UNIVERSITY OF AMSTERDAM

MASTERS THESIS

Demystifying momentum:

Time-series and cross-sectional momentum, volatility and dispersion

Author:
Johan du Plessis
(johan@johandp.com)

Supervisor (Robeco):
Dr Winfried Hallerbach

Supervisor (UvA):
Dr Peter Spreij

A thesis submitted in partial fulfilment of the requirements for the degree of MSc Stochastics and Financial Mathematics

in the

Korteweg-de Vries Institute for Mathematics
Faculteit der Natuurwetenschappen, Wiskunde en Informatica

29th June 2013

"One of the most mysterious market phenomena is momentum... "

The Economist (*Market Failure*, 12 Dec 2012)

Preface

The research in this thesis was done while the author was an intern at the investment house Robeco and was, at least partly, for the benefit of their Quantitative Strategies department. Despite this the views in this thesis are the author's own and do not necessarily reflect those of Robeco.

Interpretation of results in financial research always requires some subjectivity, particularly where results are not uniformly supportive of one particular conclusion. The results and interpretations in this thesis are no exception. Please view them as a humble attempt to make sense of the very mystical nature of momentum and do not be afraid to question them.

The author would like to thank Dr Winfried Hallerbach and Dr Peter Spreij for their effective supervision. Thanks also to the Quantitative Strategies department of Robeco and Daniel Haesen in particular for the provision of data, valuable conversations and feedback. Any remaining errors are my own.

Abstract

Variations of several momentum strategies are examined in an asset-allocation setting as well as for a set of industry portfolios. Simple models of momentum returns are considered. The difference between time-series momentum and cross-sectional momentum, with particular regard to the sources of profit for each, is clarified both theoretically and empirically. Theoretical and empirical grounds for the efficacy of volatility weighting are provided and the relationship of momentum with cross-sectional dispersion and volatility is examined.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Momentum is a pervasive feature of seemingly all major asset markets. It effectively defies the efficient market hypothesis and despite an extensive body of empirical literature and momentum's apparent simplicity (it is simply the tendency for assets that perform well, or poorly, to continue to do so), it is still not well understood.

Momentum has been observed in two flavours. The first is cross-sectional momentum, which has been an extensively studied feature of asset returns ever since the seminal paper of Jegadeesh and Titman (1993). The second (and similarly pervasive) form of momentum is time-series momentum (Moskowitz *et al.*, 2012).

In this thesis I examine some variations on time-series and cross-sectional momentum. I do so for a set of asset class data, confirming the presence of both forms of momentum in an asset-allocation setting, as well as for a set of industry portfolios as compiled by Fama and French. I consider some simple models of momentum returns and also perform empirical analyses. The relationship between the two basic forms of momentum and the relationship of momentum strategies with cross-sectional dispersion and volatility are examined. In particular the effect of volatility weighting is considered both theoretically and empirically.

1.1 The main concepts

1.1.1 Momentum strategies

Momentum refers (loosely) to the tendency of assets (often stocks) that performed well to continue performing well and assets that performed poorly to continue to do so in the near future. Time-series (TS) momentum refers to such a tendency for a single asset in isolation: the tendency to continue performing positively (negatively) if past returns were positive or high (negative or low). On the other hand cross-sectional (XS) momentum refers to such a tendency relative to other assets: the tendency for past winning stocks to outperform past losing stocks in the near future.

A momentum strategy is one that attempts to benefit from momentum in asset returns. In a very general sense a momentum strategy determines a trend (or signal) for each asset or for assets relative to each other, assigns a strength or score to this trend and then allocates capital based on this. In this thesis I consider several variations on time-series and cross-sectional strategies. Some terminology in connection with such strategies needs to be explained. The *formation period* is the period over which past returns are looked at in order to determine how to bet (based on the momentum signal and strength). The period over which a bet is held is called the *holding period.*[1] Sometimes there is a gap between the formation and holding periods due to a short-term reversal, which is found in stock data.[2]

The typical cross-sectional strategy studied in the literature ranks assets by their past returns and creates a zero-investment long-short portfolio, the long-leg equally weighting the top quantile (often the top decile for stocks) and the short-leg equally weighting the bottom quantile.[3] I refer to these as *quantile cross-sectional* strategies. A variation of this strategy, which has been considered in the theoretical analysis of cross-sectional momentum[4] invests an amount proportional to each asset's (past) deviation from the cross-sectional average return. Because the weights are linear in the assets' past returns, I refer to this

[1]The formation and holding period terminology is used in for instance Jegadeesh and Titman (1993)
[2]See for instance Jegadeesh and Titman (1993). I find no such reversal in the data I consider.
[3]See for instance Jegadeesh and Titman (1993).
[4]For instance in Jegadeesh and Titman (1993) and Lewellen (2002).

Table 1.1: ***A taxonomy of momentum strategies.*** *Here are provided abbreviations and classifications for seven different types of momentum strategies. Strategies can be time-series (TS) or cross-sectional (XS), linear or signed. Linear strategies may be scaled or unscaled. Additionally a cross-sectional strategy may be based on quantiles of ranked returns.*

	TS	XS
linear		
unscaled	ults	ulxs
scaled	slts	slxs
signed	sts	sxs
quantile	N/A	qxs

as an (unscaled) *linear cross-sectional* strategy. I consider also a linear strategy with the weights scaled so that the gross amount in each leg is equal to the notional capital available. Such scaling was done by Lewellen (2002). I refer to this as a *scaled linear cross-sectional* strategy (even though strictly speaking it is no longer linear in the past returns).

The basic form of time-series momentum that I consider buys assets with positive (excess) returns in the past and sells those with negative returns. (By *excess returns* I refer to returns in excess of the risk-free rate). This is similar to the strategy considered by Moskowitz *et al.* (2012), but even simpler as I do not perform any volatility weighting (at least not initially). I refer to this simply as a *signed time series* strategy. I also consider a linear version, as proposed by Moskowitz *et al.* (2012), in which the bet size is proportional to the past return, which I refer to as an (unscaled) *linear time-series* strategy. I further define a scaled version that adjusts bet-sizes so the gross amount invested is equal to the capital available, which is then a *scaled linear time-series* strategy. The preceding strategies consider individual assets in a market; they are 'local' strategies. One may also consider time-series momentum on a 'global'[5] level, that is by considering a time-series strategy on the market rather than on individual assets that make up the market. This essentially means using the same signal and strength for all the assets.

I further define a *signed cross-sectional* strategy which is similar to a signed time-series strategy, but investing in deviations from the average market return.

The signed and (unscaled) linear strategies are useful for theoretical work. I shall refer to or provide several results based on these strategies. For empirical work the most relevant strategies are the signed time-series and quantile cross-sectional strategies.

In table 1.1 I provide abbreviations and a hierarchical classification for all the momentum strategies that I have mentioned thus far. More details on the strategies I consider can be found in appendix B.

1.1.2 Volatility weighting

Volatility-weighting (closely linked to volatility targeting) refers to the adjustment of bet-sizes by some measure of volatility (either of the underlying assets or of the strategy itself) in the hope of creating a more stable series of returns (in fact the goal is to lower the variability of volatility – the volatility of volatility), hopefully with a higher Sharpe ratio.[6] Volatility weighting may also be useful because of a volatility timing effect, if returns are adversely affected when volatility increases (Hallerbach, 2012). There are both empirical and theoretical grounds for believing that volatility weighting can improve strategies. Volatility weighting is useful for risk-budgeting as it allows the proportion of risk placed in a strategy to be controlled. There are (at least) two ways of doing volatility weighting. The first is to consider a momentum strategy and scale the entire strategy by an ex-ante measure of its volatility. The second is to scale each asset under consideration by its ex-ante volatility. I will refer to the latter form of volatility weighting as using *normalised returns.*

1.1.3 Cross-sectional dispersion and volatility

Cross-sectional dispersion refers to the variation (dispersion) between asset returns over a given period or at a point in time. For instance, considering a single month, we can compute the standard deviation of returns *across* stocks. This is a measure of uncertainty in the market as well as of how differently stocks

[5] I take this terminology from Duyvesteyn and Martens (2013).
[6] See Hallerbach (2012) for a short discussion of volatility weighting.

behave from each other (Stivers and Sun, 2010). There is a theoretical (positive) link between dispersion and market volatility (Yu and Sharaiha, 2007). There is also a link between cross-sectional momentum and dispersion. Empirically it is negative and somewhat weak (Stivers and Sun, 2010). However, this negative relationship is counter-intuitive as cross-sectional momentum essentially invests *in* cross-sectional dispersion.

Volatility comes in three forms: strategy volatility, individual asset volatility and market volatility. The relationship of momentum with market volatility has been found to be negative (Wang and Xu, 2009). Ex-ante volatility estimates attempt to measure future volatility from past asset returns. Two important ways of measuring ex-ante volatility are realised variances (RV) and exponentially weighted moving averages (EWMA). Both of these are explained in appendix C.5.

1.2 Research goals

I will attempt to answer to some extent the following questions.

1.2.1 Is there momentum and where is it strongest?

It is of interest firstly to see whether momentum does exist at an industry and asset class level and whether it behaves consistently with previous literature, for instance when looking at various formation and holding periods. Whether time-series or cross-sectional momentum outperforms is also interesting, as well as how the performance of the various specifications compare.

1.2.2 Where do time-series and cross-sectional momentum come from and where do they differ?

The relationship between time-series and cross-sectional momentum is of interest as these have rarely been studied together. If the two strategies differ we want to know where they differ (including whether there is scope for a hybrid strategy that includes signals from both types of momentum). Given that momentum profits exist, the source of these profits becomes of interest, including how this may differ for time-series and cross-sectional momentum. I am interested interested in verifiable relationships between asset returns rather than behavioural explanations for these relationships. These issues are interesting from a theoretical and an empirical perspective.

1.2.3 Does volatility weighting work?

I hope to confirm the efficacy of volatility weighting for momentum strategies and provide some theoretical and empirical guidance as to why (or why not) volatility weighting is useful and how it works.

1.2.4 What is the relationship of momentum with dispersion and volatility?

The relationship of momentum, in particular cross-sectional momentum, with dispersion is interesting. A negative relationship has been found empirically, but the linear decomposition suggests a positive relationship and if there were no dispersion there would be no cross-sectional momentum profits. I would thus like to consider this empirically and test the relationship for industries and asset classes.

The relationship of cross-sectional momentum with market volatility has been found to be negative. I want to confirm this and also test it for time-series momentum. The interaction of dispersion and volatility in this context is also of interest, particularly as it may help to explain the efficacy (or otherwise) of volatility weighting. Different forms of volatility are of interest in different contexts: the volatility of the momentum strategy itself, the volatility of individual assets and the volatility of the market.

1.3 Literature

The literature on momentum is diverse (but mostly empirical in nature). I provide very brief highlights of areas of interest to this thesis.

1.3.1 Cross-sectional and time-series momentum

A large amount of literature, starting most prominently with Jegadeesh and Titman (1993), examines the cross-sectional form of momentum empirically. Jegadeesh and Titman (1993) first analyse what I have termed a quantile cross-sectional momentum strategy for US stock returns, finding it to be economically significant and reversing after about a year (Jegadeesh and Titman, 1993). They note a short-term reversal in stocks and skip a week between the formation and holding period.

A linear cross-sectional strategy was first defined (for contrarian strategies) by Lo and MacKinlay (1990), then considered for momentum by Jegadeesh and Titman (1993) and notably by Lewellen (2002). The linear strategies are more tractable theoretically and are used mainly to decompose profits from momentum strategies into those from auto-covariance, cross-serial covariance and from cross-sectional dispersion in mean returns (note here that this dispersion contributes positively to momentum profits). Lewellen (2002)'s finds that cross-sectional momentum profits come from cross-serial covariance. However, Moskowitz *et al.* (2012) find the source to be auto-covariance.[7]

Moskowitz *et al.* (2012) introduced time-series momentum in contrast to cross-sectional momentum for the first time. However, time-series momentum is really just a trend-following strategy, many variations of which have been examined in the past and for more complicated means of evaluating momentum. For instance technical trading rules are studied in Brock *et al.* (1992) and Thomas *et al.* (2012) use a moving average rule. I take a simpler approach closely related to that of Moskowitz *et al.* (2012).

Moskowitz *et al.* (2012) analyse time-series momentum strategies for a very wide range of assets (using futures data) and find that past returns positively predict future returns. They also find that time-series momentum (at least partially) reverses after a year. Moskowitz *et al.* (2012) find that their time-series strategy outperforms typical cross-sectional strategies and that these strategies are highly correlated. Moskowitz *et al.* (2012) also introduce the linear time-series strategy I also consider. Duyvesteyn and Martens (2013) consider briefly the distinction between global and local time-series momentum in government bonds. They hypothesise that in this market global momentum outperforms as there is mean-reversion between assets, indicated by poor cross-sectional momentum performance. The results of Thomas *et al.* (2012), however, considering trend-following within and across asset classes suggest that local momentum is better.

Antonacci (2013b), in the context of stocks, and Thomas *et al.* (2012), in the context of asset allocation, consider combining trend-following[8] and cross-sectional signals in the context of long-only momentum. Both find that this combination is better than a cross-sectional strategy, but the latter find that it is not better than a trend-following strategy.

1.3.2 Skewness

Martin (2012) and Martin and Bana (2012) perform some theoretical work directed mainly at the skewness of trend-following (i.e. not cross-sectional) strategies – the strategies analysed are very general and include both linear and non-linear versions. They, however, do so under the assumption that (normalised) asset returns are iid with zero mean and find that these strategies tend to have positive skew. This contrasts with empirical findings of negative skew for cross-sectional momentum (Daniel and Moskowitz, 2011; Barroso and Santa-clara, 2012).

1.3.3 Industry and asset class momentum

Blitz and Van Vliet (2008) consider (as one part of their paper) cross-sectional momentum strategies across asset classes using a dataset similar to the one in this thesis. They confirm that cross-sectional momentum exists not only within but across asset classes. They surmise that this arises because investors perceive tactical asset allocation on a global level to be challenging and so refrain from it. Professional managers tend to focus in allocation within asset-classes, leaving tactical allocation to end investors. They find no short-term reversal. Antonacci (2013a) looks at a long-only time-series momentum for several indices (much like the asset classes I consider). Thomas *et al.* (2012) examine both time-series and cross-sectional momentum (but long only) in the context of asset class allocation. They find that time-series momentum outperforms.

Lewellen (2002) includes, among others, industry portfolios and he confirms the presence of cross-sectional momentum there. Moskowitz and Grinblatt (1999) consider industry momentum in some depth.

[7] Lewellen (2002) considers industry portfolios and finds this result for holding periods of larger than a month, whereas Moskowitz and Grinblatt (1999) consider individual stocks and a 1 month holding period.

[8] Thomas *et al.* (2012) use moving average signals.

They find no short-term reversal and that a one month-one month momentum strategy is the strongest (but, not after transaction costs are taken into account). Furthermore they find at least as large a contribution from long positions as short positions.

1.3.4 Volatility weighting

Moskowitz *et al.* (2012) use volatility weighting (using normalised returns) for their time-series strategy (although they do so merely because they use a large number of assets and wish to make them more comparable – they do not study the effect of this weighting). Barroso and Santa-clara (2012) consider volatility weighting (using the strategy's own volatility) for a traditional cross-sectional momentum strategy and find such weighting improves the risk-profile of the strategy considerably. They find in particular that the risk of momentum (in the form of a realised variance estimate) is forecastable (positively autocorrelated) and that the volatility weighted momentum avoids the extremely large (but infrequent) momentum crashes investigated in Daniel and Moskowitz (2011). Some theoretical work on volatility weighting can be found in Hallerbach (2011) and Hallerbach (2012) (the latter result is approximate in general). Asness *et al.* (2012) report results for volatility-weighted momentum portfolios (in an internet appendix), but they do so to justify ex-post volatility procedures they use in their main results.

Thomas *et al.* (2012) consider also the use of normalised 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 normalised returns is beneficial for trend-following. They find little gain (within asset classes) from basing momentum rankings on volatility-weighted returns.[9]

1.3.5 Volatility and dispersion

Stivers and Sun (2010) find a negative relationship between cross-sectional momentum and two ex-ante measures of cross-sectional dispersion. In particular they find a strong link with a *change* in momentum profits. They argue that momentum is pro-cyclical and that dispersion acts as a leading countercyclical indicator. They report this for stock data and mention that it remains true for industry level portfolios. Lillo *et al.* (2001) do some theoretical and empirical work on cross-sectional dispersion (which they call *variety*). Yu and Sharaiha (2007) look at dispersion in a different context (alpha budgeting), but do obtain some useful theoretical results, in particular clarifying the positive link between dispersion and volatility. Wang *et al.* (2013) find that timing momentum by reducing exposure whenever cross-sectional dispersion is far from its mean is advantageous. They also show that (the rank of) momentum returns are positively related to stock sensitivity to cross-sectional dispersion in the so-called ZCAPM model. This latter finding seems to suggest a positive link of momentum with dispersion.

Wang and Xu (2009) find a negative link between cross-sectional momentum (in US stocks) and an ex-ante measure of market volatility. They find this predictability to be especially strong during negative market states and mostly from the loser portfolio. They also report a relatively large positive correlation between volatility and cross-sectional dispersion and find that dispersion's predictive power is lost in the presence of volatility. Daniel and Moskowitz (2011) find that momentum crashes tend to occur after market declines and when ex-ante volatility estimates are high.

1.4 Data

I use mainly two sets of data. The first is a set of 49 industry portfolios compiled by Fama and French (which I will sometimes refer to as *FF49* in my tables for the sake of brevity). This data is freely available, regularly updated and well-known to researchers already. My preliminary analysis is done using data over the period July 1969 to June 1994. The first date is the first time that all industry portfolios are available and remain available to the end of the dataset. I focus on the monthly returns though I also use the daily returns to estimate volatilities. The out of sample period is from July 1994 to Dec 2012.

The second set of data is provided by Robeco and is used in its Multi-Asset Allocation (MAA) model. It contains a set of asset classes, proxied via a number of indices. Investment into these indices is mainly via futures. This set of data is of most practical relevance to Robeco who include momentum as one of the variables in their MAA model. My preliminary analysis uses the period 4 January 1979 to 5 December 2002. The first date is the first time that at least 8 asset classes are available. This dataset contains weekly

[9] Here it seems that the bet sizes were not adjusted for volatility, however. In this case the risk of the strategy will still be dominated by the most volatile assets that end up in the winner portfolio.

data. The out of sample period is from 5 December 2002 to 17 April 2013. It is common to normalise asset class data by dividing by an ex-ante measure of volatility because of the widely varying volatilities of the assets. I do not do this initially in order to treat both datasets consistently and because I want to study the effect of this normalisation. Many conclusions are unchanged for normalised returns and unweighted returns.

More details regarding the data can be found in appendix A.

1.5 Structure

The rest of this thesis is organised as follows. Chapter 2 expounds some theoretical ways of looking at momentum. This includes looking at when time-series or cross-sectional momentum will outperform and the sources of profit for each, volatility weighting, and linking momentum with dispersion. Chapter 3 looks at evidence for the existence of momentum with six different types of strategies (three cross-sectional and three time-series) and examines the relationship between them. Chapter 4 examines sources of momentum profits for both time-series and cross-sectional strategies. Chapter 5 examines volatility weighting empirically and chapter 6 looks at the empirical relationship between momentum, dispersion and market volatility. Chapter 7 then examines many of the conclusions from the preceding chapter on a hold-out sample of data. The final chapter concludes and provides suggestions for further research.

Chapter 2

Theoretical perspectives

I present some disparate results using various assumptions and in different settings. I test many (but not all) of the conclusions in this chapter in the empirical work of later chapters. The goal is to build an intuition for how we expect momentum to behave, which can then be compared with the actual behaviour of momentum.

In section 2.1 I provide a very general way of thinking about the construction of a momentum strategy and in section 2.2 I consider the difference between time-series and cross-sectional momentum in a two asset market. Section 2.3 introduces four momentum strategies that I will consider in my theoretical analysis. Section 2.4 is concerned with the theory behind volatility weighting and section 2.5 considers the link between momentum and dispersion. Section 2.6 sketches a simple link between volatility and momentum and section 2.7 looks at the relationship between global and local time-series momentum and cross-sectional momentum.

I assume throughout that returns are excess returns (in excess of the risk free rate). See appendix C.1 for more on excess returns and appendix C.2 for a short discussion of the return definitions that can be used.

2.1 Conceptual framework

We may consider a momentum strategy as consisting of three elements. These are three functions of the available data and I have illustrated them in figure 2.1.

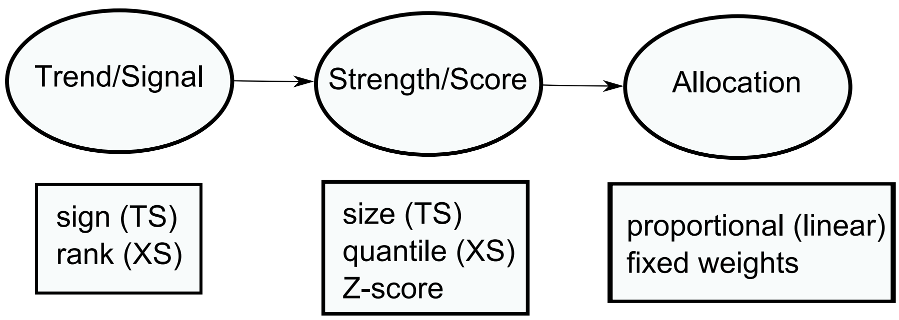


Figure 2.1: ***A conceptual framework for a momentum strategy.*** *The three components of trend/signal, strength/score and allocation are depicted.*

First the momentum of asset returns must be determined, whether up or down for instance. We call this the *trend* or *signal*. For instance for a simple time-series strategy this is simply the sign of the most recent return. For a quantile cross-sectional strategy it would be determined by rank, for instance whether the asset is in the top (up) or bottom (down) quantile or neither (neutral).

Next the *strength* or *score* of the trend can be determined. For instance the size of the past return, the rank or quantile of the return, etc. A Z-score would take into account volatility in the strength.[1] In principle a momentum strategy could invest more in assets with a stronger trend. This is the last part of

[1]A Z-score would typically divide asset returns by their standard deviation in order to make returns between assets more comparable. This is a form of volatility weighting if the investment is also in normalised returns.

the strategy, the allocation: deciding how to invest given the trend and strength. For instance a linear time-series strategy may invest proportionally to the size of the past return. The quantile cross-sectional strategy only invests in assets in the top and bottom quantiles.

2.2 A highly stylised conception of momentum

In this section I consider a very intuitive and graphical representation of momentum assuming a market of only two assets.[2] This drastic simplification provides some insight into the difference between cross-sectional and time-series momentum which I will attempt to test later. I then make the analysis somewhat more formal and consider how auto-covariances and lead-lag relationships (cross-serial covariances) influence which of time-series and cross-sectional momentum performs better.

A simple cross-sectional strategy would buy the asset that has a higher return and sell the asset with a lower return[3] and a simple time-series strategy would buy an asset that has gone up and sell an asset that has gone down.[4] We can distinguish between three cases:

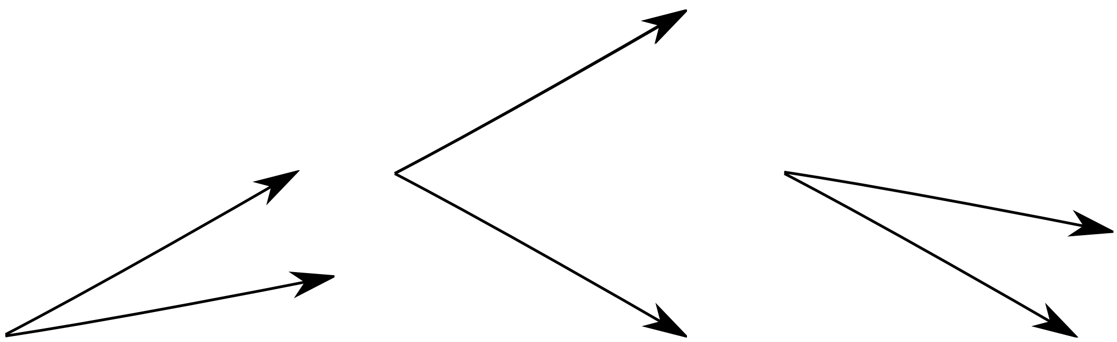
(1) both assets are trending up

(2) one asset is trending up and one is trending down

(3) both assets are trending down

These three cases are depicted graphically in figure 2.2. I will take for granted (for now) that the strategies are able to identify the trend.

In the second case, if one asset is trending up and one is trending down, time-series and cross-sectional strategies will have the same positions. If, however, both assets are trending up (or down) then a time-series strategy would for both assets go long (short), investing in the direction of the trend, and profit from it. However, a cross-sectional strategy would go long one asset and go short the other. It would lose on the asset that it shorts (longs) if the trend is up (down). By this reasoning one should never invest cross-sectionally as it has the unhelpful restriction that it must always be short one asset and long the other.



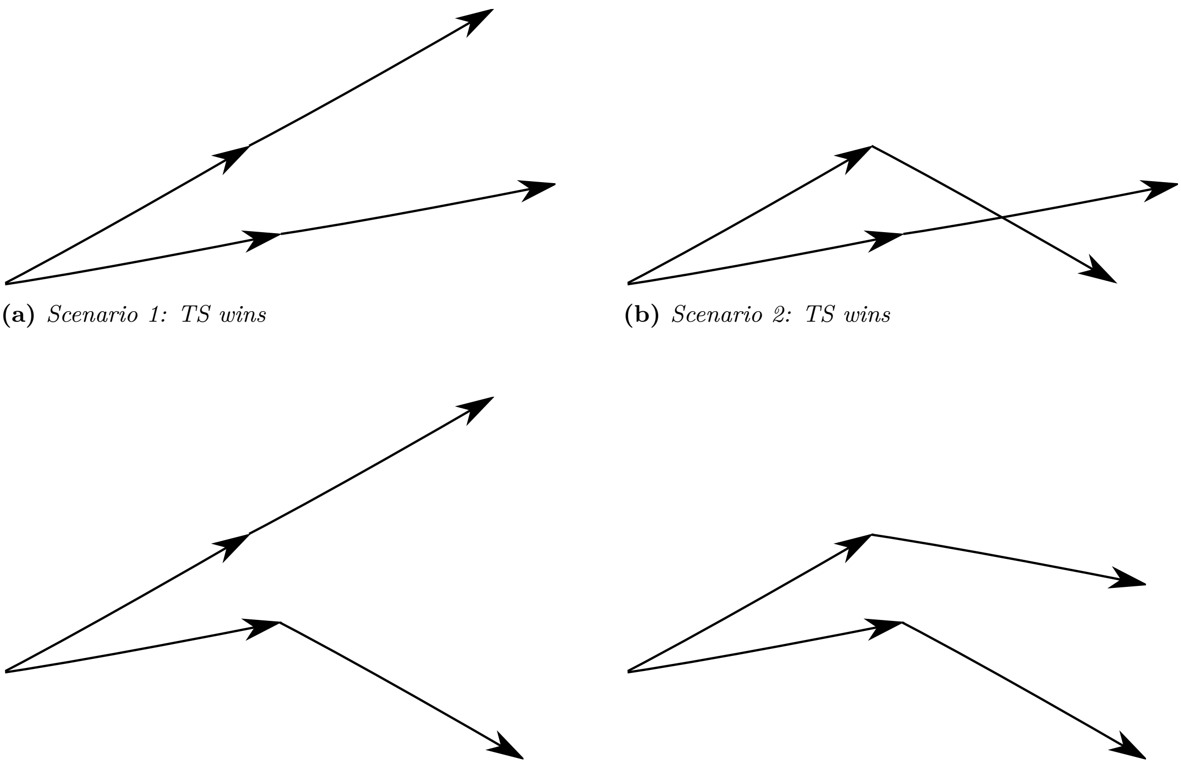
(a) *up* | *up* **(b)** *up* | *down* **(c)** *down* | *down*

Figure 2.2: ***Stylised trend cases.*** *In a two asset market both assets can trend up (up|up), both can trend down (down|down) or one can trend up and the other down (up|down).*

Of course trends do not necessarily continue and so the above stylised example misses out on cases of reversal. Let us then consider instead two periods, the formation period and the holding period. For the formation period we still have the previous 3 cases. For each of these cases there are now four possibilities as each asset may go up or down in the holding period. I graphically depict these possibilities for the case where both assets went up in the formation period in figure 2.3. The case where both assets went down is symmetric[5] and is shown in figure 2.4. The case where one asset went up and the other down is not interesting as both time-series and cross-sectional strategies take the same position.

We now see that for each of the up|up and down|down cases there are two scenarios (figures 2.3a, 2.3b for up|up) in which the time-series strategy outperforms and two scenarios (figures 2.3c, 2.3d for up|up)

[2] The initial thoughts behind this section are due to Dr Hallerbach.
[3] This is a quantile cross-sectional strategy with two quantiles.
[4] This is a signed time-series strategy.
[5] Just flip each image about the horizontal axis.



(a) *Scenario 1: TS wins* **(b)** *Scenario 2: TS wins*

(c) *Scenario 3: XS wins* **(d)** *Scenario 4: XS wins*

Figure 2.3: ***Two-period scenarios (up|up).*** *In a two asset market with two periods and both assets going up in the formation period we distinguish four scenarios as each asset can go up or down in the holding period. In two scenarios a time-series strategy outperforms and in two scenarios a cross-sectional strategy outperforms.*

in which the cross-sectional strategy outperforms. The overall performance of the two strategies would thus depend on the frequency of these cases and the magnitude of the profits/losses in each case.

Time-series wins if both trends continue but also if the top asset reverses (but not the bottom asset) – the cross-sectional strategy will in the latter case make a loss on both assets. Cross-sectional wins if bottom asset reverses (but not the top) or if both assets reverse – in the latter case the time-series strategy makes a loss on both assets.

Scenarios 2.3a and 2.3d (and their counterparts in figure 2.4) seem to suggest that time-series will outperform in the presence of strong continuation of trend and cross-sectional in the presence of mean-reversion (though it does not imply that cross-sectional strategies will profit in the latter case). Scenario 2.3b seems unlikely – it requires the asset with the stronger trend to revert.[6]

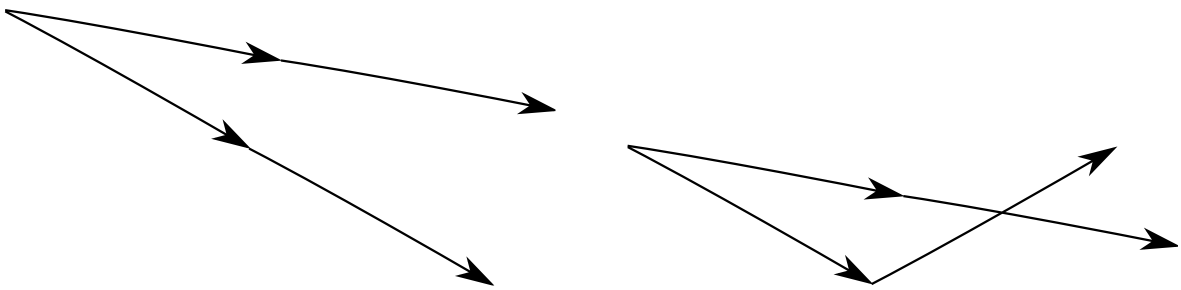
Now let us introduce some mathematics. Denote the returns on the two assets in period t as $r_{i,t}$, $i = 1, 2$ and let r^X and r^T represent the returns on the cross-sectional and time-series strategies respectively. The difference between the two strategies is given by (assuming return distributions are continuous on a support that contains zero)

$$\begin{aligned} r^X - r^T =& 2r_{1,t}\mathbf{1}_{\{r_{1,t-1}<0, r_{2,t-1}<0, r_{1,t-1}>r_{2,t-1}\}} + 2r_{2,t}\mathbf{1}_{\{r_{1,t-1}<0, r_{2,t-1}<0, r_{1,t-1}<r_{2,t-1}\}} \\ &- 2r_{1,t}\mathbf{1}_{\{r_{1,t-1}>0, r_{2,t-1}>0, r_{1,t-1}<r_{2,t-1}\}} - 2r_{2,t}\mathbf{1}_{\{r_{1,t-1}>0, r_{2,t-1}>0, r_{1,t-1}>r_{2,t-1}\}} \text{ a.s.} \end{aligned} \quad (2.1)$$

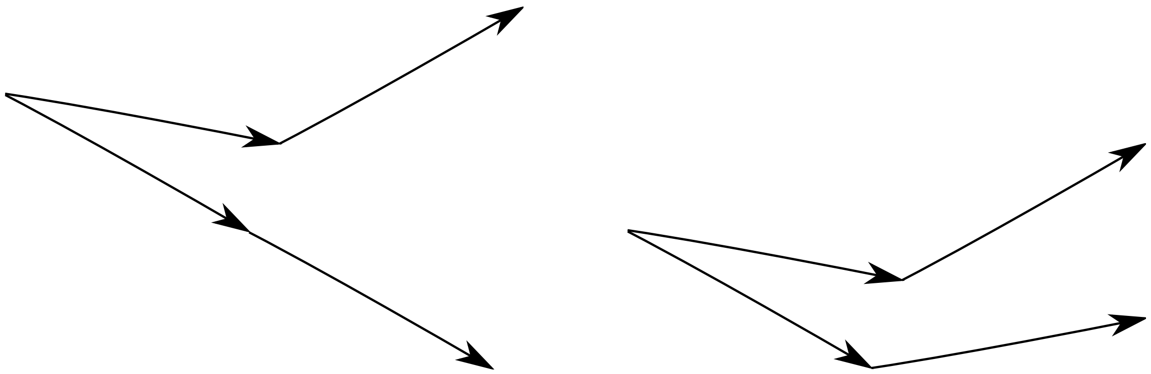
The first line is the down|down case (which has two terms, which depend on which asset is on top) and the second line is the up|up case. The cross-sectional strategy outperforms if the top (bottom) asset reverses in the down|down (up|up) scenario and the outperformance is equal to twice the (absolute) return on the asset. This is easy to see because the strategies take the same position in one asset, but opposite positions in the other, so the difference between them is twice the return of the asset on which they differ.

I will consider two special cases. These cases are considered for their simplicity rather than realism in order to build up an intuition for the behaviour of momentum strategies. In the first case momentum is caused by auto-covariance in asset returns and in the second case the momentum is caused by cross-serial covariances, also called *lead-lag relationships* between assets because one asset leads the other

[6] It would be plausible if there were overreaction in the most extreme assets.



(a) *Scenario 1: TS wins* **(b)** *Scenario 2: TS wins*



(c) *Scenario 3: XS wins* **(d)** *Scenario 4: XS wins*

Figure 2.4: ***Two-period scenarios (down/down).*** *In a two asset market with two periods and both assets going down in the formation period we distinguish four scenarios as each asset can go up or down in the holding period. In two scenarios a time-series strategy outperforms and in two scenarios a cross-sectional strategy outperforms.*

(a nice decomposition of momentum profits according to these sources is given in section 2.3.1). One may expect the former to favour time-series momentum and the latter to favour cross-sectional momentum.

Case I (auto-covariances):

I suppose that asset returns follow stationary AR(1) processes with

$$r_{i,t} = c + \phi r_{i,t-1} + \epsilon_{i,t},$$

where the innovations are independent of the past and have the same distribution for both assets. For simplicity I equalise the parameters of both of the AR(1) processes (so both assets have the same distribution) and I also assume the assets are independent, completely eliminating lead-lag relationships. This also means that the probability of each asset being on top is 0.5 for each configuration 2.2a, 2.2c, 2.2b in figure 2.2. We then have that

$$\begin{aligned}
&\mathrm{E}[r_{1,t}\mathbf{1}_{\{r_{1,t-1}<0, r_{2,t-1}<0, r_{1,t-1}>r_{2,t-1}\}}] \\
&\quad = \mathrm{E}[r_{1,t}|r_{1,t-1}<0, r_{2,t-1}<0, r_{1,t-1}>r_{2,t-1}]\mathrm{P}(r_{1,t-1}<0, r_{2,t-1}<0, r_{1,t-1}>r_{2,t-1}) \\
&\quad = \frac{1}{2}\mathrm{E}[r_{1,t}|r_{1,t-1}<0, r_{1,t-1}>r_{2,t-1}]\mathrm{P}(r_{1,t-1}<0)\mathrm{P}(r_{2,t-1}<0) \\
&\quad = \frac{1}{2}\mathrm{E}[r_{1,t}|r_{1,t-1}<0, r_{1,t-1}>r_{2,t-1}]\mathrm{P}(r_{1,t-1}<0)^2 \qquad (2.2)
\end{aligned}$$

and similarly for the other terms. Then we have

$$
\begin{aligned}
&E[r^X - r^T] \\
&= 2\mathrm{E}[r_{1,t}|r_{1,t-1} < 0, r_{1,t-1} > r_{2,t-1}]\mathrm{P}(r_{1,t-1} < 0)^2 - 2\mathrm{E}[r_{1,t}|r_{1,t-1} > 0, r_{1,t-1} < r_{2,t-1}]\mathrm{P}(r_{1,t-1} > 0)^2 \qquad (2.3) \\
&= 2\underbrace{c(\mathrm{P}(r_{1,t-1} < 0)^2 - \mathrm{P}(r_{1,t-1} > 0)^2)}_{\mathbf{A}} + \qquad (2.4) \\
&\underbrace{2\phi\left(\mathrm{E}[r_{1,t-1}|r_{1,t-1} < 0, r_{1,t-1} > r_{2,t-1}]\mathrm{P}(r_{1,t-1} < 0)^2 - \mathrm{E}[r_{1,t-1}|r_{1,t-1} > 0, r_{1,t-1} < r_{2,t-1}]\mathrm{P}(r_{1,t-1} > 0)^2\right)}_{\mathbf{B}}
\end{aligned}
$$

Note that when ϕ is positive, i.e. when there is positive auto-covariance, then **B** is negative. Whether time-series or cross-sectional outperforms then depends on the value c. If our distributions are symmetric (about their mean) then **A** will be negative whenever c is positive (it is unlikely that c will be negative as assets generally have a positive mean). Thus the time-series strategy will outperform the cross-sectional strategy if momentum is from auto-covariances under these simplifying assumption. (Note also that for $\phi < 0$, that is if there is mean-reversion, **B** is positive and cross-sectional will outperform if the effect of a positive mean is not too large).

Looking only at equation (2.3) (the second line in the panel above) we see that if the cross-sectional strategy outperforms the time-series strategy (we can no longer assume symmetric distributions – it would appear that a large left tail would be needed) then we need c to be large enough that the first term is positive, i.e. so that the expected return after a negative return remains positive (but if c is too large $\mathrm{P}(r_{1,t-1} < 0)$ will be small, counteracting this). In this case the cross-sectional strategy will outperform in the down|down case as we then have reversals of the negative returns, but will lose out in the up|up case.

Case II (lead-lag relationship):

Now I assume that asset 1 has the following returns

$$r_{1,t} = c + \phi r_{2,t-1} + \epsilon_t$$

where again the innovations are assumed to be independent of the past. I will also assume that $r_{2,t}$ is independent of the past and $r_{1,t}$ is independent of its own past. Thus the only momentum effect must come from the (asymmetric) lead-lag relationship between asset 1 and 2. Asset 2 leads asset 1. Under these assumptions we have, denoting $\mathrm{P}_{i<j} := \mathrm{P}(r_{i,t-1} < r_{j,t-1})$, $U_1 := \{r_{2,t-1} < 0, r_{1,t-1} > r_{2,t-1}\}$, $U_2 := \{r_{2,t-1} > 0, r_{1,t-1} < r_{2,t-1}\}$.

$$
\begin{aligned}
&E[r^X - r^T] \\
&= 2\mathrm{E}[r_{1,t}|r_{2,t-1} < 0, r_{1,t-1} > r_{2,t-1}]\mathrm{P}(r_{1,t-1} < 0)\mathrm{P}(r_{2,t-1} < 0)\mathrm{P}_{1>2} \\
&\quad - 2\mathrm{E}[r_{1,t}|r_{2,t-1} > 0, r_{1,t-1} < r_{2,t-1}]\mathrm{P}(r_{1,t-1} > 0)\mathrm{P}(r_{2,t-1} > 0)\mathrm{P}_{1<2} \\
&\quad + 2\mathrm{E}[r_{2,t}](\mathrm{P}(r_{1,t-1} < 0)\mathrm{P}(r_{2,t-1} < 0)\mathrm{P}_{1<2} - \mathrm{P}(r_{1,t-1} > 0)\mathrm{P}(r_{2,t-1} > 0)\mathrm{P}_{1>2} \\
&= \underbrace{2c(\mathrm{P}(r_{1,t-1} < 0)\mathrm{P}(r_{2,t-1} < 0)\mathrm{P}_{1>2} - \mathrm{P}(r_{1,t-1} > 0)\mathrm{P}(r_{2,t-1} > 0)\mathrm{P}_{1<2})}_{\mathbf{A}} + \\
&\quad \underbrace{2\mathrm{E}[r_{2,t}](\mathrm{P}(r_{1,t-1} < 0)\mathrm{P}(r_{2,t-1} < 0)\mathrm{P}_{1<2} - \mathrm{P}(r_{1,t-1} > 0)\mathrm{P}(r_{2,t-1} > 0)\mathrm{P}_{1>2})}_{\mathbf{B}} + \qquad (2.5) \\
&\quad \underbrace{2\phi\,(\mathrm{E}[r_{2,t-1}|U_1]\mathrm{P}(r_{1,t-1} < 0)\mathrm{P}(r_{2,t-1} < 0)\mathrm{P}_{1>2} - \mathrm{E}[r_{2,t-1}|U_2]\mathrm{P}(r_{1,t-1} > 0)\mathrm{P}(r_{2,t-1} > 0)\mathrm{P}_{1<2})}_{\mathbf{C}}
\end{aligned}
$$

Note that if ϕ is negative then **C** will be positive. It seems reasonable to suppose that c and $\mathbf{E}r_2, t$ are positive and so it may be that **A** and **B** are negative, if the probability $\mathrm{P}(r_{1,t-1} < 0)\mathrm{P}(r_{2,t-1} < 0)$ is small. The cross-sectional strategy will thus outperform if ϕ is large enough (in absolute value). Thus a strong negative lead-lag relationship means the cross-sectional strategy will outperform. Scenarios 3 and 4 become more important.

Now I want to examine what happens if we increase the means of both assets by the same amount. For this suppose that $r_{2,t} = c_2 + \epsilon_{2,t}$ and $c = -(\phi - 1)c_2 + c_3$ and consider what happens if $c_2 \to \infty$. $\mathrm{P}_{1<2}$ and $\mathrm{P}_{2<1}$ remain constant, but $\mathrm{P}(r_{1,t-1} < 0)\mathrm{P}(r_{2,t-1} < 0)$ goes to zero. The result is that **A** decreases to negative infinity. The first term in **B** and **C** vanish. Now we are left with

$$\begin{aligned}
& -2(\mathrm{E}[r_{2,t}]\mathrm{P}_{1>2} - \phi \mathrm{E}[r_{2,t-1}|U_2]P_{1<2})\mathrm{P}(r_{1,t-1}>0)\mathrm{P}(r_{2,t-1}>0) \\
& = -2(\mathrm{E}[\epsilon_{2,t}]\mathrm{P}_{1>2} - \phi \mathrm{E}[\epsilon_{2,t-1}|U_2]P_{1<2})\mathrm{P}(r_{1,t-1}>0)\mathrm{P}(r_{2,t-1}>0) \\
& \to 2\phi \mathrm{E}[\epsilon_{2,t-1}|U_3]P_{1<2}\mathrm{P}(r_{1,t-1}>0)\mathrm{P}(r_{2,t-1}>0)
\end{aligned} \tag{2.6}$$

where now

$$\begin{aligned}
U_2 &= \{c_2 + \epsilon_{2,t-1} > 0, \phi\epsilon_{2,t-2} + \epsilon_{1,t-1} - \epsilon_{2,t-1} + c_3 < 0\} \\
U_3 &= \{\phi\epsilon_{2,t-2} + \epsilon_{1,t-1} - \epsilon_{2,t-1} + c_3 < 0\}.
\end{aligned}$$

The limit above in (2.6) is finite. Thus in the presence of a strong mean component of trend scenario 1 becomes more and more likely and the cross-sectional strategy underperforms.

Based on the reasoning in this section we can formulate the following proposition:

Proposition 2.1. *Consider a simple time-series and cross-sectional strategy in a two asset market with return difference as in (2.1). Assume below that innovations $\epsilon_{i,t}$ are independent of the past.*

I. *(auto-covariances) If the assets are independent and follow stationary AR(1) processes $r_{i,t} = c + \phi r_{i,t} + \epsilon_{i,t}$ with $c, \phi > 0$ and have equal symmetric distributions then $E[r^X - r^T] \leq 0$.*

II. *(lead-lag relationship) If $r_{1,t} = c + \phi r_{2,t} + \epsilon_{1,t}$, $r_{1,t}$ is independent of its own past return and $r_{2,t}$ is independent of the past, then for ϕ negative and large enough (in absolute value), $E[r^X - r^T] \geq 0$.*

2.3 Linear and signed strategies

Here I define the two sets of strategies on which the majority of my theoretical analysis will be based. The first are two linear strategies for which some theory is already available in the literature and which are examined briefly in Moskowitz *et al.* (2012). These strategies' expected returns have a simple decomposition which will be examined empirically later. The second is a set of strategies based on the sign of past returns which are specific cases of market timing strategies as examined in Hallerbach (2011).

2.3.1 Linear strategies

We suppose we have N assets and that at the start of the period from t to $t+k$ we invest $w_{i,t}$ of our notional capital in asset i, which has (excess) return $r_{i,t,t+k}$. Denote $\boldsymbol{r}_{t,t+k} = (r_{i,t,t+k})_i$, $\boldsymbol{\mu}_{t,t+k} = \mathrm{E}[\boldsymbol{r}_{t,t+k}]$.

Lewellen (2002) (following Lo and MacKinlay (1990)) defines a simple linear cross-sectional momentum strategy with

$$w_{i,t} = \frac{1}{N}(r_{i,t-j,t} - \bar{r}_{t-j,t})$$

and $\bar{r}_{t-j,t} = \frac{1}{N}\sum_i r_{i,t-j,t}$.

This strategy invests proportionally to an asset's deviation from the mean (market) return over the previous period. It invests more heavily (in total in terms of gross exposure – net exposure is of course zero) when cross-sectional dispersion is higher.

Lewellen (2002) (still following Lo and MacKinlay (1990)) shows that the strategy has (unconditional) expected profits of

$$\begin{aligned}
E[r^X_{t,t+k}] &= \frac{\mathrm{tr}(\Omega_{t-j,t,t+k})}{N} - \frac{\mathbf{1}'\Omega_{t-j,t,t+k}\mathbf{1}}{N^2} + \frac{\boldsymbol{\mu}'_{t-j,t}\boldsymbol{\mu}_{t,t+k}}{N} - \frac{\sum_i \boldsymbol{\mu}_{t-j,t}(i)\sum_i \boldsymbol{\mu}_{t,t+k}(i)}{N^2} \\
&= \frac{N-1}{N^2}\mathrm{tr}(\Omega_{t-j,t,t+k}) - \frac{1}{N^2}(\mathbf{1}'\Omega_{t-j,t,t+k}\mathbf{1} - \mathrm{tr}(\Omega_{t-j,t,t+k})) + \sigma^2_{\boldsymbol{\mu}_{t-j,t},\boldsymbol{\mu}_{t,t+k}}
\end{aligned} \tag{2.7}$$

I have used an X to indicate cross-sectional momentum and $\Omega_{t-j,t,t+k} = \mathrm{Cov}(\mathbf{r}_{t-j,t}, \mathbf{r}_{t,t+k})$. This decomposes the profits from the strategy into three terms in (2.7), which are a portion due to auto-covariances (of an asset with its past), cross-serial covariances (of an asset with the past of other assets) and a term depending on mean asset returns. This last term is the (cross-sectional) covariance of mean returns.

Moskowitz *et al.* (2012) similarly define a time-series momentum strategy with $w_{i,t} = \frac{1}{N}r_{i,t-j,t}$ and this has a decomposition

$$E[r^T_{t,t+k}] = \frac{\mathrm{tr}(\Omega_{t-j,t,t+k})}{N} + \frac{\boldsymbol{\mu}'_{t-j,t}\boldsymbol{\mu}_{t,t+k}}{N} \tag{2.8}$$

I use a T to denote time-series momentum. Here we see that the cross-serial covariances do not matter any more – we are concerned only with an asset's returns relative to its own past. If mean returns tend to be of the same sign (either positive or negative, so $\boldsymbol{\mu}'_{t-j,t}\boldsymbol{\mu}_{t,t+k} > 0$) this contributes to the profits of the strategies (and of course the opposite is true if the signs differ). Time-series momentum profits from a general tendency of trend in the market (generally positive or negative returns).

Cross-sectional momentum profits from cross-sectional dispersion in mean returns (Moskowitz *et al.*, 2012). This can most clearly be seen if we take $j = k = 1$ and assume stationarity so that $\boldsymbol{\mu}_{t-k,t} = \boldsymbol{\mu}_{t,t+k} = \boldsymbol{\mu}$, then we get

$$E[r^X_{t,t+1}] = \frac{N-1}{N^2}\mathrm{tr}(\Omega) - \frac{1}{N^2}(\mathbf{1}'\Omega\mathbf{1} - \mathrm{tr}(\Omega)) + \sigma^2_{\boldsymbol{\mu}} \tag{2.9}$$

$$E[r^T_{t,t+1}] = \frac{\mathrm{tr}(\Omega)}{N} + \frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{N} \tag{2.10}$$

The last term in (2.9) is the (positive) variance of the cross-sectional means (more generally in (2.7) it is a covariance, which could be positive or negative). Time-series momentum depends on squared mean returns. I will examine this again later. Empirically, the relationship with dispersion and squared returns appears to be more complicated.

Both strategies benefit from positive auto-covariances and cross-sectional momentum also benefits from generally negative serial auto covariances or lead-lag effects – where a poor performance in one stock predicts higher returns in another stock (and vice versa) (Lewellen, 2002).

If assets prices are uncorrelated across time then only the effect due to mean returns remains. This may not be considered true momentum as it indicates no real relationship between past and future asset prices. In particular if (log) asset prices follow a random walk with drift $\boldsymbol{\mu}$, then (log) returns are

$$E[r^X_{t,t+k}] = \frac{jk\boldsymbol{\mu}'\boldsymbol{\mu}}{N} - \frac{jk(\sum_i \boldsymbol{\mu}(i))^2}{N^2}$$
$$E[r^T_{t,t+k}] = \frac{jk\boldsymbol{\mu}'\boldsymbol{\mu}}{N}.$$

2.3.2 Signed strategies

Time-series

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). This is a simple time-series strategy, much like that defined by Moskowitz *et al.* (2012).

The investment makes a positive return (i.e. the prediction is correct) if $S_t := \mathrm{sign}(r_{t-1}r_t) = 1$. I suppose this signed variable is independent of $|r_t|$.[7] 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 variously be written as

$$\begin{aligned} r^T_t &= S_t|r_t| \\ &= (\mathbf{1}_{\{S=1,r_t>0\}} - \mathbf{1}_{\{S=1,r_t<0\}} + \mathbf{1}_{\{S=-1,r_t<0\}} - \mathbf{1}_{\{S=-1,r_t>0\}})r_t \end{aligned}$$

The expected return of the strategy is (as shown in Hallerbach (2011))

$$(p-q)\mathrm{E}|r_t|$$

The return is positive if the frequency of correct predictions is larger than that of false predictions. This contrasts with the findings of Potters and Bouchaud (2005) who conclude that a trend-following strategy's potential profits can be independent of the frequency of winning trades and that in fact the frequency of winning trades is lower than that of losing trades. This can happen if the gain per winning trade is larger than the loss per losing trade, for instance. In our strategy where being correct holds no prediction for the size of the return this is not the case.

The variance (from the same source) is

$$(p+q)\mathrm{Var}(|r_t|) + (p+q-(p-q)^2)(\mathrm{E}|r_t|)^2$$

[7]This may not be a realistic assumption, but it simplifies the math.

Now consider a market of N assets $r_{i,t}$ $i = 1...N$ and denote $S_{i,t} := \text{sign}(r_{i,t-1}r_{i,t})$. A simple time-series strategy would simply diversify across all the assets available in the market, so

$$r_t^T = \frac{1}{N}\sum_{i=1}^{N} S_{i,t}|r_{i,t}|$$

Moskowitz *et al.* (2012) define just such a strategy, but based on normalised returns (and a longer formation period than holding period).

Cross-sectional

I also define a cross-sectional strategy based on the same idea as above. Here I suppose the asset returns to be invested in are $\tilde{r}_{i,t} := r_{i,t} - \bar{r}_{i,t}$ with $\tilde{S}_{i,t} := \text{sign}(\tilde{r}_{i,t-1}\tilde{r}_{i,t})$ defined in the obvious manner. That is we 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. 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 asset's own past returns or from the deviation of the return from the average return.

We again diversify over all the strategies of this type to get

$$\begin{aligned} r_t^X &= \frac{2}{N}\sum_{i=1}^{N} \tilde{S}_{i,t}|\tilde{r}_{i,t}| \\ &= \frac{2}{N}\sum_{i=1}^{N} \text{sign}(\tilde{r}_{i,t})(r_{i,t} - \frac{1}{N}\sum_j r_{j,t}) \\ &= \frac{2}{N}\sum_{i=1}^{N} \text{sign}(\tilde{r}_{i,t})r_{i,t} - \frac{2}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{N}\text{Sign}(\tilde{r}_{i,t})r_{j,t} \\ &= \frac{2}{N}\sum_{i=1}^{N} r_{i,t}\left(\text{sign}(\tilde{r}_{i,t}) - \frac{1}{N}\sum_j \text{sign}(\tilde{r}_{j,t})\right) \qquad (2.11) \end{aligned}$$

In agreement with previous nomenclature I refer to this as a *signed cross-sectional* strategy. This should be closely related to the quantile cross-sectional strategies. For instance if half the assets are below average and half above average we get a 2 quantile strategy. I have multiplied by 2 merely so that the scale is then similar (1 in each of the long and short legs). This strategy buys all the stocks above the average (and places an equal weight on them) and sells all the stocks below the average (and places an equal weight on them, which may differ from the weight on bought stocks). So we should expect differences in this strategy and the quantile strategy to result (mostly) from a skew in the cross-sectional dispersion of stocks.

