ARCO VAN OORD

Essays on Momentum Strategies in Finance

Essays on Momentum Strategies in Finance

Essays on Momentum Strategies in Finance

Verhandelingen over momentum strategieën in de financiële economie

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. H.A.P. Pols

and in accordance with the decision of the Doctorate Board.

The public defence shall be held on
Thursday 12 May 2016 at 09:30 hrs

by

Jacobus Arie van Oord
born in Gorinchem.

Erasmus University Rotterdam

Erasmus

Doctoral Committee

Promoter: Prof.dr. H.K. van Dijk

Other members: Prof.dr. M.A. van Dijk
Dr.ir. M.P.E. Martens
Dr. L.F. Hoogerheide

Erasmus Research Institute of Management - ERIM
The joint research institute of the Rotterdam School of Management (RSM)
and the Erasmus School of Economics (ESE) at the Erasmus University Rotterdam
Internet: http://www.erim.eur.nl

ERIM Electronic Series Portal: http://repub.eur.nl/pub

ERIM PhD Series in Research in Management, 380
ERIM reference number: EPS-2016-380-F&A
ISBN 978-90-5892-444-5

Design: B&T Ontwerp en advies www.b-en-t.nl

This publication (cover and interior) is printed by haveka.nl on recycled paper, Revive®.
The ink used is produced from renewable resources and alcohol free fountain solution.
Certifications for the paper and the printing production process: Recycle, EU Flower, FSC, ISO14001.
More info: http://www.haveka.nl/greening

Acknowledgements

After a little less than ten years I am happy to present my PhD thesis. Taking this amount of time to finish my PhD thesis does definitely not imply I did it all on my own. On the contrary, I am thankful to a lot of people that contributed to finishing this thesis and I take this opportunity to express my gratitude.

First and foremost I would like to thank my promotor Herman van Dijk. If it was not for you, this thesis would not have been finished, definitely not yet. I am grateful for all your support, motivation and patience during the writing of this thesis. I enjoyed our more conceptual discussions on the inclusion of uncertainty in decision making processes in practice and on the usefulness of, mixed, models in theses decision processes. I would like to thank Martin Martens for his supervision during my four years at the Erasmus University and beyond. Your ideas and rigor have been of great help when doing the research and getting Chapter 3 actually published. Furthermore, I enjoyed our discussions on pensions and pension funds as well as on how to train young football players. I am thankful for the support of Lennart Hoogerheide; even if some topics were not closely related to your own research you always had ideas for further improvements. Last, but definitely not least, I thank Nalan Baştürk for jointly working on two chapters of this thesis. Without your contributions, in particular at the end, I would not have been able to finish my PhD thesis. In this respect, I am also thankful to Stefano Grassi for the joint work on one of the chapters. Furthermore, I would like to thank all members of the doctoral committee, in particular Mathijs van Dijk, for evaluating my thesis on a relatively short notice and providing me quickly with feedback to improve this thesis. I would also like to thank the staff of both ERIM and the Econometric Institute for all their support.

I would also like to thank my fellow PhD students, in particular, our 'lunch group' consisting of Kar Yin, Hans, Peter, Yuri, Eelco and Mathijn. Although most of us no longer meet at a daily basis we are still in touch and now meet at each other's weddings and PhD thesis defenses. I hope, although my PhD thesis defense is the last and the majority of us is already married, that we keep in touch. I thank you all for your neverending subtle and less subtle encouragements to finish my PhD. I especially thank Kar Yin, my roommate during the four years at the Erasmus University: I could not have wished for more pleasant company and I am happy that we are colleagues again. I thank Hans for being my best man, for futsal and for being there to discuss topics beyond a PhD, Peter for bringing laughter every time and every place we meet, Eelco for his sharp mind in any discussion, Yuri for proposing and organizing a wide range of fun activities and Mathijn for his, mostly just, skepticism on any topic.

I am grateful to my colleagues at DNB. In particular, Olaf Sleijpen, Bert Boertje, Lode Keijser, Cindy van Oorschot, Petra Hielkema and Mark Kruidhof for allowing me time off to work on my PhD thesis and encouraging me to finish it. Furthermore, I am indebted to all my current and former colleagues at the Risk and Asset Liability Management (RALM) and Insurance Supervision Policy departments for stepping in when I was working on this thesis. I would like to thank two colleagues more specifically; David Rijsbergen for his valuable comments on one of the chapters of this thesis and Jan Smit for showing his interest in my research and pointing me to relevant publications.

