Volatility weighting applied to momentum strategies

Johan du Plessis[1]
Quantitative Analyst, Balance Sheet Management, Old Mutual Emerging Markets,
Capetown, South Africa

Winfried G. Hallerbach[2]
Quantitative Investment Research, Robeco Asset Management
Weena 850, 3014 DA Rotterdam, The Netherlands

Abstract

We consider two forms of volatility weighting (own volatility and underlying asset volatility) applied to cross-sectional and time-series momentum strategies. We present some simple theoretical results for the Sharpe ratios of weighted strategies and show empirical results for momentum strategies applied to US industry portfolios. We find that both the timing effect and the stabilizing effect of volatility weighting are relevant for the improvement in Sharpe ratios. We also introduce a *dispersion weighting* scheme which treats cross-sectional dispersion as (partially) forecastable volatility. Although dispersion weighting improves the Sharpe ratio, it seems to be less effective than volatility weighting.

Key words: momentum, trend following, volatility weighting, dispersion

JEL classification: C58, G10, G11

[1] Much of the research presented was initially done while the author was an intern in the Quantitative Investment Research department of Robeco.

[2] Corresponding author, w.hallerbach@robeco.com. Views expressed in the paper are the authors' own and do not necessarily reflect those of Old Mutual or Robeco. We thank an anonymous referee for critical comments on an earlier draft and helpful suggestions. Of course, we are responsible for any remaining errors.

1 Introduction

Volatility weighting is a form of risk management for investment strategies commonly applied in practice: when volatility is high (low) the positions are scaled down (up). In this paper we consider two main forms of volatility weighting. The first is weighting an investment strategy by its own volatility. The second is weighting each of the strategy's underlying assets with their volatilities, which is equivalent to using *normalized returns*. We provide some theory for the efficacy of both forms of volatility weighting, focusing on the effect on the Sharpe ratio of a strategy. In addition, we consider both versions empirically for time-series and cross-sectional momentum strategies based on US industry portfolios. We extend the concept of signed or directional momentum strategies from the time-series domain to the cross-sectional domain and we introduce a novel form of volatility weighting, *dispersion weighting*, which treats cross-sectional dispersion as a volatility.

Some theory and empirical work on volatility weighting can be found in Hallerbach (2012, 2014) for weighting a strategy by its own volatility as well as for using normalized returns and we build on this work. Barroso & Santa-Clara (2015) find that weighting cross-sectional equity momentum with its own volatility is very effective for improving its risk-adjusted performance: they find that implementing volatility-weighting doubles the Sharpe ratio. Moskowitz et al. (2012) study time-series momentum in futures markets and use normalized returns, but do not study the effect of doing so. Clare et al. (2016) also consider the use of normalized returns for equally-weighted portfolios and momentum strategies in the context of asset allocation. They find volatility weighting to be useful across (but not within) asset classes for the equal-weighted portfolio and that using normalized returns is beneficial for trend-following.

We distinguish between two effects that seem to contribute to the efficacy of volatility weighting: volatility stabilizing (smoothing) and volatility timing. The former exploits a convexity effect when variances are either time-varying or random: the less variation in variances, the lower the aggregate volatility will be. The latter is important when the relationship between returns and volatility is negative. These effects are hard to disentangle, but we find that both are important. Our empirical results confirm that weighting a strategy with its own volatility as well as using normalized returns adds value: the Sharpe ratio increases and the return kurtosis and downside risk decrease. Weighting a strategy with its own volatility seems to work at least when the relationship with volatility is negative and using normalized returns is almost always effective. Dispersion weighting, however, seems to be less effective, though still improving the Sharpe ratio.

This paper is organized a follows. We first present a theoretical framework in section 2 and discuss the data and strategies considered in section 3. Section 4 details the empirical results and section 5 concludes. Technical details are contained in the Appendix.

2 Theoretical framework

In this section, we first define two types of momentum strategies and next we develop some theory for volatility weighting.

2.1 Signed strategies

A *signed* momentum strategy is a directional strategy: depending on the sign of past returns or deviations from the cross-sectional average return a unit long or short position is taken. Consider a market of N assets with period-t returns $\mathrm{r}_{\mathrm{i,t}}$ $(\mathrm{i} = 1 \ldots \mathrm{N})$ and a strategy that invests $\frac{1}{\mathrm{N}}$ at t if $\mathrm{r}_{\mathrm{i,t-1}}$ is positive and $\frac{-1}{\mathrm{N}}$ if it is negative (and 0 otherwise). This is a simple time-series strategy, much like that defined by Moskowitz et al. (2012). It is not hard to define a cross-sectional version of this by considering instead deviations $d_{i,t}$ from the assets cross-sectional average return $\bar{r}_t$, $d_{i,t} = r_{i,t} - \bar{r}_t$.[3] We then go long the asset if it has a higher than average return and simultaneously short the (equal-weighted) market and do the opposite if it has a lower than average return.[4] The bet is that an asset with a higher than average return will continue to have a higher return in the next period. The only difference between the times-series and cross-sectional strategy lies in whether prediction is possible from an assets own past returns or from the deviation of the return from the cross-sectional average return.

For simplicity we will focus on the case of investment in a single asset. The directional momentum set-up is a specific case of the market timing strategies as examined in Hallerbach (2014), so we can use his results. Specifically, it can be shown that for a time-series strategy the returns and their expectation and variance can be written as (see the appendix for a derivation):

$$r_t^T = S_t|r_t|,$$

$$Er_t^T = (p - q)E|r_t|,$$

$$Var(r_t^T) = (p + q)Var(|r_t|) + (p + q - (p - q)^2)(E|r_t|)^2.$$

where $S_t = \text{sign}(r_{t-1}r_t)$ is assumed to be independent of r_t, and where p and q are the success and failure ratios (viz. the probabilities of a correct and incorrect prediction, respectively). For the cross-sectional strategy we substitute the deviation from the cross-sectional return average d_t for r_t in the preceding display.

[3] Another possibility would be to consider the median return instead of the average.
[4] The signed cross-sectional strategy will be long the top half of assets and short the bottom half of the assets if the mean and median are equal. Thus we may expect this strategy to be closely related to the typical cross-sectional strategies studied in the literature (starting with Jegadeesh & Titman (1993)) which buy the top quantile of assets and sell the bottom quantile.

2.2 Weighting with own volatility

One form of volatility weighting is to scale an entire strategy by an estimate of its future volatility. This corresponds to the empirical work done for stock momentum in Barroso & Santa-Clara (2015) (and in a more general context, by Zakamulin (2014)). Here we consider theoretically the effect of volatility weighting on the Sharpe ratio. We assume throughout that processes are adapted to some filtration $(\mathcal{F}_t)_t$. Consider a strategy with the following return decomposition:[5]

$$\mathbf{r_t} = \boldsymbol{\alpha} + \boldsymbol{\gamma\sigma_t} + \boldsymbol{\varepsilon_t\sigma_t} \tag{1}$$

where α and γ are constants, σ_t is a positive predictable process representing the conditional volatility of the strategy returns at time t, and where the variate ε_t is assumed to have zero mean and unit variance conditioned on $\mathcal{F}_{t-1}$. The return process has a predictable portion of which $\boldsymbol{\gamma\sigma_t}$ is related to the predictable volatility and where $\boldsymbol{\alpha}$ is the constant part of the (mean) return which does not depend on volatility. The unpredictable portion of the process is $\boldsymbol{\varepsilon_t\sigma_t}$, of which the volatility is dictated by $\boldsymbol{\sigma_t}$.

The volatility-weighted strategy is then of the form:

$$r_t^* = \frac{r_t}{\sigma_t} = \frac{\alpha}{\sigma_t} + \gamma + \varepsilon_t \tag{2}$$

and the Sharpe ratios of the unweighted and weighted strategies are easily seen to be (with CV denoting the coefficient of variation and assuming all the expectations exist):

$$\mathrm{Sharpe}(r_t) = \frac{\alpha + \gamma E\sigma_t}{\sqrt{\gamma^2 \mathrm{Var}(\sigma_t) + E\sigma_t^2}},$$

[5] This is similar to the return generating process discussed in Hallerbach (2012) – see note 4 of the referenced article.

$$\mathrm{Sharpe}(r_t^*) = \frac{\alpha + \gamma \mathrm{E}\left[\frac{1}{\sigma_t}\right]^{-1}}{\sqrt{\alpha^2 \mathrm{CV}\left[\frac{1}{\sigma_t}\right]^2 + \mathrm{E}\left[\frac{1}{\sigma_t}\right]^{-2}}}.$$

See the appendix for a derivation of the ratios above. The numerators have been scaled to be comparable and so we can consider the effect of volatility weighting on the numerator and the denominator. By Jensen's inequality we have that $\mathrm{E}\sigma_t \geq \mathrm{E}\left[\frac{1}{\sigma_t}\right]^{-1}$ where the strict equality holds when σ_t is deterministic, e.g. $\sigma_t = \sigma$, for all t. Hence, the numerator increases for $\gamma < 0$ and decreases for $\gamma > 0$. For the denominator we find a range of α over which the denominator falls and the expected return of the strategy is positive. This range is:

$$\alpha \in [-\gamma \mathrm{E}[\sigma_t], \Lambda] \qquad \text{for } \gamma < 0$$

$$\alpha \in [\max(-\gamma \mathrm{E}[\sigma_t], -\Lambda), \Lambda] \qquad \text{for } \gamma > 0 \tag{3}$$

where:

$$\Lambda = \sqrt{\frac{\left(\gamma^2 \mathrm{Var}(\sigma_t) + \mathrm{E}[\sigma_t^2]\right)\mathrm{E}\left[\frac{1}{\sigma_t}\right]^2 - 1}{\mathrm{Var}(\frac{1}{\sigma_t})}}.$$

See the appendix for a derivation of the above ranges. We have then for $\gamma < 0$ a sufficient condition for the Sharpe ratio to improve (that is α must be in the relevant range) and for $\gamma > 0$ we only have a necessary condition. This makes intuitive sense: if volatility is negatively related to returns then we can profit from a volatility timing effect by investing less when volatility is high and more when volatility is low. Note, however, that this intuition only goes so far: if the part of returns not related to volatility is too large (viz. when the average return is too large), then volatility weighting is not beneficial. The volatility of $\frac{1}{\sigma_t}$ is amplified, introducing a new source of volatility in place of the old. It is also the case that volatility weighting may even be beneficial if returns are positively related to volatility. It is the upper part of these ranges that is most interesting (the bottom part guarantees a positive expected return). We see that the range increases when multiplying volatility by some $c > 1$, so a larger volatility in this sense is good for volatility weighting (i.e. volatility stabilizing is at work). The range is also increasing in $|\gamma|$, so a larger dependence on volatility is good for volatility weighting.