2.4 Volatility weighting

Here I consider some theory regarding the practice of volatility weighting. The analysis is heavily influenced by the work in Hallerbach (2011) and Hallerbach (2012). I particularly study the effect of volatility weighting on the Sharpe ratio of a strategy, which appears to be the most tractable measure to study. I consider weighting a strategy by its own volatility or weighting each of the underlying assets (normalised returns).

We shall see that volatility weighting is in fact optimal under various sets of assumptions. However, these assumptions are not very realistic. Though reality may be close enough to these ideal assumptions to expect the theory to hold in many cases, it should be clear that volatility weighting is not a panacea and that it may not always improve a strategy. For instance, in the analysis I shall assume volatility is perfectly forecastable. However, in practice volatility is not even directly observable and so the efficacy of this strategy will be hampered by the ability to forecast volatility. It shall also be seen that if a large portion of the returns is independent of volatility then volatility weighing may not work. Empirically, we shall see in later chapters, neither of these violations of the assumptions seems to be so large so as to invalidate the results completely.

I will assume throughout that returns are (discrete time) processes adapted to some filtration, $(\mathcal{F}_t)_t$.

2.4.1 By a strategy's own volatility

I look at a simple description of asset returns depending on their ex-ante volatility and consider conditions under which the Sharpe ratio is improved by volatility weighting.

Consider an asset (or strategy) with the following return decomposition

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

where σ_t is a positive predictable process and ϵ_t has zero mean and unit variance given $\mathcal{F}_{t-1}$. σ_t is then the conditional volatility of the returns. We may expect for a momentum strategy that γ is negative.[8] Of course if a higher ex-ante volatility is associated with an increased risk-premium, one would expect a positive γ. The mean return, variance and Sharpe ratio of this strategy are then given by

$$\begin{aligned}
\mathrm{E}[r_t] &= \alpha + \gamma\mathrm{E}[\sigma_t] \\
\mathrm{Var}(r_t) &= \gamma^2\mathrm{Var}(\sigma_t) + \mathrm{E}(\sigma_t^2) \\
\mathrm{Sharpe}(r_t) &= \frac{\alpha + \gamma\mathrm{E}[\sigma_t]}{\sqrt{\gamma^2\mathrm{Var}(\sigma_t) + \mathrm{E}(\sigma_t^2)}}
\end{aligned} \tag{2.12}$$

Now consider the alternative volatility-weighted strategy

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

Interestingly, if α is positive, we have now *induced* a negative relationship with volatility. We should not therefore necessarily be surprised if volatility weighting does not negate a negative relationship with volatility. The volatility weighted strategy has mean, variance and Sharpe ratio given by

$$\begin{aligned}
\mathrm{E}[r_t^*] &= \alpha\mathrm{E}[\frac{1}{\sigma_t}] + \gamma \\
\mathrm{Var}(r_t^*) &= \alpha^2\mathrm{Var}(\frac{1}{\sigma_t}) + 1 \\
\mathrm{Sharpe}(r_t^*) &= \frac{\alpha\mathrm{E}[\frac{1}{\sigma_t}] + \gamma}{\sqrt{\alpha^2\mathrm{Var}(\frac{1}{\sigma_t}) + 1}} \\
&= \frac{\alpha + \gamma\mathrm{E}[\frac{1}{\sigma_t}]^{-1}}{\sqrt{\alpha^2(\frac{\mathrm{E}[\frac{1}{\sigma_t^2}]}{\mathrm{E}[\frac{1}{\sigma_t}]^2} - 1) + \mathrm{E}[\frac{1}{\sigma_t}]^{-2}}}
\end{aligned} \tag{2.13}$$

Now firstly note that $\mathrm{Var}(\frac{1}{\sigma_t})$ can be arbitrarily large. In fact it can be infinite. Thus it is not immediate that volatility weighting improves the Sharpe ratio. In (2.13), I have made the numerator and denominator of the weighted strategy's Sharpe ratio more comparable with (2.12). We can compare the numerator and denominator. I consider first two special cases: in the first the portion of returns that do not depend on volatility is zero and in the second the there is no dependence on volatility at all.

Case I: $\alpha = 0$ (no portion of returns independent of volatility)

$$\begin{aligned}
\mathrm{Sharpe}(r_t) &= \frac{\gamma\mathrm{E}[\sigma_t]}{\sqrt{\gamma^2\mathrm{Var}(\sigma_t) + \mathrm{E}(\sigma_t^2)}} = \frac{\gamma}{\sqrt{(\gamma^2 + 1)\frac{\mathrm{Var}(\sigma_t)}{\mathrm{E}[\sigma_t]^2}}} \\
\mathrm{Sharpe}(r_t^*) &= \gamma
\end{aligned}$$

Here we have that $\frac{\mathrm{E}[\sigma_t]}{\sqrt{\gamma^2\mathrm{Var}(\sigma_t)+\mathrm{E}(\sigma_t^2)}} \leq 1$ and the Sharpe ratio improves (regardless of the sign of γ). Note that volatility weighting is more effective the more variable the volatility, i.e. the larger the coefficient of variation $\frac{\sqrt{\mathrm{Var}(\sigma_t)}}{\mathrm{E}[\sigma_t]}$, which is as noted in Hallerbach (2012). However, it might be unreasonable to

[8]This is not tested directly in the literature – Wang and Xu (2009) show a negative relationship with market volatility and Barroso and Santa-clara (2012) show empirically that volatility weighting improves (cross-sectional) momentum returns (and reduces risk) using the momentum strategy's own volatility.

suppose that $\alpha = 0$.

Case II: $\gamma = 0, \alpha > 0$ (no dependence of returns on volatility)

$$\begin{aligned}\text{Sharpe}(r_t) &= \frac{\alpha}{\sqrt{\mathrm{E}(\sigma_t^2)}} \\ \text{Sharpe}(r_t^*) &= \frac{\alpha}{\sqrt{\alpha^2\left(\frac{\mathrm{E}[\frac{1}{\sigma_t^2}]}{\mathrm{E}[\frac{1}{\sigma_t}]^2} - 1\right) + \mathrm{E}[\frac{1}{\sigma_t}]^{-2}}}\end{aligned}$$

Now there is a range

$$\alpha \in \left[0, \sqrt{\frac{\mathrm{E}[\sigma_t^2]\mathrm{E}[\frac{1}{\sigma_t}]^2 - 1}{\mathrm{Var}(\frac{1}{\sigma_t})}}\right] = \left[0, \sqrt{\frac{\mathrm{E}[\sigma_t^2]}{\mathrm{CV}(\frac{1}{\sigma_t})^2} - \frac{1}{\mathrm{Var}(\frac{1}{\sigma_t})}}\right] \tag{2.14}$$

over which the Sharpe ratio is improved by volatility weighting, where I have used CV to denote a coefficient of variation. Outside this range it is worsened. We therefore cannot have α too large. If σ_t is constant the range is not valid, but then the Sharpe ratio is unaffected by volatility weighting. Note that the given range is increasing in $\mathrm{E}[\sigma_t^2]$ which is related to the size and variability of volatility. The range is also decreasing in the variability of $\frac{1}{\sigma_t}$, through the variance (in the first interval shown) or the coefficient of variation (in the second interval). These elements are, however, difficult to manipulate separately and so a clearer conclusion than this is hard to draw. We do, however, have that increasing volatility by some constant factor $c > 1$, i.e. using $c\sigma_t$, increases the size of the range by the same factor (this increases the mean and variance of volatility, but not the coefficient of variation).

I now consider two more general cases. In the first the dependence on volatility is negative and in the second the dependence on volatility is positive.

Case III: $\gamma < 0$ (negative dependence on volatility)

I will assume that expected returns are positive, which means $\alpha \geq -\gamma\mathrm{E}[\sigma_t]$. We have (by Jensen's inequality) $\mathrm{E}[\frac{1}{\sigma_t}]^{-1} \leq \mathrm{E}[\sigma_t]$ and so $\alpha + \gamma\mathrm{E}[\frac{1}{\sigma_t}]^{-1} \geq \alpha + \gamma\mathrm{E}[\sigma_t]$. Thus volatility weighting will improve the Sharpe ratio if the α is such that the numerator decreases. Thus a sufficient (but not necessary) condition for the Sharpe ratio to improve is that

$$\alpha \in [-\gamma\mathrm{E}[\sigma_t], \Lambda] \tag{2.15}$$

with

$$\Lambda := \sqrt{\frac{(\gamma^2\mathrm{Var}(\sigma_t) + \mathrm{E}[\sigma_t^2])\mathrm{E}[\frac{1}{\sigma_t}]^2 - 1}{\mathrm{E}[\frac{1}{\sigma_t^2}] - \mathrm{E}[\frac{1}{\sigma_t}]^2}} \tag{2.16}$$

Case IV: $\gamma > 0$ (positive dependence on volatility)

In this case the denominator of the Sharpe ratio decreases. However, by considering the range of α for which the numerator decreases, we derive a necessary (but not sufficient) condition for improvement, again assuming expected returns are positive:

$$\alpha \in [\max(-\Lambda, -\gamma\mathrm{E}[\sigma_t]), \Lambda] \tag{2.17}$$

For the above two cases the size of the ranges in (2.15) and (2.17) again increase when volatility is increased by a constant factor. We also see that Λ is increasing in $|\gamma|$, meaning volatility weighting is more likely to be effective if the dependence on volatility is high. In practice, of course, volatility weighting will only be effective if we can forecast σ_t, so a dependence of volatility on the past is needed (here I have assumed that volatility is predictable).

The above results are in a similar context to that of Hallerbach (2012). The results above 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 normalised returns, which will not be the case if α is non-zero. Some claim can still be made that the result should hold approximately. I 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.

Using a similar approach to Hallerbach (2012) we can say more for case I above. Let v_t be a predictable process giving the position taken in the weighted return $r_t^* = \gamma + \epsilon_t$. Suppose $\gamma > 0$. Note that we can take $\mathrm{E}[v_t] > 0$ as otherwise we get a negative (or zero) Sharpe-ratio. Then the Sharpe ratio of the strategy is

$$\frac{\gamma \mathrm{E}[v_t]}{\sqrt{\gamma^2 \mathrm{Var}(v_t) + \mathrm{E}[v_t^2]}} = \frac{\gamma}{\sqrt{\gamma^2 \frac{\mathrm{Var}(v_t)}{(\mathrm{E}[v_t])^2} + \frac{\mathrm{E}[v_t^2]}{(\mathrm{E}[v_t])^2}}} = \frac{\gamma}{\sqrt{(\gamma^2+1)\frac{\mathrm{Var}(v_t)}{(\mathrm{E}[v_t])^2} + 1}}$$

The numerator is minimised when v_t is deterministic, say $v_t = \varsigma_t$. The position in the unweighted return is then $\frac{\varsigma_t}{\sigma_t}$ where ς_t is non-negative and non-random. In practice one would usually set $\varsigma_t = \varsigma$ for all t and choose ς to achieve some desired volatility target. For $\gamma < 0$ it is similarly easy to show that the optimal strategy is $v_t = -\varsigma_t$ – i.e. short the strategy before applying volatility weighting.

The above gives an of example where the result of Hallerbach (2012) holds (not just approximately) for a strategy with non-zero mean returns. This case is still not very realistic but it does give a better indication of (approximately) what is needed in order for volatility weighting to work well.

Based on the reasoning in this section we can formulate the following proposition:

Proposition 2.2. *Let $r_t = \alpha + \gamma\sigma_t + \epsilon_t\sigma_t$ be the returns of a strategy with σ_t a strictly positive predictable process, ϵ_t having zero mean and unit variance given $\mathcal{F}_{t-1}$. Then the following conditions are necessary (sufficient) for the volatility-weighted strategy $r_t^* = \gamma + \frac{\alpha}{\sigma_t} + \epsilon_t$ to have a higher Sharpe ratio and positive expected returns for $\gamma > 0$ ($\gamma \le 0$):*

$$\alpha \in [\max(-\gamma \mathrm{E}[\sigma_t], -\Lambda), \Lambda] \;\; \textit{if } \gamma > 0$$
$$\alpha \in [-\gamma \mathrm{E}[\sigma_t], \Lambda] \;\; \textit{if } \gamma \le 0$$

with

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

In the case of $\alpha = 0$ and $\gamma > 0$ volatility weighting is optimal (in that it maximises the Sharpe ratio) among all predictable weightings $w_t = \frac{v_t}{\sigma_t}$ of r_t (if $\gamma < 0$ consider shorting the strategy). Volatility weighting corresponds to choosing v_t deterministic.

2.4.2 By the volatility of the underlying assets

Time-series

Consider a signed time-series strategy as in section 2.3.2 on an asset with returns $r_t = \epsilon_t\sigma_t$ where ϵ_t has mean zero and unit variance (given $\mathcal{F}_{t-1}$). σ is a positive predictable process as before. We would also like to have $\mathrm{E}[\sigma_t|\epsilon_t|] = \mathrm{E}[\sigma_t]\mathrm{E}[|\epsilon_t|]$ and in proposition 2.3 I provide some assumptions that guarantee this. Note that this implies that the underlying has zero expected returns. We choose this for simplicity, not realism (the proposition below relaxes this somewhat). The Sharpe ratio of a timing strategy under these assumptions is

$$\begin{aligned} &\frac{(p-q)\mathrm{E}\sigma_t \mathrm{E}|\epsilon_t|}{\sqrt{(p+q)(\mathrm{E}[\sigma_t^2] - \mathrm{E}[\sigma_t]^2\mathrm{E}[|\epsilon_t|]^2) + (p+q-(p-q)^2)(\mathrm{E}[\sigma_t])^2(\mathrm{E}|\epsilon_t|)^2}} \\ &= \frac{(p-q)\mathrm{E}|\epsilon_t|}{\sqrt{(p+q)(\frac{\mathrm{E}[\sigma_t^2]}{(\mathrm{E}[\sigma_t])^2} - \mathrm{E}[|\epsilon_t|]^2) + (p+q-(p-q)^2)(\mathrm{E}|\epsilon_t|)^2}} \end{aligned} \tag{2.18}$$

Now consider instead an investment in the volatility weighted returns $r_t^* = \frac{r_t}{\sigma_t} = \epsilon_t$, which is a process with constant unit variance. The Sharpe ratio of the timing strategy is now

$$\frac{(p-q)\mathrm{E}|\epsilon_t|}{\sqrt{(p+q)(1 - \mathrm{E}[|\epsilon_t|]^2) + (p+q-(p-q)^2)(\mathrm{E}|\epsilon_t|)^2}} \tag{2.19}$$

Because $\frac{\mathrm{E}[\sigma_t^2]}{(\mathrm{E}[\sigma_t])^2}$ is larger than 1 we see that volatility weighting does improve the Sharpe ratio of the strategy (assuming the strategy is profitable) and the more variable the (conditional) volatility of the returns the greater the improvement.

Note that as we saw in the previous section, including a mean that does not depend on volatility in the asset returns can probably change this conclusion. However, we may expect that as long as the mean is small compared to the volatility the effects shown above will dominate.

We in fact have a stronger result than just that volatility weighting improves the Sharpe ratio. We can take a similar approach to section 2.4.1 and suppose we can weight the normalised returns $r_t^* = \frac{r_t}{\sigma_t} = \epsilon_t$ by some predictable process v_t. We suppose that this does not change our ability to predict the sign of the returns. For instance if v_t is strictly positive this will be the case. We also want to have $\mathrm{E}[v_t|\epsilon_t|] = \mathrm{E}[v_t]\mathrm{E}[|\epsilon_t|]$ and conditions for this are in proposition 2.3. Note, however, that for our strategy to have timing ability, we cannot assume ϵ_t independent of the past (not if our timing is derived from past returns).

Thus we are now running our timing strategy on $v_t r_t^*$. Under the above assumptions the Sharpe ratio becomes

$$\begin{aligned} &\frac{(p-q)\mathrm{E}|v_t|\mathrm{E}|\epsilon_t|}{\sqrt{(p+q)\left(\mathrm{E}[v_t^2] - (\mathrm{E}|v_t|)^2(\mathrm{E}|\epsilon_t|)^2\right) + (p+q-(p-q)^2)(\mathrm{E}|v_t|)^2(\mathrm{E}|\epsilon_t|)^2}} \\ &= \frac{(p-q)}{\sqrt{(p+q)\left(\frac{\mathrm{E}[v_t^2]}{(\mathrm{E}|v_t|)^2(\mathrm{E}|\epsilon_t|)^2} - 1\right) + (p+q-(p-q)^2)}} \end{aligned} \tag{2.20}$$

Again this is maximised when v_t is deterministic (provided the strategy is profitable). Thus we have shown that volatility weighting is optimal (at least within a certain class of weighting strategies). The results above can also be easily extended to the case where ϵ_t has a non-zero mean. We now have the following proposition:

Proposition 2.3. *Suppose an asset has returns $r_t = d_t\sigma_t$ with d_t having unit variance given $\mathcal{F}_{t-1}$, σ_t strictly positive and predictable and one of the following*

- *for each t d_t (or $|d_t|$) independent of σ_t*
- *for each t $E[|d_t||\mathcal{F}_{t-1}]$ independent of σ_t*
- *for each t $E[|d_t||\mathcal{F}_{t-1}] = k_t$ for some deterministic k_t.*

Consider a signed time-series strategy as in section 2.3.2 with success and failure probabilities p and q. Consider all weightings $w_t = \frac{v_t}{\sigma_t}$ that are predictable and such that a signed time-series strategy applied to $r_t^* = w_t r_t$ has the same success and failure ratios (in particular this will hold if v_t is strictly positive). In case of the first two conditions above we also require that v_t satisfy these conditions in place of σ_t. In this case volatility weighting, corresponding to v_t being deterministic, is optimal (in that it maximises the Sharpe ratio of the strategy).

In appendix D I provide an example of a process that satisfies the assumptions of the proposition above. The example serves to illustrate that the assumptions are very strict indeed and not likely to be satisfied by traditional momentum models.

Note that assuming independence of the volatility from the ϵ's is not compatible with a GARCH(p, q) process (unless $q = 0$), as ϵ_{t-1} and σ_t are not necessarily independent. The empirical work I do later (based on EWMA volatility estimates) is not compatible with this independence assumption either.

A comparison with the results in Hallerbach (2011) is also in order. The above result is similar but different to the result in the cited paper. Firstly, my result is in a slightly different context: Hallerbach (2011)'s result is for the Sharpe ratio of an aggregate of N returns over a time-period (for example looking at a yearly Sharpe ratio for returns consisting of a sum of 12 monthly returns), whereas the result above considers no such aggregation. The period over which the Sharpe ratio is measured is the same as the period of the returns. In the former context, an improvement in the Sharpe ratio may be seen if volatility is deterministic but varying over time and in the latter a random volatility is necessary. Secondly, my result holds for general distributions (provided they satisfy the assumptions above). Hallerbach (2011) reports results for normal and t-distributions (with at least 4 degrees of freedom) and which are easily seen to also hold for a family of elliptical distributions.

We can take a similar approach for the linear time-series strategy. Suppose we run the strategy on $v_t\epsilon_t$ where v_t is the predictable weighting that we use. The expectation, variance and Sharpe ratio of a time-series strategy are

$$\begin{aligned}
\mathrm{E}[v_{t-1}\epsilon_{t-1}v_t\epsilon_t] &= \mathrm{E}[v_{t-1}v_t]\mathrm{E}[\epsilon_{t-1}\epsilon_t] \\
\mathrm{Var}[v_{t-1}\epsilon_{t-1}v_t\epsilon_t] &= \mathrm{E}[v_{t-1}^2v_t^2]\mathrm{E}[\epsilon_{t-1}^2\epsilon_t^2] - \mathrm{E}[v_{t-1}v_t]^2\mathrm{E}[\epsilon_{t-1}\epsilon_t]^2 \\
\mathrm{Sharpe}(v_{t-1}\epsilon_{t-1}v_t\epsilon_t) &= \left(\frac{\mathrm{E}[v_{t-1}^2v_t^2]}{\mathrm{E}[v_{t-1}v_t]^2}\frac{\mathrm{E}[\epsilon_{t-1}^2\epsilon_t^2]}{\mathrm{E}[\epsilon_{t-1}\epsilon_t]^2} - 1\right)^{-\frac{1}{2}} \qquad (2.21)
\end{aligned}$$

where we need to assume for instance that $v_{t-1}v_t$ is independent of $\epsilon_{t-1}\epsilon_t$ or that the σ_t process is independent of the v_t process. Again the Sharpe ratio is maximised if $v_{t-1}v_t$ is deterministic, corresponding to volatility weighting. Again the independence assumption is not compatible with a GARCH process and the results do not require a zero mean. We have the following proposition:

Proposition 2.4. *Suppose an asset has returns $r_t = d_t\sigma_t$ with d_t having unit variance given $\mathcal{F}_{t-1}$, σ_t strictly positive and predictable and one of the following*

- *for all t $\sigma_{t-1}\sigma_t$ is independent of $d_{t-1}d_t$)*
- *the process $(\sigma_t)_t$ is independent of the process $(d_t)_t$*

Consider all weightings $w_t = \frac{v_t}{\sigma_t}$ with v_t satisfying the relevant condition above in place of σ_t. Consider then a linear time-series strategy as in section 2.3.1 on w_tr_t. In this case volatility weighting, corresponding to v_t being deterministic, is optimal (in that it maximises the Sharpe ratio of the strategy).

One thing worth noting about the above results is that the relationship of returns with volatility is not negative, but positive. Here it is not a volatility timing effect that results in an improved Sharpe ratio, but rather a volatility stabilising effect. In practice, the volatility timing may also be important.

Cross-sectional

I consider a signed cross-sectional strategy, looking at only one component, i.e. considering a timing strategy on $\tilde{r}_{i,t} = r_{i,t} - \bar{r}_t$. We can weight by the volatility of $\tilde{r}_{i,t}$, which is covered in the previous result for time-series strategies. Here I consider a different approach. Suppose that $r_{i,t} = \epsilon_{i,t}\sigma_{i,t}$ as before. We can also consider replacing each underlying asset with some investment $v_{i,t}$ in the volatility weighted asset $\frac{r_{i,t}}{\sigma_t} = \epsilon_{i,t}$. The analysis below is again similar in style to that of Hallerbach (2011).

The Sharpe ratio of the timed strategy on $v_{i,t}\epsilon_{i,t} - \frac{1}{N}\sum_j v_{j,t}\epsilon_{j,t} = \frac{1}{N}((N-1)v_{i,t}\epsilon_{i,t} - \sum_{j\neq i} v_{j,t}\epsilon_{j,t})$ is then

$$\frac{p-q}{\sqrt{(p+q)\Lambda + (p+q-(p-q)^2)}} \qquad (2.22)$$

with

$$\begin{aligned}
\Lambda &= \frac{\mathrm{Var}(|(N-1)v_{i,t}\epsilon_{i,t} - \sum_{j\neq i} v_{j,t}\epsilon_{j,t}|)}{\mathrm{E}[|(N-1)v_{i,t}\epsilon_{i,t} - \sum_{j\neq i} v_{j,t}\epsilon_{j,t})|]^2} \\
&= \frac{\mathrm{E}[|(N-1)v_{i,t}\epsilon_{i,t} - \sum_{j\neq i} v_{j,t}\epsilon_{j,t}|^2]}{\mathrm{E}[|(N-1)v_{i,t}\epsilon_{i,t} - \sum_{j\neq i} v_{j,t}\epsilon_{j,t})|]^2} - 1 \qquad (2.23)
\end{aligned}$$

Now suppose the $\epsilon_{i,t}$ are uncorrelated (conditional on $\mathcal{F}_{t-1}$). Denote $\Upsilon_t^2 := (N-1)^2v_{i,t}^2 + \sum_{j\neq i} v_{j,t}^2$

$$\begin{aligned}
\mathrm{E}[|(N-1)v_{i,t}\epsilon_{i,t} - \sum_{j\neq i} v_{j,t}\epsilon_{j,t}|^2] &= \mathrm{E}[\mathrm{E}[((N-1)v_{i,t}\epsilon_{i,t} - \sum_{j\neq i} v_{j,t}\epsilon_{j,t})^2|\mathcal{F}_{t-1}]] \\
&= \mathrm{E}[(N-1)^2v_{i,t}^2 + \sum_{j\neq i} v_{j,t}^2] = \mathrm{E}[\Upsilon_t^2] \\
\mathrm{E}[|(N-1)v_{i,t}\epsilon_{i,t} - \sum_{j\neq i} v_{j,t}\epsilon_{j,t})|] &= \mathrm{E}\left[\Upsilon_t\mathrm{E}\left[\left|\frac{(N-1)v_{i,t}\epsilon_{i,t} - \sum_{j\neq i} v_{j,t}\epsilon_{j,t}}{\Upsilon_t}\right||\mathcal{F}_{t-1}\right]\right] \qquad (2.24)
\end{aligned}$$

Now if we have that, conditional on $\mathcal{F}_{t-1}$, $\frac{\sum_i a_i\epsilon_{i,t}}{\sqrt{\sum_i a_i^2}}$ follows the same distribution for any (a_i) and that this conditional distribution is the same for a subset of the sample space with probability 1, then we have that

$$\kappa := \mathrm{E}\left[\left|\frac{(N-1)v_{i,t}\epsilon_{i,t} - \sum_{j\neq i} v_{j,t}\epsilon_{j,t}}{\Upsilon}\right||\mathcal{F}_{t-1}\right]$$

is a constant. This condition holds, for instance, if the $\epsilon_{i,t}$ are iid standard normal distributions.

Then we have

$$\Lambda = \frac{\mathrm{E}[\Upsilon_t^2]}{\mathrm{E}[\Upsilon_t]^2}\frac{1}{\kappa^2} - 1 \tag{2.25}$$

which is minimised if Υ_t is deterministic, for instance if the $v_{i,t}$ are all deterministic, which corresponds to volatility weighting each asset.

Note, however, that the above analysis does not take into account the effect that volatility weighting may have on the predictive ability of the strategy, i.e. on the probabilities p and q. It is far from obvious that they should be unaffected.[9] We can formulate the following proposition:

Proposition 2.5. *Suppose a market of N assets has returns of the form $r_{i,t} = \sigma_{i,t}\epsilon_{i,t}$ with $\boldsymbol{\sigma}_t = (\sigma_{i,t})_{i=1,..N}$ being a strictly positive predictable process and the $\epsilon_{i,t}$ uncorrelated, having zero mean and unit variance, conditional on $\mathcal{F}_{t-1}$. Suppose further that, conditional on $\mathcal{F}_{t-1}$, the distribution of*

$$\frac{\sum_i a_i \epsilon_{i,t}}{\sqrt{\sum_i a_i^2}}$$

does not depend on the sequence $(a)_{i=1,...,N}$ and is a.s. the same (this holds for instance if the $(\epsilon_{i,t})_i$ are iid Normal).

Now consider some set of predictable weightings $w_{i,t} = \frac{v_{i,t}}{\sigma_{i,t}}$ of the assets which include weightings with the $v_{i,t}$ deterministic and a signed cross-sectional strategy that invests in $\tilde{r}^*_{i,t} = w_{i,t}r_{i,t} - \frac{1}{N}\sum_{i=1}^N w_{i,t}r_{i,t}$. Suppose that the success and failure ratio does not depend on the weightings. In this case volatility weighting, corresponding to choosing $v_{i,t}$ deterministic is optimal (in that it maximises the Sharpe ratio of the strategy).

The above proposition is rather strict and I do not expect interesting examples that satisfy it to be plentiful.

2.4.3 Diversifying over strategies

The previous results are for a single series of returns or for strategies based on a single asset (or a single asset's deviations from the average return). In practice we will diversify across such strategies. We would expect the results above to continue to hold (at least approximately) for such diversified strategies.

For instance suppose we have N uncorrelated strategies or assets with expected returns $\mu_{i,t}$, volatilities $\sigma_{i,t}$, and volatilities after applying a volatility weighting scheme $\varsigma_{i,t}$. We can suppose that the returns after and before volatility weighting are equal (by adjusting the size of the investment).[10] Then $\varsigma_{i,t} \leq \sigma_{i,t}$ if volatility weighting is effective. The Sharpe ratios of strategies investing equally in each of the assets or strategies are then

$$\frac{\sum_i \mu_{i,t}}{\sqrt{\sum_i \varsigma_{i,t}^2}} \geq \frac{\sum_i \mu_{i,t}}{\sqrt{\sum_i \sigma_{i,t}^2}}. \tag{2.26}$$

This applies to for instance a signed time-series strategy diversified across all the available assets. It could also apply to an investment in the underlying assets themselves. In chapter 5 we will consider investing in the equal-weighted market. If it is effective to apply volatility weighting to individual assets then it should also be better to invest equally in volatility-weighted asset returns.

We can say more about diversification across signed time-series strategies. Hallerbach (2011) shows that for normal and t-distributions with at least 4 degrees of freedom (with mean 0) and uncorrelated timing signals, it is optimal to weight the underlyings by their volatility (and use the same volatility target) for diversified timing strategies.

Under similar assumptions to those I have used until now I can formulate a slightly more general result, but using the exact same principle. Suppose we have asset returns $r_{i,t} = \sigma_{i,t}\epsilon_{i,t}$ and that we run timing strategies on each of these assets and diversify across timing strategies. We will assume that success and failure ratios are the same across all strategies, i.e. $p_i = p$, $q_i = q$ and the $\epsilon_{i,t}$ has the same distribution for all assets (more generally we could also assume that $\mathrm{E}[|\epsilon_{i,t}|]$ is the same across all assets). Denote $\kappa := \mathrm{E}[|\epsilon_{i,t}|]$. If the timing strategies are uncorrelated the Sharpe ratio of the diversified strategy is

[9] We see later that this assumption is empirically justifiable.

[10] We could also equalise the volatilities. This equalises the risk-contributions of the strategies (risk-parity) and would be appealing from a risk-budgeting perspective.

$$
\begin{aligned}
\mathrm{Sharpe}(r_t^T) &= \frac{(p-q)\sum_i \mathrm{E}[|\sigma_{i,t}\epsilon_{i,t}|]}{\sqrt{(p+q)\sum_i \mathrm{Var}(|\sigma_{i,t}\epsilon_{i,t}|) + (p+q-(p-q)^2)\sum_i \mathrm{E}[|\sigma_{i,t}\epsilon_{i,t}|]^2}} \\
&= \frac{(p-q)\kappa\sum_i \mathrm{E}[\sigma_{i,t}]}{\sqrt{(p+q)\sum_i \mathrm{E}[\sigma_{i,t}^2] - (p-q)^2\kappa^2\sum_i \mathrm{E}[\sigma_{i,t}]^2}} \\
&= \frac{(p-q)\kappa\Lambda_1}{\sqrt{(p+q)\Lambda_2 - (p-q)^2\kappa^2}} \qquad (2.27)
\end{aligned}
$$

with (and using Jensen's inequality)

$$
\Lambda_1 = \frac{\sum_i \mathrm{E}[\sigma_{i,t}]}{\sqrt{\sum_i \mathrm{E}[\sigma_{i,t}]^2}} \leq \sqrt{N} \qquad (2.28)
$$

$$
\Lambda_2 = \frac{\sum_i \mathrm{E}[\sigma_{i,t}^2]}{\sum_i \mathrm{E}[\sigma_{i,t}]^2} \geq 1 \qquad (2.29)
$$

Λ_1 attains its maximum if the volatilities are equal, that is if the same volatility target is used for all the timing strategies. Λ_2 attains its minimum if the volatilities are constant, that is if volatility weighting is applied. We have the following proposition:

Proposition 2.6. *Suppose a market of N assets has returns of the form $r_{i,t} = \sigma_{i,t}\epsilon_{i,t}$ with $\boldsymbol{\sigma}_t = (\sigma_{i,t})_{i=1,..N}$ being a strictly positive predictable process and the $\epsilon_{i,t}$ each have zero mean and unit variance, conditional on $\mathcal{F}_{t-1}$. Suppose further that (for every t) their absolute values $|\epsilon_{i,t}|$ are uncorrelated across assets and have the same expectation and that the $S_{i,t} = \mathrm{sign}(r_{i,t-1}r_{i,t})$ are also uncorrelated across assets. We suppose further that $(S_{i,t})_i$ is independent of $\boldsymbol{r}_t = (r_{i,t})_i$ for every t.*

Also suppose one of the following conditions is satisfied:

- *for each t $\boldsymbol{\epsilon}_t = (\epsilon_{i,t})_i$ is independent of $\boldsymbol{\sigma}_t$*
- *for each t $(E[|\epsilon_{i,t}||\mathcal{F}_{t-1}])_i$ is independent of $\boldsymbol{\sigma}_t$*
- *for each i, t $E[|\epsilon_{i,t}||\mathcal{F}_{t-1}] = k_t$ for some deterministic k_t.*

Consider signed time-series strategies on the assets in this market and suppose the success and failure probabilities p and q are the same for all the strategies. Consider further signed time-series strategies $r_{i,t}^T$ on weighted assets $r_{i,t}^* = w_{i,t}r_{i,t}$ with $w_{i,t} = \frac{v_{i,t}}{\sigma_{i,t}}$ and $\boldsymbol{v_t} = (v_{i,t})_i$ predictable. Consider only weightings that do not affect the success and failure probabilities and with $\boldsymbol{v}_t$ that satisfy (when relevant) the chosen condition above in place of $\boldsymbol{\sigma}_t$. Under the above assumptions the strategies are uncorrelated for any given weighting scheme.

A diversified strategy with returns $r_t^T = \frac{1}{N}\sum_i r_{i,t}^T$ obtains its maximal Sharpe ratio when the $v_{i,t}$ are deterministic (volatility weighting) and $v_{i,t} = v_t$ for all i (equal volatility targets).

2.4.4 Effect of volatility weighting on skew and kurtosis

Here I consider the effect of volatility weighting on the skew and kurtosis of a distribution.

First consider returns of the form $r_t = \sigma_t\epsilon_t$ with ϵ_t having zero mean and unit variance given $\mathcal{F}_{t-1}$. Further suppose that $\mathrm{E}[\epsilon_t^3|\mathcal{F}_{t-1}]$ and $\mathrm{E}[\epsilon_t^4|\mathcal{F}_{t-1}]$ are independent of $\mathcal{F}_{t-1}$.

The skewness and kurtosis are

$$
\mathrm{Skew}(r_t) = \frac{\mathrm{E}[\sigma_t^3]}{\mathrm{E}[\sigma_t^2]^{\frac{3}{2}}}\mathrm{E}[\epsilon_t^3] \qquad (2.30)
$$

$$
\mathrm{Kurt}(r_t) = \frac{\mathrm{E}[\sigma_t^4]}{\mathrm{E}[\sigma_t^2]^2}\mathrm{E}[\epsilon_t^4] \qquad (2.31)
$$

As $\frac{\mathrm{E}[\sigma_t^3]}{\mathrm{E}[\sigma_t^2]^{\frac{3}{2}}}$ and $\frac{\mathrm{E}[\sigma_t^4]}{\mathrm{E}[\sigma_t^2]^2}$ are greater than 1 volatility weighting would reduce both kurtosis and skewness. Note however, that both positive and negative skewness would be reduced.[11]

[11] It would not explain volatility weighting resulting in positively skewed returns becoming more positively skewed for instance.

The assumptions above are rather restrictive, but it does not appear that they can easily be relaxed. The results are applicable to weighting strategies with their own volatility as in section 2.4.1 and to the distribution of individual asset returns when weighted. This says nothing about what happens when diversifying across strategies or assets – skew and kurtosis do not necessarily behave nicely even when assets are uncorrelated. Nevertheless, we may expect that a momentum strategy run on assets with a lower skew or kurtosis may also display a lower skew or kurtosis much of the time.

2.5 Dispersion

In this section I consider the link between momentum strategies and dispersion in a theoretical manner. This is of most interest (and theoretically attractive) for cross-sectional strategies. However, I also include similar results for time-series strategies.

2.5.1 Links with dispersion and sums of returns

I attempt to link the returns of momentum strategies with returns in the previous period via dispersion (for cross-sectional strategies) or a sum of returns (for time-series strategies). I assume returns or deviations (or their absolute values) are AR(1) processes adapted to some filtration to achieve this. Without this or an assumption with a similar effect it is not clear that one can create a predictive (as opposed to contemporaneous) link between momentum profits and these quantities. These models predict a positive relationship between cross-sectional momentum and dispersion, which is not found empirically. A link with volatility may be the missing element (dispersion is related to volatility).

Signed strategies

I consider the diversified signed strategies defined in 2.3.2 and look at a model for each. The time-series model neglects lead-lag effects and the cross-sectional model looks at a kind of lead-lag effect between each stock and the market average.[12]

Time-series
Suppose that absolute returns follow an AR(1) process with $|r_{i,t}| = |c|_i + \phi_i|r_{i,t-1}| + \epsilon_t$ where errors are white noise independent of the returns (vertical bars on $|c|$ are just notational). This is a form volatility clustering and it has been shown that the correlation between absolute returns is strong, at least in stock markets.[13] I also need that $S_{i,t}$ is independent of $r_{i,t-1}$. The time-series strategy returns can be written as

$$r_t^T = \frac{1}{N}\sum_{i=1}^{N} S_{i,t}(\phi_i|r_{i,t-1}| + |c|_i + \epsilon_t).$$

Then the expectation conditional on time $t-1$ is

$$\begin{aligned}\mathrm{E}[r_t^T|\mathcal{F}_{t-1}] &= \frac{1}{N}\sum_{i=1}^{N}[\phi_i(p_i - q_i)|r_{i,t-1}| + |c|_i(p_i - q_i)] \\ &= \phi(p-q)\frac{1}{N}\sum_{i=1}^{N}|r_{i,t-1}| + (p-q)\frac{1}{N}\sum_{i=1}^{N}|c|_i. \qquad (2.32)\end{aligned}$$

The latter equality follows if $\phi_i = \phi,\ p_i = p,\ q_i = q$ for all the assets. In this case it follows that the return of the strategy is related to the sum of absolute returns.

[12] I do not claim that these two models are consistent (we are not attempting to find a unified theory of momentum). In fact they are not.
[13] See for instance Mandelbrot and Hudson (2004).

Cross-sectional
We can do the same as above, but now assuming that the $|\tilde{r}_{i,t}|$ follow an AR(1) process. Then

$$\begin{aligned}
\mathrm{E}[r_t^X|\mathcal{F}_{t-1}] &= \frac{1}{N}\sum_{i=1}^{N}[\phi_i(p_i-q_i)|\tilde{r}_{i,t-1}| + |\tilde{c}|_i(p_i-q_i)] \\
&= \phi(p-q)\frac{1}{N}\sum_{i=1}^{N}|\tilde{r}_{i,t-1}| + (p-q)\frac{1}{N}\sum_{i=1}^{N}|\tilde{c}|_i
\end{aligned} \tag{2.33}$$

Where I again assume that the success and failure probabilities are equal for all the individual strategies and the AR(1) parameter is the same. In this case the momentum return has a dependence on the sum of the absolute deviations (a measure of dispersion) from the (equal-weighted) market.

Linear strategies

Time-series
Now suppose that returns are AR(1): $r_{i,t} = c_i + \phi r_{i,t-1} + \epsilon_t$. Then the return on a linear time-series strategy is

$$\begin{aligned}
r_t^T &= \frac{1}{N}\sum_{i=1}^{N} r_{i,t-1}r_{i,t} \\
&= \frac{1}{N}\sum_{i=1}^{N}\left[\phi(r_{i,t-1})^2 + c_i r_{i,t-1} + \epsilon_{i,t}r_{i,t-1}\right] \\
\mathrm{E}[r_t^T|\mathcal{F}_{t-1}] &= \phi\frac{1}{N}\sum_{i=1}^{N}(r_{i,t-1})^2 + \frac{1}{N}\sum_{i=1}^{N} c_i r_{i,t-1}.
\end{aligned}$$

Under the assumption that absolute returns are AR(1) the above becomes

$$\begin{aligned}
r_t^T &= \frac{1}{N}\sum_{i=1}^{N}\left[S_{i,t}\phi(r_{i,t-1})^2 + S_{i,t}c_i|r_{i,t-1}| + S_i\epsilon_{i,t}|r_{i,t-1}|\right] \\
\mathrm{E}[r_t^T|\mathcal{F}_{t-1}] &= \phi(p-q)\frac{1}{N}\sum_{i=1}^{N}(r_{i,t-1})^2 + (p-q)\frac{1}{N}\sum_{i=1}^{N} c_i|r_{i,t-1}|.
\end{aligned} \tag{2.34}$$

Here a dependence on squares of returns[14] is evident (variability in the returns will be amplified by the squares).

Cross-sectional
By replacing $r_{i,t}$ with $\tilde{r}_{i,t} := r_{i,t} - \bar{r}_t$ and noting that

$$\sum_i (r_{i,t-1} - \bar{r}_{t-1})r_{i,t} = \sum_i (r_{i,t-1} - \bar{r}_{t-1})(r_{i,t} - \bar{r}_t)$$

we get analogous results to those above under analogous assumptions. Now we see dependence on squared deviations from the (equal-weighted) average of the form (2.35) below.

I draw some Inspiration from Yu and Sharaiha (2007) and use a (simple) result they derive below. They show that $\sum_{i=1}^{N}\tilde{r}_{i,t-1}r_{i,t-1} = \sum_{i=1}^{N}\tilde{r}_{i,t-1}^2$.

Under the assumption that returns follow an AR(1) process we get

$$\begin{aligned}
r_t^X &= \frac{1}{N}\sum_i\left[\phi\tilde{r}_{i,t-1}r_{i,t-1} + c_i\tilde{r}_{i,t-1} + \epsilon_{i,t}\tilde{r}_{i,t-1}\right] \\
&= \phi\frac{1}{N}\sum_{i=1}^{N}\tilde{r}_{i,t-1}^2 + \frac{1}{N}\sum_{i=1}^{N}c_i\tilde{r}_{i,t-1} + \frac{1}{N}\sum_{i=1}^{N}\tilde{r}_{i,t-1}\epsilon_i \\
\mathrm{E}[r_t^X|\mathcal{F}_{t-1}] &= \phi\frac{1}{N}\sum_{i=1}^{N}\tilde{r}_{i,t-1}^2 + \frac{1}{N}\sum_{i=1}^{N}c_i\tilde{r}_{i,t-1}.
\end{aligned} \tag{2.35}$$

[14] $(r_{i,t-1})^2$ can be thought of as the instantaneous variance of the return process. It is, for instance, an EWMA estimate with only one observation.

This is of exactly the same form as assuming that deviations $\tilde{r}_t$ follow an AR(1) process.[15] Thus we see dependence on square deviations even with no lead-lag effects between stocks. The average square deviation seen here is a measure of dispersion (if one takes the square root one gets the cross-sectional standard deviation, which is the usual estimate of dispersion).

2.5.2 Dispersion as volatility

We can model dispersion as a volatility around a mean return which gives some insights into its role. This is so particularly for the cross-sectional strategies. With time-series strategies this perspective is less useful. In both cases it does, however, appear that a more stable dispersion is better.

Following Solnik and Roulet (2000) we can model cross-sectional dispersion as a predictable process σ_t such $r_{i,t} = f_t + \sigma_t \epsilon_{i,t}$ where we take for f_t the (equal-weighted) market. Then $\tilde{r}_t = \sigma_t \epsilon_{i,t}$. A signed timing strategy or a linear momentum strategy on these deviations can then be improved by weighting by the cross-sectional dispersion. This follows from the results for volatility weighting of signed and linear time-series strategies in 2.4.2 but applied to deviations (we need to make the corresponding assumptions for the deviations). The mean and variance for a linear strategy in this case are (one can calculate these for a signed strategy too):

$$\begin{aligned}
\mathrm{E}[r_t^X] &= \mathrm{E}[\sigma_t \sigma_{t-1}]\mathrm{E}[\epsilon_{i,t}\epsilon_{i,t-1}] \\
\mathrm{Var}[r_t^X] &= \mathrm{E}[\sigma_t^2 \sigma_{t-1}^2]\mathrm{E}[\epsilon_{i,t}^2 \epsilon_{i,t-1}^2] - \mathrm{E}[\sigma_t \sigma_{t-1}]^2 \mathrm{E}[\epsilon_{i,t}\epsilon_{i,t-1}]^2.
\end{aligned}$$

This is, of course, not a very realistic model. It assumes a beta of 1 with the market for all assets. In practice this will not be the case. The model also does not predict a negative relationship with dispersion (in fact the relationship is still positive). For instance $\sigma_t \sigma_{t-1} \mathrm{E}[\epsilon_{i,t}\epsilon_{i,t-1} | \mathcal{F}_{t-1}]$ is increasing in dispersion (provided it is positive). What this model shows is that despite cross-sectional momentum depending on dispersion for its profit, stabilising the dispersion can result in an improved risk profile for the strategy.

Using the same model we can calculate the mean and variance of a time-series strategy. To simplify things we additionally assume that the factor has mean zero, that f and ϵ are independent and that their distributions are Normal or, more generally, elliptical (this last assumption is so that for instance $\mathrm{E}[\epsilon_t \epsilon_{t-1}^2] = 0$ – we could assume this directly of course).

We now have

$$\begin{aligned}
\mathrm{E}[r_t^T] &= \mathrm{E}[f_{t-1} f_t] + \mathrm{E}[\sigma_t \sigma_{t-1}]\mathrm{E}[\epsilon_{i,t}\epsilon_{i,t-1}] \\
\mathrm{Var}[r_t^T] &= \mathrm{Var}(f_{t-1} f_t) + 2\mathrm{E}[f_t f_{t-1}]\mathrm{E}[\sigma_t \sigma_{t-1}]\mathrm{E}[\epsilon_{i,t}\epsilon_{i,t-1}] + \mathrm{E}[\sigma_t^2 \sigma_{t-1}^2]\mathrm{E}[\epsilon_{i,t}^2 \epsilon_{i,t-1}^2] - \mathrm{E}[\sigma_t \sigma_{t-1}]^2 \mathrm{E}[\epsilon_{i,t}\epsilon_{i,t-1}]^2 \\
&= \mathrm{Var}(f_{t-1} f_t) + 2\mathrm{E}[f_t f_{t-1}]\mathrm{E}[\sigma_t \sigma_{t-1}]\mathrm{E}[\epsilon_{i,t}\epsilon_{i,t-1}] + \mathrm{Var}(\sigma_t \sigma_{t-1} \epsilon_{i,t}\epsilon_{i,t-1}).
\end{aligned}$$

Note that this decomposes the variance into a systematic portion, an idiosyncratic portion and a portion due to auto-covariance of the of the factor and residuals. Now the effect of dispersion on the Sharpe ratio is not so clear. Suppose momentum is entirely due to persistence in the factor, so $\mathrm{E}[\epsilon_t \epsilon_{t-1}] = 0$ and $\mathrm{E}[f_t f_{t-1}] > 0$.

The Sharpe ratio is then

$$\mathrm{Sharpe}(r_t^T) = \frac{\mathrm{E}[f_t f_{t-1}]}{\sqrt{\mathrm{Var}[f_t f_{t-1}] + \mathrm{E}[\sigma_t^2 \sigma_{t-1}^2]\mathrm{E}[\epsilon_{t-1}^2 \epsilon_t^2]}}. \tag{2.36}$$

This decreases if $\mathrm{E}[\sigma_{t-1}^2 \sigma_t^2]$ increases, for instance if volatility is increased by some factor $c > 1$ or if the variance of $\sigma_{t-1}^2 \sigma_t^2$ increases without affecting the mean (which would increase the coefficient of variation).

Now suppose instead that all of the momentum is idiosyncratic, so that $\mathrm{E}[\epsilon_t \epsilon_{t-1}] > 0$ and $\mathrm{E}[f_t f_{t-1}] = 0$, then the Sharpe ratio is

$$\mathrm{Sharpe}(r_t^T) = \frac{\mathrm{E}[\epsilon_{t-1}\epsilon_t]}{\sqrt{\frac{\mathrm{E}[f_t^2 f_{t-1}^2]}{\mathrm{E}[\sigma_t \sigma_{t-1}]^2} + \frac{\mathrm{E}[\sigma_t^2 \sigma_{t-1}^2]}{\mathrm{E}[\sigma_t \sigma_{t-1}]^2}\mathrm{E}[\epsilon_{t-1}^2 \epsilon_t^2]}}. \tag{2.37}$$

This increases if you multiply dispersion by some constant $c > 1$ but it is decreasing in $\mathrm{E}[\sigma_{t-1}^2 \sigma_t^2]$ as before so that a more variable dispersion is not beneficial.

[15]Thus I only report the equation once.

2.5.3 Link between dispersion and volatility

There is a positive link between dispersion and market volatility and negative link with market correlation.

Yu and Sharaiha (2007) derive the following result. Denoting $\chi_t^2 = \frac{1}{N}\sum_i (r_{i,t} - \bar{r}_{i,t})^2$ we have

$$\mathrm{E}[\chi_t^2|\mathcal{F}_{t-1}] = \frac{1}{N}\sum_i \sigma_{i,t}^2 - \sigma_{mt}^2 + \frac{1}{N}\sum_i (\mu_{i,t} - \mu_{mt})^2 \tag{2.38}$$

$$= \frac{N-1}{N}(\overline{\sigma^2}_t - \overline{\mathrm{cov}}_t) + \frac{1}{N}\sum_i (\mu_{i,t} - \mu_{mt})^2 \tag{2.39}$$

where $\sigma_{i,t}^2,\ \mu_{i,t}, \sigma_{mt}^2,\ \mu_{mt}$ are conditional variances and means of individual assets and the (equal-weighted) market and $\overline{\sigma_t^2} = \frac{1}{N}\sum_{i=1}^N \sigma_{i.t}^2,\ \overline{\mathrm{cov}}_t = \frac{1}{(N-1)N}\sum_{i\neq j}\sigma_{ij}$ are average variances and covariances.

This links dispersion positively to the average market volatility and negatively to average covariance. The positive link will be lessened if covariances also increase in times of high volatility (which is probably the case empirically). Also of note is that dispersion is related to the difference between average market variance and the variance of the market (Yu and Sharaiha, 2007). A portion of dispersion is also due to cross-sectional dispersion in (conditional) mean returns. This corresponds to the last term in the decomposition (2.7) for linear cross-sectional strategies, the (unconditional) cross-sectional dispersion in mean returns. In section 6.4 I examine these two portions of dispersion.

If we normalise all the assets then the conditional volatilities all become 1. If we suppose assets are uncorrelated we see that the dependence of dispersion on volatility (of normalised returns) is removed. We may expect the dispersion to be more stable then too, which as we have seen is likely to be good for the risk-profile of momentum. Note also that with normalised returns the conditional means are now normalised, which means they are negatively related to ex-ante volatility. Thus we could expect a negative relationship of dispersion (in normalised returns) with market volatility (in unweighted returns).

2.6 Volatility, prediction accuracy and profits

I attempt to model a link between volatility and profits or predictive accuracy of momentum strategies. I propose these models merely as a first step toward explaining the results in the prediction analysis in section 4.3 and to give some idea for why previous empirical studies have found a negative link between momentum and volatility. I do not test these models formally.

2.6.1 Predictive accuracy

Consider an AR(1) model for stock returns (or deviations) $r_{i,t} = \phi_i r_{i,t-1} + \sigma_{i,t}\epsilon_{i,t}$ with $\sigma_{i,t}$ a positive predictable process and deviations independent of the past. For the timing strategies we are interested in the sign of

$$r_{i,t}r_{i,t-1} = \phi_i r_{i,t-1}^2 + \sigma_{i,t}\epsilon_{i,t}r_{i,t-1}. \tag{2.40}$$

Consider the above conditional on $\mathcal{F}_{t-1}$. The probability that the sign is positive is lower (and if $\epsilon_{i,t}$ is symmetric always at least 50 %) if volatility is higher. We may, however, expect a positive relationship between σ_t and $r_{i,t-1}^2$ so that in fact there is a counteracting tendency for the predictive accuracy to be high when volatility is high, which may dominate.

Now consider the case where the mean is positive: $r_{i,t} = c + \phi_i r_{i,t-1} + \sigma_{i,t}\epsilon_{i,t}$. Now

$$r_{i,t}r_{i,t-1} = \phi_i r_{i,t-1}^2 + c r_{i,t-1} + \sigma_{i,t}\epsilon_{i,t}r_{i,t-1}. \tag{2.41}$$

The previous conclusion holds as long as $\phi_i r_{i,t-1}^2 + c r_{i,t-1} > 0$.

2.6.2 Profits

I posit a weakening link between past and future returns in high volatility states as follows: $r_{i,t} = \frac{\phi_i}{\sigma_{i,t}} r_{i,t-1} + c + \sigma_{i,t}\epsilon_{i,t}$. This retains the features of predictive accuracy in the previous section (in fact it is more pronounced).

Consider a linear (time-series) momentum strategy on these returns. We get

$$\begin{aligned} r_t^T &= \frac{\phi_i}{\sigma_{i,t}} r_{i,t-1}^2 + c_i r_{i,t-1} + \epsilon_{i,t} r_{i,t-1} \sigma_{i,t} \\ \mathrm{E}[r_t^T | \mathcal{F}_{t-1}] &= \frac{\phi_i}{\sigma_{i,t}} r_{i,t-1}^2 + c_i r_{i,t-1} \end{aligned} \tag{2.42}$$

Thus this model predicts lower profit given higher volatility. Results can be carried over to the cross-sectional strategies by considering deviations from the mean.

2.7 Global vs local time-series momentum

One can consider running a time-series strategy on the market return (the equal-weighted market for instance) rather than on individual assets. Reasons to do so may include there being less noise in the market trend (Duyvesteyn and Martens, 2013). Duyvesteyn and Martens (2013) call this a global momentum strategy, as opposed to a local strategy and find this seems to perform better for government bonds, where cross-sectional momentum is weak. The reason for this can be seen below and relates to the sources of profit for time-series and cross-sectional momentum. Given a market of N assets, running a linear time-series strategy on the market average gives profits that can be decomposed as follows:

$$\mathrm{E}[r_t^{GT}] = \mathrm{E}[\bar{r}_{t-1}\bar{r}_t] = \frac{\mathrm{E}[r_t^T]}{N} + \frac{1}{N^2}(\mathbf{1}'\Omega_{t-1,t}\mathbf{1} - \mathrm{tr}(\Omega_{t-1,t})) + \frac{1}{N^2}(\mathbf{1}'\boldsymbol{\mu}_{t-1}\boldsymbol{\mu}_t'\mathbf{1} - \mathrm{tr}(\boldsymbol{\mu}_{t-1}\boldsymbol{\mu}_t')). \tag{2.43}$$

Now the first term in the expression is the returns from a linear time-series strategy, but divided by N. The second term is the cross-serial covariances term in the linear cross-sectional strategy, but with the opposite sign. Here it benefits if assets have positive correlations rather than negative.