Finishing a PhD comes at the cost of spare time, but it cannot be finished without the distraction of spending some spare time with friends. I thank Edwin, Jan-Joost, Niels and Hendrik for the distraction of various long game nights, Robert and Joep for a week of distraction in Bulgaria and subsequently remembering that week on multiple occasions over one or more beers, Vincent for discussing financial markets over several good meals, Kenneth for the distraction of all the good to very bad movies we have enjoyed and Erik for discussing football.

I thank my parents in law, Ries and Leny, for providing me with extra time to work on this thesis by letting Finn and Mila stay over for multiple nights at their grandpa and grandma as well as the holidays to recharge. This recharging would of course not have been sufficient if you, Vincent and Emma, would not have taken me to that aqua theme park in Portugal. I also thank 'tante' Jo for looking after Finn and Mila on several Mondays. Many thanks to Iwan and Nena, you are much more than my brother and sister in law: you were always there to help and to provide some distraction with football. I am very grateful to my parents, Arie and Matty, for their support and encouragement: you are the best parents I can imagine and I can only wish that I can offer Finn and Mila what you have offered me.

In the end, I am most indebted to my wife Céline and Finn and Mila. You are the best that ever happened to me. I am grateful for all your love, support and cheerful distraction when I needed it the most. I look forward to everything that will come our way.

Arco van Oord
Gorinchem, 31 March 2016

Contents

Chapter 1

Introduction

1.1 Motivation

Equity momentum is the phenomenon that the stocks that recently have outperformed other stocks continue to outperform these stocks in the coming months. Likewise, stocks that have underperformed continue to underperform in the coming months. Jegadeesh and Titman (1993) show that a strategy that is long in the stocks that have recently performed well and short in the stocks that have underperformed has positive average returns.

Equity momentum has turned out to be a profitable strategy, available to many investors. Using just historical returns of the stocks in one's investment universe is sufficient to exploit equity momentum. Academics struggle to explain momentum's performance as a reward for risk and as such equity momentum challenges the efficient market hypothesis. Nobel laureate Eugene Fama even refers to momentum as the biggest challenge to market efficiency[1].

Despite momentum's positive average returns, it exhibits substantial crash risk. A crash of just several months wipes out years of positive average returns; in this sense, momentum is not an attractive strategy for a risk-averse investor, or even any investor at all. Next to this crash risk, momentum also exhibits time-varying risks and returns, in particular over the business cycle.

[1]See, for example, `http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2013/fama-lecture.html` for Eugene Fama's Nobel prize lecture.

All these aspects make momentum less attractive for investors and support the efficient market hypothesis that momentum's positive average returns are a compensation for specific types of risks. In this thesis we explore several models and methods that are able to reduce momentum's crash risk as well as its time-varying risks and returns. Moreover, these strategies are able to improve momentum's performance by increasing the Sharpe ratio. As such, results are interesting for practitioners as well as academics who study momentum's challenge of the efficient market hypothesis.

In the three essays in this thesis we look at equity momentum in different ways. In Chapter 2 we compare and combine alterations of the equity momentum strategy, while in Chapter 3 our approach is to leave the equity momentum strategy as it is and investigate if, and how, hedging is able to improve momentum's performance. Both chapters aim at reducing momentum's time-varying risks and returns as well as its crash risk while improving, or at least maintaining, momentum's returns. In Chapter 4 we look at the time variation in the comovements of stock returns using a Bayesian latent factor model and subsequently apply this model to a momentum strategy. This strategy performs relatively well in times of market turbulence when equity momentum itself experiences negative returns. As such, all essays address time variation in general, and that of momentum more specifically, in different ways.

1.2 Summary of Methods and Results

Equity momentum thus exhibits substantial crash risk. Daniel and Moskowitz (2013) and Daniel, Jagannathan, and Kim (2012) among others state that momentum's average positive returns are a compensation for its crash risk. In Chapter 2 and Chapter 3 we show that optimization, combination and hedging strategies are able to substantially reduce momentum's crash risk without affecting its positive average returns. As such we weaken this crash risk explanation of momentum.