The above results are in a similar context to that of Hallerbach (2012). They are less general in that they do not attempt to show that volatility weighting is optimal among a class of strategies. However, the result in Hallerbach (2012) holds under the condition that volatility is independent of the normalized returns, which will not be the case if α is non-zero.[6] Some claims can still be made that the result should hold approximately. We have made more precise the intuition that the mean (specifically the part not depending on volatility) should have only a small effect by explicitly giving ranges of α over which volatility weighting may work and have also included explicitly a dependence of returns on volatility. If α is zero then the result of Hallerbach (2012) applies and volatility weighting is in fact optimal for the Sharpe ratio.[7] In this case the effectiveness is related to the variability of volatility by minimizing $\mathrm{CV}(\sigma_t)$. Thus the smaller the effect of the part of returns not depending on volatility the more effective volatility weighting is likely to be.

2.3 Weighting with underlying volatility (normalized returns)

Instead of weighting the entire strategy by its volatility, we can weight each of the underlying assets, creating a new (notional) set of normalized assets. We refer to this as using *normalized returns*. We assume again that all processes are adapted to some filtration $(\mathcal{F}_t)_t$. First consider a

[6] See also Hallerbach (2012), footnote 4.
[7] We should assume $\gamma > 0$ for positive expected returns.

signed time-series strategy on a single asset with returns of the form $\mathrm{r_t} = c_t \sigma_t$. Here we suppose c_t has a conditional variance of 1, σ_t is a positive predictable process and for every t $|c_t|$ and σ_t are independent.[8] Now consider weighting the asset by $w_t = \frac{v_t}{\sigma_t}$ to get a new asset $r_t^* = c_t v_t$ on which the time-series strategy can now be run. Picking v_t deterministic corresponds to volatility targeting. The multiplicative form is assumed as this allows us to make statements about optimality (otherwise we get complications as in the previous section).

If we suppose that the weighting does not influence the probability of making correct or incorrect predictions (for instance restricting v_t to be positive) then the Sharpe ratio of the time-series strategy on the normalized asset can be derived as (see the appendix for details):

$$\mathrm{Sharpe}(\mathrm{R_t^T}) = \frac{p-q}{\sqrt{(p+q)(\mathrm{CV}(|v_t|)^2+1)(\mathrm{CV}(|c_t|)^2+1)-(p-q)^2}}.$$

This is maximized by choosing v_t deterministic (i.e. volatility weighting). Note that the improvement relies on reducing the coefficient of variation of volatility [illegible] the variability of volatility, much as noted in Hallerbach (2012). It is not hard to construct some simple processes that display momentum and to which this result (or a minor variation of it) can be applied. Note that this result applies to a random volatility for a single defined period of time. It does not apply to a deterministic volatility, even if this deterministic volatility is time-varying. Hallerbach (2014) provides results for the latter case.

The above can be applied in a cross-sectional context by weighting with the volatility of the deviations from the cross-sectional average return. A specific case of this would be what we term dispersion weighting, which we will look at empirically. Following Solnik & Roulet (2000) we can model cross-sectional dispersion as a predictable process σ_t such $\mathrm{r_{i,t}} = \mathrm{f_t} + \sigma_t \epsilon_{i,t}$ where we take for $\mathrm{f_t}$ the (equal-weighted) market. That is, we consider dispersion as a cross-sectional

[8] There are slightly weaker assumptions that can be made in the last instance, but this is the most convenient.

volatility. Now weighting by cross-sectional dispersion would improve a time-series strategy run on deviations (i.e. a cross-sectional strategy), which can be seen by replacing returns with deviations, now of the form $\sigma_t \epsilon_{i,t}$ in the previous argument.

One thing worth noting about the above results is that the relationship of returns with volatility (or dispersion) is not negative, but positive. The conditional expectation of returns under the timing strategy is $(p - q)\sigma_t E|c_t|$, which is increasing in volatility. Here it is not a volatility timing effect that results in an improved Sharpe ratio, but rather a volatility stabilizing effect. In practice, the volatility timing may also be important. Of course, it is also necessary to be able to forecast volatility (we have assumed volatility is perfectly predictable) and thus the efficacy of volatility weighting in practice will also depend on how effectively this can be done.

3 Data and implemented strategies

We consider the set of 49 US industry portfolios, as compiled by French[9] with two basic strategies: a signed time-series strategy and a quantile cross-sectional strategy. Each strategy starts with a notional capital of 1. These strategies invest each month based on the past J months of returns (formation period) and hold their positions for 1 month (holding period). J is chosen to be either twelve months (long formation) or one month (short formation).[10] The signed time-series strategy invests a proportion $\frac{1}{N}$ of the (notional) capital available in each asset in each period, positively (negatively) if the formation period return was positive (negative). The quantile strategy invests positively in the top quarter (12 assets) of assets in the formation period

[9] Download from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[10] Individual stocks are characterized by short-term reversal; for this reason their momentum is measured over the past 12 months, excluding the most recent month (“12-1 momentum”). Industries, however, exhibit both short-term and long-term momentum. See Moskowitz & Grinblatt (1999) and our Table 3.

and negatively in the bottom quarter (each leg equal to the capital available). We consider two time periods July 1969 – June 1994 and July 1994 – December 2012. The latter period gives an indication of whether markets behave differently more recently, whereas the former provides a look at behavior over a relatively long time period.

4 Empirical results

We examine empirically the effect of volatility weighting. We consider weighting signed time-series and quantile cross-sectional strategies by their own volatility as well using normalized industry returns. We also consider a strategy that simply invests equally in each asset available each month (the equal-weighted market) for comparison. For the cross-sectional strategies only we consider the effect of dispersion weighting. Similarly to Moskowitz et al. (2012) we use an exponentially weighted moving average (EWMA) with a persistence parameter of 0.9836 based on daily data, scaled by $\sqrt{21}$ to estimate both strategy and individual asset (ex-ante) volatilities.[11]

In this section we use robust regressions with a bisquare weighting function. The reported p-values are based on normality and can only be seen as approximate (even should the underlying data be normally distributed). We sometimes report percentage R-squared values as well – it should be noted that these do not have the same interpretation as for an OLS regression and can sometimes be negative.

[11] The EWMA technically gives a daily volatility estimate, thus the scaling and also the use of a relatively persistent EWMA. The persistence coincides with N=120 days, with a weighted average time lag of about 61 days and a half-life (the time for the weight for a specific return to halve) of about 42 days.

4.1 Own volatility

First we consider whether strategy volatility is predictable. Following Barroso & Santa-Clara (2015) we run AR(1) regressions of the square root of the 21 day realized variances for our strategies in order to analyze the predictability of the volatility.[12]

Table 1 reports the slope coefficients (along with t-statistics in parentheses and p-values in square brackets) and R-squared values for robust AR(1) regressions. The intercepts are all positive and highly significant and thus not reported. We find that volatility is quite predictable with significant AR(1) coefficients of over 0.4 (with one exception) and R-squared of close to 20% or even higher. Volatility appears to be even more predictable in the more the recent period with higher t-values and R-squareds.

Table 1. Predictability of volatility (robust) *Reported are estimated AR(1) coefficients and R-squared values for monthly volatility estimates of various momentum strategies computed as a square root of the realized variance from the past 21 daily returns of these strategies. Industry data for the period Jul1969-Jun1994 and Jul 1994 – Dec 2012 are used to compute the strategy returns. Robust regressions with a bisquare weighting function with a parameter of 4.685 are used. "sts" and "qxs" refer to the signed time-series and the quantile cross-sectional strategy, respectively. t-statistics are in parentheses and p-values in square brackets.*

	Jul 1969 - Jun 1994		Jul 1994 - Dec 2012	
strategy	slope	R-sq %	slope	R-sq %
12 mo qxs	0.44 (11.19) [0.000]	22.03	0.58 (22.33) [0.000]	46.44
1 mo qxs	0.42 (9.47) [0.000]	20.95	0.42 (12.48) [0.000]	24.25
12 mo sts	0.63 (20.42) [0.000]	15.19	0.66 (22.18) [0.000]	58.20
1 mo sts	0.38 (8.21) [0.000]	18.03	0.51 (12.69) [0.000]	32.56
equal-weighted	0.50 (16.94) [0.000]	17.45	0.61 (22.48) [0.000]	52.82

We now try to assess whether, in the framework of section 2, a strategy's return is dependent on its own (ex-ante) volatility and predict whether volatility weighting may be beneficial. First we regress normalized strategy returns on the inverse of volatility as follows:[13]

[12] Barroso & Santa-Clara (2015) run their regressions on the variance estimates, not the volatility estimates. We also ran such regressions. However, these results (unreported) are weaker and in any case of less interest since it is volatility not variance that we wish to forecast.

$$\frac{r_t}{\sigma_t} = \gamma + \alpha\frac{1}{\sigma_t} + \epsilon_t$$

The above regression is based on equation (1) and allows us to estimate the dependence on volatility (γ) and the portion of returns not depending on volatility (α). We can also calculate Λ from (2) and the associated ranges for α over which volatility weighting may work. These figures are reported in Table 2 for robust regressions.

