No. 2307 [EN]

IMI Working Paper

How Does the Volatility-Timing Strategy Perform in Mutual Funds Portfolios?

Yin Zhida, Jiang Jilin and Qian Zongxin

INTERNATIONAL MONETARY INSTITUTE

For further information, please visit
http://www.imi.ruc.edu.cn/en/

Weibo

WeChat

IMI Working Paper No. 2307 [EN]

How Does the Volatility-Timing Strategy Perform in Mutual Funds Portfolios?*

By* Yin Zhida, Jiang Jilin and Qian Zongxin

March 2023

Abstract

Literature suggests that a volatility-timing strategy improves the performance of factor portfolios in the stock market and currency carry trade. This paper shows that the performance of this strategy is mixed when applied to mutual fund portfolios. More specifically, its performance not only depends on the investment style of the mutual funds but also the time periods when it is applied.

JEL Classification: G11, G23
Keywords: mutual fund, volatility-timing, factor model, skew

---

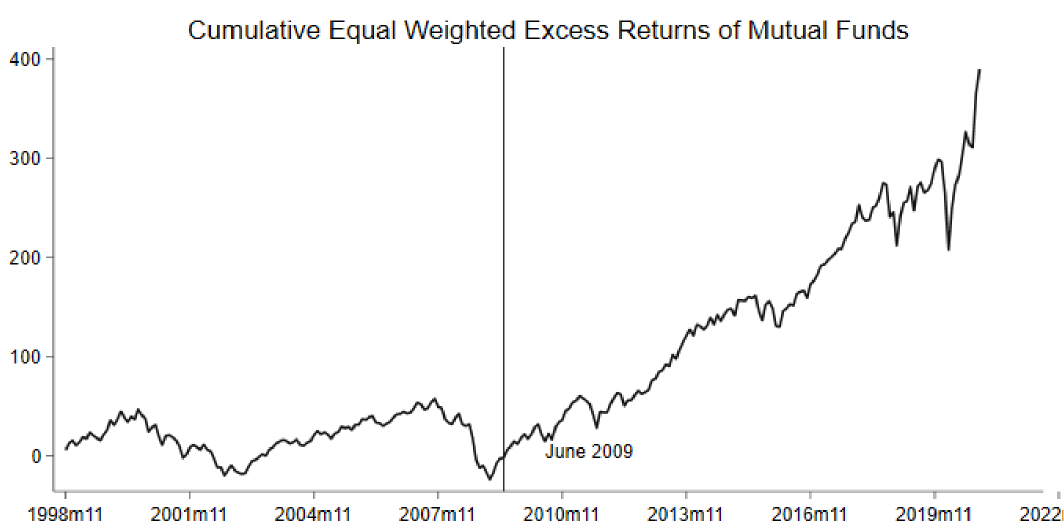
* Published in International Review of Finance 2022, 1-16.
* Yin Zhida, Renmin University of China. Jiang Jilin, Global Decision Science Department, American Express. Qian Zongxin, Senior Research Fellow of IMI, School of Finance, Renmin University of China.

1. Introduction

Recent studies on stock and currency market find that scaling the investment positions by past volatility from time to time could help reduce the uncertainty and improve performance. For example, Barroso and Santa-Clara (2015) show that by scaling the position of momentum strategy on the stock market, the Sharpe ratios are doubled and huge losses are avoided during the stock market crashes. Moreira and Muir (2017) further demonstrate that scaling positions by volatility also increase Sharpe ratios of other equities portfolios and the currency carry trade. Therefore, it is interesting to know whether the same risk management by volatility timing (VT) could improve the performance of investments in mutual funds. Wang et al. (2021) apply the VT technique to mutual funds and find that VT also improve the performance of mutual fund portfolios.

In our study, we extend the analysis of Wang et al. (2021) in two aspects. First, we compare the performance of VT in mutual fund portfolios with different investment styles. More specifically, we follow Hunter *et al*. (2014) to divide the mutual funds into different active peer groups by investment styles, small capitalization (cap) core, small cap value, small cap growth, mid cap growth, large cap core, large cap value, large cap growth, and compare the performance of VT for those different groups.[1] Second, we compare the performance of VT in different sample periods. Figure 1 plots the cumulative excess return of equal weighted equity mutual funds from January 1998 to December 2020. Obviously, the trend of the cumulative mutual fund excess return changes after June 2009. More specifically, the return grows faster. A formal structural break test (Chow test) confirms that the excess return of the equal weighted equity mutual funds portfolio experienced a mean shift in June 2009 (the test p value is 0.042). Hence, it is interesting to compare the performance of VT before and after June 2009.

Figure 1 Cumulative equal-weighted excess returns of equity mutual funds



Source: WRDS.

[1] Hunter *et al* (2014) also consider the mid cap value and mid cap core groups. However, there are insufficient observations of those two groups for some of our OOS tests, so we do not include those groups in our analysis.

We follow the methodology by Barroso and Santa-Clara (2015) to dynamically scale the position in mutual fund investments by the return volatility of the past six-month and compare the excess return series of the scaled portfolios of mutual funds with the unscaled portfolios. The results are mixed. In the period from September 1998 to June 2009, the volatility timing (VT) technique significantly raises the Sharpe ratios of mutual fund portfolios which invest in growth stocks. It also raises Fama-French five-factor alphas of those portfolios. Return distributions of volatility-managed portfolios are also less left-skewed than the original portfolios. However, the performance of VT is mixed when applied to mutual fund portfolios which is balanced in growth and value stocks or focused on value stocks. Particularly, VT increases Sharpe ratio and skewness in these fund portfolios and reduces their kurtosis, but VT also reduces the Fama-French five-factor alphas of mutual fund portfolios which is balanced or focus on value stocks. More importantly, VT fails in the period from July 2009 to December 2020. It reduces the Sharpe ratios and Fama-French five-factor alphas of all mutual fund portfolios while raises the crash risk.

Moreira and Muir (2017) suggest that the fact that volatility timing improves portfolio return challenges the risk-based theory of asset pricing. Economically speaking, taking on more risk when volatility is high reduces investor utility, therefore, the volatility timing strategy which reduces risk-taking should require lower risk premium. Our tests using the mutual fund portfolio somehow reduces the puzzle posed by Moreira and Muir (2017). In some mutual fund portfolio groups, volatility timing does decrease the Fama-French five-factor alphas. Moreover, VT fails to improve performance in more recent years. What drives the change in VT performance is an interesting question for future theoretical research.

To further explore the reasons behind this difference, we investigate the out-of-sample (OOS) exposures to factor risks of the mutual fund portfolios. The exposures to factor risks significantly differ between the scaled portfolios and the unscaled ones.

We find that the changes in the exposure to market risk brought about by scaling tend to increase the Sharpe ratios of the growth stocks-focused mutual fund portfolios before June 2009. After June 2009, this effect weakens. Actually, the ability of VT to assume market risk in proper time is worse for all fund portfolios after June 2009.

Before June 2009, the changes in the exposure to Fama-French SMB factor and HML factor brought about by scaling increase the Sharpe ratios of all mutual fund portfolios, except for the SCG group in which scaling improves its Sharpe ratio through changing loadings to SMB factors, but does not through the HML factor. After June 2009, the effect remains in most mutual fund portfolios but not in the small-cap stocks-focused mutual fund portfolio for SMB factor and not in the LCV group for HML factor.

The changes in the exposure to the Fama-French RMW factor only increase Sharpe ratios in LCV group before June 2009, and LCG and MCG groups after June 2009. While changes in the exposure to the Fama-French CMA factor brought about by scaling do not improve fund portfolio performance before June 2009, they improve the Sharpe ratios of all mutual fund portfolios after June 2009.

Those results suggest that the contrasting performance of VT before and after June 2009 is affected by its time-varying ability to assume different factor risks. In particular, the poorer ability of VT to assume market risk after June 2009 is an important reason why it could not improve the Sharpe ratios of the mutual fund portfolios.

Wang and Yan (2021) propose another way to understand the performance of the VT technique. First, following Moreira and Muir (2017), they regress the volatility-managed portfolio return on the unmanaged return to obtain the alpha. Second, they decompose the alpha into a volatility timing (VT^*) component and a return timing (RT) component. The VT^* component is positive if volatility exhibts positive serial correlation. The RT component is positive if lagged volatility is negatively correlated to futurn returns. Interestingly, the return timing ability of the volatiltiy management technique is superior in the sample before June 2009, but contributes negatively to the portfolio return after June 2009. This result suggests that past volatility is negatively correlated to future returns before June 2009, but positively correlated to future returns after June 2009. The VT^* component is negative in many cases before June 2009 while postive after June 2009. However, the RT component dominates both before and after June 2009, which explains why the volatility management strategy performs better before June 2009.