Momentum's crash risk is a consequence of its time-varying exposures. Grundy and Martin (2001) show that, by construction, momentum has time-varying exposures to the three Fama and French (1993) equity risk factors. Although momentum's average exposures to these risk factors are close to zero, the momentum strategy's exposures to these factors do vary substantially over time, depending on the recent returns on the equity risk factors. If the recent returns on a particular factor have been negative, momentum is negatively exposed to this particular factor. If the return on this factor then reverses to a positive return the momentum strategy suffers from its negative exposure to this factor. Actually, momentum's largest losses occurred in July 1932 and March 2009 after equity markets strongly recovered from two of the largest losses on equity markets ever. In Chapter 3 we show that hedging momentum's time-varying exposures to the equity risk factors indeed reduces momentum's crash risk.

Another consequence of momentum's time-varying exposures is that its performance varies over the business cycle, as shown by Chordia and Shivakumar (2002). Momentum suffers when the return on the equity risk factors reverses and these reversals occur more often during specific stages of the business cycle. Most prominent are momentum's large losses at the end of recessions when equity markets already start to recover from their previous losses, even before the real economy recovers and starts expanding again. Hedging momentum's time-varying exposures thus also reduces momentum's return variation over the business cycle, see Chapter 3.

The standard equity momentum strategy does not take into account the risks of the constituent stocks; it is a strategy that simply attaches equal weights to the stocks in the top and bottom of momentum's ranking. It would be straightforward to apply Markowitz (1952) mean-variance optimization to make a trade-off between the expected returns and risks in the momentum strategy. However, several claim that mean-variance optimization does not provide a better risk-adjusted performance. For example, Michaud (1989) refers to mean-variance optimization as 'error maximization' and DeMiguel, Garlappi, and Uppal (2009) find that equally weighted equity portfolios outperform optimized portfolios. Nevertheless, in Chapter 2 we find that applying mean-variance optimization to the equity momentum strategy does improve momentum's performance, in particular, when combined with other improvements of equity momentum.

Momentum's top-bottom phenomenon is key to take into account when applying standard mean-variance optimization. The momentum phenomenon of continuing outperformance of recent outperforming stocks only holds for the stocks that are ranked high according to their recent performance. Stocks in the middle of momentum's ranking actually do not outperform other stocks in the middle of this ranking, although their recent performance might be higher than that of other stocks in the middle of momentum's ranking. Similarly, only the lowest ranked stocks continue to underperform in the coming months. When using the stocks' recent performance as input for the expected returns, mean-variance optimization suffers since the differences in expected returns of stocks in the middle of momentum's ranking are not reflected in the stocks' realized returns. In Chapter 2 we find that taking into account momentum's top-bottom phenomenon further improves momentum's performance.

Optimization of equity momentum also reduces momentum's time-varying exposures and thereby its crash risk and return variation over the business cycle. Mean-variance optimization is a trade-off between the risks and returns of the stocks in the investment universe. As such, it does take into account the risks associated with the stocks in the momentum strategy; at times when the equally weighted momentum strategy is more exposed to the equity risk factors, standard mean-variance optimization selects the stocks with relatively less risk and less exposures to the equity risk factors. In Chapter 2 we indeed find that the optimized momentum strategy has less time-varying exposures and therefore also exhibits less crash risk and performs more stable over the business cycle.

Recent alterations of the equity momentum strategy have also improved momentum's performance. In particular, residual momentum by Blitz, Huij, and Martens (2011) and constant-risk momentum by Barroso and Santa-Clara (2015) outperform the standard momentum strategy. Residual momentum is also an equally weighted zero net-investment strategy. Rather than ranking stocks on their recent total returns, residual momentum ranks stocks on their residual returns from the regression of the stocks' total returns on the Fama and French (1993) equity risk factors. Constant-risk momentum adjusts the exposure of the standard momentum strategy to address its time-varying risks. If standard momentum's recent returns have been volatile, constant-risk momentum reduces the exposure to the standard momentum strategy, while it increases the exposure when momentum returns have recently been less volatile. In Chapter 2 we compare these alterations of equity momentum as well as our optimized momentum to the standard momentum strategy. We find that residual momentum performs best in terms of Sharpe ratio, but optimized momentum exhibits the least crash risk. In Chapter 2 we also show that combining alterations of the equity momentum strategy outperforms both standard momentum as well as our optimized and the residual and constant-risk alternatives, both in terms of Sharpe ratio as in terms of crash risk. Using a bootstrap test on the differences between the Sharpe ratios of the different strategies we find that the Sharpe ratio of the combined strategy is also significantly higher than those of standard, optimized, residual and constant-risk momentum strategies over the full sample. Although mean-variance optimization often increases transaction costs we find that the outperformance of specific optimized and combined momentum strategies does not come at the cost of substantial additional transactions compared to standard and residual momentum. The combined strategies consistently outperform the standard and constant-risk momentum strategies over different subsamples between 1929 to 2014, but the residual momentum strategy performs best both in terms of Sharpe ratio and crash risk in the most recent subsample starting in 1973; however, residual momentum's Sharpe ratio is not significantly higher than those of the combined strategies during these subsamples.