The α estimates fall comfortably into the relevant ranges and the upper value of the range, Λ, is large compared to α, which provides some empirical justification for Hallerbach's (2012) approximation. One would expect volatility weighting to work at least where γ is negative (recall that here the condition was sufficient and not just necessary). The relationship found with volatility is, however, weak and not always negative as one may have expected. In

Table 2. Strategy relationship with own volatility. *Reported are estimates of α (scaled by 100) and γ as in eq.(2), the upper and lower bounds of the ranges in eq.(3) and the estimated error variance of the regression eq.(2) for selected momentum strategies and an equal-weighted market. Return data over the periods Jul1969-Jun1994 and Jul1994-Dec2012 are used to estimate strategy returns, the first 21 days of which are used to provide initial volatility estimates. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used. "sts" and "qxs" refer to the signed time-series and the quantile cross-sectional strategy, respectively. t-statistics are in parentheses and p-values in square brackets.*

Panel (a) Jul 1969 - Jun 1994

strategy	γ	100α	low	high	error var
12 mo qxs	-0.47 (-1.31) [0.189]	2.15 (2.78) [0.005]	1.13	4.5	2.02
1 mo qxs	-0.08 (-0.27) [0.786]	1.05 (1.68) [0.092]	0.19	4.01	1.8
12 mo sts	-0.19 (-1.23) [0.218]	0.90 (3.09) [0.002]	0.52	4.06	1.38
1 mo sts	0.39 (1.89) [0.059]	-0.43 (-0.97) [0.333]	-1.04	4.66	1.19
equal-weighted	0.47 (1.65) [0.100]	-1.20 (-1.26) [0.208]	-1.79	7.92	1.61

Panel (b) Jul 1994 - Dec 2012

strategy	γ	100α	low	high	error var
12 mo qxs	0.01 (0.05) [0.961]	0.54 (0.89) [0.373]	-0.04	8.41	0.96
1 mo qxs	-0.07 (-0.38) [0.705]	0.37 (0.67) [0.503]	0.28	6.88	0.84
12 mo sts	-0.16 (-0.86) [0.390]	0.97 (1.95) [0.051]	0.56	8.36	0.87
1 mo sts	0.00 (0.02) [0.984]	0.29 (0.92) [0.357]	-0.01	6.29	0.52
equal-weighted	-0.02 (-0.09) [0.925]	1.00 (1.55) [0.120]	0.08	9.16	0.81

[13] This allows the use of standard regression techniques (otherwise we need to compensate for conditional heteroskedasticity).

particular the short formation time-series strategy has a positive relationship with volatility (as does the equal-weighted market). The other three momentum strategies do appear to have a negative relationship, with the caveat that the long formation cross-sectional strategy has a very weakly positive γ estimate in the more recent period (this is negative under an (unreported) OLS regression).

Table 3. ***Performance statistics of volatility weighted and unweighted strategies, Jul1969-Jun1994.*** *Descriptive statistics are calculated for quantile cross-sectional and signed-time series strategies with 12 month and 1 month formation periods with a 1 month holding period as well as an equal-weighted market strategy. Industry data over the period Jul1969-Jun1994 are used to calculate the strategy returns and volatility estimates, using the first 21 days of strategy returns or asset returns for initial volatility estimates. Panel (a) reports results for unweighted strategies, starting with the first full monthly return after the first 21 days for industry data. In Panel (b) statistics are reported for strategies weighted with their own volatility. In Panel (c) statistics are reported for strategies run on normalized returns. Annualization is as mentioned previously. "sts" and "qxs" refer to the signed time-series and the quantile cross-sectional strategy, respectively.*

Panel (a) unweighted

strategy	mean	stdev	mean less med	kurtosis	Sharpe	avge drawdowns
12 mo qxs	11.11	13.65	-2.11	1.77	0.77	-5.12
1 mo qxs	10.91	10.17	1.20	0.35	1.02	-4.22
12 mo sts	1.59	13.91	-6.64	6.93	0.11	-8.77
1 mo sts	6.28	12.49	1.06	2.62	0.49	-5.37
equal-weighted	5.56	18.32	0.06	2.42	0.30	-5.75

Panel (b) own volatility

strategy	mean	stdev	mean less med	kurtosis	Sharpe	avge drawdowns
12 mo qxs	15.86	15.75	0.58	0.47	0.94	-4.70
1 mo qxs	15.27	12.78	2.86	-0.19	1.12	-3.77
12 mo sts	6.01	14.45	-5.11	4.83	0.40	-5.82
1 mo sts	6.5	13.37	-0.07	1.63	0.47	-5.66
equal-weighted	2.91	14.24	-1.26	2.50	0.20	-4.72

Panel (c) underlying volatility (normalized returns)

strategy	mean	sd	mean less med	kurtosis	Sharpe	avge drawdowns
12 mo qxs	7.6	6.59	-1.63	0.19	1.12	-4.14
1 mo qxs	7.34	4.97	-0.46	-0.14	1.43	-3.10
12 mo sts	1.35	7.91	-3.06	8.01	0.17	-6.67
1 mo sts	3.77	7.06	0.70	1.29	0.52	-5.53
equal-weighted	2.82	10.78	-0.93	2.79	0.26	-6.87

Table 4. ***Performance statistics of volatility weighted and unweighted strategies, Jul1994-Dec2012.*** *Descriptive statistics are calculated for quantile cross-sectional and signed-time series strategies with 12 month and 1 month formation periods with a 1 month holding period as well as an equal-weighted market strategy. Industry data over the period Jul1994-Dec2012 are used to calculate the strategy returns and volatility estimates, using the first 21 days of strategy returns or asset returns for initial volatility estimates. Panel (a) reports results for unweighted strategies, starting with the first full monthly return after the first 21 days for industry data. In Panel (b) statistics are reported for strategies weighted with their own volatility. In Panel (c) statistics are reported for strategies run on normalized returns. Annualization is as before. “sts” and “qxs” refer to the signed time-series and the quantile cross-sectional strategy, respectively.*

Panel (a) unweighted						
strategy	mean	stdev	mean less med	kurtosis	Sharpe	avge drawdowns
12 mo qxs	6.00	18.05	2.14	6.19	0.32	-4.91
1 mo qxs	2.49	14.73	0.66	1.42	0.17	-5.54
12 mo sts	4.02	12.84	-6.72	5.45	0.31	-4.85
1 mo sts	6.40	11.96	3.68	6.46	0.52	-3.98
equal-weighted	8.34	16.60	-4.46	2.59	0.48	-5.82
Panel (b) own volatility						
strategy	mean	stdev	mean less med	kurtosis	Sharpe	avge drawdowns
12 mo qxs	5.78	11.06	1.77	0.36	0.51	-4.66
1 mo qxs	2.74	11.18	1.26	0.63	0.24	-5.05
12 mo sts	5.46	9.64	-4.43	1.03	0.55	-4.85
1 mo sts	6.60	10.12	2.93	5.05	0.63	-3.58
equal-weighted	5.59	9.90	-3.82	1.76	0.55	-5.09
Panel (c) underlying volatility (normalized returns)						
strategy	mean	stdev	mean less med	kurtosis	Sharpe	avge drawdowns
12 mo qxs	2.76	5.31	0.05	0.68	0.51	-5.04
1 mo qxs	1.50	4.70	1.94	0.33	0.32	-5.15
12 mo sts	2.34	4.86	-0.64	0.88	0.48	-5.05
1 mo sts	2.92	4.43	0.97	2.91	0.65	-4.13
equal-weighted	4.12	6.89	-2.78	1.49	0.59	-5.31

Tables 3 and 4 report descriptive statistics for both unweighted strategies and weighted strategies over the two sub-periods.[14] The descriptive statistics reported are the mean, standard deviation, mean less median (to reflect skewness), kurtosis, Sharpe ratio, and the average of the largest 5 (normalized) drawdowns. The volatility weighted strategies use strategy returns

[14] Figures are annualized as follows. Mean: $(1+r)^{12}-1$, Standard deviation and Sharpe ratio: multiply by $\sqrt{12}$, Mean less median: multiplied by 12

$r_t^* = \frac{\varsigma}{\sigma_t} r_t$ where ς is a (conditional) volatility target of $\frac{0.1}{\sqrt{12}}$, corresponding to an annual volatility of 10%.[15] Figures 1 and 2 plot the log cumulative returns for the weighted and unweighted strategies, where the returns (for the sake of comparability) are normalized with their ex-post volatility.

In the earlier period, all one-month industry momentum strategies have a higher Sharpe ratio than the corresponding long formation strategies; this confirms the findings of Moskowitz & Grinblatt (1991). However, for the second period this is only true for the time series strategies.

It appears that weighting with own volatility is effective (i.e. it increases the Sharpe ratio) at least when the relationship with volatility is negative. For the earlier period this is so for the cross-sectional strategies and the long formation time-series strategy (for the latter strategy, the weighting even flips the Sharpe ratio from below to above the equal-weighted market's). For the market and the short formation time-series strategy where the relationship was positive we see a deterioration: this suggests that (negative) volatility timing is at work and it subsumes volatility smoothing. This is reflected in the cumulative return graphs as well. For the more recent period, the weighting improves the Sharpe ratio for all five of the strategies.

We would expect the standard deviations of weighted strategies to be close to the 10% target (or perhaps a little higher due to effect of a non-zero conditional mean). This is the case in the more recent period, but in the earlier period the standard deviation is much higher, indicating that the volatility estimate does not capture all of the future volatility. The weighting also lowers the kurtosis of the strategies, which suggests further beneficial aspects of volatility weighting. A

[15] The target is arbitrary, except that if it is too large the strategy may end up with negative capital when there are large negative returns. The unconditional volatility will, however, be higher (even under a perfect volatility weighting scheme) because of the effect of a non-zero conditional mean. $\mathrm{Var}(\frac{r_t}{\sigma_t}) = 1 + \mathrm{Var}(\frac{\mu_t}{\sigma_t})$ where μ_t is the conditional (on time t-1) mean.

lower kurtosis would naturally result from stabilizing volatility.[16] The effect on skewness is less clear but the normalized drawdowns tend to decrease, which points at a reduction of downside risk.

