Is it possible to control for momentum's huge crashes?

Scaled residual momentum – A modified approach to momentum investing

Luís Saias

Dissertation written under the supervision of Joni Kokkonen

Dissertation submitted in partial fulfilment of requirements for the MSc in Finance, at the Universidade Católica Portuguesa, 25/01/2017

Is it possible to control for momentum's huge crashes?

Scaled residual momentum – A modified approach to momentum investing

Abstract

Many researchers find that information travels slowly and an apparently under-reaction to news. As a response, several relative strength strategies appear. Total momentum is one of the most known and widely used nowadays. However, total momentum has huge crashes from times to times. The early 1930s and the late 2000s were the two darkest periods for the strategy. In fact, an investor entering in the beginning of these two periods would see 60% to 75% of his investment wiped out. Trying to control for these crashes we study 3 alternative strategies: residual momentum, scaled momentum and scaled residual momentum. The last strategy scales its exposure to residual momentum, relying on the risk predictability of residual momentum. Our main findings state that overall scaled residual momentum has a Sharpe ratio slightly lower than scaled momentum (0,91 vs 1,00), but in turbulent times this reverts completely. In the early 1930s and the late 2000s it becomes the strategy with highest Sharpe ratio (ranging from 0,36 to 0,66) and highest cumulative return. For the same periods, where total momentum has huge losses, an investor would see a valorization of 20% (in the late 2000s) and of 60% (in the early 1930s). The reason for this is the evident superiority shown by scaled residual momentum in terms of controlling for the crashes - lowest kurtosis, less negative skewness and lower and shorter drawdowns.

Muitos investigadores notam que a informação viaja devagar e que parece existir reacção insuficiente às notícias. Como resposta, muitas estratégias de força relativa aparecem. O total momentum é uma das mais conhecidas e usadas actualmente. No entanto, o total momentum tem grandes quedas de tempos a tempos. O ínicio dos anos 30 e o final da década de 2000 foram os dois períodos mais negros da estratégia. Um investidor que entrasse no ínicio desses dois períodos veria entre 60% a 75% do seu investimento desaparecer. Tentando controlar para estas quedas nós estudamos 3 estratégias alternativas: residual momentum, scaled momentum e scaled residual momentum. A última varia a sua exposição ao residual momentum, confiando na previsibilidade do risco do residual momentum. As nossas principais descobertas demonstram que na amostra completa o scaled residual momentum tem um Sharpe ratio ligeiramente menor que o scaled momentum (0,91 vs 1,00), revertendo-se totalmente em tempos turbulentos. No ínicio dos anos 30 e no final da década de 2000 torna-se a estratégia com maior Sharpe Ratio (desde 0,36 a 0,66) e maior retorno cumulativo. Para os mesmos períodos, onde o total momentum tem enormes quedas, um investidor observa valorização de 20% (no final da década de 2000) e de 60% (no início dos anos 30). A razão é a evidente superioridade demonstrada pelo scaled residual momentum em termos de controlar para as quedas – menor kurtosis, skewness menos negativa e menores e mais curtos drawdowns.

Acknowledgments

In this section I want to recognize and thank those people who helped me the most during this period of my life.

First, I want to thank Professor Joni Kokkonen for his never ending patience, motivation, expertise and leadership. Professor Joni Kokkonen was crucial during my research and writing of the thesis, guiding me all along the process.

Second I want to thank my family: my mother Madalena Saias, my father Luís Saias, my sister Rita Saias and all other members who helped me to get to where I am today. Their constant availability and encouraging were, are and always will be very important to me.

Thirdly, I want to thank Inês Menezes for her support, positivity, availability, interest and proud shown in my masters, my thesis and in life.

Finally, I want to thank Fundação para a Ciência e a Tecnologia (FCT) for financial support.

1. Introduction and Literature review

How information travels across investors and how fast they react to it is a much debated topic in behavioral finance. Many authors believe information is not immediately incorporated in stock prices, which can result in an under-reaction to news (especially to firm specific news). For example, Barberies, Schleifer and Vishny (1998) state that investors suffer from a conservative bias under-reacting to firm-specific news, such as earning announcements. Hong and Stein (1999) argue that each "newswatcher" is able to capture some information but not all, which results in gradual-information-diffusion. Hong, Lim and Stein (2000) findings are clearly consistent with Hong and Stein (1999) and with the view that firm specific information diffuses gradually across investors (especially negative information).

Supporting these findings, several relative strength strategies try to take advantage of the anomaly. In one of the first researches related to relative strength strategies, Levy (1967) state that investing in stocks that have been "relatively strong in price movements" (or in more common language, that have had relative high returns) seems to be profitable. Despite being somewhat contested due to a selection bias when choosing the trading rule (almost 70 different trading rules have been considered), Levy opened a path for the following researchers.

In 1993 appears the first worldwide known notion of momentum, followed by many others. Jegadeesh and Titman (1993) select stocks according to their returns in the past J months and hold the portfolio for K months. The strategy is tested with J and K ranging from 3 to 12 months. Grundy and Martin (2001) use a 6-month formation period to select the stocks and skip a month between the formation and the holding period. However, nowadays the broadest used momentum definition is to select the stocks using a 11-month formation period, skipping a month between the formation and the holding period. This procedure will be presented in detail further in the thesis.

Momentum does not appear only in academic research. It is also widely used in practice by the big players in the market. In fact, Grinblatt and Titman (1989, 1993) find that several mutual funds show a tendency to bet in assets that performed well in the past. Also, many of the mutual funds incur in window dressing, which is basically selling assets that performed badly in the past in order to avoid showing them in the quarterly report.

Momentum has been heavily studied and employed because it seems to be a strategy that delivers outstanding returns, with a good Sharpe ratio. However, this superior performance has to come aligned with some kind of downside. In fact, momentum has a big problem. It can be delivering great returns for years, even decades and suddenly experience a huge crash.

One of the explanations for these crashes comes from the momentum strategy having time-varying exposure to the market, size and value factors. As showed in Grundy and Martin (2001) after bear (bull) markets momentum has a negative (positive) beta. After a bear (bull) market stocks that performed well (badly) in the past usually have a low (high) beta and stocks that performed badly (well) have a high (low) beta. As momentum is constructed as buying the winners (stocks with a good past performance) and selling the losers (stocks with a bad past performance) after bear (bull) markets momentum has negative (positive) beta. So, we can observe severe crashes when the market recovers rapidly after a bear market. In fact, the worst crashes for the strategy happened in 1932 and in 2009, right when the markets were recovering from the two most pronounced recessions that modern society has gone through. In a broader way, when the factor (market, size or value) has an opposite sign during the formation and the holding period momentum experiences the worst performances.

Several attempts have been done to avoid or minimize these crashes. Grundy and Martin (2001) hedge the crashes using betas that rely in information that investors couldn´t access at the time they were investing. However, Daniel and Moskowitz (2015) show this strategy doesn't work when using only available information at the time of investment. Blitz, Huitz and Martens (2011) select stocks according to their residual returns instead of regular returns. Barroso and Santa-Clara (2015) scale their exposure to momentum to have constant volatility.

With the purpose of testing to what extent the crashes can be controlled our research compares the traditional total momentum strategy with three strategies that try to solve the problem: residual momentum from Blitz, Huitz and Martens (2011), scaled momentum from Barroso and Santa-Clara (2015) and scaled residual momentum. The last strategy scales its exposure to residual momentum (instead of scaling to total momentum as done in scaled momentum from Barroso and Santa-Clara (2015)). The scaling relies on the risk predictability of residual momentum (as it uses the realized variance of residual momentum in the last 6 months in the scaling process).

First we test all the four strategies in a full sample from November 1930 to December 2015 and then we test those same strategies for the darkest periods for momentum: the early 1930s and the late 2000s (referred in the remaining thesis as the 1930s ant the 2005s respectively). It seems quite obvious that scaled strategies perform better than non-scaled ones in every situation. Overall scaled residual momentum has a Sharpe ratio slightly lower than scaled momentum (0,91 vs 1,00), but in turbulent times this reverts completely. In turbulent times it becomes the strategy with highest Sharpe ratio, lowest volatility and highest cumulative return. The reason for this is the evident superiority shown by scaled residual momentum in terms of controlling for the crashes - lowest kurtosis, less negative skewness and lower and shorter drawdowns. It seems obvious that the strategy is able to virtually eliminate the dangers of the momentum crashes. Also, scaled residual momentum is able to deliver the highest annualized alpha among the strategies (14,74%).

The rest of the thesis is organized as follows: Section 2 describes the data. Section 3 explains the methodology used in the construction of the strategies. Section 4 presents the results. Section 5 concludes and delivers recommendation for future research.

2. Data

The data is extracted from two different sources: the Center for Research in Security Prices (CRSP) data base and Kenneth French's library[1].

Stock returns are extracted from CRSP data base. ADRs, REITs, financials, close-end-funds and foreign shares are excluded (filter by share codes 10 and 11). Only shares listed on NYSE, NASDAQ and AMEX are considered (filter by exchange codes 1, 2 and 3). As in Blitz, Huitz and Martens (2011) only periods (months or days) where stock price is above $1 are considered (to avoid concerns related to microstructure).