For the last term first note that

$$\sigma^2_{\boldsymbol{\mu_{t-1}},\boldsymbol{\mu_t}} = \frac{N-1}{N^2}\mathrm{tr}(\boldsymbol{\mu}_{t-1}\boldsymbol{\mu}') - \frac{1}{N^2}(\mathbf{1}'\boldsymbol{\mu}_{t-1}\boldsymbol{\mu}_t'\mathbf{1} - \mathrm{tr}(\boldsymbol{\mu}_{t-1}\boldsymbol{\mu}_t')). \tag{2.44}$$

The second term above occurs, again with the opposite sign, in (2.43). Thus the global trend downweights the typical time-series strategy and instead includes a short position in two aspects of cross-sectional momentum: the cross-serial covariances and a portion of the mean term (the covariance of means), consisting of the off-diagonal elements of the matrix $\boldsymbol{\mu}_{t-1}\boldsymbol{\mu}_t'$. Intuitively, if assets tend to follow each other, i.e. cross-serial covariances are large (probably when correlations are high), this strategy will do better than a simple time-series strategy. (Also the last term in (2.43) is positive if means are generally positive, which is probably the case).

The decomposition here could also potentially explain why no short-term reversal is found in industry and asset-allocation momentum. A time-series strategy on an industry or market index is essentially a global time-series strategy on the underlying industries or stocks making up the market. Even if there are strong negative auto-covariances causing a short-term reversal in underlying assets, these will be downweighted compared to cross-serial correlations. If the latter are positive they could dominate and eliminate the short-term reversal.

2.8 Conclusions

- By considering initially a stylised market with only two assets we see that time-series and cross-sectional momentum differ where a market is mostly going up or mostly going down (in the formation period) and that they are similar when some (about half) of the assets go up and the rest go down.
- Time-series momentum is likely to outperform in the presence of strong *positive* auto-covariances or a large mean component of returns. Cross-sectional momentum may outperform if there are strong *negative* lead-lag relationships between assets.
- One may expect the efficacy of time-series (cross-sectional) momentum to depend on the ability to predict future returns (deviations) from past returns (deviations). For signed strategies it is particularly the ability to predict their sign that is important.
- There are theoretical grounds for the efficacy of volatility weighting when either

  - weighting a strategy by its own volatility, or
  - weighting the underlying assets by their volatilities

  under some idealised assumptions.

- Volatility weighting requires any part of returns not dependent on volatility to not be too large and is more likely to be effective when there is a negative relationship with volatility.
- Volatility weighting is more effective when volatility is highly variable. That is it works by reducing the volatility of volatility. Thus it may work even in the absence of a negative relationship with volatility (a volatility timing effect).
- Volatility weighting may in fact induce a negative relationship with volatility (of unweighted returns) rather than removing it.
- When diversifying across signed strategies there is a contribution from volatility weighting in terms of equalising volatility across underlyings and a contribution from stabilising volatility within each underlying.
- Naïve AR(1) type models of returns and deviations prediction a positive relationship of time-series (cross-sectional) strategies with absolute and squared returns (absolute and squared deviations). The absolute and squared deviations are measures of cross-sectional dispersion.
- Dispersion can be seen modelled as a volatility and there is some reason to believe dispersion-weighting may be effective for cross-sectional strategies. Reducing variability of dispersion (or exposure to it) can be good for both cross-sectional and time-series strategies.
- There is some reason to believe that higher volatility negatively impacts prediction accuracy of signed strategies and an explicit weakening of the link between successive returns in higher volatility states gives momentum profits that depend negatively on volatility.
- A global trend strategy invests negatively in that part of cross-sectional momentum that depends on cross-serial covariances and will benefit when this part of the cross-sectional strategy hinders cross-sectional momentum profits.

In the remaining chapters many of these conclusions will be considered empirically.

Chapter 3

Evidence from six momentum strategies

I consider two questions. Firstly, the question of whether there are in fact momentum profits in the data under consideration and secondly where these profits are the strongest.

As momentum strategies have been defined in different ways in previous literature, including some specifications (such as linear strategies) useful mainly for theoretical analysis I examine various specifications. I consider here six types of momentum strategies, which allows some initial remarks about the difference between these forms of strategies to be made. The six strategies are

- quantile cross-sectional (qxs)
- (unscaled) linear cross-sectional (ulxs)
- scaled linear cross-sectional (slxs)
- signed time-series (sts)
- (unscaled) linear time-series (ults)
- scaled linear time-series (slts)

Of course we expect the different specifications to be closely related. However, differences give insight into the behaviour of momentum and we need a close relationship whenever extending theoretical conclusions based on a convenient specification of momentum to a more practically relevant specification.

I firstly examine in section 3.1 momentum strategies with a 12 month (52 week) formation period for the industry (MAA) data as this has been examined in previous literature, including in Moskowitz *et al.* (2012) and consider the pattern of returns over the following 18 months (18 weeks). In section 3.2 I then vary both the formation and holding period length and look at the returns and other measures. In the final section I use regressions to examine the linear relationship between the strategies. I do so to confirm that there is a close relationship between the strategies, to identify possible sources of diversification (particularly between times-series and cross-sectional momentum) and to confirm any outperformance of any one type of strategy over another as seen in the other sections.

3.1 Returns over the holding period

I examine the pattern of returns over an 18 month (18 week) holding period for momentum strategies with a 12 month (52 week) formation period for industry (MAA) data[1] I choose a 1 year formation period as this has been used in the literature before, for instance in Moskowitz *et al.* (2012).

I calculate returns over every possible combination of formation period and holding period. This means that the data are based on overlapping formation periods (even though for each month/week in the holding period observations are separate). I report results for strategies where the portfolios are reweighted[2] (with their original weights, i.e. whichever weights were chosen at the start of the holding

[1]I also considered month-long periods (defined as 4 weeks for the MAA data). This gave similar results.

[2]Buy and hold strategies can more easily run into trouble when the capital becomes negative and our return calculations become invalidated (see appendix C.4). Reweighting the portfolios each period is a kind of risk management strategy in any case. It ensures that the gross exposure (and thus the implied leverage) of the positions does not grow too large.

period) every month (or every week). For datasets where not all assets are available at the start, I consider at each point only the assets that are available and have sufficient history. The portfolio thus formed includes only these assets, even should another asset come available during the holding period.

Table 3.1 records the annualised Sharpe ratios of the returns for 6 different strategies for each set of data. (Note that the time period for the industry data is in months and for the MAA data in weeks.) The Sharpe ratio is based on the monthly or weekly (i.e. not annualised) returns. See appendix C.3 for details regarding how I calculated the returns.[3]

For the quantile cross-sectional strategies I use 4 quantiles for industries and 3 for MAA. The scaled linear strategies are linear strategies with bets scaled so that the gross-exposure is 1 for each leg of the cross-sectional strategy and total gross exposure is 1 for the time-series strategy. Scaling equalises gross exposure but also reduces the effect of cross-sectional dispersion (particularly for the cross-sectional strategy). It could in fact be seen as a form of volatility weighting and thus foreshadows the results of chapter 5. Formulas for the weights given to each asset for all these strategies can be found in appendix B.

I find positive momentum returns that are strongest in the first period after formation for all the strategies (which suggests that looking at larger holding periods is suboptimal[4]). This suggests that there is no short-term reversal (which agrees with what Blitz and Van Vliet (2008) and Moskowitz and Grinblatt (1999) report). The Sharpe ratio (and mean returns) also generally decrease over the holding period, also agreeing with previous findings.

Due to the different scales of the bets taken by the (unscaled) linear strategies versus the other strategies, the magnitude of the returns cannot really be compared directly and one should rather look at the Sharpe ratio (or possibly other risk-adjusted measures – we shall see a little later that even the Sharpe ratio provides a misleading comparison). It does seem, however, that the additional variability of linear strategies is not compensated for by sufficiently higher returns. One may have expected that investing relatively more in assets with stronger momentum would improve performance.[5]

For industries the profits and Sharpe ratio are much higher for the cross-sectional strategies. Scaling improves performance – which is not surprising if momentum is in fact negatively related to cross-sectional dispersion. For cross-sectional momentum the linear strategies fare worse but for the time-series strategy the scaled linear strategy performs best (looking at the first period after formation). The scaled linear strategy may perform better here because, like a cross-sectional strategy, it more explicitly plays assets against each other.

For MAA we see that time-series strategies tend to have a higher Sharpe ratio (in direct contrast to the industry data), except for the scaled linear strategies. The linear strategies appear to fare worse than the non-linear strategies and scaling also seems to improve performance.

3.2 Varying formation and holding periods

I calculate statistics as in the previous section, but for returns over the full holding period and now varying the length of the formation period (J) and holding period (K). For the industry data, these are measured in months and for the MAA data in weeks. Results are annualised again. Tables 3.2a and 3.2b report the Sharpe ratios for the industry and MAA data respectively.[6] Tables 3.3a and 3.3b report more detailed statistics for strategies with a 12 month or 1 month (52 week or 4 week) formation and 1 month (1 week) holding period for the industry (MAA) data. The last column is the average of the largest 5 drawdowns (or fewer if there are not as many drawdowns[7]) normalised by the standard deviation of returns. I also report the mean less the median as a more robust measure of skewness.

We see that shorter holding periods generally produce stronger momentum profits, reinforcing the conclusions in the previous section. However, we also see stronger momentum in a shorter formation

[3] Gupta *et al.* (2013) consider the effect of different means of calculating returns for stock momentum in several countries.

[4] Note, however, that the shorter formation periods may well have higher turnover (and thus higher transaction costs) as the signal changes more quickly. Moskowitz and Grinblatt (1999) find this erases the profitability of short-term momentum for industries.

[5] One possible reason for this not being the case is overreaction in the most extreme returns (i.e. assets being overbought or oversold).

[6] This table is somewhat like one to be found in Antonacci (2013a) but here I vary both formation and holding period and do not look at different indices, but rather different strategies. Note that the Sharpe ratios reported here are not for investable strategies (except for the shortest holding period) as the observations overlap. This introduces a dependence between the observations which should be borne in mind. The returns are over different horizons and despite being annualised are thus not strictly comparable for an investor interested in returns over one particular horizon – one cannot for instance convert weekly into equivalent annual returns just by annualising the average weekly return. Note also that the method of estimating returns used here is different from the overlapping portfolio method of Jegadeesh and Titman (1993).

[7] This happens once only for the industry 12 month formation unscaled linear time-series strategy (only 3 drawdowns).

Table 3.1: ***Sharpe ratios over holding period.*** *Annualised Sharpe ratios are calculated from monthly (weekly) returns for six different momentum strategies over the period Jul1969-Jun1994 (4Jan1979 - 5Dec2002) for the industry (MAA) data considering every possible combination of 12 month (52 week) formation and 18 month (18 week) holding period. Sharpe ratios are reported for every month (week) in the holding period. Holdings are rebalanced monthly (weekly) according to the weights assigned by each strategy. Sharpe ratios are annualised by multiplying by* $\sqrt{12}$ ($\sqrt{52}$).

(a) *Industries*

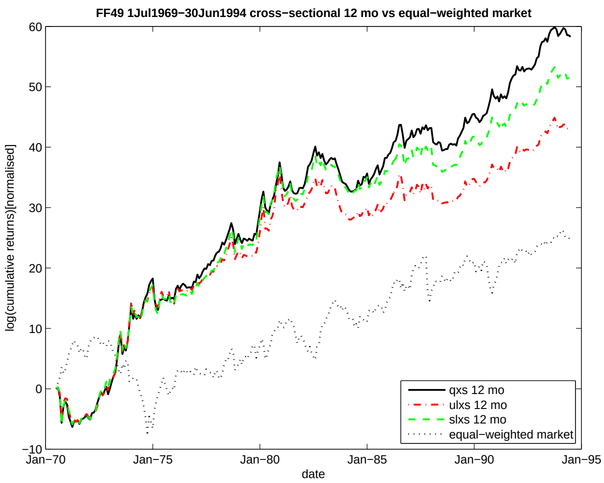
	month								
strategy	1	2	3	4	5	6	7	8	9
FF49 qxs	0.77	0.70	0.63	0.62	0.60	0.59	0.60	0.50	0.42
FF49 ulxs	0.51	0.46	0.47	0.52	0.40	0.43	0.48	0.42	0.37
FF49 slxs	0.69	0.64	0.64	0.66	0.51	0.57	0.55	0.47	0.45
FF49 sts	0.09	-0.00	0.02	0.05	0.09	0.03	0.02	0.04	0.04
FF49 ults	0.00	-0.05	-0.11	-0.01	-0.05	-0.02	0.01	-0.04	-0.01
FF49 slts	0.25	0.13	0.14	0.14	0.15	0.09	0.07	0.10	0.10
	month								
strategy	10	11	12	13	14	15	16	17	18
FF49 qxs	0.27	0.02	-0.06	-0.25	-0.38	-0.37	-0.28	-0.22	-0.38
FF49 ulxs	0.22	-0.01	-0.12	-0.23	-0.44	-0.40	-0.34	-0.28	-0.27
FF49 slxs	0.26	0.05	-0.04	-0.16	-0.35	-0.31	-0.23	-0.20	-0.23
FF49 sts	-0.10	-0.32	-0.43	-0.28	-0.31	-0.15	-0.05	-0.06	-0.13
FF49 ults	-0.10	-0.21	-0.31	-0.29	-0.35	-0.14	-0.08	-0.07	-0.09
FF49 slts	-0.06	-0.28	-0.36	-0.27	-0.31	-0.19	-0.08	-0.12	-0.13

(b) *MAA*

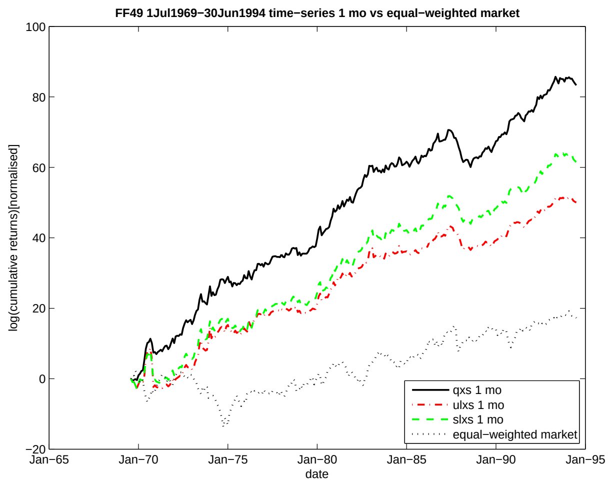
	week								
strategy	1	2	3	4	5	6	7	8	9
MAA qxs	0.91	0.79	0.75	0.60	0.56	0.55	0.52	0.54	0.60
MAA ulxs	0.40	0.41	0.26	0.25	0.24	0.23	0.21	0.19	0.24
MAA slxs	0.71	0.69	0.56	0.54	0.47	0.45	0.42	0.39	0.38
MAA sts	1.10	0.92	0.83	0.79	0.71	0.67	0.63	0.69	0.69
MAA ults	0.47	0.44	0.27	0.24	0.22	0.21	0.18	0.18	0.23
MAA slts	0.60	0.54	0.40	0.40	0.35	0.31	0.29	0.27	0.28
	week								
strategy	10	11	12	13	14	15	16	17	18
MAA qxs	0.61	0.61	0.61	0.53	0.58	0.48	0.46	0.43	0.38
MAA ulxs	0.27	0.35	0.38	0.40	0.35	0.25	0.23	0.21	0.21
MAA slxs	0.39	0.38	0.41	0.39	0.37	0.27	0.24	0.20	0.15
MAA sts	0.65	0.63	0.55	0.49	0.56	0.52	0.47	0.43	0.34
MAA ults	0.27	0.35	0.36	0.41	0.38	0.29	0.26	0.22	0.22
MAA slts	0.31	0.34	0.37	0.38	0.35	0.26	0.24	0.18	0.16

period.

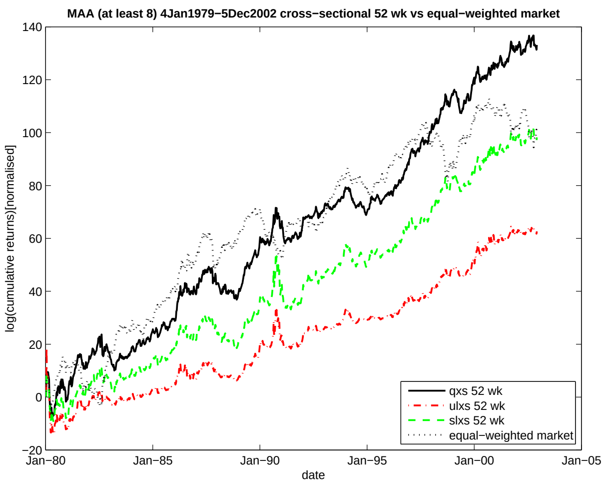
For the industry data the 1 month formation and holding period momentum strategies are the strongest.[8] We still see that the cross-sectional strategies perform better. However, the time-series momentum is far more profitable for the shorter formation period than for the longer period. There is thus something different about short and long momentum. In table 3.3a the 1 month formation appears to have lower drawdowns (and often lower kurtosis) than the 12 month formation. Table 3.3a shows mostly negative skews. Here we also see slightly higher normalised drawdowns for the unscaled (vs the scaled) linear strategies.



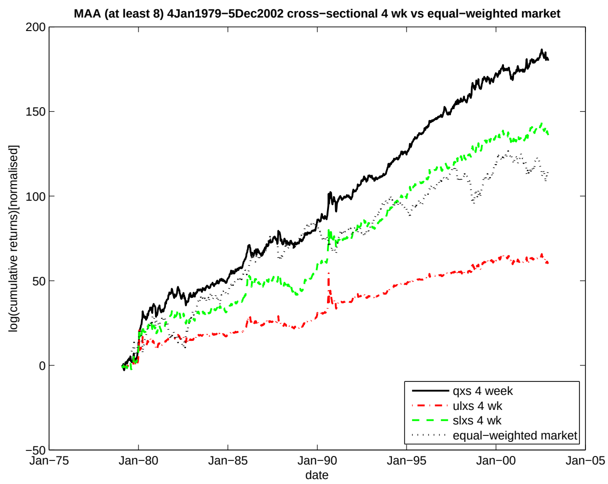
(a) *industry 12 month formation*



(b) *industry 1 month formation*



(c) *MAA 52 week formation*

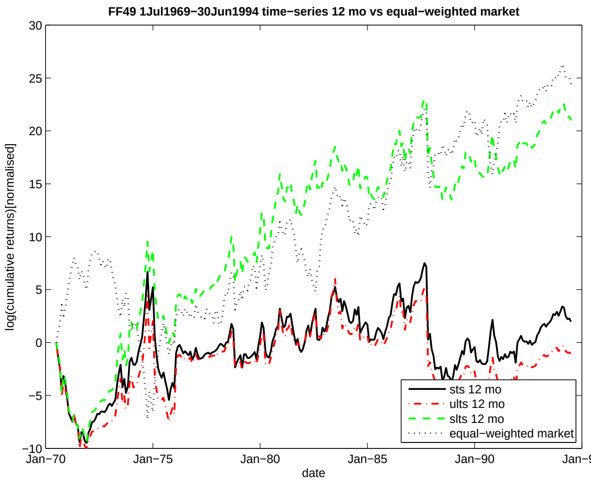


(d) *MAA 4 week formation*

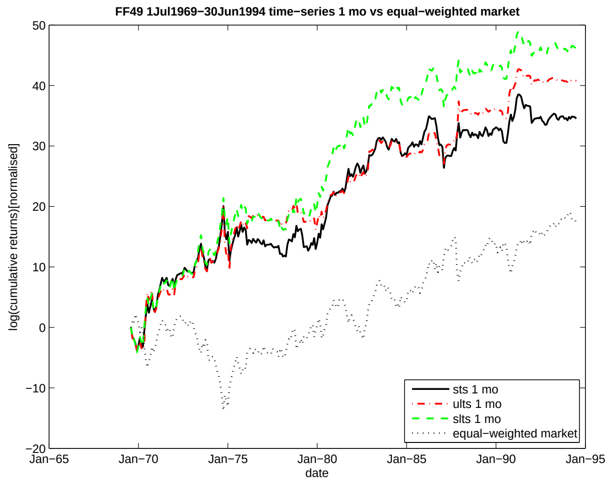
Figure 3.1: ***Cumulative performance graphs of cross-sectional momentum strategies and the market return.*** *The standardised logarithm of cumulative returns is plotted for three types of cross-sectional strategies with 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period using asset returns over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for the industry (MAA) data. The series for each strategy is standardised by dividing by the ex-post standard deviation of log returns. The same is done for a strategy that invests equally in each asset every month (week), the equal-weighted market strategy.*

For the MAA data we see that a 1 week holding period displays the highest Sharpe ratio (except for a 1 week formation period). We still see a higher Sharpe ratio for the time-series strategy and the 4 week[9] formation has the strongest momentum. Scaling still improves the Sharpe ratio. In table 3.3b the 4 week formation usually has lower drawdowns than the 52 week but the kurtosis is often higher. It

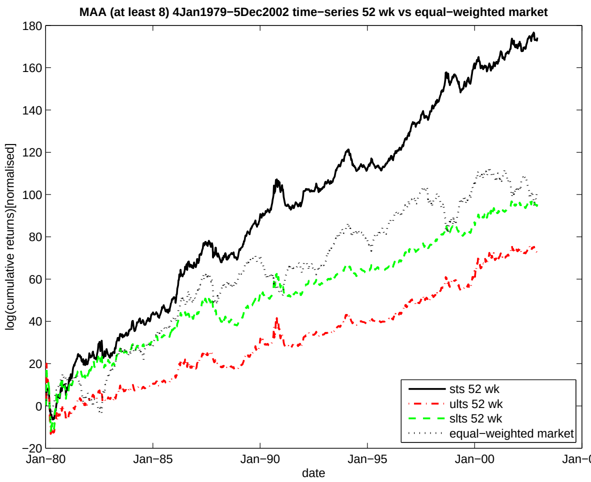
[8]12 month and 3 month formation also have strong profits for cross-sectional and time-series strategies respectively.
[9]This is followed by 8 week and 52 week formation periods.



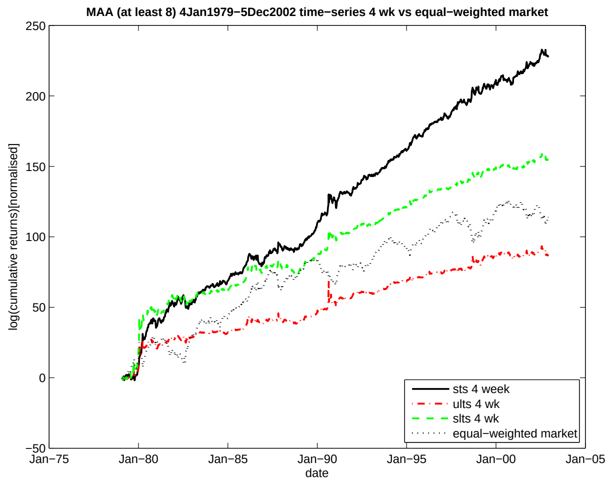
(a) *industry 12 month formation*



(b) *industry 1 month formation*



(c) *MAA 4 week formation*



(d) *MAA 52 week formation*

Figure 3.2: ***Cumulative performance graphs of time-series momentum strategies and the market return.*** *The standardised logarithm of cumulative returns is plotted for three types of time-series strategies with 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period using asset returns over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for the industry (MAA) data. The series for each strategy is standardised by dividing by the ex-post standard deviation of log returns. The same is done for a strategy that invests equally in each asset every month (week), the equal-weighted market strategy.*

now at first seems that negative skewness is not necessarily an ubiquitous characteristic of momentum strategies. Skew is often positive for this data. However, the mean less median figure is still often negative (indicating that some extreme positive returns influence the skewness coefficient). This can be seen in table 3.3b. We see very extreme kurtosis for especially the shorter formation and holding periods for the linear strategies – this is caused by just one or two extreme returns.[10]. This can be seen most clearly in table 3.3b. Such kurtosis is also seen for the scaled linear time-series strategy (though it can be avoided by choosing a period with 2 more assets in the data).

It would appear that linear strategies amplify volatility. If extreme returns tend to be followed by extreme returns (which seems to be a stylised fact of financial markets) then the linear strategy will invest most exactly when returns are most extreme or when volatility is highest (remember that volatility and dispersion are positively linked). The lesson to take from this is that Sharpe ratios are not sufficient as risk-adjusted measures and that the linear strategies, though of academic interest, are not suited for practical implementation. In later chapters I shall pay little (or no) attention to linear strategies for this reason.

In figures 3.1 and 3.2 I plot log cumulative returns of cross-sectional and time-series strategies compared with the equal-weighted market. I have standardised the plots by dividing by the (ex-post) standard deviation of log returns in order to make the graphs more comparable. Noticeable in these graphs is the gap between the scaled and unscaled strategies. For all except the industry 1 month formation time-series strategies one also sees a moderate to marked outperformance for the non-linear strategies, also over the market.

3.3 Relationship between the strategies

In this section I examine the linear relationships between the six different momentum strategies. I run some (simple 1 dimensional) regressions of the different strategy variations on each other. I also calculate correlation coefficients. Because of the presence of outliers and the non-normality of the data I report (iteratively reweighted) robust regression estimates. Again I consider formation periods of 12 months (52 weeks) with a 1 month (1 week) holding period for the industry (MAA) data.

I run three types of regressions. In each case I attempt to compare like with like. I run regressions in both directions in each case as there is no clear reason to make one return series the dependent or independent variable. The three types of regressions are

(1) Cross-sectional vs Time-series: I run quantile cross-sectional against signed time-series (and vice versa), unscaled linear cross-sectional versus unscaled linear time-series, scaled linear cross-sectional vs scaled linear time-series;

(2) Scaled vs unscaled: I regress scaled strategies on unscaled and vice versa; and

(3) Linear vs non-linear: I regress quantile cross-sectional on the linear cross-sectional strategies (and vice versa) and similarly for the time-series strategies.

The first type of regression is perhaps the most interesting. The goals of these regressions are to confirm (or not) the close relationship between the different momentum strategies and to confirm whether the higher performance observed for some strategies in the previous sections is also seen statistically. First, however, I consider briefly the correlations between the strategies.

I report selected correlation estimates[11] in table 3.4 Here I show the correlations of the 6 different types of strategies with the signed time-series and quantile cross-sectional strategies.

Unsurprisingly all the correlations are positive and relatively high. The correlations within time-series and within cross-sectional strategies are very high indeed. We see that the scaled cross-sectional (time-series) strategy is more highly correlated with the quantile cross-sectional (signed time-series) strategy than the unscaled version – this agrees with intuition.

The correlations between time-series and cross-sectional strategies are lower for the industry data (look at the correlation of qxs with sts for instance) and indicate a possible diversification benefit from the

[10] These extreme returns for linear strategies are present in the industry data as well, but not evident here. Using the industry data with 17 portfolios (also composed by Fama and French) going back to 1926 one can see remarkably higher kurtosis from these strategies, which seems to result mostly from extreme returns in the 1930s. Scaling mostly eliminates this. I consider this dataset now because all 17 of the industries have data going back as far as 1926.

[11] As a check I also grouped returns into quarters (months) for the industry (MAA) data and recalculated the correlations (unreported). The pattern is the same, but the correlations are somewhat higher (probably due to noise from each strategy begin reduced).

Table 3.2: ***Sharpe ratios varying formation and holding period.*** *Annualised Sharpe ratios are calculated from monthly (weekly) returns for six different momentum strategies over the period Jul1969-Jun1994 (4Jan1979 - 5Dec2002) for the industry (MAA) data considering every possible combination of J month (week) formation and K month (week) holding period. Sharpe ratios are reported for the complete holding period. Holdings are rebalanced monthly (weekly) according to the weights assigned by each strategy. Sharpe ratios are annualised by multiplying by* $\sqrt{\frac{12}{K}}$ $(\sqrt{\frac{52}{K}})$.

(a) *Industries*

		K			
Strategy	J	1	3	6	12
FF49 qxs	1	1.02	0.34	0.26	0.31
	3	0.48	0.28	0.25	0.42
	6	0.53	0.36	0.44	0.51
	12	0.78	0.66	0.64	0.49
FF49 ulxs	1	0.58	0.21	0.16	0.24
	3	0.33	0.19	0.17	0.35
	6	0.27	0.19	0.35	0.42
	12	0.52	0.47	0.52	0.39
FF49 slxs	1	0.78	0.31	0.22	0.29
	3	0.42	0.29	0.22	0.40
	6	0.39	0.32	0.41	0.50
	12	0.70	0.65	0.64	0.49
FF49 sts	1	0.47	0.10	0.07	0.05
	3	0.22	0.02	-0.05	0.01
	6	0.08	-0.06	-0.06	-0.01
	12	0.10	0.06	0.07	-0.02
FF49 ults	1	0.48	0.12	0.04	0.03
	3	0.23	0.05	-0.05	0.01
	6	0.09	-0.08	-0.10	-0.01
	12	0.01	-0.04	-0.02	-0.06
FF49 slts	1	0.54	0.14	0.07	0.08
	3	0.25	0.06	-0.00	0.08
	6	0.14	0.02	0.02	0.08
	12	0.26	0.18	0.16	0.05

(b) *MAA*

		K				
Strategy	J	1	4	8	26	52
MAA qxs	1	0.52	0.61	0.42	0.21	0.22
	4	1.12	0.82	0.50	0.24	0.29
	8	0.98	0.68	0.43	0.26	0.33
	26	0.76	0.55	0.44	0.49	0.38
	52	0.87	0.72	0.62	0.49	0.28
MAA ulxs	1	-0.12	0.27	0.20	0.09	0.12
	4	0.35	0.46	0.24	0.12	0.19
	8	0.35	0.33	0.17	0.12	0.21
	26	0.27	0.21	0.13	0.34	0.24
	52	0.38	0.33	0.26	0.30	0.10
MAA slxs	1	0.21	0.50	0.32	0.18	0.19
	4	0.88	0.67	0.40	0.22	0.27
	8	0.79	0.60	0.40	0.24	0.32
	26	0.55	0.43	0.37	0.46	0.38
	52	0.68	0.62	0.53	0.36	0.21
MAA sts	1	0.99	0.76	0.51	0.23	0.23
	4	1.35	0.91	0.54	0.26	0.30
	8	1.19	0.83	0.51	0.29	0.36
	26	0.84	0.64	0.53	0.54	0.38
	52	1.07	0.82	0.71	0.48	0.23
MAA ults	1	0.08	0.36	0.24	0.11	0.12
	4	0.51	0.54	0.27	0.14	0.20
	8	0.46	0.38	0.16	0.14	0.23
	26	0.36	0.24	0.15	0.36	0.24
	52	0.45	0.34	0.25	0.29	0.07
MAA slts	1	0.47	0.55	0.36	0.20	0.17
	4	0.90	0.67	0.40	0.25	0.25
	8	0.81	0.57	0.34	0.26	0.29
	26	0.57	0.40	0.33	0.46	0.29
	52	0.57	0.46	0.37	0.29	0.05

Table 3.3: ***Strategy performance statistics.*** *The mean, annualised standard deviation, skewness, mean less the median, excess kurtosis, annualised Sharpe ratio, and the average of the largest five drawdowns (normalised with the ex-post standard deviation of returns) are calculated for six types of momentum strategies with 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period using asset returns over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for the industry (MAA) data. The mean is annualised as* $r_a = (1 + r_w)^N - 1$ *where* N *is 12 (52) for the industry (MAA) data. The standard deviation and Sharpe ratio are scaled by* $\sqrt{N}$ *and the mean less median figure by* N*.*

(a) *Industries*

strategy	mean	sd	skew	mean less med	kurtosis	Sharpe	ave drawdowns
FF49 qxs 12 mo	11.10	13.63	-0.53	-1.80	1.77	0.78	-5.13
FF49 qxs 1 mo	10.79	10.17	0.01	1.19	0.33	1.01	-4.23
FF49 ulxs 12 mo	0.52	0.99	-0.60	-0.13	3.31	0.52	-4.95
FF49 ulxs 1 mo	0.13	0.22	-1.29	-0.03	16.81	0.58	-5.43
FF49 slxs 12 mo	10.37	14.30	-0.53	-1.93	2.41	0.69	-5.01
FF49 slxs 1 mo	9.49	11.76	-0.28	-1.73	2.46	0.77	-4.81
FF49 sts 12 mo	1.36	13.93	-1.32	-6.75	6.79	0.10	-9.07
FF49 sts 1 mo	6.02	12.54	-0.49	0.81	2.52	0.47	-5.48
FF49 ults 12 mo	0.05	4.66	-1.82	-1.25	13.46	0.01	-10.16
FF49 ults 1 mo	0.52	1.10	0.37	0.23	10.50	0.48	-5.52
FF49 slts 12 mo	0.09	0.36	-1.04	-0.09	4.24	0.25	-7.51
FF49 slts 1 mo	0.17	0.31	-0.37	0.00	2.12	0.54	-5.39
FF49 equal-weighted	5.19	18.38	-0.39	-0.28	2.34	0.28	-5.74

(b) *MAA*

strategy	mean	sd	skew	mean less med	kurtosis	Sharpe	ave drawdowns
MAA qxs 52 wk	13.97	14.91	-0.12	-3.40	6.39	0.88	-11.64
MAA qxs 4 wk	17.19	14.33	0.56	-1.11	8.07	1.11	-8.83
MAA ulxs 52 wk	0.71	1.85	3.74	-0.26	112.15	0.39	-12.51
MAA ulxs 4 wk	0.18	0.51	-2.14	-0.04	123.02	0.35	-10.36
MAA slxs 52 wk	14.65	19.83	0.36	-4.58	14.45	0.69	-11.22
MAA slxs 4 wk	18.79	19.70	1.25	-2.50	19.12	0.88	-8.13
MAA sts 52 wk	6.01	5.41	-0.35	-1.97	3.41	1.08	-11.30
MAA sts 4 wk	7.48	5.37	0.34	0.53	6.27	1.34	-8.13
MAA ults 52 wk	1.00	2.21	4.07	-0.43	108.21	0.45	-13.11
MAA ults 4 wk	0.29	0.58	-0.23	0.02	93.63	0.50	-10.12
MAA slts 52 wk	0.61	1.05	3.24	-0.21	77.32	0.58	-13.35
MAA slts 4 wk	0.90	1.01	5.42	0.14	107.72	0.89	-8.05
MAA equal-weighted	4.06	5.68	-0.36	-2.21	2.32	0.70	-17.23

use of both time-series and cross-sectional strategies. These latter correlations are lower for the shorter formation period.

For the MAA data the correlations between the strategies are higher than for the FF49 data. Thus there is a lower diversification benefit. This suggests that possibly the cross-sectional strategy is mostly a time-series strategy on each asset – the assets being shorted being the ones to have gone down and vice versa. This is examined directly in section 4.2.

Table 3.4: ***Selected correlations between momentum strategies.*** *Correlations are reported for the quantile cross-sectional and signed-time series strategies with six types of momentum strategies. 12 month and 1 month (52 week and 4 week) formation period periods with a 1 month (1 week) holding period are considered for the industry (MAA) data over the period Jul1969-Jun1994 (4Jan1979-5Dec2002).*

		strategy					
strategy		qxs	ulxs	slxs	sts	ults	slts
FF49 12 mo	qxs	1.00	0.95	0.96	0.59	0.58	0.68
	sts	0.59	0.54	0.55	1.00	0.95	0.98
FF49 1 mo	qxs	1.00	0.86	0.92	0.31	0.37	0.43
	sts	0.31	0.31	0.32	1.00	0.87	0.97
MAA 52 wk	qxs	1.00	0.79	0.89	0.90	0.81	0.84
	sts	0.90	0.70	0.82	1.00	0.76	0.78
MAA 4 wk	qxs	1.00	0.73	0.90	0.89	0.76	0.81
	sts	0.89	0.67	0.81	1.00	0.73	0.76

In table 3.5 I report the intercepts of the aforementioned robust regressions along with p-values in square brackets. The OLS regression intercepts can be found in appendix E.1. I do not report the slope coefficients as these were all positive and very highly significant (unsurprisingly).

I expressed the return series in per cent and annualised by multiplying by 52 or 12 for the MAA and industry data respectively in order to make the intercepts more interpretable.

Table 3.5 (a) reports the intercepts of the time-series vs cross-sectional (and vice versa) regressions. For the industry data there appears to be a positive alpha[12] for cross-sectional vs time-series, which agrees with previous remarks. The magnitude is quite large for the nonlinear and scaled strategies. However, because of the varying (and small) bet-size of the unscaled linear strategies, the magnitude is hard to evaluate for these regressions. The outperformance of cross-sectional is most clearly shown for the 12 month formation, where time-series also has a negative alpha when regressed on cross-sectional. For the MAA data one would expect a positive alpha of time-series vs cross-sectional and this does appear to be the case (at least for the shorter formation period), but the evidence is only strong for quantile cross-sectional vs signed time-series.

The R-squared estimates for these regressions provide a similar conclusion to that drawn from the correlations. The R-squareds are higher for MAA, indicating a stronger relationship. For industries the 1 month R-squared is lower than the 12 month.[13] It is interesting that the relationship between the time-series and cross-sectional strategies is stronger for the MAA data than the industries. One may posit that this is due to a stronger relationship between the assets in the MAA dataset. However, looking at the average correlation between the assets in each dataset we see that this is 0.64 for the industries and 0.07 for the MAA data. For the MAA data some of the correlations are also negative so that the average absolute correlation is 0.14, still much lower than for the industries. This suggests that the higher correlation between time-series and cross-sectional strategies for the MAA data is more due to the positions being taken by time-series and cross-sectional strategies being similar (the assets that go up are the ones that the cross-sectional strategy longs and similarly the ones that go down are shorted). I will explore this issue later. In fact, if assets are strongly positively related, as in the industry data, one would expect them to tend to move in the same direction, which, if this is persistent would result in time-series outperforming cross-sectional, which we also do not see.

In table 3.5 (b) I report the intercepts of regression of scaled vs unscaled linear strategies.[14] Here

[12]I shall refer to the outperformance of one strategy over another as "alpha" in order to avoid confusion with the symbol α which occurs in other contexts in this thesis.

[13]For instance for industries qxs-sts R-squared values are 0.34 and 0.1 for the 12 month and 1 month formations respectively. For MAA these are 0.82 and 0.79 for the 52 week and 4 week formations respectively.

[14]The R-squareds for these regressions are mostly above 50% and often above 80%.

Table 3.5: ***Intercepts of regressions of momentum strategies on each other (robust).*** *Intercepts (alphas) and their p-values (in square brackets) are reported for robust regressions of monthly (weekly) momentum strategy returns on each other. The format for the column headings is "var1-var2" for var1 regressed on var2. The intercept for the opposite regression "var2-var1" is also reported in each case. In table (a) one of var1 and var2 is a time-series strategy and one is a cross-sectional strategy. In table (b) one of the variables is an unscaled linear strategy and the other a scaled linear strategy. In table (c) one of the variables is a non-linear strategy (either quantile cross-sectional or signed time-series) and the other is a scaled or unscaled linear strategy. The strategies were run with industry (MAA) data over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) and the regression used a bisquare weighting function with a parameter of 4.685. Returns are annualised by multiplying by 12 (52) and expressed in per cent.*

(a) *XS vs TS*

	regression					
formation	qxs-sts	sts-qxs	ulxs-ults	ults-ulxs	slxs-slts	slts-slxs
FF49 12 mo	9.66[0.000]	-1.88[0.460]	0.55[0.003]	-0.37[0.672]	6.86[0.004]	-0.03[0.664]
FF49 1 mo	9.49[0.000]	2.24[0.405]	0.12[0.006]	0.07[0.738]	7.56[0.001]	0.05[0.447]
MAA 52 wk	-0.25[0.862]	1.34[0.010]	0.06[0.554]	0.12[0.346]	1.37[0.602]	0.05[0.665]
MAA 4 wk	-2.07[0.156]	2.18[0.000]	-0.01[0.651]	0.03[0.317]	-0.92[0.668]	0.10[0.328]

(b) *Scaled vs unscaled*

	regression			
formation	slxs-ulxs	ulxs-slxs	slts-ults	ults-slts
FF49 12 mo	1.63[0.027]	-0.08[0.129]	0.06[0.057]	-0.49[0.248]
FF49 1 mo	0.98[0.345]	-0.01[0.609]	0.02[0.594]	-0.00[0.973]
MAA 52 wk	2.45[0.321]	-0.09[0.645]	0.05[0.463]	0.03[0.832]
MAA 4 wk	4.84[0.194]	-0.05[0.418]	0.12[0.420]	-0.01[0.837]

(c) *Linear vs non-linear*

	regression			
formation	qxs-ulxs	ulxs-qxs	qxs-slxs	slxs-qxs
FF49 12 mo	3.48[0.000]	-0.16[0.021]	0.92[0.266]	0.02[0.977]
FF49 1 mo	3.04[0.020]	-0.02[0.403]	1.74[0.044]	-0.65[0.518]
MAA 52 wk	5.21[0.038]	-0.02[0.953]	1.54[0.327]	1.71[0.402]
MAA 4 wk	7.26[0.025]	-0.07[0.390]	3.88[0.007]	-1.04[0.604]

	regression			
formation	sts-ults	ults-sts	sts-slts	slts-sts
FF49 12 mo	0.14[0.895]	0.12[0.732]	-2.06[0.002]	0.06[0.000]
FF49 1 mo	0.30[0.854]	0.03[0.825]	-1.09[0.091]	0.03[0.041]
MAA 52 wk	2.55[0.007]	-0.17[0.619]	1.96[0.032]	-0.06[0.694]
MAA 4 wk	3.69[0.000]	-0.07[0.452]	2.51[0.005]	-0.06[0.677]

we see an alpha for scaled strategies, though it is not always significant. For the robust regression it is not entirely clear, however, where the alpha for MAA 52 week formation time-series strategy lies (it is with the scaled strategy in the OLS regression). It appears that scaling is particularly beneficial for the cross-sectional strategies, indicating again that removing the effect of dispersion is good.

In table 3.5 (c) I report the intercepts of the regressions of the non-linear strategies vs the scaled and unscaled linear strategies.[15]

Here we find a clear alpha for the quantile cross-sectional strategies versus the unscaled linear strategies, which is not surprising as the quantile strategy reduces the exposure to dispersion in much the same way as scaling. There also appears to be an alpha for the quantile cross-sectional strategies versus the scaled linear strategies. This is very clear in the OLS regressions, but not for the 12 month (52 week) formation strategies for industries (MAA) in the robust regressions. It is, however, not surprising that there should be less of a difference between these two strategies.

For the time-series strategies we see an alpha for the signed strategy for the MAA data, but for industries it is the scaled linear time-series strategy that has best performance with no meaningful difference between the unscaled and signed strategies. The benefits of scaling are still evident here, but it seems there might be a reason to take into account the strength of a trend for allocation in time-series strategies.

3.4 Conclusions

- Momentum has a decreasing strength over the holding period.
- Shorter holding and formation periods are better (although a 1 week formation period is too short)
- In both the industry and MAA (except for a 1 week formation) data the strongest momentum is immediately after formation. There is no short-term reversal.
- Cross-sectional (time-series) momentum has stronger performance for industries (MAA).
- A momentum strategy appears to be superior to just investing in the equal-weighted market (at least when ignoring the linear (scaled and unscaled) strategies) – the only exception is the industry 12 month formation time-strategy, which is very weak.
- Unscaled linear strategies perform poorly and scaling improves performance on a risk-adjusted basis, suggesting that counteracting the effect of dispersion is positive for a momentum strategy.
- Linear strategies introduce somewhat extreme tail risks by amplifying volatility. This means Sharpe ratios are not adequate for evaluating these strategies.
- The distributions are generally negatively skewed, though this does not always show in the skewness coefficient.
- Correlations between the industry and time-series strategies (and also the R-squareds for the cross-sectional vs time-series regressions) are much higher for the MAA data. For industries there is a possible diversification between in combining time-series and cross-sectional strategies.
- It is possible that for MAA a cross-sectional strategy tends to short assets that went down and buy assets that went up, effectively mimicking a time-series strategy.

[15]The R-squareds for these regressions are perhaps worth noting very briefly. They are above 90% for the industry 12 month formation strategies and above 60% (but as high as 94% for sts-slts regressions) for the 1 month formation. For the MAA regressions they are mostly above 40% and always lower than the industry regression R-squared.

Chapter 4

Source of momentum profits

In this chapter I look more closely at where momentum profits come from. I am not so much interested in behavioural explanations as in relationships between asset returns that could explain why momentum strategies produce profits. Any differences between time-series and cross-sectional momentum here are of particular interest.

Section 4.1 considers empirically the decomposition for linear strategies from section 2.3.1. The next section examines the performance differences between time-series and cross-sectional momentum using the ideas from section 2.2 by looking at the long and short positions of the strategies. Section 4.3 examines the predictive abilities of signed strategies (as defined in section 2.3.2) run on individual assets. Section 4.4 considers the predictive power of the rank of formation period returns and section 4.5 examines the validity of the AR(1) assumptions suggested in section 2.5.1. Lastly, section 4.6 examines the performance of global and local time-series strategies.

In section 5.2 I reconsider many of the conclusions in this chapter for normalised returns and find that they mostly remain unchanged.

4.1 Empirical decomposition

I consider the decomposition in section 2.3.1 and denote the terms as follows

$$E[r^X_{t,t+k}] = \underbrace{\frac{N-1}{N^2}\mathrm{tr}(\Omega_t)}_{\text{auto}} \underbrace{-\frac{1}{N^2}(\mathbf{1}'\Omega_t\mathbf{1} - \mathrm{tr}(\Omega_t))}_{\text{cross}} + \underbrace{\frac{\boldsymbol{\mu}'_{t-j,t}\boldsymbol{\mu}_{t,t+k}}{N} - \frac{\sum_i \boldsymbol{\mu}_{t-j,t}(i)\sum_i \boldsymbol{\mu}_{t,t+k}(i)}{N^2}}_{\text{mean}} \tag{4.1}$$

$$E[r^T_{t,t+k}] = \underbrace{\frac{\mathrm{tr}(\Omega_t)}{N}}_{\text{auto}} + \underbrace{\frac{\boldsymbol{\mu}'_{t-j,t}\boldsymbol{\mu}_{t,t+k}}{N}}_{\text{mean}} \tag{4.2}$$

I calculate the above terms for each type of strategy (implicitly assuming stationarity) in order to gauge the sources of momentum profits. I do this for both the MAA data and industry portfolio data, considering all possible combinations of formation period and holding period. 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period are considered for the industry (MAA) data. For the MAA data I consider only 11 assets and a period over which data is available for all these assets. The decompositions are reported in table 4.1. Figures are in per cent and annualised (multiplied by 52 for MAA data and by 12 for the industry data).

In all cases there are positive momentum profits from both strategies. Note that because bet sizes are proportional and of varying size it is somewhat hard to compare (total) profits between strategies and also to gauge whether they are economically large.[1]

For the industry data there is a clear difference between the long and short formation. For the 12 month formation, the auto-covariance component is negative, whereas the cross-serial covariance component is

[1]For instance, there are seemingly weaker profits for the shorter formation periods. However, where we took into account variability in returns in chapter 3, this conclusion was reversed. The 12 month asset returns are naturally larger than the 1 month returns and so the 12 month strategy naturally takes larger bets, making this an unfair comparison. A more fair comparison would be to scale the shorter formation returns by 12 (13) for the industry (MAA) data. One gets 1.56, 6.3 (0.793, 1.612) for the cross-sectional and time series strategies for the industries (MAA) data. Thus we do see a larger profit for shorter formation with this scaling.

Table 4.1: ***Empirical decomposition for linear strategies.*** *Estimates for the decomposition for linear strategies in section 2.3.1 are reported from returns from all possible combinations of formation period and holding period. 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period are considered for industry (MAA) data over the period Jul1969-Jun1994 (31Dec81-5Dec2002). For the MAA data a set of 11 assets (for which data is available over the whole period) is considered. Figures are in per cent and annualised by multiplying by 12 (52). Percentage of total profits is indicated in brackets.*

strategy	auto	cross	mean	total
FF49 12 mo XS	-0.43(-83.99)	0.88(170.27)	0.07(13.72)	0.52(100.00)
FF49 12 mo TS	-0.44(-905.56)		0.49(1005.56)	0.05(100.00)
FF49 1 mo XS	0.49(372.13)	-0.36(-276.61)	0.01(4.48)	0.13(100.00)
FF49 1 mo TS	0.50(94.41)		0.03(5.59)	0.52(100.00)
MAA11 52 wk XS	0.50(82.50)	-0.11(-17.92)	0.22(35.42)	0.61(100.00)
MAA11 52 wk TS	0.55(63.36)		0.32(36.64)	0.87(100.00)
MAA11 4 wk XS	0.09(150.27)	-0.05(-74.67)	0.01(24.40)	0.06(100.00)
MAA11 4 wk TS	0.10(81.76)		0.02(18.24)	0.12(100.00)

positive.[2] This agrees with the decomposition seen by Lewellen (2002) (but for longer holding periods) but not Moskowitz *et al.* (2012) (the latter look at individual stock momentum not industries). The long formation cross-sections momentum is driven by lead-lag relations, whereas for the time-series momentum there is in fact a reversal effect. The only profit is due to mean returns being positive.

For the MAA data we see most of the profits coming from auto-correlations and a negative contribution from cross-serial covariances for cross-sectional momentum. The mean contribution is also not inconsiderable.

It should not surprise, perhaps, that we find the 12 month formation time-series momentum to be very weak in chapter 3. The reason for the outperformance of the cross-sectional strategy here is clear. There are strong lead-lag relationships as in case II in section 2.2 and autocovariances even contribute negatively.

For the industry 1 month formation strategies and the MAA strategies, it would seem to follow from the negative contribution of the cross term (lead-lag relations) that time-series strategies should perform better as in case I in section 2.2. We do see this for the MAA data. However, the industry 1 month cross-sectional strategy outperforms despite the pattern of the decomposition and this is somewhat mysterious. The decomposition explains why the time-series momentum profits are better for the 1 month formation than the 12 month, but not why the cross-sectional strategy still outperforms. The decomposition even predicts a higher profit for the time-series strategy.[3] I will examine this issue further in section 4.2. It may then just be coincidental that Moskowitz *et al.* (2012)'s decomposition agreed with their finding of an alpha for time-series versus cross-sectional strategies as this decomposition does not take into account variability or the scale of bets at all.

The contribution of means is not dominant for any of the momentum strategies, except for the industry 12 month time-series momentum, indicating that it is more than just large means or a dispersion in mean returns causing momentum profits – there is a clear relationship between past and future returns.

4.2 Long-short analysis

The tests in this section are inspired by the intuitive and theoretical insights in section 2.2. That is, I consider the long and short positions of momentum strategies and particularly where these differ between the time-series and cross-sectional strategies. The goal is to see where the outperformance of one strategy over the other comes from. I will use as basis for these tests a simple signed time-series strategy and a quantile cross-sectional strategy with 2 quantiles (i.e. one that sells the bottom 50% and buys the top 50% of assets). For each of the 12 month and 1 month (52 week and 4 week) formation strategies I construct three different tables.

[2] An aside: the pattern of profits seen here is not necessarily constant. Taking a longer view (using the 17 industry portfolios) for data from 1926 we see that the decomposition of 12 month formation now more closely resembles what we saw for the 1 month momentum.

[3] The absolute (not risk adjusted) profits of 1 month unscaled linear time-series strategy are in fact higher than the cross-sectional strategy profits but with a higher variance, which gives the Sharpe ratio figures in table 3.3a.

The first table (*long and short portfolios*) considers four different portfolios in each period. The first (last) portfolio, long|long (short|short), considers all the assets that a time-series strategy as well as a 2 quantile cross-sectional strategy would buy (sell). The other two portfolios consider those assets that are treated differently by the two strategies. The long|short portfolio consists of those assets the time-series strategy would buy and the cross-sectional strategy would sell (these are like the bottom asset in the up|up case in figure 2.2 (a)) and similarly the short|long portfolio considers those assets the time-series strategy would sell and the quantile cross-sectional strategy would buy (these are like the top asset in the down|down case in figure 2.2 (2.2c)). Note that in any given period at most one of the two portfolios long|short and short|long can be non-empty. Any outperformance for one of the two strategies must come from these portfolios. I report the return and Sharpe ratio during the holding period for each portfolio (considering only those cases where the portfolio is non-empty) as well as the average proportion of assets in each portfolio (including times where the portfolio is empty). These are in tables 4.2 (a1), (b1), (c1) and (d1).

The second table (*two quantile strategies*) divides the assets into two quantiles based on each formation period (if necessary omitting one asset) and constructs two portfolios, the top and the bottom portfolio which equally weight the assets in each quantile. This is designed to mimic the two-asset scenarios of section 2.2. A time-series strategy would buy any of the two portfolios with a positive return and a cross-sectional strategy would buy the top and sell the bottom as before. I report the average return and Sharpe ratio in the holding period for each of the three cases (up|up, up|down, down|down) considered in figure 2.2. It is in the up|up and down|down cases that any difference between the two strategies must occur. I also report the proportion of periods in which each of the three cases occurs. This second set of tables consists of tables 4.2 (a2), (b2), (c2) and (d2).

The third table (*two quantile strategies: four scenarios*) considers in more detail the up|up and down|down cases in the previous table. It divides each of these into the four scenarios in figure 2.3 and reports the proportion of periods in which each scenario occurs (for each case) and the Sharpe ratio for the time-series and cross-sectional strategies. This third set of tables consists of tables 4.2 (a3), (b3), (c3) and (d3).