Throughout this thesis equity risk factors play a role in different places: in the factor covariance matrix used for the optimization strategy, in the time-varying risks and exposures of momentum and in the different altered and combined momentum strategies. As such, we zoom also zoom on factor models in Chapter 4, in particular, on a Bayesian latent factor model. We show that conditional posteriors of the latent factors and the factor loadings are well-known distributions if we use flat or conjugate priors. We discuss the well-known issue that the number of factors in a static factor model are typically fixed. Using predictive likelihoods we are able to determine the optimal number of latent factors. We find that the optimal number of latent factors varies considerably over time.

We also apply the Bayesian latent factor model in a momentum based strategy and make use of the fact that this Bayesian model takes the uncertainty with respect to the parameters and the unknown latent factors into account when forecasting. In particular, we apply the Bayesian latent factor model to a momentum based strategy for ten industry portfolios. We find that selecting the best industry every year from ranking the stocks on their recent residuals from the latent factor model with time-varying number of factors performs well, in particular in comparison to a similar strategy based on static latent factors. We find that the choice for the size of the training sample is important for the performance of the strategy.

The Bayesian latent factor model application to residual industry momentum particularly outperforms other residual industry moment strategies in turbulent times, for example, during the crisis that started in 2008. We hypothesize that a Bayesian latent factor model adjust more quickly to big shocks than static models with known rather than latent factors, like, for example, the Fama and French (1993) equity risk factor model. During the more quiet times, like in the nineties, the applications of other static models with known factors also perform well. We therefore hypothesize that time-varying combinations of model structures and momentum strategies improve momentum's performance, though at the cost of increased complexity.

1.3 Outline

This thesis consists of two parts. The first part consists of Chapter 2 and Chapter 3 and contains empirical research on the improvement of the equity momentum strategy. Chapter 2 is based on Baştürk, Hoogerheide, Van Dijk, and Van Oord (2016). Chapter 3 has been published as Martens and Van Oord (2014). In these two chapters we explore whether optimization and hedging as well as combining different momentum strategies improve momentum's performance. The second part, in Chapter 4, is based on Baştürk, Grassi, Hoogerheide, Van Dijk, and Van Oord (2016) and it contains the derivation of a Bayesian latent factor model and its application to a momentum based strategy on ten industry portfolios.

The contribution of Chapter 2 is twofold: first, it shows the impact on momentum's performance if we apply mean-variance optimization and, secondly, it combines this optimization together with other alterations of the equity momentum strategy like residual and constant-risk momentum. We evaluate the performance of the optimized and combined momentum strategies in terms of standard equity performance measures. like the Sharpe ratio. as well as their crash risk characteristics. like the maximum monthly loss and the maximum drawdown. The crash risk characteristics play an important role in evaluating momentum strategies since standard momentum is known for its crash risk. With respect to combining different momentum strategies into new strategies we find that specific combinations outperform the standard momentum strategy as well as residual, constant-risk and optimized momentum strategies, without significantly increasing transaction costs.

Chapter 3 focuses on hedging momentum's time-varying risks. We apply two feasible hedging strategies to reduce momentum's time-varying exposures to the equity risk factors and as such aim at reducing momentum's time-varying risks. Since standard momentum's time-varying exposures result in substantial crash risk as well as variation in the performance over the business cycle we particularly focus on these characteristics after hedging. We indeed find that hedging the time-varying exposures of equity momentum reduces its crash risk as well as its performance variation over the business cycle. It turns out that a conditional hedging strategy using momentum's recent time-varying returns to estimate the time-varying exposures performs better than the commonly used hedging strategy aggregating the individual stocks' exposures that constituent the momentum strategy at that particular moment.