The rest of this paper proceeds as follows: Section 2 describes our empiric model; Section 3 outlines the dataset in this study; Section 4 illustrates and discusses the empirical results; Section 5 concludes the paper.

2. Empiric model: an APB-augmented five-factor model

Following the literature (Carhart, 1997; Hunter *et al*., 2014), we construct mutual funds portfolios based on the ranking of alphas from an estimated factor model. Jordan and Riley (2015) documented that the Fama and French (2015) five-factor model has an advantage of eliminating the volatility anomaly to avoid mis-measurement of fund alphas. However, there is potentially another source of mis-measurement that is not addressed by Fama and French (2015): the tendency of mutual funds to use correlated investment strategies. This phenomenon typically leads to correlated residuals from commonly used factor models and reduces estimation efficiency. A study by Hunter *et al*. (2014) proposes an active peer benchmark (APB) to eliminate the residual correlation. It shows that the APB-augmented Carhart four-factor model (Carhart, 1997) generates high out-of-sample alphas. In this study, we combine the Fama-French five-factor model with the APB so that our benchmark is free from measurement errors due to both the volatility anomaly and correlated residuals. Doing so can not only improve estimation efficiency by correcting the cross-sectional correlation in the error term but also eliminate the estimation bias from the omitted variable bias in the four-factor model. This combination is important because we do not want our results to come from econometric problems caused by model specification errors.

Specifically, this study follows Hunter et al. (2014) to estimate the APB of a certain fund group by the following equation:

$$r_{APB,t} = \alpha_{APB,t} + \beta_{APB,t,rmrf} r_{rmrf,t} + \beta_{APB,t,SMB} SMB_t + \beta_{APB,t,HML} HML_t + \beta_{APB,t,RMW} RMW_t + \beta_{APB,t,CMA} CMA_t + \varepsilon_{APB,t} \quad (1)$$

where, $r_{APB,t}$ is the averaged excess return of the active peer group of funds in period t; $r_{rmrf.t}$, SMB_t, HML_t, RMW_t and CMA_t are the five factors in Fama and French (2015)[2]. All the betas in Equation (1) represent the corresponding (time-varying) factor loadings, $\alpha_{APB,t}$ is a group intercept and $\varepsilon_{APB,t}$ is the error term (the APB factor). We then use the APB factor to revise the five-factor model as follows:

$$r_{i,t} = \alpha_{i,t} + \beta_{i,t,rmrf} r_{rmrf,t} + \beta_{i,t,SMB} SMB_t + \beta_{i,t,HML} HML_t + \beta_{i,t,RMW} RMW_t + \beta_{i,t,CMA} CMA_t + \lambda_{i,t} \varepsilon_{APB_i,t} + e_{i,t} \quad (2)$$

where $r_{i,t}$ is an excess return of fund i at time t. According to Hunter *et al*. (2014), the term $\lambda_{i,t}\varepsilon_{APBi,t}$ helps eliminate commonalities in risk-taking by funds in the same active peer

[2] The five factors in Fama and French (2015) are: the market ($r_{rmrf.t}$), size (SMB_t), book-to-market (HML_t), profitability (RMW_t) and investment (CMA_t) factors.

benchmark group. Hence the APB adjustment as shown by Equation (2) should improve the estimation efficiency relative to a standard five-factor model.

Hunter *et al*. (2014) suggest that forming mutual fund portfolios on the basis of their investment styles and in-sample alphas generate superior out-of-sample (OOS) performance. More specifically, their strategy is to choose the quartile of mutual funds which has the highest in-sample alphas from each category of investment style. In this paper, we use their strategy to form mutual fund portfolios and evaluate the performance of the VT technique. Section 4.4.1 shows that our main results remain if we form fund portfolios using other quartile of mutual funds.

3. Dataset and fund categorization

We collect a monthly dataset from September 1998 to December 2020 for model estimation. The excess fund return is defined as the return of dividend-adjusted net asset value (NAV) minus the rate of six-month Treasury bills. Following the study of Hunter *et al*. (2014), we add back 1/12 of the annual expense ratio to the NAV return to obtain the prior-expense fund return. Time series of dividend-adjusted NAV returns and the expense ratios of mutual funds are obtained from the Center for Research in Security Prices (CRSP) database. Following Hunter *et al*. (2014), we only include no-load mutual fund share-classes. Table 1 reports the summary statistics on fund characteristics. The five factors of Fama and French (2015) are taken from French's webpage.[3] All funds are categorized using an approach proposed by Hunter *et al*. (2014) which is based on the best-fitting benchmark of Cremers and Petajisto (2009). Specifically, funds in our sample are categorized into seven groups: small-cap core (SCC), small-cap value (SCV), small-cap growth (SCG), mid-cap growth (MCG), large-cap core (LCC), large-cap value (LCV), and large-cap growth (LCG).[4]

Table 1 Summary statistics on fund characteristics

This table report summary statistics on fund characteristics. The sample period starts from 1998 and ends by 2020.

Characteristics	Mean	Std. Dev	10th percentile	Median	90th percentile
Expense Ratio as of Fiscal Year-End (%)	1.15	0.52	0.64	1.12	1.68
Fund Turnover Ratio (%)	73.23	80.24	13.00	54.00	151.00
Total Net Assets as of Month End (millions)	1,935.01	6,691.79	29.10	330.40	3,783.10

4. Empirical results

As discussed before, past literature (Barroso and Santa-Clara, 2015; Moreira and Muir, 2017) has documented that the technique of scaling investment positions in the equity market (and the currency market) by past volatility reduces the OOS volatility and increases the Sharpe ratio at the same time. It also helps reduce crash risk. In this subsection, we investigate whether this volatility timing (VT) technique works well in terms of improving performance (Sharpe ratios)

3 http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
4 Appendix A describes how those groups are formed.

while mitigating crash risk. Therefore, a volatility-managed[5] OOS holding-period return of a fund portfolio is defined as follows:

$$r_{i,t}^{*} = \frac{k}{\sigma_t} r_{i,t} \tag{3}$$

where σ_t is the sample standard deviation of the portfolio returns in the past 125 trading days (six calendar months). The parameter k is a target level of standard deviation set by the investors. Following the literature (Moreira and Muir, 2017; Wang *et al*. 2021), we impose a value for k so that the portfolios tuned by the volatility timing (VT), which we shall call volatility-managed portfolios, have the same full sample volatility as the portfolios absent of VT (the original portfolios).

4.1 Volatility scaled vs. unscaled fund portfolio performance

4.1.1 Results before June 2009

Empirical results for the sample period from September 1998 to June 2009 are listed in Panel A of Table 2. The columns labeled "VT" in Table 2 reports the main characteristics of the returns of the volatility-managed portfolios. In each "VT" column, we highlight the cases in which the volatility timing (VT) improves the performance by assigning # as upper notations to the numbers. In general, the VT technique improves the portfolio performance in large-cap (LC) and mid-cap (MC) groups. It is clear in Panel A of Table 2 that all volatility-managed portfolios within the Growth groups, compared with their original counterparts, enjoy higher Sharpe ratio, larger skewness, smaller kurtosis, and higher mean alpha. In particular, the VT procedure turns the Sharpe ratios from negative or zero to positive numbers. Core groups and Value groups also get better return distribution characteristics after volatility-timing, though their mean alphas decrease. The phenomenon indicates that in these groups volatility-timing works through market-timing. Specifically, time-varying factor loadings induced by scaling bring higher Sharpe ratio, larger skewness, smaller kurtosis to these groups. The improvement brought by volatility-timing is smaller in the small-cap groups compared with large-cap and mid-cap groups. While in all the large-cap and mid-cap groups, the rise of Sharpe ratio brought by scaling is 0.5, the enhancement of Sharpe ratio is only 0.2 in SCC, SCG group, and 0.4 in SCV group. Besides, in SCC and SCG groups the enhancement of skewness and decrement of kurtosis caused by volatility-timing are also lower than in other groups compared with the original ones.

Figure 2 presents the cumulative returns of the unscaled and scaled portfolios during the subprime crisis. VT helps reduce the loss at the trough for all fund portfolios. Its loss-reducing effects are particularly strong for those portfolios which heavily invest in value stocks. The loss-reducing effects are weaker for those portfolios which have a focus on growth stocks.