Market, size, value and risk free returns are extracted from Kenneth French's library.

All data is extracted monthly and daily from July 1926 to December 2015.

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

3. Methodology

This section discusses all the methodologies used to implement and evaluate the four strategies analyzed in the thesis. Section 3.1 addresses total momentum, Section 3.2 residual momentum and Section 3.3 scaled momentum and scaled residual momentum.

3.1. Total Momentum

Consistent with the majority of the recent literature on the subject (see for example Daniel and Moskowitz (2016), Barroso and Santa Clara (2015) or Asness, Moskowitz and Pedersen (2013)) at the beginning of each month *t* the stocks are ranked according to their cumulative return in the past 12 month excluding the last month (from month t-12 to month *t*-2). Then, according to their ranking are assigned to mutually exclusive decile portfolios. The top portfolio is called "the winners" and the bottom portfolio is called "the losers". Stocks are equal weighted within these portfolios. Total momentum then buys "the winners" and sells "the losers" and holds the position for 1 month.

To analyze the performance of the strategies we consider the returns, risks, Sharpe ratios, kurtosis and skewness as well as the alphas relative to the market, size and value and the drawdowns of each strategy. In order to compute the alphas and in line with Blitz, Huij and Martens (2011) and Grundy and Martin (2001) we use the following regression:

$$r_{i,t} = \alpha_i + \beta_{1,i}RMRF_t + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}RMRF_{UP_t} + \beta_{5,i}SMB\_UP_t \quad (1)$$

$$+\beta_{6,i}HML\_UP_t + \varepsilon_{i,t},$$

where $r_{i,t}$ corresponds to the excess return of stock *i* in month *t*, $RMRF_t$, SMB_t and HML_t correspond to the excess return of the market, size and value at time *t*, $\mathrm{RMRF\_UP}_t$, $\mathrm{SMB\_UP}_t$ and $\mathrm{HML\_UP}_t$ correspond to the excess return of the market, size and value at time *t* when the cumulative return for these factors is positive over the period from month

t-12 to month t-2 and is zero when the cumulative return over the same period is either zero or negative.

The drawdown of each strategy measures the dimension of the cumulative crashes and allows observing their length, or in other words, the time it takes to overcome them. As in Blitz, Huitz and Martens (2011) to determine the drawdown first we compute the ratio between cumulative return in month t and the maximum cumulative return up to month t. Then, we subtract 1 to this ratio. Thus, by construction the drawdown is never positive. It is zero when the cumulative return of the last month t is also the highest cumulative return up to month t and negative otherwise. We compute both the alpha and drawdown in the same manner for all the four strategies analyzed in this thesis.

3.2. Residual Momentum

Residual momentum methodology is very similar to total momentum methodology. Stocks are ranked according to their cumulative returns in the last 12 months excluding the most recent month and assigned to equal weighted and mutually exclusive decile portfolios. Then, the strategy buys "the winners" and sells "the losers", holding the portfolio for the period of 1 month. However, the main difference is on how the stocks are ranked. While with total momentum they were ranked on total returns, with residual momentum they are ranked on residual returns. Using this different approach we are able to exclude the returns attributable to known sources of risk such as the market, size and value.

To compute the residual returns we use the Fama and French three-factor model given by:

$$r_{i,t} = \alpha_i + \beta_{1,i} RMRF_t + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + \varepsilon_{i,t}, \tag{2}$$

where $r_{i,t}$ corresponds to the total excess return of firm i in month t, $RMRF_t$, SMB_t and HML_t correspond respectively to market, size and value in month t and $\varepsilon_{i,t}$ is the residual return of firm i in month t. However, $\varepsilon_{i,t}$ is different from the residual return we use to rank the stocks and determine residual momentum. While $\varepsilon_{i,t}$ includes the estimated alpha (α_i), the residual return we use to rank the stocks and determine residual momentum excludes it. As the estimated alpha is calculated with a 36-month window more than two thirds of its value is relate with a period prior to the 11-month formation period. So, it would not make sense to include the alpha when ranking the stocks. Otherwise we probably would be ranking low stocks with good returns in the period from t-36 until t-13.

We estimate the regression every month *t* for each stock *i* using a 36-month rolling window. This means that the market, size and value betas are estimated over the period from *t*-36 until *t*-1. Stocks that do not present a complete return history over the 36-month period are excluded.

Also, and in line with Blitz, Huitz and Martens (2011) who state that raw residual returns are noisier than standardized residual returns and Guitierrez and Pirinsky (2007) who believe that, when residual returns are standardized the decrease in the noise allows to reliantly connect the residual returns to firm-specific news, we standardize the cumulative returns over the period from *t*-12 to *t*-2 by its standard deviation over that same period. We do this hoping to obtain an improved measure, which in fact happens[2].

3.3. Scaled Momentum and Residual Scaled Momentum

Both scaled strategies (scaled momentum and scaled residual momentum) aim at having constant risk over time and use estimates of the non-scaled strategies' volatility (total momentum and residual momentum) to try to accomplish this. So, the volatility of the strategy being scaled must have at least some predictive power.

To compute realized and forecasted variances (and volatilities) we use daily returns.

[2] If the cumulative returns are not standardized (raw residual returns), then residual momentum has a maximum and a minimum monthly return of 19,55%, and -44,82% respectively, average return of 64,48%, volatility of 16,88%, Sharpe ratio of 0,52, kurtosis of 18,45, skewness of -2,61 and annualized alpha of 8,26%. In table 2 are presented the results for residual momentum with standardized cumulative returns, which seem more appealing than these not only in terms of the risk-return tradeoff, but also in the higher moments (skewness and kurtosis – highly responsible for the crashes).

Total momentum daily returns follows exactly the same procedure reported in Section 3.1 , with the exception that the stocks are ranked according to their cumulative return in the past 250 days excluding the last 21 days and the portfolio is held for one day.

Residual Momentum daily returns follow a similar procedure to the one reported in Section 3.2, also with some exceptions. To compute the residual returns we use the Fama and French three-factor model in equation (3):

$$r_{i,t} = \beta_{1,i}RMRF_t + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \varepsilon_{i,t} \tag{3}$$

Equation (3) is similar to equation (2). However, in equation (3) no coefficients are estimated (α or β) but instead $\beta_{1,i}$, $\beta_{2,i}$ and $\beta_{3,i}$ are the coefficients from equation (2) for the month to which day *t* belongs. $r_{i,t}$ corresponds to the total excess return of firm *i* in day *t*, $RMRF_t$, SMB_t and HML_t correspond respectively to market, size and value in day *t* and $\varepsilon_{i,t}$ is the residual return of firm *i* in day *t*.

Also, the stocks are ranked according to their cumulative residual returns in the past 250 days excluding the last 21 days and the portfolio is held for one day.

Scaled strategies aim at having constant risk over time. They try to accomplish this by scaling their exposure to the non-scaled strategies. This scaling is computed with time-varying weights (explained in detail further in this section), dependent on the realized variance of the non-scaled strategies in the past 6 months. Then, there is a need to check the risk predictability of the non-scaled strategies. Realized variance in month *t* is calculated as the sum of the squared returns of the last 21 trading sessions:

$$RV_{i,t} = \sum_{j=0}^{20} r_{i,d_t-j}^2 \tag{4}$$

where $\mathrm{r}_{\mathrm{i,d_t-j}}^2$ is the daily squared return for each non-scaled strategy *i* (total momentum and residual momentum) in each trading session from d_0 until d_{-20}. Realized Volatility is computed as the realized variance squared root.

Then, an AR(1), auto regressive process, is computed to ensure the predictive power of Realized Variance. In other words, we want to find how dependent this month' realized variance is on last month' Realized Variance:

$$RV_{i,t} = \alpha + \beta RV_{i,t-1} + \varepsilon_t \tag{5}$$

where $\mathrm{RV}_{\mathrm{i,t}}$ and $\mathrm{RV}_{\mathrm{i,t-1}}$ correspond to the realized variances of each non-scaled strategy in month *t* and month *t*-1 respectively.

An out-of-sample R^2 is also computed to check if the predictive power persists out-of-sample. We run the first AR(1) process with a sample of 240 moths. Then, each month we use the coefficients generated from the autoregressive process and the last observation of the realized variance to forecast the variance of the next month. As we use an expanding window each new month's forecast uses more one observation than the forecast of the previous month. In order to compare the accuracy of this OOS forecast with the accuracy of the realized variance historical mean it is used the following measure:

$$R^2_{i,OOS} = 1 - \frac{\sum_{t-s}^{T-1}(\widehat{\alpha}_t + \widehat{\beta}_t RV_{i,t} - RV_{i,t+1})^2}{\sum_{t=s}^{T-1}(\overline{RV}_{i,t} - RV_{i,t+1})^2} \tag{6}$$

where s is the initial sample (240 months), T is the size of the sample, $\hat{\alpha}_t$ and $\hat{\beta}_t$ are the coefficients estimated in the AR(1) in each month *t* and $\overline{RV}_{i,t}$ corresponds to the realized variance historical mean of each non-scaled strategy in month *t*.

