2.2)

A.2 GARCH Models

The parameters of three GARCH models below are estimated with the same initial in-sample period of 228 months as estimating conditional excess return, from Jan. 1996 to Dec. 2014.

1. GARCH Model 1: a GARCH framework

$$R^{\text{MomEx}}_{t+1} = \gamma_0 + \gamma_3 \mathbb{I}_{B,t} s^{\text{daily}}_{\mathsf{MKT},t} + \sqrt{h_{t+1}}\varepsilon_{t+1}, \quad (29)$$

$$h_{t+1} = \alpha_0 + \alpha_1 h_t + \alpha_2 \left(\sqrt{h_t}\varepsilon_t - \alpha_3\right)^2. \quad (30)$$

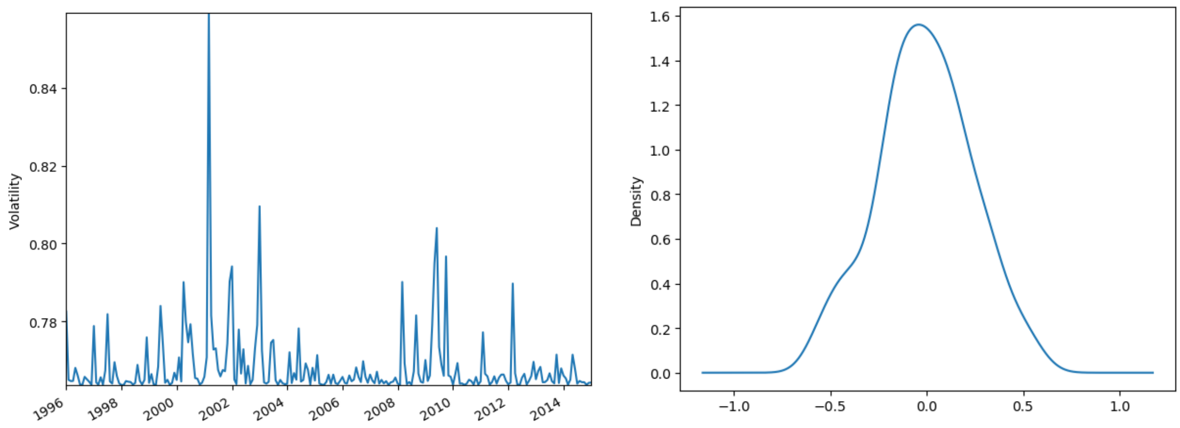
Therefore, the variance of GARCH framework is given by:

$$h_{t+1} = \alpha_0 + \alpha_1 h_t + \alpha_2 \left(R^{\text{MomEx}}_t - \gamma_0 + \gamma_3 \mathbb{I}_{B,t-1} s^{\text{daily}}_{\mathsf{MKT},t-1} - \alpha_3\right)^2,$$

which is an estimate of the monthly excess return variance. The volatility is often presented in a annualized version, that is,

$$\sqrt{h^{\text{annualized}}_{\mathsf{MomEx},t}} = \sqrt{12}\sqrt{h^{\text{monthly}}_{\mathsf{MomEx},t}}.$$

Figure A1: Volatility of GARCH Model 1



PANEL 1 PANEL 2

This graph presents the annualized volatility of GARCH framework: Equation (29), (30), and the residual distribution.

The parameters are estimated by maxmimum likelihood method (MLE), that is

$$\theta = \arg\max -\frac{1}{2}\left[\sum_{t=0}^{T-1}\left(\log(2\pi) + \log(h_{t+1}) + \frac{R_{t+1}^{\text{MomEx}} - \gamma_0 + \gamma_3 \mathbb{I}_{B,t} s_{\mathsf{MKT},t}^{\text{daily}}}{h_{t+1}}\right)\right]$$

with restrictions on parameters

$$\begin{aligned}\alpha_1 + \alpha_2 &< 1,\\ \alpha_0 + \alpha_2\alpha_3^2 &> 0,\\ \alpha_1 > 0, \alpha_2 &> 0.\end{aligned}$$

The estimation result of this model is presented in Column(1) of Table A2, and Figure A1 reports the annualized excess return volatility and the distribution of residuals. The t-statistic results show that most of the parameters are statistically significant except for the parameters α_0. The residuals are statistically normally distributed. DW test statistics is constructed to study the autocorrelation between residuals and their one-period lag, that is

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

where u_t follows standard normal distribution. From the result of the DW test in column (1) of Table A1, we see no autocorrelation from the DW test result.

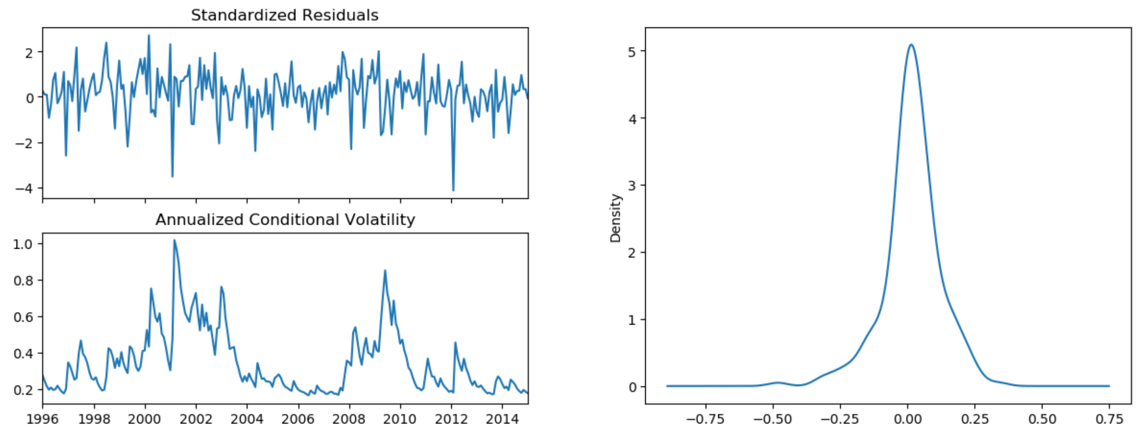
2. GARCH Model 2: GARCH(1,1)
In this model, the excess momentum return process is fitted by Equation (29) as before,