Table 2 Performance of scale versus unscaled mutual fund portfolios (baseline)

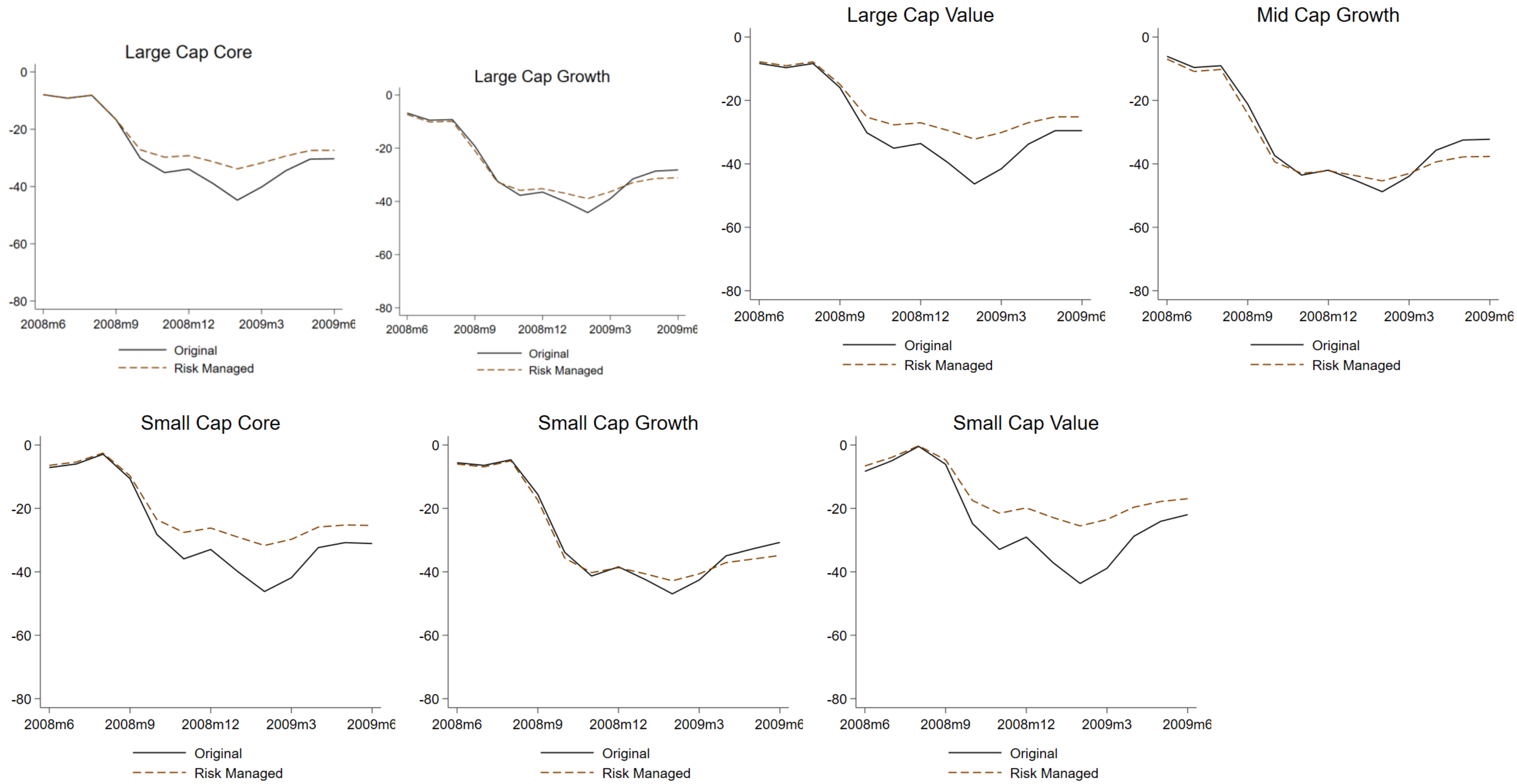
Numbers labeled "Original" ("VT") are performance indicators of the portfolios without (with) volatility timings. #s indicate that VT contributes to improve the performance. Fund portfolios are sorted into 7 investment style groups: Small-cap core (SCC); Small-cap value (SCV); Small-cap growth (SCG), Mid-cap growth (MCG), Large-cap core (LCC), Large-cap value (LCV), and Large-cap growth (LCG). And in each group of investment style, we only report the portfolios

[5] This study mainly focuses on the volatility timing (VT) as a technique to manage the portfolio risk, so the "volatility-managed" portfolios and portfolios with "VT" are used interchangeably in all discussions.

with the **highest in-sample t-value of alphas** in that group. The sample starts from September 1998 and ends by December 2020.

Portfolio	Sharpe Ratio		Skewness		Kurtosis		Mean Alpha	
	Original	VT	Original	VT	Original	VT	Original	VT
	Panel A. September 1998-June 2009							
LCC	-0.06	-0.01#	-0.63	-0.39#	3.71	2.57#	-0.02	-0.05
LCG	-0.07	-0.02#	-0.47	-0.31#	3.39	2.58#	0.30	0.39#
LCV	0.02	0.07#	-0.58	-0.41#	4.30	2.72#	0.10	0.05
MCG	-0.04	0.01#	-0.67	-0.47#	4.06	2.94#	0.23	0.27#
SCC	0.10	0.12#	-0.61	-0.58#	4.38	3.62#	0.22	0.15
SCG	0.00	0.02#	-0.47	-0.45#	3.74	3.06#	-0.14	-0.07#
SCV	0.12	0.16#	-0.70	-0.45#	4.72	3.46#	0.07	-0.01
	Panel B. July 2009-December 2020							
LCC	0.29	0.25	-0.41	-0.93	4.28	5.02	-0.02	-0.08
LCG	0.34	0.31	-0.18	-0.43	3.37	4.32	0.17	0.06
LCV	0.24	0.18	-0.41	-1.17	4.98	5.86	-0.04	-0.13
MCG	0.37	0.31	-0.34	-1.12	4.66	6.31	0.07	-0.14
SCC	0.19	0.10	-0.46	-1.46	5.37	8.10	-0.09	-0.35
SCG	0.27	0.21	-0.37	-1.44	4.93	7.66	0.27	0.14
SCV	0.20	0.11	-0.73	-1.48	7.11	7.94	0.09	-0.24

Figure 2 Cumulative mutual fund portfolio returns by investment style (June 2008-June 2009)



Notes: Authors' calculations, %.

4.1.2 Results after June 2009

Empirical results for the sample period from July 2009 to December 2020 are listed in Panel B of Table 2. Contrasting to the results in Table 2 and previous literature, VT does not work well.[6] It not only reduces Sharpe ratios and alphas of all fund categories, but also increases the crash risk. The poor performance is especially dramatic in the fund portfolios which focus on small-cap funds. The Sharpe ratio of the SCC group is reduced by 0.09. Its mean alpha is reduced by 0.26. Its skewness is decreased by 1 to -1.46. The Sharpe ratio of the SCV group is reduced by 0.09. Its mean alpha is reduced by 0.33. Its skewness is decreased by 0.75 to -1.48. The Sharpe ratio of the SCG group is reduced by 0.06. Its skewness is decreased by 0.97 to -1.44. Its mean alpha is reduced by 0.13.

4.2 Contributions of factor exposures brought about by VT to changes in the Sharpe ratios

According to Equation (2), apart from the alpha, the other part of the OOS portfolio return is generated by a combination of factor exposures. In order to examine the contribution of each factor loading, we carry out the following experiment. First, we keep the factor loading of that specific factor at the same level as the original portfolios. Second, we set all the other factor loading and the alpha to the values that are equal to their counterparts in the volatility-managed portfolios. The above two steps enable us to build a restricted model. Third, we use this restricted model to simulate the portfolio returns. Finally, we calculate the corresponding Sharpe ratios from the simulated portfolio returns. The difference between the Sharpe ratios of the simulated returns and the volatility-managed portfolios are then compared. A formal description of the above steps can be described as follows:

$$E\left(r_{o,t}\right)=\alpha_{o,t}+\beta_{o,t,rmrf}r_{rmrf,t}+\beta_{o,t,SMB}SMB_t+\beta_{o,t,HML}HML_t+\beta_{o,t,RMW}RMW_t+\beta_{o,t,CMA}CMA_t+\lambda_{o,t}\varepsilon_{APB_1,t} \quad (4)$$

$$E\left(r_{v,t}\right)=\alpha_{v,t}+\beta_{v,t,rmrf}r_{rmrf,t}+\beta_{v,t,SMB}SMB_t+\beta_{v,t,HML}HML_t+\beta_{v,t,RMW}RMW_t+\beta_{v,t,CMA}CMA_t+\lambda_{v,t}\varepsilon_{APB_1,t} \quad (5)$$

where $r_{o,t}$ denotes the original portfolio return and $r_{v,t}$ denotes the volatility-managed portfolio (portfolio adjusted by VT) return and there are restrictions:

$$\beta_{v,t,j}=\beta_{o,t,j} \text{ or } \lambda_{v,t}=\lambda_{o,t} \quad (6)$$

where j stands for different factors (market, SMB, HML, RMW or CMA factors). So, if the restriction as in Equation (6) are imposed to Equation (5), a restricted model can then be constructed as follows[7]:

$$E\left(r_{R,t}\right)=\alpha_{v,t}+\beta_{o,t,rmrf}r_{rmrf,t}+\beta_{v,t,SMB}SMB_t+\beta_{v,t,HML}HML_t+\beta_{v,t,RMW}RMW_t+\beta_{v,t,CMA}CMA_t+\lambda_{v,t}\varepsilon_{APB_1,t} \quad (7)$$

where $r_{R,t}$ is the portfolio excess return of the volatility-managed portfolio by assuming: $\beta_{v,t,rmrf} = \beta_{o,t,rmrf}$. The Sharpe ratio calculated by using the simulated $E(r_{R,t})$ series is denoted by $S_{R,t}$, and the Sharpe ratio of the volatility-managed portfolio is denoted by $S_{v,t}$. The gap between the two ratios, $(S_{R,t} - S_{v,t})$, is then used as an indicator to measure how much does the VT technique contribute to improve the performance, if any, through changing the loading of the market factor. Here, a negative value of $(S_{R,t} - S_{v,t})$ implies that if the VT technique had not changed the loading of the market factor in the original portfolio, the Sharpe ratio would be

[6] Table A1 presents the overall results from September 1998 to December 2020. There, we could see that VT's overall performance is weaker than in the first subsample while stronger than in the second subsample.