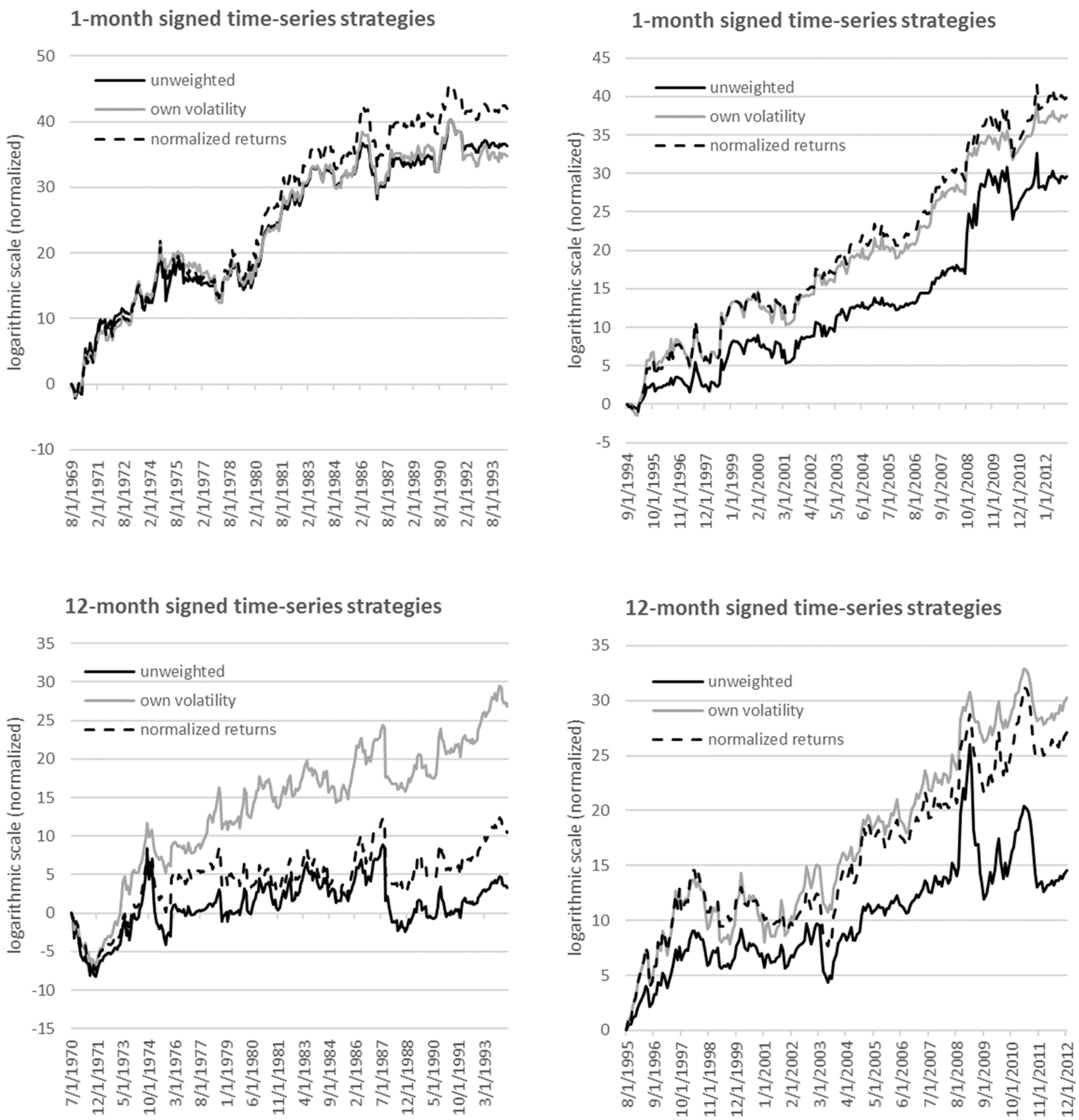

Figure 1: Standardised logarithm of cumulative returns for weighted and unweighted signed time-series strategies.

[16] Consider for instance a GARCH model where it is the random nature of conditional volatility that induces excess kurtosis in the unconditional distribution.

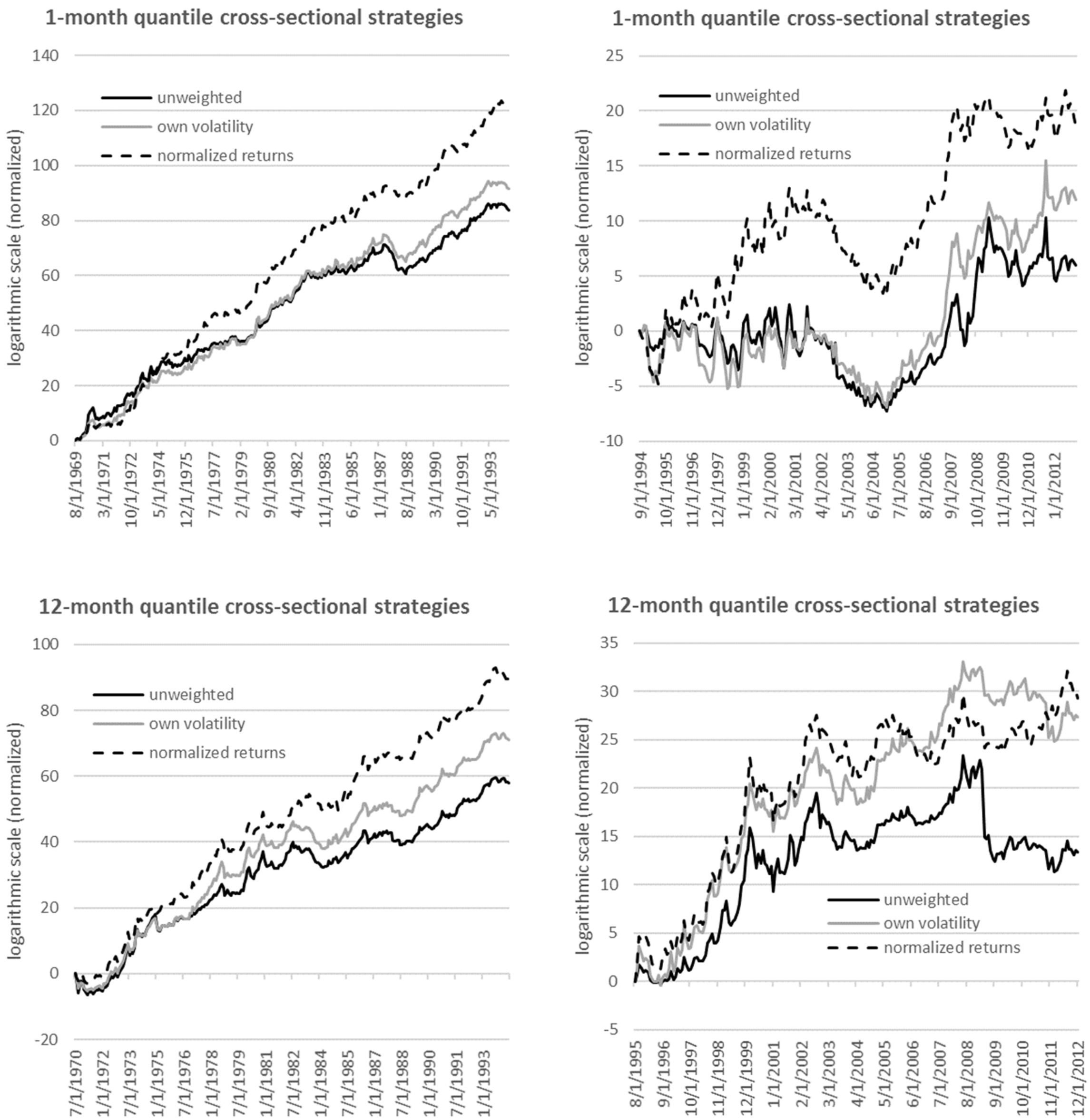

Figure 2: Standardised logarithm of cumulative returns for weighted and unweighted quantile cross-sectional strategies.

Table 5 reports the intercepts for regressions of weighted vs unweighted strategies. This offers a statistical test for outperformance (i.e. alpha).[17] Notable is that the intercepts are positive

[17] The slope coefficient of the regression (not reported) would automatically indicate the (ex-post) relative weighting needed of one strategy vs the other to ensure comparability in cases where the strategies have very different volatility.

(indicating outperformance) albeit only weakly so, in all cases, except for the short formation time-series strategy and the equal-weighted market in the earlier period (where the Sharpe ratios also deteriorated).

Table 5. ***Intercepts of regressions of volatility weighted strategies on unweighted strategies.*** *Reported are intercepts (alphas) for robust regressions (with a bisquare weighting function and a parameter of 4.685) of volatility weighted strategies vs unweighted strategies. In Panel (a) the strategies are weighted with their own volatility and in Panel (b) normalized returns are used. Return data over the period Jul1969-30Jun1994 and Jul 1994-Dec2012 are used to obtain strategy returns and volatility estimates. ‘sts’ and ‘qxs’ refer to the signed time-series and the quantile cross-sectional strategy, respectively. t-statistics are in parentheses and p-values in square brackets.*

Panel (a) weighted with own volatility vs unweighted

	intercept*100	
strategy	Jul 1969 - Jun 1994	Jul 1994 - Dec 2012
qxs 12 mo	1.14 (1.85) [0.064]	0.53 (0.54) [0.588]
qxs 1 mo	0.50 (1.03) [0.301]	0.52 (0.60) [0.550]
sts 12 mo	1.86 (2.72) [0.007]	0.07 (0.11) [0.911]
sts 1 mo	-0.56 (-0.96) [0.339]	0.30 (0.46) [0.645]
equal-weighted	-0.99 (-1.79) [0.073]	0.46 (0.68) [0.499]

Panel (b) normalized returns vs unweighted

	intercept*100	
strategy	Jul 1969 - Jun 1994	Jul 1994 - Dec 2012
qxs 12 mo	2.32 (3.54) [0.000]	0.79 (1.17) [0.244]
qxs 1 mo	3.06 (5.07) [0.000]	0.25 (0.38) [0.707]
sts 12 mo	-0.21 (-0.68) [0.499]	0.59 (1.58) [0.114]
sts 1 mo	-0.24 (-0.95) [0.343]	0.66 (2.18) [0.029]
equal-weighted	-0.07 (-0.21) [0.836]	0.66 (1.41) [0.157]

In order to further investigate the ability of weighting to reduce downside risk (negative skewness or drawdowns), we investigate the possible asymmetry (non-linearity) of the relation between weighted strategy returns and the corresponding unweighted strategy returns (and between the weighted market and the unweighted market). For this purpose, we perform a multivariate regression in the spirit of the asymmetric response market timing model of Henriksson & Merton (1981). However, because we are interested in the downside, we regress the weighted strategy returns on (1) the corresponding unweighted strategy returns, and (2) the *negative* part of the unweighted strategy returns (defined as min[return, 0]). When this second slope coefficient is negative, this implies that the relation between the weighted and unweighted strategy returns is convex: the weighted strategy's sensitivity to the unweighted strategy is smaller for negative returns than for positive returns. Consequently, the weighting reduces the left tail of the strategy's return distribution (and hence reduces downside risk).[18]

Table 6 (a) reports the robust regression coefficients and the corresponding statistics. For all strategies we observe a negative slope coefficient for the negative part of the unweighted strategy return. Notable exceptions are the short formation time-series strategy and the market, which also both showed a deterioration of the Sharpe ratio from weighting in the first period (but not in the second). The negative slopes point at a convex relation between the weighted and unweighted strategies' returns, thus confirming that weighting with own volatility tends to reduce the downside risk of the weighted strategies.[19] Although this particular regression set-up does reveal non-linearities, it can obscure out- or underperformance. In particular, to judge the outperformance of the weighted versus the unweighted strategies, the sign of the intercept should be evaluated together with the sign of the slope coefficient of the negative part of the unweighted strategy. After all, a positive alpha can outweigh a less negative (or even positive) slope and *vice versa*. In this perspective, only the combinations of (1) negative slope and positive alpha (i.e. outperformance) and (2) negative alpha and positive slope (i.e. underperformance) are unambiguous.