4.2.1 Long and short portfolios

In the *long and short portfolios* tables we see that time-series outperforms in the long|short case and cross-sectional in short|long case for the industries. The outperformance of the cross-sectional strategies must come from the latter case. This, however, remains puzzling because the proportion of assets is higher for the case where time-series outperforms and for the 1 month formation the returns and Sharpe ratio are higher as well. We can, however, conjecture that there is a tendency for positive returns to continue and for negative returns to reverse (possibly the effect of positive mean returns is important), which causes time-series to outperform in the one portfolio and cross-sectional in the other. This will be seen again in the next section (on prediction analysis).

For the MAA data we see that the time-series strategy outperforms in both the long|short and short|long portfolios. We must conclude there is a much stronger tendency for positive or negative returns to continue (the effect of a positive mean return is less important) in this case, which harms the cross-sectional strategy compared to the time-series strategy.

We see that most of the time the two strategies seem to take similar positions. The proportion of assets in the long|short and short|long positions are the lowest and are lower for the MAA data.

Note also that the long|long portfolio outperforms the short|short portfolio (in fact for the industries the latter portfolio contributes negatively), which suggests that most of the profits come from the long positions and long-only strategies may do better (note that this is different from findings of some studies in equity momentum where the short portfolio is the largest contributor, for instance in Bhootra (2011)).

It does not appear that a time-series strategy could be improved by taking into account the cross-sectional signal, for instance when an asset has negative time-series momentum, the strategy would not do better by only shorting the assets that are also in the bottom quantile (except for the industry 12 month formation). It does, however, appear that a cross-sectional strategy could be improved by taking into account the time-series signal. For instance by only shorting those assets in the bottom quantile that have negative times-series momentum.

4.2.2 Two quantile strategies

In the *two quantile strategies* tables we see that for the industry 12 month formation in both the up|up and down|down case the cross-sectional strategy seems to outperform. However, for the 1 month strategy we again see that time-series does better in the up|up case and cross-sectional does better in the down|down case. The outperformance of the cross-sectional strategy overall is still surprising because the down|down case is less frequent and the profits do not appear to be extremely high compared to the other case.

For the MAA strategies we see the same pattern for the 52 week formation – time-series outperforms in the up|up case and cross-sectional in the down|down case. However the latter occurs very infrequently and so it is clear that time-series should outperform. For the 4 week formation time-series outperforms in both cases.

We again see that most of the time the strategies invest similarly, with the up|down case occurring with the greatest frequency. This frequency is particularly high for the MAA data. In the up|down case also see that the long position outperforms the short in all cases and the latter contributes negatively for the industry 12 month formation.

4.2.3 Four scenarios

In the last set of tables, *two quantile strategies: four scenarios*, we see for the industry strategies that scenarios two and three hardly occur. The outperformance of the cross-sectional strategies must come mostly from scenario 4, which is where both assets reverse. In particular we may conjecture that it is mostly from the down|down case, i.e. a reversal when both assets went down in the formation period.

For the 12 month formation scenarios 3 and 4 combined (where cross-sectional outperforms) occur more frequently in both the up|up and down|down cases, but for the 1 month formation this is only so for the down|down case. Focussing on this case we see for both the 12 month and 1 month formation that the time-series strategy seems to lose most of its profits from scenario 1 in scenario 4 (scenario 1 and 4 occur with almost equal frequency).

Looking at the MAA data we see that here scenario 1 is the most frequent (except in the down|down case for the 12 month formation, but this case is very infrequent). I conclude that it is mostly from positive return continuations that time-series outperforms. Note also that for both the MAA and industry data scenario 2 is the least frequent (in fact in the down|down case it is so infrequent that a Sharpe ratio cannot always be computed). This is as was expected in section 2.2 – this scenario requires only the asset with the strongest trend to revert.

4.2.4 Overall

Cross-sectional strategies seem to outperform for industries due to a tendency for negative returns to reverse, whereas time-series outperforms for the MAA data due to a stronger tendency for returns to continue and it is particularly the continuation of positive returns that is important. Most of the time, however, the two strategies invest similarly, particularly for the MAA data. This explains why the two strategies were more closely related in chapter 3.3 for the MAA data and confirms the conjecture made there. It appears that long positions are more profitable than short positions, suggesting that long-only strategies may be useful. In fact there seems to be evidence for this in Antonacci (2013a) (although the author does not directly compare the long-only and long-short strategies).

Overall the outperformance of the cross-sectional strategy for the industry 1 month formation remains a mystery.[4] However, as we noted in section 2.2, for the cross-sectional strategy to outperform in the presence of auto-covariances and a positive mean one would expect the strategy to outperform only in the down|down case, which is exactly what we see.

[4]One may conjecture that the outperformance may come from using a four quantile cross-sectional strategy rather than a two quantile strategy. However, the latter strategy has a Sharpe ratio of 1.03 with a 1 month formation and this is even a little higher than the Sharpe ratio with four quantiles

Table 4.2: ***Long-short analysis.*** *The* long and short portolfolio tables *report returns, Sharpe ratios and the average proportion of assets in each of four portfolios dividing assets into those that a times-series and two quantile strategy would buy or sell. The* two quantile strategies *tables divide assets into a top and bottom half (possible omitting one asset) each formation period and reports the returns and Sharpe ratio for each half in three cases (both quantiles going up in the formation period; one going up and one down; and both going down) as in figure 2.2 and the proportion of occurrences for each case. The* four scenarios *tables reports time-series and cross-sectional Sharpe ratios for each of four scenarios (in up|up and down|down cases) as defined in figures 2.3 and 2.4 along with the proportion of occurrences of each scenario. Industry (MAA) data for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) is used. Where there is not enough data to calculate an entry, this is indicated by* NaN. *Sharpe ratios are annualised.*

(a) *Industry 12 month formation*

(a1) *Long and short portfolios*

portfolio	return	Sharpe	proportion
long|long	13.25	0.64	0.39
long|short	5.07	0.31	0.19
short|long	10.41	0.51	0.10
short|short	4.57	0.24	0.30

(a2) *Two quantile strategies*

case	portfolio	return	Sharpe	proportion
up|up	top	0.73	0.04	0.37
	bottom	-4.07	-0.26	
up|down	top	18.39	1.09	0.42
	bottom	9.55	0.53	
down|down	top	13.28	0.67	0.20
	bottom	7.70	0.29	

(a3) *Two quantile strategies: four scenarios*

		scenario			
case	statistic	sc1	sc2	sc3	sc4
up|up	proportion	0.45	0.03	0.07	0.45
	TS Sharpe	1.41	0.06	0.36	-0.95
	XS Sharpe	0.56	-1.43	1.69	-0.20
down|down	proportion	0.47	0.00	0.05	0.47
	TS Sharpe	1.76	NaN	0.47	-1.55
	XS Sharpe	0.94	NaN	1.90	-0.40

(b) *Industry 1 month formation*

(b1) *Long and short portfolios*

portfolio	return	Sharpe	proportion
long|long	7.69	0.40	0.35
long|short	13.10	0.86	0.16
short|long	3.07	0.14	0.14
short|short	0.61	0.03	0.33

(b2) *Two quantile strategies*

case	portfolio	return	Sharpe	proportion
up|up	top	16.89	1.17	0.31
	bottom	12.25	0.87	
up|down	top	4.91	0.25	0.45
	bottom	-1.18	-0.06	
down|down	top	7.53	0.35	0.24
	bottom	-4.96	-0.22	

(b3) *Two quantile strategies: four scenarios*

		scenario			
case	statistic	sc1	sc2	sc3	sc4
up|up	proportion	0.54	0.02	0.07	0.36
	TS Sharpe	1.65	0.57	0.23	-1.60
	XS Sharpe	0.25	-1.63	1.81	0.08
down|down	proportion	0.46	0.01	0.07	0.45
	TS Sharpe	1.50	NaN	-0.47	-1.41
	XS Sharpe	0.77	NaN	1.79	0.03

Table 4.2: ***Long-short analysis.*** *(continued from previous page)*

(c) *MAA 52 week formation*

(c1) *Long and short portfolios*

portfolio	return	Sharpe	proportion
long|long	6.79	0.65	0.45
long|short	4.82	0.79	0.12
short|long	-5.77	-0.75	0.04
short|short	-3.90	-0.41	0.37

(c2) *Two quantile strategies*

case	portfolio	return	Sharpe	proportion
up|up	top	13.18	1.88	0.08
	bottom	5.13	1.17	
up|down	top	7.40	1.03	0.91
	bottom	-1.03	-0.12	
down|down	top	12.01	2.37	0.01
	bottom	-33.24	-5.56	

(c3) *Two quantile strategies: four scenarios*

		scenario			
Scenario	statistic	sc1	sc2	sc3	sc4
up|up	proportion	0.40	0.14	0.23	0.23
	TS Sharpe	1.75	-0.71	0.54	-1.76
	XS Sharpe	0.47	-2.40	2.18	-0.70
down|down	proportion	0.27	0.09	0.45	0.18
	TS Sharpe	2.10	NaN	0.63	-4.64
	XS Sharpe	-0.02	NaN	2.45	3.71

(d) *MAA 4 week formation*

(d1) *Long and short portfolios*

portfolio	return	Sharpe	proportion
long|long	11.31	1.40	0.44
long|short	5.14	0.71	0.09
short|long	-5.91	-0.68	0.05
short|short	-2.80	-0.31	0.40

(d2) *Two quantile strategies*

case	portfolio	return	Sharpe	proportion
up|up	top	25.03	2.95	0.05
	bottom	9.02	2.09	
up|down	top	9.23	1.28	0.92
	bottom	-2.26	-0.28	
down|down	top	-5.90	-1.34	0.03
	bottom	-12.33	-0.89	

(d3) *Two quantile strategies: four scenarios*

		scenario			
case	statistic	sc1	sc2	sc3	sc4
up|up	proportion	0.44	0.17	0.24	0.15
	TS Sharpe	1.34	-0.08	0.79	-1.41
	XS Sharpe	0.53	-1.32	1.40	-0.19
down|down	proportion	0.34	0.21	0.21	0.24
	TS Sharpe	1.26	-0.59	0.20	-0.96
	XS Sharpe	0.71	-1.73	1.26	-0.47

4.3 Prediction analysis

Here I consider the signed strategies of section 2.3.2 run on individual assets. That is I consider timing strategies on $r_{i,t}$ for time-series and $\tilde{r}_{i,t}$ (being deviations from the cross-sectional average) for cross-sectional. The goal is to see how much of momentum profits is derived from

- accurately predicting the sign of the next return, and
- the size of the profit(loss) given a correct(incorrect) prediction

and to investigate whether momentum profits or prediction accuracy are

- better or worse for positive or negative predictions, and
- better or worse in high or low volatility conditions.

I also consider the interaction of the time-series and cross-sectional signals. The tables I report are inspired by a simpler table in Wang and Xu (2009), based on overall stock momentum.

First I consider the accuracy of the predictions, i.e. how often a positive or negative return (or deviation) is followed by a return (or deviation) of the same sign. In the models of 2.3.2 we need at least a 50% accuracy for the strategies to be profitable. In table 4.3 (a) I report the proportion of accurate predictions for these strategies. I consider both overall prediction accuracy and accuracy when the prediction is positive or negative and the volatility is high or low. I report only the overall prediction accuracy across all assets (i.e. aggregating all the data points for each asset in each category and then calculating the proportion of correct predictions).

In order to determine low/high volatility states, I do the following. I calculate both a fast and a slow EWMA volatility of the return series (either $r_{i,t}$ or $\tilde{r}_{i,t}$). A high volatility state is defined as one where the fast EWMA is larger than the short EWMA and a low volatility state is the opposite. Note that for the cross-sectional strategies the volatility is the volatility of the deviations, not of the asset returns and that the volatility is per-asset (I have not used an overall market volatility).

For the industry data I use daily data for the volatility estimates. The short EWMA has $\lambda = 0.9836$ (effective history 61 days) and the long EWMA has $\lambda = 0.996$ (effective history of 250 days or 1 year). For the MAA data the short EWMA has $\lambda = 0.97$ (effective history $33^1/_3$ weeks) and the long EWMA has $\lambda = 0.9936$ (effective history of 156 weeks or 3 years). Please see appendix C.5 for details on the volatility estimates.

In order to get a better idea of where the prediction accuracy of the strategies comes from I determine the prediction accuracy (p_A) less the prediction accuracy expected under independence of returns (p_I). The latter is defined as $P(r_{t-1} > 0)P(r_t > 0) + P(r_{t-1} < 0)P(r_t < 0)$, which I estimate by aggregating the return data across all assets and estimating the relevant probabilities. I also report a t-statistic based on the normal approximation to the binomial distribution, defined as

$$\frac{\sqrt{n}(p_A - p_I)}{\sqrt{p_I(1 - p_I)}}$$

I refer to $p_A - p_I$ as the excess accuracy of the strategy. These figures make no distinction between assets (implicitly it assumes the signs of asset returns all have the same distribution). The excess accuracy figures are in table 4.4.

In table 4.3 (b) I report Sharpe ratios of returns divided into the same categories as the prediction accuracy table. I calculate the Sharpe ratio for each asset and then report a weighted average across the assets.[5] In table 4.3 (c) I calculate the average profit (across all assets) given that the prediction is correct and the average loss (across all assets) given that the prediction is wrong and report the ratio of these figures.

I further calculate Sharpe ratios for the signed strategies when the time-series and cross-sectional signals agree (the return is positive (negative) and the deviation from the average return is positive (negative)) and disagree (when the return is positive (negative) and the deviation from the average return is negative (positive)). One would expect stronger profits when the signals coincide than when they disagree. If so, there may be scope for taking into account both signals as done by Antonacci (2013b) for long-only stock momentum. These Sharpe ratios are in table 4.5. These Sharpe ratios are calculated by first aggregating returns across all assets (as opposed to averaging Sharpe ratios calculated for each asset as was done previously).

[5]I weight by the number of observations used to calculate each Sharpe ratio. I do so in order to avoid allowing the average to be dominated by Sharpe ratios calculated from few observations, which can easily become very extreme and turned out to be problematic in the more recent data in chapter 7.

Table 4.3: *Prediction analysis for signed strategies (accuracy, annualised Sharpe ratios, profits over losses).* *Tables (a), (b) and (c) consider, respectively, the proportion of correct predictions, the (weighted average) Sharpe ratio and ratio of average profits when a prediction is correct to losses when wrong, for signed time-series (cross-sectional) strategies run on individual asset returns (deviations). Positive and negative predictions and low and high volatility states are considered separately. Volatilities are estimated with a slow and fast EWMA per asset and a high (low) volatility state is when the fast EWMA is above (below) the slow EWMA. Returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for industries (MAA) and the first 21 days (52 weeks) of data for each asset are used to obtain initial volatility estimates (so the first available return is then for the first full calendar month after 21 days for industries).*

(a) *Accuracy*

	+ prediction			− prediction			overall		
strategy	σ low	σ high	all +	σ low	σ high	all −	σ low	σ high	all
FF49 12 mo XS	53.52	50.50	52.31	53.18	53.18	53.18	53.34	51.90	52.76
FF49 1 mo XS	54.35	51.69	53.27	53.94	54.14	54.02	54.14	52.92	53.65
FF49 12 mo TS	51.34	61.01	54.22	51.00	46.18	48.94	51.22	53.78	52.12
FF49 1 mo TS	51.83	60.93	55.08	52.45	47.86	50.90	52.10	54.86	53.06
MAA 52 wk XS	53.51	52.57	53.10	54.06	52.07	53.23	53.81	52.31	53.17
MAA 4 wk XS	53.69	52.86	53.23	54.77	52.51	53.86	54.25	52.68	53.56
MAA 52 wk TS	56.83	55.58	56.38	53.86	50.62	52.28	55.70	53.19	54.64
MAA 4 wk TS	57.59	55.82	56.91	52.92	51.05	52.26	55.45	53.61	54.77

(b) *Sharpe ratios*

	+ prediction			− prediction			overall		
strategy	σ low	σ high	all +	σ low	σ high	all −	σ low	σ high	all
FF49 12 mo XS	0.33	0.09	0.22	0.23	0.14	0.19	0.27	0.12	0.20
FF49 1 mo XS	0.37	0.14	0.26	0.29	0.23	0.25	0.32	0.19	0.26
FF49 12 mo TS	0.17	0.69	0.34	-0.13	-0.45	-0.28	0.06	0.09	0.07
FF49 1 mo TS	0.34	0.90	0.56	0.15	-0.16	0.02	0.23	0.31	0.27
MAA 52 wk XS	0.34	0.32	0.32	0.48	0.26	0.37	0.41	0.27	0.33
MAA 4 wk XS	0.50	0.31	0.39	0.69	0.23	0.47	0.59	0.24	0.41
MAA 52 wk TS	0.77	0.55	0.65	0.58	-0.01	0.25	0.68	0.25	0.46
MAA 4 wk TS	0.97	0.70	0.86	0.58	0.14	0.38	0.76	0.40	0.60

(c) *Profits over losses*

	+ prediction			− prediction			overall		
strategy	σ low	σ high	all +	σ low	σ high	all −	σ low	σ high	all
FF49 12 mo XS	1.12	1.07	1.09	1.08	0.97	1.02	1.09	1.01	1.05
FF49 1 mo XS	1.08	1.05	1.06	1.05	0.99	1.02	1.07	1.02	1.04
FF49 12 mo TS	1.06	1.07	1.09	0.90	0.83	0.86	1.00	0.92	0.97
FF49 1 mo TS	1.18	1.24	1.22	1.04	0.96	0.99	1.10	1.04	1.08
MAA 52 wk XS	0.99	1.05	1.02	1.02	0.99	1.00	1.01	1.01	1.00
MAA 4 wk XS	1.06	1.00	1.03	1.08	0.94	0.99	1.07	0.97	1.01
MAA 52 wk TS	0.96	1.02	0.99	1.06	0.96	0.99	1.00	0.99	0.98
MAA 4 wk TS	1.01	1.02	1.01	1.12	0.96	1.02	1.06	0.99	1.01

Table 4.4: ***Prediction analysis for signed strategies (excess accuracy).*** *The proportion of correct predictions (in excess of those expected under independence) along with t-statistic in brackets are reported for data over the period the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for industries (MAA).*

industries		MAA	
strategy	excess accuracy	strategy	excess accuracy
FF49 12 mo XS	2.71(6.44)	MAA 52 wk XS	3.04(7.91)
FF49 1 mo XS	3.59(8.69)	MAA 4 wk XS	3.57(9.53)
FF49 12 mo TS	1.40(3.32)	MAA 52 wk TS	4.12(10.73)
FF49 1 mo TS	2.86(6.92)	MAA 4 wk TS	4.63(12.35)

4.3.1 Prediction accuracy

The overall prediction accuracy in table 4.3 (a) is above 50% for all the strategies, indicating that at least a part of the momentum profitability is from this source. Looking at table 4.4 we see that the prediction accuracy is higher than under independence and that this appears to be significant. In fact, it appears this excess accuracy accounts for virtually all of the accuracy of the strategies exceeding 50%. This indicates a dependence between successive returns (deviations) that allows a momentum strategy to profit from predicting the sign of the next return. Note also that the prediction accuracy (and excess accuracy) of the cross-sectional (time-series) signal is marginally higher for the industry (MAA) data in agreement with the strategy performances seen. It is also notable that the weakest momentum strategy, the industry 12 month formation time-series strategy, has the weakest and smallest excess accuracy.

Except for the industry time-series strategies volatility seems to negatively affect predictive accuracy as suggested by the comments in section 2.6.1 (note also it remains above 50%). The industry time-series strategies behave very differently though. They have a very high predictive accuracy (over 60%) in high volatility states with positive predictions. However, with negative predictions high volatility is again hurtful, even lowering the accuracy to below 50%. It may be that for these strategies the correlation between σ_t and r_{t-1}^2 dominates as discussed in section 2.6.1. However, it seems that for these strategies an asymmetric model for positive and negative predictions is needed. The other strategies lack such an asymmetric effect of volatility on prediction accuracy.[6] It is harder to make a clear conclusion on the overall difference between positive and negative predictions. Negative predictions seem to be more accurate for the cross-sectional strategies and positive predictions for the time-series. The latter confirms that negative returns tend to reverse.

4.3.2 Sharpe ratios and profits over losses

In table 4.3 (b) we see a higher Sharpe ratio for cross-sectional (time-series) strategies for the industry (MAA) data. We also see that with the exception of the industry time-series strategies the Sharpe ratio is lower in the high volatility state. Again there is an asymmetry in the positive and negative predictions only for the industry time-series strategies. The negative effect of volatility could possibly be explained by the model in section 2.6.2 but an asymmetric model would be needed for the industry time-series strategies. It is also useful to note that profits are positive in almost all states: only some of the time-series negative predictions are not.

Except for the MAA cross-sectional strategies we see a higher Sharpe ratio for positive predictions. In particular it seems that the outperformance of cross-sectional strategies in the industries stems from a poor performance of time-series momentum for negative predictions. Negative momentum is short-lived. However, for the MAA data, the time-series strategies also perform poorly with negative momentum. This would suggest that a long-only strategy would do better. This again agrees with the conclusions in section 4.2.

We also see an asymmetry in that average profits in table 4.3 (c) when correct are generally larger than the losses when wrong. This asymmetry is not taken into account in the models for the signed strategies in section 2.3.2 and is an additional source of profit. However the ratio is still quite close to 1, which suggests the assumption of the models is not entirely unreasonable. The ratio of profits to losses is also lower in the high volatility state (this time for all the strategies) – this stems from the negative prediction states where this is seen most acutely.

[6]The industry 1 month XS strategy has a slightly higher accuracy for high volatility under negative predictions, which goes against the grain of the other results.

4.3.3 Agreeing and opposing signals

Table 4.5: ***Prediction analysis for signed strategies (agreeing and opposing signal Sharpe ratios)*** *Annualised Sharpe ratios for signed time-series and cross-sectional strategies (based on aggregated returns across all assets) are reported for cases where the time-series and cross-sectional signals agree and disagree. The time-series (cross-sectional) signal for an asset is the sign of the formation period return (deviation). Data for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for industry (MAA) data.*

(a) *Opposing signals*

	TS neg XS pos		TS pos XS neg		all	
data and formation	TS	XS	TS	XS	TS	XS
FF49 12 mo	-0.45	0.37	0.39	0.16	0.02	0.25
FF49 1 mo	-0.29	0.38	0.70	0.09	0.16	0.23
MAA 52 wk	0.18	-0.35	0.45	0.04	0.38	-0.05
MAA 4 wk	0.89	-0.37	0.49	0.33	0.64	0.06

(b) *Agreeing signals*

	TS neg XS neg		TS pos XS pos		all	
data and formation	TS	XS	TS	XS	TS	XS
FF49 12 mo	0.04	0.22	0.34	0.16	0.19	0.19
FF49 1 mo	0.15	0.28	0.46	0.17	0.29	0.23
MAA 52 wk	0.17	0.35	0.60	0.39	0.39	0.37
MAA 4 wk	0.16	0.37	0.75	0.49	0.45	0.43

The results in table 4.5 are not particularly conclusive. For industries (MAA) we see that the time-series (cross-sectional) profit is higher when the two signals agree. That is, when the signal of the weaker strategy agrees with signal of the stronger strategy, the weaker strategy does better (which is not all that surprising). However, the stronger strategy does not do better. The source of this appears to be in the left side of the top table where the only negative profits occur, when the weaker strategy disagrees with the stronger one. The table does not seem to indicate that a cross-sectional strategy that avoids going long when the time-series signal is negative will perform better for industries. However, it may do better for MAA. A cross-sectional strategy that avoids going short with positive time-series momentum may, however, do better on both counts, which again just illustrates that short positions are weak. The results here seem to be consistent with those of Thomas *et al.* (2012) who find one can improve (long only) cross-sectional momentum for asset allocation with trend-following signals, but that this is not better than a trend-following strategy.

4.3.4 Overall

With the notable exception of the industry time-series strategies it appears that momentum is negatively affected by volatility. It seems that predictive accuracy (including an excess accuracy) and an asymmetry in profits and losses contribute to momentum profits. There is no real evidence that combining cross-sectional and time-series signals would be useful here. However, we again see that short positions are weak and could potentially be useful to avoid.

4.4 Ranked returns

Here I look at whether the rank of an asset's return in the formation period has predictive power for its return in the holding period. I rank assets by their formation period return for all possible pairs of holding period and formation period and calculate the average return over the holding and formation periods for each rank. The results here indicate the presence of momentum but do not seem to allow for a nice distinction between cross-sectional and time-series momentum.

For the MAA data where some assets are missing in the beginning I still keep rankings from 1 to 17 (the total number of assets), but I move assets to the extremes (so for instance, if only 8 assets are available the asset with the lowest rank would be placed under rank 17 and the highest under rank 1). The missing ranks are treated as missing data and my averages only include returns actually available.[7]

I group the above average returns into quantiles (by rank over the formation period) and calculate average return over the quantiles (the average of the average returns in each quantile). This reduces the size of the tables and makes the results less “noisy”. For the industry data I use quantiles of 4 assets (1 asset in the middle is omitted). For the MAA data with 17 assets I use quantiles of 3 assets (2 assets in the middle are omitted).[8] I report the returns in these quantiles in table 4.6 for 1 year and 1 month (52 week and 4 week) formation periods and for 1 month (1 week) holding periods for the industry (MAA) data. All return figures are annualised.

Table 4.6: ***Returns in formation and holding period by quantile in formation period.*** *Average returns for each rank are calculated. Ranks are then grouped into quantiles and the annualised average of these average returns is reported. Data over the period Jul1969-Jun1994 MAA (4Jan1979-5Dec2002) for industry (MAA) data are used.*

(a) *FF49 12 month formation quantiles of 4*

quantile	form	hold
1	38.75	12.71
2	23.68	14.80
3	17.57	8.98
4	13.49	9.53
5	10.26	8.00
6	7.32	9.22
7	3.81	9.29
8	0.88	7.34
9	-2.54	4.43
10	-6.53	1.76
11	-11.98	0.83
12	-22.25	0.24

(b) *FF49 1 month formation quantiles of 4*

quantile	form	hold
1	156.99	11.48
2	73.59	10.61
3	48.32	9.61
4	32.42	8.92
5	20.34	9.47
6	10.43	5.32
7	-0.50	5.20
8	-8.61	3.64
9	-16.94	3.21
10	-26.42	3.88
11	-37.53	-2.04
12	-58.51	-2.09

(c) *MAA 52 week formation quantiles of 3*

quantile	form	hold
1	27.81	12.65
2	9.76	5.15
3	2.89	2.13
4	-3.62	-1.28
5	-15.00	-1.22

(d) *MAA 4 week formation quantiles of 3*

quantile	form	hold
1	93.90	16.16
2	23.60	6.09
3	5.36	1.03
4	-15.16	-1.62
5	-43.95	-2.66

There is a generally decreasing pattern of holding periods returns (it is more noisy looking at individual assets (unreported) rather than quantiles) from top to bottom of each table. The lower ranks have less extreme returns than the higher ranks, indicating again that, probably, most of the profit from the momentum strategies is from the long positions. For the industry data we see that most of the assets (and quantiles) have positive holding period returns (even if the formation period returns are negative), indicative of a reversal of negative returns as before. Both time-series and cross-sectional momentum would have lost out by shorting these assets. The momentum strategies could possibly be improved with a momentum indicator that does not necessarily predict negative returns from a negative holding period return. I also calculated the average rank correlation between the rank of the assets in the holding period and the formation period and found this to be about 8 - 9% in all cases, which is positive but not extremely high.

[7] Since this is not ideal, I also calculated the average ranked returns for a set of MAA data with exactly ten assets over a similar time period – the results are similar and I do not report them.

[8] For the MAA data with 10 assets I used quantiles of 2 assets.

4.5 Predictive regressions

I check the AR(1) assumptions of section 2.5.1 by regressing holding period returns and absolute returns on returns and absolute returns (respectively) of the formation period. I do the same for deviations of returns from the mean and absolute deviations from the mean. I consider the same formation and holding periods as before.

I run separate robust regressions for each asset in each dataset and report in table 4.7 the average slope, the average t-statistic and the average p-value[9] across all the assets in each dataset. I have scaled the slopes in order to make the them more comparable and interpretable – the 1 month on 12 month slope is scaled by 12, the 1 week on 4 week slope by 4 and the 1 week on 52 week slope by 52. The average slopes for OLS regressions can be found in table E.2 in appendix E.2.

The R-squareds are generally very low (even slightly negative for the robust regression) and many of the relationships are not at all significant. However, the regressions generally indicate the presence of momentum and verify (sometimes only weakly[10]) the assumptions of the models in 2.5.

The average slopes shown are positive and the slopes of individual assets are also generally positive. A few features are of note. For industries the deviation (and absolute deviation) regressions have lower average p-values (and higher average t-stats except for the 1 month on 1 month regression for returns and deviations). For MAA we see the opposite: lower p-values and higher t-statistics for the returns and absolute returns. This agrees with the stronger cross-sectional (time-series) momentum seen in the industry (MAA) data. The (average) slopes of the absolute returns and deviations are larger indicating more predictability here (this is of course not necessarily predictability one can profit from as it does not include the sign of the returns). For the industry (MAA) data we see larger slopes for the return and absolute return (deviation and absolute deviation) regressions which agrees with the stronger cross-sectional (time-series) profits in these datasets.

Similarly comparing the slopes for the long and short formation strategies does not give a clear conclusion. For industries the 1 month on 1 month return and deviation regressions are on average more often significant, which helps explain why this momentum is stronger (however, this is not true for the absolute returns and deviations). For MAA the 1 week on 4 week slopes are have on average lower p-values (and only the deviation regression average t-statistic is not higher), which also supports the stronger momentum results with the 4 week formation.

Table 4.7: ***Average scaled slopes of predictive regressions (robust).*** *Average slopes, t-statistics (in round brackets) and p-values (in square brackets) are reported for per-asset predictive regressions for various metrics (returns and absolute returns, deviations and absolute deviations) in the holding period versus the same metric in the formation period. The 1 month on 12 month, 1 week on 4 week and 1 week on 52 week regressions slopes are scaled by 12, 4, and 52 respectively. Data for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for industries (MAA) are used. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used.*

	metric			
regression	returns	|returns|	deviations	|deviations|
FF49 1 mo vs 12 mo	0.01(0.10)[0.499]	0.18(0.88)[0.426]	0.16(0.94)[0.404]	0.25(1.29)[0.286]
FF49 1 mo vs 1 mo	0.10(1.58)[0.262]	0.05(0.76)[0.479]	0.08(1.24)[0.209]	0.07(1.12)[0.353]
MAA 1 wk vs 52 wk	0.33(1.67)[0.234]	0.69(2.85)[0.058]	0.27(0.79)[0.436]	0.37(0.98)[0.225]
MAA 1 wk vs 4 wk	0.14(2.17)[0.202]	0.22(3.18)[0.008]	0.04(0.46)[0.254]	0.19(1.66)[0.196]

4.6 Global and local time-series momentum

I consider global time-series strategies (i.e. strategies on the equal-weighted market) for three different datasets: the industry and MAA data used until now and also a subset of the MAA data consisting only of 3 bond indices (US, EU, Japan).

The reason for evaluating the difference between global and local time-series momentum is that in light of the decomposition in section 2.7 this may shed light on the relationship between time-series and

[9]Note that this means the reported average p-value does not need to (in fact it won't) correspond to a tail probability for the reported average t-statistic.

[10]Two sets of regressions do have very significant slopes. These are the absolute return regressions for the MAA data.

Table 4.8: ***Empirical decomposition for linear global time-series strategy.*** *Estimates for the decomposition for global linear time-series strategies in section 2.7 are reported from returns from all possible combinations of formation period and holding period. 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period are considered. Returns are used for industry data over the period Jul1969-30Jun1994, all MAA data over the period 4Jan1979-5Dec2002 and a subset of three bond indices from the MAA data over the period 31Dec81-5Dec2002. Figures are in per cent and annualised by multiplying by 12 or 52. Percentage of total profits is indicated in brackets.*

strategy	TS	cross	mean	total
FF49 12 mo	0.00(-0.21)	-0.88(188.08)	0.41(-87.86)	-0.47(100.00)
FF49 1 mo	0.01(2.72)	0.36(91.47)	0.02(5.81)	0.39(100.00)
MAA11 52 wk	0.08(30.12)	0.11(41.45)	0.07(28.43)	0.26(100.00)
MAA11 4 wk	0.01(17.99)	0.05(73.08)	0.01(8.93)	0.06(100.00)
MAA bonds 52 wk	0.08(34.50)	0.07(33.29)	0.07(32.21)	0.22(100.00)
MAA bonds 4 wk	0.02(36.63)	0.03(54.66)	0.01(8.71)	0.06(100.00)

cross-sectional strategies and the sources of their profits.

As noted before the global time-series strategies is essentially an investment in the cross-serial covariances (with the opposite sign to that of a cross-sectional strategy) plus a downweighted version of the time-series return. We would therefore expect the 12 month global time-series strategy for industries to do poorly because here the decomposition is unfavourable for it (the cross-serial part contributes negatively and so does the auto-covariance part). If the global strategy is to fare better we may expect it for the MAA data where there are fewer assets and where cross-sectional did less well (so cross-serial covariances may be more important). The subset of 3 bond indices is chosen because it may exaggerate this: it has even fewer assets and one may expect the bond indices to be closely related (high cross-serial correlations).

I will only examine signed time-series strategies (even though the decomposition really only applies to linear strategies) as we have already noted that linear strategies are not of practical interest.

I report firstly, in table 4.8, the empirical decomposition of global time-series profits as in section 2.7. I denote terms as follows

$$\mathrm{E}[r_t^{GT}] = \underbrace{\frac{\mathrm{E}[r_t^T]}{N}}_{\text{TS}} + \underbrace{\frac{1}{N^2}(\mathbf{1}'\Omega_{t-1,t}\mathbf{1} - \mathrm{tr}(\Omega_{t-1,t}))}_{\text{cross}} + \underbrace{\frac{1}{N^2}(\mathbf{1}'\boldsymbol{\mu}_{t-1}\boldsymbol{\mu}_t'\mathbf{1} - \mathrm{tr}(\boldsymbol{\mu}_{t-1}\boldsymbol{\mu}_t'))}_{\text{mean}} \tag{4.3}$$

This yields no surprises. Except for the 12 month formation industry strategy we see a positive contribution from time-series, larger where there are fewer assets, a positive contribution from cross-serial covariances (this is now the opposite sign to that in section 4.1) and a positive contribution from means. The total profits for the industry and 11 asset MAA datasets are lower than those report in table 4.1 for time-series strategies, which does not look promising for the performance of a global strategy.

I also report the Sharpe ratios of the local and global signed time-series strategies in table 4.9. We see that the 12 month industry profits are positive but weak, which agrees with expectations.

We do, however, also see that (with one exception, where the improvement is marginal) the global time-series strategy does worse. Only the 52 week formation strategy on the bond indices performs marginally better. We note also that for the bond data cross-sectional momentum profits (with 3 quantiles – i.e. buy/sell the top/bottom asset) are positive but weak (the Sharpe ratios are 0.16 (0.15) for the 52 week (4 week) formation).

We must conclude that auto-covariances are the most important part of momentum profits (at least for the data considered here). The negative effect of cross-serial correlations on cross-sectional momentum is not significant enough to consider trying to (implicitly) capture profit from it via global time-series momentum. The latter strategy downweights the most important part of momentum profits and gives more weight to the seemingly inferior profits from cross-sectional covariances. One would need assets with a stronger relationship between them (probably that means higher cross-correlations) in order for global time-series momentum to outperform.

Table 4.9: ***Sharpe ratios for local and global signed time-series strategies.*** *The local strategies are run on (and diversified over) individual assets. The global strategy is a signed strategy run on an equal-weighted market of all assets in the dataset. Returns are used for industry data over the period Jul1969-30Jun1994, all MAA data over the period 4Jan1979-5Dec2002 and a subset of three bond indices from the MAA data over the period 31Dec81-5Dec2002. Sharpe ratios are annualised.*

data and formation	local	global
FF49 12 mo	0.11	0.10
FF49 1 mo	0.49	0.40
MAA 52 wk	1.12	0.48
MAA 4 wk	1.33	0.98
MAA bonds 52 wk	0.90	0.91
MAA bonds 4 wk	1.38	1.19

4.7 Conclusions

- The linear decomposition reveals profits from a link between past and future returns rather than just from means or a dispersion in means (except for the industry 12 month formation time-series strategy). The auto-covariances term is positive and the cross-serial covariance term negative for the industry 1 month, MAA 52 week and MAA 4 week strategies. For the industry 12 month strategies the opposite is the case, which shows momentum can come from different sources.
- The outperformance of cross-sectional strategies in the industry data comes from negative cross-serial correlations in the 12 month formation. For the 1 month formation it seems to come from a tendency for negative returns to reverse. The outperformance is still puzzling as the magnitudes of the returns when the two strategies differ still suggests that time-series should outperform.
- The outperformance of time-series for the MAA data seems to come from a stronger tendency for returns (in particular positive returns) to continue, and from positive cross-serial covariances (lead-lag relationships). When the strategies differ, the time-series strategy outperforms.
- The stronger relationship between time-series and cross-sectional strategies for the MAA data seems to come from a larger probability of the two strategies taking very similar positions, that is from being in a scenario analogous to the up|down case in 2.2.
- With the exception of the industry time-series strategies momentum does appear to have a negative relationship with volatility in terms of prediction accuracy, Sharpe ratio and the ratio of profits to losses.
- A predictive accuracy above 50%, an excess accuracy (compared to the case where returns are independent) and an asymmetry in profits and losses all seem to contribute to momentum profits.
- There is not clear evidence that combining time-series and cross-sectional signals would be useful, but it appears that a cross-sectional strategy could potentially be improved by considering the time-series signal.
- Ranking by returns in the formation period also shows evidence of momentum − higher (lower) formation period returns are followed by higher (lower) holding period returns.
- Predictive regressions (weakly) validate the AR(1) assumptions of section 2.5.1. There is a positive link between past and future returns, absolute returns, deviations and absolute deviations. The relative strength of these relationships seems to agree with the findings of outperformance for cross-sectional (time-series) strategies for industries (MAA) which suggests they do have something to do with the momentum profits.
- There is evidence that profits are concentrated in the long positions and it may in fact be beneficial to avoid short positions.
- Global time-series momentum downweights auto-covariances compared to cross-serial correlations. The latter appear to be a lesser form of profit and thus the global strategy does not do as well as a

local time-series strategy. It appears that cross-serial covariances must be large enough to wipe out cross-sectional profits entirely for the global time-series strategy to outperform.

Chapter 5

Volatility weighting

This chapter empirically examines the efficacy of volatility weighting in improving momentum strategies. As a comparison the effect on a strategy that invests equally in all available assets (the equal-weighted market) is also considered.

There are (at least) two ways of doing volatility weighting. The first is to consider a momentum strategy and scale the entire strategy by an ex-ante measure of its volatility. Some theory on this is seen in section 2.4.1. The second is to run momentum strategies on normalised assets, weighting each asset by its volatility (using normalised returns). Some theory for the latter weighting is seen in section 2.4.2. I consider weighting a strategy in section 5.1 and using normalised returns in section 5.2. I consider also the effect of volatility weighting on skew and kurtosis as looked at theoretically in section 2.4.4.

5.1 Weighting with own volatility

In this section I consider empirically the results of section 2.4.1. I try to assess whether a strategy's return is dependent on its own (ex-ante) volatility and predict whether volatility weighting may be beneficial. I also consider the predictability of volatility for the industry data, where I have daily returns. I then perform volatility weighting on several strategies and report the results.

5.1.1 Relationship with own volatility

I first consider the predictability of volatility for momentum strategies on the industry data.

Following Barroso and Santa-clara (2012) I run an AR(1) regression of the square root of the 21 day realised variances for certain strategies in order to analyse the predictability of the volatility.[1] I do this for the quantile cross-sectional and signed time-series strategies on the industry data as well as a strategy that equally weights all the assets in each month (a proxy for the market). An explanation of the volatility estimates used can be found in appendix C.5. I do not perform similar regressions for the MAA data as there I do not have data at a higher frequency.

In order to be able to calculate the volatility estimates for the industry data, I calculate a daily return series for each strategy. I simply take the portfolio weights of the strategy at the start of each month and then calculate the daily returns over the month using these weights (the portfolio is not rebalanced during the month).[2]

I report the slope coefficients (along with t-statistics in round brackets and p-values in square brackets) and R-squared values for robust AR(1) regressions in table 5.1 (the OLS results are in table E.3 in appendix E.3). The intercepts are all positive and highly significant and thus I do not report them. For the industry data 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% (even the equal-weighted market has very predictable volatility). This is, however, still lower than the coefficient and R-squared reported by Barroso and Santa-clara (2012) in the context of stock momentum.[3]

[1] Barroso and Santa-clara (2012) run their regressions on the variance estimates, not the volatility estimates. I also ran such regressions. The results here (unreported) are weaker and in any case of less interest since it is volatility not variance that we wish to forecast.

[2] There are small discrepancies between the monthly returns and a monthly return series calculated from the daily returns due to rounding in the original FF49 daily and monthly datasets.

[3] This remains true when considering OLS regressions.

Table 5.1: ***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 realised variance from the past 21 daily returns of these strategies. Industry data for the period Jul1969-Jun1994 are used to compute the strategy returns. Robust regressions with a bisquare weighting function with a parameter of 4.685 are used.*

strategy	slope	R-squared
12 mo qxs	0.44(7.54)[0.000]	21.83
1 mo qxs	0.42(7.64)[0.000]	20.88
12 mo sts	0.63(7.89)[0.000]	15.30
1 mo sts	0.38(6.42)[0.000]	18.07
equal-weighted	0.49(6.66)[0.000]	17.36

I now try to assess whether, in the framework of section 2.4.1, a strategy's return is dependent on its own (ex-ante) volatility and predict whether volatility weighting may be beneficial. I perform regressions of (quantile) cross-sectional and (signed) time-series strategies on ex-ante estimates of their volatility based on their own past returns. In particular, for a strategy with returns r_t and volatility estimates σ_t I regress normalised returns on the inverse of volatility.[4]

$$\frac{r_t}{\sigma_t} = \gamma + \alpha \frac{1}{\sigma_t} + \epsilon_t. \tag{5.1}$$

I shall report results for EWMA volatility estimates with an effective history of 61 days or $\lambda = 0.9836$ ($33^1/_3$ weeks or $\lambda = 0.97$) for industry (MAA) data.[5] Details on the volatility estimates are in appendix C.5. I consider quantile cross-sectional and signed time-series strategies with 12 month and 1 month (52 week and 4 week) formation periods for the industry (MAA) data. I also consider a portfolio that equally weights all the available assets in each period (either weekly or monthly) as a proxy for a market portfolio.

In table 5.2 I report the α and γ estimates and associated t-statistics and p-values for robust regressions[6] (OLS regressions are in table E.4 in appendix E.3). I scale the α estimate by 100 merely to ease the presentation. I also report the estimate of the error variance as, if the volatility estimates are adequate, this should be close to 1.

In order to gain more insight into whether volatility weighting may be effective I calculate for the robust regression estimates the range of α over which volatility weighting may be effective (and the strategies have positive profits), as in section 2.4.1. Specifically, for $\hat{\gamma} < 0$ I report

$$[-\hat{\gamma}\hat{\sigma}_t, \Lambda] \tag{5.2}$$

and for $\hat{\gamma} > 0$

$$[\max(-\hat{\gamma}\hat{\sigma}_t, -\Lambda), \Lambda] \tag{5.3}$$

where

$$\Lambda = \sqrt{\frac{(\hat{\gamma}^2 \widehat{\mathrm{Var}(\sigma_t)} + \widehat{\mathrm{E}[\sigma_t^2]})(\widehat{\mathrm{E}[\sigma_t^{-1}]})^2 - 1}{\widehat{\mathrm{Var}(\frac{1}{\sigma_t})}}} \tag{5.4}$$

I estimated the values indicated by symbols with "hats" in the obvious way using the estimated volatility series. I report these figures (as "low" and "high") in table 5.2 as well and scale the estimates by 100 for ease of presentation.

There is a generally negative (but not significant) relationship with ex-ante volatility. Exceptions are the equal-weighted market strategies and for the industry data the 1 month formation time-series strategy. The 52 week cross-sectional strategy on the MAA data also has a positive γ. However, it is the smallest and most insignificant coefficient and is negative with an OLS regression.[7]

[4] I do so in order to be able to use standard regression techniques. Otherwise I must compensate for heteroscedasticity.

[5] I also considered 21 and 126 day RVs (realised variance) for the industry data and 26 week RVs for the MAA data (unreported). These (almost) always gave the same conclusions.

[6] The R-squareds of the regressions range from less than 1% (even slightly negative) to just over 3%, indicating that relationships are not extremely strong.

[7] It is also negative when using a 26 week RV instead.

Table 5.2: ***Strategy relationship with own volatility (robust).*** *Reported are estimates of α (scaled by 100) and γ as in (5.1), the upper and lower bounds of the ranges in proposition 2.2 and the estimated error variance of the regression (5.1) for selected momentum strategies and an equal-weighted market. Return data over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data to estimate strategy returns, the first 21 days (52 weeks) of which are used to provide initial volatility estimates. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used.*

strategy	γ	100α	low	high	error var
FF49 12 mo qxs	-0.46(-1.23)[0.220]	2.15(2.61)[0.010]	1.13	4.50	2.43
FF49 1 mo qxs	-0.08(-0.27)[0.786]	1.05(1.68)[0.094]	0.19	4.01	1.62
FF49 12 mo sts	-0.19(-1.05)[0.296]	0.90(2.62)[0.009]	0.52	4.07	2.03
FF49 1 mo sts	0.39(1.63)[0.103]	-0.44(-0.84)[0.403]	-1.04	4.66	1.79
FF49 equal-weighted	0.46(1.49)[0.138]	-1.19(-1.13)[0.259]	-1.79	7.92	2.03
MAA 52 wk qxs	0.02(0.15)[0.878]	0.26(1.37)[0.172]	-0.03	3.39	0.98
MAA 4 wk qxs	-0.21(-1.76)[0.079]	0.66(3.30)[0.001]	0.39	3.48	1.00
MAA 52 wk sts	-0.15(-1.16)[0.246]	0.24(2.80)[0.005]	0.11	1.27	0.98
MAA 4 wk sts	-0.12(-1.07)[0.283]	0.21(2.88)[0.004]	0.09	1.30	1.00
MAA equal-weighted	0.08(0.55)[0.581]	0.04(0.36)[0.721]	-0.06	1.33	1.06

The α estimates are generally positive and often significant. The α's fall quite comfortably into the ranges over which volatility weighting may work. 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 upper end of the range (Λ) is large compared to α, which provides some empirical justification for Hallerbach (2012)'s approximation. The estimated error variances are too high for the industry data but are quite close to 1 for the MAA data.

5.1.2 Effect of volatility weighting

I run volatility-weighted versions of cross-sectional and time-series momentum strategies as well as the equal-weighted market and compare them to the unweighted versions. The volatility estimates are based on the past returns of the unweighted strategy and are thus estimates of the volatility of the strategy itself. This fits into the framework of section 2.4.1, Hallerbach (2012) and the empirical work in Barroso and Santa-clara (2012). The goal is see whether volatility-weighting improves performance and whether this agrees with both intuition and the predictions of the model in 2.4.1.

I use 61 day and 33⅓ week EWMA volatility estimates for the industry and MAA data respectively (see appendix C.5). The weighted return series is simply $r_t^* = \frac{\varsigma}{\sigma_t} r_t$ where ς is a (conditional) volatility target.[8] I choose a (conditional) volatility target of $\frac{0.1}{\sqrt{12}}$ for the industry data and $\frac{0.1}{\sqrt{52}}$ for the MAA data, which corresponds to an annualised volatility of 10%. 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.

In tables 5.3 (a) and (b) I report descriptive statistics (mean, standard deviation, skew, mean less median, kurtosis, Sharpe ratio, average of the largest 5 drawdowns) for both the weighted and the unweighted strategies.[9] In figures 5.1, 5.2 and 5.3 I plot the standardised logarithm of cumulative returns for the weighted and unweighted strategies, as well as strategies run on normalised returns, which I examine in the next section.

The results show an improvement in the Sharpe ratio whenever the relationship with volatility is negative and a deterioration for the equal-weighted market strategies and industry 1 month time-series strategy, where it was positive. There is also an improvement in the MAA 52 week quantile cross-sectional

[8]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 a conditional mean.

[9]Note that the statistics for the unweighted strategies may differ a little bit from those reported earlier because the first few returns of the sample are not included (because these are used to calculate the first volatility estimate). The MAA statistics differ somewhat (the kurtosis statistics differ a lot) because of the exclusion of a very extreme return at the beginning of the sample.

Table 5.3: ***Performance statistics of volatility weighted and unweighted strategies.*** *Descriptive statistics (as in table 3.3) are calculated for quantile cross-sectional and signed-time series strategies with 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period as well as an equal-weighted market strategy. Industry (MAA) data over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used to calculate the strategy returns and volatility estimates, using the first 21 days (52 weeks) of strategy returns or asset returns for initial volatility estimates. Table (a) reports results for unweighted strategies, starting with the first full monthly return after the first 21 days for industry data and after 52 weeks for the MAA data. In table (b) statistics are reported for strategies weighted with their own volatility. In table (c) statistics are reported for strategies run on normalised returns. Annualisation is as before.*

(a) *unweighted*

strategy	mean	sd	skew	mean less med	kurtosis	Sharpe	ave drawdowns
FF49 qxs 12 mo	11.11	13.65	-0.53	-2.11	1.76	0.77	-5.12
FF49 qxs 1 mo	10.91	10.17	0.00	1.20	0.34	1.02	-4.22
FF49 sts 12 mo	1.59	13.91	-1.34	-6.64	6.89	0.11	-8.77
FF49 sts 1 mo	6.28	12.49	-0.50	1.06	2.60	0.49	-5.37
FF49 equal-weighted	5.56	18.32	-0.39	0.06	2.40	0.30	-5.75
MAA qxs 52 wk	14.42	13.64	-0.42	-4.15	1.85	0.99	-11.42
MAA qxs 4 wk	15.73	13.72	0.16	-1.87	5.66	1.07	-9.21
MAA sts 52 wk	5.96	5.18	-0.43	-2.09	2.60	1.12	-10.56
MAA sts 4 wk	7.25	5.26	0.29	0.63	6.15	1.33	-8.30
MAA equal-weighted	3.65	5.67	-0.31	-2.23	2.31	0.63	-17.27

(b) *weighted with own volatility*

strategy	mean	sd	skew	mean less med	kurtosis	Sharpe	ave drawdowns
FF49 qxs 12 mo	15.86	15.75	-0.34	0.58	0.45	0.94	-4.70
FF49 qxs 1 mo	15.27	12.78	0.09	2.86	-0.20	1.12	-3.77
FF49 sts 12 mo	6.01	14.45	-0.94	-5.11	4.80	0.40	-5.82
FF49 sts 1 mo	6.50	13.37	-0.26	-0.07	1.62	0.47	-5.66
FF49 equal-weighted	2.91	14.24	-0.59	-1.26	2.48	0.20	-4.72
MAA qxs 52 wk	11.80	9.90	-0.10	-2.56	1.22	1.13	-11.12
MAA qxs 4 wk	13.57	10.07	0.41	-0.52	4.38	1.27	-8.24
MAA sts 52 wk	13.56	9.91	-0.18	-2.53	1.85	1.28	-9.42
MAA sts 4 wk	16.60	10.07	0.49	2.30	4.49	1.53	-7.20
MAA equal-weighted	6.22	10.28	-0.46	-4.16	1.81	0.59	-16.91

(c) *weighted with underlying volatility (normalised returns)*

strategy	mean	sd	skew	mean less med	kurtosis	Sharpe	ave drawdowns
FF49 qxs 12 mo	7.60	6.59	-0.32	-1.63	0.18	1.12	-4.14
FF49 qxs 1 mo	7.34	4.97	0.07	-0.46	-0.15	1.43	-3.10
FF49 sts 12 mo	1.35	7.91	-1.41	-3.06	7.97	0.17	-6.67
FF49 sts 1 mo	3.77	7.06	-0.10	0.70	1.28	0.52	-5.53
FF49 equal-weighted	2.82	10.78	-0.62	-0.93	2.77	0.26	-6.87
MAA qxs 52 wk	2.98	4.35	-0.23	-0.37	1.01	1.41	-9.71
MAA qxs 4 wk	3.87	4.31	0.34	0.11	2.62	1.84	-6.84
MAA sts 52 wk	1.34	1.82	-0.14	-0.07	1.75	1.52	-11.02
MAA sts 4 wk	1.75	1.82	0.46	0.25	4.60	1.99	-5.97
MAA equal-weighted	0.76	1.84	-0.44	-0.47	1.71	0.86	-18.86

strategy.[10] The graphs support this. There is a widening gap between the weighted and unweighted strategies wherever the Sharpe ratio improved and clear deterioration for the industry market strategy which showed quite a large drop in the Sharpe ratio.

The pattern agrees with intuition: you want to invest less when the return is likely to be less, when volatility is high in this case. However, the model does not predict that this will necessarily be the case. It is useful to note that the results do provide some evidence that volatility weighting is not a panacea. It does not necessarily improve a strategy. It is also interesting to note that the weighting increases the Sharpe ratio of the 12 month formation time-series strategy on the industry data above the Sharpe ratio of the market (and despite this strategy having showed a positive relationship with volatility in section 4.3), whereas the unweighted Sharpe ratio is well below it. The standard deviation estimates in table 5.3 (b) are of note. The MAA strategy standard deviations are quite close to the 10% target, but the industry data estimates are not. There are two possible effects that could cause this: firstly the volatility estimate for the industry data may be less effective, not capturing as much of the volatility in the next period, secondly the effect of a non-zero conditional mean may be larger.[11]

Volatility weighting also generally results in a more positive (or less negative) skew, a lower kurtosis and also lower normalised drawdowns, indicating a benefit of volatility weighting beyond improving the Sharpe ratio. A lower kurtosis would naturally result from stabilising volatility. See section 2.4.4 for some theory that predicts volatility weighting should result in less extreme skew and a lower kurtosis under somewhat restrictive assumptions. We do see a lower kurtosis mostly, but an increase in positive skew is not predicted.

I also regress the weighted strategies on the unweighted strategies in order to statistically test for outperformance (alpha). I report intercepts of robust regressions in table 5.4 (a) (the slopes are all positive and very significant). The OLS regression intercepts are in table E.5 (a) in appendix E.3. For these regressions I scaled the returns by 12 and 52 (respectively, for the industry and MAA data) and expressed them in per cent in order to make the alphas easier to interpret.