[7] Here, for illustration purpose, we restrict the loading of the market factor ($\beta_{v,t,rmrf} = \beta_{o,t,rmrf}$) in this example.

lower. And we can repeat the procedure for every factor loading. All the results are reported in Table 3.

Table 3 The effects of changes in factor loading caused by VT on the Sharpe ratios

The column "Factor loading" indicates that we keep all other parameters at the same level as the volatility-managed portfolios with volatility timing but change the factor loading indicated in that column to the same level as the original portfolios. The numbers outside the parentheses are the simulated Sharpe ratios. The numbers inside the parentheses are the difference between the simulated Sharpe ratios and the Sharpe ratios of the volatility-managed portfolios. Therefore, a negative number in the parentheses suggests that the Sharpe ratio would be lower if the volatility timing had not changed the factor loading indicated in the column "Factor Loading". #s indicate that VT contributes to improve the performance. 7 investment style groups are considered: Small-cap core (SCC); Small-cap value (SCV); Small-cap growth (SCG), Mid-cap growth (MCG), Large-cap core (LCC), Large-cap value (LCV), and Large-cap growth (LCG).

Factor loading	LCC	LCG	LCV	MCG	SCC	SCG	SCV
	Panel A. September 1998-June 2009						
MKT	-0.04	-0.04	0.05	-0.02	0.13	0.00	0.17
	(-0.03)#	(-0.03)#	(-0.02)#	(-0.04)#	(0.01)	(-0.03)#	(0.01)
SMB	-0.03	-0.03	0.03	0.00	0.11	0.02	0.12
	(-0.02)#	(-0.02)#	(-0.04)#	(-0.01)#	(-0.01)#	(0.00)	(-0.04)#
HML	-0.03	-0.04	0.04	0.00	0.09	-0.02	0.10
	(-0.02)#	(-0.02)#	(-0.03)#	(-0.01)#	(-0.03)#	(-0.04)#	(-0.06)#
RMW	-0.01	0.00	0.05	0.01	0.11	0.04	0.16
	(0.00)	(0.02)	(-0.02)#	(0.00)	(0.00)	(0.02)	(0.00)
CMA	0.00	0.00	0.07	0.05	0.17	0.07	0.19
	(0.01)	(0.02)	(0.00)	(0.04)	(0.05)	(0.05)	(0.03)
	Panel B. July 2009-December 2020						
MKT	0.24	0.31	0.19	0.32	0.12	0.20	0.13
	(-0.01)#	(0.00)	(0.00)	(0.02)	(0.02)	(-0.01)#	(0.01)
SMB	0.22	0.28	0.17	0.28	0.10	0.21	0.11
	(-0.03)#	(-0.03)#	(-0.01)#	(-0.03)#	(0.00)	(0.00)	(0.00)
HML	0.22	0.30	0.19	0.29	0.09	0.19	0.10
	(-0.03)#	(-0.01)#	(0.00)	(-0.02)#	(-0.01)#	(-0.02)#	(-0.02)#
RMW	0.26	0.30	0.20	0.28	0.10	0.22	0.12
	(0.01)	(-0.01)#	(0.02)	(-0.03)#	(0.00)	(0.00)	(0.01)
CMA	0.22	0.27	0.15	0.26	0.08	0.17	0.10
	(-0.03)#	(-0.04)#	(-0.03)#	(-0.04)#	(-0.02)#	(-0.04)#	(-0.01)#

Results in Panel A of Table 3 suggest that if we were to restrict the loadings on the market factor to the identical values as in the original portfolios, the simulated Sharpe ratios would be lower in the LCC, LCG, MCG and SCG groups before June 2009. This suggests that VT increases the Sharpe ratios of the all mutual fund portfolios which invest mainly in growth stocks or large-cap stocks before June 2009. Comparing the row labeled "MKT" in Panel A and Panel B of Table 3, it is easy to notice that changes in the exposure to the market risk brought about by VT has weaker ability to increase the Sharpe ratio in all fund categories after June 2009 except for the SCV group. The contrast is most dramatic in the LCG and MCG groups. Changes made to the loading to the market factor by VT are the most important source of increase in the Sharpe ratios of the LCG and MCG groups in the period before June 2009. After June 2009, changes in the exposure to the market risk brought about by VT alone have no impact on the Sharpe ratio of the LCG group. Those changes even reduce the Sharpe ratio of the MCG group. Although those changes still improve the Sharpe ratio of the LCC and SCG groups after June 2009, its impact is much smaller than before June 2009. They also no longer improves the Sharpe ratio of the LCV group.

Changes in the loading to the SMB factor brought about by VT increase the Sharpe ratio for LCC, LCG, LCV, MCG groups both before and after June 2009. They help improve the Sharpe ratio for the SCC and SCV groups before June 2009 but no longer work for these groups through the SMB factor after June 2009. VT increases the Sharpe ratio through changing HML factor loadings in all groups before June 2009, and it also works through the channel to improve the performance of mutual fund portfolios after June 2009 except for the LCV group.

Before June 2009, VT improves the Sharpe ratios of the LCV group through the RMW factor. In the period from July 2009 to December 2020, LCG and MCG groups have a higher Sharpe ratio through RMW after scaling. While VT does not improve the Sharpe ratios through the CMA factor before June 2009, it works through the CMA factor after June 2009.

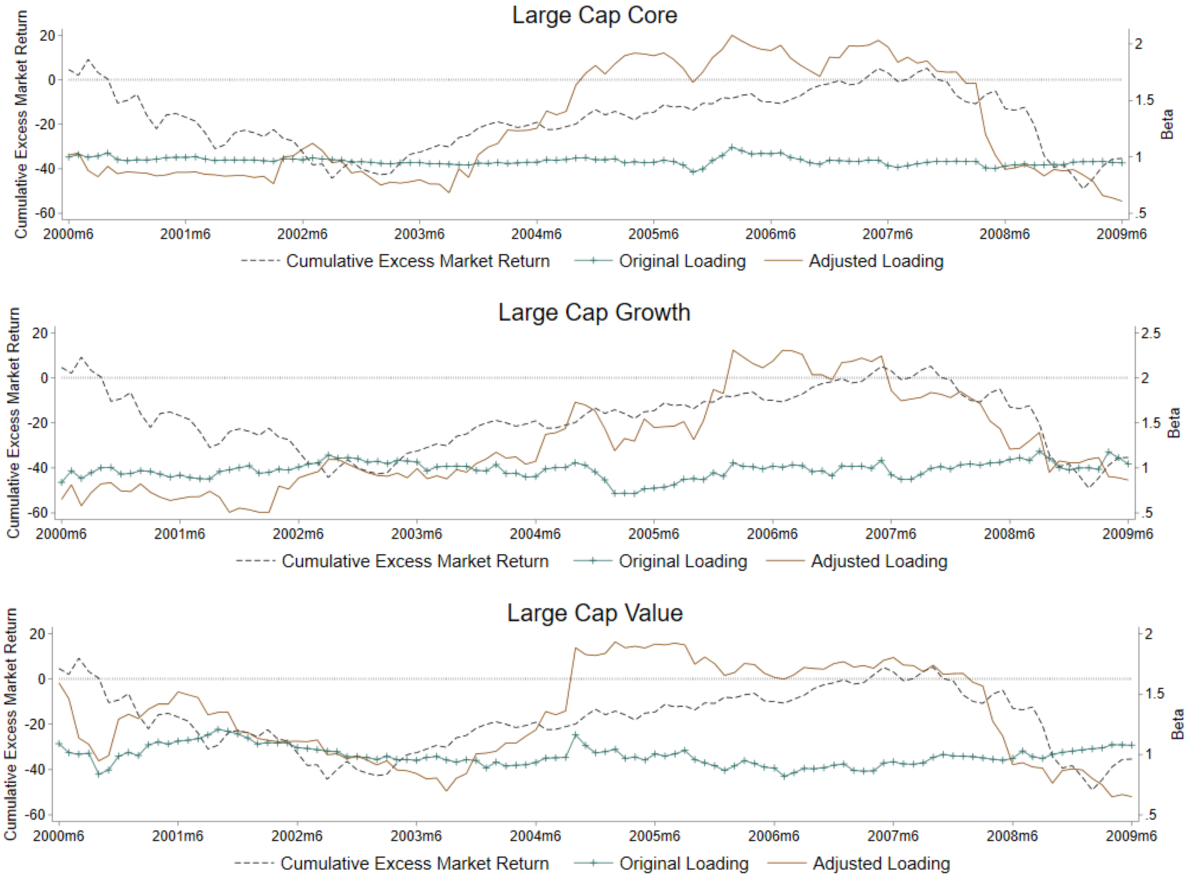
In summary, VT still improves the Sharpe ratios of many mutual fund portfolios, through the SMB, HML, RMW and CMA factor after June 2009. However, the net effects of VT on all portfolios are not positive. This indicates that VT's contrasting effects on the loading of the Market factor is a crucial reason why it stops to improve the Sharpe ratios.