[18] When the slope of the negative return part is significant, this implies that the slope of the negative part is significantly different from the slope of the positive return part (i.e. significant non-linearity).
[19] Note that the Sharpe ratio does not reflect any changes in the form of the distribution as long as the mean and standard deviation do not change.

Table 6. ***Asymmetric response regressions of volatility weighted strategies on unweighted strategies and their negative part.*** *Reported are intercepts (alphas) and slopes for robust regressions (with a bisquare weighting function and a parameter of 4.685) of volatility weighted strategies. In Panel (a) the strategies are weighted with their own volatility and in Panel (b) normalized returns are used. Return data over the period Jul1969-30Jun1994 and Jul 1994-Dec2012 are used to obtain strategy returns and volatility estimates. "sts" and "qxs" refer to the signed time-series and the quantile cross-sectional strategy, respectively. t-statistics are in parentheses and p-values in square brackets.*

Panel (a) weighted with own volatility vs unweighted (-neg)			
	Jul 1969 - Jun 1994		
strategy	intercept*100	unweighted	unweighted neg
12 mo qxs	-0.13 (-1.61) [0.107]	1.270 (53.39) [0.000]	-0.300 (-7.36) [0.000]
1 mo qxs	-0.08 (-1.31) [0.191]	1.334 (62.20) [0.000]	-0.165 (-3.78) [0.000]
12 mo sts	0.06 (0.81) [0.420]	1.062 (37.36) [0.000]	-0.082 (-2.01) [0.044]
1 mo sts	0.01 (0.18) [0.855]	1.006 (44.01) [0.000]	0.075 (1.95) [0.051]
equal-weighted	-0.01 (-0.17) [0.862]	0.740 (47.73) [0.000]	0.055 (2.11) [0.035]
	Jul 1994 - Dec 2012		
strategy	intercept*100	unweighted	unweighted neg
12 mo qxs	0.03 (0.27) [0.790]	0.559 (19.99) [0.000]	-0.009 (-0.20) [0.842]
1 mo qxs	-0.37 (-4.11) [0.000]	1.146 (46.84) [0.000]	-0.583 (-13.59) [0.000]
12 mo sts	-0.11 (-1.29) [0.196]	1.039 (36.51) [0.000]	-0.280 (-6.07) [0.000]
1 mo sts	0.14 (1.83) [0.068]	0.748 (30.84) [0.000]	0.213 (4.51) [0.000]
equal-weighted	0.18 (2.12) [0.034]	0.571 (26.06) [0.000]	0.126 (3.57) [0.000]
Panel (b) normalized returns vs unweighted (-neg)			
	Jul 1969 - Jun 1994		
strategy	intercept*100	unweighted	unweighted neg
12 mo qxs	0.15 (1.78) [0.075]	0.452 (18.89) [0.000]	-0.035 (-0.85) [0.394]
1 mo qxs	0.11 (1.46) [0.143]	0.463 (17.84) [0.000]	-0.138 (-2.62) [0.009]
12 mo sts	0.05 (1.28) [0.201]	0.540 (41.98) [0.000]	0.066 (3.60) [0.000]
1 mo sts	0.05 (1.49) [0.136]	0.531 (52.66) [0.000]	0.075 (4.44) [0.000]
equal-weighted	0.03 (0.75) [0.454]	0.570 (55.91) [0.000]	0.030 (1.72) [0.085]
	Jul 1994 - Dec 2012		
strategy	intercept*100	unweighted	unweighted neg
12 mo qxs	-0.00 (-0.04) [0.969]	0.292 (15.07) [0.000]	-0.042 (-1.37) [0.172]
1 mo qxs	-0.01 (-0.12) [0.908]	0.259 (11.55) [0.000]	-0.023 (-0.58) [0.564]
12 mo sts	-0.08 (-1.81) [0.070]	0.492 (33.18) [0.000]	-0.145 (-6.04) [0.000]
1 mo sts	0.03 (0.79) [0.432]	0.419 (40.63) [0.000]	-0.042 (-2.09) [0.037]
equal-weighted	0.17 (2.88) [0.004]	0.399 (26.21) [0.000]	0.091 (3.71) [0.000]

A final question is whether the performance difference between weighted and unweighted strategies is due to different exposures to established priced risk factors.[20] We consider the

[20] We thank the referee for suggesting this extension.

Carhart (1997) 4-factor model, comprising the Fama & French (1992) US market factor (MKT), size factor (small minus big, SMB) and value factor (high minus low, HML), supplemented with the 12-1 momentum factor (MOM).[21] We include the momentum factor to adjust industry momentum for individual stock momentum (cf. Moskowitz & Grinblatt (1991)).

Table 7 shows the results for the unweighted and weighted strategies separately. Starting with the significant exposures of the unweighted strategies, we see that the equal-weighted market has a positive loading on the size factor[22] in both periods, a negative stock momentum exposure in the first period and a positive value exposure in the second. In both periods, the equal-weighted market has no significant alpha. The cross-section strategies have no average market exposures (except for the 1-month strategy in the more recent period), positive loadings on the momentum factor and tend to have negative exposures to the size factor. The 12-month unweighted time series strategy shows significant positive loadings on the momentum, value and market factors in both periods; the 1-month strategy has a negative market exposure in the first period and a positive market loading in the second. The only significant alphas are in the first period: negative for the 12-month time-series strategy and positive for both 1-month strategies.

For the strategies weighted with their own volatility, Table 7 (b) shows the same pattern of significant exposures as for the unweighted strategies, with again significant positive alphas for the 1-month strategies and a negative alpha for the 12-month time series strategy. We conclude that in this multi-factor context, the short formation weighted momentum strategies maintain their risk-adjusted outperformance (non-withstanding any improvements in the downside risk profile of the 12-month strategies as shown in Table 6, which we cannot capture in the linear Carhart model).

[21] These data are also taken from French' website, see footnote 9.

[22] This can be explained from the fact that the Fama-French market factor is value-weighted. The equal-weighted market puts more weight on smaller industries.

The question arises whether the weighting with the own volatility improves the risk-adjusted performance. If the volatilities of the weighted and unweighted strategies are very different (this is especially true for the normalized returns strategies, see Tables 3 and 4), then their return difference will be dominated by the strategy with the largest volatility and this confounds the results. So we cannot simply regress the return difference on the four factors. Instead, we augment the 4-factor model with the unweighted strategys return as an additional regressor. Next, we regress the weighted strategy return on this augmented 4-factor model. The resulting alphas are excess alphas in the sense that they reflect the difference between the unweighted and weighted strategies alphas, while at the same time taking into account any differences in volatility (scaling) between the unweighted and weighted strategies.

Table 8 shows the augmented regression alphas. For both periods, the signs of the strategy alphas agree with the signs of the univariate regression alphas as reported in Table 5, although their statistical significance is nowhere near accepted levels. We observe the biggest change for the 12-month time series strategy in the earlier period: its largest, positive and highly significant univariate alpha is now the smallest and least significant.[23] In the earlier period, the 1-month time series strategy has a negative alpha (significant at the 10% level). For this strategy, Table 2 already showed that weighting reduced the Sharpe ratio. We conclude that weighting with own volatility provides a positive tilt to linear multi-factor alphas (except for the 1-month time series strategy), but the improvement appears not significant.

[23] The augmented regression factor exposures (not shown here for the sake of brevity) reveal that this is also the only strategy that has (highly) significant positive exposures to the market, value and stock momentum factors. These exposures have apparently reduced the strategy's alpha.

Table 7. Fama-French-Carhart regressions. *Reported are intercepts (alphas) and slopes for robust regressions (with a bisquare weighting function and a parameter of 4.685) of unweighted and volatility weighted strategies on the 4 Fama-French-Carhart factors. In Panel (a) the strategies are weighted with their own volatility and in Panel (b) normalized returns are used. Return data over the period Jul1969-30Jun1994 and Jul 1994-Dec2012 are used to obtain strategy returns and volatility estimates. "sts" and "qxs" refer to the signed time-series and the quantile cross-sectional strategy, respectively. t-statistics are in parentheses and p-values in square brackets.*

Panel (a) unweighted strategies

	Jul 1969 - Jun 1994				
strategy	100*alpha	MKT	SMB	HML	MOM
12 mo qxs	0.018 (0.14) [0.886]	0.000 (0.01) [0.989]	-0.058 (-1.31) [0.190]	0.102 (2.15) [0.031]	0.968 (28.70) [0.000]
1 mo qxs	0.674 (4.03) [0.000]	-0.017 (-0.45) [0.656]	-0.205 (-3.43) [0.001]	0.107 (1.69) [0.090]	0.257 (5.61) [0.000]
12 mo sts	-0.494 (-3.33) [0.001]	0.245 (7.15) [0.000]	0.031 (0.58) [0.562]	0.250 (4.43) [0.000]	0.618 (15.37) [0.000]
1 mo sts	0.673 (3.33) [0.001]	-0.296 (-6.39) [0.000]	0.002 (0.04) [0.970]	-0.066 (-0.86) [0.387]	-0.019 (-0.34) [0.735]
equal-wtd	0.026 (0.62) [0.538]	1.034 (107.37) [0.000]	0.310 (20.55) [0.000]	-0.014 (-0.87) [0.387]	-0.044 (-3.85) [0.000]
	Jul 1994 - Dec 2012				
Strategy	100*alpha	MKT	SMB	HML	MOM
12 mo qxs	0.031 (0.19) [0.852]	-0.004 (-0.10) [0.918]	-0.098 (-1.98) [0.047]	0.007 (0.13) [0.893]	0.872 (27.80) [0.000]
1 mo qxs	0.224 (0.94) [0.348]	-0.227 (-4.02) [0.000]	-0.125 (-1.74) [0.082]	0.028 (0.37) [0.712]	0.241 (5.26) [0.000]
12 mo sts	-0.131 (-0.85) [0.395]	0.294 (8.18) [0.000]	0.038 (0.82) [0.411]	0.164 (3.36) [0.001]	0.383 (13.20) [0.000]
1 mo sts	0.203 (1.24) [0.216]	0.123 (3.18) [0.001]	0.039 (0.80) [0.423]	0.045 (0.86) [0.388]	0.023 (0.75) [0.455]
equal-wtd	0.016 (0.23) [0.822]	0.973 (57.74) [0.000]	0.206 (9.61) [0.000]	0.287 (12.51) [0.000]	-0.006 (-0.46) [0.647]