The second part of this thesis consists of Chapter 4 and contains the derivation of a Bayesian latent factor model. In particular, we focus on how to determine the optimal number of factors in a latent factor model. We use the predictive likelihoods of the latent factor models, that only differ in the number of factors applied, to determine the optimal number of factors. We find that the optimal number of latent factors varies considerably over time. We also apply the latent factor model to a residual momentum based strategy of ten industry portfolios. Compared to static models with known factors the Bayesian latent factor model performs particularly well in turbulent times.

1.4 Declaration of Contribution

In this section I declare my contribution to the different chapters of this thesis and acknowledge the contributions of others.

Chapter 1: I have written this chapter independently and implemented feedback that I received from my supervisor.

Chapter 2: This chapter is based on Baştürk, Hoogerheide, et al. (2016). The ideas on mean-variance optimization of standard momentum and combining different strategies were discussed in joint meetings, where I participated as a full member. I have performed the data gathering, programming and much of the analyses of this chapter. The writing is done in collaboration with my co-authors. I have implemented valuable comments from the members of the doctoral committee which also inspired me to add further empirical analyses. A modified version of this chapter will be submitted for publication in an international journal.

Chapter 3: This chapter is based on Martens and Van Oord (2014) that appeared in the Journal of Empirical Finance. Developing the ideas of this chapter was a joint project with the co-author, where I fully participated. I have programmed all analyses in this chapter. The writing of this chapter is joint work with the co-author.

Chapter 4: This chapter is based on Baştürk, Grassi, et al. (2016). Preliminary ideas were a continuation of a project of my co-authors on cointegration models. I fully participated in the derivation and double checking of the new theoretical results. I have participated in the writing jointly with my co-authors. I also contributed by implementing the valuable comments that we received from the members of the doctoral committee. The results of this chapter will be published in a book that will appear by Springer Verlag.

Chapter 2

Equity momentum strategies at work *

2.1 Introduction

Exploiting momentum in a portfolio of stocks has been a profitable and therefore popular strategy among investors over a long period of time. At the same time, explaining momentum's positive performance in connection with the efficient market hypothesis has turned out to be a great challenge for academics. Nobel laureate Eugene Fama refers to momentum in his Nobel prize lecture as the biggest challenge to market efficiency [1].

The standard equity momentum strategy ranks stocks on their recent returns, skips a particular short-term period to overcome short-term return reversals and then buys the stocks in the top of the ranking and short-sells the stocks in the bottom of this ranking. The strategy applies equal weights to the stocks. Jegadeesh and Titman (1993) show that this zero net-investment strategy yields positive average returns.

In this chapter we explore the working of several extensions of standard momentum. We start with mean-variance optimization of equity momentum with a focus on momentum's top-bottom phenomenon. This top-bottom phenomenon refers to the fact that stocks in the top and bottom of momentum's ranking out- and underperform the other stocks, while stocks ranked in the middle do not outperform each other consistently over the total

[*] This chapter is based on Baştürk, Hoogerheide, et al. (2016), see also Van Oord (2015)

[1] See `http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2013/fama-lecture.html` for Eugene Fama's Nobel Prize lecture

time period considered.

Apart from the proposed mean-variance optimization we investigate two other recent extensions of equity momentum: the residual momentum strategy by Blitz et al. (2011) and, as we call it, the constant-risk momentum strategy by Barroso and Santa-Clara (2015).

The working of these different momentum strategies is evaluated by using common return characteristics like the average realized return, realized volatility and Sharpe ratio, but also using several crash risk features. Equity momentum is well-known for its crash risk: in 1932 and 2009 this strategy lost 95 and 83 percent, respectively, in just three months. For this reason, Daniel et al. (2012) and Daniel and Moskowitz (2013) state that momentum fits the efficient market hypothesis as its performance is a reward for this crash risk.

A first conclusion from our empirical analysis is that the three suggested strategies outperform standard momentum in terms of higher Sharpe ratios and less crash risk. When one compares these strategies among each other a mixed result appears. In terms of a higher Sharpe ratio residual momentum performs best, while in terms of crash risk, the optimization strategy is best.

Therefore, we continue the analysis using several combinations of the mean-variance optimized, residual and constant-risk momentum strategies. Empirical results indicate further improvements in terms of return and crash risk characteristics for these combined momentum strategies over the whole period 1929 - 2014.

Given these improvements in momentum's Sharpe ratio, we also investigate whether these improvements are statistically significant. We do not only test if a Sharpe ratio is significantly larger than zero, but also whether the Sharpe ratio of each strategy is significantly higher than the Sharpe ratios of standard momentum and other improvements of momentum. The Sharpe ratios of thecombination strategies turn out to be higher while they also havebetter performance in terms of crash risk. However, we note that this holds forthe whole period considered.