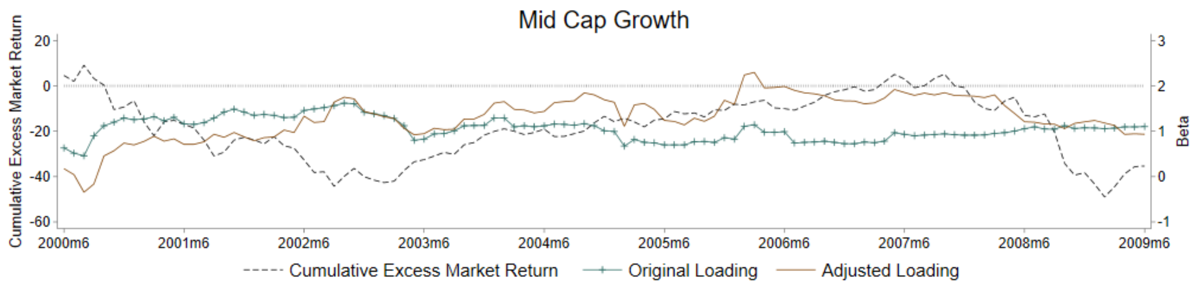
The role of VT's contrasting effects on the loading of the Market factor is also visible from Figure 3 and 4.[8] The dashed lines are the cumulative market returns. The solid lines with no marker are the loading of the volatility-managed portfolios on the market factor. The solid lines with "+" markers are the loading of the original portfolios on the market factor. The first notable pattern is that the loading of the volatility-managed portfolios on the market factor is much more volatile than the original load both before and after June 2009. Second, it is also easy to observe that in our sample before June 2009, cumulative market return is volatile and low, being negative most of the time. The volatility management successfully increases the loading of the portfolios on the market factor when the market return starts to climb after the first half of 2003. It also successfully decreases the loading of the portfolios on the market factor before the market return starts to decline in 2007. Those good timing cases should have contributed to VT's superior performance before June 2009. After June 2009, the cumulative market return has an upward trend. VT keeps the loading of the portfolios on the market factor too low at the beginning of the

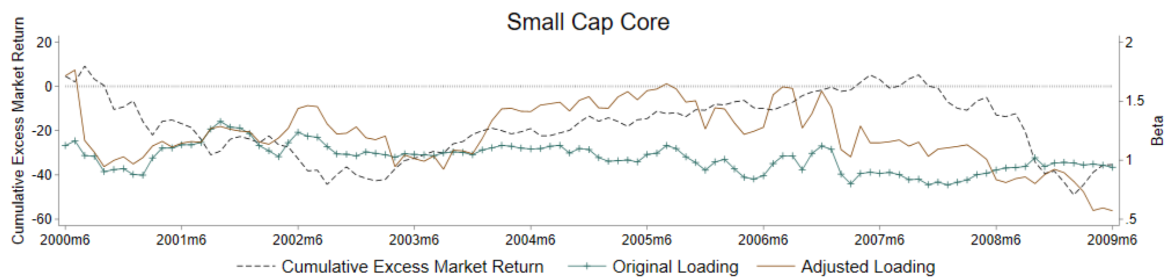
[8] As it is required to have at least 30 observations for a fund in a 3-year rolling window to be included in a fund portfolio and a portfolio is required to have at least five funds in a month to implement our trading strategy, there are some months that the strategy cannot be executed. In Figure 4, if the lines of loadings become dotted, it indicates that in these months the strategy are not implemented due to few observations.

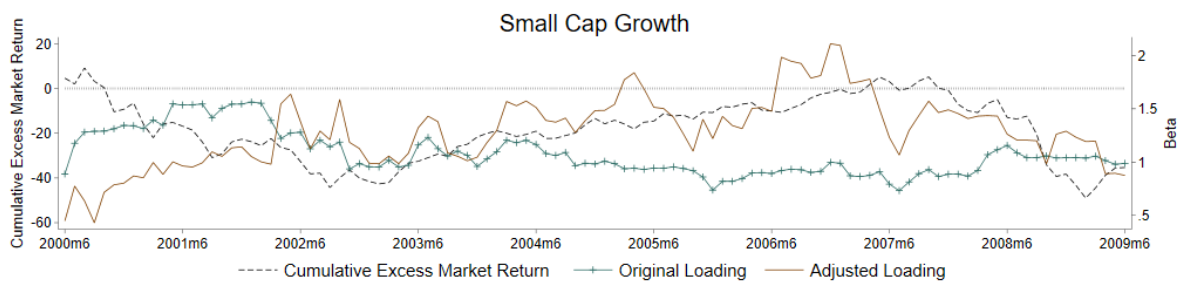
sample period. It also reduces the loading of the portfolios on the market factor at the end of the sample period when the growth of market price accelerates. For those reasons, volatility management gives up a large share of the profits from the original portfolios. This at least partially explains the inferior performance of the VT technique after June 2009.

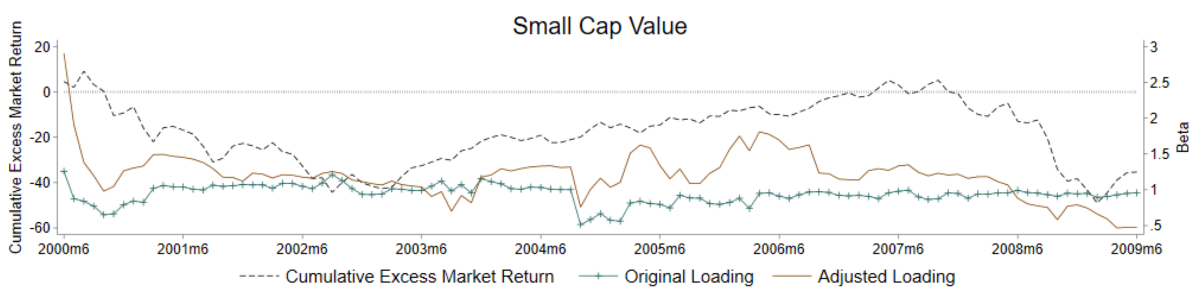
Figure 3 Cumulative returns of the market factor and loading of the scaled and unscaled portfolios on the market factor (before June 2009)