Panel (b) weighted with own volatility

	Jul 1969 - Jun 1994				
Strategy	100*alpha	MKT	SMB	HML	MOM
12 mo qxs	0.210 (1.24) [0.216]	0.021 (0.54) [0.587]	0.043 (0.70) [0.482]	0.140 (2.17) [0.030]	1.074 (23.41) [0.000]
1 mo qxs	0.931 (4.23) [0.000]	0.003 (0.06) [0.954]	-0.219 (-2.78) [0.005]	0.151 (1.81) [0.070]	0.291 (4.83) [0.000]
12 mo sts	-0.373 (-2.22) [0.026]	0.283 (7.26) [0.000]	0.110 (1.82) [0.069]	0.311 (4.87) [0.000]	0.695 (15.27) [0.000]
1 mo sts	0.669 (3.08) [0.002]	-0.319 (-6.42) [0.000]	-0.025 (-0.32) [0.745]	-0.054 (-0.66) [0.508]	0.002 (0.03) [0.978]
equal-wtd	-0.110 (-1.67) [0.094]	0.830 (55.22) [0.000]	0.200 (8.53) [0.000]	0.005 (0.18) [0.855]	0.007 (0.38) [0.706]
	Jul 1994 - Dec 2012				
Strategy	100*alpha	MKT	SMB	HML	MOM
12 mo qxs	0.210 (1.41) [0.158]	0.060 (1.75) [0.081]	-0.101 (-2.30) [0.021]	-0.009 (-0.18) [0.854]	0.469 (16.83) [0.000]
1 mo qxs	0.207 (0.97) [0.332]	-0.122 (-2.42) [0.016]	-0.088 (-1.37) [0.172]	-0.068 (-0.99) [0.322]	0.086 (2.11) [0.035]
12 mo sts	0.170 (1.18) [0.237]	0.182 (5.45) [0.000]	0.040 (0.95) [0.344]	0.057 (1.27) [0.205]	0.315 (11.70) [0.000]
1 mo sts	0.316 (1.82) [0.069]	0.007 (0.18) [0.856]	0.025 (0.48) [0.628]	-0.049 (-0.87) [0.383]	0.002 (0.06) [0.952]
equal-wtd	0.098 (1.10) [0.273]	0.559 (26.67) [0.000]	0.043 (1.61) [0.107]	0.168 (5.90) [0.000]	0.025 (1.45) [0.146]

Table 7. Fama-French-Carhart regressions - continued

Panel (c) normalized returns					
	Jul 1969 - Jun 1994				
Strategy	100*alpha	MKT	SMB	HML	MOM
12 mo qxs	0.191 (2.51) [0.012]	0.025 (1.42) [0.156]	0.002 (0.08) [0.933]	0.086 (2.97) [0.003]	0.430 (20.75) [0.000]
1 mo qxs	0.476 (5.64) [0.000]	0.003 (0.13) [0.898]	-0.104 (-3.44) [0.001]	0.061 (1.92) [0.055]	0.132 (5.71) [0.000]
12 mo sts	-0.254 (-2.84) [0.004]	0.165 (7.97) [0.000]	0.028 (0.87) [0.383]	0.183 (5.42) [0.000]	0.319 (13.22) [0.000]
1 mo sts	0.380 (3.25) [0.001]	-0.194 (-7.25) [0.000]	-0.011 (-0.25) [0.803]	-0.055 (-1.24) [0.216]	-0.009 (-0.29) [0.774]
equal-wtd	-0.027 (-0.64) [0.522]	0.640 (66.06) [0.000]	0.128 (8.43) [0.000]	0.013 (0.83) [0.407]	0.001 (0.09) [0.925]
	Jul 1994 - Dec 2012				
Strategy	100*alpha	MKT	SMB	HML	MOM
12 mo qxs	0.100 (1.32) [0.188]	0.010 (0.54) [0.590]	-0.025 (-1.13) [0.257]	-0.028 (-1.15) [0.250]	0.207 (14.52) [0.000]
1 mo qxs	0.077 (0.83) [0.404]	-0.023 (-1.05) [0.294]	-0.050 (-1.79) [0.073]	-0.017 (-0.56) [0.575]	0.04 6 (2.62) [0.009]
12 mo sts	0.050 (0.63) [0.532]	0.111 (5.97) [0.000]	0.007 (0.31) [0.759]	0.028 (1.10) [0.272]	0.132 (8.79) [0.000]
1 mo sts	0.203 (2.55) [0.011]	0.005 (0.29) [0.776]	0.008 (0.32) [0.746]	-0.025 (-0.98) [0.328]	0.003 (0.17) [0.862]
equal-wtd	0.059 (1.04) [0.299]	0.411 (30.55) [0.000]	0.045 (2.63) [0.008]	0.152 (8.32) [0.000]	0.023 (2.05) [0.040]

Table 8. Augmented Carhart regression alphas. *Reported are intercepts (alphas) and slopes for robust regressions (with a bisquare weighting function and a parameter of 4.685) of volatility weighted strategies on the unweighted strategy and the 4 Fama-French-Carhart factors. In Panel (a) the strategies are weighted with their own volatility and in Panel (b) normalized returns are used. Return data over the period Jul1969-30Jun1994 and Jul 1994-Dec2012 are used to obtain strategy returns and volatility estimates. “sts” and “qxs” refer to the signed time-series and the quantile cross-sectional strategy, respectively. t-statistics are in parentheses and p-values in square brackets.*

Panel (a) weighted with own volatility		
strategy	Jul 1969 - Jun 1994	Jul 1994 - Dec 2012
12 mo qxs	0.109 (1.81) [0.070]	0.032 (0.39) [0.699]
1 mo qxs	0.035 (0.80) [0.426]	0.052 (0.67) [0.504]
12 mo sts	0.029 (0.45) [0.652]	0.026 (0.45) [0.651]
1 mo sts	-0.094 (-1.79) [0.074]	0.056 (0.93) [0.352]
equal-weighted	-0.136 (-2.82) [0.005]	-0.006 (-0.10) [0.922]
Panel (b) normalized returns		
strategy	Jul 1969 - Jun 1994	Jul 1994 - Dec 2012
12 mo qxs	0.100 (1.78) [0.075]	0.058 (1.00) [0.319]
1 mo qxs	0.207 (4.03) [0.000]	-0.008 (-0.13) [0.893]
12 mo sts	-0.002 (-0.06) [0.952]	0.053 (1.89) [0.059]
1 mo sts	-0.008 (-0.35) [0.724]	0.054 (2.12) [0.034]
equal-weighted	-0.041 (-1.36) [0.174]	0.003 (0.07) [0.941]

4.2 Underlying volatility

We now run momentum strategies on normalized asset returns. For each underlying asset i a normalized return series $\frac{\varsigma}{\sigma_{i,t}} r_{i,t}$ is used where the conditional volatility target ς is again chosen to be $\frac{0.1}{\sqrt{12}}$.[24] The strategies are then run on this new set of normalized assets.

Descriptive statistics are also reported in Table 4 and the cumulative returns plotted in Figures 1 and 2. Here the Sharpe ratio increases in almost all cases (the only exception is the equal-weighted market in the earlier period). This is also reflected in the graphs. The kurtosis also mostly decreases, but the effect on drawdowns and skewness is less clear.

Table 5 (b) reports intercepts for the normalized return strategies regressed on the unweighted strategies. Here we see for the more recent period intercepts that are positive, albeit not significant. For the earlier period, however, only the cross-sectional strategies show positive intercepts; for these strategies we also observe the largest increase of Sharpe ratios (see Table 3).

Turning to the asymmetric response regressions reported in Table 6 (b), we first see that the slopes for the full unweighted strategy returns are well below unity; this reflects the diversification across normalized industries. As before, we observe negative slope coefficients for the negative part of the unweighted strategy return, except for the two time-series strategies in the earlier period (which also showed only modest increases in their Sharpe ratios) and the equal-weighted market again in both periods. This confirms that also weighing with the underlying volatility tends to reduce the downside risk of the momentum strategies.

In order to better understand these results we consider some additional tests. Firstly we regress the weighted and unweighted strategies on market volatility and also calculate correlations. The

[24] Results in Hallerbach (2014) suggest it is best to use the same target for all the assets. These targets are for the underlying assets, not the strategy. Unlike with weighting strategies by their own volatility it is not clear what to expect for the standard deviation of the strategy (except being lower because of diversification effects).

slopes of these regressions and the correlations are in Table 9. Three of the four momentum strategies betray a negative relationship with market volatility. From similar findings by Wang & Xu (2015) the negative relationship is not unexpected. It is thus interesting that the short formation time-series strategy shows a positive relationship. In the more recent period the regression estimate is weakly negative, but the correlation is still positive. The equal-weighted market shows a positive relationship with volatility.[25]

Table 9. ***Slopes of momentum vs market volatility regressions and correlations.*** *Reported are slope coefficients (with t-value in round brackets and p-value in square brackets) of regressions of selected momentum strategies against ex-ante market volatility. The first 21 days (52 weeks) of strategy or asset returns are used for the initial volatility estimate. Correlations are also reported. Asset returns for the period Jul1969-Jun1994 and Jul 1994 – Dec 2012 are used. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used. "sts" and "qxs" refer to the signed time-series and the quantile cross-sectional strategy, respectively.*

	Jun 1969 - Jun 1994				Jul 1994 - Dec 2012			
strategy	slope			correlation	slope			correlation
12 mo qxs	-0.26	(-1.77)	[0.077]	-0.14	-0.10	(-0.81)	[0.420]	-0.18
1 mo qxs	-0.32	(-2.68)	[0.007]	-0.15	-0.09	(-0.91)	[0.362]	0.00
12 mo sts	-0.18	(-1.39)	[0.165]	-0.10	-0.38	(-4.68)	[0.000]	-0.10
1 mo sts	0.31	(2.38)	[0.017]	0.01	-0.06	(-0.96)	[0.336]	0.11
equal-weighted	0.43	(2.15)	[0.031]	0.11	0.23	(2.03)	[0.042]	0.03

A negative relationship with market volatility suggests that some of the improvement from volatility weighting should be attributed to volatility timing. The total exposure of the strategy would fall if average market volatility is high and increase when it is low (though changes in correlations would not be adjusted for). The positive relationship with market volatility would at least explain the only small improvement in the Sharpe ratio (from 0.49 to 0.52) and the negative intercept for the short formation time-series strategy.

[25] Note, however, this regression is ill-specified: it regresses on the market's volatility without considering that this volatility changes. The previous results in Table 2 for the market strategy are more appropriate in this respect.