With the risk predictability checked it is finally possible to implement the scaled strategies. First, a volatility forecast is computed as:

$$\hat{\sigma}_{i,t} = \sqrt{21 \sum_{j=0}^{125} r^2_{i,d_{t-1}-j} / 126} \tag{7}$$

where $r_{i,d_{t-1}-j}^2$ corresponds to the daily squared return of each non-scaled strategy in each of the last 126 trading sessions.

Then, we scale the returns on the forecasted volatility as followed:

$$r_{i^*,t} = \frac{\sigma_{target}}{\hat{\sigma}_{i,t}} r_{i,t} \tag{8}$$

where $r_{i^*,t}$ corresponds to the return in month *t* of the scaled strategy, σ_{target} is the target volatility and $r_{i,t}$ is the return of the non-scaled strategy in month *t*. As in Barroso and Santa-Clara (2015), we also pick a value of 12% for target volatility.

The methodology of the scaled strategies aims at having constant risk, using varying weights to do so. With the objective of having constant volatility, the scaled strategies scale their exposure to the non-scaled strategies, relying on their risk predictability.

4. Results

This section comprises the results of all the four strategies for different time periods. Section 4.1 addresses the volatilities and the risk predictability of both daily non-scaled strategies (total momentum and residual momentum), Section 4.2 the performance measures, results and cumulative returns for the complete sample (from November 1930 until December 2015) and Section 4.3 the strategies' performance in the two most critical periods of the momentum risk factor.

4.1. Risk predictability of non-scaled strategies

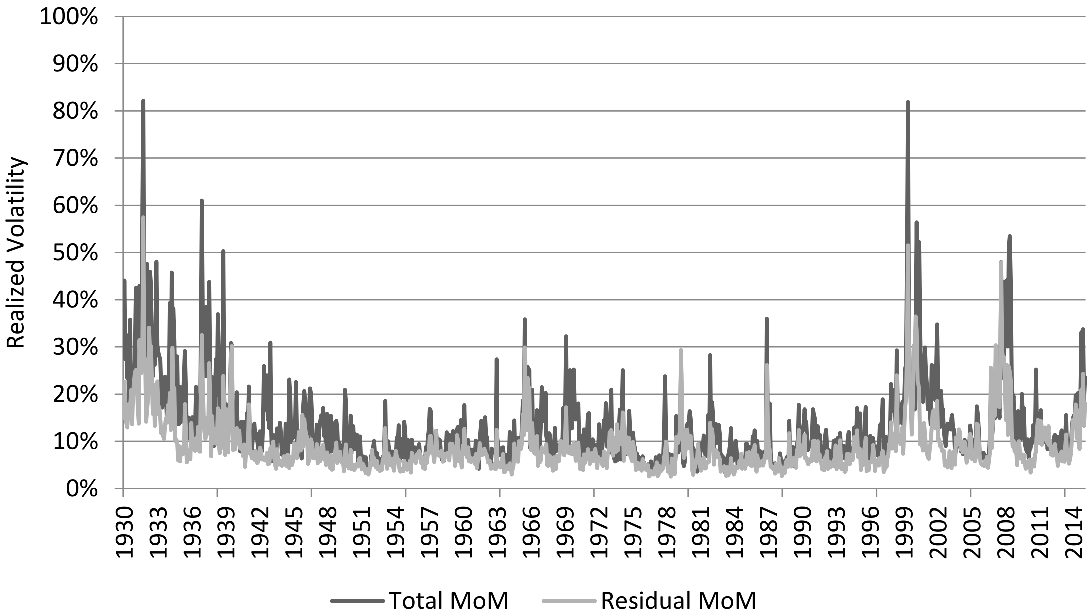
Both scaled strategies being studied in this thesis, scaled momentum from Barroso and Santa-Clara (2015) and scaled residual momentum, scale their exposures to their peer non-scaled strategies based on their past 6 month volatility. This exposure is not constant.

In fact, the weight allocated to the non-scaled strategies can take very disparate values. The weights are calculated as a ratio between a target volatility and the forecast volatility of the non-scaled strategies (see equation (8)), so it is crucial to observe risk predictability in non-scaled strategies. In this section we address the volatilities and risk predictability of total momentum and residual momentum.

First, Figures 1 presents the monthly realized volatilities of total momentum and residual momentum. These monthly volatilities are obtained using daily data, as explained in detail in Section 3.3.

Figure 1 – Realized volatility

Figure 1 shows the monthly realized volatility of total momentum and residual momentum. The volatilities are the square root of the realized variances as in equation (4). Stock returns are extracted from CRSP data base. ADRs, REITs, financials, close-end-funds and foreign shares are excluded. Only shares listed on NYSE, NASDAQ and AMEX are considered. Only periods (months or days) where stock price is above $1 are considered. Our sample period goes from November 1930 to December 2015. The volatilities are annualized and daily returns are used in their computation.



Looking in more detail at the distribution over time of the monthly realized volatility it is easily spotted that it is far from being constant. It can take very disparate values, ranging from minimums of roughly 3% (for both strategies) to maximums of 82% and

57% for total momentum and residual momentum respectively. The volatility peaks coincide with the crashes. In fact, the maximum value for both strategies occur in August 1932 (the worst month for the momentum factor).Also, presenting a correlation of about 85% it is quite evident that the realized volatility of both strategies move in a very similar way over time. However, the realized volatility of total momentum is predominately higher (presenting an average of 13% against 9% for residual momentum). This can be partially explained by residual momentum being a steadier strategy, due to its hedging of the fama-french factors.

Second, in order to ascertain the risk predictability Table 1 presents the results of the autoregressive process (AR(1)) regressions of the monthly realized variance of total momentum and residual momentum. These results are summarized in Table 1.

Table 1 – AR(1) regressions of the monthly realized variance

Table 1 presents the AR(1) regression of the monthly realized variance for both strategies: total momentum and residual momentum.

The variance is computed as in equation (4) and using daily returns. The AR(1) process regresses the variance on its own lagged variable. Columns 1 and 2 present the alphas, the betas and the t-statistics between brackets. Columns 3 and 4 present measures of fitness of the model: R-Squared (R^2) and Out-of-sample R-Squared (OOS R^2). R-Squared (R^2) uses the full sample. For the computation of the Out-of-sample R-Squared (R^2) it is used an expanding window. An initial regression of 240 monthly observations is used to set the initial forecast. Then each month the forecast uses one more observation until the end of the sample. Columns 5 and 6 present the average volatility ($\bar{\sigma}$) and the volatility standard deviation (σ_σ). $\bar{\sigma}$ and σ_σ are annualized. Stock returns are extracted from CRSP data base. ADRs, REITs, financials, close-end-funds and foreign shares are excluded. Only shares listed on NYSE, NASDAQ and AMEX are considered. Only periods (months or days) where stock price is above $1 are considered. Our sample period goes from November 1930 to December 2015.

Strategy	α	B	R^2	OOS R^2	$\bar{\sigma}$	σ_σ
Total MoM	0,0009	0,56	31,17	33,33	13,14	9,24
	(7,77)	(21,52)				
Residual MoM	0,0004	0,56	31,83	35,26	8,89	5,90
	(7,79)	(21,86)				

The coefficients (β) of the variance of total momentum and residual momentum are both 0,56 and clearly significant (with t-statistics above 20). Thus, both strategies seem to present persistence or predictable risk. Also, and reinforcing the idea both measures of fitness of the model (R^2 and OOS R^2) show values above 30%. Not only, the risk predictability works in sample, but it also works out-of-sample.

Total momentum presents an average volatility somewhat higher than residual momentum (13% against 9%). The same is also true for the volatility's standard deviation (9% for total momentum against 6% for residual momentum), which means that the risk of total momentum is in some way more variable. However, residual momentum presents slightly higher R^2 and OOS R^2. Overall, we might conclude that both strategies show risk predictability, which means that their peer scaled strategies can take advantage of this predictive power to scale their exposures.

4.2. Main Results: Performance evaluation - full sample

We start our analysis of the main results with an evaluation of some key performance measures: average return, volatility, Sharpe ratio, kurtosis, skewness, alpha, minimum and maximum monthly return and the probability of having a positive monthly return in a certain month. These results are summarized in Table 2.

Table 2 – Key performance measures (full sample)

Table 2 presents the average return, volatility, Sharpe ratio, kurtosis, skewness, alpha, minimum and maximum monthly return and the probability of having a positive monthly return in a certain month for total momentum, residual momentum, scaled momentum and scaled residual momentum. Alphas are estimated as in equation (1). Stock returns are extracted from CRSP data base. ADRs, REITs, financials, close-end-funds and foreign shares are excluded. Only shares listed on NYSE, NASDAQ and AMEX are considered. Only periods (months or days) where stock price is above $1 are considered. Our sample period goes from November 1930 to December 2015. All values are annualized except for the minimum and maximum monthly returns.