Table 5.4: ***Intercepts of regressions of volatility weighted strategies on unweighted strategies (robust).*** *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 table (a) the strategies are weighted with their own volatility and in table (b) normalised returns are used. Return data over the period Jul1969-30Jun1994 (4Jan1979-5Dec2002) for industries (MAA) is used to obtain strategy returns and volatility estimates.*

(a) *weighted with own volatility vs unweighted*

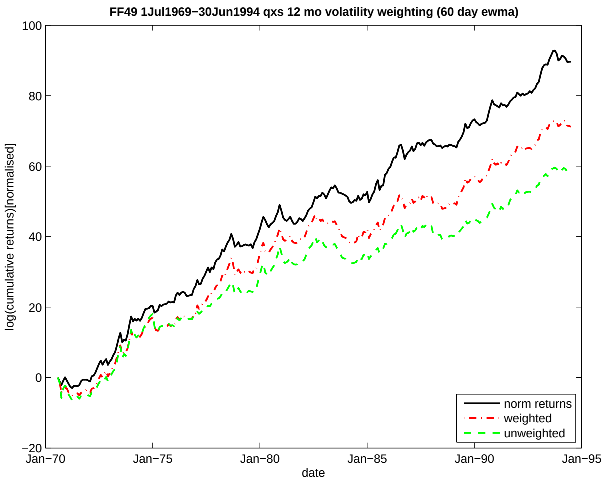
industries		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	1.16(1.06)[0.290]	MAA qxs 52 wk	1.35(2.12)[0.034]
FF49 qxs 1 mo	0.41(0.53)[0.595]	MAA qxs 4 wk	0.94(1.42)[0.156]
FF49 sts 12 mo	1.77(1.48)[0.139]	MAA sts 52 wk	1.41(2.44)[0.015]
FF49 sts 1 mo	-0.56(-0.51)[0.612]	MAA sts 4 wk	0.88(1.29)[0.197]
FF49 equal-weighted	-0.99(-1.19)[0.233]	MAA equal-weighted	0.51(0.93)[0.353]

(b) *normalised returns vs unweighted*

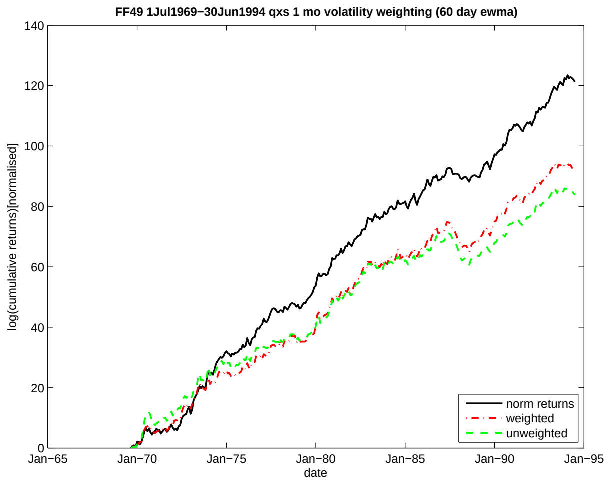
industries		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	2.32(3.10)[0.002]	MAA qxs 52 wk	6.33(4.59)[0.000]
FF49 qxs 1 mo	3.06(4.74)[0.000]	MAA qxs 4 wk	8.39(6.39)[0.000]
FF49 sts 12 mo	-0.23(-0.48)[0.633]	MAA sts 52 wk	2.08(4.87)[0.000]
FF49 sts 1 mo	-0.25(-0.54)[0.588]	MAA sts 4 wk	2.67(6.78)[0.000]
FF49 equal-weighted	-0.06(-0.11)[0.909]	MAA equal-weighted	1.16(3.11)[0.002]

[10] This strategy arguably also has a weakly negative relationship with volatility despite the positive estimate reported as the OLS regression estimate and the estimates with a 26 week RV were negative.

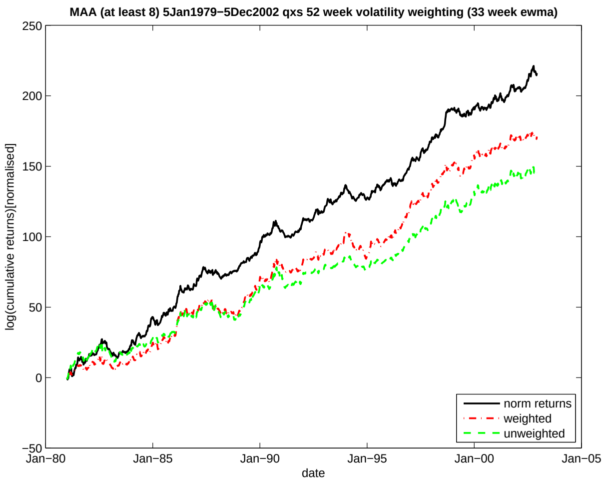
[11] The latter effect should be larger for the less frequent monthly data (volatility dominates with higher frequencies), but probably is not so large as to account for the large differences in the ex-post standard deviations.



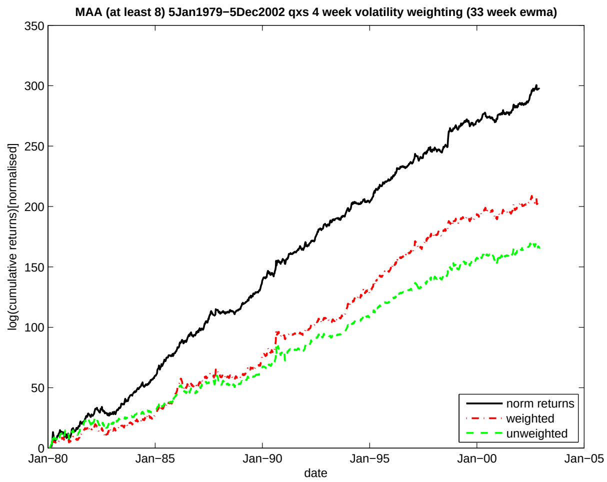
(a) *Industry 12 month formation*



(b) *Industry 1 month formation*

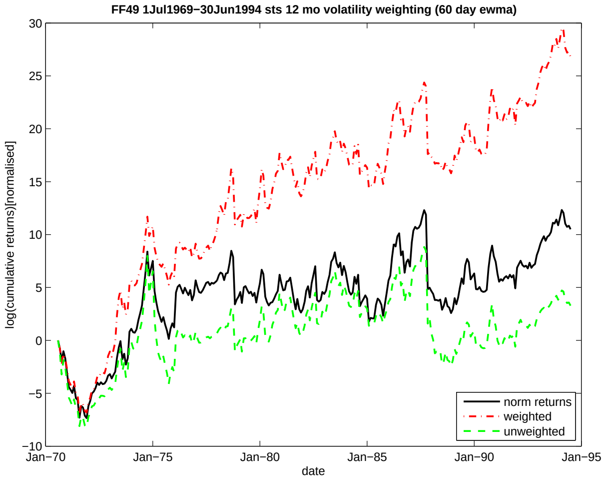


(c) *MAA 52 week formation*

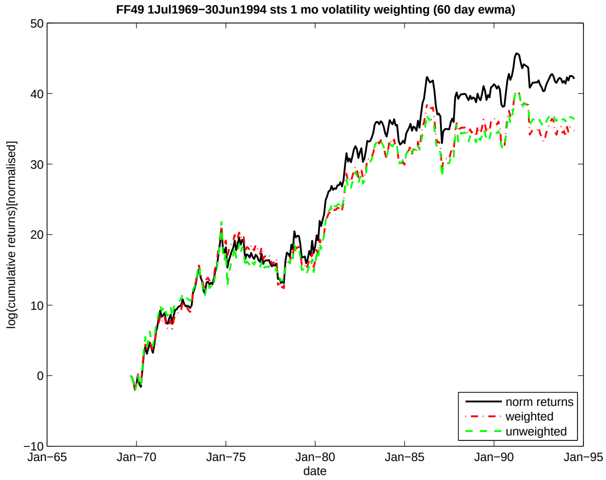


(d) *MAA 4 week formation*

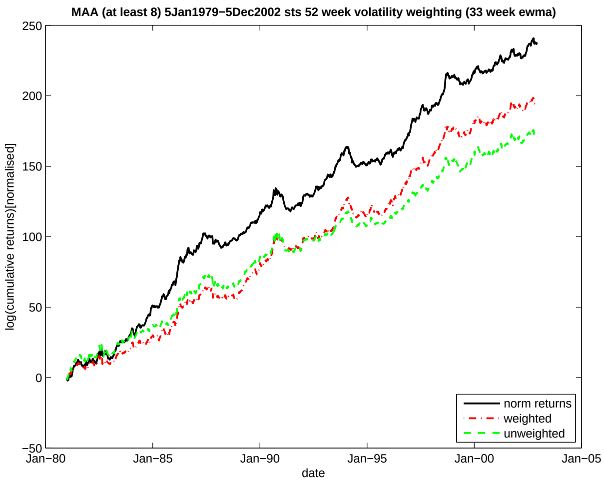
Figure 5.1: ***Cumulative performance graphs of volatility-weighted strategies: cross-sectional.*** *The standardised logarithm of cumulative returns is plotted for quantile cross-sectional strategies weighted with their own volatility and run on normalised returns along with an unweighted strategy. Asset returns over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for the industry (MAA) data are used to calculate strategy returns and volatility estimates for 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period. The series for each strategy is standardised by dividing by the ex-post standard deviation of log returns.*



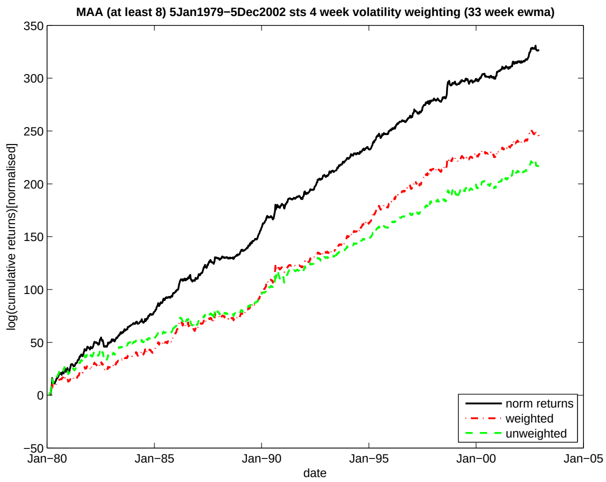
(a) *Industry 12 month formation*



(b) *Industry 4 week formation*



(c) *MAA 52 week formation*



(d) *MAA 4 week formation*

Figure 5.2: ***Cumulative performance graphs of volatility-weighted strategies: time-series.*** *The standardised logarithm of cumulative returns is plotted for signed time-series strategies weighted with their own volatility and run on normalised returns along with an unweighted strategy. Asset returns over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for the industry (MAA) data are used to calculate strategy returns and volatility estimates for 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period. The series for each strategy is standardised by dividing by the ex-post standard deviation of log returns.*

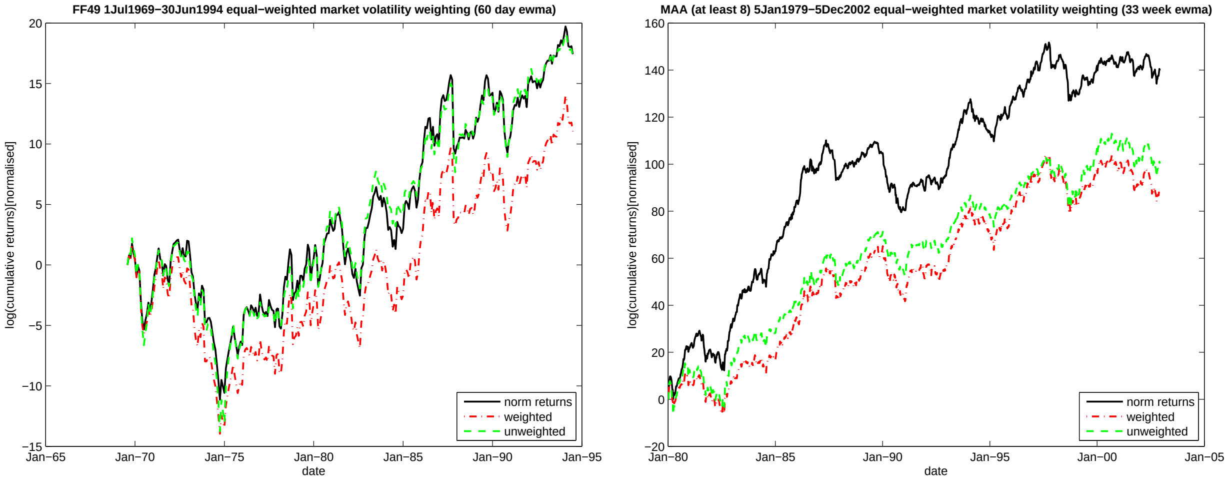


Figure 5.3: ***Cumulative performance graphs of volatility weighted strategies: equal-weighted market.*** *The standardised logarithm of cumulative returns is plotted for (equal-weighted) market strategies weighted with their own volatility and run on normalised returns along with an unweighted strategy. Asset returns over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for the industry (MAA) data are used to calculate strategy returns and volatility estimates. The series for each strategy is standardised by dividing by the ex-post standard deviation of log returns.*

The alphas (generally) agree with the conclusions from the Sharpe ratios. An exception is the MAA equal-weighted market (the alpha is positive), but the estimated alpha is small and insignificant and is the correct sign for the OLS regression. The robust regression alphas are generally not significant (the OLS estimates often are).

To gain some further insight into why volatility weighting works (or not) and test the models in section 2.4.1 I calculate

$$n = \frac{\widehat{\mathrm{E}r_t^*}}{\widehat{\varsigma\mathrm{E}[\frac{1}{\sigma_t}]}} \tag{5.5}$$

which I then annualise. This is an estimate of $\alpha + \gamma\mathrm{E}[\frac{1}{\sigma_t}]^{-1}$ in the numerator of the Sharpe ratio (2.13) which avoids using the estimated coefficients (I do not report these figures). One would expect this to be greater than the unweighted mean return where the dependence on volatility is negative and the opposite where it is positive. In particular, if this is so in the latter case it would help to explain a weak or negative performance of volatility weighting despite the α being in the appropriate range.

Disappointingly, the estimates of the numerator of the Sharpe ratio are very often contrary to expectations. In fact, the mean is almost always lower than for the unweighted strategy. It is larger for the two MAA cross-sectional strategies, which the model predicts, and for the industry market strategy, which is not what it predicts. The contrary estimates may be as a result of the inadequacy of the volatility estimates[12] or a non-stationary volatility distribution. There may also be something not captured by the model.

5.2 Weighting with underlying volatility

I consider the effect of normalising all the assets under consideration by an ex-ante measure of their volatility (using normalised returns) as examined theoretically in section 2.4.2. This is more closely associated with the work in Hallerbach (2011) and this is also what Moskowitz *et al.* (2012) do for their time-series strategies. I also examine more closely the differences between normalised and unweighted returns that may account for any effect of volatility weighting.

[12]The estimates are also not independent which makes them less than ideal for estimating moments.

5.2.1 Effect of volatility weighting

I now run momentum strategies on normalised asset returns,[13] i.e. weighting each underlying asset by its own (ex-ante) volatility rather than weighting the strategy by its (ex-ante) volatility. I consider again time-series and cross-sectional strategies as well as an equally-weighted market proxy.

The construction of the new strategies is simple. I compute EWMA volatility estimates (using the same parameters as before) for each underlying asset and create a normalised return series $\frac{\varsigma}{\sigma_{i,t}} r_{i,t}$. I then run the strategies on these normalised returns. For instance for cross-sectional momentum both the ranking and investment are then based on normalised returns.

I choose a (conditional) volatility target of $\frac{0.1}{\sqrt{12}}$ for the industry data and $\frac{0.1}{\sqrt{52}}$ for the MAA data, which corresponds to an annualised volatility of 10%. Note, however, that these are targets for the underyling assets, not the strategies (we observe the strategies have a lower standard deviation). The results in 2.4.3 suggest that using the same volatility target for all assets is preferable. From this theory we can expect normalising returns to benefit the strategy by equalising volatility within and across assets. This form of weighting would, however, also be effective if there is a negative relationship with market volatility. 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).

I calculate descriptive statistics as before and report these in table 5.3 (c). Cumulative performance graphs are in figures 5.1, 5.2 and 5.3.

Volatility weighting improves the Sharpe ratio in almost all cases and appears to be most effective for the cross-sectional strategies. The single exception is the industry equal-weighted market strategy. The increase in Sharpe ratio is larger than for weighting by the strategy's own volatility except for the 12 month formation time-series strategy on the industry data. The cumulative performance graphs also reflect the effect on Sharpe ratios, with a widening gap between the normalised return and unweighted strategies where the Sharpe ratio improved. This gap is, however, less impressive for the industry time-series strategies. Recall that these betrayed a positive relationship with volatility in section 4.3. The volatility weighting also usually improves skewness, kurtosis and drawdowns as well. One should resist the urge to compare the standard deviations of the tables 5.3 (b) and (c). For the former we had a target for the volatility. For the latter, however, the underlying assets have been normalised and I have not made any predictions for the effect of this on standard deviation of the momentum strategy based on these returns.

I again regress the strategies with normalised returns on the unweighted strategies in order to statistically test for outperformance (alpha). The robust intercepts are in table 5.4 (b). The OLS regression results are in table E.5 (b) in appendix E.3. For the regressions I scaled the returns by 12 (52) for the industry (MAA) data and expressed them in per cent in order to make the alphas easier to interpret.

The regressions show a significant alpha for all the MAA strategies in agreement with the effect on the Sharpe ratios. So do the industry cross-sectional strategies. However, the industry time-series strategies show a weak negative (weakly positive under OLS) alpha and are thus the wrong sign. The industry equal-weighted strategy has a very weak negative alpha. The cross-sectional strategies (for both the MAA and industry data) have somewhat larger and more significant alphas. It is perhaps not surprising that the cross-sectional strategies are most improved as these invest only in the most extreme assets. Using normalised returns can result in a large change in the composition of the long/short portfolio. The volatility weighting is also clearly more effective for the MAA data, possibly because there are very large differences in volatility across these assets. The weaker effect of volatility weighting for the industry time-series strategies seems to agree with anomalous results for these strategies in section 4.3 where the performance depended positively on volatility.

The observation that using normalised returns for the equal-weighted strategies is effective for the MAA data, but not for the industry data is consistent with the finding of Thomas *et al.* (2012) that volatility-weighting is effective across but not within asset classes.

Moskowitz *et al.* (2012) found an alpha for their time-series strategies (except for equity momentum). However, as it appears their cross-sectional momentum strategy included no volatility weighting the conclusion was perhaps premature – they were not comparing like with like.

[13] At least for the quantile cross-sectional strategy one may also wish to consider ranking by normalised returns but investing in the unweighted assets. One may wish to do so to reduce turnover, to avoid going net short or to avoid blow-ups when volatility is very low. However, the risk of the strategy will still be dominated by the most volatile assets that enter the long-short portfolios.

5.2.2 Digging deeper

I consider here in more detail what the effect of using normalised returns is on the behaviour of momentum, including the sources of profit in chapter 4. I repeat many of the results for the unweighted returns for ease of comparison.

In the first instance one can compare the holdings of the strategies on normalised and unweighted returns. One would expect the unweighted strategies to be more heavily concentrated in the volatile assets which would tend to have the most extreme returns. This is so particularly for the cross-sectional strategy where these stocks will be in the top/bottom quantiles. Intuitively one would expect the normalised return strategies to downweight the more volatile assets and upweight the less volatile assets (even for the time-series strategy) and this is in fact the case. I calculate the change in average gross exposure to each asset in each dataset, ranked by volatility and find a downweighting of the most volatile assets as expected.[14] As these results are entirely as expected I do not report them.

One may also wish to consider whether normalising returns changes the conclusions we saw in chapter 3. As such I calculate again Sharpe ratios for the six strategies varying the formation and holding period as section 3.2. These new results are in tables F.1 (a) and (b) in appendix F. Most of the earlier conclusions remain unchanged. We still see that shorter holding periods are good and the strongest momentum is for the 1 month (4 week) formation for industries (MAA). Cross-sectional is stronger for industries and time-series for MAA. Scaling is still helpful for performance and at least for the MAA data the linear strategies fare worse (for industries the scaled strategies perform the best and the unscaled do not fare much worse). Note that for the MAA data the unscaled linear strategies still have very high kurtosis (unreported) but far less extreme than before and are thus still practically infeasible.

I also decompose the linear momentum profits as in section 4.1, now for the normalised returns. I report these and the earlier decomposition for unweighted returns in table 5.5. The basic pattern of the decomposition for the normalised returns is mostly unchanged. The only real difference is for the 12 month formation for industries where we now see a positive contribution from autocorrelations (however, the performance improvement for these strategies was not greater than for the others).

I also perform a long-short analysis as in section 4.2 and report the three sets of tables for each of the four strategies in table F.2 in appendix F. We reach similar conclusions. Most of the time the strategies still take the same positions and this is more so for the MAA data. The relative proportions of assets in the long and short portfolio tables and the frequency of the up|up, up|down and down|down cases are virtually unchanged. The proportions of the four scenarios are also more or less the same. The outperformance of the cross-sectional strategies for industries still seems to come from the short|long portfolio and the down|down case, particularly scenario 4. The MAA data still show a greater tendency for return continuation and the time-series strategy outperforms whenever the two strategies differ. In particular most of the outperformance seems to come from scenario 1 in the up|up case. We also still see that long positions outperform short positions and that a cross-sectional strategy may be improved by taking into account the time-series signal.

I further consider the effect of normalised returns on a prediction analysis as in section 4.3. Most of the conclusions drawn earlier remain unchanged.

In tables F.3 and F.4 in appendix F I report the prediction accuracy and excess accuracy (respectively) of signed cross-sectional strategies on normalised returns and unweighted returns. I do not report time-series strategies as here the prediction accuracy is not affected by normalising. The division into low and high volatility states is as before and is based on the unweighted returns in all cases. Normalising has a seemingly negligible (but positive overall) impact on the prediction accuracy of the cross-sectional strategies, which empirically validates the corresponding assumption in section 2.4.2. Normalising also does not change the impact of volatility – accuracy remains lower in the high volatility states. We do see a somewhat higher and more significant excess accuracy for the normalised returns. This (and the higher prediction accuracy above) gives some indication (but weak) of stronger cross-sectional momentum in the normalised returns in this sense. The patterns for the Sharpe ratios of signed strategies (table F.5) and profits over losses (table F.6) seem to remain unchanged. We do see at least a higher Sharpe ratio and profit to loss ratio overall for the normalised returns.

In table 5.6 I report the Sharpe ratios when the time-series and cross-sectional signals agree and

[14] I consider for each strategy the amounts invested in the (unweighted) asset at the start of each period as a proportion of the notional capital available. I then take the absolute value in order to get a gross exposure to each asset and calculate the average over time for each asset. I then take the vector of these averages and normalise it (i.e. express each figure as a proportion of the total). I do this because for the strategies with normalised returns the gross exposure to the unweighted assets is not constant. I then subtract the unweighted strategy vector from the normalised strategy vector, giving the change in proportion of exposure to each asset.

Table 5.5: ***Empirical decomposition for linear strategies (unweighted and normalised returns).*** *Estimates for the decomposition for linear strategies in section 2.3.1 are reported from returns from all possible combinations of formation period and holding period. 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period are considered for industry (MAA) data over the period Jul1969-Jun1994 (31 Dec 81 to 5Dec2002). In table (a) the returns are unweighted and in table (b) they are normalised with the first 21 days (52 weeks) used to calculate the initial volatility estimates. For the MAA data a set of 11 assets (for which data is available over the whole period) is considered. Figures are in per cent and annualised by multiplying by 12 (52). Percentage of total profits is indicated in brackets.*

(a) *Unweighted returns*

strategy	auto	cross	mean	total
FF49 12 mo XS	-0.434(-84.0)	0.880(170.3)	0.071(13.7)	0.517(100)
FF49 12 mo TS	-0.443(-905.6)	-	0.492(1005.6)	0.049(100)
FF49 1 mo XS	0.485(372.1)	-0.361(-276.6)	0.006(4.5)	0.130(100)
FF49 1 mo TS	0.495(94.4)	-	0.029(5.6)	0.525(100)
MAA11 52 wk XS	0.503(82.5)	-0.109(-17.9)	0.216(35.4)	0.610(100)
MAA11 52 wk TS	0.553(63.4)	-	0.320(36.6)	0.873(100)
MAA11 4wk XS	0.092(150.3)	-0.046(-74.7)	0.015(24.4)	0.061(100)
MAA11 4wk TS	0.101(81.8)	-	0.023(18.2)	0.124(100)

(b) *Normalised returns*

strategy	auto	cross	mean	total
FF49 12 mo XS	0.184(96.2)	-0.032(-16.9)	0.040(20.7)	0.191(100)
FF49 12 mo TS	0.188(42.1)	-	0.259(57.9)	0.447(100)
FF49 1 mo XS	0.168(580.5)	-0.142(-491.1)	0.003(10.6)	0.029(100)
FF49 1 mo TS	0.172(89.2)	-	0.021(10.8)	0.193(100)
MAA11 52 wk XS	0.571(90.6)	-0.158(-25.0)	0.217(34.4)	0.631(100)
MAA11 52 wk TS	0.629(65.1)	-	0.337(34.9)	0.965(100)
MAA11 4 wk XS	0.170(118.7)	-0.043(-30.1)	0.016(11.3)	0.143(100)
MAA11 4 wk TS	0.187(87.19)	-	0.026(12.1)	0.213(100)

disagree. It seems most of the improvement from volatility weighting occurs where the signals agree. In fact, all strategies now exhibit stronger returns when they agree (overall) than when they disagree. Here there does seem to be scope for taking into account both signals (there is only exception to this – the industry 1 month cross sectional strategies with a positive signal).

I also look at ranked returns as in section 4.4 but with normalised returns. These results are in table F.7 in appendix F. One can draw the same conclusions as before. The positive returns are still more extreme than the negative ones. However, this appears to be less pronounced. Rank correlations are still positive and of a similar magnitude (they are a little higher in fact which does suggest a slightly stronger momentum in the normalised returns[15]).

In order to further clarify the source of the improved returns from volatility weighting I ran predictive regressions such as in section 4.5 for normalised returns and report the average slopes in table 5.7 (the OLS slopes are in table E.2b in appendix E.2). One may expect from the improved performance of the volatility weighted strategies to find stronger evidence of momentum in these regressions. However, this is not the case. Results are similar or weaker. This suggests that the improved performance of the strategies based on normalised returns are in fact due to a stabilising of volatility or volatility timing, rather than stronger momentum in normalised returns. In fact the absolute return regressions for industry 1 month (time-series) momentum are now even negative. This strategy's Sharpe ratio improved only marginally, however, from using normalised returns and the alpha was weakly negative (weakly positive) under robust (OLS) regressions. There is another interpretation for weaker regressions for the absolute returns and deviations. This in fact indicates weaker volatility clustering, i.e. a more stable volatility, which could be a source of improved risk-adjusted performance.

A linear decomposition for a global time-series strategy with normalised returns and a comparison of the Sharpe ratios of global and local strategies reveals the same conclusions as earlier (the global strategies do not perform better). I report the decomposition and the Sharpe ratios in tables F.8 and F.9 in appendix F.

5.3 Conclusions

- Weighting strategies by their own volatility seems to be effective at least when the relationship with volatility is negative.
- The part of returns independent of volatility is always small enough that one may expect volatility weighting to be beneficial.
- Momentum strategies mostly have a negative relationship with their volatility (the industry 1 month time-series strategies do not).
- Running strategies on normalised returns is apparently even more effective than weighting a strategy by its own volatility, particularly for the cross-sectional strategies (which naturally invest only in the most extreme assets).
- Both forms of volatility weighting considered generally increase skew (more positive, less negative) and decrease kurtosis, indicating additional benefits of volatility weighting.
- A part of the improved return from running strategies on normalised returns does seem to follow from investing less in the most extreme assets.
- The improved performance from normalised returns does not seem to follow from stronger momentum in these returns - in fact it is weaker (although predictive accuracy is a little higher for the cross-sectional strategies and rank correlations are also higher). One concludes that it follows from stabilising volatility, including possibly a reduction in volatility clustering, and/or from volatility timing.
- Improved performance of normalised returns is mostly from when both the cross-sectional and time-series signals agree.
- Normalising returns does not, however, seem to give a different relationship between time-series and cross-sectional strategies. The conclusions from chapter 3 and 4 for unweighted returns remain

[15] Rank correlations for industries (MAA) increase from 8.45 and 8.93 (8.26 and 8.75) to 9.15 and 9.9 (8.8 and 9.91) for the long and short formation respectively.

Table 5.6: ***Prediction analysis for signed strategies (agreeing and opposing signal Sharpe ratios) with unweighted and normalised returns .*** *Annualised Sharpe ratios for signed time-series and cross-sectional strategies (based on aggregated returns across all assets) are reported for cases where the time-series and cross-sectional signals agree and disagree. The time-series (cross-sectional) signal for an asset is the sign of the formation period return (deviation). Data for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for industry (MAA) data. In tables (a) and (b) the results for unweighted returns are repeated. Tables (c) and (d) contain the results for normalised returns where the first 21 days (52 weeks) of asset returns are used for the initial volatility estimates.*

(a) *Opposing signals – unweighted returns*

	TS neg XS pos		TS pos XS neg		all	
strategy	TS	XS	TS	XS	TS	XS
FF49 12 mo	-0.46	0.37	0.39	0.16	0.02	0.25
FF49 1 mo	-0.26	0.39	0.70	0.09	0.17	0.24
MAA 52 wk	0.18	-0.35	0.38	0.14	0.33	0.01
MAA 4 wk	0.89	-0.37	0.46	0.30	0.63	0.03

(b) *Agreeing signals – unweighted returns*

	TS neg XS neg		TS pos XS pos		all	
strategy	TS	XS	TS	XS	TS	XS
FF49 12 mo	0.04	0.22	0.34	0.15	0.19	0.18
FF49 1 mo	0.16	0.29	0.46	0.17	0.30	0.23
MAA 52 wk	0.19	0.35	0.62	0.41	0.40	0.38
MAA 4 wk	0.18	0.37	0.73	0.48	0.45	0.42

(c) *Opposing signals – normalised returns*

	TS neg XS pos		TS pos XS neg		all	
strategy	TS	XS	TS	XS	TS	XS
FF49 12 mo	-0.35	0.17	0.00	0.09	-0.11	0.11
FF49 1 mo	-0.26	0.39	0.59	0.14	0.13	0.26
MAA 52 wk	0.27	-0.19	0.30	0.22	0.29	0.11
MAA 4 wk	0.58	-0.04	0.36	0.52	0.44	0.31

(d) *Agreeing signals – normalised returns*

	TS neg XS neg		TS pos XS pos		all	
strategy	TS	XS	TS	XS	TS	XS
FF49 12 mo	-0.16	0.39	0.46	0.32	0.19	0.35
FF49 1 mo	0.21	0.40	0.48	0.29	0.34	0.35
MAA 52 wk	0.33	0.54	0.82	0.54	0.59	0.54
MAA 4 wk	0.44	0.63	1.05	0.70	0.74	0.66

Table 5.7: ***Average scaled slopes of predictive regressions with normalised and unweighted returns (robust).*** *Average slopes, t-statistics (in round brackets) and p-values (in square brackets) are reported for per-asset predictive regressions for various metrics (returns and absolute returns, deviations and absolute deviations) in the holding period versus the same metric in the formation period. The 1 month on 12 month, 1 week on 4 week and 1 week on 52 week regressions slopes are scaled by 12, 4, and 52 respectively. Data for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for industries (MAA) are used. Table (a) repeats the results for unweighted returns and table (b) is based on normalised returns where the first 21 days (52 weeks) of asset returns have been used for the initial volatility estimates. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used.*

(a) *Unweighted returns*

	metric			
regression	returns	|returns|	deviations	|deviations|
FF49 1 mo vs 12 mo	0.01(0.10)[0.499]	0.18(0.88)[0.426]	0.16(0.94)[0.404]	0.25(1.29)[0.286]
FF49 1 mo vs 1 mo	0.10(1.58)[0.262]	0.05(0.76)[0.479]	0.08(1.24)[0.209]	0.07(1.12)[0.353]
MAA 1 wk vs 52 wk	0.33(1.67)[0.234]	0.69(2.85)[0.058]	0.27(0.79)[0.436]	0.37(0.98)[0.225]
MAA 1 wk vs 4 wk	0.14(2.17)[0.202]	0.22(3.18)[0.008]	0.04(0.46)[0.254]	0.19(1.66)[0.196]

(b) *Normalised returns*

	metric			
regression	returns	|returns|	deviations	|deviations|
FF49 1 mo vs 12 mo	0.05(0.31)[0.534]	0.09(0.45)[0.505]	0.22(1.30)[0.308]	0.21(1.18)[0.386]
FF49 1 mo vs 1 mo	0.08(1.31)[0.275]	-0.02(-0.40)[0.617]	0.09(1.48)[0.232]	0.03(0.52)[0.504]
MAA 1 wk vs 52 wk	0.35(1.95)[0.198]	0.48(1.50)[0.290]	0.25(0.77)[0.435]	0.41(1.05)[0.335]
MAA 1 wk vs 4 wk	0.15(2.43)[0.152]	0.11(1.68)[0.186]	0.06(0.61)[0.278]	0.11(0.91)[0.395]

largely unchanged for normalised returns. However we now see more evidence that combining the cross-sectional and time-series signals may be useful.

Chapter 6

Relationship with dispersion and volatility

In this chapter I consider the relationship of cross-sectional momentum with dispersion, including the efficacy of dispersion weighting, and of time-series momentum with similar measures as in section 2.5. I also study briefly the relationship of momentum with market volatility and the relationship between dispersion and volatility, including how this is affected by normalising returns.

In section 6.1 I regress momentum strategies on dispersion measures and test the predictions of the models in section 2.5.1. In section 6.2 I regress momentum strategies on market volatility. In section 6.3 I study the effect of dispersion weighting and in section 6.4 I consider the relationship between volatility and dispersion, also considering the effect of normalising returns.

6.1 Regressions on dispersion measures

I run regressions of the various momentum strategy variations on 5 lagged measures,

(1) the average of absolute returns in the formation period: $\frac{1}{N_t}\sum_{i\in\mathcal{N}_t}|r_{i,t-j,t}|$

(2) the average of squared returns in the formation period: $\frac{1}{N_t}\sum_{i\in\mathcal{N}_t}(r_{i,t-j,t})^2$

(3) the average of absolute deviations from the (equal weighted) market: $\frac{1}{N_t}\sum_{i\in\mathcal{N}_t}|\tilde{r}_{i,t-j,t}|$

(4) the average of squared deviations from the (equal weighted) market: $\frac{1}{N_t}\sum_{i\in\mathcal{N}_t}(\tilde{r}_{i,t-j,t})^2$

(5) the cross-sectional standard deviation (which I denote by xsd for short): $\sqrt{\frac{1}{N_t}\sum_{i\in\mathcal{N}_t}(\tilde{r}_{i,t-j,t})^2}$

Here tildes indicate deviations from the average return, $\mathcal{N}_t$ is the set of assets available at time t and N_t is the number of assets available. The first four measures are suggested by my theoretical decomposition in section 2.5.1 and the last measure is the typical way of defining cross-sectional dispersion. It is studied for instance in Stivers and Sun (2010), but for individual stock momentum (and they use a different set of portfolios for calculating the dispersion). Only the last three measures are actually measures of dispersion. The first two measures could be seen as dispersion away from zero. One would expect a positive relation of time-series momentum with the first two factors and a positive relation of cross-sectional momentum with the last three factors according to the models in section 2.5. However, we have already noted that empirically the relationship with cross-sectional standard deviation is negative.

I report the slopes (with t-value and p-value) of robust regressions in table 6.1 (OLS regressions can be found in table E.6 in appendix E.4).

The R-squareds are generally very low (even slightly negative for some of the robust regressions) and many of the relationships are not at all significant. Like Stivers and Sun (2010) I find a negative rather than positive relationship for cross-sectional strategies with dispersion. It is in fact a bit of a puzzle that this link is negative as the models do not seem *prima facie* implausible and the predictive regressions in section 4.5 did not give such contradictory results. Possibly this is because of the close relationship between volatility and dispersion – I examine this in section 6.4.

The evidence for the AR(1) models is weak (for the time-series model) at best and at worst contradictory (for the cross-sectional models). We must conclude that ex-ante dispersion measures cannot explain momentum and that there is more at work. The negative relationship with cross-sectional dispersion rings true if dispersion is in fact a countercyclical indicator as proposed by Stivers and Sun (2010).

One set of regressions in table 6.1 gives nicely the results expected from the models in section 2.5.1. These are the industry strategies with a 1 month formation regressed on the absolute and squared returns. We see significant positive relationships with the time-series strategies. The other unscaled linear time-series strategies also have a positive relationship with the absolute and squared returns (but it is weak and is in fact negative for the OLS regressions). For the rest of the time-series strategies we see mostly weakly negative relationships (all negative under OLS). If one ignores the scaled linear time-series strategy (the scaling could arguably introduce a negative relationship) then more slopes are positive than negative, but this is not overwhelming.

For the cross-sectional strategies regressed on the deviations measures the estimates are all negative and sometimes even significant for the OLS regressions. For the robust regression this remains mostly true, except that some of the unscaled linear cross-sectional regressions show a weakly positive relationship. The models seem to work better for unscaled linear strategies (but still poorly), but still removing the effect of dispersion is good. This may be because the scaled strategies take advantage of more reliable profits when dispersion is low (if dispersion is countercyclical and because these strategies invest relatively more than the linear strategy when dispersion is low). However, they still invest even when dispersion is high and make relatively less profit at this time. The linear strategy in contrast could possibly make most of its profit when dispersion is high as it bets the most then. In a later section we will see what happens when removing the effect of dispersion more directly.

I run the regressions for strategies on normalised returns as well in order to see if they behave similarly or possibly closer to what was predicted in the models of section 2.5.1. Because of the weaker predictive regression results in section 5.2.2 there seems to be little hope of this. I report the robust slopes in table 6.2 and the OLS slopes in table E.6 in appendix E.4.

The results for the normalised returns are in fact (a little) more supportive of the AR(1) models for the cross-sectional regressions. Here the slopes are now often positive (particularly in the OLS regressions), particularly for the unscaled linear strategies. The pattern for the time-series robust regressions is much the same as before (there are more positive slopes for the OLS regressions now).

6.2 Regressions on market volatility

I regress the strategies on market volatility in order to confirm (or reject) the hypothesis that momentum is negatively related to market volatility. I consider two ways of defining the market. The first is an equal-weighted position in each (unweighted) asset in each period and the second is an equal-weighted position in each normalised asset in each period.

I run three types of regressions for quantile cross-sectional and signed time-series strategies where volatilities are estimated as before:

(1) returns from the strategy on unweighted assets vs the volatility of the equal-weighted market based on unweighted assets (unweighted|unweighted)

(2) returns from the strategy on normalised assets vs the volatility of the equal-weighted market based on unweighted assets (normalised|unweighted)

(3) returns from the strategy on normalised assets vs the volatility of the equal-weighted market based on normalised assets (normalised | normalised)

If there is an aspect of volatility timing to volatility weighting that contributes positively to returns, then one would expect a negative relationship for the first regression. If volatility weighting is effective in reducing the exposure to volatility, the second regression may be expected to show no (or a weaker) relationship with volatility, with the caveat that as we saw in section 2.4.1 volatility weighting (with a strategy's own volatility) can in fact induce a negative relationship with volatility if there is a portion of returns not dependent on volatility, and something similar could happen here.

One would perhaps also expect a weak relationship for the third regression based on the fact that a negative relationship here would suggest that a second round of volatility weighting (weighting the weighted assets) would be beneficial and this seems implausible.

It should be noted that the volatility weighting I have applied does not take into account covariances between assets at all and this of course contributes to market volatility (it most likely contributes positively as assets are often positively correlated). Volatility weighting using normalised returns would thus not necessarily reduce exposure to the effect of changes in correlation.

Table 6.1: ***Slopes of dispersion regressions (robust).*** *Reported are slope estimates (along with t-values in round brackets and p-values in square brackets) of regressions of returns of time-series momentum strategies on average past absolute and squared asset returns (table (a)), and returns of cross-sectional momentum strategies on past absolute and squared deviations and cross-sectional standard deviation (table (b)). Asset returns over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used.*

(a) *Time-series*

strategy		measure: |returns|	measure: (returns)^2
FF49 12 mo	sts	-0.03(-1.01)[0.313]	-0.04(-1.04)[0.299]
	ults	0.01(0.64)[0.521]	0.02(1.08)[0.281]
	slts	-0.00(-1.19)[0.234]	-0.00(-1.21)[0.228]
FF49 1 mo	sts	0.18(2.36)[0.019]	0.82(2.40)[0.017]
	ults	0.02(2.41)[0.016]	0.24(7.21)[0.000]
	slts	0.00(2.00)[0.047]	0.02(2.05)[0.041]
MAA 52 wk	sts	0.00(0.09)[0.925]	-0.00(-0.53)[0.598]
	ults	0.00(0.70)[0.487]	0.00(0.48)[0.630]
	slts	-0.00(-0.16)[0.876]	-0.00(-0.71)[0.481]
MAA 4 wk	sts	-0.01(-0.48)[0.633]	0.00(0.01)[0.993]
	ults	0.00(1.08)[0.280]	0.03(1.79)[0.073]
	slts	-0.00(-1.17)[0.242]	-0.03(-1.31)[0.192]

(b) *Cross-sectional*

strategy		measure: |deviations|	measure: (deviations)^2	measure: xsd
FF49 12 mo	qxs	-0.12(-1.64)[0.103]	-0.29(-1.91)[0.057]	-0.08(-1.55)[0.122]
	ulxs	-0.01(-1.00)[0.317]	-0.02(-1.60)[0.110]	-0.00(-1.08)[0.281]
	slxs	-0.13(-1.73)[0.084]	-0.33(-2.14)[0.033]	-0.11(-1.87)[0.062]
FF49 1 mo	qxs	-0.01(-0.06)[0.954]	-0.82(-0.53)[0.596]	-0.05(-0.29)[0.771]
	ulxs	0.00(0.45)[0.656]	0.02(0.53)[0.595]	0.00(0.46)[0.647]
	slxs	-0.01(-0.04)[0.966]	0.28(0.14)[0.886]	-0.00(-0.00)[0.998]
MAA 52 wk	qxs	-0.00(-0.26)[0.793]	-0.01(-0.54)[0.590]	-0.00(-0.06)[0.949]
	ulxs	0.00(0.09)[0.927]	-0.00(-0.32)[0.747]	0.00(0.24)[0.814]
	slxs	-0.02(-0.97)[0.332]	-0.04(-1.25)[0.213]	-0.01(-0.79)[0.431]
MAA 4 wk	qxs	-0.05(-1.01)[0.311]	-0.21(-0.68)[0.494]	-0.03(-0.70)[0.481]
	ulxs	0.00(0.22)[0.825]	0.00(0.23)[0.815]	0.00(0.19)[0.853]
	slxs	-0.19(-2.41)[0.016]	-2.08(-4.43)[0.000]	-0.14(-2.37)[0.018]

Table 6.2: ***Slopes of dispersion regressions with normalised returns (robust).*** *Reported are slope estimates (along with t-values in round brackets and p-values in square brackets) of regressions of returns of time-series momentum strategies on average past absolute and squared asset returns (table (a)), and returns of cross-sectional momentum strategies on past absolute and squared deviations and cross-sectional standard deviation (table (b)). The strategies are based on normalised returns. Asset returns over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data and the first 21 days (52 weeks) for each asset are used for initial volatility estimates. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used.*

(a) *Time-series*

strategy		measure	
		|returns|	$(\text{returns})^2$
FF49 12 mo	sts	-0.02(-0.64)[0.524]	-0.05(-0.57)[0.570]
	ults	0.01(1.17)[0.241]	0.03(1.70)[0.090]
	slts	-0.00(-0.83)[0.406]	-0.00(-0.72)[0.472]
FF49 1 mo	sts	0.07(0.96)[0.340]	0.46(0.84)[0.402]
	ults	0.01(3.02)[0.003]	0.07(3.03)[0.003]
	slts	0.00(0.66)[0.510]	0.01(0.60)[0.548]
MAA 52 wk	sts	0.00(0.65)[0.517]	-0.00(-0.31)[0.757]
	ults	0.00(2.49)[0.013]	0.00(2.06)[0.040]
	slts	0.00(1.01)[0.312]	0.00(0.62)[0.534]
MAA 4 wk	sts	-0.00(-0.15)[0.879]	-0.10(-0.54)[0.588]
	ults	0.00(1.06)[0.288]	0.00(0.07)[0.944]
	slts	-0.00(-0.00)[1.000]	-0.01(-0.53)[0.598]

(b) *Cross-sectional*

strategy		measure		
		|deviations|	$(\text{deviations})^2$	xsd
FF49 12 mo	qxs	0.08(1.23)[0.219]	0.22(0.82)[0.411]	0.06(1.08)[0.281]
	ulxs	0.01(3.21)[0.002]	0.04(3.14)[0.002]	0.01(2.96)[0.003]
	slxs	0.06(1.03)[0.303]	0.16(0.62)[0.535]	0.04(0.88)[0.381]
FF49 1 mo	qxs	0.32(1.67)[0.096]	3.82(1.18)[0.237]	0.22(1.39)[0.166]
	ulxs	0.01(2.85)[0.005]	0.08(2.54)[0.012]	0.00(2.66)[0.008]
	slxs	0.33(1.78)[0.075]	4.35(1.38)[0.168]	0.25(1.63)[0.104]
MAA 52 wk	qxs	0.01(0.65)[0.518]	-0.00(-0.08)[0.938]	0.00(0.20)[0.844]
	ulxs	0.00(1.67)[0.096]	0.00(0.92)[0.356]	0.00(1.28)[0.202]
	slxs	0.00(0.16)[0.874]	-0.02(-0.46)[0.646]	-0.00(-0.15)[0.883]
MAA 4 wk	qxs	-0.06(-1.21)[0.227]	-0.88(-1.60)[0.110]	-0.05(-1.18)[0.237]
	ulxs	0.00(0.54)[0.591]	0.00(0.26)[0.798]	0.00(0.51)[0.607]
	slxs	-0.05(-0.90)[0.368]	-0.48(-0.86)[0.391]	-0.03(-0.73)[0.466]

Table 6.3: ***Slopes of momentum vs market volatility regressions (robust)****. 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. In the first column the momentum strategy and the market are defined on unweighted returns. In the second column the strategy is run on normalised returns. In the last column the market is also based on normalised returns. The first 21 days (52 weeks) of strategy or asset returns are used for the initial volatility estimate. Asset returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used.*

strategy	unweighted | unweighted	normalised | unweighted	normalised | normalised
FF49 12 mo qxs	-0.27(-1.53)[0.127]	-0.22(-2.70)[0.007]	-0.18(-0.92)[0.357]
FF49 1 mo qxs	-0.31(-2.49)[0.013]	-0.13(-2.10)[0.036]	-0.19(-1.34)[0.182]
FF49 12 mo sts	-0.15(-0.80)[0.423]	-0.09(-0.91)[0.365]	0.05(0.23)[0.820]
FF49 1 mo sts	0.33(2.01)[0.045]	0.04(0.44)[0.661]	0.03(0.17)[0.867]
MAA 52 wk qxs	-0.29(-0.84)[0.398]	-0.50(-2.19)[0.028]	-1.35(-2.78)[0.005]
MAA 4 wk qxs	-0.45(-1.39)[0.164]	-0.44(-2.11)[0.035]	-0.70(-1.48)[0.139]
MAA 52 wk sts	-0.10(-0.77)[0.442]	-0.24(-2.49)[0.013]	-0.41(-1.95)[0.052]
MAA 4 wk sts	-0.09(-0.76)[0.447]	-0.14(-1.56)[0.120]	-0.22(-1.10)[0.272]

Table 6.4: ***Correlations of momentum and market volatility****. Reported are estimated correlation coefficients (with t-value in round brackets and p-value in square brackets) of selected momentum strategies and ex-ante market volatility. In the first column the momentum strategy and the market are defined on unweighted returns. In the second column the strategy is run on normalised returns. In the last column the market is also based on normalised returns. The first 21 days (52 weeks) of strategy or asset returns are used for the initial volatility estimate. Asset returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data.*

strategy	unweighted | unweighted	normalised | unweighted	normalised | normalised
FF49 12 mo qxs	-0.14	-0.15	-0.06
FF49 1 mo qxs	-0.15	-0.11	-0.07
FF49 12 mo sts	-0.09	-0.05	0.01
FF49 1 mo sts	0.02	-0.01	0.00
MAA 52 wk qxs	-0.06	-0.08	-0.07
MAA 4 wk qxs	-0.05	-0.06	-0.04
MAA 52 wk sts	-0.05	-0.09	-0.06
MAA 4 wk sts	-0.03	-0.05	-0.05

In table 6.3 I report the slopes of these (robust) regressions in the same format as before.[1] The slopes of the OLS regressions can be found in table E.8 in appendix E.4. Table 6.4 gives the corresponding correlations of the momentum strategies with market volatility.

With one notable exception it appears that the relationship of momentum with market volatility is negative and that this contributes to the improvement in returns from volatility weighting.

Except for the 1 month industry time-series momentum strategy the relationship with volatility (unweighted | unweighted) is negative. This strategy and the 12 month industry time-series momentum betrayed a positive relationship with volatility in section 4.3. The latter strategy displays a weakly negative relationship with market volatility here (it is the weakest of the industry strategies). Furthermore we saw that for these strategies volatility weighting with normalised returns gave a weakly negative alpha and particularly the 1 month strategy had only a small improvement in the Sharpe ratio.

This together with the negative relationships with the other strategies where volatility weighting gave a far more impressive improvement suggests that volatility timing is an important aspect of the effect of volatility weighting and may even be the most important aspect – though it is hard to separate the effect of stabilising volatility from timing volatility.

It would seem that cross-sectional strategies are more affected by volatility than time-series strategies

[1] The R-squareds of these regressions are quite low (mostly less than 1 percent).

Table 6.5: ***Predictability of dispersion (robust).*** *Reported are the estimated AR(1) coefficients of dispersion (defined as the cross-sectional standard deviation of asset returns in each period) for two datasets. Asset returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data. A robust regression with a bisquare weighting function and a parameter of 4.685 is used.*

dataset	slope	R-squared
FF49	0.35(6.06)[0.000]	12.07
MAA	0.22(6.85)[0.000]	8.45

(they have more significant slope coefficients and larger correlations). Cross-sectional strategies are also more naturally related to dispersion, which we have seen is related to volatility. (The fact that the quantile strategy only invests in the extreme assets may also play a role)

We see that the strategies with normalised returns (excluding the industry 1 month time-series strategy) continue to have a significantly negative (in fact it is often even more significant) relationship with the market volatility (of unweighted returns) and in fact the correlation is often more negative. This could indicate ineffective volatility estimates (which overestimate volatility when it is high for instance). I also indicated in section 2.4.1 that volatility weighting can induce a negative relationship with volatility when weighting a strategy by its own returns. Something similar may be at work here because of the portion of the returns that are not related to volatility. It is also interesting that the MAA data have smaller correlations with volatility (for the unweighted market) and yet volatility weighting was most effective here, possibly the result of a more variable volatility (i.e. a stabilising effect) – the variability of volatility is, however, hard to compare with EWMA estimators due to the dependence between successive estimates.[2]

The regressions of the normalised return strategies against the volatility of normalised returns still indicate negative relationships except for the two industry time-series strategies. However only two of the slopes are significant. The relationship with volatility does seem to weaken for the industry data when considering the volatility of normalised returns (the last column in tables 6.3 and 6.4), but we do not see this for the MAA data. Despite this one would expect an additional round of volatility weighting to be less effective if the variability of volatility has already been greatly reduced (I will examine this in a later section). This would lower the potential benefits both from stabilising volatility and from volatility timing.

6.3 Dispersion weighting

In this section I will attempt to treat dispersion as a form of volatility and determine whether it can be forecasted (in subsection 6.3.1) and whether weighting with dispersion improves cross-sectional momentum strategies (in subsection 6.3.2). I consider only cross-sectional strategies as it is for these strategies that a link with dispersion is predicted by my models. I weight cross-sectional strategies with both a forecast of dispersion and the actual dispersion for a period.

6.3.1 Forecasting

I run AR(1) regressions for the monthly (industry) and weekly (MAA) dispersion estimates (estimated as a cross-sectional standard deviation). The robust slopes and R-squareds of these regressions are reported in table 6.5 (the intercepts are positive and highly significant). The OLS results are in table E.9 in appendix E.4. The AR(1) regressions have a significantly positive slope coefficient, but this and the R-squared are lower than for the realised variance regressions in section 5.1.1 for the industry market, indicating that dispersion is harder to forecast (at least with an AR(1) model) than market volatility (at least for industries).

In order to forecast dispersion I run OLS AR(1) regressions with a moving window of length 36 months and 52 weeks for the industry and MAA data respectively. My dispersion forecast is then just an AR(1) forecast from this regression. I plot both the forecast and the actual dispersion on the same axes in figure 6.1. 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. Better dispersion forecasts can probably improve dispersion weighting, but it is beyond the scope of this thesis to find such forecast

[2]The adaptiveness, λ, of the estimator will affect the variability.