Table A2: GARCH Framework Summary

This table presents the estimation result of GARCH models. Column (1) reports the estimation result of Equation (29), (30), Column (2) reports the estimation result of GARCH(1,1) model and Column (3) reports the estimation result of GJR-GARCH(1,1,1) model. The parameters are estimated using an initial in-sample period of 229 months, from Dec 1995 to Dec 2014. The t-statistics results are shown in parentheses.

Coefficient	MLE Estimation		
	(1)	(2)	(3)
γ_0	0.0196	0.0036	0.0132*
	(1.413)	(0.682)	(2.548)
γ_3	−0.5325*	−0.4172*	−0.4388*
	(-1.055)	(-1.912)	(-3.890)
α_0	0.0465*	0.0007	0.0006*
	(6.389)	(1.321)	(1.692)
α_1	0.0434	0.6490*	0.6977*
	(0.293)	(3.801)	(4.305)
α_2	0.0440*	0.3145*	0.5367*
	(1.473)	(1.719)	(1.919)
α_3	0.0440*		
	(1.254)		
α_4			−0.5367*
			(-2.198)
DW	1.908	1.959	2.000

Figure A2: GARCH(1,1) Model



PANEL 1 PANEL 2

This graph presents the annualized volatility of GARCH(1,1) model, the standardized residuals and the distribution of residuals. The standardized residuals are given by: $SRes = \frac{Res-\overline{Res}}{Var(Res)}$.

while the volatility process of excess momentum returns is fitted by a GARCH (1,1), that is,

$$R_{t+1}^{\text{MomEx}} = \gamma_0 + \gamma_3 \mathbb{I}_{B,t} s_{\mathsf{MKT},t}^{\text{daily}} + \sqrt{h_{t+1}}\varepsilon_{t+1}, \tag{31}$$

$$h_{t+1} = \alpha_0 + \alpha_1 h_t + \alpha_2 \left(\sqrt{h_t}\varepsilon_t\right)^2. \tag{32}$$

Therefore, the variance of GARCH (1,1) is an estimate of the monthly excess return variance:

$$h_{t+1} = \alpha_0 + \alpha_1 h_t + \alpha_2 \left(R_t^{\text{MomEx}} - \gamma_0 + \gamma_3 \mathbb{I}_{B,t-1} s_{\mathsf{MKT},t-1}^{\text{daily}}\right)^2,$$

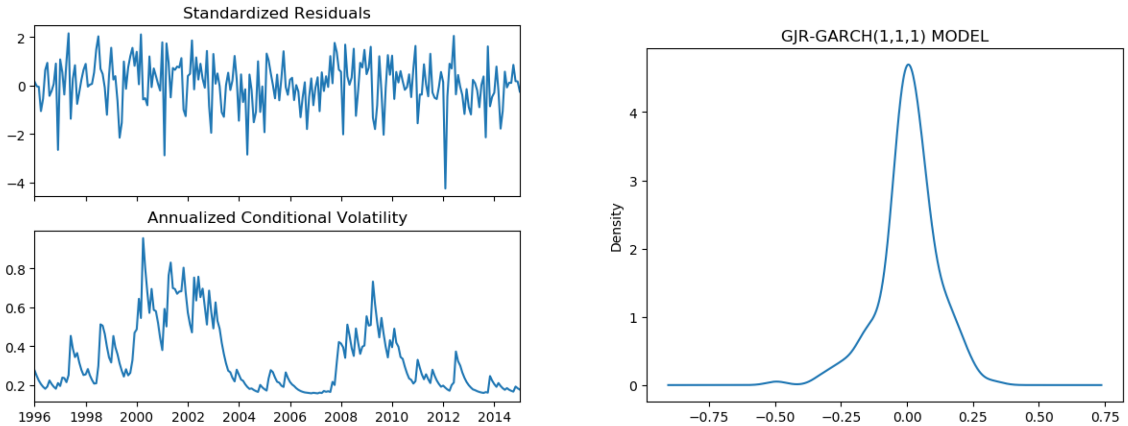
The estimation result is presented in Column(2) of Table A2, and Figure A2 reports the annualized excess return volatility and the residual distribution. The t-statistic results show that most of the parameters are statistically significant except for the parameters γ_0 and α_0. The residuals are statistically normally distributed, and there is no autocorrelation from the DW test result.

3. GARCH Model 3: GJR-GARCH(1,1,1)
In this model, the excess momentum return process is fitted by Equation (29) as before, while the volatility process of excess momentum returns is fitted by a GJR-GARCH (1,1,1), that is,

$$R_{t+1}^{\text{MomEx}} = \gamma_0 + \gamma_3 \mathbb{I}_{B,t} s_{\mathsf{MKT},t}^{\text{daily}} + \sqrt{h_{t+1}}\varepsilon_{t+1}, \tag{33}$$

$$h_{t+1} = \alpha_0 + \alpha_1 h_t + \alpha_2 \left(\sqrt{h_t}\varepsilon_t\right)^2 + \alpha_4 \left(\sqrt{h_t}\varepsilon_t\right)^2 \mathbb{I}_{\varepsilon_t<0} \tag{34}$$

Figure A3: GJR-GARCH(1,1,1) Model



PANEL 1 PANEL 2

This graph presents the annualized volatility of GJR-GARCH(1,1,1) model, the standardized residuals and the distribution of residuals. The standardized residuals are given by: $SRes = \frac{Res - \overline{Res}}{Var(Res)}$.