In order to explore the effect of transaction costs, we check whether the different proposed strategies and their combinations lead to an increase in the transaction costs that may adversely effect their profitability. We find that the performance of most strategies is not very sensitive to reasonable changes in the amounts of transaction costs. Specifically, the number and average size of the transactions of the different strategies do not substantially increase compared to the standard momentum strategy, except for the constant-risk strategies.

The results, described so far, hold for the whole period 1929 - 2014. A major concern with respect to the performances of equity momentum strategies is their sensitivity to special data features as turbulence and volatility that may occur in particular subperiods.

We explore the time variation in the exposures of the different strategies to the Fama and French (1993) equity risk factors as already indicated by Grundy and Martin (2001). Another concern is the variation in return over the business cycle. A reduction in the time-varying exposures to the equity risk factors is obtained for the optimized momentum and the combination strategies. This also implies lower variation in the returns over the business cycle.

The sensitivity of the performance of the different strategies is also explored in a time series and episode analysis. Here we investigate how the strategies differ over the pre-World-War-II period where a major economic crisis occurred; over the post World-War-II period where a restoration of the economies of the industrialized nations took place and as a third episode we take the period after the first oil price shock in 1973.

Our results indicate that different strategies have different performances over these three episodes. The Sharpe ratios of the combination strategies are rather robust for shocks and turbulence and these strategies also exhibit less crash risk. The Sharpe ratio of the combination strategies is also significantly higher than those of standard momentum and other momentum improvements. For moderate transaction costs the combination strategies also keep the highest Sharpe ratio. However, the residual momentum strategy performsbest in terms of Sharpe ratio and several crash risk criteria in the thirdepisode after the oil price shock in 1973, although its Sharpe ratio isnot significantly higher than those of the combination strategies.

The contents of this chapter is organized as follows. Section 2.2 presents the data, the standard equity momentum strategy and the stylized fact regarding momentum's top-bottom phenomenon. Section 2.3 continues with different implementations of mean-variance optimization of momentum taking account of this top-bottom phenomenon. Section 2.4 contains a comparison of the performance of the optimized, residual and constant-risk momentum strategies. Section 2.5 contains the results of several combination strategies. In Section 2.6 we deal with the effect of introducing transaction costs. Section 2.7 presents the analysis and results of the time-varying exposures of the different strategies and Section 2.8 contains an analysis of the different momentum strategies over three important episodes. Section 2.9 contains conclusions.

In the appendices we present some technical details and some robustness checks. Appendix 2.A contains the details of the closed form solution used in the mean-variance optimization. In Appendix 2.B we discuss the robustness of the results if we use different quantiles for the selection of the stocks in the top and bottom deciles of the momentum strategies. Appendix 2.C describes a bootstrap test that we apply to the Sharpe ratios of the different strategies as well as their differences. We make use of a block bootstrap which is a simplification of the Ledoit and Wolf (2008) bootstrap test for Sharpe ratios. We use this bootstrap rather than the Jobson and Korkie (1982) test with the Ledoit and Wolf (2003) adjustment since this latter test is not valid for non-normally distributed time series like we have.

The uncertainty of our results is an interesting issue. A major analysis of parameter and model uncertainty and their effect on the performance of the different strategies is outside the scope of this chapter. A first attempt to deal with this feature is presented in Chapter 4. We take only an initial step in this direction by investigating whether the Sharpe ratios of the different strategies are significantly larger than zero.

2.2 Data, standard momentum strategy and top-bottom phenomenon

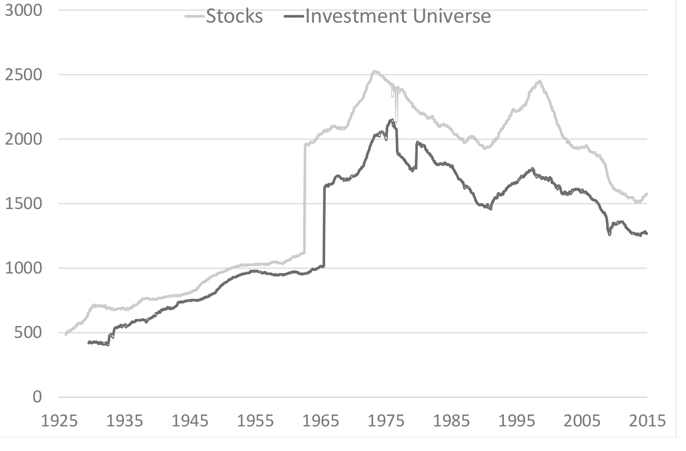
The data contain all monthly NYSE and AMEX common equity returns from the CRSP database[2] from January 1926 until December 2014. In addition we use the three Fama and French (1993) equity risk factors: market, size and value, from July 1926 until December 2014 and the 1 month T-bill return from January 1926 until December 2014. Every