Table 6.6: ***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 period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data.*

	weighting		
strategy	unweighted	forecast	actual
FF49 qxs 12 mo	0.88	0.92	1.05
FF49 qxs 1 mo	0.97	1.00	1.00
MAA qxs 52 wk	0.88	1.02	1.28
MAA qxs 4 wk	1.08	1.28	1.39

techniques. It would be sufficient to know that a dispersion forecast has a similar behaviour to weighting with the actual dispersion for my purposes.[3]

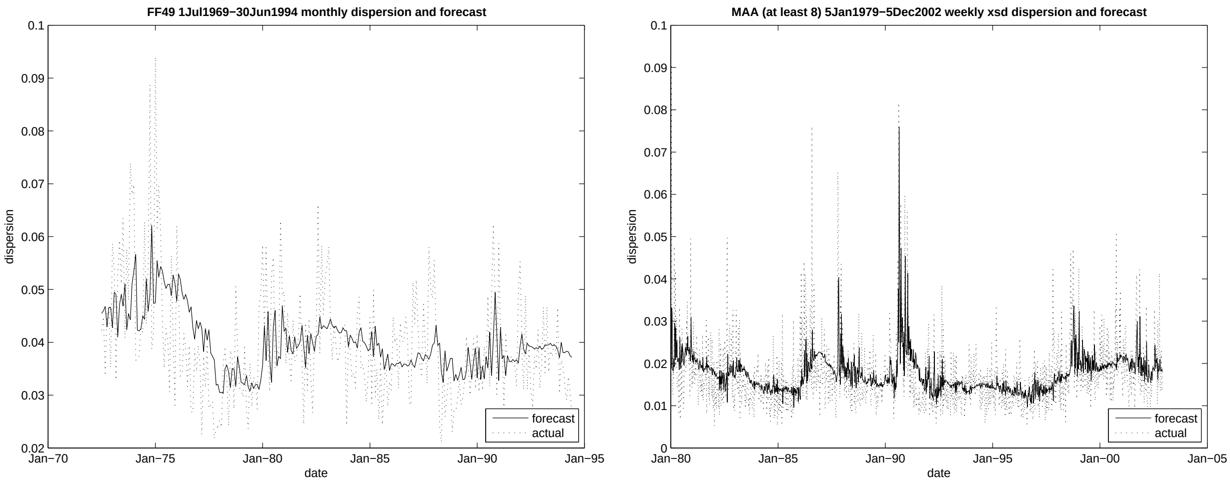


Figure 6.1: ***Actual dispersion and AR(1) forecast of dispersion****. The actual dispersion (measured as a cross-sectional standard deviation) and an (OLS) AR(1) forecast based on a moving window regression of 36 months (52 weeks) for the industry (MAA) data. Asset returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used.*

6.3.2 Effect of dispersion weighting

I run cross-sectional momentum strategies scaled with the dispersion forecast, report statistics 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 I also consider strategies scaled by the actual dispersion in the holding period. This assumes a perfect forecast ability and so these latter strategies cannot be implemented in actual investing, but they do provide a useful base for comparison. They represent a kind of upper bound for the possible value of dispersion weighting.

The efficacy of dispersion weighting would not mean that dispersion is bad for cross-sectional momentum. In a sense momentum invests in the dispersion. It is variability in dispersion that is problematic – controlling for this variability results in more stable profits. Of course if there is in fact a negative link with dispersion this may well serve to improve the effect of dispersion weighting even further. I report performance Sharpe ratios in table 6.6 and regression intercepts in table 6.7. OLS intercepts are in table E.10 in appendix E.4. Cumulative performance graphs are in figure 6.2.

The results of dispersion weighting are weak for the FF49 data, but promising for the MAA data. At least for MAA it seems there is value in dispersion-weighting and in particular in finding better dispersion

[3] A look at the autocorrelation function (unreported) of the dispersion shows a more or less decreasing pattern (the first lag is the second highest for industries and highest for MAA) suggesting an AR(1) forecast should be sufficient. One possible improvement would be to base the regression on a (moving) average of the past n months' dispersion values in order to reduce the noise in the estimation.

Table 6.7: ***Intercepts of regressions of dispersion weighted strategies on unweighted strategies (robust).*** *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. In table (a) the strategies are weighted with an AR(1) dispersion forecast and in table (b) with actual dispersion. Return data over the period Jul1969-30Jun1994 (4Jan1979-5Dec2002) for industries (MAA) is used to obtain strategy returns and dispersion estimates.*

(a) *Weighted with dispersion forecast*

industry		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	0.20(0.56)[0.578]	MAA qxs 52 wk	1.08(1.78)[0.075]
FF49 qxs 1 mo	0.10(0.40)[0.689]	MAA qxs 4 wk	0.89(1.30)[0.193]

(b) *Weighted with actual dispersion*

industry		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	0.94(1.50)[0.135]	MAA qxs 52 wk	2.15(1.89)[0.060]
FF49 qxs 1 mo	-0.26(-0.56)[0.575]	MAA qxs 4 wk	2.47(2.28)[0.023]

forecasts.

For the industries we see an improvement in the Sharpe ratio for the 12 month formation when weighting by the forecast and a further improvement with actual dispersion. For the 1 month formation the improvement with either weighting is marginal.

The regressions show a small and insignificant alpha for the 12 month weighted strategies and the 1 month strategy weighted with the forecast. The 1 month strategy weighted with actual dispersion shows a weak negative alpha (weak positive under OLS).

For the MAA data we see a sizeable improvement in the Sharpe ratio when weighting with the forecast, which is improved upon when weighting with the actual dispersion. The alpha for these strategies positive and larger than for the industry data. The alpha is larger and more significant when weighting with actual dispersion. However, with a robust regression only the 1 month strategy weighted with actual dispersion has a significant alpha (under OLS they all are).

The results are reflected in the cumulative performance graphs, which show a widening gap between the weighted and unweighted strategies for the MAA strategies particularly, and also the industry 12 month formation strategy. The performance graphs of industry 1 month strategies are all close together.

Dispersion weighting, though theoretically interesting is perhaps not practically feasible as dispersion is not easy to predict and this weighting is less effective than using normalised returns (even in the utopian case). However, this weighting does show promise for the MAA data, especially if dispersion forecasts can be improved. It may be more advantageous,though, to use strategies that naturally reduce dependence on dispersion or stabilise dispersion. A quantile cross-sectional strategy already reduces dependence compared to a linear strategy. However, using normalised returns may have a similarly beneficial effect.

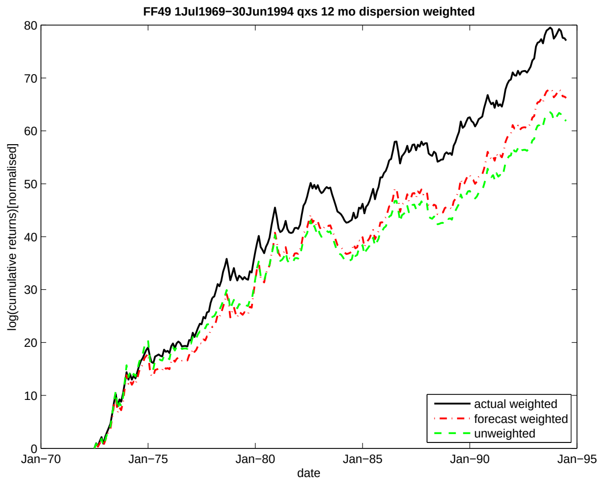
6.4 Dispersion, volatility and normalised returns

I briefly examine the relationship between dispersion and volatility and the effect of normalised returns.

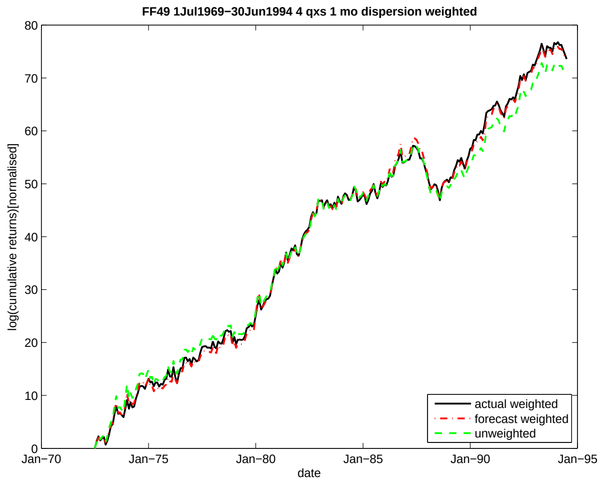
6.4.1 Variability of market volatility and dispersion

The dispersion in the MAA dataset is more variable as it has a larger coefficient of variation and kurtosis . It is then, perhaps, not surprising that dispersion weighting is more effective for MAA given the larger variability. Given that there are fewer assets in the MAA dataset it is also not surprising that the dispersion is more variable.

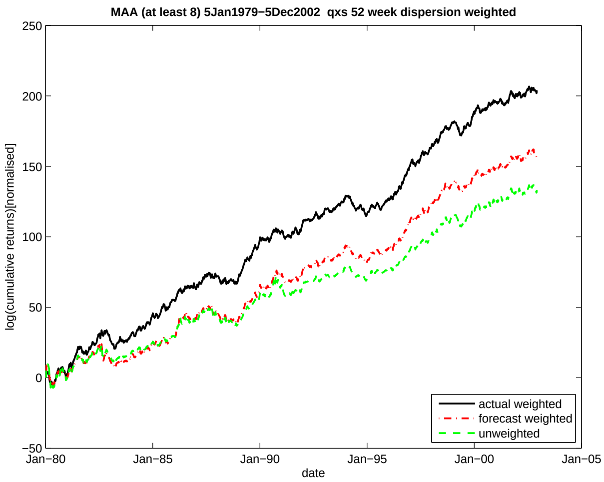
The intuition that normalising returns should stabilise dispersion appears to be correct, though it is not clear how much this contributes to improving the strategies. The coefficient of variation and kurtosis for the industry data decrease from 0.3 and 8.9 to 0.25 and 4.25 when using normalised returns. For the



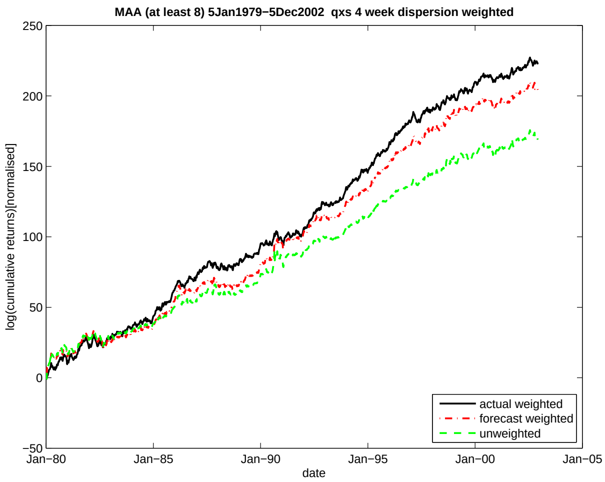
(a) *Industry 12 month formation*



(b) *Industry 1 month formation*



(c) *MAA 52 week formation*



(d) *MAA 4 week formation*

Figure 6.2: ***Cumulative performance graphs of dispersion-weighted cross-sectional strategies****. The standardised logarithm of cumulative returns is plotted for momentum strategies weighted with an AR(1) dispersion forecast and actual dispersion, as well as for an unweighted strategy. Asset returns over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for the industry (MAA) data are used to calculate strategy returns and dispersion estimates. The series for each strategy is standardised by dividing by the ex-post standard deviation of log returns.*

Table 6.8: ***Dispersion vs volatility regressions (robust).*** *Reported are slope estimates (with t-statistics in round brackets and p-values in square brackets) of regressions of dispersion versus market volatility. Dispersion and market volatility are calculated based either on unweighted or normalised returns. Asset returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used.*

(a) *Dispersion (unweighted returns) vs market volatility (unweighted returns)*

dataset	slope	R-squared
FF49	0.23(4.66)[0.000]	11.28
MAA	0.89(6.32)[0.000]	2.21

(b) *Dispersion (normalised returns) vs market volatility (unweighted returns)*

dataset	slope	R-squared
FF49	-0.17(-7.89)[0.000]	18.26
MAA	-0.24(-2.84)[0.005]	-1.87

(c) *Dispersion (normalised returns) vs market volatility (normalised returns)*

dataset	slope	R-squared
FF49	-0.28(-5.41)[0.000]	9.69
MAA	0.08(0.40)[0.687]	-1.48

MAA data the coefficient of variation decreases from 0.46 to 0.39, but kurtosis increases from 13.61 to 14.56.

One can also compare the variability of market volatility for unweighted and normalised returns. One would expect that if volatility weighting is effective it would reduce the variability of volatility for the normalised returns market vs the unweighted returns market. This is indeed the case. The coefficient of variation of volatility (estimated with EWMA's as before) for industries drops from 0.37 to 0.27 and for MAA from 0.24 to 0.15.[4] This suggests volatility weighting succeeds in stabilising volatility to some degree. However, this does not mean that this is the main source of any profits from volatility weighting.

6.4.2 Dispersion and volatility

I confirm that there is a positive link between dispersion and volatility, which suggests that dispersion weighting may include an aspect of volatility timing, which would benefit momentum strategies were there a negative link with (ex-ante) market volatility. In table 6.8 (a) I report the slope and R-squared of a simple robust regression of dispersion on market volatility measured with a 61 day (33⅓ week) EWMA for industry (MAA) data. The OLS slopes are in table E.11 in appendix E.4. Dispersion does have a significant positive relationship with volatility and a high positive correlation (0.37 for industries, 0.23 for MAA). The relationship is stronger for the industry data and is in fact somewhat weak in terms of R-squared for MAA.

I consider also dispersion in normalised returns and the volatility of the unweighted market and the market with normalised returns. These are in table 6.8 (b) and (c). The relationship of normalised dispersion vs the market volatility (with unweighted returns) is as expected in section 2.5.3. The dispersion is now negatively related to market volatility (the correlations are -0.43 for industries and -0.05 for MAA). The slope estimates are now negative (and significant). The dispersion of the normalised returns increases when market volatility is lower, which could allow the strategy to benefit if returns are now higher. Note, however, that the induced negative relationship between dispersion of normalised returns and the unweighted market volatility could also be from a poor volatility estimator.

For the regression of normalised return dispersion on the volatility of the normalised market we see for the MAA data what was predicted in section 2.5.3, an insignificant relationship with volatility.[5] However, for the industry data we see a significant negative relationship, which could be due to the effect of positive covariances (which contribute negatively to dispersion). In any case it is clear that dispersion behaves quite differently with normalised assets.

I also investigate the decomposition of dispersion into a portion related to conditional volatilities and a portion related to conditional means as in section 2.5.3. It would be interesting to see firstly how important the volatility portion is and whether it has a strong (presumably negative) link with momentum

[4] Note again that comparing the volatilities of industry and MAA is not possible here because we have used different parameters and data frequencies for the EWMA estimates.

[5] It is significantly positive under OLS, but less significant and with a smaller slope than for the unweighted returns.

Table 6.9: ***Slopes of regressions of momentum vs (log) volatility proportion of dispersion (robust).*** *Reported are slope estimates (with t-statistics in round brackets and p-values in square brackets) of regressions of momentum strategies versus the logarithm of the volatility portion of (squared) dispersion. The first column reports results for unweighted returns, the second for momentum strategies with normalised returns, and the third for both the strategy and dispersion based on normalised returns. Asset returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data are used with the first 21 days (52 weeks) for each asset or strategy used for initial volatility estimates. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used.*

strategy	unweighted | unweighted	normalised | unweighted	normalised | normalised
FF49 12 mo qxs	-0.02(-1.76)[0.080]	-0.00(-1.13)[0.257]	-0.01(-3.23)[0.001]
FF49 1 mo qxs	-0.01(-2.18)[0.030]	-0.00(-1.52)[0.130]	-0.01(-2.57)[0.011]
MAA 52 wk qxs	-0.00(-2.34)[0.020]	-0.00(-2.61)[0.009]	-0.00(-1.68)[0.094]
MAA 4 wk qxs	-0.00(-2.06)[0.039]	-0.00(-2.69)[0.007]	-0.00(-3.64)[0.000]

profits. As such I calculate

$$\nu_t^2 = \frac{1}{N}\sum_i \sigma_{i,t}^2 - \sigma_{mt}^2 \tag{6.1}$$

and consider the proportion of (squared) dispersion (χ_t^2) taken up by this, $\frac{\nu_t^2}{\chi_t^2}$. This latter figure is on average 0.36 (0.66) for industries (MAA).[6] It is interesting that the dispersion weighting was more effective where volatility took up a larger portion of the dispersion. A caveat is, however, necessary. The portion of dispersion taken up by volatility is quite variable and for the MAA data the estimate often exceeds 200%, whereas in theory it should not exceed 100%. This points to the inadequacy of the estimators. As such the results here must be treated with caution.

The correlation between the volatility portion and the mean portion (what remains) of dispersion is high and negative, -0.78 (-0.8) for industries (MAA), suggesting they tend to offset each other. This high correlation means it is hard to distinguish the separate effects of the two portions – they will work against each other. It does, however, suggest a possible explanation for why a nice empirical relationship with dispersion was not forthcoming.[7]

Looking at the normalised asset market I calculate the figures above once more and I find that now the volatility portion is even higher, 0.67 (0.81) and the negative correlation also higher, -0.92 (-0.86).

In order to see whether it is important how much of dispersion is taken up by the volatility portion I regress momentum returns on the logarithm of $\frac{\nu_t^2}{\chi_t^2}$ and report in table 6.9 the (robust) slopes of these regressions. The OLS slopes can be found in table E.12 in appendix E.4.

What one immediately notices in table 6.9 is that all the estimates are negative.[8] This suggests that it is in fact the volatility portion of dispersion that is harmful to momentum profits. The relationships remain negative when considering a strategy based on normalised returns with the dispersion of the unweighted market and the dispersion of the market of normalised returns. This latter observation is interesting as it also suggests that volatility weighting does not change the basic relationship of volatility and momentum.

6.5 Conclusions

- With one exception (the 1 month industry time-series strategies) the models of section 2.5.1 do not seem to hold. In fact the relationship of cross-sectional momentum with dispersion appears to be negative rather than positive (though often weak).
- Momentum does appear to be negatively related to volatility (with a notable exception), particularly for cross-sectional strategies and this negative relationship may well be important for volatility timing. The negative relationship persists when normalising returns.

[6] In contrast to the approximate arguments in Yu and Sharaiha (2007) it does not appear that the conditional mean portion of dispersion is negligible.

[7] Another means of separating the effect of volatility from dispersion is to orthogonalise the two series with a moving window regression and then to consider the residual of dispersion regressed on volatility.

[8] Some of the OLS estimates are positive or much weaker (smaller t-stats) because of the presence of influential observations easily identified in a scatterplot.

- Dispersion appears to be somewhat forecastable, but not as much as volatility and dispersion weighting is effective for the MAA data, which has the most variable dispersion. However, dispersion weighting is only marginally effective for the industry cross-sectional strategies.
- Dispersion is less variable for the normalised returns and so is market volatility, suggesting these are possible sources of improvement from this form of volatility weighting.
- The conditional volatility and conditional mean portions of dispersion are negatively related. The former seems to be inimical to momentum profits. This lends credence to the idea that volatility timing is at working in dispersion weighting and helps explain why a nice empirical relationship with dispersion could not be found.

Chapter 7

Out of sample

In this chapter I will test several conclusions, formed in the preceding chapters, on out-of-sample data. I list the relevant conclusions at the start of each section and indicate confirmation of a conclusion (or part of one) with a ✓ and disconfirmation with a ✗. There is one section for each of chapters 3, 4, 5, and 6. For ease of comparison I will sometimes repeat earlier results from the in-sample period.

7.1 Evidence of momentum and the relationship between strategies

(1) Shorter formation periods (down to a month) are better. ✗

(2) Shorter holding periods are stronger, there is no short-term reversal. ✓

(3) Cross-sectional (time-series) momentum performs better for industries (MAA). ✗

(4) Reducing the effect of dispersion via scaling improves linear strategies. ✓

(5) Momentum is a superior strategy to the market. ✗

I report a table of Sharpe ratios varying the formation and holding period as in section 3.2. A noticeable feature of these new tables is that, with the exception of industry time-series momentum, the momentum is much weaker in the more recent period, indicating that possibly the momentum phenomenon is being arbitraged away. It may also relate to the much weaker performance of momentum in the crisis period.[1]

In the in-sample period we saw that a 1 month (4 week) formation with a 1 month (1 week) holding period had the strongest performance. This is still true for the industry time-series strategies and the MAA quantile cross-sectional and signed time-series strategies. However, for the industry cross-sectional strategies and the MAA linear and scaled linear strategies a longer formation period is now better.

We still see, however, that shorter holding periods perform better. This does, however, have two caveats. Firstly, for the MAA 1 week formation strategies there is a short-term reversal before there is momentum. This is still consistent with the in-sample data where the performance in the first week was mostly weaker, though still positive. Secondly, the decreasing pattern of momentum returns by holding period is no longer clear for the industry cross-sectional strategies and the MAA linear strategies (scaled and unscaled) except for longer formation periods. I still consider the conclusion validated as the industry cross-sectional momentum is now clearly weaker for shorter formation periods and the linear strategies are not practically relevant.

We see that time-series momentum still outperforms cross-sectional momentum for MAA. But this is now the case for industries as well when looking at the shorter holding periods and all but the longest formation period. The nature of the market appears to be changing to favour time-series momentum, which showed less dramatic reductions and even improvements in profitability.

We still see that scaling the linear strategies results in an improved performance.

In table 7.10 (a) I also provide Sharpe ratios for the market and here only one (unweighted) momentum strategy (industry 1 month time-series) now has a Sharpe ratio higher than that of the market, which calls into question the ability of momentum to earn superior returns.

[1] Consider for instance the weak performance in 2009 for stock momentum mentioned in Daniel and Moskowitz (2011).

Table 7.1: ***OOS Sharpe ratios varying formation and holding period.*** *Annualised Sharpe ratios are calculated from monthly (weekly) returns for six different momentum strategies over the period Jul1994-Dec2012 (5Dec2002-17Apr2013) for the industry (MAA) data considering every possible combination of J month (week) formation and K month (week) holding period. Sharpe ratios are reported for the complete holding period. Holdings are rebalanced monthly (weekly) according to the weights assigned by each strategy. Sharpe ratios are annualised by multiplying by* $\sqrt{\frac{12}{K}}$ $(\sqrt{\frac{52}{K}})$.

(a) *Industries*

		K			
Strategy	J	1	3	6	12
FF49 qxs	1	0.17	0.11	0.09	0.17
	3	0.15	0.17	0.18	0.24
	6	0.17	0.25	0.27	0.25
	12	0.32	0.26	0.22	0.16
FF49 ulxs	1	0.02	0.01	0.02	0.08
	3	0.05	0.06	0.11	0.12
	6	0.08	0.13	0.21	0.15
	12	0.27	0.21	0.21	0.08
FF49 slxs	1	0.17	0.07	0.06	0.12
	3	0.14	0.11	0.15	0.18
	6	0.18	0.21	0.25	0.22
	12	0.32	0.25	0.22	0.13
FF49 sts	1	0.52	0.18	0.06	0.13
	3	0.36	0.10	0.11	0.15
	6	0.33	0.21	0.23	0.23
	12	0.31	0.23	0.19	0.13
FF49 ults	1	0.36	0.13	0.03	0.05
	3	0.23	0.14	0.04	0.04
	6	0.09	0.06	0.03	0.06
	12	0.18	0.06	0.06	0.09
FF49 slts	1	0.48	0.20	0.07	0.14
	3	0.33	0.11	0.09	0.14
	6	0.29	0.23	0.27	0.25
	12	0.39	0.30	0.28	0.15

(b) *MAA*

		K				
Strategy	J	1	4	8	26	52
MAA qxs	1	-0.18	0.15	0.13	0.13	0.09
	4	0.41	0.14	0.19	0.09	0.13
	8	0.34	0.31	0.34	0.15	0.17
	26	0.35	0.25	0.21	0.05	0.09
	52	0.31	0.33	0.25	0.03	0.13
MAA ulxs	1	-0.54	-0.00	0.02	0.06	-0.01
	4	0.00	0.11	0.15	0.04	-0.01
	8	0.05	0.16	0.25	0.01	-0.03
	26	0.25	0.14	0.09	-0.16	-0.15
	52	0.01	-0.01	-0.10	-0.29	-0.16
MAA slxs	1	-0.39	0.08	0.04	0.08	0.04
	4	0.17	0.04	0.13	0.07	0.07
	8	0.16	0.19	0.30	0.12	0.12
	26	0.46	0.32	0.23	-0.04	-0.03
	52	0.30	0.24	0.12	-0.15	-0.09
MAA sts	1	0.10	0.31	0.22	0.18	0.13
	4	0.69	0.31	0.29	0.17	0.18
	8	0.49	0.45	0.44	0.24	0.24
	26	0.52	0.45	0.46	0.28	0.28
	52	0.62	0.58	0.52	0.29	0.28
MAA ults	1	-0.43	0.04	0.07	0.08	0.02
	4	0.08	0.20	0.20	0.11	0.04
	8	0.13	0.22	0.29	0.09	0.04
	26	0.33	0.27	0.24	-0.00	-0.01
	52	0.13	0.13	0.05	-0.12	-0.02
MAA slts	1	-0.31	0.18	0.14	0.14	0.09
	4	0.28	0.20	0.26	0.16	0.15
	8	0.33	0.38	0.44	0.23	0.21
	26	0.68	0.56	0.51	0.22	0.17
	52	0.45	0.43	0.34	0.04	0.03

7.2 Source of momentum profits

(1) *Except for the 12 month formation for industries autocovariances are the main source of momentum profits.* ✗

(2) *The outperformance of the cross-sectional strategies for industries come from negative cross-serial correlations or the reversal of negative returns* ✓ *and for time-series strategies in the MAA data it comes from return continuation mostly of positive returns* ✓.

(3) *There is a contribution to momentum from accuracy* ✓, *asymmetry in profits* ✗ *and excess accuracy*✓.

(4) *There is a lower accuracy and Sharpe ratio in high volatility states* ✓ *except for the industry time-series momentum strategies which show an asymmetric relationship with volatility* ✗.

(5) *Long positions contribute more than short positions.* ✓

(6) *Performance is not stronger when time-series and cross-sectional strategies agree* ✓, *but it appears that a cross-sectional strategy could potentially be enhanced by considering the time-series signal* ✓.

(7) *There are positive relationships in the predictive regressions with absolute and squared returns and deviations.* ✗

(8) *Global time-series momentum does not do better when cross-serial covariances are not large enough to eliminate cross-sectional momentum profits* ✓.

Empirical decomposition

In table 7.2 I report the empirical decomposition for linear strategies over the out-of-sample period (now using the full 17 assets for the MAA data). For convenience I repeat the in-sample decomposition as well. There are clear differences in the decomposition from what was observed in the in-sample period.

The industry 12 month formation period decomposition now has positive auto-correlations, but these are small and thus this decomposition is still consistent with the in-sample data where we saw that cross-serial correlations were the most important source of momentum for cross-sectional and means for time-series momentum. However, it does show why time-series has managed to perform better and why cross-sectional still outperforms (except for the scaled linear strategies).

The industry 1 month formation period decomposition is also consistent with the in-sample decomposition, but drag caused by a negative contribution to cross-serial covariances is larger, which explains why time-series now outperforms. In fact the outperformance of the cross-sectional strategies in the in-sample period given the pattern of the decomposition seemed anomalous. Now things are rather more like what one would expect.

The drag of cross-serial covariances has also become greater for the MAA 4 week strategy. It is in the MAA 52 week strategies that the greatest difference lies. The auto-covariances are now negative and the only contribution to momentum is from the mean terms. It is still consistent with the outperformance of time-series as this strategy does not have the negative contribution of cross-serial covariances. Note also that except for the industry 12 month time-series strategies the contribution of the mean terms has become more important, which is consistent with a weakening momentum.

Long-short analysis

In table 7.3 I report figures for a long-short analysis as in section 4.2 for the new data. For the industry 12 month formation as we saw the quantile cross-sectional strategy still (barely) outperforms the signed time-series strategy. It is still the case that for the short|long portfolios and down|down case with 2 quantiles that the cross-sectional strategy outperforms from a reversal of the negative returns and that this seems to be mostly from scenario 4 (where both quantiles going down reverse). However, the frequency of these cases where XS outperforms has dropped and profits fallen and in the up|up case in table 7.3 (a2) we see time-series outperforming where cross-sectional outperformed before. The somewhat unimpressive outperformance of the cross-sectional strategy here suggests that despite the negative lead-lag relationships present, the mean return is high enough, as reasoned in section 2.2, to negate much of the gain to the cross-sectional strategy.

For the industry 1 month formation we no longer see an outperformance for cross-sectional strategies for industries and this is reflected here. We now see that time-series outperforms due to a continuation of positive returns. There is also a continuation of negative returns (which was not there before), but

Table 7.2: ***OOS empirical decomposition for linear strategies.*** *Estimates for the decomposition for linear strategies in section 2.3.1 are reported from returns from all possible combinations of formation period and holding period. 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period are considered for industry (MAA) data over the period Jul1994-Dec2012 (5Dec2002-17Apr2013) are considered for the hold-out sample in table (b). In-sample results are repeated in table (a). For the MAA data a set of 11 assets (for which data is available over the whole period) is considered. Figures are in per cent and annualised by multiplying by 12 (52). Percentage of total profits is indicated in brackets.*

(a) *In-sample*

strategy	auto	cross	mean	total
FF49 12 mo XS	-0.43(-83.99)	0.88(170.27)	0.07(13.72)	0.52(100.00)
FF49 12 mo TS	-0.44(-905.56)		0.49(1005.56)	0.05(100.00)
FF49 1 mo XS	0.49(372.13)	-0.36(-276.61)	0.01(4.48)	0.13(100.00)
FF49 1 mo TS	0.50(94.41)		0.03(5.59)	0.52(100.00)
MAA11 52 wk XS	0.50(82.50)	-0.11(-17.92)	0.22(35.42)	0.61(100.00)
MAA11 52 wk TS	0.55(63.36)		0.32(36.64)	0.87(100.00)
MAA11 4 wk XS	0.09(150.27)	-0.05(-74.67)	0.01(24.40)	0.06(100.00)
MAA11 4 wk TS	0.10(81.76)		0.02(18.24)	0.12(100.00)

(b) *Hold-out sample*

strategy	auto	cross	mean	total
FF49 12 mo XS	0.03(6.51)	0.24(56.01)	0.16(37.47)	0.43(100.00)
FF49 12 mo TS	0.03(3.28)		0.84(96.72)	0.86(100.00)
FF49 1 mo XS	0.39(5434.56)	-0.40(-5489.62)	0.01(155.06)	0.01(100.00)
FF49 1 mo TS	0.40(85.59)		0.07(14.41)	0.47(100.00)
MAA17 52 wk XS	-0.17(-502.64)	-0.03(-97.93)	0.23(700.57)	0.03(100.00)
MAA17 52 wk TS	-0.18(-51.05)		0.52(151.05)	0.35(100.00)
MAA17 4 wk XS	0.01(343.15)	-0.03(-710.33)	0.02(467.18)	0.00(100.00)
MAA17 4 wk TS	0.02(25.47)		0.05(74.53)	0.06(100.00)

these cases are less frequent and the profits smaller. The outperformance seems to be from scenario 1 in the up|up case (both quantiles go up and continue to do so), which has become more frequent, whereas scenario 4 (in which cross-sectional outperforms) is less frequent.

For the MAA data we still see the outperformance of time-series for the long|short portfolios and in the up|up case with two quantiles. However, the cross-sectional strategy outperforms for the short|long portfolios and in the down|down case. The latter, however, are so infrequent that they seem to make very little difference. Given the negative auto-covariances in the decomposition for the 52 week formation strategy, it would appear that here the outperformance of time-series is merely the result of a high mean return (which makes the continuation of negative returns less likely).

In both cases most of the profits still seem to come from the long positions: these are more frequent and have higher Sharpe ratios. We also see that a cross-sectional strategy performs better in the portfolios where it has the same position as a time-series strategy.

Table 7.3: ***OOS Long-short analysis***. *The* long and short portolfolio tables *report returns, Sharpe ratios and the average proportion of assets in each of four portfolios dividing assets into those that a times-series and two quantile strategy would buy or sell. The* two quantile strategies *tables divide assets into a top and bottom half (possible omitting one asset) each formation period and reports the returns and Sharpe ratio for each half in three cases (both quantiles going up in the formation period; one going up and one down; and both going down) as in figure 2.2 and the proportion of occurrences for each case. The* four scenarios *tables reports time-series and cross-sectional Sharpe ratios for each of four scenarios (in up|up and down|down cases) as defined in figures 2.3 and 2.4 along with the proportion of occurrences of each scenario. Industry (MAA) data for the period Jul1994-Dec2012 (5Dec2002-17Apr2013) is used. Where there is not enough data to calculate an entry, this is indicated by* NaN.

(a) *Industry 12 month formation*

(a1) *Long and short portfolios*

portfolio	return	Sharpe	proportion
long|long	9.69	0.57	0.42
long|short	3.35	0.25	0.22
short|long	5.98	0.29	0.07
short|short	3.82	0.17	0.27

(a2) *Two quantile strategies*

case	portfolio	return	Sharpe	proportion
up|up	top	12.45	0.85	0.45
	bottom	6.10	0.49	
up|down	top	10.46	0.63	0.46
	bottom	3.20	0.16	
down|down	top	2.40	0.10	0.09
	bottom	18.28	0.42	

(a3) *Two quantile strategies: four scenarios*

		scenario			
case	statistic	sc1	sc2	sc3	sc4
up|up	proportion	0.57	0.02	0.06	0.35
	TS Sharpe	1.72	-0.02	0.98	-1.40
	XS Sharpe	0.39	-1.73	1.37	0.02
down|down	proportion	0.33	0.00	0.11	0.56
	TS Sharpe	1.42	NaN	2.31	-1.60
	XS Sharpe	0.72	NaN	1.53	-0.79

(b) *Industry 1 month formation*

(b1) *Long and short portfolios*

portfolio	return	Sharpe	proportion
long|long	8.94	0.53	0.39
long|short	15.98	1.08	0.16
short|long	-1.58	-0.09	0.10
short|short	8.36	0.40	0.33

(b2) *Two quantile strategies*

case	portfolio	return	Sharpe	proportion
up|up	top	21.34	1.46	0.31
	bottom	19.21	1.54	
up|down	top	6.15	0.42	0.53
	bottom	4.96	0.29	
down|down	top	-4.02	-0.18	0.16
	bottom	0.40	0.01	

(b3) *Two quantile strategies: four scenarios*

		scenario			
case	statistic	sc1	sc2	sc3	sc4
up|up	proportion	0.59	0.10	0.06	0.25
	TS Sharpe	1.31	0.51	-0.84	-1.55
	XS Sharpe	0.38	-2.00	1.34	-0.38
down|down	proportion	0.34	0.06	0.09	0.51
	TS Sharpe	1.35	-0.46	0.55	-1.56
	XS Sharpe	0.63	-1.29	1.91	-0.59

Table 7.3: ***OOS Long-short analysis*** *(continued from previous page)*

(c) *MAA 52 week formation*

(c1) *Long and short portfolios*

portfolio	return	Sharpe	proportion
long|long	6.61	0.75	0.45
long|short	2.04	0.40	0.17
short|long	4.14	0.26	0.02
short|short	0.52	0.04	0.30

(c2) *Two quantile strategies*

case	portfolio	return	Sharpe	proportion
up|up	top	10.96	1.09	0.31
	bottom	5.58	1.35	
up|down	top	5.13	0.65	0.69
	bottom	-0.54	-0.04	
down|down	top	NaN	NaN	0.00
	bottom	NaN	NaN	

(c3) *Two quantile strategies: four scenarios*

		scenario			
case	statistic	sc1	sc2	sc3	sc4
up|up	proportion	0.41	0.19	0.24	0.17
	TS Sharpe	1.79	-0.65	1.11	-1.37
	XS Sharpe	0.92	-1.08	1.87	-0.62
down|down	proportion	NaN	NaN	NaN	NaN
	TS Sharpe	NaN	NaN	NaN	NaN
	XS Sharpe	NaN	NaN	NaN	NaN

(d) *MAA 4 week formation*

(d1) *Long and short portfolios*

portfolio	return	Sharpe	proportion
long|long	8.54	1.02	0.44
long|short	6.02	0.84	0.10
short|long	0.02	0.00	0.03
short|short	2.02	0.15	0.37

(d2) *Two quantile strategies*

case	portfolio	return	Sharpe	proportion
up|up	top	4.38	0.48	0.07
	bottom	11.24	4.66	
up|down	top	8.34	1.07	0.92
	bottom	0.69	0.05	
down|down	top	16.59	5.22	0.01
	bottom	106.59	2.73	

(d3) *Two quantile strategies: four scenarios*

		scenario			
case	statistic	sc1	sc2	sc3	sc4
up|up	proportion	0.44	0.31	0.19	0.06
	TS Sharpe	2.66	-0.92	0.51	-3.36
	XS Sharpe	1.05	-1.49	1.37	-1.90
down|down	proportion	0.00	0.25	0.25	0.50
	TS Sharpe	NaN	NaN	NaN	-12.85
	XS Sharpe	NaN	NaN	NaN	-0.92

Prediction analysis

In table 7.4 and 7.5 I report some of the prediction analysis results as in section 4.3 for the new datasets. Prediction accuracy is still above 50% overall (sometimes only barely) and also in most other states. There is also still an excess accuracy, which for the most part is still significant,[2] but lower than before. It is interesting to note it is now the time-series strategies that have the highest excess accuracy in all cases (previously cross-sectional had the advantage for industries). The asymmetry in profits seen earlier now seems less clear. Half the strategies now show lower profits when right than losses when wrong (overall). However, the ratio remains close to 1. Momentum strategies now do appear to have a negative relationship with volatility, with the aberrant behaviour of the industry time-series strategies disappearing. In table 7.4 (a) three strategies[3] show a (marginally) higher accuracy in the high volatility state overall. For all strategies except the MAA 4 week cross-sectional strategy we see a higher Sharpe ratio overall in the low volatility state in table 7.4 (b). We now see some aberrant behaviour in the MAA 4 week strategies which show a higher Sharpe ratio and a higher profit to loss ratio in the high volatility state, which they did not do before. Within the positive and negative prediction states conclusions are now harder to draw. There are now many exceptions to the general rule that higher volatility negatively affects momentum in the negative prediction state, but few in the positive prediction state. Previously this seemed rather to be the other way round.

In table 7.6 I report the Sharpe ratios for signed strategies when the cross-sectional and time-series signals agree and disagree. We see here that the cross-sectional strategy does better when signals agree (which suggests again that the cross-sectional strategy is weaker) and that only the MAA 52 week time-series strategy performs better when the signals agree. Thus I conclude again that it may only be useful to consider the time-series signal for a cross-sectional strategy.

Predictive regressions

In table 7.7 I report the (scaled) average slopes of predictive regressions of returns, absolute returns, deviations and absolute deviations. Here the slopes of the absolute returns and deviations remain positive, and many are now larger. However, the relationship of deviations with past deviations is now negative except for the industry 1 month vs 12 month regressions where the slope is less positive. This is consistent with a weakening cross-sectional momentum. We also see that the relationship of returns is negative for the industry (MAA) 1 month vs 12 month (1 week vs 52 week) regressions and less positive in the other two cases.

Local and global time-series momentum

In table 7.8 I report Sharpe ratios for signed local and global strategies as in section 4.6. Here we see a similar picture to earlier. The global time-series strategy does not perform better, either because the cross-serial covariances are not positive or they are not strong enough. In this case, however, the larger part of time-series profits comes not from auto-covariances for MAA but from positive means and this is being downweighted.

7.3 Volatility weighting

(1) α, the portion of returns independent of strategy volatility is in the appropriate range. ✓

(2) Momentum strategies typically have a negative relationship with their own volatility. ✓

(3) Weighting a strategy by its own volatility is effective when the relationship with volatility is negative. ✓

(4) Using normalised returns is more effective than weighting a strategy by its own volatility ✓.

(5) Volatility weighting lowers kurtosis ✓ *and gives a more positive skew* ✗ *and lower normalised drawdowns* ✗.

In table 7.9 I report estimates of γ, α and the upper and lower limits of the range over which volatility weighting could be expected to work as in table 5.2 in section 5.1.

[2]It is not surprising that it is less significant than before as we now have less data.

[3]These are the industry cross-sectional strategies and the MAA 4 week cross-sectional strategy.

Table 7.4: ***OOS Prediction analysis for signed strategies (accuracy, Sharpe ratios, profits over losses).*** *Tables (a), (b) and (c) consider, respectively, the proportion of correct predictions, the (weighted average) Sharpe ratio and ratio of profits when a prediction is correct to losses when wrong, for signed time-series (cross-sectional) strategies run on individual asset returns (deviations). Positive and negative predictions and low and high volatility states are considered separately. Volatilities are estimated with a slow and fast EWMA per asset and a high (low) volatility state is when the fast EWMA is above (below) the slow EWMA. Returns for the period Jul1994-Dec2012 (Dec2002 -17Apr2013) are used for industries (MAA) and the first 21 days (52 weeks) of data for each asset are used to obtain initial volatility estimates.*

(a) *Accuracy*

	+ prediction			− prediction			overall		
strategy	σ low	σ high	all +	σ low	σ high	all −	σ low	σ high	all
FF49 12 mo XS	51.35	50.44	50.94	50.99	52.23	51.48	51.15	51.33	51.23
FF49 1 mo XS	50.36	49.58	50.03	50.31	51.25	50.72	50.34	50.43	50.38
FF49 12 mo TS	57.64	57.29	57.51	42.53	47.75	45.60	53.73	53.01	53.40
FF49 1 mo TS	60.15	58.49	59.44	45.72	46.77	46.26	54.48	52.85	53.72
MAA 52 wk XS	52.51	52.17	52.36	54.10	51.84	53.30	53.44	52.00	52.89
MAA 4 wk XS	50.46	51.42	50.82	52.46	52.07	52.32	51.50	51.76	51.59
MAA 52 wk TS	56.89	56.09	56.66	51.89	48.34	50.18	55.57	52.62	54.53
MAA 4 wk TS	56.76	56.49	56.67	48.03	48.22	48.10	53.22	52.69	53.04

(b) *Sharpe ratios*

	+ prediction			− prediction			overall		
strategy	σ low	σ high	all +	σ low	σ high	all −	σ low	σ high	all
FF49 12 mo XS	0.14	0.02	0.09	0.14	0.09	0.12	0.15	0.06	0.11
FF49 1 mo XS	-0.00	-0.09	-0.04	0.03	-0.01	0.01	0.01	-0.04	-0.01
FF49 12 mo TS	0.49	0.40	0.45	-0.43	-0.10	-0.19	0.22	0.13	0.17
FF49 1 mo TS	0.69	0.58	0.63	-0.25	0.01	-0.09	0.28	0.27	0.27
MAA 52 wk XS	0.15	0.18	0.19	0.36	0.00	0.22	0.28	0.07	0.16
MAA 4 wk XS	0.01	0.26	0.14	0.33	0.17	0.26	0.14	0.19	0.16
MAA 52 wk TS	0.60	0.54	0.57	-0.01	-0.23	-0.08	0.44	0.14	0.28
MAA 4 wk TS	0.78	0.78	0.75	-0.10	0.02	-0.01	0.37	0.36	0.36

(c) *Profits over losses*

	+ prediction			− prediction			overall		
strategy	σ low	σ high	all +	σ low	σ high	all −	σ low	σ high	all
FF49 12 mo XS	1.11	1.03	1.06	1.05	1.00	1.03	1.08	1.01	1.05
FF49 1 mo XS	1.03	0.97	1.00	0.98	0.96	0.97	1.00	0.96	0.98
FF49 12 mo TS	1.07	0.98	1.03	1.02	1.02	1.04	1.02	0.97	0.99
FF49 1 mo TS	1.09	1.09	1.09	0.99	1.16	1.09	1.01	1.10	1.05
MAA 52 wk XS	1.03	0.96	0.99	0.96	0.95	0.95	0.99	0.96	0.97
MAA 4 wk XS	1.03	1.05	1.04	0.96	1.00	0.98	0.99	1.02	1.01
MAA 52 wk TS	0.96	0.88	0.92	0.94	1.04	0.98	0.95	0.93	0.92
MAA 4 wk TS	0.98	0.98	0.98	0.97	1.15	1.06	0.96	1.04	1.00

Table 7.5: ***OOS prediction analysis for signed strategies (excess accuracy).*** *Proportion of correct predictions (in excess of those expected under independence) along with t-statistic in brackets are reported for data over the period the period Jul1994-Dec2012 (5Dec2002-17Apr2013) for industries (MAA).*

industries		MAA	
strategy	excess accuracy	strategy	excess accuracy
FF49 12 mo XS	1.15(2.33)	MAA 52 wk XS	2.84(5.18)
FF49 1 mo XS	0.47(0.98)	MAA 4 wk XS	1.95(3.73)
FF49 12 mo TS	1.46(2.96)	MAA 52 wk TS	3.36(6.13)
FF49 1 mo TS	2.77(5.76)	MAA 4 wk TS	2.90(5.53)

Table 7.6: ***OOS prediction analysis for signed strategies (agreeing and opposing signal Sharpe ratios)*** *Annualised Sharpe ratios for signed time-series and cross-sectional strategies (based on aggregated returns across all assets) are reported for cases where the time-series and cross-sectional signals agree and disagree. The time-series (cross-sectional) signal for an asset is the sign of the formation period return (deviation). Data for the period Jul1994-Dec2012 (5Dec2002-17Apr2013) are used for industry (MAA) data.*

(a) *Opposing signals*

	TS neg XS pos		TS pos XS neg		all	
data and formation	TS	XS	TS	XS	TS	XS
FF49 12 mo	0.00	-0.08	0.69	0.02	0.39	-0.02
FF49 1 mo	0.09	-0.07	0.79	-0.02	0.49	-0.04
MAA 52 wk	-1.02	0.43	0.46	0.06	0.25	0.10
MAA 4 wk	-0.53	0.33	0.70	-0.10	0.35	0.03

(b) *Agreeing signals*

	TS neg XS neg		TS pos XS pos		all	
data and formation	TS	XS	TS	XS	TS	XS
FF49 12 mo	-0.14	0.02	0.49	0.04	0.17	0.03
FF49 1 mo	-0.14	0.01	0.50	0.02	0.18	0.01
MAA 52 wk	0.05	0.24	0.54	0.20	0.29	0.22
MAA 4 wk	-0.01	0.24	0.66	0.19	0.30	0.21

Table 7.7: ***OOS average scaled slopes of predictive regressions (robust).*** *Average slopes, t-statistics (in round brackets) and p-values (in square brackets) are reported for per-asset predictive regressions for various metrics (returns and absolute returns, deviations and absolute deviations) in the holding period versus the same metric in the formation period. Data for the period Jul1994-Dec2012 (5Dec2002 -17Apr2013) for industries (MAA) are used. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used.*

	metric			
regression	returns	|returns|	deviations	|deviations|
FF49 1 mo vs 12 mo	-0.04(-0.13)[0.527]	0.25(0.99)[0.387]	0.03(0.24)[0.399]	0.34(1.37)[0.306]
FF49 1 mo vs 1 mo	0.04(0.54)[0.497]	0.09(1.21)[0.325]	-0.00(-0.03)[0.465]	0.11(1.45)[0.274]
MAA 1 wk vs 52 wk	-0.10(-0.17)[0.491]	0.73(1.92)[0.149]	-0.20(-0.51)[0.519]	0.73(1.83)[0.217]
MAA 1 wk vs 4 wk	0.00(0.08)[0.321]	0.22(2.22)[0.147]	-0.02(-0.16)[0.418]	0.26(2.54)[0.107]

Table 7.8: ***OOS Sharpe ratios for local and global time-series strategies (norm returns).*** *The local strategies are run on (and diversified over) individual assets. The global strategy is a signed strategy run on an equal-weighted market of all assets in the dataset. Results are reported for industry data, all MAA data and and subset of three bond indices from the MAA data. Return data over the period Jul1994-Dec2012 (5Dec2002-17Apr2013) for industries (MAA) are used.*

data and formation	local	global
FF49 12 mo	0.31	0.05
FF49 1 mo	0.52	0.36
MAA 52 wk	0.50	0.40
MAA 4 wk	0.60	0.49
MAA bonds 52 wk	0.79	0.76
MAA bonds 4 wk	0.46	0.36

Table 7.9: ***OOS strategy relationship with own volatility (robust).*** *Reported are estimates of α (scaled by 100) and γ as in (5.1), the upper and lower bounds of the ranges in proposition 2.2 and the estimated error variance of the regression (5.1) for selected momentum strategies and an equal-weighted market. Return data over the period Jul1994-Dec2012 (5Dec2002-17Apr2013) are used for the industry (MAA) data to estimate strategy returns, the first 21 days (52 weeks) of which are used to provide initial volatility estimates.*

strategy	γ	100α	low	high	error var
FF49 12 mo qxs	0.01(0.06)[0.953]	0.53(0.83)[0.406]	-0.05	8.41	1.22
FF49 1 mo qxs	-0.07(-0.35)[0.726]	0.37(0.61)[0.541]	0.28	6.88	1.26
FF49 12 mo sts	-0.16(-0.81)[0.420]	0.98(1.82)[0.071]	0.58	8.36	0.93
FF49 1 mo sts	0.00(0.01)[0.994]	0.30(0.74)[0.457]	-0.00	6.29	1.03
FF49 equal-weighted	-0.02(-0.09)[0.928]	1.02(1.44)[0.152]	0.08	9.16	0.99
MAA 52 wk qxs	-0.34(-2.07)[0.039]	1.00(3.15)[0.002]	0.83	5.40	1.13
MAA 4 wk qxs	-0.09(-0.58)[0.562]	0.30(0.94)[0.348]	0.22	4.63	1.12
MAA 52 wk sts	-0.27(-1.68)[0.094]	0.35(2.84)[0.005]	0.26	2.19	1.04
MAA 4 wk sts	-0.08(-0.56)[0.573]	0.13(1.25)[0.213]	0.08	1.76	1.12
MAA equal-weighted	0.05(0.31)[0.759]	0.11(0.86)[0.391]	-0.05	1.90	1.12

We see a similar pattern to before. Most of the momentum strategies have a negative relationship with their ex-ante volatility. The 1 month industry time-series strategy still has a positive relationship, but it appears to be far weaker. The industry 12 month XS strategy is the only other momentum strategy displaying a positive (but very weak) relationship, but this is again negative under an OLS regression (unreported). The industry market strategy now also has a (weak) negative relationship with its volatility, so it is not clear that the markets should have a positive relationship with volatility. Again most of the relationships are not significant. Alpha is once again within the upper and lower limits with the upper limit being large. It is also encouraging that the error variance is quite close to 1 in these regressions.

In table 7.10 I report various statistics for unweighted strategies and volatility-weighted strategies as in table 5.3 in chapter 5. Table 7.11 reports intercepts of regressions of volatility-weighted strategies on unweighted strategies as in table 5.4. I repeat for convenience the in-sample regression results.

Both forms of volatility weighting now show convincing improvements in Sharpe ratios. There is only one exception, weighting the MAA market strategy with its own volatility, where there is a negligible decrease in the Sharpe ratio. Nevertheless the intercepts of the regressions in table 7.11 are all positive. I conclude that volatility weighting remains effective and in fact appears to be more effective in this more recent period. It is encouraging that volatility weighting at least raises the Sharpe ratio of the time-series strategies above that of a simple equal-weighted market strategy. Note, however, that weighting assets by their underlying volatility also improves the Sharpe ratio of the market strategy so that only the industry 1 month and MAA 4 week time-series strategies performs better when using normalised returns. The standard deviations of the strategies weighted with their own volatility are now around the target of 10% for both datasets, which suggests the volatility estimates are somewhat effective.

Using normalised returns still seems to perform better in most cases (but there are a number of exceptions). It is interesting to note that the industry 1 month formation time-series strategy and the industry market strategy have gone from being hurt by weighting with their ex-ante volatility to being improved. The former strategy seems to have a weaker positive relationship with its volatility and the latter now has a negative relationship. It is also interesting that the time-series strategies now have a positive alpha in table 7.11d and betray a negative relationship with volatility in table 7.4. This lends credence to the idea that a negative relationship with volatility is important (but not necessary) for volatility weighting. We again note that volatility weighting appears to be most effective for the MAA data.

Volatility weighting still seems to reduce kurtosis (it does so in all cases now), but is not clear that it improves skewness or lowers drawdowns.

7.4 Relationship with dispersion and market volatility

(1) The models of section 2.5.1 do not seem to hold and the relationship with dispersion is negative rather than positive (at least for unweighted returns). ✓

(2) Momentum has a mostly negative relationship with market volatility, which persists under normalised returns. ✓

(3) Using normalised returns reduces variability in dispersion and stabilises market volatility.

(4) Dispersion weighting is effective for the MAA data. ✓

(5) The conditional volatility portion of dispersion is negatively related to momentum profits ✗.

In table 7.12 I report the slopes of regressions of time-series strategies on absolute and squared returns and cross-sectional strategies on measures of dispersion. For the former we once again see little evidence of positive relationships and for the latter the majority of slopes are again negative.

In tables 7.13 and 7.14 I report slopes of regressions of momentum strategies on market volatility and correlations of momentum and market volatility respectively. We see mostly negative slopes and correlations which reinforce our earlier conclusion (the MAA 4 week time-series strategy now has a weakly positive rather than negative slope).

Consider the industry time-series strategies. The slopes of the 1 month formation strategy is now negative (it was positive before) in the first column of table 7.13 and the 12 month formation has a significant negative slope (it was weaker and smaller before). At first glance this seems to reinforce the conclusion that a negative relationship with market volatility is important for the effectiveness of volatility weighting. However, the industry 1 month strategy still has a positive correlation (it is even larger) with market volatility and the MAA four week time-series strategy has a positive slope in the first column of table 7.13 and here volatility still seems quite effective (note that it still has a negative correlation and a negative slope in the OLS regression(unreported)).

In tables 7.15 and 7.16 I report Sharpe ratios for cross-sectional strategies weighted with dispersion and intercepts of regressions of dispersion-weighted vs unweighted strategies. The results are very similar to those we had before.

Dispersion weighting increases the Sharpe ratio in all cases (both for the forecast and actual dispersion) and we see a positive alpha for the weighted strategies, even for the industry 1 month cross-sectional strategy weighted with actual dispersion, which had negative alpha before. However, this is still the only strategy for which weighting with actual dispersion is not more profitable than weighting with the forecast.

The coefficient of variation of dispersion is again lower for normalised returns – 0.28 vs 0.39 (0.46 vs 0.58) for industry (MAA) data (the kurtosis for MAA data does not drop now). The coefficient of variation of market volatility also still drops for normalised returns – 0.24 vs 0.52 (0.2 vs 0.43).

In table 7.17 I report the slopes of regressions of momentum vs the logarithm of the volatility portion of dispersion as in table 6.9 in section 6.4. Here we no longer see a negative relationship with the volatility portion of dispersion (it is weakly positive) and the earlier conclusion is cast into doubt.