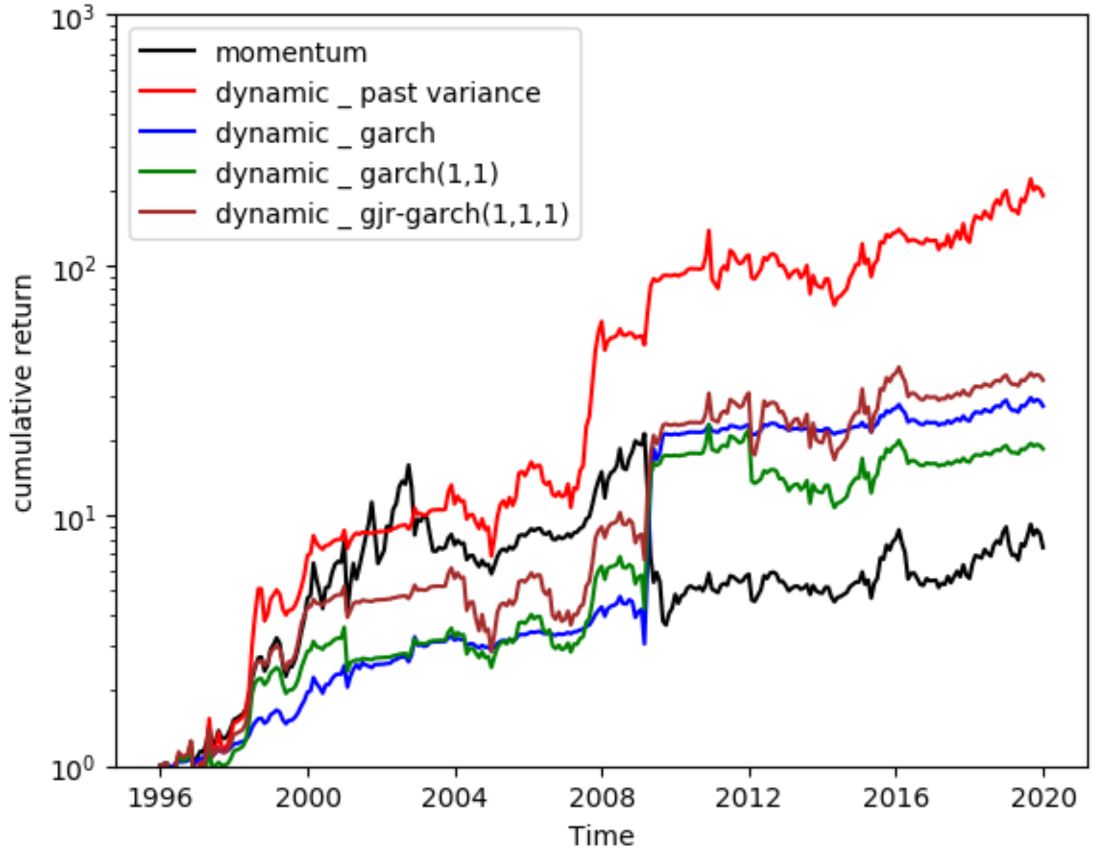
Therefore, the variance of GJR-GARCH(1,1,1) is an estimate of the monthly excess return variance:

$$h_{t+1} = \alpha_0 + \alpha_1 h_t + \alpha_2 \left(R_t^{\text{MomEx}} - \gamma_0 + \gamma_3 \mathbb{I}_{B,t-1} s_{\mathsf{MKT},t-1}^{\text{daily}} \right)^2 + \alpha_4 \mathbb{I}_{\epsilon_t < 0},$$

The estimation result is presented in Column(3) of Table A2, and Figure A3 reports the annualized excess return volatility and the distribution of residuals. The t-statistic results show that all the parameters are statistically significant. The residuals are statistically normally distributed, and there is no autocorrelation from the DW test result.

Figure A4 compares the cumulative return of benchmark momentum strategy, dynamic strategy with conditional variance estimated by historical volatility of momentum excess return, GARCH framework of Equation (29) and (30), GARCH(1,1) model and GJR-GARCH(1,1,1) model. From the result, the dynamic strategy with conditional variance estimated by historical volatility of momentum excess return performs best.

Figure A4: Comparison of Dynamic Strategies



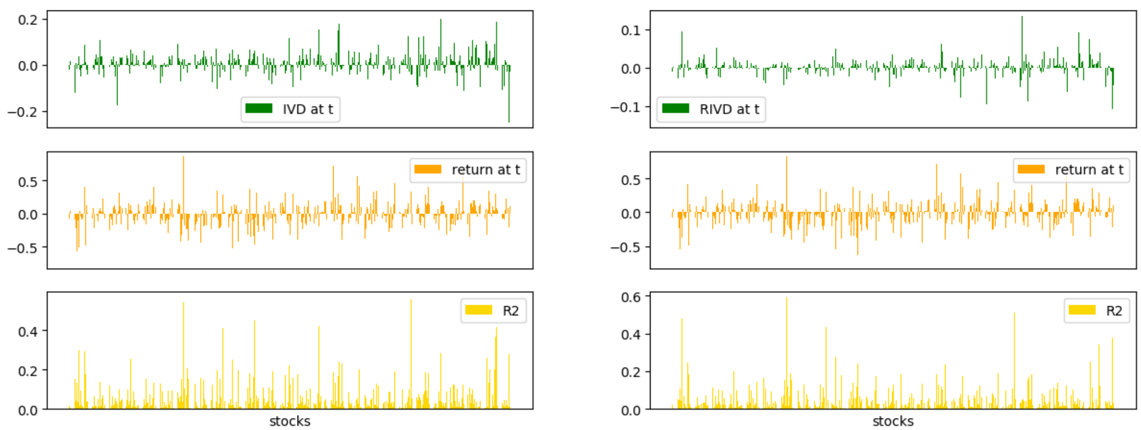
This graph presents the comparison of cumulative return of benchmark momentum strategy (black curve), dynamic strategy with conditional variance estimated by historical volatility of momentum excess return (red curve), GARCH framework of Equation (29) and (30) (blue curve), GARCH(1,1) model (green curve) and GJR-GARCH(1,1,1) model (brown curve).