Strategy	Max	Min	P(ret>0)	Mean	Vol	Sharpe	Kurtosis	Skewness	Alpha
Total MoM	28,48	-70,26	66,05	12,24	24,78	0,49	26,04	-3,23	9,88
Residual MoM	21,48	-38,31	65,07	9,01	14,45	0,62	14,34	-1,84	7,95
Scaled MoM	19,45	-38,69	66,05	18,19	18,15	1,00	5,89	-1,06	13,90
Scaled Residual MoM	20,43	-31,14	65,07	16,22	17,85	0,91	2,77	-0,60	14,74

Total Momentum, as it has been showed in previous work by several researchers, from time to times experiences huge crashes (see for example Daniel and Moskowitz (2016), Barroso and Santa Clara (2015), Chaves (2012) or Blitz, Huitz and Martens (2011)). It's minimum monthly return is -70,26% in August of 1932. All the three other strategies manage to decrease these huge negative monthly returns. While the worst monthly return of residual momentum and scaled momentum are -38,31% and -38,69% respectively, scaled residual momentum in its worst month has a loss of -31,14%. When we look at the best monthly returns we observe that total momentum´s maximum return corresponds to 28,48% in February of 2000, while the others´ strategies maximum ranges from 19,45% for scaled momentum to 21,48% for residual momentum. Scaled residual momentum delivers 20,43% in its´ best month and is the strategy with the tighter range of values.

In what regards the distribution of monthly returns between positive and negative returns, all the four strategies seem very similar. The probability of having positive monthly returns is either 65% or 66% for all the four strategies under analysis.

In terms of the risk reward relation both scaled strategies are much superior to the non-scaled ones. Total momentum is the strategy with the higher volatility (24,78%) and it's return (12,24%) clearly does not compensate for the risk, which results in the lowest Sharpe ratio among all the four strategies. Despite having the lowest volatility (14,45%) residual momentum also has the lowest profitability (9,01%). So, delivering much better profits with very reasonable volatilities, the scaled strategies show much more attractive relations of risk-return. Scaled momentum and scaled residual momentum deliver Sharpe ratios of 1,00 and 0,91 respectively. These, compared with the 0,49 (of total momentum) and 0,62 (of residual momentum) constitute a big increase.

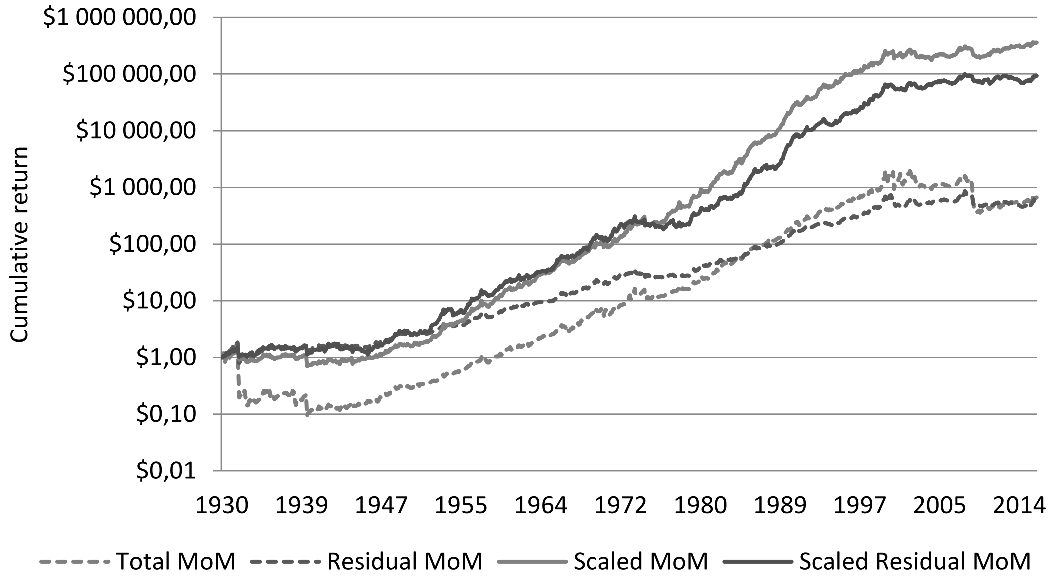
But, where the scaled strategies make really a great difference is in the higher moments of the distribution. Total momentum tends to have heavy tails, which means that it presents more extreme values than a normal distribution. These extreme values in the left side of the distribution (momentum crashes) are the ones that have been worrying the investors in the last decades. Also, total momentum is a strategy that is left skewed, being more concentrated in the left side of the mean. We find that total momentum has a kurtosis of 26,04 and a skewness of -3,23.

Residual momentum is able to reduce both these measures but only to some extent. It presents a kurtosis of 14,34 and a skewness of -1,84. However is with the scaled strategies, and especially with scaled residual momentum, that we see skewness and kurtosis reduced to minimum values. Scaled Momentum has a kurtosis of 5,89 and skewness of -1,06. Scaled residual Momentum goes even further and practically eliminates the problem of the momentum huge crashes. In fact, this is one of the features that distances this new strategy from the others already developed. It has kurtosis of 2,77 (even lower than the normal distribution, which has a kurtosis of 3) and skewness of -0,60 (slightly left skewed).

The other feature that makes of scaled residual moment a superior strategy is the alpha it delivers. Relatively to the factor model in equation (1) this is the strategy that delivers the highest alpha, which means that when controlling for the three known sources of risk (market, size and value) scaled residual momentum is the most profitable strategy of all the four strategies being analyzed. While residual momentum and total momentum deliver annualized alphas of 7,95% and 9,88% respectively, the scaled strategies deliver alphas of at least 1,5 times these values. Scaled momentum delivers an alpha of 13,90% and scaled residual momentum is the strategy delivering the highest alpha, 14,74%.

Figure 2 - Cumulative returns (full sample)

Figure 2 shows the cumulative returns of total momentum, residual momentum, scaled momentum and scaled residual momentum. Stock returns are extracted from CRSP data base. ADRs, REITs, financials, close-end-funds and foreign shares are excluded. Only shares listed on NYSE, NASDAQ and AMEX are considered. Only periods (months or days) where stock price is above $1 are considered. Our sample period goes from November 1930 to December 2015.



In terms of cumulative returns it is clearly observable in Figure 2 that when considering the whole period under analysis (from November 1930 to December 2015) the scaled strategies are delivering much better returns than the non-scaled ones. In fact, while the scaled strategies show cumulative returns with 6 digits, both non-scaled strategies show cumulative returns for the period below a thousand percent. Also, total momentum is the strategy with more profound and visible crashes, which take a lot of time to recover from. In the early 30s and in the second half of the first decade of the 21st century these crashes are monstrously obvious.

In order to observe the dimension and length of the crashes it is presented the drawdown of all the four strategies. Once the drawdown curves of residual momentum, scaled momentum and residual scaled momentum move in a very similar way we report two Figures: Figure 3 shows the drawdown of both non-scaled strategies and Figure 4 the drawdown of total momentum and both scaled strategies.

Figures 3 and 4 – Drawdowns (full sample)

Figures 3 and 4 present the drawdowns of total momentum, residual momentum, scaled momentum and scaled residual momentum. The drawdown is computed as the ratio between cumulative return in month t and the maximum cumulative return up to month *t*, to which is then subtracted 1. Stock returns are extracted from CRSP data base. ADRs, REITs, financials, close-end-funds and foreign shares are excluded. Only shares listed on NYSE, NASDAQ and AMEX are considered. Only periods (months or days) where stock price is above $1 are considered. Our sample period goes from November 1930 to December 2015. For the sake of better understanding, Figure 3 shows the drawdown of total momentum and residual momentum and Figure 4 shows the drawdown of total momentum, scaled momentum and scaled residual momentum.