The other elements of volatility weighting that would contribute to its effectiveness are stabilizing of volatility — both within and across assets. This means that volatility weighting may be effective despite a positive relationship with volatility. We can surmise that this stabilizing is the more important effect for the short formation time-series strategy, but it is not clear for the other strategies. A simple means of considering stabilization is to compute volatility estimates for both the unweighted market and the market of normalized returns. The coefficient of variation of the volatility estimate falls from 0.37 to 0.27 in the earlier period and from 0.52 to 0.24 in the more recent period. Similarly the coefficient of variation of dispersion (which we noted is also a kind of volatility), calculated as a cross-sectional standard deviation, drops from 0.30 to 0.25 in the earlier period and from 0.39 to 0.28 in the more recent period. This at least suggests that stabilization is in fact an important element of volatility weighting.

We finally turn to the risk-adjusted performance statistics from the (augmented) Fama-French-Carhart regressions as reported in Tables 7 (c) and 8 (b). Regarding significant factor loadings in the Carhart model in Table 7 (c), the normalized returns strategies all show the same patterns as the strategies weighted with own volatility (and the unweighted strategies); compare with panels (a) and (b). The stock momentum exposures of the cross-section strategies and the 12-month time series strategy are positive in both periods. The two cross-section strategies now both have a significant positive alpha in the first period, the 1-month time series strategy in both periods. The 12-month time series strategy again has a negative alpha in the first period.

Regarding the improvement in risk-adjusted performance of using normalized returns versus the unweighted strategies, Table 8 (b) shows that the normalized cross-section strategies have positive excess alphas in the first period (where the 1-month strategy's alpha becomes highly significant) but effectively zero alphas in the second period. The normalized time-series strategies show a reversed pattern: effectively zero in the first period but significantly (at the 10% level) positive in the second. So there is no consistent improvement in risk-adjusted excess performance over both periods.

We note that this linear multi-factor context may ignore any improvements in downside risk by weighting. Table 6 (b) shows that the normalized time-series strategies in the first period exhibit significant concavity with respect to the unweighted time-series strategies and convexity in the second (positive and negative slopes for the negative part of the unweighted strategy, respectively). Hence, the undesirable concavity is paired with zero excess alphas and the desirable convexity is paired with positive excess alphas (Table 8 (b)). This suggests that the linear factor model does subsume non-linearities to some extent via factor exposures. However, these are selected examples and our results do not warrant any generalization.

4.3 Dispersion as volatility (dispersion weighting)

We now test the effect of dispersion weighting for cross-sectional strategies. In order for this to work we need dispersion to be forecastable. Table 10 gives the slope and R-squared of robust AR(1) regressions of dispersion (measured as a cross-sectional standard deviation). We do see a significant AR(1) coefficient and a moderate R-squared value, suggesting some ability to forecast dispersion, particularly in the more recent period.

Table 10. ***Predictability of dispersion.*** *Reported are the estimated AR(1) coefficients of dispersion (defined as the cross-sectional standard deviation of industry returns in each period) over the periods Jul1969-Jun1994 and Jul 1994 Dec 2012. A robust regression with a bisquare weighting function and a parameter of 4.685 is used. t-statistics are in parentheses and p-values in square brackets.*

Period	Slope	R-squared
Jul 1969 - Jun 1994	0.35 (7.63) [0.000]	12.03
Jul 1994 - Dec 2012	0.62 (13.63) [0.000]	37.73

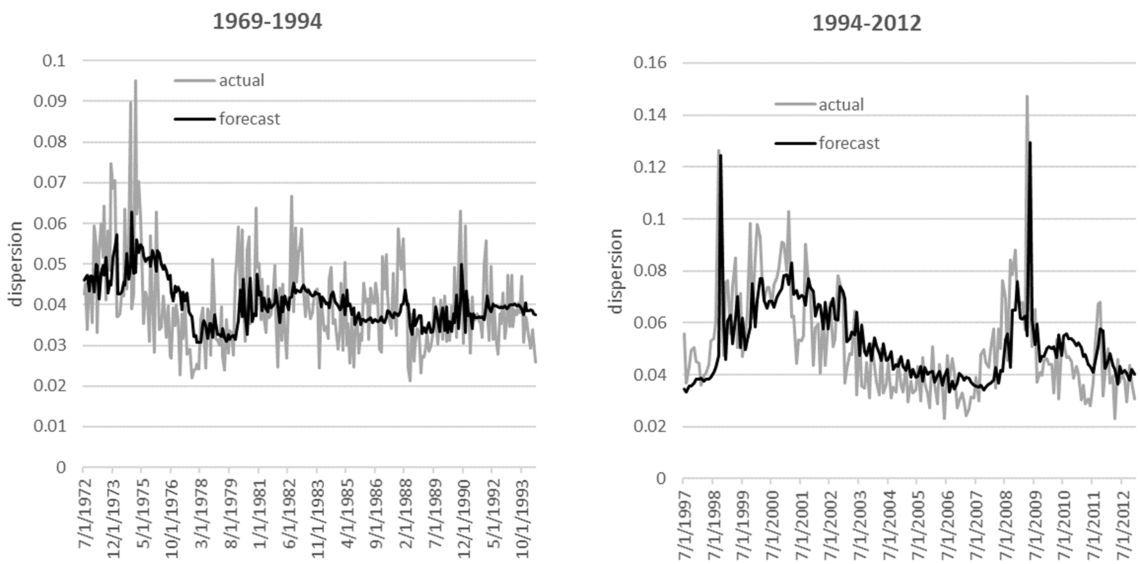

Figure 3. Forecast and actual dispersion over two periods of time.

For our dispersion forecast we run a moving window (OLS) AR(1) regression with a window of 36 months and then do a one step ahead forecast. The forecast and the actual dispersion in each period are plotted in Figure 3. The dispersion forecast is not nearly as wild as the actual dispersion and seems to lag behind it. There is a lot of unanticipated dispersion that the forecast cannot capture. This is, of course, a naïve forecast and could potentially be much improved.

We run quantile cross-sectional momentum strategies scaled with the dispersion forecast and report Sharpe ratios for these strategies, and regress them on the unweighted strategies. In order to disentangle the effect of dispersion from the (in-)accuracy of the forecast we also consider strategies scaled by the actual dispersion in the holding period. This utopian case assumes perfect forecast ability. Although these latter strategies cannot actually be implemented, they do provide a useful benchmark for comparison: they represent the upper bound on the potential value-added of dispersion weighting. The Sharpe ratios are in Table 11 and the intercepts in Table 12.

Table 11. ***Sharpe ratios of cross-sectional strategies weighted with dispersion.*** *The first column reports the Sharpe ratio for an unweighted strategy, the second for a strategy weighted with an AR(1) forecast and the third for a strategy weighted with contemporaneous dispersion in each period. Asset returns for the periods Jul1994-Dec2012 and Jun 1994 - Dec 2012 are used. "qxs" refers to the quantile cross-sectional strategy.*

Panel (a) Jul 1969 - Jun 1994			
Strategy	unweighted	Forecast	actual
qxs 12 mo	0.88	0.92	1.05
qxs 1 mo	0.97	1	1
Panel (b) Jun 1994 - Dec 2012			
Strategy	unweighted	Forecast	actual
qxs 12 mo	0.32	0.41	0.47
qxs 1 mo	0.21	0.34	0.25

Table 12. ***Intercepts of regressions of dispersion weighted strategies on unweighted strategies.*** *Reported are intercepts (alphas) for robust regressions (with a bisquare weighting function and a parameter of 4.685) of dispersion weighted strategies vs unweighted strategies. Return data over the periods Jul1969 - Jun1994 and Jul1994-Dec2012 are used to obtain strategy returns and dispersion estimates. Both a forecast of dispersion and actual dispersion in each period are considered. "sts" and "qxs" refer to the signed time-series and the quantile cross-sectional strategy, respectively. t-statistics are in parentheses and p-values in square brackets.*

Panel (a) Jul 1969 - Jun 1994		
	intercept	
strategy	forecast	actual
qxs 12 mo	0.19 (1.06) [0.291]	0.78 (2.24) [0.025]
qxs 1 mo	0.10 (0.78) [0.436]	-0.31 (-1.23) [0.218]
Panel (b) Jun 1994 - Dec 2012		
	intercept	
strategy	forecast	actual
qxs 12 mo	0.28 (0.62) [0.538]	0.62 (1.03) [0.303]
qxs 1 mo	0.62 (1.53) [0.125]	0.55 (0.93) [0.351]

We do see an increase in the Sharpe ratio for the dispersion weighted strategies. But the increase is mild, particularly in the early period. For the long formation we see that weighting with actual dispersion gives the greatest improvement as expected, but this is not the case for the short formation strategy where weighting with actual dispersion actually did worse than the forecast in the more recent period. This latter result is unexpected. We do, however, see positive intercepts for the regressions (indicating outperformance) except in one case.

Any positive effect from dispersion weighting would also flow from a stabilization of dispersion and possibly a timing effect from a negative relationship with dispersion. Stivers & Sun (2010) for instance find a negative relationship between dispersion and cross-sectional stock momentum. Since volatility and dispersion are positively related (this is seen for stock momentum in Wang & Xu (2015)) there may also be an aspect of volatility timing involved.

5 Summary and conclusions

We defined two types of signed momentum strategies, a time-series and a cross-sectional strategy, the former already well-known in the literature. We consider two forms of volatility weighting: weighting a strategy by its own (predicted) volatility and weighting each of the underlying assets (normalized returns). In the former case we see that the intuition that if returns are negatively related to ex-ante volatility, then volatility weighting will be beneficial is only partially accurate. We consider using normalized returns in both a time-series and cross-sectional setting, deriving some simple results.

We distinguish between a timing effect and a stabilizing effect in volatility weighting. The latter is important when the relationship between returns and volatility is negative. These effects are hard to disentangle, but we find that both are important. Our empirical results confirm that weighting a strategy with its own volatility as well as using normalized returns adds value: the Sharpe ratio increases, the kurtosis and downside risk decrease, and risk-adjusted performance as measured against the 4-factor Fama-French-Carhart model tends to be positive. Regarding the latter there is substantial variation over the two time periods studied, however, and the interplay between improved downside risk and positive alpha prevents conclusive statements. We can conclude, however, that weighting a strategy with its own volatility seems to work at least when the relationship with volatility is negative and using normalized returns is almost always effective. Dispersion weighting, however, seems to be less effective, though still improving the Sharpe ratio.

The results on the industry portfolios have limitations in that it is hard to verify whether strategies could actually have been traded efficiently. Market frictions may erode many of the benefits of volatility weighting and this certainly needs further investigation. Also further disentangling the effects of volatility stabilizing and timing is a fruitful avenue for future research.