B The Prediction Effect of IVD and RIVD

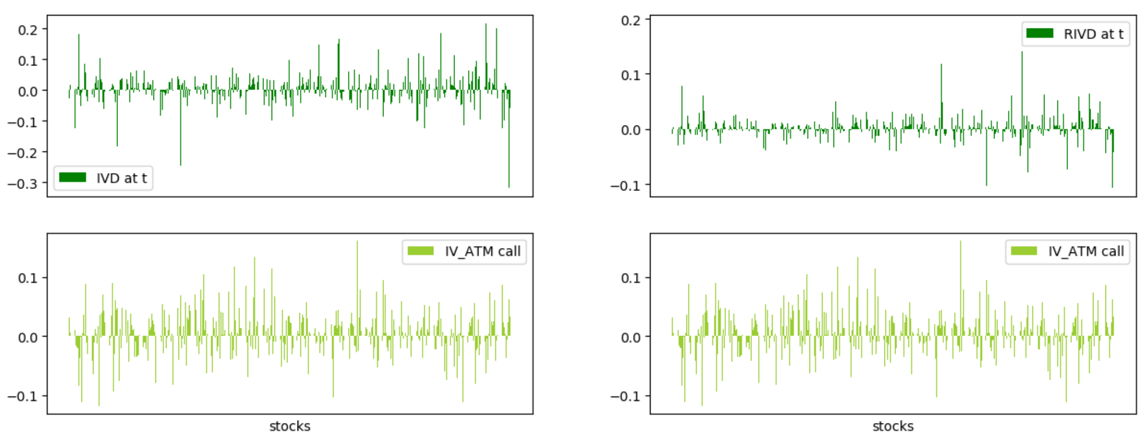
B.1 Estimation Results of Two Linear Models

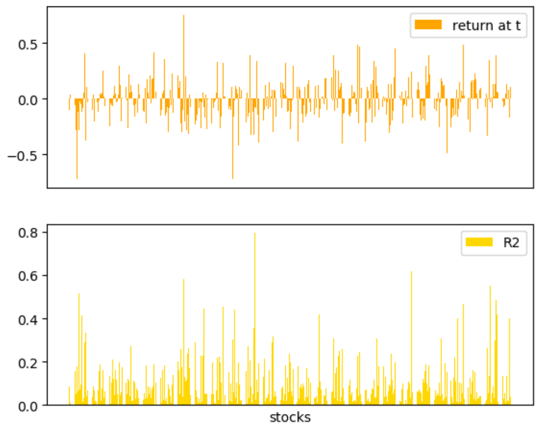
As illustrated in section 6.3.1, the dependent variables for all four regressions are the daily stock return at day $t+1$. The explainatory variables in the first regression (equation (21) and (25) for call and put respectively) are daily IVD at day t and stock return at day t; those in the second regression (equation (22) and (26) for call and put respectively) are RIVD at day t and stock return at day t; those in the third regression (equation (23) and (27) for call and put respectively) are IVD at day t, IV of ATM option at day t and stock return at day t; and those in the fourth regression (equation (24) and (28) for call and put respectively) are RIVD at day t, IV of ATM option at day t and stock return at day t. Figure B1 reports the coefficients and R^2 of the four regressions for call options, while Figure B2 reports those for put options.