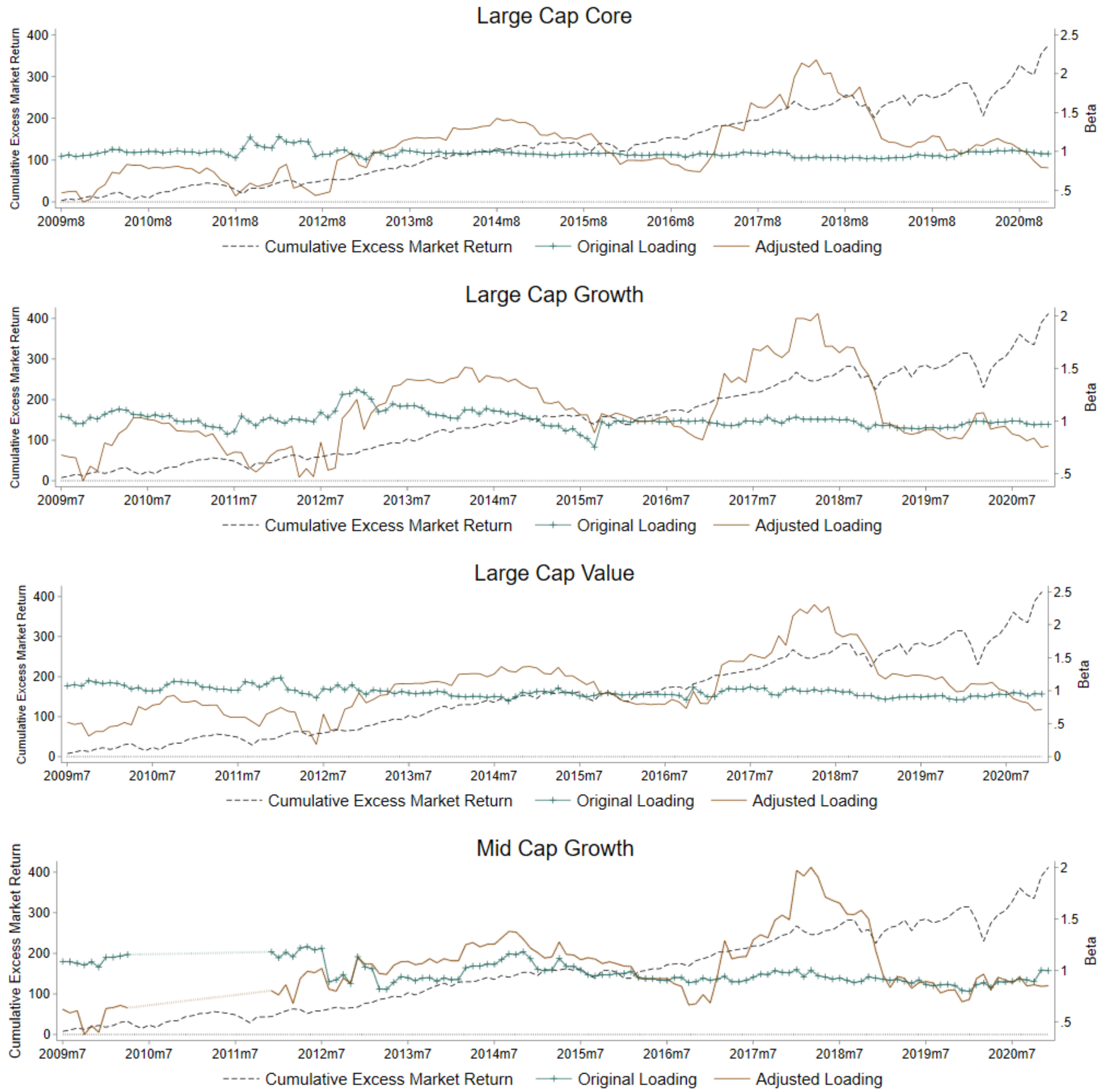


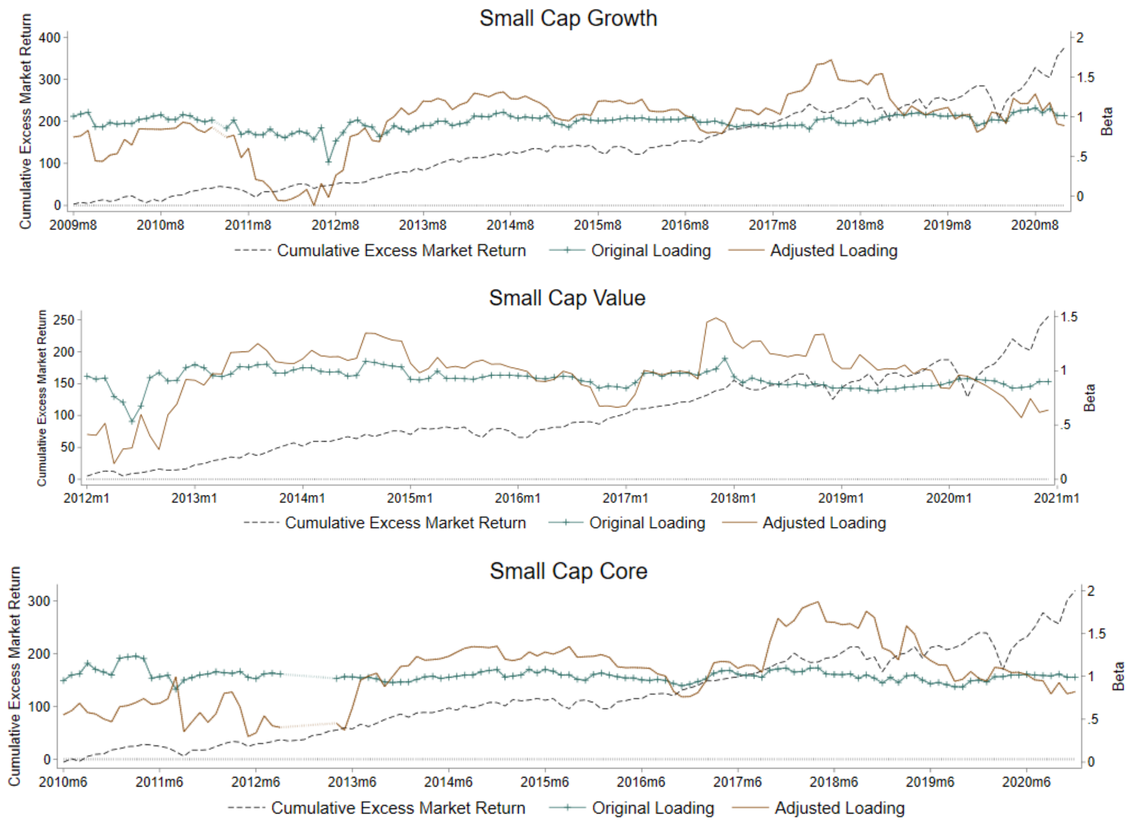




Notes: Authors' calculations, cumulative market return on the left vertical axis (%), loading on the right vertical axis .

Figure 4 Cumulative returns of the market factor and loading of the scaled and unscaled portfolios on the market factor (after June 2009)





Notes: Authors' calculations, cumulative market return on the left vertical axis (%), loading on the right vertical axis . Sample starts from June 2009 except when there are insufficient fund observations to implement the strategy.

4.3 Return timing versus volatility timing

Wang and Yan (2021) suggest that we could better understand the performance of the VT strategy by decomposing the alpha from a spanning regression of the volatiltiy managed portfolio return on the unmanaged portfolio return. More specifically, they run the following regression:

$$r_{v,t} = \alpha_d + \beta_d r_{o,t} + \varepsilon_{d,t}, \tag{8}$$

where $\varepsilon_{d,t}$ is the error term. Wang and Yan (2021) show that

$$\alpha_d = RT + VT^*, \tag{9}$$

where $RT \equiv \left(1+\frac{E^2(r_{o,t})}{Var(r_{o,t})}\right)cov(\frac{k}{\sigma_t}, r_{o,t})$, $VT^* \equiv -\frac{E(r_{o,t})}{Var(r_{o,t})}cov(\frac{k}{\sigma_t}, r^2_{o,t})$. RT and VT^* measure the relatively contribution of return timing and volatility timing ability of the volatility management strategy to the alpha of regression (8), respectively. A positve RT suggests that high past volatility predicts low future return. A positive VT^* suggests that high past volatility predicts high future volatility.

Table 4 reports the decomposition results. It is easy to notice that the alpha of the spanning regression is positive in all style groups before June 2009 while turns negative in all groups after June 2009, confirming our key result that VT performs better in the pre-June 2009 subsample. The decomposition provides two explanations for the contrasting performance of VT before and after June 2009. First, the RT component are positive before June 2009 while negative after June 2009. This suggests that after June 2009, high past volatility actually predicts high future return. Hence, reducing investment position conditioning on high past volatility decreases rather than increases portfolio return. Second, the VT^* component is negative in the LCC, LCG, MCG, and SCG group before June 2009. This result suggests that before June 2009, high past volatility in those groups predicts low future volatility. By contrast, VT^* is postive in all groups after June 2009, implying a better volatility timing ability of the volatility management strategy in the second subsample. However, the RT component dominates both before and after June 2009, leading to a higher alpah in all groups before June 2009.

Table 4 Results of alpha decomposition

Fund portfolios are sorted into 7 investment style groups: Small-cap core (SCC); Small-cap value (SCV); Small-cap growth (SCG), Mid-cap growth (MCG), Large-cap core (LCC), Large-cap value (LCV), and Large-cap growth (LCG). And in each group of investment styles, we report the relative performance of volatility-time portfolios compared to the baseline portfolios without volatility timing and alpha decomposition results. Numbers labeled “Alpha” are the estimated constant term from a regression of volatility-timing portfolios' returns to baseline portfolios' returns. Numbers labeled “RT” ("VT^*") are the return-timing (volatility-timing) components of estimated alphas. The sample periods are from September 1998 to June 2009 and from July 2009 to December 2020 in panel A and panel B, respectively.