Figure 3 – Total momentum and residual momentum

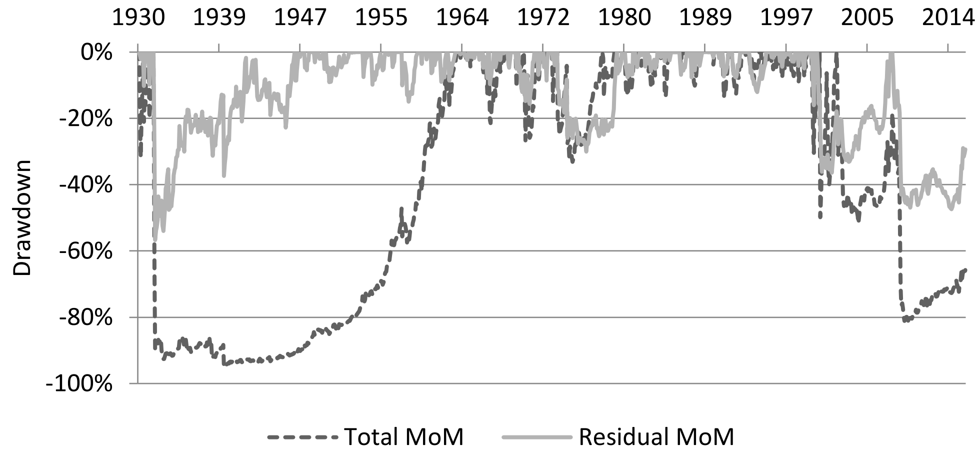
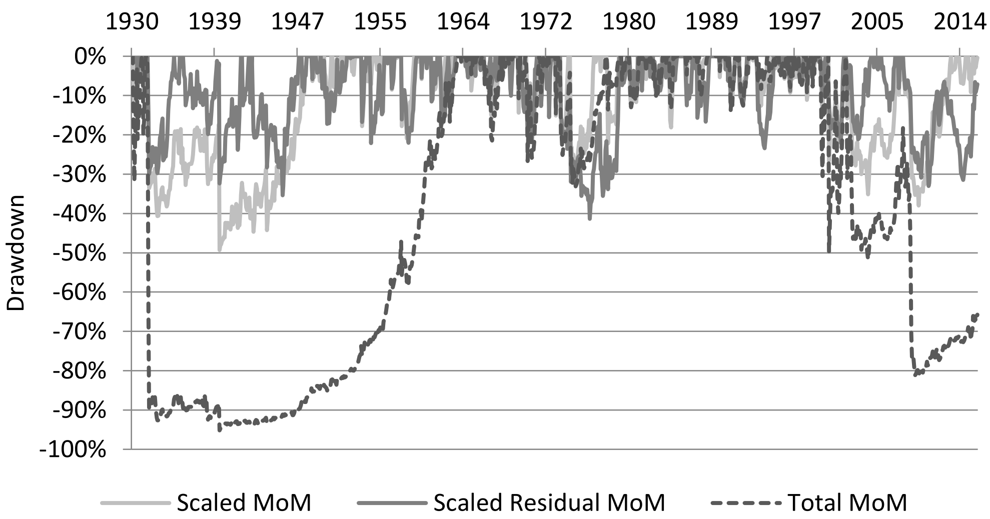


Figure 4 – Total momentum, scaled momentum and scaled residual momentum



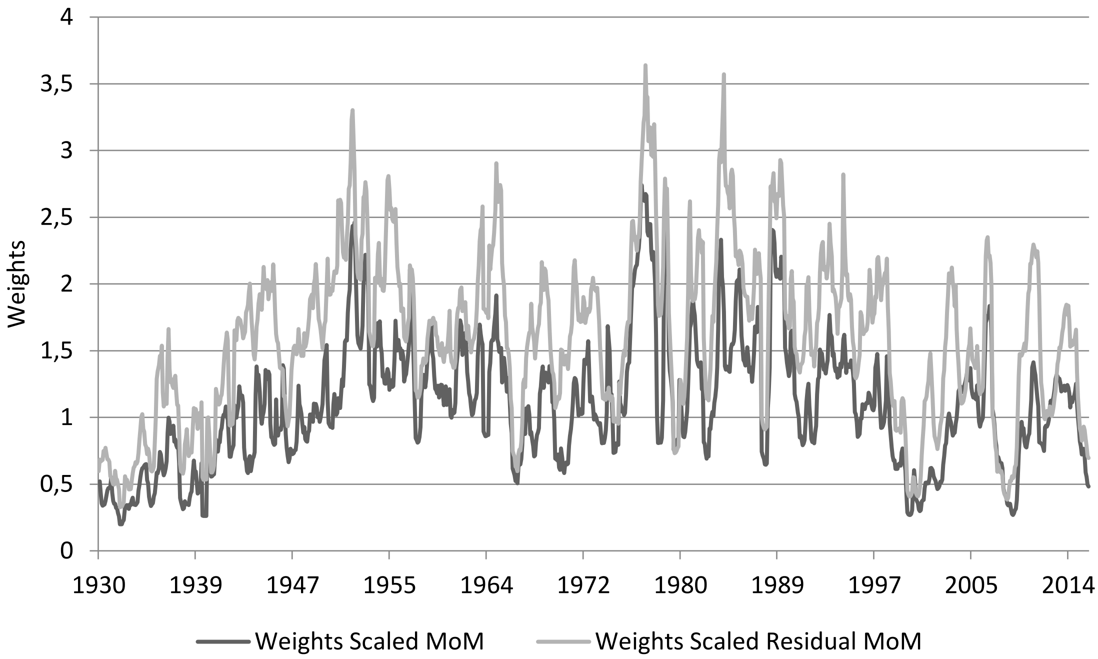
It seems quite obvious that the strategy experiencing the worst drawdowns is total momentum. In fact it´s worst drawdown reaches a negative peak of -95,18% in the 30s, which is only fully recovered in the 60s (more than 31 years later). In the first decade of the 21st century total momentum reaches its second worst drawdown (-81,57%), from which in December 2015 it is still recovering.

The remaining strategies and especially scaled residual momentum minimize a lot the dimension and length of the drawdowns. While, residual momentum´s worst peak is -56,71%, scaled momentum and scaled residual momentum present negative records of -49,30% and -41,37% respectively. Also, while residual momentum and scaled momentum present drawdowns that take 14 to 16 years to recover from, the maximum length recorded for any scaled residual momentum' drawdown is of 7 years so far (it is still happening in December 2015, with a drawdown of 7,11%)

As explained in detail in Section 3.3, the scaled strategies differ from the non-scaled ones in terms of the way they are constructed. These strategies aim at having constant risk, varying their weights on the non-scaled strategies as in equation (8). In other words, their exposure to the non-scaled strategies varies across time. Thus, it is interesting to see how this exposure is distributed. Figure 5 presents the weights of scaled momentum and scaled residual momentum on total momentum and residual momentum respectively.

Figure 5 – Weights (full sample)

Figure 5 shows the weight scaled momentum has on total momentum and the weight scaled residual momentum has on residual momentum at any given month *t*. Stock returns are extracted from CRSP data base. ADRs, REITs, financials, close-end-funds and foreign shares are excluded. Only shares listed on NYSE, NASDAQ and AMEX are considered. Only periods (months or days) where stock price is above $1 are considered. Our sample period goes from November 1930 to December 2015.



The weights of both scaled strategies follow a very similar distribution, with a correlation of 0,86. However, scaled residual momentum reaches higher maximums and does not reach minimums as low as scaled momentum, being more exposed to its non-scaled strategy. While scaled momentum´s lowest is 0,20, scaled residual momentum´s lowest is 0,33. The highest weights are 2,74 and 3,64 for scaled momentum and scaled residual momentum respectively. As it would be expected the lowest weights occur in the 30s and in the first decade of the 21st century, which are the period with more instability and also the periods with the huge crashes of momentum. The average weights are 1,09 for scaled momentum and 1,59 for scaled residual momentum, which means that on average scaled momentum has virtually full exposure to momentum, while scaled residual momentum leverages quite extendedly its position.

Finally, it is important to refer that as stated in Barroso and Santa-Clara (2015) the turnover of the scaled strategies is similar to the turnover of the non-scaled strategies, not being a constraint, due to the scaled strategies' high profitability.

4.3. Main results: Performance Evaluation – the crashes

As already stated in this thesis and in previous work by several authors momentum has huge crashes from time to times (see for example Daniel and Moskowitz (2016), Barroso and Santa Clara (2015), Chaves (2012) or Blitz, Huitz and Martens (2011)), which result in the darkest times for the strategy. The two worst periods for momentum were the beginning of the 30s and the second half of the decade of 2000. Thus, in this section we examine in more detail the performance of all the four strategies in these two more problematic periods. In order to have a more realistic view we set an investment horizon of 60 months (5 years). Our samples are from November 1930 until October 1935 and from January 2005 until December 2009.

We start our analysis of the main results with an evaluation of some of the key performance measures: average return, volatility, Sharpe ratio, kurtosis, skewness, minimum and maximum monthly return and the probability of having a positive monthly return in a certain month. These results are summarized in Tables 3 and 4 (1930s and 2005s respectively).

Tables 3 and 4 - Key performance measures (1930s and 2005s)

Table 3 and 4 present the average return, volatility, Sharpe ratio, kurtosis, skewness, alpha, minimum and maximum monthly return and the probability of having a positive monthly return in a certain month for total momentum, residual momentum, scaled momentum and scaled residual momentum. Alphas are estimated as in equation (1). Stock returns are extracted from CRSP data base. ADRs, REITs, financials, close-end-funds and foreign shares are excluded. Only shares listed on NYSE, NASDAQ and AMEX are considered. Only periods (months or days) where stock price is above $1 are considered. Table 3' sample period goes from November 1930 to October 1935. Table 4' sample period goes from January 2005 to December 2009. All values are annualized except for the minimum and maximum monthly returns.