Figure B1: Regression Summary of Call Options

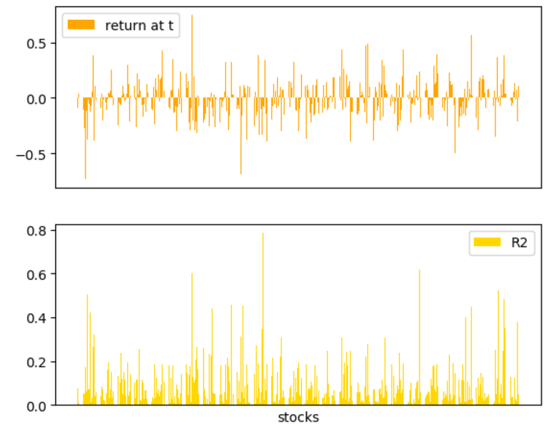


PANEL 1 PANEL 2





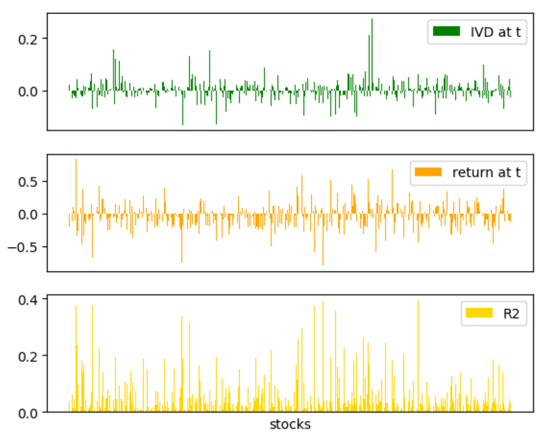
PANEL 3



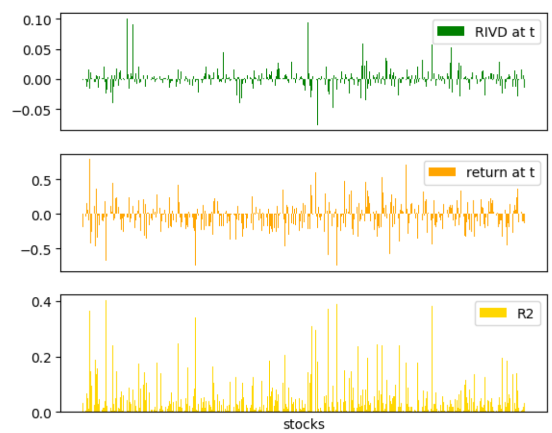
PANEL 4

Panel 1 reports the estimated coefficients and R^2 of equation (21) for each stock; Panel 2 reports those of equation (22); Panel 3 reports those of equation (23); and Panel 4 reports those of equation (24). The green bars are the parameter β_1; the orange bars are parameter β_2; the orange bars are parameter β_3; and the gold bars are the R^2. The x-axis is all the stocks with option data.

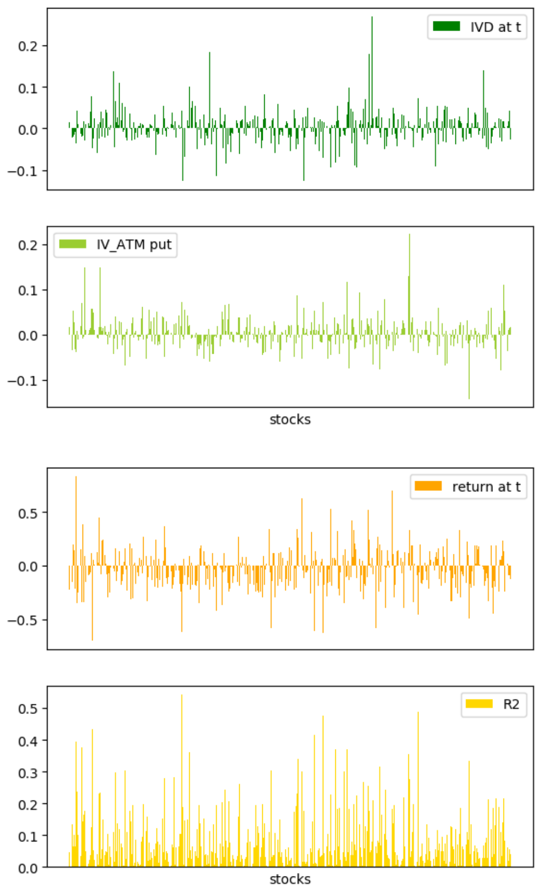
Figure B2: Regression Summary of Put Options



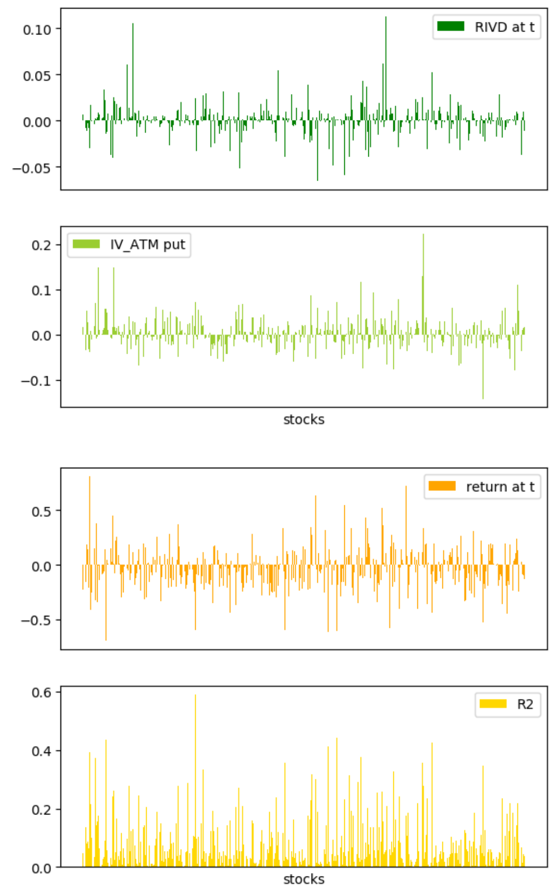
PANEL 1



PANEL 2



PANEL 3



PANEL 4

Panel 1 reports the estimated coefficients and R^2 of equation (25) for each stock; Panel 2 reports those of equation (26); Panel 3 reports those of equation (27); and Panel 4 reports those of equation (28). The green bars are the parameter β_1; the orange bars are parameter β_2; the orange bars are parameter β_3; and the gold bars are the R^2.

B.2 XGBoost Model Principle

Assume y is the actual value of target data, which is momentum return at time $t+1$ in our dataset, x is input features, $\hat{y} = f(x)$ is the predicted value using XGBoost, s is the current base learner index, Q is the current base learner leaf node number, I_j is the sample index set of the jth node, andω_j is the weight of jth node. If the sample x is on the j th node, then $f(x) = \omega_j$. The objective function of XGBoost contains a loss function $L(y, \hat{y})$ and regularization $\Omega(f_s) = \gamma Q + \frac{1}{2}\lambda \sum_{j=1}^{Q} \omega_j^2$, that is

$$\begin{aligned} Obj_s &= \sum_{i=1}^{m} L(y^{(i)}, \hat{y}^{(i)}) + \Omega(f_s) \\ &= \sum_{i=1}^{m} \left[L\left(y^{(i)}, \hat{y}_{s-1}^{(i)} + f_s\left(x^{(i)}\right)\right)\right] + \gamma Q + \frac{1}{2}\lambda \sum_{j=1}^{Q} \omega_j^2 \end{aligned}$$