Portfolio	LCC	LCG	LCV	MCG	SCC	SCG	SCV
	Panel A. September 1998-June 2009						
Alpha	0.19	0.18	0.26	0.26	0.15	0.12	0.25
RT	0.32	0.32	0.24	0.34	0.07	0.14	0.13

VT*	-0.13	-0.14	0.02	-0.08	0.08	-0.02	0.13
	Panel B. July 2009-December 2020						
Alpha	-0.07	-0.01	-0.17	-0.13	-0.34	-0.18	-0.37
RT	-0.31	-0.31	-0.35	-0.43	-0.47	-0.42	-0.44
VT*	0.23	0.30	0.18	0.30	0.13	0.24	0.07

4.4 Robustness of baseline results

4.4.1 Results for the funds with alphas in other quartile

In our baseline study, we adopt the mutual fund investment strategy of Hunter *et al*. (2014) to choose the quartile of mutual funds which has the highest in-sample alphas from each category of investment style. The evaluation of VT is based on this strategy. However, our main results are not sensitive to this specific choice of fund investment strategy. Table 5 presents the performance of scaled and unscaled fund portfolios which are formed using alphas in other quartiles.[9] VT still does not work after June 2009 while work quite well in improving distribution characters of returns before June 2009, with some exceptions in the SCG and SCC groups. More specifically, VT increases the Sharpe ratios, skewness while reduces the kurtosis in most groups before June 2009. It reduces the skewness of all SCG groups formed with other quartiles. There is only one slight difference. When the fund portfolios are formed in the lowest alpha quartile, VT has one more failure for improving Sharpe ratio, skewness, and kurtosis before June 2009. It decreases the skewness of the SCC group. Besides, VT only enhances the mean alpha of portfolios in the Growth groups.

[9] The table only reports the Sharpe ratios and alphas. Results on skewness and kurtosis are available upon request.

Table 5 Performance of scale versus unscaled mutual fund portfolios

(portfolios formed in the second highest, second-lowest, and lowest quartiles by in-sample t-value of alphas)

In each group of investment style, we report the portfolios with the second highest, second-lowest, and lowest in-sample t-value of alphas of that group in panel A, panel B, and panel C, respectively. Numbers labeled “Original” (“VT”) are performance indicators of the portfolios without (with) volatility timings. #s indicate that VT contributes to improving performance. The sample starts from September 1998 and ends by December 2020.

Portfolio	Panel A. Second Highest				Panel B. Second Lowest				Panel C. Lowest			
	Sharpe Ratio		Mean Alpha		Sharpe Ratio		Mean Alpha		Sharpe Ratio		Mean Alpha	
	Original	VT	Original	VT	Original	VT	Original	VT	Original	VT	Original	VT
	September 1998-June 2009											
LCC	-0.07	-0.03#	0.05	0	-0.07	-0.03#	0	-0.06	-0.07	-0.02#	-0.02	-0.08
LCG	-0.09	-0.06#	-0.05	-0.06	-0.07	-0.03#	0.22	0.23#	-0.09	-0.05#	0.04	0.02
LCV	0.02	0.07#	-0.04	-0.13	0.01	0.06#	-0.1	-0.19	-0.01	0.05#	0	-0.09
MCG	-0.04	0.00#	0.25	0.22	-0.01	0.02#	0.32	0.2	-0.04	0.00#	0.2	0.14
SCC	0.07	0.10#	0.01	-0.07	0.06	0.08#	-0.04	-0.09	0.04	0.06#	-0.15	-0.21
SCG	-0.02	0.00#	-0.01	0.05#	-0.02	0.00#	-0.15	-0.08#	-0.02	0.00#	-0.14	-0.13#
SCV	0.05	0.09#	-0.16	-0.31	0.11	0.15#	0.18	0.03	0.07	0.10#	-0.15	-0.34
	July 2009-December 2020											
LCC	0.29	0.26	-0.03	-0.07	0.31	0.26	0.03	-0.03	0.29	0.24	-0.01	-0.04
LCG	0.34	0.31	0.18	0.07	0.35	0.32	0.19	0.1	0.31	0.27	-0.02	-0.15
LCV	0.23	0.17	-0.01	-0.18	0.23	0.17	-0.07	-0.13	0.22	0.17	-0.08	-0.18
MCG	0.35	0.3	0.11	-0.07	0.32	0.28	-0.07	-0.24	0.33	0.27	0.07	-0.19
SCC	0.19	0.12	0.01	-0.25	0.18	0.1	-0.05	-0.26	0.19	0.12	0.09	-0.16
SCG	0.2	0.19	-0.37	-0.27#	0.24	0.19	-0.13	-0.32	0.24	0.18	0.12	-0.04
SCV	0.18	0.1	-0.08	-0.28	0.17	0.09	-0.06	-0.33	0.16	0.09	-0.09	-0.3

4.4.2 Results for the downside volatility-managed strategy

Wang *et al*. (2020) show that the performance improvement from volatility scaling is even greater when fund returns are scaled by past downside volatility. In this subsection, we check if the baseline results changes when we switch to downside volatility timing (DVT). Table 6 summarizes the results.[10] Similar to VT, DVT does not work after June 2009 but increases the Sharpe ratios of all groups before June 2009. It turns the Sharpe ratio of the LCC, LCG, and MCG groups from negative to positive. It also increases the skewness of all groups before June 2009. Compared with VT, DVT appears to be more effective in reducing crash risk after June 2009. While VT fails to increase the skewness of the all groups after June 2009, DVT successfully increases the skewness of the lap-cap and MCG groups.

Table 6 Performance of scale versus unscaled mutual fund portfolios (downside volatility scaling)

Numbers labeled "Original" ("DVT") are performance indicators of the portfolios without (with) downside volatility timings. #s indicate that DVT contributes to improve the performance. Fund portfolios are sorted into 7 investment style groups: Small-cap core (SCC); Small-cap value (SCV); Small-cap growth (SCG), Mid-cap growth (MCG), Large-cap core (LCC), Large-cap value (LCV), and Large-cap growth (LCG). And in each group of investment style, we only report the portfolios with the **highest in-sample t-value of alphas** in that group. The sample starts from September 1998 and ends by December 2020.

Portfolio	Sharpe Ratio		Skewness		Kurtosis		Mean Alpha	
	Original	DVT	Original	DVT	Original	DVT	Original	DVT
	Panel A. September 1998-June 2009							
LCC	-0.06	0.05#	-0.63	-0.24#	3.71	3.61#	-0.02	-0.38
LCG	-0.07	0.05#	-0.47	0.02#	3.39	4.15	0.30	0.16
LCV	0.02	0.11#	-0.58	-0.18#	4.30	3.01#	0.10	-0.22
MCG	-0.04	0.07#	-0.67	-0.15#	4.06	4.72	0.23	0.00
SCC	0.10	0.12#	-0.61	-0.43#	4.38	4.01#	0.22	-0.06
SCG	0.00	0.05#	-0.47	-0.32#	3.74	4.82	-0.14	-0.19
SCV	0.12	0.16#	-0.70	-0.13#	4.72	3.81#	0.07	-0.23
	Panel B. July 2009-December 2020							
LCC	0.29	0.21	-0.41	0.49#	4.28	14.78	-0.02	-0.10
LCG	0.34	0.28	-0.18	1.50#	3.37	13.80	0.17	0.01
LCV	0.24	0.14	-0.41	-0.24#	4.98	13.22	-0.04	-0.31
MCG	0.37	0.25	-0.34	0.43#	4.66	11.95	0.07	-0.28
SCC	0.19	0.03	-0.46	-2.22	5.37	12.07	-0.09	-0.78
SCG	0.27	0.19	-0.37	-1.62	4.93	9.06	0.27	0.05
SCV	0.20	0.04	-0.73	-1.99	7.11	10.70	0.09	-0.64

[10] Here we only reports the results for the portfolios with the highest alphas. Other results are similar and available unpon request.

4.4.3 Results for scaling using volatility over the prior month

Previous literature uses volatility over the prior month to scale the investment position. In this subsection, we show that our baseline results are robust if we use the same window as in the literature to calculate past volatility. Table 7 reports the performance of unscaled portfolios versus portfolios scaled using volatility estimated over the past one month. VT still generates higher Sharpe ratios in all groups before June 2009. It also increases the skewness of portfolios in all style groups and reduces kurtosis in most groups before June 2009. By contrast, VT seldom improve the original portfolio after June 2009. Therefore, our key results are robust to the window choice for past volatility calculation.

Table 7 Performance of scale versus unscaled mutual fund portfolios

(volatility estimated over the past one month)

Numbers labeled "Original" ("VT") are performance indicators of the portfolios without (with) volatility timings. #s indicate that VT contributes to improve the performance. Fund portfolios are sorted into 7 investment style groups: Small-cap core (SCC); Small-cap value (SCV); Small-cap growth (SCG), Mid-cap growth (MCG), Large-cap core (LCC), Large-cap value (LCV), and Large-cap growth (LCG). And in each group of investment style, we only report the portfolios with the **highest in-sample t-value of alphas** in that group. The sample starts from September 1998 and ends by December 2020.