Table 3 – 1930s

Strategy	Max	Min	P(ret>0)	Mean	Vol	Sharpe	Kurtosis	Skewness	Alpha
Total MoM	18,83	-70,26	65,00	-2,35	56,86	-0,04	8,13	-2,44	9,88
Residual MoM	14,43	-38,31	71,67	14,32	31,27	0,46	7,24	-2,29	28,29
Scaled MoM	10,23	-18,40	65,00	4,30	18,88	0,23	3,16	-1,33	8,16
Scaled Residual MoM	10,13	-17,73	71,67	11,53	17,50	0,66	3,92	-1,62	18,56

Table 4 – 2005s

Strategy	Max	Min	P(ret>0)	Mean	Vol	Sharpe	Kurtosis	Skewness	Alpha
Total MoM	12,35	-46,23	63,33	-11,67	28,35	-0,41	15,45	-3,21	-13,37
Residual MoM	12,54	-20,32	55,00	0,16	16,77	0,01	5,35	-1,35	-1,55
Scaled MoM	8,45	-14,32	63,33	4,25	14,13	0,30	2,07	-0,80	2,49
Scaled Residual MoM	7,48	-10,64	55,00	4,41	12,15	0,36	0,91	-0,47	2,39

Total momentum is the strategy with the worst minimum monthly return in both samples. In August 1932 it has a return of -70,26% and in April 2009 it has a return of –46,23%. It is also the strategy presenting the highest maximum return in the 30s (18,83%) and the second highest in the 2005s (12,35%). Residual Momentum and, especially both scaled strategies display much less pronounced minimums. In the 30s, while residual momentum has a minimum monthly return of -38,31%, scaled momentum and scaled residual momentum present minimums of -18,40% and -17,73% respectively. In the 2005s residual momentum's minimum is -20,32% and scaled strategies' minimums range between -10,64% for scaled residual momentum and -14,32% for scaled momentum. In both decades scaled strategies' maximums are similar (around 10% in the 30s and 8% in the 2005s), while residual momentum's maximums range between 12,54% in the 2005s and 14,43% in the 30s.

Despite followed closely by scaled momentum, scaled residual momentum is in both samples the strategy with the tighter range of values for monthly returns and also the strategy presenting the lowest minimums. Later in this section, with the assessment of higher moments it will become even clearer that scaled residual momentum is able in both problematic periods to hedge against extreme monthly returns.

In what regards the distribution of monthly returns between positive and negative returns, total momentum and scaled momentum show that the probability of having positive monthly returns is 65,00% in the 30s and 63,33% in the 2005s . Residual momentum and scaled residual momentum show that the probability of having positive monthly returns is 71,57% in the 30s and 55,00% in the 2005s. So the probability of having positive returns is always higher in the 30s than in the 2005s, although this difference being more pronounced for strategies involving residual returns. However, as also happens in the full sample the probability of having positive returns is always considerately higher than 50%.

In terms of the risk reward relation, scaled residual momentum is superior to all the other four strategies in both samples. In the 30s it delivers more than 11,5% of return, with a volatility of 17,5% and in the 2005s it delivers 4,41% of return with a volatility of 12,15%. Despite these being the worst periods for momentum, scaled residual momentum has a 0,36 Sharpe ratio in the 2005s and an incredible 0,66 Sharpe ratio in the 30s. The worst strategy and the only delivering negative Sharpe ratio is total momentum. In the 30s it has return of -2,35 and huge volatility (56,86%), while in the 2005s it presents return of -11,67% and volatility of 28,35%. While in the 30s residual momentum performs quite well, with the highest return (14,32%) and the second highest Sharpe ratio (0,46) of the period, in the 2005s it's return of 0,16% keeps its' Sharpe ratio near to 0. Scaled momentum's returns are about 4% in both periods, its' volatility ranges from 14,33 to 18,88% and its' Sharpe ratio from 0,23 to 0,30 in the 2005s and in the 30s respectively.

As stated in Section 4.2, the higher moments of the distribution play a major role in the assessment of strategies trying to minimize the crashes (extreme returns in the left side of the distribution).

All the four strategies are negatively skewed. However, we find that skewness is less negative for scaled strategies than for non-scaled ones in both decades. Also, heavy tails (or high kurtosis) are no longer a huge problem, as in both decades, scaled strategies are able to minimize them. Our findings in these periods confirm the results in the full sample, which state that scaling results in much improved higher moments. Also, and in spite of showing low values for skewness and kurtosis in both samples, scaled residual momentum superiority is more visible in the 2005s, where it shows skewness and kurtosis of barely half of the ones shown by scaled momentum.

Looking at the annualized alphas we can take two main conclusions. First, the scaled strategies are the only ones capable of delivering positive alphas independent of the period of the sample. Secondly, scaled residual momentum is the strategy delivering the most interesting alphas independent of the period – while in the 2005s it delivers virtually the same alpha as scaled momentum in the 1930s it is able to deliver a much superior one, 18,56%.

Figure 6 and 7 - Cumulative returns (1930s and 2005s)

Figure 6 and 7 show the cumulative returns of total momentum, residual momentum, scaled momentum and scaled residual momentum. Stock returns are extracted from CRSP data base. ADRs, REITs, financials, close-end-funds and foreign shares are excluded. Only shares listed on NYSE, NASDAQ and AMEX are considered. Only periods (months or days) where stock price is above $1 are considered. Figures 6' sample period goes from November 1930 to October 1935. Figures 7' sample period goes from January 2005 to December 2009.

Figure 6 – 1930s

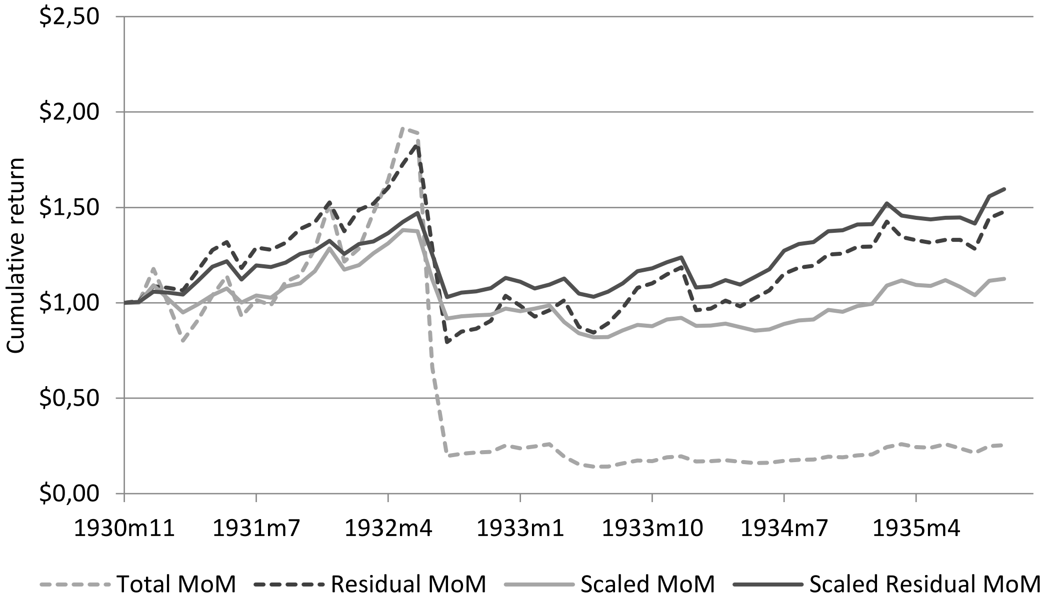
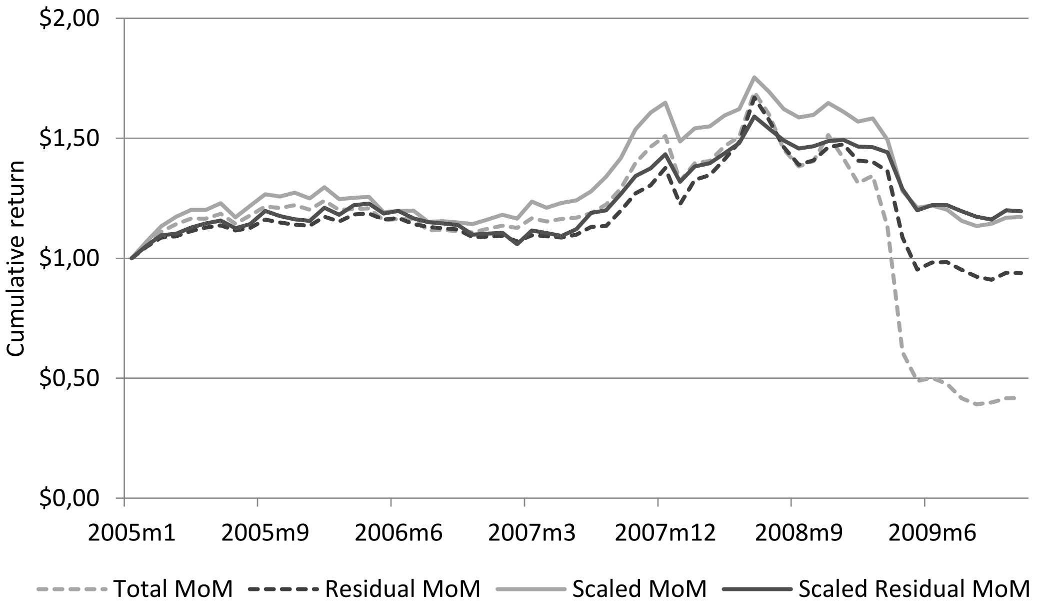


Figure 7 – 2005s



As observable in Figures 6 and 7, market timing is crucial for an investor in total momentum. An investor entering in the beginning of these two periods would see 75% of his investment wiped out in the 1930s and 60% in the 2005s. However, this is no longer true for the other 3 strategies. Both scaled strategies end the periods above the watermark. Scaled residual momentum is visibly the strategy performing better, never going below the initial investment and ending both periods with the highest cumulative return among the four strategies. In fact, with scaled residual momentum, an investor would see a valorization of 60% in the 1930s and of 20% in the 2005s. Where total momentum has huge losses, scaled residual momentum is able to make serious money.