[2]Stocks with exchange code equal to 1 or 2 (NYSE and AMEX respectively) and share code equal to 10 or 11 (representing common shares and as such ruling out ADRs, REITs etc.)

investment month the investment universe contains stocks that have a stock price higher than 1$ to overcome possible noise from the relative large tick sizes for these stocks[3]. We also require stocks to have at least a return history of 36 consecutive months as we require 36 months to estimate the stocks' loadings on the Fama and French (1993) equity risk factors.

The number of stocks included in the analysis varies over time. Figure 2.1 presents the number of stocks at each time period, with and without the requirement of a 36 month return history. Without adjusting the universe for the 36 month history the number of available stocks runs from 484 in January 1926 to 1576 stocks in December 2014; when adjusting the investment universe for a 36 month return history for each stock the number of stocks runs from 418 in July 1929 to 1269 in December 2014.

Figure 2.1: **Investment universe.** Number of stocks in the investment universe after adjusting for the 36 month return history requirement versus the total number of stocks from January 1926 - December 2014.



2.2.1 Standard momentum strategy

The standard momentum strategy we apply sorts the stocks in the investment universe based on their past six months performance on returns, skips one month and subsequently invests in the top decile of this ranking and short-sells the stocks in the bottom decile.

[3]We follow Pastor and Stambaugh (2003) who exclude stocks with prices lower than 5$; our results are robust for the exclusion of stocks with prices below 5$, below 1$ or no exclusion at all.

An equal weighting scheme is used for the stocks in the top and bottom deciles.

Table 2.1 presents the return and crash risk characteristics of this strategy between July 1929 and December 2014. The period of analysis starts in July 1929 since we require the stocks in the investment universe to have at least 36 months of return history. Standard equity momentum itself does not require a particular return history, but we do in order to obtain a fair comparison with other strategies. Both mean-variance optimization and residual momentum strategies require, for instance, the estimation of the stocks' factor loadings using the Fama and French (1993) equity risk factors and 36 months are required to do so with sufficient accuracy. July 1929 is the first month that 36 months of equity risk factor returns are available.

Table 2.1 displays several performance characteristics of the standard momentum strategy. We specifically look into two types: standard characteristics and crash risk ones. Standard performance characteristics are the annualized mean, annualized volatility and annualized Sharpe ratio, which practitioners commonly use. Crash risk characteristics are particularly important for the evaluation of equity momentum performance. Daniel and Moskowitz (2013) and Cooper, Jr., and Hameed (2004) state, for instance, that equity momentum's performance is a reward for its crash risk. We make use of the following four crash risk characteristics: largest monthly loss, defined as the largest loss in a single month over the whole period; realized 1% shortfall, defined as the average loss over one percent of the months with the worst returns (this reduces to the average over the 10 worst months of our sample of 1026 months); maximum drawdown, defined as the largest cumulative loss over the whole period; and finally, maximum recovery period, defined as the maximum time, measured in months, that it takes to recover from a drawdown to the level of the previous peak. The number of stocks in each case are in line with Figure 2.1.

Table 2.1 indicates that the obtained Sharpe-ratio is not very large, mainly due to the large volatility and that standard equity momentum exhibits considerable crash risk. A maximum drawdown of around 300 percent and a maximum recovery period of more than 400 months, approximately 35 years, are clear indicators of crash risk. We note that requiring a shorter return history for the stocks to be eligible for the momentum strategy gave minor differences in terms of risk and return characteristics.

These results are such that it appears worthwhile to investigate improvements over standard equity momentum both from an academic and a practical perspective.

Table 2.1: **Return and crash risk characteristics of the standard momentum strategy.** The return characteristics are the annualized mean, annualized volatility and annualized Sharpe ratio over the period July 1929 to December 2014. The crash risk characteristics are the largest loss in a single month, the maximum drawdown over the whole period and the maximum period to recover from a drawdown to the level of the prior peak.