Portfolio	Sharpe Ratio		Skewness		Kurtosis		Mean Alpha	
	Original	VT	Original	VT	Original	VT	Original	VT
	Panel A. September 1998-June 2009							
LCC	-0.05	0.00#	-0.55	-0.22#	3.58	2.41#	-0.02	-0.16
LCG	-0.05	-0.01#	-0.37	-0.20#	3.23	2.60#	0.30	0.27
LCV	0.02	0.09#	-0.54	-0.28#	4.30	3.21#	0.10	0.05
MCG	-0.01	0.04#	0.13	0.20#	4.98	4.12#	0.22	0.21
SCC	0.11	0.14#	-0.55	-0.19#	4.28	6.35	0.22	0.09
SCG	0.00	0.04#	0.02	0.13#	4.05	3.73#	-0.12	-0.18
SCV	0.12	0.16#	-0.70	-0.61#	4.85	5.01	0.08	-0.13
	Panel B. July 2009-December 2020							
LCC	0.29	0.27	-0.41	-1.21	4.28	6.62	-0.02	-0.05
LCG	0.34	0.35#	-0.18	-0.63	3.37	4.74	0.17	0.25#
LCV	0.24	0.23	-0.41	-0.72	4.98	4.25#	-0.04	-0.13
MCG	0.37	0.33	-0.34	-0.99	4.66	5.41	0.07	-0.24
SCC	0.19	0.10	-0.46	-1.06	5.37	6.88	-0.09	-0.42
SCG	0.27	0.25	-0.37	-0.96	4.93	5.07	0.27	0.24
SCV	0.20	0.10	-0.73	-0.84	7.11	5.96#	0.09	-0.31

5. Conclusion

In this paper, we apply the volatility timing (VT) technique to mutual fund portfolios of different investment styles. We show that VT is useful not only in the stock and currency markets as documented in the literature but also in mutual fund investments before June 2009. Its effects differ across fund investment styles. For fund portfolios which focus on large-cap and mid-cap stocks, VT universally increases the Sharpe ratio, and skewness while reduces the kurtosis before June 2009. It also increases the Sharpe ratios of the small-cap stocks focused fund portfolios, but the enhancement is smaller compared with large-cap and mid-cap groups. While all the groups have higher Sharpe ratio, higher skewness, and lower kurtosis before June 2009, VT only increases the mean alphas of portfolios focused on growth funds. The fact shows that VT improving portfolios' performance through market-timing and changing factor loadings but not by anomalies.

Moreira and Muir (2017) suggest that the superior performance of the volatility management strategy poses a challenge to the risk-based asset pricing theory. In our mutual fund context, a similar puzzle exists in the sample before June 2009. However, the puzzle disappear after June 2009. Volatility management by scaling the mutual fund positions by past volatility can no longer improve the performance of the fund portfolios. A notable difference between the sample before and after June 2009 is that the stock market return grows much faster after June 2009. We find that the ability of VT to improve Sharpe ratios through changing the portfolio loading on the market factor weakens after June 2009. Hence, how the changing stock market trend affects the effectiveness of the VT technique deserve further theoretical and empirical research.

References

Barroso, P., Santa-Clara, P., 2015. Momentum has its moments. Journal of Financial Economics, 116, 111-120.

Carhart, M., 1997. On persistence in mutual fund performance. The Journal of Finance, 52, 57-82.

Cremers, K. M., Petajisto, A., 2009. How active is your fund manager? a new measure that predicts performance. The Review of Financial Studies 22, 3329–3365.

Fama, E., French, K., 2015. A five-factor asset pricing model. Journal of Financial Economics, 116, 11-22.

Hunter, D., Kandel, E., Kandel, S., Wermers, R., 2014. Mutual fund performance evaluation with active peer benchmarks. Journal of Financial Economics, 112, 1–29.

Jordan, B., Riley, T., 2015. Volatility and mutual fund manager skill. Journal of Financial Economics, 118, 289-298.

Moreira A., Muir, T., 2017. Volatility managed portfolios. The Journal of Finance, 72 (4), 1611–1644..

Wang, F., Yan, X., Zheng, L., 2021. Should mutual fund investors time volatility? Financial Analysts Journal, 77, 30–42

Wang, F., Yan, X. S., 2021. Downside risk and the performance of volatility-managed portfolios. Journal of Banking and Finance, 131, 106198.

Appendix A:

As pointed out by Hunter *et al*. (2014), mutual funds with similar investment strategies will have correlated residuals from factor models like they share a common idiosyncratic risk factor.

This in-group correlation will reduce the efficiency of estimation. Hunter et al. (2014) proposed an approach to group mutual funds with similar investment strategies and use the commonalities of funds' returns in the group to deal with the problem about the in-group correlation of residuals.

Specifically, they use the methodology proposed by Cremers and Petajisto (2009) to measure the deviation between the portfolio held by a fund and a passive equity index's constituents, which is called active share of the fund concerning the index as the benchmark, and categorize funds into different styles groups based on their closet matching equity index. The groups are defined as active peer group benchmark (APB). After that, the authors estimate the residuals of the regression of equal-weighted return of each styles groups to Carhart 4 factors (similar to formula 1 in the main text), and the residuals are augmented to the Carhart 4-factor model to adjust residual correlation between individual funds and enhance the efficiency of estimation (similar to formula 2 in the main text). Their empirical results show that the APB-augmented Carhart 4-factor model substantially reduces time series in-group correlation of residuals compared with the original model.

We follow Hunter et al. (2014) to deal with the econometric problem and group funds into seven groups: small-cap core (SCC), small-cap value (SCV), small-cap growth (SCG), mid-cap growth (MCG), large-cap core (LCC), large-cap value (LCV), and large-cap growth (LCG).

Before September 2009, data of active share is obtained from Petajisto's website. From October 2009 to December 2020, when the data of the best fit index is not provided by the author, we use a similar method as Hunter et al. (2014) and Cremers and Petajisto (2009) to construct the data. More Specifically, we use Russell 1000, Russell 1000 Value, Russell 1000 Growth, Russell Midcap, Russell Midcap Value, Russell Midcap Growth, Russell 2000, Russell 2000 Value, Russell 2000 Growth as the benchmark passive indexes, and for every fund, every benchmark passive index, and in every quarter, we calculate active share with the data of fund's portfolio holdings and the data of the index's constituents and weights in that quarter, the benchmark index that makes a fund have minimum active shares is defined as the best fit index of the fund, we categorize funds according to their best fit indexes. For example, funds whose best-fitted index is Russell 1000 Growth will be categorized into the group of large-cap growth, since Russell 1000 Growth is an index heavily weighted in large-cap growth stocks. Groups with Russell 1000 series index best-fitted funds are named Large-Cap. Groups with Russell 2000 series index best-fitted funds are named Small-Cap. Index constituents are download from Bloomberg and the weights are estimated by us following the methodology manual of Russell index. Data on mutual funds' holdings are downloaded from WRDS.

Table A1 Performance of scale versus unscaled mutual fund portfolios (from September 1998 to December 2020)

Numbers labeled "Original" ("VT") are distributional characteristics and Sharpe ratios of the excess returns of the portfolios without (with) volatility timings. #s indicate that VT contributes to improve the performance. And in each group of investment style, we only report the portfolios with the highest OOS alphas in that group. The sample starts from September 1998 and ends by December 2020.

Portfolio	Sharp Ratio		Skewness		Kurtosis		Mean Alpha	
	Original	VT	Original	VT	Original	VT	Original	VT
LCC	0.12	0.14#	-0.56	-0.68	4.09	4.14	-0.02	-0.07
LCG	0.13	0.17#	-0.48	-0.34#	3.74	3.55#	0.23	0.19

LCV	0.14	0.14#	-0.52	-0.90	4.69	4.98	0.02	-0.06
MCG	0.12	0.17#	-0.81	-0.79#	5.09	4.56#	0.14	0.01
SCC	0.14	0.11	-0.53	-1.12	4.91	6.51	0.06	-0.13
SCG	0.13	0.13	-0.54	-0.98	4.49	5.38	0.09	0.06
SCV	0.16	0.14	-0.72	-1.08	5.78	6.46	0.08	-0.13

Note: a. The performance of a given portfolio becomes better if either one (or a combination) of the followings happens (conditional on other variables holding constant): the mean increases, standard deviation decreases, skewness increases, or the Sharpe ratio increases. b. Small-cap core (SCC); Small-cap value (SCV); Small-cap growth (SCG), Mid-cap growth (MCG), Large-cap core (LCC), Large-cap value (LCV), and Large-cap growth (LCG).