where m is the total number of samples. Assume that the first-order derivative and second-order derivative of loss function are $g_i = \frac{\partial L(y^{(i)}, \hat{y}_{s-1}^{(i)})}{\partial \hat{y}_{s-1}^{(i)}}$, $h_i = \frac{\partial^2 L(y^{(i)}, \hat{y}_{s-1}^{(i)})}{\partial \left(\hat{y}_{s-1}^{(i)}\right)^2}$, applying second-order Taylor expansion on the Obj_s,

$$Obj_s = \sum_{i=1}^{m} \left[L\left(y^{(i)}, \hat{y}_{s-1}^{(i)}\right) + g_i f_s\left(x^{(i)}\right) + \frac{1}{2} h_i \left(f_s\left(x^{(i)}\right)\right)^2 \right] + \gamma Q + \frac{1}{2}\lambda \sum_{j=1}^{Q} \omega_j^2$$

then combining the samples on each nodes and removing the constant, the objective funtion is

$$Obj_s = \sum_{j=1}^{Q} \left[\left(\sum_{i \in I_j} g_i \right) \omega_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) \omega_j^2 \right] + \gamma Q$$

Therefore, for each tree structure, the weight of its leaf nodes is

$$\begin{aligned} \frac{\partial Obj_s}{\partial \omega_j} &= \left(\sum_{i \in I_j} g_i \right) + \left(\sum_{i \in I_j} h_i + \lambda \right) \omega_j = 0 \\ \omega_j &= -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \end{aligned}$$

and its objective function is

$$Obj_s = -\frac{1}{2} \sum_{j=1}^{Q} \frac{\left(\sum_{i \in I_j} g_i\right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma Q$$

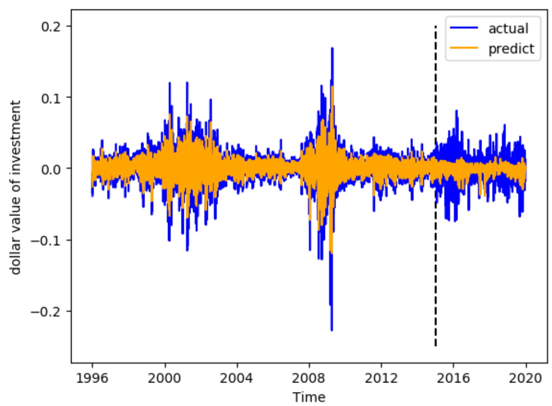
By calculating the Obj_s of each tree and comparing Obj_s for all tree structures, the tree structure with minimize Obj_s is selected as the new learner at s.

B.3 Prediction Results of RF and XGBoost

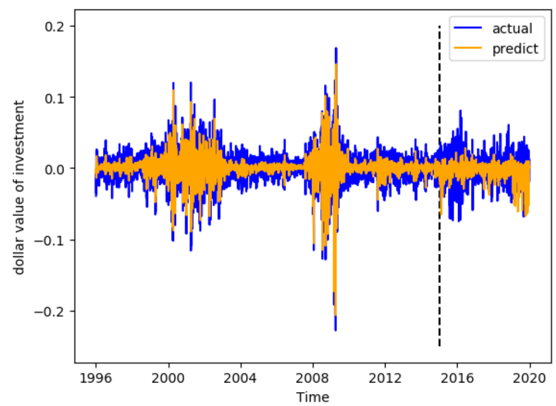
Panel 1 and Panel 2 in Figure B3 report the comparison of actual momentum return and predicted momentum return by RF model and XGBoost model. In addition, Panel 3 and Panel 4 in Figure B3 present the comparison of actual dynamic momentum return and predicted dynamic momentum return by RF model and XGBoost model.

Figure B4 compares the feature importance of RF and XGBoost on the first and second datasets respectively. In RF model on the first dataset, 5 features with the highest weight are historical momentum return, the put option IV difference of Brown & Brown, Inc. (BRO), the call option IV difference of CSX Corporation (CSX), the call option IV difference of MGM Resorts International (MGM), and the call option IV difference of Massey Energy Company (MEE); which are different from those in XGBoost model of the first dataset, the put option IV difference of Abbott Laboratories (ABT), the put option IV difference of Brown & Brown, Inc. (BRO), the call option IV difference of Deere & Company (DE), the put option IV difference of Principal Financial Group Inc. (PFG) and the call option IV difference of LKQ Corporation (LKQX). For the second dataset with the dynamic momentum return, three of the five features with the highest weight are the same as RF model on the first dataset, with the other two features being option IV difference and option IV difference in level of Alaska Air Group, Inc.. It should be remarked that the 5 features with the highest weight of XGBoost on the second dataset are completely different than those on the first dataset.

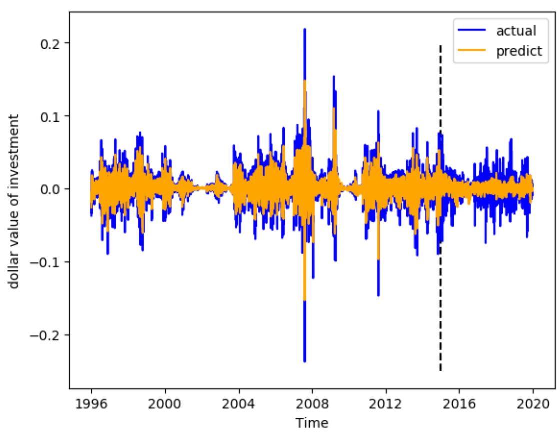
Figure B3: Return Comparison by RF and XGBoost