References

Barroso, P. & Santa-Clara, P., 2015. Momentum has its moments. *Journal of Financial Economics*, 116(1), pp. 111-120.

Carhart, M. M., 1997. On Persistence in Mutual Fund Performance. *The Journal of Finance* 52(1), pp. 57-82

Clare, A., J. Seaton, P.N. Smith & S. Thomas, 2016. The Trend is Our Friend: Risk Parity, Momentum and Trend Following in Global Asset Allocation. *Journal of Behavioral and Experimental Finance*, 9, p. 63-80.

Fama, E., & K. French, 1992.The Cross Section of Expected Stock Returns. *The Journal of Finance* 47(2), pp. 427-465.

Hallerbach, W.G., 2012. A Proof of the Optimality of Volatility Weighting Over Time. *Journal of Investment Strategies*, 1(4), pp. 87-99.

Hallerbach, W.G., 2014. On the Expected Performance of Market Timing Strategies. *The Journal of Portfolio Management*, 40(4), Summer, pp. 42-51.

Henriksson, R.D. & R.C. Merton, 1981. On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills. *The Journal of Business*, 54(4), pp. 513-533.

Jegadeesh, N. & Titman, S., 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, 48(1), pp. 65-91.

Daniel, K. D. & T.J. Moskowitz, 2013. Momentum Crashes. Columbia Business School Research Paper No. 14-6. Available at SSRN: http://ssrn.com/abstract=2371227

Moskowitz, T.J. & M. Grinblatt, 1991. Do Industries Explain Momentum? *The Journal of Finance* 54(4), pp. 1249-1290.

Moskowitz, T.J., Ooi, Y.H. & Pedersen, L.H., 2012. Time series momentum. *Journal of Financial Economics*, 104(2), pp. 228-250.

Solnik, B. & Roulet, J., 2000. Dispersion as cross-sectional correlation. *Financial Analysts Journal*, 56(1), pp. 54-61.

Stivers, C. & Sun, L., 2010. Cross-Sectional Return Dispersion and Time Variation in Value and Momentum Premiums. *Journal of Financial and Quantitative Analysis*, 45(04), pp. 987-1014.

Wang, K.Q. & Xu, J., 2015. Market Volatility and Momentum. *Journal of Empirical Finance*, 30(C), pp. 79-91.

Zakamulin, V., 2014. Dynamic Asset Allocation Strategies Based on Unexpected Volatility. *Journal of Alternative Investments*, 16(4) (Spring), pp. 37-50.

6 Appendix

For interested readers we derive some of the results in section 2 of the paper.

6.1 Signed strategies

We derive the results in section 2.1. First consider a single asset with returns r_t and a strategy that invests 1 at t if r_{t-1} is positive and -1 if it is negative (0 otherwise). The investment makes a positive return (i.e. the prediction is correct) if $S_t := \text{sign}(r_{t-1}r_t) = 1$. Suppose this signed variable is independent of $|r_t|$. Denote $P(S_t = 1) = p$, $P(S_t = -1) = q$. For continuous variables these will add up to 1.

The return on the strategy can be written as

$$R_t^T = S_t|r_t|$$

The expected return of the strategy is $E[R_t^T] = (p - q)E|r_t|$, noting that $E[S_t] = 1 \times p - 1 \times q$.

The variance of the strategy is

$$\begin{aligned}
\text{Var}(r_t^T) &= E[(s_t r_t)^2] - E[s_t r_t]^2 \\
&= (p+q)E[r_t^2] - (p-q)^2 E[|r_t|]^2 \\
&= (p + q)\text{Var}(|r_t|) + (p + q - (p - q)^2)(E|r_t|)^2.
\end{aligned}$$

6.2 Weighting with own volatility

We derive the results in section 2.2. Consider an asset (or strategy) with the following return decomposition

$$r_t = \alpha + \gamma\sigma_t + \varepsilon_t\sigma_t$$

where α and γ are constants, σ_t is a positive predictable process representing the conditional volatility of the strategy returns at time t, and where the variate ε_t is assumed to have zero mean and unit variance conditioned on $\mathcal{F}_{t-1}$. The expected return is $\alpha + \gamma E\sigma_t$ and the variance is

$$\mathrm{Var}(\gamma\sigma_t + \varepsilon_t\sigma_t) = \mathrm{E}[(\gamma + \varepsilon_t)^2\sigma_t^2] - \mathrm{E}[(\gamma + \varepsilon_t)\sigma_t]^2$$

$$= \gamma^2\mathrm{E}[\sigma_t^2] + 2\gamma\mathrm{E}[\varepsilon_t\sigma_t^2] + \mathrm{E}[\varepsilon_t^2\sigma_t^2] - \gamma^2\mathrm{E}[\sigma_t]^2$$

$$= \gamma^2\mathrm{Var}[\sigma_t] + 2\gamma\mathrm{E}[\mathrm{E}[\varepsilon_t|\mathcal{F}_{t-1}]\sigma_t^2] + \mathrm{E}[\mathrm{E}[\varepsilon_t^2|\mathcal{F}_{t-1}]\sigma_t^2]$$

$$= \gamma^2\mathrm{Var}[\sigma_t] + \mathrm{E}[\sigma_t^2]$$

Thus the Sharpe ratio is

$$\mathrm{Sharpe}(r_t) = \frac{\alpha + \gamma\mathrm{E}\sigma_t}{\sqrt{\gamma^2\mathrm{Var}(\sigma_t) + \mathrm{E}\sigma_t^2}},$$

Now consider the volatility weighted strategy

$$r_t^* = \frac{r_t}{\sigma_t} = \frac{\alpha}{\sigma_t} + \gamma + \varepsilon_t$$

The expected return is $\mathrm{E}[\frac{\alpha}{\sigma_t}] + \gamma$ and the variance is

$$\mathrm{Var}\left(\frac{\alpha}{\sigma_t} + \varepsilon_t\right) = \mathrm{E}\left[\left(\frac{\alpha}{\sigma_t} + \varepsilon_t\right)^2\right] - \mathrm{E}\left[\frac{\alpha}{\sigma_t} + \varepsilon_t\right]^2$$

$$= \mathrm{E}\left[\left(\frac{\alpha}{\sigma_t}\right)^2\right] + 2\mathrm{E}\left[\frac{\alpha}{\sigma_t}\varepsilon_t\right] + E[\varepsilon_t^2] - \mathrm{E}\left[\frac{\alpha}{\sigma_t}\right]^2$$

$$= \mathrm{Var}\left(\frac{\alpha}{\sigma_t}\right) + 1$$

Thus the Sharpe ratio is

$$\mathrm{Sharpe}(r_t^*) = \frac{\alpha\mathrm{E}\left[\frac{1}{\sigma_t}\right] + \gamma}{\sqrt{\mathrm{Var}\left(\frac{\alpha}{\sigma_t}\right) + 1}}$$

$$= \frac{\alpha + \gamma\mathrm{E}\left[\frac{1}{\sigma_t}\right]^{-1}}{\sqrt{\alpha^2\mathrm{CV}\left[\frac{1}{\sigma_t}\right]^2 + \mathrm{E}\left[\frac{1}{\sigma_t}\right]^{-2}}}.$$

We now derive a range for α over which the denominator above falls when performing volatility weighting and the expected return of the strategy is positive. We need $\alpha + \gamma \mathrm{E}\sigma_t \geq 0$ so $\alpha \geq -\gamma \mathrm{E}\sigma_t$. We also need

$$\alpha^2 \mathrm{CV}\left[\frac{1}{\sigma_t}\right]^2 + \mathrm{E}\left[\frac{1}{\sigma_t}\right]^{-2} \leq \gamma^2 \mathrm{Var}(\sigma_t) + \mathrm{E}\sigma_t^2$$

$$\alpha^2 \leq \frac{(\gamma^2 \mathrm{Var}(\sigma_t) + \mathrm{E}\sigma_t^2)\mathrm{E}\left[\frac{1}{\sigma_t}\right]^2 - 1}{\mathrm{Var}\left(\frac{1}{\sigma_t}\right)} = \Lambda^2$$

From this one easily obtains the ranges in (3).

6.3 Weighting with underlying volatility (normalized returns)

We derive the results in section 2.3. Consider a signed time-series strategy on a single asset with returns of the form $r_t = c_t \sigma_t$ with c_t having a conditional variance of 1 and σ_t a positive predictable process. Further assume that for every t $|c_t|$ and σ_t are independent. Now consider weighting the asset by $w_t = \frac{v_t}{\sigma_t}$ to get a new asset $r_t^* = c_t v_t$ on which the time-series strategy can now be run. We suppose the success and failure rates p and q are the same for both strategies and $|c_t|$ and v_t are also independent. The expected return and variance of the timing strategy on r_t^* (with S_t defined as in the text)

Are

$$\mathrm{E}[R_t^T] = \mathrm{E}[S_t|r_t^*|] = (p-q)\mathrm{E}[|r_t^*|] = (p-q)\mathrm{E}[|c_t|]\mathrm{E}[|v_t|]$$

$$\mathrm{Var}[R_t^T] = (p+q)\mathrm{Var}(|r_t^*|) + (p+q-(p-q)^2)(\mathrm{E}|r_t^*|)^2$$

$$= (p+q)(\mathrm{E}[v_t^2] - \mathrm{E}[|c_t|]^2\mathrm{E}[|v_t|]^2) + (p+q-(p-q)^2)\mathrm{E}[|c_t|]^2\mathrm{E}[|v_t|]^2$$

Thus the Sharpe ratio is

$$\mathrm{Sharpe}(R_t^T) = \frac{(p-q)\mathrm{E}[|c_t|]\mathrm{E}[|v_t|]}{\sqrt{(p+q)(\mathrm{E}[v_t^2] - \mathrm{E}[|c_t|]^2\mathrm{E}[|v_t|]^2) + (p+q-(p-q)^2)\mathrm{E}[|c_t|]^2\mathrm{E}[|v_t|]^2}}$$

$$= \frac{(p-q)}{\sqrt{(p+q)\left(\frac{\mathrm{E}[v_t^2]}{\mathrm{E}[|c_t|]^2\mathrm{E}[|v_t|]^2}-1\right)+(p+q-(p-q)^2)}}$$

$$= \frac{(p-q)}{\sqrt{(p+q)[(\mathrm{CV}(|v_t|)^2+1)(\mathrm{CV}(|c_t|)^2+1)-1]+(p+q-(p-q)^2)}}$$

$$= \frac{p-q}{\sqrt{(p+q)(\mathrm{CV}(|v_t|)^2+1)(\mathrm{CV}(|c_t|)^2+1)-(p-q)^2}}.$$