In order to observe the dimension and length of the huge crashes that happen in these two periods it is presented the drawdown of all the four strategies in both samples. Figure 8 displays the drawdowns of the strategies in the 1930s and Figure 9 displays the drawdowns of the strategies in the 2005s.

Figures 8 and 9 – Drawdowns (1930s and 2005s)

Figures 8 and 9 present the drawdowns of total momentum, residual momentum, scaled momentum and scaled residual momentum. The drawdown is computed as the ratio between cumulative return in month *t* and the maximum cumulative return up to month *t*, to which is then subtracted 1. Stock returns are extracted from CRSP data base. ADRs, REITs, financials, close-end-funds and foreign shares are excluded. Only shares listed on NYSE, NASDAQ and AMEX are considered. Only periods (months or days) where stock price is above $1 are considered. Figures 8' sample period goes from November 1930 to October 1935. Figures 9' sample period goes from January 2005 to December 2009.

Figure 8 – 1930s

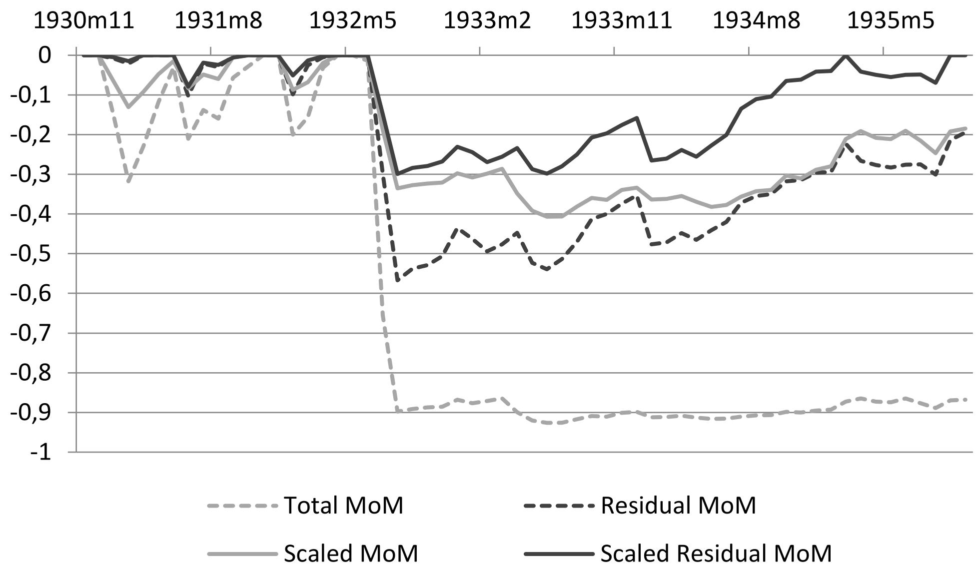
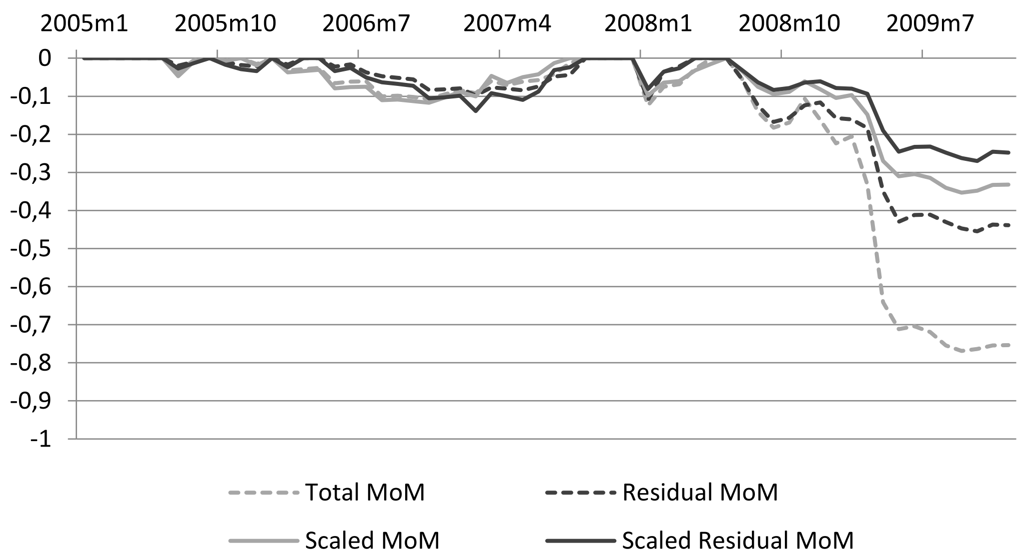


Figure 9 – 2005s



Figures 8 and 9 confirm the results found in the full sample. Total momentum is the strategy presenting the worst drawdowns in both samples. In the 1930s it reaches a minimum of -92,6% and in the 2005s a minimum of -76,9%. All the remaining 3 strategies are able to diminish these drawdowns quite extensively. The strategy doing this more successfully is scaled residual momentum followed by scaled momentum and residual momentum. In fact, scaled residual momentum is not only the strategy with the lowest maximum's drawdown in both samples, but it is also the strategy recovering from them faster. In 1935 it has already fully recovered from the crash of 1932 (only strategy able to do this) and by the end of the 2005s it is the strategy presenting the lowest drawdown.

As said in Section 4.2, the exposures of the scaled strategies to theirs peer non-scaled ones vary over time. Thus, it is interesting to see how these exposures are distributed. In Figures 10 and 11 are plotted the weights of scaled momentum and scaled residual momentum on total momentum and residual momentum (respectively) for the 1930s and the 2005s.

Figures 10 and 11 – Weights (1930s and 2005s)

Figures 10 and 11 show the weight scaled momentum has on total momentum and the weight scaled residual momentum has on residual momentum at any given month *t*. Stock returns are extracted from CRSP data base. ADRs, REITs, financials, close-end-funds and foreign shares are excluded. Only shares listed on NYSE, NASDAQ and AMEX are considered. Only periods (months or days) where stock price is above $1 are considered. Figures 10' sample period goes from November 1930 to October 1935. Figures 11' sample period goes from January 2005 to December 2009.

Figure 10 – 1930s

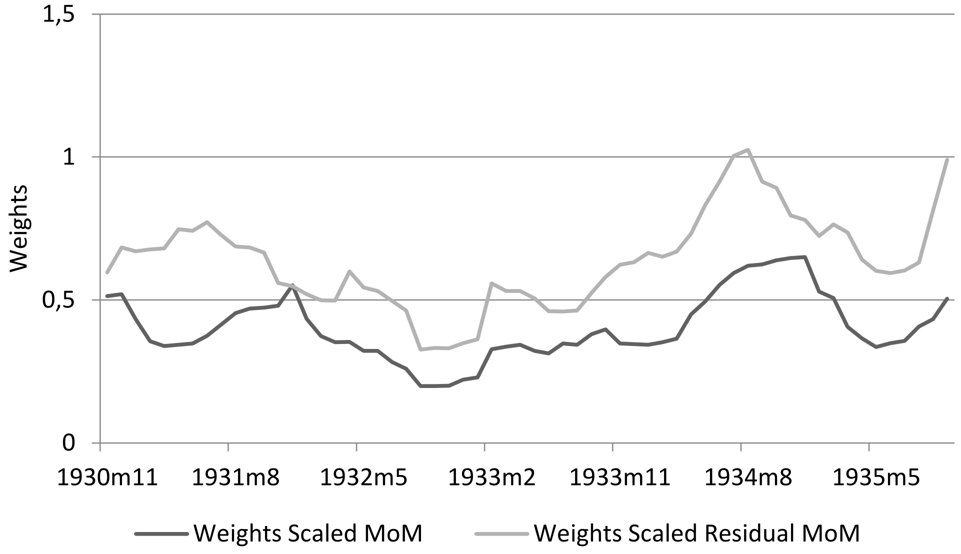
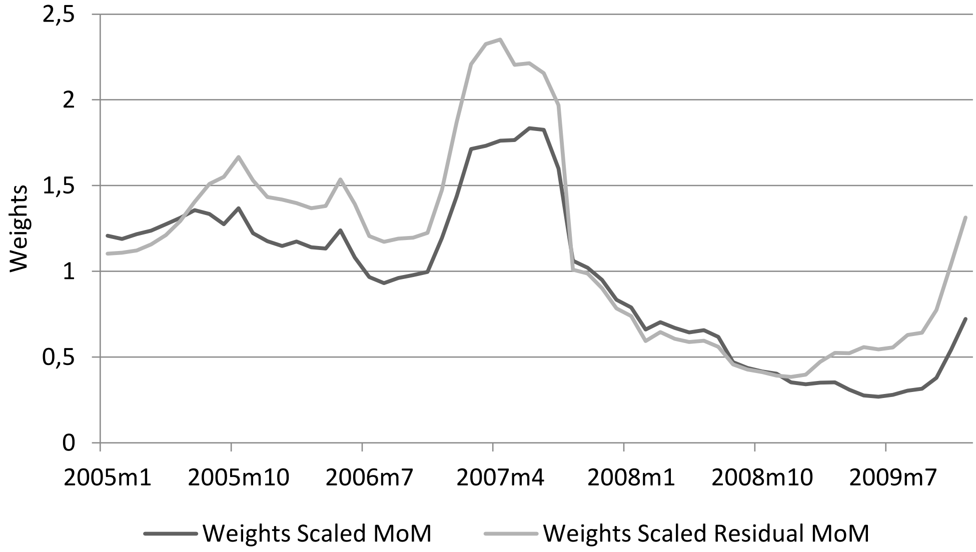