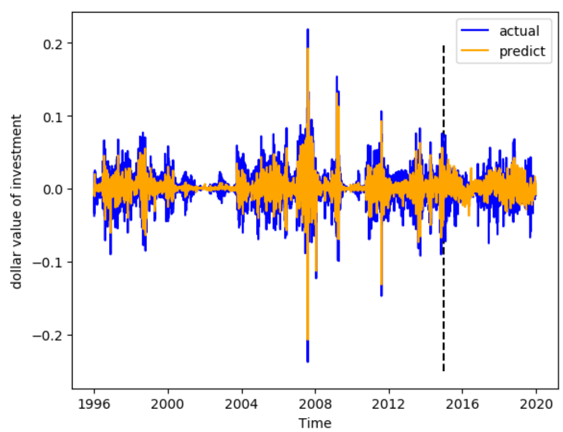
PANEL 1



PANEL 2



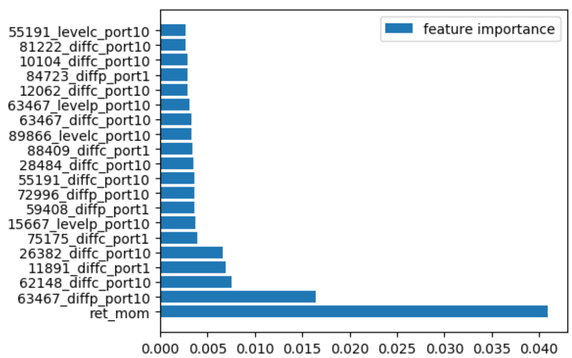
PANEL 3



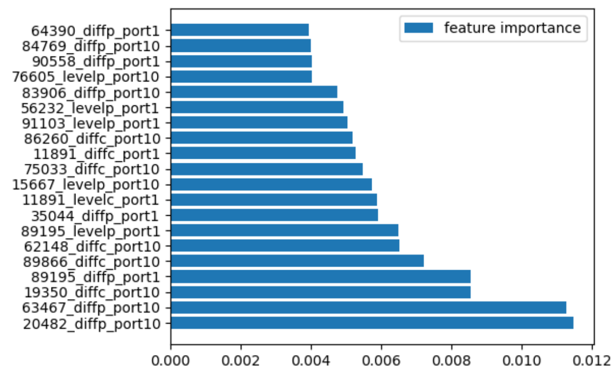
PANEL 4

Panel 1 reports the actual momentum return and predicted momentum return by RF, Panel 2 reports those by XGBoost; Panel 3 reports the actual dynamic momentum return and predicted dynamic momentum return by RF, Panel 4 reports those by XGBoost.

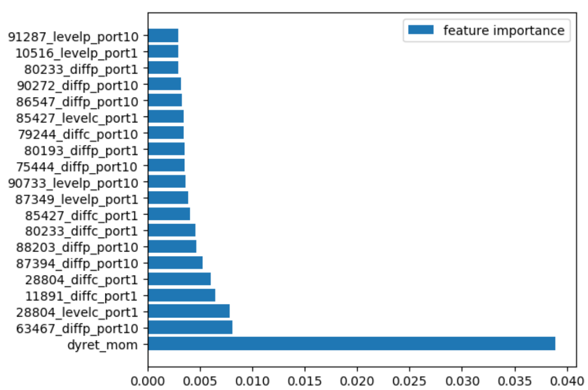
Figure B4: Feature Importance



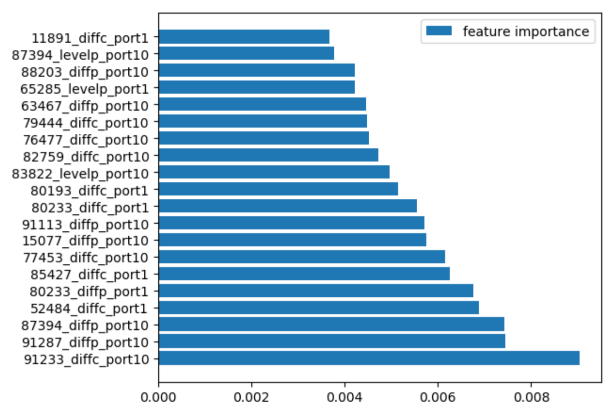
PANEL 1



PANEL 2



PANEL 3



PANEL 4

Panel 1 reports 20 features with highest weight of RF model on the first dataset; Panel 2 reports those of XGBoost model on the first dataset; Panel 3 reports those of RF model on the second dataset; and Panel 4 reports those of XGBoost model on the second dataset.

C The Dynamic Strategy Based on Option Information on Past Winners and Past Losers

Figure C1 presents the performance of this option information strategy on past winner portfolio and past loser portfolio respectively. The purple curve is the cumulative return of past winner portfolio; the deep pink curve is that of the option information strategy on past winners; the grey curve is that of past loser portfolio; and the orange curve is that of the option information strategy on past losers. In Panel 1, we assume that 1$ is invested on Jan. 1996. While Panel 2 compares the accomplishments over the test set.