7.5 Conclusions

- Momentum is weaker in the more recent past and conditions have changed it seems so as to favour time-series momentum.

Table 7.10: ***OOS performance statistics of volatility weighted and unweighted strategies.*** *Descriptive statistics (as in table 3.3) are calculated for quantile cross-sectional and signed-time series strategies with 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period as well as an equal-weighted market strategy. Industry (MAA) data over the period Jul1994-Dec2012 (5Dec2002-17Apr2013) are used to calculate the strategy returns and volatility estimates, using the first 21 days (52 weeks) of strategy returns or asset returns for initial volatility estimates. Table (a) reports results for unweighted strategies, starting with the first full monthly return after the first 21 days for industry data and after 52 weeks for the MAA data. In table (b) statistics are reported for strategies weighted with their own volatility. In table (c) statistics are reported for strategies run on normalised returns. Annualisation is as before.*

(a) *unweighted*

strategy	mean	sd	skew	mean less med	kurtosis	sharpe	ave drawdowns
FF49 qxs 12 mo	6.00	18.05	-0.97	2.14	6.14	0.32	-4.91
FF49 qxs 1 mo	2.60	14.70	0.15	0.59	1.41	0.17	-5.55
FF49 sts 12 mo	4.02	12.84	-0.11	-6.72	5.41	0.31	-4.85
FF49 sts 1 mo	6.45	11.99	0.91	3.73	6.38	0.52	-3.97
FF49 equal-weighted	8.34	16.60	-0.67	-4.46	2.56	0.48	-5.82
MAA qxs 52 wk	5.42	20.08	0.11	-10.90	9.50	0.26	-7.39
MAA qxs 4 wk	7.15	18.66	0.41	-4.74	6.48	0.37	-6.59
MAA sts 52 wk	3.98	7.74	0.14	-3.60	9.58	0.50	-8.17
MAA sts 4 wk	4.46	7.26	0.78	-1.37	9.45	0.60	-6.06
MAA equal-weighted	5.08	7.38	-0.71	-1.66	8.42	0.67	-8.62

(b) *weighted with own volatility*

strategy	mean	sd	skew	mean less med	kurtosis	sharpe	ave drawdowns
FF49 qxs 12 mo	5.78	11.06	-0.03	1.77	0.35	0.51	-4.66
FF49 qxs 1 mo	2.96	11.19	0.29	1.24	0.58	0.26	-5.04
FF49 sts 12 mo	5.46	9.64	-0.15	-4.43	1.01	0.55	-4.85
FF49 sts 1 mo	6.68	10.14	1.14	2.98	4.98	0.64	-3.57
FF49 equal-weighted	5.59	9.90	-0.83	-3.82	1.74	0.55	-5.09
MAA qxs 52 wk	6.63	10.63	-0.02	-6.19	5.47	0.60	-8.28
MAA qxs 4 wk	5.13	10.57	0.84	-3.59	5.66	0.47	-6.32
MAA sts 52 wk	9.54	10.22	0.11	-3.03	4.81	0.89	-6.98
MAA sts 4 wk	8.09	10.58	0.96	-1.90	7.16	0.74	-5.88
MAA equal-weighted	7.13	10.51	-1.40	-5.96	6.80	0.66	-9.31

(c) *weighted with underlying volatility (normalised returns)*

strategy	mean	sd	skew	mean less med	kurtosis	sharpe	ave drawdowns
FF49 qxs 12 mo	2.76	5.31	0.10	0.05	0.66	0.51	-5.04
FF49 qxs 1 mo	1.50	4.70	0.43	1.94	0.32	0.32	-5.15
FF49 sts 12 mo	2.34	4.86	-0.13	-0.64	0.86	0.48	-5.05
FF49 sts 1 mo	2.92	4.43	0.56	0.97	2.88	0.65	-4.13
FF49 equal-weighted	4.12	6.89	-0.83	-2.78	1.47	0.59	-5.31
MAA qxs 52 wk	6.00	10.65	0.03	-4.79	4.35	0.55	-8.05
MAA qxs 4 wk	7.03	10.62	0.73	4.00	3.59	0.64	-7.06
MAA sts 52 wk	3.65	4.51	-0.07	-1.58	4.71	0.80	-8.11
MAA sts 4 wk	4.44	4.36	0.83	-1.13	6.36	1.00	-6.17
MAA equal-weighted	3.45	3.86	-1.45	-2.79	6.01	0.88	-10.90

Table 7.11: ***OOS intercepts of regressions of volatility weighted strategies on unweighted strategies (robust).*** *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. Return data over the period Jul1994-Dec2012 (5Dec2002-17Apr2013) for industries (MAA) is used to obtain strategy returns and volatility estimates for the hold-out sample. In table (c) the strategies are weighted with their own volatility and in table (d) normalised returns are used. Table (a) and (b) repeat earlier results for the in-sample period.*

(a) *in-sample: weighted with own volatility vs unweighted*

industries		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	1.16(1.06)[0.290]	MAA qxs 52 wk	1.35(2.12)[0.034]
FF49 qxs 1 mo	0.41(0.53)[0.595]	MAA qxs 4 wk	0.94(1.42)[0.156]
FF49 sts 12 mo	1.77(1.48)[0.139]	MAA sts 52 wk	1.41(2.44)[0.015]
FF49 sts 1 mo	-0.56(-0.51)[0.612]	MAA sts 4 wk	0.88(1.29)[0.197]
FF49 equal-weighted	-0.99(-1.19)[0.233]	MAA equal-weighted	0.51(0.93)[0.353]

(b) *in-sample: normalised returns vs unweighted*

industries		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	2.32(3.10)[0.002]	MAA qxs 52 wk	6.33(4.59)[0.000]
FF49 qxs 1 mo	3.06(4.74)[0.000]	MAA qxs 4 wk	8.39(6.39)[0.000]
FF49 sts 12 mo	-0.23(-0.48)[0.633]	MAA sts 52 wk	2.08(4.87)[0.000]
FF49 sts 1 mo	-0.25(-0.54)[0.588]	MAA sts 4 wk	2.67(6.78)[0.000]
FF49 equal-weighted	-0.06(-0.11)[0.909]	MAA equal-weighted	1.16(3.11)[0.002]

(c) *hold-out sample: weighted with own volatility vs unweighted*

industries		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	0.51(0.37)[0.714]	MAA qxs 52 wk	1.58(0.66)[0.512]
FF49 qxs 1 mo	0.52(0.42)[0.674]	MAA qxs 4 wk	1.39(0.84)[0.401]
FF49 sts 12 mo	1.48(0.98)[0.328]	MAA sts 52 wk	1.34(0.57)[0.567]
FF49 sts 1 mo	0.45(0.30)[0.761]	MAA sts 4 wk	1.50(0.85)[0.397]
FF49 equal-weighted	0.44(0.37)[0.714]	MAA equal-weighted	0.72(0.44)[0.660]

(d) *hold-out sample: normalised returns vs unweighted*

industries		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	0.79(1.06)[0.290]	MAA qxs 52 wk	1.65(0.63)[0.528]
FF49 qxs 1 mo	0.26(0.33)[0.738]	MAA qxs 4 wk	1.34(0.65)[0.517]
FF49 sts 12 mo	0.59(0.92)[0.361]	MAA sts 52 wk	1.94(2.28)[0.023]
FF49 sts 1 mo	0.67(1.31)[0.191]	MAA sts 4 wk	1.57(2.22)[0.027]
FF49 equal-weighted	0.65(0.93)[0.355]	MAA equal-weighted	2.12(3.12)[0.002]

Table 7.12: ***OOS slopes of dispersion regressions (robust).*** *Reported are slope estimates (along with t-values in round brackets and p-values in square brackets) of regressions of returns of time-series momentum strategies on average past absolute and squared asset returns (table (a)), and returns of cross-sectional momentum strategies on past absolute and squared deviations and cross-sectional dispersion (table (b)). Asset returns over the period Jul1994-Dec2012 (5Dec2002-17Apr2013) are used for the industry (MAA) data. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used.*

(a) *Time-series*

strategy		measure	
		|returns|	$(\text{returns})^2$
FF49 12 mo	sts	-0.03(-0.84)[0.403]	0.01(0.28)[0.780]
	ults	-0.03(-1.65)[0.100]	-0.05(-1.92)[0.057]
	slts	-0.00(-1.22)[0.225]	-0.00(-0.26)[0.799]
FF49 1 mo	sts	-0.03(-0.32)[0.749]	0.16(0.32)[0.747]
	ults	0.01(0.93)[0.351]	0.08(1.35)[0.177]
	slts	-0.00(-0.57)[0.569]	-0.00(-0.04)[0.968]
MAA 52 wk	sts	-0.01(-1.28)[0.200]	-0.02(-0.93)[0.355]
	ults	-0.00(-0.01)[0.992]	-0.00(-0.44)[0.658]
	slts	-0.00(-1.36)[0.174]	-0.00(-1.14)[0.253]
MAA 4 wk	sts	0.05(1.56)[0.121]	0.28(1.42)[0.157]
	ults	0.00(0.60)[0.550]	0.01(0.26)[0.797]
	slts	0.00(0.83)[0.406]	0.03(1.56)[0.120]

(b) *Cross-sectional*

strategy		measure		
		|deviations|	$(\text{deviations})^2$	xsd
FF49 12 mo	qxs	0.02(0.23)[0.816]	-0.03(-0.22)[0.829]	-0.00(-0.05)[0.958]
	ulxs	0.01(1.35)[0.179]	-0.02(-1.39)[0.165]	0.00(0.01)[0.993]
	slxs	-0.00(-0.02)[0.985]	-0.09(-0.63)[0.531]	-0.02(-0.36)[0.716]
FF49 1 mo	qxs	-0.40(-1.79)[0.075]	-2.52(-2.03)[0.044]	-0.33(-2.01)[0.045]
	ulxs	-0.01(-0.94)[0.347]	-0.16(-4.29)[0.000]	-0.01(-1.21)[0.229]
	slxs	-0.42(-1.79)[0.075]	-3.21(-2.42)[0.017]	-0.40(-2.28)[0.024]
MAA 52 wk	qxs	-0.07(-2.31)[0.021]	-0.14(-1.99)[0.047]	-0.05(-1.89)[0.060]
	ulxs	-0.00(-0.98)[0.329]	-0.01(-1.69)[0.091]	-0.00(-1.31)[0.192]
	slxs	-0.09(-2.74)[0.006]	-0.23(-3.01)[0.003]	-0.07(-2.78)[0.006]
MAA 4 wk	qxs	0.05(0.46)[0.645]	0.41(0.60)[0.548]	0.02(0.21)[0.834]
	ulxs	0.00(0.29)[0.773]	0.02(0.77)[0.440]	-0.00(-0.07)[0.941]
	slxs	-0.06(-0.53)[0.600]	-0.54(-0.79)[0.429]	-0.05(-0.56)[0.575]

Table 7.13: ***OOS slopes of momentum vs market volatility regressions (robust).*** *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. In the first column the momentum strategy and the market are defined on unweighted returns. In the second column the strategy is run on normalised returns. In the last column the market is also based on normalised returns. The first 21 days (52 weeks) of strategy or asset returns are used for the initial volatility estimates. Asset returns for the period Jul1994-Dec2012 (5Dec2002-17Apr2013) are used for the industry (MAA) data. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used.*

strategy	unweighted | unweighted	normalised | unweighted	normalised | normalised
FF49 12 mo qxs	-0.11(-0.71)[0.479]	-0.04(-0.99)[0.325]	-0.13(-0.61)[0.546]
FF49 1 mo qxs	-0.07(-0.57)[0.570]	-0.04(-1.01)[0.314]	-0.02(-0.13)[0.898]
FF49 12 mo sts	-0.38(-2.80)[0.006]	-0.08(-1.96)[0.051]	-0.15(-0.73)[0.465]
FF49 1 mo sts	-0.07(-0.61)[0.540]	0.01(0.31)[0.753]	0.14(0.72)[0.470]
MAA 52 wk qxs	-0.91(-2.36)[0.019]	-0.49(-2.73)[0.007]	-1.29(-1.76)[0.080]
MAA 4 wk qxs	-0.03(-0.09)[0.924]	-0.12(-0.66)[0.511]	-0.64(-0.85)[0.398]
MAA 52 wk sts	-0.27(-1.84)[0.066]	-0.22(-2.87)[0.004]	-0.80(-2.55)[0.011]
MAA 4 wk sts	0.12(0.89)[0.376]	-0.05(-0.66)[0.508]	-0.07(-0.22)[0.828]

Table 7.14: ***OOS correlations of momentum and market volatility***. *Reported are estimated correlation coefficients (with t-value in round brackets and p-value in square brackets) of selected momentum strategies and ex-ante market volatility. In the first column the momentum strategy and the market are defined on unweighted returns. In the second column the strategy is run on normalised returns. In the last column the market is also based on normalised returns. The first 21 days (52 weeks) of strategy or asset returns are used for the initial volatility estimates. Asset returns for the period Jul1994-Dec2012 (5Dec2002-17Apr2013) are used for the industry (MAA) data.*

strategy	unweighted | unweighted	normalised | unweighted	normalised | normalised
FF49 12 mo qxs	-0.18	-0.10	-0.06
FF49 1 mo qxs	-0.00	-0.09	-0.02
FF49 12 mo sts	-0.10	-0.09	-0.02
FF49 1 mo sts	0.11	0.01	0.01
MAA 52 wk qxs	-0.11	-0.10	-0.05
MAA 4 wk qxs	-0.03	-0.04	-0.02
MAA 52 wk sts	-0.11	-0.08	-0.03
MAA 4 wk sts	-0.01	-0.05	-0.03

Table 7.15: ***OOS 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 period Jul1994-Dec2012 (5Dec2002-17Apr2013) are used for the industry (MAA) data.*

	weighting		
strategy	unweighted	forecast	actual
FF49 qxs 12 mo	0.32	0.41	0.47
FF49 qxs 1 mo	0.21	0.34	0.25
MAA qxs 52 wk	0.33	0.60	1.00
MAA qxs 4 wk	0.40	0.51	0.54

Table 7.16: ***OOS intercepts of regressions of dispersion weighted strategies on unweighted strategies (robust).*** *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 period Jul1994-Dec2012 (5Dec2002-17Apr2013) for industries (MAA) is used to obtain strategy returns and dispersion estimates for the hold-out sample. In table (c) the strategies are weighted with an AR(1) dispersion forecast and in table (d) with actual dispersion. In tables (a) and (b) the earlier in-sample results are repeated.*

(a) *in-sample: weighted with dispersion forecast*

industry		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	0.20(0.56)[0.578]	MAA qxs 52 wk	1.08(1.78)[0.075]
FF49 qxs 1 mo	0.10(0.40)[0.689]	MAA qxs 4 wk	0.89(1.30)[0.193]

(b) *in-sample: weighted with actual dispersion*

industry		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	0.94(1.50)[0.135]	MAA qxs 52 wk	2.15(1.89)[0.060]
FF49 qxs 1 mo	-0.26(-0.56)[0.575]	MAA qxs 4 wk	2.47(2.28)[0.023]

(c) *hold-out sample: weighted with dispersion forecast*

industry		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	0.29(0.45)[0.653]	MAA qxs 52 wk	1.80(0.82)[0.414]
FF49 qxs 1 mo	0.60(0.97)[0.334]	MAA qxs 4 wk	0.52(0.26)[0.797]

(d) *hold-out sample: weighted with actual dispersion*

industry		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	0.60(0.66)[0.511]	MAA qxs 52 wk	7.02(2.94)[0.003]
FF49 qxs 1 mo	0.55(0.82)[0.411]	MAA qxs 4 wk	2.46(1.13)[0.260]

Table 7.17: ***OOS slopes of regressions of momentum vs (log) volatility proportion of dispersion (robust).*** *Reported are slope estimates (with t-statistics in round brackets and p-values in square brackets) of regressions of momentum strategies versus the logarithm of the volatility portion of (squared) dispersion. The first column reports results for unweighted returns, the second for momentum strategies with normalised returns, and the third for both the strategy and dispersion based on normalised returns. Asset returns for the period Jul1994-Dec2012 (5Dec2002-17Apr2013) are used for the industry (MAA) data with the first 21 days (52 weeks) for each asset or strategy used for initial volatility estimates. Robust regressions with a bisquare weighting function and a parameter of 4.685 are used.*

strategy	unweighted | unweighted	normalised | unweighted	normalised | normalised
FF49 12 mo qxs	0.00(0.02)[0.987]	0.00(0.03)[0.979]	-0.00(-0.25)[0.802]
FF49 1 mo qxs	-0.01(-1.13)[0.258]	-0.00(-1.28)[0.203]	-0.00(-1.60)[0.112]
MAA 52 wk qxs	0.00(0.98)[0.328]	-0.00(-0.00)[0.998]	-0.00(-0.04)[0.967]
MAA 4 wk qxs	0.00(0.63)[0.528]	0.00(0.69)[0.489]	-0.00(-0.14)[0.892]

- The conclusion that shorter formation periods have stronger momentum has also partially reversed.
- This draws attention to the dangers of data mining. The best momentum strategy in the past can easily drop in the rankings. It also shows that quite possibly momentum is not a static phenomenon. It changes over time.
- The weaker momentum is now visible in that momentum no longer seems to outperform the market; mean terms that are more important in the linear decomposition and even some negative auto-covariances; weaker and sometimes negative predictive regression results.
- We still see a strong continuation of positive returns where time-series outperforms (positively from positive mean returns). A reversal of negative returns and negative cross-serial correlations still seem to be important where the cross-sectional strategy outperforms.
- Another mystery is the change in behaviour of the industry time-series strategies, which had a positive dependence on volatility before, but now often show a negative dependence. It is not at all clear why this has happened.
- We still see evidence that taking into account the time-series signal for a cross-sectional strategy could be useful, but not necessarily the other way round.
- Global time-series strategies continue to perform less well than local strategies.
- Volatility weighting appears to be effective for momentum strategies and is in fact more effective in the more recent period.
- Volatility weighting improves (i.e. lowers) kurtosis, but it is not clear if it improves skew and drawdowns.
- Dispersion weighting seems to be effective for cross-sectional momentum strategies, and is in fact more effective in the more recent period.
- Momentum has a negative relationship with both its own volatility and market volatility, which is clearer in the more recent period. It is, however, no longer clear that there is a negative relationship with the volatility portion of dispersion.

Chapter 8

Summary, conclusions and further research

8.1 Summary and conclusions

I set out to answer a number of questions. Below I briefly relate what I learned for each of these.

8.1.1 Is there momentum and where is it strongest?

In order to answer this question I considered six different types of momentum strategies in chapter 3 (three cross-sectional and three time-series strategies). I found positive momentum profits. However, these were only better than an equal-weighted strategy in the in-sample period.

It seemed at first that shorter formation and holding periods produced stronger momentum, but this was called into question in the out of sample period, particularly for the industry cross-sectional strategies.

It was seen that cross-sectional momentum was stronger in industries and time-series momentum stronger in MAA in the in-sample period. This was confirmed with linear regressions in section 3.3. However, the former conclusion was partially reversed in the out of sample period, where it seems conditions now favoured time-series momentum.

The linear strategies performed performed suboptimally due to a wildly varying bet size and scaling improved their performance. The latter was our first inkling that volatility weighting and dispersion weighting should be beneficial.

8.1.2 Where do time-series and cross-sectional momentum come from and where do they differ?

Theoretical

In chapter 2 I studied some of the sources of momentum theoretically. From the linear decomposition in section 2.3.1 we learnt that time-series momentum invests in auto-covariances and a square of mean returns. Cross-sectional momentum invests in auto-covariances, negative cross-serial correlations and a dispersion in mean returns. From the signed strategies in section 2.3.2 we learnt that momentum can come from a persistence in returns or in a deviation from the average return in such a way that their sign can be predicted. We also learnt that a global time-series momentum strategy is a downweighted local time-series strategy with an investment in positive cross-serial correlations (i.e. the opposite of a cross-sectional strategy) and an additional mean term.

I concluded that time-series momentum is likely to outperform (vs cross-sectional momentum) in the presence of strong auto-covariances or high mean returns and that cross-sectional momentum would outperform if lead-lag relationships are very strong and negative. We learnt that time-series and cross-sectional momentum differ where the market is mostly going up or mostly going down and that they are similar when some (about half) of the assets go up and half go down.

Empirical

We saw that in our initial sample auto-covariances were the most important source of momentum profits for MAA and the industry 1 month formation strategies. For the 12 month formation strategies it was (negative) cross-serial covariances. In the later out of sample period, however, means became a far more important source of return and were even the only source of profit for the 52 week MAA strategies. As expected time-series momentum outperformed where means or auto-covariances were most important and cross-sectional outperformed where cross-serial correlations were negative, with the exception of the industry 1 month formation strategies in our initial sample. Why the cross-sectional strategy outperformed in the latter case remains a mystery.

We did, however, see that most of the time time-series momentum and cross-sectional momentum take similar positions (meaning that mostly the market is split into a number of assets going up and and a number going down) and that this is particularly so for the MAA data where these strategies have a far closer relationship. Where these strategies do differ it is a continuation of positive returns that results in the outperformance of time-series strategies (MAA data and 1 month formation industry data in the more recent sample). Cross-sectional strategies outperform from negative cross-serial covariances, or from a reversal of negative returns in their absence (industry strategies in the in-sample period).

Momentum strategies benefit from a predictive accuracy above 50 % and this accuracy is higher than it would have been if returns were independent. The ratio of profits to losses may be an additional source of return for some strategies (this is not clear after considering the out of sample data), but is in any case quite close to 1.

Momentum seemed evident in a persistence of returns, absolute returns, deviations and absolute deviations when running predictive regressions in section 4.5. However, this persistence only remained for absolute returns and deviations out of sample.

Momentum strategies seemed to benefit more from long positions than short positions and it seemed that at least a cross-sectional strategy could be improved by taking into account the time-series signal for each asset.

We saw that global time-series strategies do not outperform local time-series strategies, probably because positive cross-serial correlations were not large enough to eliminate cross-sectional profits.

8.1.3 Does volatility weighting work?

Theoretical

In chapter 2 we saw theoretical results for two forms of volatility weighting:

(1) weighting a strategy with its own ex-ante volatility; and

(2) weighting each of the underlying assets (normalised returns).

We saw that weighting a strategy by its own volatility is likely to work if the portion of returns that do not depend on volatility is not too large and if the relationship with volatility is negative. We may expect this to hold for using normalised returns as well. weighting a strategy with its own ex-ante volatility and weighting each of the underlying assets (normalised returns).weighting a strategy with its own ex-ante volatility and weighting each of the underlying assets (normalised returns). Here I was able to provide a number of results giving somewhat strict conditions under which volatility weighting (with normalised return) is optimal.

Empirical

Volatility weighting works by stabilising volatility and, possibly, via volatility timing. There is some reason to believe that volatility weighting may reduce kurtosis and reduce skewness by stabilising volatility.

We find mostly positive evidence for the efficacy of volatility weighting, particularly in the out of sample period. Weighting strategies by their own volatility is effective particularly when the relationship with volatility is negative and using normalised returns appears to be more effective in most cases. Volatility weighting does seem to improve kurtosis, but the effect on skewness and drawdowns is not as clear. Using normalised returns reduces both the variability of volatility and dispersion and this may well be important for the efficacy of this form of volatility weighting.

8.1.4 What is the relationship of momentum with dispersion and volatility?

Theoretical

In section 2.5 we saw that if returns or deviations were persistent then one may expect time-series momentum to have a positive relationship with past absolute or squared returns and cross-sectional momentum to have a positive relationship with absolute or squared deviations. The latter are measures of dispersion.

Dispersion can also be seen as a form of volatility and thus weighting with dispersion may improve cross-sectional strategies as it is a form of volatility weighting. Dispersion and market volatility are positively related and dispersion can be decomposed into a portion consisting of conditional volatilities of assets and a portion consisting of conditional means.

Empirical

The predictions of a positive relationship with dispersion and past returns for cross-sectional and time-series strategies respectively are not borne out. There is little evidence to believe these models. In fact the relationship with dispersion is negative rather than positive (for unweighted assets).

Dispersion weighting appears to be at least somewhat effective, particularly for the MAA data and particularly in the more recent period.

The relationship of momentum with both its own volatility and market volatility appears to be mostly negative (the most notable exception is the industry 1 month time-series strategy). We also see a negative effect on predictive accuracy, Sharpe ratios and the ratio of profits to losses. Exceptions to this are the industry time-series strategies which in the in-sample period sometimes betray a positive relationship. The reason for the exceptional behaviour of the industry time-series strategies and the 1 month strategy in particular is a mystery.

It furthermore seems that this negative relationship with volatility is important (but not necessary) for the efficacy of volatility weighting, indicating a volatility timing effect.

The conjecture that the (conditional) volatility portion of dispersion is negatively related to momentum and thus responsible for the counter-intuitive relationship of momentum with dispersion seemed to hold in-sample, but no longer held in the out-of-sample period and thus this mystery remains unexplained.

8.2 Further research

Despite the wide range of theoretical and empirical analyses considered in this thesis, it is clear that much about momentum still remains a mystery. A few questions have been answered, but others have been raised. A number of robustness checks for the conclusions drawn here are also necessary for them to be generally accepted.

The models in chapter 2 are still immature. In particular the models in section 2.6, which link momentum with volatility, are infantile at best. There is a need for models that

- clarify the relationship for time-series and cross-sectional strategies for more than two assets;
- explain the effect of volatility weighting more realistically;
- provide plausible links between momentum and dispersion and momentum and market volatility and intuition for why these should be so; and that
- have a reasonable statistical fit with observed data.

Of particular interest would be to separate the effect of volatility timing and volatility stabilising when performing volatility weighting.

A particular mystery is the relationship of the industry time-series strategies with volatility, which seemed to be very clearly positive at first, but then changed.

I have not considered transaction costs at all as my main goal was merely to understand momentum. To profit from it, however, one will need to consider transaction costs.[1]

I have not studied the effectiveness of different volatility estimators. It may well be that there are better volatility estimates (for instance ones based on intraday data such as in Baltas and Kosowski (2012)) which may give better results. The sensitivity of the results to the chosen volatility estimate is also important.

For the MAA data I have not taken into account the different closing times of the markets, effectively assuming they all close at the same time. An adjustment to take this into account would be to include a lag (of at least a day) between the formation period and holding period. Given that momentum persists for some time after the formation period, I do not expect that this will affect the relationships found.

[1]Momentum may just be one of many factors considered in an overall investment strategy (notably in Robeco's MAA model) and considering the transaction costs of a momentum strategy in isolation may unfairly penalise it if there are other factors with low transaction costs that also contribute to the model.

8.3 The last word

I investigated theoretically and empirically the differences between time-series and cross-sectional momentum, explaining much (but not all) of the differing performance observed. I provided theoretical and empirical grounds for the effectiveness of volatility weighting and made an initial essay into the relationship of momentum with dispersion and volatility. Here mysteries remain, particularly for the relationship with dispersion.

Bibliography

Andreu, L., Swinkels, L. and Tjong-A-Tjoe, L. (2013). Can exchange traded funds be used to exploit industry and country momentum?, *Financial Markets and Portfolio Management* **27**(2): 127–148. [Accessed: 2013-06-03].
URL: *http://link.springer.com/10.1007/s11408-013-0207-8*

Antonacci, G. (2013a). Absolute Momentum: A Simple Rule-Based Strategy and Universal Trend-Following Overlay. [Accessed: 2013-04-24].
URL: *http://papers.ssrn.com/abstract=2244633*

Antonacci, G. (2013b). Risk Premia Harvesting Through Dual Momentum. [Accessed: 2013-01-14].
URL: *http://papers.ssrn.com/abstract=2042750*

Asness, C. S., Moskowitz, T. J. and Pedersen, L. H. (2012). Value and Momentum Everywhere. [Accessed: 2013-03-13].
URL: *http://faculty.chicagobooth.edu/tobias.moskowitz/research/papers/Value-and-Momentum-Everywhere.pdf*

Baltas, A.-N. and Kosowski, R. (2012). Improving Time-Series Momentum Strategies: The Role of Trading Signals and Volatility Estimators, *SSRN Electronic Journal* . [Accessed: 2013-01-14].
URL: *http://papers.ssrn.com/abstract=2140091*

Barroso, P. and Santa-clara, P. (2012). Managing the Risk of Momentum, *Business* (April): 1–26. [Accessed: 2012-11-16].
URL: *http://ssrn.com/paper=2041429*

Bhootra, A. (2011). Are momentum profits driven by the cross-sectional dispersion in expected stock returns?, *Journal of Financial Markets* **14**(3): 494–513. [Accessed: 2013-01-21].
URL: *http://dx.doi.org/10.1016/j.finmar.2010.12.002*

Blitz, D. C. and Van Vliet, P. (2008). Global Tactical Cross-Asset Allocation: Applying Value and Momentum Across Asset Classes, *The Journal of Portfolio Management* **35**(October 2007): 23–38. [Accessed: 2013-01-23].
URL: *http://www.iijournals.com/doi/abs/10.3905/JPM.2008.35.1.23*

Brock, W., Lakonishok, J. and LeBaron, B. (1992). Simple Technical Trading Rules and the Stochastic Properties of Stock Returns, *Journal of Finance* **47**(5): 1731–1764. [Accessed: 2013-01-21].
URL: *http://www.jstor.org/stable/2328994?origin=crossref*

Daniel, K. and Moskowitz, T. J. (2011). Momentum Crashes. [Accessed: 2013-01-08].
URL: *http://bus.miami.edu/umbfc/_common/files/papers/Daniel.pdf*

Duyvesteyn, J. and Martens, M. P. (2013). Emerging government bond market timing.

Gupta, K., Locke, S. and Scrimgeour, F. (2013). Profitability of momentum returns under alternative approaches, *International Journal of Managerial Finance* **9**(3): 4. [Accessed: 2013-04-25].
URL: *http://www.emeraldinsight.com/journals.htm?articleid=17086210*

Hallerbach, W. G. (2011). On the Expected Performance of Market Timing Strategies, *SSRN Electronic Journal* . [Accessed: 2013-01-29].
URL: *http://papers.ssrn.com/abstract=1887264*

Hallerbach, W. G. (2012). A Proof of the Optimality of Volatility Weighting Over Time, *Journal of Investment Strategies* **1**(4): 87–99. [Accessed: 2013-01-29].
URL: *http://www.ssrn.com/abstract=2008176*

Jegadeesh, N. and Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *The Journal of Finance* **48**(1): 65–91. [Accessed: 2013-01-10].
URL: *http://doi.wiley.com/10.1111/j.1540-6261.1993.tb04702.x*

Lewellen, J. (2002). Momentum and Autocorrelation in Stock Returns, *Review of Financial Studies* **15**(2): 533–564. [Accessed: 2013-01-25].
URL: *http://rfs.oupjournals.org/cgi/doi/10.1093/rfs/15.2.533*

Lillo, F., Mantegna, R. N., Bouchaud, J.-P. and Potters, M. (2001). Introducing Variety in Risk Management, *Quantitative Finance* (1): 12. [Accessed: 2013-01-29].
URL: *http://arxiv.org/abs/cond-mat/0107208*

Lo, A. W. and MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction?, *Review of Financial Studies* **3**(2): 175–205. [Accessed: 2013-01-17].
URL: *http://rfs.oupjournals.org/cgi/doi/10.1093/rfs/3.2.175*

Mandelbrot, B. and Hudson, R. (2004). *The (Mis)behaviour of Markets*, Profile Books, London.

Martin, R. (2012). Momentum trading: 'skews me, *Risk* . [Accessed: 2013-01-08].
URL: *http://www.risk.net/digital_assets/5656/risk_0812_martin.pdf*

Martin, R. and Bana, A. (2012). Non-linear momentum strategies, *Risk* pp. 60 – 65. [Accessed: 2013-01-11].
URL: *http://www.risk.net/risk-magazine/technical-paper/2221212/nonlinear-momentum-strategies*

Moskowitz, T. J. and Grinblatt, M. (1999). Do Industries Explain Momentum?, *Journal of Finance* **54**(4): 1249–1290. [Accessed: 2013-05-10].
URL: *http://www.blackwell-synergy.com/doi/abs/10.1111/0022-1082.00146*

Moskowitz, T. J., Ooi, Y. H. and Pedersen, L. H. (2012). Time series momentum, *Journal of Financial Economics* **104**(2): 228–250. [Accessed: 2013-01-08].
URL: *http://linkinghub.elsevier.com/retrieve/pii/S0304405X11002613*

Potters, M. and Bouchaud, J.-P. (2005). Trend followers lose more often than they gain, *Science* **1**(4): 7. [Accessed: 2013-01-09].
URL: *http://arxiv.org/abs/physics/0508104*

Solnik, B. and Roulet, J. (2000). Dispersion as cross-sectional correlation, *Financial Analysts Journal* . [Accessed: 2013-03-18].
URL: *http://www.jstor.org/stable/10.2307/4480223*

Stivers, C. and Sun, L. (2010). Cross-Sectional Return Dispersion and Time Variation in Value and Momentum Premiums, *Journal of Financial and Quantitative Analysis* **45**(04): 987–1014. [Accessed: 2013-02-13].
URL: *http://journals.cambridge.org/abstract_S0022109010000384*

Thomas, S. H., Clare, A. D., Smith, P. N. and Seaton, J. (2012). The Trend is Our Friend: Global Asset Allocation Using Trend Following, *SSRN Electronic Journal* . [Accessed: 2013-06-03].
URL: *http://papers.ssrn.com/abstract=2126478*

Wang, K. Q. and Xu, J. (2009). Market Volatility and Momentum, *SSRN Electronic Journal* . [Accessed: 2013-01-15].
URL: *http://papers.ssrn.com/abstract=1342719*

Wang, Y., Mesomeris, S., Salvini, M., Avettand-Fenoel, J.-R. and Beceren, M. (2013). Momentum: Riding a bumpy road.

Yu, W. and Sharaiha, Y. M. (2007). Alpha budgeting - Cross-sectional dispersion decomposed, *Journal of Asset Management* **8**(1): 58–72. [Accessed: 2013-01-19].
URL: *http://www.palgrave-journals.com/jam/journal/v8/n1/full/2250060a.html*

Appendices

Appendix A

A description of the data used

I use mainly two datasets. One is a set of asset class data for the Robeco multi-asset allocation model (MAA). The other is a set of industry data.

The MAA data consist of weekly excess returns for 17 asset classes (based on indices). These include four equity indices (US, European Union, Japan, emerging markets), 4 government bond indices (US, European Union, Japan, emerging markets), a US real estate index, two investment styles (Small minus large, Value minus growth), 2 corporate bond indices (US investment grade, US high yield), 2 commodities (oil and gold) and 2 cash indices (US and Japan). The data are from close on Wednesday to close on the next Wednesday. Data are available till 17 April 2013 and we have a different amount of history for each class.

All the asset class data was kindly provided (and compiled) by Robeco. Data are mostly constructed as follows. A total return index for the asset return under consideration is constructed, say TR_t, and a risk-free rate,[1] RF_t, is chosen covering the period t to $t+1$. The excess return is then $\frac{TR_t}{TR_{t-1}} - 1 - RF_{t-1}$. I will denote $\mathrm{ret}(TR_t) := \frac{TR_t}{TR_{t-1} - 1}$.

I indicate in table A.2 the various data series used in the construction of the excess returns and in table A.3 I give details as to the calculations, including the risk-free rate where this is explicitly subtracted and not already subsumed in the calculation of the asset return. Much of the data is constructed from more than one index and I have indicated all the indices in the calculation. The indices used change over time for many of the classes and I have indicated each of the different calculations used over time (though I have not indicated the splice dates). Most of the assets considered here can be invested in indirectly via futures contracts (which have low transaction costs). For the High yield US data credit default swap (CDS) data was used for the most recent period rather than bond indices. CDS contracts are more liquid and thus representative of actual investment opportunities, but do not have much history and so the earlier part of the series consists of an excess return of bonds over US treasuries. I provide some descriptive statistics for the data in table A.1.

The industry data I consider are the Fama and French value-weighted industry portfolios of US stocks, specifically the 49 industry portfolios, data for which can be found on Ken French's website. I construct excess returns by (multiplicatively) subtracting the 1 month T-bill rate, which can also be found on Ken French's website. That is I calculate excess returns as $r_e = \frac{1+r}{1+r_f} - 1$. Descriptive statistics of the data can be found in table A.4. Industry portfolios are hard to invest in directly (one would need to buy individual stocks), but exposure could be obtained via ETFs. Andreu *et al.* (2013) report that this can allow cross-sectional momentum profits to be captured even after transaction costs.

[1] The risk-free rate generally needs to be scaled (divided by 5200) to obtain the correct weekly rate – I take this for granted. It is also necessary for some indices to add 100 to get to a form appropriate for the calculations here – again I take this for granted.

Table A.1: ***Descriptive statistics: MAA.*** *Annualised means, annualised volatility, skew, excess kurtosis and annualised Sharpe ratios are reported. These are based on weekly returns from the start date of every series to 10 Aug 2011. Mean returns are annualised as* $r_a = (1 + r_w)^{52} - 1$ *and volatility and Sharpe ratios are annualised by multiplying by* $\sqrt{52}$.

asset class	mean	volatily	skew	kurtosis	Sharpe
Equity US	5.94	16.76	-0.71	8.64	0.34
Equity Europe	6.54	16.80	-0.64	4.94	0.38
Equity Japan	1.75	18.54	-0.23	3.37	0.09
Equity Emerging	11.79	21.58	-0.76	3.51	0.52
Real estate US	12.90	25.78	0.94	18.07	0.47
Small minus Large US	0.52	10.18	-0.41	4.57	0.05
Value minus Growth US	-1.04	8.84	0.38	4.42	-0.12
Bonds US	3.26	7.44	0.28	4.32	0.43
Bonds Europe	2.48	4.75	-0.02	2.40	0.52
Bonds Japan	3.26	4.63	-0.79	4.92	0.69
Investment grade US	-0.10	3.55	-2.59	30.29	-0.03
High yield US	1.48	8.53	-0.75	7.27	0.17
Emerging debt	9.94	13.38	-1.59	12.67	0.71
Oil	11.89	31.53	-0.04	3.38	0.36
Gold	3.14	19.41	1.11	13.22	0.16
USD cash	1.91	10.14	-0.12	3.03	0.19
JPY cash	1.52	11.08	0.91	5.43	0.14

Table A.2: ***Indices and data series used in the construction of MAA data.***

designation	index	notes
A	BARCLAYS US CORPORATE - TOT RETURN IND	
B	BARCLAYS US HY - TOT RETURN IND	
C	CDX HY UNFUNDED - TOT RETURN IND	
D	BD TOTAL 7-10 YEARS DS GOVT. INDEX - TOT RETURN IND	
E	EMU-DS Market - TOT RETURN IND (~FL)	
F	EURO STOXX 50 - DIVIDEND YIELD	
G	EURO STOXX 50 - PRICE INDEX	
H	FRANK RUSSELL 2000 (FRC) - TOT RETURN IND	
I	FTSE/NAREIT ALLREITS $ - TOT RETURN IND	
J	GERMANY EU-MARK 1M (FT/ICAP/TR) - MIDDLE RATE	
K	GERMANY EU-MARK 1WK (FT/ICAP/TR) - MIDDLE RATE	
L	JAPAN EURO-YEN 1 WK (FT/ICAP/TR) - MIDDLE RATE	
M	JAPANESE YEN TO EURO (ECB) - EXCHANGE RATE	
N	JPM EMBI+ BRADY BROAD- TOT RETURN IND	
O	JPM EMBI+ COMPOSITE - TOT RETURN IND	
P	JPM GBI GERMANY7-10Y (E) - TOT RETURN IND	Germand bond return
Q	JPM GBI JAPAN 7-10Y (Y) - TOT RETURN IND	Japanese bond return
R	JPM GBI US 7-10Y (U$) - TOT RETURN IND	US Treasury return
S	JP TOTAL 7-10 YEARS DS GOVT. INDEX - TOT RETURN IND	
T	MSCI EUROPE EX UK - DIVIDEND YIELD	
U	[ROBECO constructed DIV YIELD]	Constructed by Robeco
V	S&P 500/CITIGROUP VALUE - TOT RETURN IND	
W	S&P 500 COMPOSITE - DS TOT RETURN IND	
X	S&P 500 COMPOSITE - TOT RETURN IND	
Y	SPGI BMI US :G U$ - TOT RETURN IND	
Z	SPGI BMI US :V U$ - TOT RETURN IND	
AA	S&P GSCI Gold Excess Return - RETURN IND. (OFCL)	
AB	S&P GSCI Petroleum Excess Return - RETURN IND. (OFCL)	
AC	S&P/IFCI D COMPOSITE - TOT RETURN IND	
AD	TOPIX - TOT RETURN IND	
AE	US-DS R/E Ivst Trust - TOT RETURN IND	
AF	US EURO-$ 1 WEEK (FT/ICAP/TR) - MIDDLE RATE	
AG	US $ TO EURO (DS SYNTHETIC) - EXCHANGE RATE	
AH	US $ TO EURO (ECB) - EXCHANGE RATE	
AI	US TOTAL 7-10YEARS DS GOVT. INDEX - TOT RETURN IND	
AJ	MSCI EUROPE EXUK - PRICE INDEX	
AK	MSCI EM U$ - PRICE INDEX	
AL	S&P 500/CITIGROUP GROWTH - TOT RETURN IND	
AM	MERILL LYNCH CORPORATE MASTER INDEX	
AN	MERILL LYNCH HY MASTER II INDEX	

Table A.3: ***Construction of MAA data.*** *Indicated are the calculations to construct the excess returns for assets in the MAA dataset (calculations were done by Robeco) and the start date of each series. Where different calculations were used at different points in time these are indicated in chronological order, but the splice dates are not indicated. The data are compiled for the period 2 Jan 1975 to 17 Apr 2013. The excess return is the asset return less the risk-free rate, unless no risk-free rate is indicated, in which case the asset return calculation is already an excess return.*

asset	asset return	risk-free	start date
Equity US	$\text{ret}(W)$	AF	4 Jan 1979
	$\text{ret}(X)$	AF	
Equity Europe	$\text{ret}(E)$	J	2 Jan 1975
	$\text{ret}(AJ) + T$	J	
	$\text{ret}(G) + T$	J	
	$\text{ret}(G) + F$	J	
Equity Japan	$\text{ret}(AD)$	L	4 Aug 1978
Equity Emerging	$\text{ret}(AK) + U$	AF	31 Dec 1987
	$\text{ret}(AC)$	AF	
Real estate US	$\text{ret}(AE)$	AF	2 Jan 1975
	$\text{ret}(I)$	AF	
Small minus large US	$\text{ret}(H) - \text{ret}(W)$	–	4 Jan 1979
	$\text{ret}(H) - \text{ret}(X)$	–	
Value minus growth US	$\text{ret}(Z) - \text{ret}(Y)$	–	6 Jul 1989
	$\text{ret}(V) - \text{ret}(AL)$	–	
Bonds US	$\text{ret}(AI)$	AF	3 Jan 1980
	$\text{ret}(R)$	AF	
Bonds Europe	$\text{ret}(D)$	K	3 Jan 1980
	$\text{ret}(P)$	K	
Bonds Japan	$\text{ret}(S)$	L	31 Dec 1981
	$\text{ret}(Q)$	L	
Investment grade US	$\text{ret}(AM)$	$\text{ret}(R)$	6 Nov 1986
	$\text{ret}(A)$	$\text{ret}(R)$	
High yield US	$\text{ret}(AN)$	$\text{ret}(R)$	6 Apr 1995
	$\text{ret}(B)$	$\text{ret}(R)$	
	$\text{ret}(C)$	–	
Emerging debt	$\text{ret}(N)$	AF	3 Jan 1991
	$\text{ret}(O)$	AF	
Oil	ret(AB)	–	6 Jan 1983
Gold	ret(AA)	–	12 Jan 1978
USD cash	$(\text{ret}(\frac{1}{AG}) + 1)\frac{1+AF}{1+K}$	–	2 Jan 1975
	$(\text{ret}(\frac{1}{AH}) + 1)\frac{1+AF}{1+K}$		
JPY cash	$(\text{ret}(\frac{1}{M}) + 1)\frac{1+L}{1+K}$	–	28 Dec 1978

Table A.4: ***Descriptive statistics: Industries.*** *Annualised means, annualised volatility, skew, excess kurtosis and annualised Sharpe ratios are reported. These are based on monthly returns from July 1969 to December 2012. Mean returns are annualised as* $r_a = (1 + r_w)^{12} - 1$ *and volatility and Sharpe ratios are annualised by multiplying by* $\sqrt{12}$.

asset class	mean	volatily	skew	kurtosis	Sharpe ratio
Agric	7.58	22.69	0.00	1.83	0.32
Food	8.37	15.81	0.08	1.99	0.51
Soda	8.97	23.64	0.16	3.73	0.36
Beer	8.55	18.79	-0.06	2.19	0.44
Smoke	13.46	22.01	-0.07	2.51	0.58
Toys	4.09	25.19	-0.24	1.39	0.16
Fun	10.51	27.87	-0.18	2.98	0.36
Books	5.43	20.76	0.04	2.16	0.26
Hshld	5.18	16.75	-0.30	1.90	0.30
Clths	7.15	23.65	-0.04	2.30	0.29
Hlth	7.31	29.44	-0.10	2.66	0.24
MedEq	6.64	18.79	-0.37	1.28	0.34
Drugs	7.85	17.78	0.18	2.75	0.43
Chems	7.27	20.04	-0.13	2.22	0.35
Rubbr	6.88	21.41	-0.23	2.91	0.31
Txtls	6.24	26.07	0.60	9.67	0.23
BldMt	6.94	22.29	-0.01	4.14	0.30
Cnstr	5.81	25.96	-0.13	0.81	0.22
Steel	3.87	26.79	-0.22	2.04	0.14
FabPr	3.07	26.35	-0.14	2.55	0.11
Mach	6.45	22.60	-0.39	2.31	0.28
ElcEq	9.22	22.48	-0.20	1.60	0.39
Autos	4.69	24.75	0.29	5.80	0.19
Aero	8.86	23.84	-0.37	1.65	0.36
Ships	7.24	26.07	0.01	1.51	0.27
Guns	9.41	23.32	-0.17	2.12	0.39
Gold	6.75	37.01	0.75	5.06	0.18
Mines	7.50	25.96	-0.42	2.33	0.28
Coal	11.60	36.56	0.33	1.82	0.30
Oil	8.12	19.49	0.02	1.16	0.40
Util	5.63	14.33	-0.13	0.99	0.38
Telcm	6.04	16.74	-0.18	1.17	0.35
PerSv	1.74	24.11	-0.28	1.55	0.07
BusSv	5.56	20.49	-0.41	2.08	0.26
Hardw	5.71	25.82	-0.16	1.58	0.22
Softw	6.24	38.80	0.60	4.21	0.16
Chips	6.38	26.99	-0.33	1.38	0.23
LabEq	6.56	25.81	-0.17	1.00	0.25
Paper	6.49	19.79	0.11	2.25	0.32
Boxes	6.67	20.35	-0.41	1.95	0.32
Trans	6.05	20.93	-0.26	1.15	0.28
Whlsl	6.20	19.59	-0.35	2.36	0.31
Rtail	7.44	19.69	-0.20	1.96	0.37
Meals	7.34	21.92	-0.56	2.47	0.32
Banks	6.17	21.71	-0.27	1.94	0.28
Insur	7.03	19.82	-0.28	1.95	0.34
RlEst	0.82	27.19	0.62	9.00	0.03
Fin	7.82	22.41	-0.42	1.14	0.34
Other	-0.20	24.17	-0.49	1.48	-0.01

Appendix B

Strategy weights

For ease of reference I provide mathematical expressions for the weight $\omega_{i,t}$ given to each asset (as a proportion of notional capital) by each momentum strategy that I consider. Let N_t be the number of assets available at time t and $\mathcal{N}_t$ the set of assets available. Consider a formation period of j and denote $\bar{r}_{t-j,t} = \frac{1}{N_t}\sum_{i\in\mathcal{N}_t} r_{i,t-j,t}$ where $r_{i,t-j,t}$ is the return of asset i over the formation period.

Quantile cross-sectional (qxs)
This strategy ranks asset returns in the holding period, buys the top quantile (equally weighted) and sells the bottom quantile (equally weighted). If there are q quantiles $n_t := \lfloor \frac{N_t}{q} \rfloor$ is the number of assets in each quantile.

$$w_{i,t} = \frac{1}{n_t}\mathbf{1}_{\{\text{rank}(r_{i,t-j,t})>N_t-n_t\}} - \mathbf{1}_{\{\text{rank}(r_{i,t-j,t})\leq n_t\}}$$

(Unscaled) linear cross-sectional (ulxs)
The investment is proportional to each assets deviation from the mean return.

$$w_{i,t} = \frac{1}{N_t}(r_{i,t-j,t} - \bar{r}_{t-j,t})$$

Scaled linear cross-sectional (slxs)
The investment above is scaled so that the gross exposure is 1 in each leg.

$$w_{i,t} = \frac{1}{N_t}(r_{i,t-j,t} - \bar{r}_{t-j,t}) \times \left[\sum_{i\in\mathcal{N}_t} |r_{i,t-j,t} - \bar{r}_{t-j,t}|\right]^{-1} \times 2$$

Signed cross-sectional (sxs)
This strategy buys all the stocks above the average (and places an equal weight on them) and sells all the stocks below the average (and places an equal weight on them, which may differ from the weight on bought stocks). Denote $\tilde{r}_{i,t-j,t} := r_{i,t-j,t} - \bar{r}_{t-j,t}$.

$$\omega_{i,t} = \frac{2}{N}\left(\text{sign}(\tilde{r}_{i,t-j,t}) - \frac{1}{N}\sum_k \text{sign}(\tilde{r}_{k,t-j,t})\right)$$

Signed time-series (sts)
This strategy longs any asset with a positive return and shorts any asset with a negative return.

$$w_{i,t} = \frac{\text{sign}(r_{i,t-j,t})}{N_t}$$

(Unscaled) linear time-series(ults)
The investment is proportional to the return in the formation period.

$$w_{i,t} = \frac{1}{N_t} r_{i,t-j,t}$$

Scaled linear time-series (slts)
The investment above is scaled so that the total gross exposure is 1.

$$w_{i,t} = \frac{1}{N_t} r_{i,t-j,t} \times \left[\sum_{i \in \mathcal{N}_t} |r_{i,t-j,t}| \right]^{-1}$$

Appendix C

Return and volatility calculations

C.1 Why returns should be excess returns

Suppose we have given capital 1 and can invest in a risky asset with return r and a non-risky asset with return r_f. If we invest ω in the in the risky asset we have a return of

$$(1-\omega)r_f + \omega r$$

This is a (multiplicative) excess return of

$$\frac{(1-\omega)r_f + \omega r + 1}{1+r_f} - 1 = \omega\left(\frac{1+r}{1+r_f} - 1\right) \tag{C.1}$$

or an additive excess return of

$$(1-\omega)r_f + \omega r - r_f = \omega(r - r_f).$$

Thus if we start off with excess returns r_e for our risky asset and we assume a self-financing strategy we can calculate excess returns of our strategies as ωr_e. This can easily be extended to multiple risky assets.

For cross-sectional momentum portfolios are long-short and so (excess) returns are of the form $r_l - r_s$ (for additive returns), i.e. the difference in returns between a long and short portfolio. As long as the benchmark return used for both these returns is the same it does not matter whether excess or raw returns are used. However, for time-series strategies long and short legs are not necessarily equal and so it does matter whether returns are excess or not. For the MAA data the benchmark risk-free rates are in fact different and so it makes a difference whether returns are excess returns even for cross-sectional strategies. It is also important when positions are taken in normalised returns as then the positions in the unweighted asset returns may no longer having equal long and short legs.

C.2 Return definitions

For given asset prices P_t it is possible to model returns as either discretely compounded returns ($r_t := \frac{P_t}{P_{t-1}} - 1$) or as log returns ($r_t := \ln(\frac{P_t}{P_{t-1}})$). I mostly avoid this distinction by not modelling asset prices directly. Many of my results are applicable under either assumption, but the assumption that returns are Normal (or follow any distribution not bounded from below) is of course not strictly compatible with discrete returns, which are bounded from below. The difference between the two definitions should generally be small if returns are not too large. However, empirical work is generally (and I also do so) undertaken with discretely compounded returns.

C.3 Return calculations

In my empirical work (unless otherwise stated) I calculate returns as

$$\frac{C_{\text{end}}}{C_{\text{start}}} - 1$$

where $C_{\text{start}}, C_{\text{end}}$ are the capital at the start and end of the period respectively.

In general

$$C_{\text{end}} = C_{\text{start}}\boldsymbol{\omega}'\boldsymbol{r}$$

where $\boldsymbol{\omega}$ is the vector of portfolio weights and $\boldsymbol{r}$ the vector of asset returns over the period.

In order to do this I must assume a notional capital is invested. I always start off with a notional capital of 1 (which may then increase or decrease if more than one period is under consideration).

For cross-sectional strategies (on unweighted assets) this is equivalent to assuming a notional capital of 1 in each period and then calculating returns as

$$\frac{1 + LS_{\text{end}}}{1 + LS_{\text{start}}} - 1$$

where LS is the value of the long-short portfolio (which should never be less than -1 if this calculation is to remain valid). I calculate returns in this way for the cross-sectional strategies in chapter 3.

C.4 The trouble with negative capital

Buy and hold strategies are not considered in this paper because they sometimes result in negative capital. Consider an initially zero-investment long-short portfolio with 1 notionally invested in the risk-free asset (or some benchmark). As time carries on the size of the long and short legs tend to increase and the gross amount invested can become large compared to the 1 notional. This makes it more likely that a negative return may exceed the available capital. Should this happen it is likely that the value of the long-short position will only be slightly less than negative one, say $-1 - \epsilon$, so that the value of the capital is $-\epsilon$. In the next time period, suppose the long-short position has some return $r \geq -1$ so that its value becomes $-(1 + \epsilon)(1 + r)$, then the new return over this period would be

$$\frac{1 - (1 + \epsilon)(1 + r)}{-\epsilon} - 1 = \frac{r}{\epsilon} + r$$

This becomes larger for ϵ small, dominating any averages and moments calculated from the return series.

Ignoring (ex-post) the buy and hold strategies because they result in negative capital does introduce some bias in our research (although arguably this same bias is present in other research that consider long-short portfolios as well).