Figure 11 – 2005s



Three main conclusions may be drawn from the analysis of Figures 10 and 11. First, as it is observable in both samples, the weights are clearly bellow the averages of the full sample (1,09 for scaled momentum and 1,59 for scaled residual momentum) especially near the crashes of 1932 and 2009. It seems evident that the risk predictability of total momentum and residual momentum is being effective in these two samples, once we can observe that the weights decrease significantly a few months before the most intense crashes. Secondly, and confirming the results obtained for the full sample it seems obvious that scaled residual momentum is much more exposed to its' peer non-scaled strategy than scaled momentum. Thirdly and finally, once again the correlation between the weights of the scaled strategies is very high. In the 1930s it is 80%, while in the 2005s it is an impressive 94%.

5. Conclusions and recommendations

Many researchers find that information travels slowly and that it seems to exist an under-reaction to news (see Barberies, Schleifer and Vishny (1998), Hong, Lim and Stein (2000), Hong and Stein (1999)). As a response to these anomalies, several relative strength strategies appear, being momentum one of the most known and widely used nowadays.

Momentum buys winners (stocks with a good past performance) and sells losers (stocks with a bad past performance). The first definition was given by Jegadeesh and Titman (1993), who select stocks according to their returns in the past J months and hold the portfolio for K months. However many variants followed.

In spite of the good performance and high Sharpe ratio, momentum has a big problem. Due to its high kurtosis, it has huge crashes from times to times. Several strategies have been tested to cope with this issue. During this thesis we study residual momentum from Blitz, Huitz and Martens (2011), scaled momentum from Barroso and Santa-Clara (2015) and scaled residual momentum. The last strategy scales its exposure to residual momentum (instead of scaling to total momentum as done in scaled momentum from Barroso and Santa-Clara (2015)). The scaling relies on the risk predictability of residual momentum (as it uses its realized variance in the last 6 months in the scaling process).

We observe that both, total momentum and residual momentum, show risk predictability. The AR (1) regressions of the monthly realized variance show that the coefficients (β) of the realized variance of total momentum and residual momentum are both 0,56 and clearly significant (with t-statistics above 20). Also, the model R^2 and OOS R^2 show values above 30%. As a result, scaled momentum and scaled residual momentum can take advantage of the variance predictive power to scale their exposures.

In the full sample we conclude that total momentum is the most volatile strategy (24,78%), with lowest Sharpe (0,49), higher kurtosis (26,04) and lowest skewness (-3,23). All the other three improve considerably these measures. Never the less, scaled strategies show results clearly more attractive. Scaled momentum is the strategy with the highest cumulative return in the end of the 85 year period and also with the highest Sharpe (1,00), followed closely by scaled residual momentum (0,91). However, scaled residual momentum has the highest annualized alpha (14,74%) and seems to be controlling more effectively for the crashes. It is not only the strategy with lowest kurtosis (2,77) and highest skewness (-0,60), but also the one presenting lower and shorter drawdowns. Also, note that an 85 year investment horizon is far too large. For smaller investment periods investors are clearly apprehensive with the possibility of seeing their money wiped out in months (due to the big crashes). That is why it is also study the performance of the strategy in the two darkest periods of momentum: the 1930s and the 2005s.

In the 1930s and the 2005s we find that only parts of the results are consistent with the ones found in the full sample. Scaled strategies continue to be superior, but scaled residual momentum seems to stand out from the others. It is in both samples the strategy with the lowest volatility and the highest Sharpe. In has impressive Sharpe of 0,66 and 0,36 in the 1930s and 2005s respectively. It is also the one presenting more interesting alphas, as it is able to deliver positive, large and significant alphas in both periods, especially in the 1930s. Like in the full sample it presents low values for kurtosis and near zero for skewness and it is the strategy with lowest and shorter drawdowns.

In terms of cumulative returns, with total momentum an investor entering in the beginning of these two periods would see 75% of his investment wiped out in the 1930s and 60% in the 2005s. The superior performance of scaled residual momentum is impressive when we verify that for the same periods in the 1930s an investor would see a valorization of 60% and in the 2005s of 20%. Where total momentum has huge losses, scaled residual momentum is able to make serious money.

We might conclude that overall scaled residual momentum has a Sharpe slightly lower than scaled momentum, but in turbulent times this reverts completely. In turbulent times it becomes the strategy with highest Sharpe, lowest volatility and highest cumulative return. The reason for this is the evident superiority shown by scaled residual momentum in terms of controlling for the crashes. Scaled residual momentum is the strategy with lowest kurtosis highest skewness and highest alpha. It seems obvious that the strategy is able to virtually eliminate the dangers of the momentum crashes.

Despite the apparent solution for the momentum crashes presented in this thesis, there is still a lot to be done in the topic. First, our research uses only American stocks (NASDAQ, NYSE and AMEX). So, as several authors have done for total momentum (see for example Fama and French (2012) or Aness. Moskowitz and Pedersen (2013)), we encourage future researchers to test if our results hold for other markets and asset classes. Second, for the sake of brevity and comparison our research sets a formation period of 12-1 months and a holding period of 1 month in the construction of the non-scaled strategies. We suggest testing if the results hold with different definitions of momentum and residual momentum. Finally, our scaled strategies set a target volatility of 12% and use the predictability of realized variance in the previous 6 months to scale the exposures. We suggest finding if other horizons and other values for target volatility corroborate and/or are able to improve our results.

References

Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen. "Value and momentum everywhere." *The Journal of Finance* 68.3 (2013): 929-985.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. "A model of investor sentiment." *Journal of financial economics* 49.3 (1998): 307-343.

Barroso, Pedro, and Pedro Santa-Clara. "Momentum has its moments." *Journal of Financial Economics* 116.1 (2015): 111-120.

Blitz, David, Joop Huij, and Martin Martens. "Residual momentum." *Journal of Empirical Finance* 18.3 (2011): 506-521.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam. "Investor psychology and security market under-and overreactions." *the Journal of Finance* 53.6 (1998): 1839-1885.

Daniel, Kent, Ravi Jagannathan, and Soohun Kim. *Tail risk in momentum strategy returns*. No. w18169. National Bureau of Economic Research, 2012.

Daniel, Kent, and Tobias J. Moskowitz. "Momentum crashes." *Journal of Financial Economics* 122.2 (2016): 221-247.

Fama, Eugene F., and Kenneth R. French. "Size, value, and momentum in international stock returns." *Journal of financial economics* 105.3 (2012): 457-472.

Grinblatt, Mark, and Sheridan Titman. "Mutual fund performance: An analysis of quarterly portfolio holdings." *Journal of business* (1989): 393-416.

Grinblatt, Mark, and Sheridan Titman. "Performance measurement without benchmarks: An examination of mutual fund returns." *Journal of business* (1993): 47-68.

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

Hong, Harrison, and Jeremy C. Stein. "A unified theory of underreaction, momentum trading, and overreaction in asset markets." *The Journal of finance* 54.6 (1999): 2143-2184.

Hong, Harrison, Terence Lim, and Jeremy C. Stein. "Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies." *The Journal of Finance* 55.1 (2000): 265-295.

Jegadeesh, Narasimhan, and Sheridan Titman. "Profitability of momentum strategies: An evaluation of alternative explanations." *The Journal of finance* 56.2 (2001): 699-720.

Jegadeesh, Narasimhan, and Sheridan Titman. "Returns to buying winners and selling losers: Implications for stock market efficiency." *The Journal of finance* 48.1 (1993): 65-91.

Levy, Robert A. "Relative strength as a criterion for investment selection." *The Journal of Finance* 22.4 (1967): 595-610.

Okunev, John, and Derek White. "Do momentum-based strategies still work in foreign currency markets?." *Journal of Financial and Quantitative Analysis* 38.02 (2003): 425-447.

Rouwenhorst, K. Geert. "International momentum strategies." *The Journal of Finance* 53.1 (1998): 267-284.