Figure C1: Port 1 and Port 10 Analysis

PANEL 1 PANEL 2

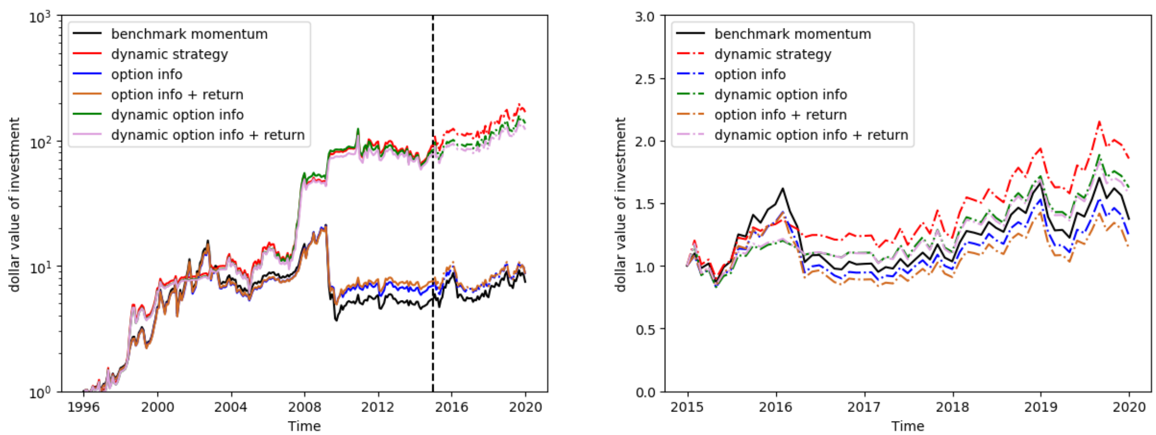
This graph presents the cumulative return of the past winner portfolio (the purple curve), the option information strategy on past winners (the deep pink curve), the past loser portfolio (the grey curve) and the option information strategy on past losers (the orange curve). The thresholds x and y for stop trading are set on training set (from Jan. 1996 to Dec. 2014). Panel 1 assumes that 1$ is invested in Jan. 1996, while Panel 2 assumes that the 1$ investment is on Jan. 2015, the beginning of test set.

D Two Methods to Test the Stop Trading Signals

D.1 The First Method: Option Threshold + Stock Return

As illustrated in section 6.4, in the first method, we consider using option thresholds and underlying return together as the signal to stop trading a stock. In other words, when investors hold a long position in past winner stocks, the option information strategy will sell the stocks that their returns at day t is lower than the risk-free rate at day t and their put option RIVD is higher than their static thresholds x; while when they hold a short position in past loser stocks, the option information strategy will buy the stocks that their returns at day t is higher than the risk-free rate at day t and their call option RIVD is higher than their static thresholds y. The static thresholds x and y are set on the training dataset (from Jan. 1996 to Dec. 2014).

Figure D1: Comparison of Different Strategies



PANEL 1 PANEL 2

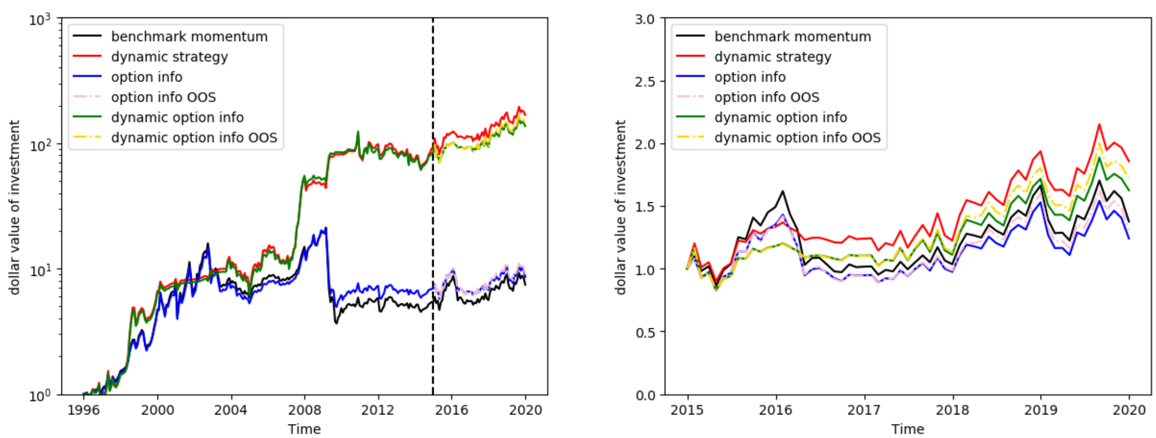
This graph presents the cumulative return of the benchmark momentum (the black curve), dynamic momentum (the red curve), option information strategy with static thresholds (the navy blue curve), dynamic option information strategy with static thresholds (the green curve), option information strategy that uses option information and underlying return at day t together as thresholds (the chocolate curve) and the dynamic option information strategy that uses option information and underlying return at day t together as thresholds (the plum curve). Panel 1 assumes that 1$ is invested in Jan. 1996, while Panel 2 assumes that the 1$ investment is on Jan. 2015, the beginning of test set.

Figure D1 presents the performance of the six strategies: benchmark momentum (the black curve), dynamic momentum (the red curve), option information strategy with static thresholds (the navy blue curve), dynamic option information strategy with static thresholds (the green curve), option information strategy that uses option information and underlying return at day t together as thresholds (the chocolate curve), dynamic option information strategy that uses option information and underlying return at day t together as thresholds (the plum curve). Panel 1 assumes that 1$ is invested at Jan. 1996, while Panel 2 presents the performance over the test set.

D.2 The Second Method: Dynamic Threshold

In the second method to test the stop trading signals, the dynamic thresholds that update the thresholds every six months over time are considered. Figure D2 compares the performance of this option information strategy with static thresholds on benchmark momentum portfolio (the navy blue curve) and on dynamic momentum portfolio (the green curve), and dynamic thresholds on benchmark momentum portfolio (the light pink dash curve) and on dynamic momentum portfolio (the gold dash curve). Panel 1 assumes that 1$ is invested at the beginning of whole dataset, Jan. 1996; while Panel 2 assumes that 1$ is invested at the beginning of test set, Jan. 2015.

Figure D2: Comparison of Different Strategies



PANEL 1 PANEL 2

This graph presents the cumulative return of the benchmark momentum (the black curve), dynamic momentum (the red curve), option information strategy with static thresholds (the navy blue curve), dynamic option information strategy with static thresholds (the green curve), option information strategy with dynamic thresholds (the light pink curve) and option information strategy with dynamic thresholds (the gold curve). Panel 1 assumes that 1$ is invested in Jan. 1996, while Panel 2 assumes that the 1$ investment is on Jan. 2015, the beginning of test set.