C.5 On the volatility estimates

I need make ex-ante estimates of volatility in several places in the thesis. Some explanation of and justification for these estimates is given in this section.

C.5.1 Realised variance (RV)

For a return series r_t the realised variance estimate using the using the k returns in the past period (say a month) is

$$\text{RV}_t = \sum_{i=1}^{k} r_{t-k}^2. \tag{C.2}$$

Taking the square root gives an estimate of the volatility (standard deviation). If more returns are available over a given period (as k tends to infinity), i.e. as we measure returns more finely over each month, say, we get a better estimate.[1]

For the industry data I have daily returns and thus for a period of a month the largest k I can use is 21 days. For a period of 6 months I can use 126 days. The latter technically gives a 6-month variance estimate, but scaling by $\frac{12}{126}$ one can use it as a monthly variance estimate.

A 21 day realized variance is useful for monthly returns as the variance estimates are then based on non-overlapping data[2]. Realised variances are also useful because no parameters need to be estimated. Barroso and Santa-clara (2012) use both a 21 day and 126 day RV in their paper.

[1] I shall not go into the mathematical justification for this.

[2] There may be a few months with fewer than 21 trading days available where this is not strictly true, but this is not worth fussing over.

A realised variance estimate has the drawback that it can give misleading results when an extreme datapoint drops out of the estimation window. If there was a very large return, this can result in a sudden and artificial drop in the volatility estimate when this return is at lag $k+1$. This is called a phantom effect.

C.5.2 Exponentially weighted moving average (EWMA)

An EWMA estimate is calculated as follows. Given an initial estimate at $t = 0$ (or assuming an infinite history of returns), for each t we have

$$\sigma_t = \sqrt{\lambda \sigma_{t-1}^2 + (1-\lambda) r_{t-1}^2} \tag{C.3}$$

with $0 < \lambda < 1$. The weighted average lag (or centre of mass) of the estimate is

$$\sum_{i=1}^{\infty} i(1-\lambda)\lambda^{i-1} = \frac{1}{1-\lambda} \tag{C.4}$$

I shall call this figure the *effective history* of the estimate.[3]

I calculate EWMA's for the industries using daily data and for MAA with weekly data, in which case the effective history is measured in days or weeks respectively. An estimate from daily data is technically an estimate of a daily volatility. However, scaling by, say $\sqrt{21}$ one can convert it into a monthly estimate.

An EWMA has the advantage that it does not suffer from the phantom effect mentioned in the previous subsection. However, it has a single parameter that needs to be chosen. I do not estimate this parameter from the data but instead make a pragmatic choice. The quality of the estimate should be robust to small changes in the effective history of the estimate. The parameter should be small enough to allow the estimate to adapt to the most recent changes in volatility. However, it should be large enough to avoid sudden and uninformative jumps of the estimate – i.e. to smooth out noise in the data.

For most of the thesis I use a λ of 0.9836 or an effective history of about 61 days for the industry data. This is the value used by Moskowitz *et al.* (2012) for their estimates.[4] I initialise the EWMA with a 21 day standard deviation and scale by $\sqrt{21}$. A volatility estimate is then only available from the first calendar month that starts after the first 21 days of the dataset – this is only an issue when the first month has fewer than 21 days (most months have 21 trading days).

Similarly I mostly use a λ of 0.97 or an effective history of 33⅓ weeks for the MAA data. This is the value used by Robeco in their MAA model. I initialise the estimate with a 52 week standard deviation

On occasion I shall also use a slower EWMA (i.e. with a larger parameter value), for instance as a measure of past volatility (as opposed to current volatility). Again, this is merely pragmatic.

[3]Note that by this definition, taking a simple standard deviation of the past T periods has an effective history of $\frac{T+1}{2}$. Thus an EWMA with an effective history of K should be roughly equivalent to a standard deviation based on 2K returns.

[4]They define the effective history a little differently and also use a slightly different equation for estimating the EWMA.

Appendix D

Mathematical appendix

I construct examples of processes to which proposition 2.3 can be applied. The examples are not realistic (and due to the strictness of the assumptions realism can only be marginally improved), but they are at least non-trivial in the sense that they display momentum in some form. It is of course easy to see that a process of iid standard Normal distributions and a constant variance would trivially satisfy the conditions, but this is entirely uninteresting. For convenience I restate the proposition here:

Proposition 2.3. *Suppose an asset has returns $r_t = d_t\sigma_t$ with d_t having unit variance given $\mathcal{F}_{t-1}$, σ_t strictly positive and predictable and one of the following*

- *for each t d_t (or $|d_t|$) independent of σ_t*
- *for each t $E[|d_t||\mathcal{F}_{t-1}]$ independent of σ_t*
- *for each t $E[|d_t||\mathcal{F}_{t-1}] = k_t$ for some deterministic k_t.*

Consider a signed time-series strategy as in section 2.3.2 with success and failure probabilities p and q. Consider all weightings $w_t = \frac{v_t}{\sigma_t}$ that are predictable and such that a signed time-series strategy applied to $r_t^* = w_t r_t$ has the same success and failure ratios (in particular this will hold if v_t is strictly positive). In case of the first two conditions above we also require that v_t satisfy these conditions in place of σ_t. In this case volatility weighting, corresponding to v_t being deterministic, is optimal (in that it maximises the Sharpe ratio of the strategy).

Basic setup

Consider three independent discrete time processes $(\sigma_t)_t$, $(|d|_t)_t$ and $(s_t)_t$. $(\sigma_t)_t$ can be any strictly positive process such that its mean and variance are defined. $|d|_t$ should also be non-negative: it represents the magnitude of (normalised) returns. $(s_t)_t$ should be binary with values 1 and -1. This latter process represents the sign of the returns. Now set $d_t := s_t|d|_t$, $r_t := \sigma_t d_t$, $S_t := \operatorname{sign}(r_t r_{t-1}) = s_t s_{t-1}$.

We have from the construction that S_t and $|r_t|$ are independent, which we assumed for signed strategies. We have also that $(d_t)_t$ and (σ_t) are independent, which satisfies the first optional condition in proposition 2.3.

We still need a filtration. Set $\mathcal{H}_t := \sigma(\sigma_0, ..., \sigma_{t+1})$ and $\mathcal{G}_t := \sigma(|d|_0, ..., |d|_t)$, $\mathcal{S}_t := \sigma(s_0, ..., s_t)$. Then set $\mathcal{F}_t := \mathcal{H}_t \vee \mathcal{G}_t \vee \mathcal{S}_t$. Then $(\sigma_t)_t$ is predictable and the return process adapted.

Example 1 (no persistence)

Suppose $|d|_t$ is independent of the past and that the s_t are iid taking on 1 with probability ϱ and -1 otherwise. Now we have

$$\mathrm{E}[s_t] = 2\varrho - 1 \tag{D.1}$$
$$\mathrm{Var}(d_t|\mathcal{F}_{t-1}) = \mathrm{E}[|d|_t^2] - (2\varrho - 1)^2\mathrm{E}[|d|_t]^2 =: \kappa_t^2 \tag{D.2}$$
$$\mathrm{E}[d_t] = (2\varrho - 1)\mathbf{E}[|d|_t] \tag{D.3}$$
$$\mathrm{E}[S_t] = (2\varrho - 1)^2 \tag{D.4}$$
$$\mathrm{E}[r_t^T] = (2\varrho - 1)^2\mathrm{E}[|d|_t] \tag{D.5}$$

We can assume wlog that $\kappa_t = 1$ as if it is not we simply define $|d|'_t = \frac{d_t}{\kappa_t}$ and $\sigma'_t = \sigma_t \kappa_t$. Thus we have that the conditional volatility of r_t is σ_t. A signed strategy will be profitable iff $\rho \neq \frac{1}{2}$, that is if either positive or negative returns dominate.

Example 2 (persistence in signs)

Suppose again $|d|_t$ is independent of the past. Choose $(s_t)_t$ to be a binary Markov chain with values 1 and -1 and transition probabilities defined by $\mathrm{P}(s_t = 1 | s_{t-1} = i) = 0.5 + \varrho i$ with $0 \leq \varrho < 0.5$. Initialise s_0 to be 1 or -1 with equal probability.

It can be shown (by induction) that

$$\begin{aligned}
\mathrm{P}(\{s_t = 1\}) &= 0.5(0.5 + \varrho) + 0.5(0.5 - \varrho) = 0.5 && \text{(D.6)}\\
\mathrm{E}[s_t] &= 0 && \text{(D.7)}\\
\mathrm{E}[s_t | \mathcal{F}_{t-1}] &= 2\varrho(\mathbf{1}_{\{s_{t-1}=1\}} - \mathbf{1}_{\{s_{t-1}=-1\}}) && \text{(D.8)}\\
\mathrm{Var}(d_t | \mathcal{F}_{t-1}) &= \mathrm{E}[d_t^2] - 4\varrho^2 \mathrm{E}[|d|_t]^2 =: \kappa_t^2 && \text{(D.9)}\\
\mathrm{E}[d_t] &= 0 && \text{(D.10)}\\
\mathrm{E}[S_t] &= 0.5(0.5 + \varrho) - 0.5(0.5 - \varrho) + 0.5(0.5 + \varrho) - 0.5(0.5 - \varrho) = 2\varrho \geq 0 && \text{(D.11)}\\
\mathrm{E}[r_t^T] &= 2\varrho \mathrm{E}[|d|_t] && \text{(D.12)}
\end{aligned}$$

We can again choose $|d|_t$ such that $\kappa_t = 1$ and thus σ_t is the conditional variance of the return process. In this example the momentum comes from the sign of returns being persistent and this results in a profitable time-series strategy despite the returns having a zero mean.

Example 3 (persistence in magnitude)

Choose s_t as in the previous example. Now suppose $|d|_t$ is a positive AR(1) process with $|d|_t = c + \phi |d|_{t-1} + \eta_t$. (It does not really matter whether it is stationary or not, but this would be more convenient.) $(\eta_t)_t$ is a white noise process with unit variance independent of the past. We have

$$\begin{aligned}
\mathrm{P}(\{s_t = 1\}) &= 0.5(0.5 + \varrho) + 0.5(0.5 - \varrho) = 0.5 && \text{(D.13)}\\
\mathrm{E}[s_t] &= 0 && \text{(D.14)}\\
\mathrm{E}[s_t | \mathcal{F}_{t-1}] &= 2\varrho(\mathbf{1}_{\{s_{t-1}=1\}} - \mathbf{1}_{\{s_{t-1}=-1\}}) && \text{(D.15)}\\
\mathrm{Var}(d_t | \mathcal{F}_{t-1}) &= 1 + (c + \phi |d|_{t-1})^2 (1 - 4\varrho^2) =: \kappa_t^2 && \text{(D.16)}\\
\mathrm{Var}(|d|_t | \mathcal{F}_{t-1}) &= 1 && \text{(D.17)}\\
\mathrm{E}[d_t] &= 0 && \text{(D.18)}\\
\mathrm{E}[S_t] &= 0.5(0.5 + \varrho) - 0.5(0.5 - \varrho) + 0.5(0.5 + \varrho) - 0.5(0.5 - \varrho) = 2\varrho \geq 0 && \text{(D.19)}\\
\mathrm{E}[r_t^T] &= 2\varrho \mathrm{E}[|d|_t] && \text{(D.20)}
\end{aligned}$$

Note that in this example the conditional variance of d_t is not constant and cannot simply be made to be constant. However, the conditional variance of $|d_t|$ is 1 and proposition 2.3 still applies if we are satisfied to weight by the volatility of $|r_t| = \sigma_t |d_t|$. In this example the momentum comes from both the sign of returns and their magnitude being persistent. However, the strategy does not capitalise on the latter persistence as it does not adjust the size of the investment to take into account the size of the past returns.

Appendix E

OLS regression results

Here I include corresponding ordinary least squares results for all the robust regression results reported in the main text.

E.1 Relationship between the strategies

Table E.1: ***Intercepts of regressions of momentum strategies on each other (OLS).*** *Intercepts (alphas) and their p-values (in square brackets) are reported for ordinary least squares regressions of monthly (weekly) momentum strategy returns on each other. The format for the column headings is "var1-var2" for var1 regressed on var2. The intercept of the opposite regression "var2-var1" is also reported in each case. In table (a) one of var1 and var2 is a time-series strategy and one is a cross-sectional strategy. In table (b) one of the variables is an unscaled linear strategy and the other a scaled linear strategy. In table (c) one of the variables is a non-linear strategy (either quantile cross-sectional or signed time-series) and the other is a scaled or unscaled linear strategy. The strategies were run with industry (MAA) data over the period Jul1969-Jun1994 (4Jan1979-5Dec2002). Returns are annualised by multiplying by 12 (52) and expressed in per cent.*

(a) *XS vs TS*

	regression					
formation	qxs-sts	sts-qxs	ulxs-ults	ults-ulxs	slxs-slts	slts-slxs
FF49 12 mo	9.80[0.000]	-5.00[0.035]	0.51[0.002]	-1.32[0.098]	7.51[0.001]	-0.07[0.202]
FF49 1 mo	8.80[0.000]	1.87[0.453]	0.08[0.044]	0.24[0.234]	6.22[0.004]	0.06[0.321]
MAA 52 wk	-1.44[0.286]	1.54[0.002]	-0.10[0.300]	0.17[0.128]	3.63[0.070]	-0.03[0.799]
MAA 4 wk	-1.22[0.372]	1.92[0.000]	-0.07[0.001]	0.09[0.000]	1.18[0.479]	0.10[0.256]

(b) *Scaled vs unscaled*

	regression			
formation	slxs-ulxs	ulxs-slxs	slts-ults	ults-slts
formation	slxs-ulxs	ulxs-slxs	slts-ults	ults-slts
FF49 12 mo	2.62[0.000]	-0.15[0.001]	0.09[0.001]	-1.06[0.002]
FF49 1 mo	2.65[0.001]	-0.03[0.027]	0.04[0.225]	0.02[0.840]
MAA 52 wk	6.74[0.000]	-0.45[0.006]	0.16[0.013]	-0.23[0.078]
MAA 4 wk	11.29[0.000]	-0.21[0.000]	0.46[0.000]	-0.16[0.009]

Table E.1: *(continued from previous page)*

(c) *Linear vs non-linear*

	regression			
formation	qxs-ulxs	ulxs-qxs	qxs-slxs	slxs-qxs
FF49 12 mo	3.83[0.000]	-0.21[0.001]	1.49[0.055]	-0.77[0.349]
FF49 1 mo	5.24[0.000]	-0.06[0.009]	3.02[0.000]	-1.90[0.043]
MAA 52 wk	8.56[0.000]	-0.57[0.017]	3.90[0.005]	-1.89[0.312]
MAA 4 wk	12.23[0.000]	-0.23[0.001]	4.64[0.000]	-2.33[0.197]

	regression			
formation	sts-ults	ults-sts	sts-slts	slts-sts
FF49 12 mo	1.21[0.161]	-0.38[0.186]	-2.12[0.000]	0.06[0.000]
FF49 1 mo	0.66[0.602]	0.08[0.474]	-0.68[0.256]	0.03[0.085]
MAA 52 wk	3.99[0.000]	-0.81[0.008]	3.39[0.000]	-0.28[0.045]
MAA 4 wk	5.23[0.000]	-0.28[0.001]	3.58[0.000]	-0.12[0.369]

E.2 Predictive regressions

Table E.2: ***Average scaled slopes of predictive regressions (OLS).*** *Average slopes, t-statistics (in round brackets) and p-values (in square brackets) are reported for per-asset predictive OLS regressions for various metrics (returns and absolute returns, deviations and absolute deviations) in the holding period versus the same metric in the formation period. Data for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for industries (MAA) are used. A second table reports results for normalised returns where the first 21 days (52 weeks) of data for each asset is used for the initial EWMA volatility estimates.*

(a) *Unweighted returns*

	metric			
regression	returns	|returns|	deviations	|deviations|
FF49 1 mo vs 12 mo	-0.11(-0.50)[0.455]	0.45(2.35)[0.110]	0.12(0.80)[0.396]	0.30(1.74)[0.216]
FF49 1 mo vs 1 mo	0.09(1.55)[0.269]	0.05(0.90)[0.379]	0.07(1.21)[0.224]	0.08(1.49)[0.281]
MAA 1 wk vs 52 wk	0.20(1.10)[0.370]	1.19(5.48)[0.010]	0.17(0.59)[0.530]	0.51(1.42)[0.215]
MAA 1 wk vs 4 wk	0.14(2.42)[0.141]	0.43(7.35)[0.000]	0.04(0.43)[0.407]	0.32(3.21)[0.106]

(b) *Normalised returns*

	metric			
regression	returns	|returns|	deviations	|deviations|
FF49 1 mo vs 12 mo	-0.04(-0.12)[0.495]	0.18(0.91)[0.417]	0.20(1.30)[0.298]	0.23(1.39)[0.297]
FF49 1 mo vs 1 mo	0.09(1.54)[0.232]	-0.05(-0.87)[0.414]	0.09(1.59)[0.211]	0.03(0.52)[0.426]
MAA 1 wk vs 52 wk	0.29(1.66)[0.187]	0.77(2.71)[0.162]	0.17(0.59)[0.499]	0.73(1.75)[0.276]
MAA 1 wk vs 4 wk	0.18(3.08)[0.093]	0.23(3.78)[0.011]	0.06(0.70)[0.269]	0.20(1.88)[0.243]

E.3 Volatility weighting

Table E.3: ***Predictability of volatility (OLS).*** *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 realised variance from the past 21 daily returns of these strategies. Industry data for the period Jul1969-Jun1994 are used to compute the strategy returns.*

strategy	slope	R-squared
12 mo qxs	0.49(9.44)[0.000]	23.80
1 mo qxs	0.47(9.32)[0.000]	22.70
12 mo sts	0.44(8.29)[0.000]	19.42
1 mo sts	0.44(8.49)[0.000]	19.57
equal-weighted	0.44(8.48)[0.000]	19.50

Table E.4: ***Strategy relationship with own volatility (OLS).*** *Reported are estimates of α (scaled by 100) and γ as in (5.1) and the estimated error variance of the regression (5.1) for selected momentum strategies and an equal-weighted market. Return data over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data to estimate strategy returns, the first 21 days (52 weeks) of which are used to provide initial volatility estimates. OLS regressions are used.*

strategy	γ	100α	error var
FF49 12 mo qxs	-0.53(-1.47)[0.142]	2.15(2.76)[0.006]	2.42
FF49 1 mo qxs	-0.09(-0.30)[0.763]	1.08(1.79)[0.075]	1.62
FF49 12 mo sts	-0.32(-1.81)[0.071]	1.05(3.15)[0.002]	2.02
FF49 1 mo sts	0.40(1.82)[0.069]	-0.52(-1.07)[0.287]	1.79
FF49 equal-weighted	0.39(1.31)[0.192]	-1.09(-1.07)[0.284]	2.03
MAA 52 wk qxs	-0.08(-0.72)[0.475]	0.41(2.24)[0.025]	0.98
MAA 4 wk qxs	-0.25(-2.21)[0.027]	0.75(3.90)[0.000]	1.00
MAA 52 wk sts	-0.18(-1.50)[0.133]	0.24(3.04)[0.002]	0.98
MAA 4 wk sts	-0.16(-1.42)[0.156]	0.25(3.48)[0.001]	1.00
MAA equal-weighted	0.13(1.02)[0.310]	-0.04(-0.39)[0.694]	1.06

Table E.5: ***Intercepts of regressions of volatility weighted strategies on unweighted strategies (OLS).*** *Reported are intercepts (alphas) for OLS regressions of volatility weighted strategies vs unweighted strategies. In table (a) the strategies are weighted with their own volatility and in table (b) normalised returns are used. Return data over the period Jul1969-30Jun1994 (4Jan1979-5Dec2002) for industries (MAA) is used to obtain strategy returns and volatility estimates.*

(a) *weighted with own volatility vs unweighted*

industries		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	3.11(3.32)[0.001]	MAA qxs 52 wk	1.74(3.02)[0.003]
FF49 qxs 1 mo	1.68(2.43)[0.016]	MAA qxs 4 wk	2.43(4.09)[0.000]
FF49 sts 12 mo	4.32(4.11)[0.000]	MAA sts 52 wk	1.98(3.75)[0.000]
FF49 sts 1 mo	0.23(0.23)[0.815]	MAA sts 4 wk	2.54(4.17)[0.000]
FF49 equal-weighted	-1.18(-1.56)[0.119]	MAA equal-weighted	-0.29(-0.58)[0.563]

(b) *normalised returns vs unweighted*

industries		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	2.95(4.21)[0.000]	MAA qxs 52 wk	6.08(4.68)[0.000]
FF49 qxs 1 mo	2.99(4.90)[0.000]	MAA qxs 4 wk	9.26(7.45)[0.000]
FF49 sts 12 mo	0.48(1.09)[0.278]	MAA sts 52 wk	2.06(5.13)[0.000]
FF49 sts 1 mo	0.40(0.96)[0.338]	MAA sts 4 wk	3.05(8.23)[0.000]
FF49 equal-weighted	-0.31(-0.60)[0.550]	MAA equal-weighted	1.13(3.19)[0.001]

E.4 Relationship with dispersion and volatility

Table E.6: ***Slopes of dispersion regressions (OLS).*** *Reported are slope estimates (along with t-values in round brackets and p-values in square brackets) of OLS regressions of returns of time-series momentum strategies on average past absolute and squared asset returns (table (a)), and returns of cross-sectional momentum strategies on past absolute and squared deviations and cross-sectional standard deviation (table (b)). Asset returns over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data.*

(a) *Time-series*

strategy		measure: |returns|	measure: (returns)^2
FF49 12 mo	sts	-0.06(-2.29)[0.023]	-0.07(-1.93)[0.054]
	ults	-0.02(-2.53)[0.012]	-0.03(-2.43)[0.016]
	slts	-0.00(-2.46)[0.014]	-0.00(-2.14)[0.033]
FF49 1 mo	sts	0.11(1.62)[0.105]	0.61(1.90)[0.058]
	ults	0.03(4.77)[0.000]	0.15(5.70)[0.000]
	slts	0.00(1.23)[0.221]	0.01(1.40)[0.162]
MAA 52 wk	sts	-0.00(-0.71)[0.479]	-0.01(-1.02)[0.307]
	ults	-0.00(-1.62)[0.104]	-0.01(-3.08)[0.002]
	slts	-0.00(-1.61)[0.107]	-0.00(-2.39)[0.017]
MAA 4 wk	sts	-0.02(-0.91)[0.362]	-0.09(-1.14)[0.254]
	ults	-0.01(-3.25)[0.001]	-0.07(-8.97)[0.000]
	slts	-0.00(-1.04)[0.299]	-0.05(-3.15)[0.002]

(b) *Cross-sectional*

strategy		measure: |deviations|	measure: (deviations)^2	measure: xsd
FF49 12 mo	qxs	-0.13(-1.83)[0.068]	-0.24(-1.75)[0.082]	-0.08(-1.67)[0.096]
	ulxs	-0.01(-1.96)[0.050]	-0.02(-2.11)[0.035]	-0.01(-1.87)[0.063]
	slxs	-0.15(-2.00)[0.047]	-0.31(-2.16)[0.032]	-0.11(-2.03)[0.044]
FF49 1 mo	qxs	-0.07(-0.37)[0.709]	-1.25(-0.94)[0.350]	-0.09(-0.64)[0.524]
	ulxs	-0.00(-0.30)[0.763]	-0.07(-2.44)[0.015]	-0.00(-1.18)[0.240]
	slxs	-0.18(-0.79)[0.431]	-2.86(-1.86)[0.065]	-0.20(-1.22)[0.223]
MAA 52 wk	qxs	-0.01(-0.92)[0.355]	-0.02(-0.83)[0.409]	-0.01(-0.54)[0.591]
	ulxs	-0.00(-2.48)[0.013]	-0.01(-4.27)[0.000]	-0.00(-2.66)[0.008]
	slxs	-0.03(-1.76)[0.078]	-0.07(-2.56)[0.011]	-0.02(-1.85)[0.065]
MAA 4 wk	qxs	-0.05(-1.07)[0.287]	-0.39(-1.60)[0.111]	-0.02(-0.67)[0.505]
	ulxs	-0.01(-5.31)[0.000]	-0.10(-12.62)[0.000]	-0.01(-5.76)[0.000]
	slxs	-0.17(-2.60)[0.009]	-1.54(-4.58)[0.000]	-0.10(-2.25)[0.024]

Table E.7: ***Slopes of dispersion regressions with normalised returns (OLS).*** *Reported are slope estimates (along with t-values in round brackets and p-values in square brackets) of OLS regressions of returns of time-series momentum strategies on average past absolute and squared asset returns (table (a)), and returns of cross-sectional momentum strategies on past absolute and squared deviations and cross-sectional standard deviation (table (b)). The strategies are based on normalised returns. Asset returns over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data and the first 21 days (52 weeks) for each asset are used for initial volatility estimates.*

(a) *Time-series*

strategy		measure	
		|returns|	$(\text{returns})^2$
FF49 12 mo	sts	-0.07(-2.41)[0.017]	-0.19(-2.29)[0.023]
	ults	-0.01(-2.55)[0.011]	-0.04(-2.57)[0.011]
	slts	-0.00(-2.34)[0.020]	-0.00(-2.23)[0.027]
FF49 1 mo	sts	0.06(0.92)[0.357]	0.48(0.92)[0.361]
	ults	0.01(3.29)[0.001]	0.07(3.37)[0.001]
	slts	0.00(0.60)[0.550]	0.01(0.62)[0.538]
MAA 52 wk	sts	0.00(1.16)[0.246]	0.00(0.33)[0.740]
	ults	0.00(2.46)[0.014]	0.00(1.86)[0.063]
	slts	0.00(1.02)[0.306]	0.00(0.60)[0.549]
MAA 4 wk	sts	0.01(0.64)[0.520]	0.06(0.35)[0.724]
	ults	0.00(3.30)[0.001]	0.02(2.30)[0.022]
	slts	0.00(0.81)[0.417]	0.00(0.31)[0.760]

(b) *Cross-sectional*

strategy		measure		
		|deviations|	$(\text{deviations})^2$	xsd
FF49 12 mo	qxs	0.02(0.29)[0.772]	-0.03(-0.10)[0.919]	0.01(0.24)[0.809]
	ulxs	0.00(0.66)[0.509]	0.00(0.23)[0.817]	0.00(0.62)[0.537]
	slxs	0.01(0.16)[0.873]	-0.06(-0.24)[0.812]	0.00(0.09)[0.927]
FF49 1 mo	qxs	0.30(1.64)[0.101]	3.39(1.11)[0.267]	0.19(1.32)[0.189]
	ulxs	0.01(3.47)[0.001]	0.08(2.98)[0.003]	0.00(3.14)[0.002]
	slxs	0.34(1.90)[0.058]	4.23(1.43)[0.155]	0.24(1.66)[0.097]
MAA 52 wk	qxs	0.01(0.75)[0.453]	0.01(0.18)[0.853]	0.00(0.34)[0.736]
	ulxs	0.00(1.70)[0.089]	0.00(0.93)[0.351]	0.00(1.19)[0.236]
	slxs	0.00(0.30)[0.763]	-0.01(-0.31)[0.758]	-0.00(-0.14)[0.891]
MAA 4 wk	qxs	0.01(0.30)[0.767]	0.19(0.41)[0.680]	0.01(0.39)[0.697]
	ulxs	0.00(2.17)[0.031]	0.01(1.97)[0.049]	0.00(2.45)[0.015]
	slxs	0.00(0.04)[0.971]	0.13(0.28)[0.781]	0.01(0.29)[0.776]

Table E.8: ***Slopes of momentum vs market volatility regressions (OLS).*** *Reported are slope coefficients (with t-value in round brackets and p-value in square brackets) for OLS regressions of selected momentum strategies against ex-ante market volatility. In the first column the momentum strategy and the market are defined on unweighted returns. In the second column the strategy is run on normalised returns. In the last column the market is also based on normalised returns. The first 21 days (52 weeks) of strategy or asset returns are used for the initial volatility estimate. Asset returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data.*

strategy	unweighted | unweighted	normalised | unweighted	normalised | normalised
FF49 12 mo qxs	-0.39(-2.41)[0.017]	-0.20(-2.53)[0.012]	-0.17(-0.93)[0.353]
FF49 1 mo qxs	-0.31(-2.63)[0.009]	-0.11(-1.94)[0.053]	-0.16(-1.19)[0.233]
FF49 12 mo sts	-0.25(-1.51)[0.132]	-0.09(-0.89)[0.373]	0.02(0.10)[0.923]
FF49 1 mo sts	0.05(0.31)[0.754]	-0.01(-0.14)[0.889]	0.00(0.02)[0.986]
MAA 52 wk qxs	-0.63(-1.93)[0.054]	-0.58(-2.68)[0.007]	-1.25(-2.71)[0.007]
MAA 4 wk qxs	-0.53(-1.74)[0.081]	-0.39(-1.99)[0.047]	-0.63(-1.41)[0.158]
MAA 52 wk sts	-0.20(-1.59)[0.112]	-0.26(-2.90)[0.004]	-0.43(-2.14)[0.033]
MAA 4 wk sts	-0.13(-1.09)[0.277]	-0.14(-1.67)[0.096]	-0.33(-1.78)[0.076]

Table E.9: ***Predictability of dispersion (OLS).****Reported are the estimated AR(1) coefficients of dispersion (defined as the cross-sectional standard deviation of stock returns in each period) for two datasets. Asset returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data. OLS regressions are used.*

strategy	slope	R-squared
FF49	0.37(6.80)[0.000]	13.49
MAA	0.36(13.54)[0.000]	12.84

Table E.10: ***Intercepts of regressions of dispersion weighted strategies on unweighted strategies (OLS).*** *Reported are intercepts (alphas) for OLS regressions of dispersion weighted strategies vs unweighted strategies. In table (a) the strategies are weighted with an AR(1) dispersion forecast and in table (b) with actual dispersion. Return data over the period Jul1969-30Jun1994 (4Jan1979-5Dec2002) for industries (MAA) is used to obtain strategy returns and dispersion estimates.*

(a) *weighted with dispersion forecast*

industry		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	0.44(1.42)[0.156]	MAA qxs 52 wk	1.84(3.33)[0.001]
FF49 qxs 1 mo	0.26(1.16)[0.249]	MAA qxs 4 wk	2.48(4.19)[0.000]

(b) *weighted with actual dispersion*

industry		MAA	
strategy	intercept	strategy	intercept
FF49 qxs 12 mo	1.77(3.31)[0.001]	MAA qxs 52 wk	4.91(5.16)[0.000]
FF49 qxs 1 mo	0.41(1.05)[0.292]	MAA qxs 4 wk	3.93(4.31)[0.000]

Table E.11: ***Dispersion vs volatility regressions (OLS).*** *Reported are slope estimates (with t-statistics in round brackets and p-values in square brackets) for OLS regressions of dispersion versus market volatility. Dispersion and market volatility are calculated based either on unweighted or normalised returns. Asset returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data.*

(a) *Dispersion (unweighted returns) vs market volatility (unweighted returns)*

dataset	slope	R-squared
FF49	0.31(6.83)[0.000]	13.59
MAA	1.04(7.99)[0.000]	5.08

(b) *Dispersion (normalised returns) vs market volatility (unweighted returns)*

dataset	slope	R-squared
FF49	-0.17(-8.31)[0.000]	18.86
MAA	-0.13(-1.69)[0.090]	0.24

(c) *Dispersion (normalised returns) vs market volatility (normalised returns)*

dataset	slope	R-squared
FF49	-0.29(-5.79)[0.000]	10.16
MAA	0.38(2.14)[0.033]	0.40

Table E.12: ***Slopes of regressions of momentum vs (log) volatility proportion of dispersion (OLS).*** *Reported are slope estimates (with t-statistics in round brackets and p-values in square brackets) for OLS regressions of momentum strategies versus the logarithm of the volatility portion of (squared) dispersion. The first column reports results for unweighted returns, the second for momentum strategies with normalised returns, and the third for both the strategy and dispersion based on normalised returns. Asset returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for the industry (MAA) data with the first 21 days (52 weeks) for each asset or strategy used for initial volatility estimates.*

strategy	unweighted | unweighted	normalised | unweighted	normalised | normalised
FF49 12 mo qxs	0.00(0.23)[0.815]	-0.00(-0.63)[0.527]	-0.01(-2.38)[0.018]
FF49 1 mo qxs	-0.01(-1.70)[0.090]	-0.00(-1.70)[0.090]	-0.01(-2.57)[0.011]
MAA 52 wk qxs	0.00(0.24)[0.809]	-0.00(-1.13)[0.258]	0.00(0.39)[0.700]
MAA 4 wk qxs	-0.00(-2.00)[0.046]	-0.00(-3.67)[0.000]	-0.00(-4.18)[0.000]

Appendix F

Normalised returns

Here I report results for normalised returns corresponding to the results for unweighted returns in chapter 4 and which I decided not to include in the main text in section 5.2.2.

Table F.1: ***Sharpe ratios varying formation and holding period with normalised returns.*** *Annualised Sharpe ratios are calculated from monthly (weekly) returns for six different momentum strategies over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for the industry (MAA) data considering every possible combination of J month (week) formation and K month (week) holding period. Returns are normalised and the first 21 days (52 weeks) of data for each asset are used for the initial volatility estimates. Sharpe ratios are reported for the complete holding period. Holdings are rebalanced monthly (weekly) according to the weights assigned by each strategy. Sharpe ratios are annualised by multiplying by* $\sqrt{\frac{12}{K}}$ ($\sqrt{\frac{52}{K}}$).

(a) *Industries*

		K			
strategy	J	1	3	6	12
FF49 qxs	1	1.44	0.51	0.36	0.43
	3	0.82	0.44	0.43	0.60
	6	0.89	0.55	0.66	0.66
	12	1.12	0.95	0.84	0.62
FF49 ulxs	1	1.41	0.49	0.38	0.40
	3	0.77	0.44	0.40	0.54
	6	0.76	0.50	0.63	0.64
	12	0.95	0.80	0.75	0.55
FF49 slxs	1	1.45	0.53	0.42	0.46
	3	0.79	0.45	0.42	0.59
	6	0.80	0.53	0.65	0.67
	12	1.09	0.91	0.82	0.62
FF49 sts	1	0.53	0.09	0.06	0.05
	3	0.18	0.01	-0.04	0.03
	6	0.08	-0.05	-0.05	0.02
	12	0.17	0.11	0.10	-0.01
FF49 ults	1	0.62	0.10	0.05	0.04
	3	0.19	0.04	-0.04	0.03
	6	0.12	-0.07	-0.07	0.02
	12	0.10	0.03	0.03	-0.02
FF49 slts	1	0.65	0.15	0.09	0.09
	3	0.22	0.04	0.01	0.08
	6	0.16	0.04	0.03	0.09
	12	0.32	0.22	0.18	0.05

(b) *MAA*

		K				
strategy	J	1	4	8	26	52
MAA qxs	1	0.84	0.84	0.52	0.21	0.23
	4	1.84	1.23	0.75	0.27	0.33
	8	1.43	0.98	0.68	0.28	0.41
	26	0.95	0.73	0.57	0.54	0.51
	52	1.40	1.11	0.91	0.73	0.50
MAA ulxs	1	0.74	0.77	0.49	0.17	0.20
	4	1.30	1.02	0.64	0.20	0.31
	8	1.14	0.86	0.56	0.19	0.37
	26	0.74	0.53	0.38	0.53	0.50
	52	1.23	1.01	0.87	0.71	0.49
MAA slxs	1	1.00	0.95	0.56	0.23	0.24
	4	1.79	1.19	0.73	0.29	0.37
	8	1.48	1.02	0.68	0.27	0.43
	26	0.88	0.65	0.52	0.59	0.54
	52	1.34	1.12	0.95	0.70	0.46
MAA sts	1	1.39	0.98	0.65	0.27	0.26
	4	1.99	1.28	0.83	0.34	0.38
	8	1.71	1.20	0.80	0.34	0.43
	26	1.26	0.98	0.80	0.67	0.47
	52	1.51	1.27	1.04	0.58	0.35
MAA ults	1	1.06	0.93	0.60	0.24	0.24
	4	1.54	1.17	0.77	0.30	0.36
	8	1.44	1.09	0.75	0.30	0.42
	26	1.08	0.78	0.60	0.58	0.47
	52	1.36	1.03	0.86	0.58	0.38
MAA slts	1	1.44	1.05	0.64	0.29	0.24
	4	2.07	1.30	0.85	0.39	0.40
	8	1.72	1.21	0.85	0.39	0.47
	26	1.24	0.92	0.74	0.66	0.50
	52	1.45	1.10	0.93	0.63	0.38

Table F.2: ***Long-short analysis with normalised returns.*** *The* long and short portolfolio tables *report returns, Sharpe ratios and the average proportion of assets in each of four portfolios dividing assets into those that a times-series and two quantile strategy would buy or sell. The* two quantile strategies *tables divide assets into a top and bottom half (possible omitting one asset) each formation period and reports the returns and Sharpe ratio for each half in three cases (both quantiles going up in the formation period; one going up and one down; and both going down) as in figure 2.2 and the proportion of occurrences for each case. The* four scenarios *tables reports time-series and cross-sectional Sharpe ratios for each of four scenarios (in up|up and down|down cases) as defined in figures 2.3 and 2.4 along with the proportion of occurrences of each scenario. Industry (MAA) data for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) is used. Returns are normalised with the first 21 days (52 weeks) of data used for initial volatility estimates for each asset. Where there is not enough data to calculate an entry, this is indicated by* NaN.

(a) *Industry 12 month formation*

(a1) *Long and short portfolios*

portfolio	return	Sharpe	proportion
long|long	6.72	0.61	0.39
long|short	1.98	0.18	0.19
short|long	4.95	0.46	0.10
short|short	0.37	0.03	0.30

(a2) *Two quantile strategies*

case	portfolio	return	Sharpe	proportion
up|up	top	0.78	0.07	0.37
	bottom	-3.63	-0.35	
up|down	top	10.22	0.94	0.44
	bottom	4.85	0.43	
down|down	top	5.87	0.63	0.20
	bottom	2.06	0.18	

(a3) *Two quantile strategies: four scenarios*

		scenario			
case	statistic	sc1	sc2	sc3	sc4
up|up	proportion	0.43	0.04	0.08	0.46
	TS Sharpe	1.46	0.20	0.02	-0.94
	XS Sharpe	0.61	-1.36	1.84	0.06
down|down	proportion	0.46	0.02	0.05	0.47
	TS Sharpe	2.00	NaN	0.19	-1.51
	XS Sharpe	1.04	NaN	3.31	-0.33

(b) *Industry 1 month formation*

(b1) *Long and short portfolios*

portfolio	return	Sharpe	proportion
long|long	5.22	0.49	0.35
long|short	6.59	0.76	0.16
short|long	1.37	0.10	0.14
short|short	-0.86	-0.07	0.33

(b2) *Two quantile strategies*

case	portfolio	return	Sharpe	proportion
up|up	top	9.89	1.27	0.33
	bottom	6.05	0.78	
up|down	top	1.47	0.12	0.43
	bottom	-3.91	-0.31	
down|down	top	6.20	0.56	0.24
	bottom	0.23	0.02	

(b3) *Two quantile strategies: four scenarios*

		scenario			
case	statistic	sc1	sc2	sc3	sc4
up|up	proportion	0.57	0.02	0.06	0.35
	TS Sharpe	1.68	-0.04	-0.14	-1.70
	XS Sharpe	0.30	-1.35	1.13	0.33
down|down	proportion	0.48	0.00	0.07	0.45
	TS Sharpe	1.53	NaN	-0.63	-1.65
	XS Sharpe	0.72	NaN	2.85	-0.00

Table F.2: *Long-short analysis with normalised returns.* *(continued from previous page)*

(c) *MAA 52 week formation*

(c1) *Long and short portfolios*

portfolio	return	Sharpe	proportion
long|long	7.95	1.41	0.44
long|short	2.30	0.33	0.11
short|long	-0.32	-0.04	0.05
short|short	-3.67	-0.63	0.37

(c2) *Two quantile strategies*

case	portfolio	return	Sharpe	proportion
up|up	top	12.90	2.51	0.13
	bottom	4.43	0.96	
up|down	top	7.36	1.40	0.84
	bottom	-2.76	-0.53	
down|down	top	-1.86	-0.53	0.03
	bottom	-3.99	-0.76	

(c3) *Two quantile strategies: four scenarios*

		scenario			
case	statistic	sc1	sc2	sc3	sc4
up|up	proportion	0.38	0.14	0.29	0.19
	TS Sharpe	1.88	0.19	0.43	-1.50
	XS Sharpe	0.01	-1.59	1.93	-0.31
down|down	proportion	0.39	0.16	0.23	0.23
	TS Sharpe	1.60	-1.18	0.28	-2.37
	XS Sharpe	0.33	-2.36	1.19	-0.50

(d) *MAA 4 week formation*

(d1) *Long and short portfolios*

portfolio	return	Sharpe	proportion
long|long	11.39	2.04	0.44
long|short	4.34	0.58	0.09
short|long	-5.11	-0.61	0.05
short|short	-3.23	-0.52	0.40

(d2) *Two quantile strategies*

case	portfolio	return	Sharpe	proportion
up|up	top	21.76	4.38	0.07
	bottom	6.57	1.60	
up|down	top	9.26	1.97	0.91
	bottom	-3.55	-0.64	
down|down	top	-8.53	-1.48	0.03
	bottom	-3.87	-0.56	

(d3) *Two quantile strategies: four scenarios*

		scenario			
case	statistic	sc1	sc2	sc3	sc4
up|up	proportion	0.42	0.20	0.31	0.07
	TS Sharpe	1.68	0.41	0.55	-1.29
	XS Sharpe	0.26	-1.97	2.16	2.85
down|down	proportion	0.38	0.25	0.22	0.16
	TS Sharpe	1.78	-0.21	0.54	-1.55
	XS Sharpe	0.44	-1.29	1.73	0.44

Table F.3: ***Prediction analysis for signed strategies (accuracy) with normalised and unweighted returns.*** *The proportion of correct predictions for signed strategies run on individual asset returns (deviations) is reported. Positive and negative predictions and low and high volatility states are considered separately. Volatilities are estimated with a slow and fast EWMA per asset and a high (low) volatility state is when the fast EWMA is above (below) the slow EWMA. Returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for industries (MAA) and the first 21 days (52 weeks) of data for each asset are used to obtain initial volatility estimates. Table (a) repeats the results for unweighted assets and table (b) reports the results for assets normalised with the fast EWMA and with low and high volatility states still determined with the volatility of the unweighted assets.*

(a) *Unweighted returns*

	+ prediction			− prediction			overall		
strategy	σ low	σ high	all +	σ low	σ high	all −	σ low	σ high	all
FF49 12 mo XS	53.52	50.50	52.31	53.18	53.18	53.18	53.34	51.90	52.76
FF49 1 mo XS	54.35	51.69	53.27	53.94	54.14	54.02	54.14	52.92	53.65
MAA 52 wk XS	53.51	52.57	53.10	54.06	52.07	53.23	53.81	52.31	53.17
MAA 4 wk XS	53.69	52.86	53.23	54.77	52.51	53.86	54.25	52.68	53.56

(b) *Normalised returns*

	+ prediction			− prediction			overall		
strategy	σ low	σ high	all +	σ low	σ high	all −	σ low	σ high	all
FF49 12 mo XS	53.68	50.70	52.47	54.96	52.47	53.98	54.34	51.59	53.24
FF49 1 mo XS	53.51	52.31	53.02	54.87	54.09	54.55	54.20	53.21	53.80
MAA 52 wk XS	55.81	53.22	54.74	54.59	53.25	54.00	53.34	52.20	54.37
MAA 4 wk XS	55.57	52.44	54.24	54.36	52.95	53.73	54.89	52.61	53.99

Table F.4: ***Prediction analysis for signed strategies (excess accuracy) with normalised and unweighted returns.*** *. The proportion of correct predictions (in excess of those expected under independence) along with t-statistic in brackets are reported for data over the period the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for industries (MAA). Table (a) repeats the results for unweighted returns and table (b) gives results for normalised returns where the first 21 days (52 weeks) of data were used for initial volatility estimates.*

(a) *unweighted returns*

industries		MAA	
strategy	excess accuracy	strategy	excess accuracy
FF49 12 mo XS	2.74(6.51)	MAA 52 wk XS	3.14(8.09)
FF49 1 mo XS	3.65(8.81)	MAA 4 wk XS	3.55(9.34)

(b) *Normalised returns*

industries		MAA	
strategy	excess accuracy	strategy	excess accuracy
FF49 12 mo XS	3.22(7.64)	MAA 52 wk XS	4.37(11.09)
FF49 1 mo XS	3.79(9.15)	MAA 4 wk XS	3.99(10.36)

Table F.5: *Prediction analysis for signed strategies (Sharpe ratios) with normalised and unweighted returns.* *The (weighted average) Sharpe ratio for signed strategies run on individual asset returns (deviations) is reported. Positive and negative predictions and low and high volatility states are considered separately. Volatilities are estimated with a slow and fast EWMA per asset and a high (low) volatility state is when the fast EWMA is above (below) the slow EWMA. Returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for industries (MAA) and the first 21 days (52 weeks) of data for each asset are used to obtain initial volatility estimates. Table (a) repeats the results for unweighted assets and table (b) reports the results for assets normalised with the fast EWMA and with low and high volatility states still determined with the volatility of the unweighted assets.*

(a) *Unweighted returns*

	+ prediction			− prediction			overall		
strategy	σ low	σ high	all +	σ low	σ high	all −	σ low	σ high	all
FF49 12 mo XS	0.30	0.07	0.21	0.18	0.12	0.17	0.27	0.11	0.20
FF49 1 mo XS	0.35	0.12	0.26	0.27	0.21	0.25	0.32	0.19	0.26
FF49 12 mo TS	0.15	0.70	0.33	-0.17	-0.47	-0.29	0.06	0.08	0.07
FF49 1 mo TS	0.32	0.91	0.55	0.13	-0.16	0.02	0.23	0.32	0.27
MAA 52 wk XS	0.21	0.00	0.15	0.37	0.25	0.30	0.43	0.23	0.31
MAA 4 wk XS	0.36	0.23	0.33	0.59	0.29	0.45	0.51	0.28	0.41
MAA 52 wk TS	0.62	0.24	0.56	0.66	-0.10	0.20	0.66	0.17	0.45
MAA 4 wk TS	0.81	0.65	0.83	0.52	0.22	0.42	0.71	0.45	0.64

(b) *Normalised returns*

	+ prediction			− prediction			overall		
strategy	σ low	σ high	all +	σ low	σ high	all −	σ low	σ high	all
FF49 12 mo XS	0.33	0.13	0.27	0.25	0.15	0.24	0.33	0.17	0.28
FF49 1 mo XS	0.34	0.23	0.31	0.30	0.28	0.30	0.33	0.28	0.31
FF49 12 mo TS	0.16	0.64	0.30	-0.14	-0.40	-0.22	0.07	0.15	0.10
FF49 1 mo TS	0.34	0.84	0.51	0.14	-0.13	0.05	0.24	0.32	0.27
MAA 52 wk XS	0.41	0.12	0.32	0.58	0.31	0.41	0.61	0.34	0.48
MAA 4 wk XS	0.66	0.39	0.59	0.80	0.42	0.63	0.78	0.43	0.64
MAA 52 wk TS	0.61	0.30	0.58	0.58	-0.04	0.27	0.66	0.22	0.53
MAA 4 wk TS	0.89	0.74	0.90	0.55	0.32	0.51	0.78	0.54	0.74

Table F.6: ***Prediction analysis for signed strategies (profits over losses) with normalised and unweighted returns.*** *The ratio of average profits when a prediction is correct to losses when wrong for signed strategies run on individual asset returns (deviations) is reported. Positive and negative predictions and low and high volatility states are considered separately. Volatilities are estimated with a slow and fast EWMA per asset and a high (low) volatility state is when the fast EWMA is above (below) the slow EWMA. Returns for the period Jul1969-Jun1994 (4Jan1979-5Dec2002) are used for industries (MAA) and the first 21 days (52 weeks) of data for each asset are used to obtain initial volatility estimates. Table (a) repeats the results for unweighted assets and table (b) reports the results for assets normalised with the fast EWMA and with low and high volatility states still determined with the volatility of the unweighted assets.*

(a) *Unweighted returns*

	+ prediction			− prediction			overall		
strategy	σ low	σ high	all +	σ low	σ high	all −	σ low	σ high	all
FF49 12 mo XS	1.12	1.07	1.09	1.08	0.97	1.02	1.09	1.01	1.05
FF49 1 mo XS	1.08	1.05	1.06	1.05	0.99	1.02	1.07	1.02	1.04
FF49 12 mo TS	1.06	1.07	1.09	0.90	0.83	0.86	1.00	0.92	0.97
FF49 1 mo TS	1.18	1.24	1.22	1.04	0.96	0.99	1.10	1.04	1.08
MAA 52 wk XS	0.99	1.05	1.02	1.02	0.99	1.00	1.01	1.01	1.00
MAA 4 wk XS	1.06	1.00	1.03	1.08	0.94	0.99	1.07	0.97	1.01
MAA 52 wk TS	0.96	1.02	0.99	1.06	0.96	0.99	1.00	0.99	0.98
MAA 4 wk TS	1.01	1.02	1.01	1.12	0.96	1.02	1.06	0.99	1.01

(b) *Normalised returns*

	+ prediction			− prediction			overall		
strategy	σ low	σ high	all +	σ low	σ high	all −	σ low	σ high	all
FF49 12 mo XS	1.14	1.10	1.13	1.04	1.06	1.05	1.09	1.08	1.09
FF49 1 mo XS	1.14	1.11	1.13	1.05	1.08	1.06	1.09	1.09	1.09
FF49 12 mo TS	1.09	1.03	1.07	0.90	0.89	0.90	1.01	0.96	0.99
FF49 1 mo TS	1.22	1.19	1.20	1.03	0.99	1.02	1.12	1.05	1.09
MAA 52 wk XS	0.98	1.01	0.99	1.03	1.01	1.02	1.08	1.05	1.00
MAA 4 wk XS	1.07	1.06	1.07	1.13	1.03	1.08	1.11	1.04	1.07
MAA 52 wk TS	1.03	0.99	1.02	1.08	1.01	1.05	1.12	1.05	1.03
MAA 4 wk TS	1.09	1.06	1.08	1.12	1.05	1.09	1.11	1.05	1.08

Table F.7: *Returns in formation and holding period by quantile in formation period with normalised returns.* *Average returns for each rank are calculated for normalised returns. Ranks are then grouped into quantiles and the annualised average of these average returns is reported. Data over the period Jul1969-Jun1994 (4Jan1979-5Dec2002) for industry (MAA) data are used. The first 21 days (52 weeks) of returns for each asset are used for the initial volatility estimates.*

(a) *FF49 12 month formation quantiles of 4*

quantile	form	hold
1	20.86	7.55
2	14.09	7.34
3	10.60	5.82
4	8.12	4.75
5	6.09	4.87
6	4.19	4.67
7	1.94	4.60
8	-0.02	3.58
9	-2.13	0.68
10	-4.61	0.20
11	-7.67	-0.78
12	-13.04	-1.39

(b) *FF49 1 month formation quantiles of 4*

quantile	form	hold
1	68.52	6.74
2	38.55	7.04
3	26.66	4.87
4	18.36	5.53
5	11.71	4.80
6	5.60	3.07
7	-0.86	2.41
8	-5.94	2.05
9	-11.18	0.57
10	-16.99	0.91
11	-24.21	-1.16
12	-36.43	-2.87

(c) *MAA 52 week formation quantiles of 3*

quantile	form	hold
1	20.37	9.18
2	9.63	6.64
3	3.50	3.79
4	-3.42	-1.86
5	-12.26	-3.07

(d) *MAA 4 week formation quantiles of 3*

quantile	form	hold
1	64.02	12.59
2	22.23	7.67
3	5.89	3.31
4	-13.71	-2.37
5	-35.99	-5.13

Table F.8: *Empirical decomposition for linear global time-series strategy with normalised returns.* *Estimates for the decomposition for global linear time-series strategies in section 2.7 are reported from normalised returns from all possible combinations of formation period and holding period. 12 month and 1 month (52 week and 4 week) formation periods with a 1 month (1 week) holding period are considered for industry (MAA) data. Returns are used for industry data over the period Jul1969-30Jun1994, all MAA data over the period 4Jan1979-5Dec2002 and a subset of three bond indices from the MAA data over the period 31Dec81-5Dec2002. The first 21 days (52 weeks) of returns for each asset used for the initial volatility estimates. For the MAA data a set of 11 assets (for which data is available over the whole period) and a set of 3 bond indices. Figures are in per cent and annualised by multiplying by 12 or 52. Percentage of total profits is indicated in brackets.*

strategy	TS	cross	mean	total
FF49 12 mo	0.00(-2.42)	-0.23(196.90)	0.11(-94.49)	-0.11(100.00)
FF49 1 mo	0.00(3.21)	0.10(91.24)	0.01(5.55)	0.11(100.00)
MAA11 52 wk	0.09(26.24)	0.16(47.11)	0.09(26.65)	0.33(100.00)
MAA11 4 wk	0.02(27.82)	0.04(61.93)	0.01(10.25)	0.07(100.00)
MAA bonds 52 wk	0.41(33.25)	0.43(34.79)	0.39(31.96)	1.23(100.00)
MAA bonds 4 wk	0.11(35.88)	0.16(54.57)	0.03(9.55)	0.29(100.00)

Table F.9: ***Sharpe ratios for local and global signed time-series strategies with normalised returns.*** *The local strategies are run on (and diversified over) individual assets. The global strategy is a signed strategy run on an equal-weighted market of all assets in the dataset. Returns are used for industry data over the period Jul1969-30Jun1994, all MAA data over the period 4Jan1979-5Dec2002 and a subset of three bond indices from the MAA data over the period 31Dec81-5Dec2002. The first 21 days (52 weeks) of data for each asset are used for the initial volatility estimates. Sharpe ratios are annualised.*

strategy	local	global
FF49 12 mo	0.17	0.08
FF49 1 mo	0.52	0.37
MAA 52 wk	1.52	0.75
MAA 4 wk	1.99	1.34
MAA bonds 52 wk	1.09	1.09
MAA bonds 4 wk	1.64	1.56