Characteristics	Results
Annualized Mean	6.5%
Annualized Volatility	22.2%
Annualized Sharpe Ratio	29.3%
Largest Monthly Loss	-57.3%
1% Expected Shortfall	-38.1%
Full Period Maximum Drawdown	323.6%
Maximum Recovery Period in Months	430

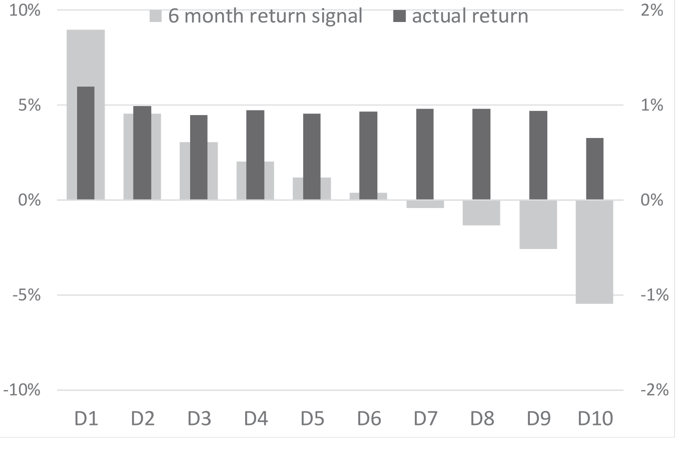
2.2.2 Momentum's top-bottom phenomenon

Momentum strategies are based on the expectation that past winners will continue to be winners, and similarly, past losers will continue to be losers. The standard strategy outlined in the previous paragraph is based on a top-bottom phenomenon, where the relation between past and future returns are assumed for the top and bottom deciles.

A justification for using this top-bottom phenomenon is presented in Figure 2.2 for the average monthly performance of ten deciles, where the averages of six months lagged returns based on the past performance of stocks are shown. These may be labeled as 'signals' and are used as inputs for standard momentum strategy. For comparison, the actual realized returns are also shown. The similarity between averages of past returns and actual returns is seen at the top and bottom deciles: Decile 1 is clearly characterized by relatively higher past returns and relatively higher actual returns. Similarly, decile 10 is characterized by relatively lower past returns and relatively lower actual returns. Using the top and bottom deciles and the past performance of stocks are surely suitable for portfolio construction.

However, Figure 2.2 further shows that all remaining deciles have a similar average realized return of approximately 0.9 percent per month but the averages of the past returns have a typical decreasing pattern. This substantial difference between the averages of past returns and actual returns for the stocks in deciles 2-9 may potentially have an adverse effect on mean-variance optimization using all deciles. In the following sections, we outline a methodology which incorporates momentum's top-bottom phenomenon in mean-variance optimization.

Figure 2.2: **Actual versus lagged returns.** Average six month lagged return signals (left axis) used for standard momentum's ranking and the actual realized returns after one skip month (right axis) over July 1929 - December 2014.



We emphasize that this data feature holds for the period from July 1929 to December 2014. In Section 2.8 we investigate the robustness of this feature for particular subperiods.

2.3 Mean-variance optimization under the top-bottom phenomenon

In this section we start with a brief description of standard mean-variance optimization. Markowitz (1952) puts forward mean-variance optimization as a way to construct equity portfolios that ex-ante trade-off the risks and returns of the stocks. This is different from, for example, an equally weighting scheme like the standard equity momentum strategy and in theory mean-variance optimization would yield better risk adjusted returns. However, practitioners are not keen on mean-variance optimization as it results in portfolios with extreme long and short positions in stocks with high and low expected returns or attractive and unattractive risk characteristics. If a stock's expected return is over- or underestimated its weight in the optimized portfolio is also over- or understated. Given this sensitivity to extreme observations, Michaud (1989), for example, refers to mean-variance optimization as 'error maximization'. Since equity momentum appears to be more or less right about the expected returns in the top and bottom deciles, we explore whether mean-variance optimization with expected returns based on this data feature results in better risk-adjusted returns.

2.3.1 Mean-variance optimization

This section presents mean-variance optimization for the construction of momentum portfolios. Every month t, we construct a portfolio h_t, a $N_t \times 1$ vector of weights for the N_t stocks in the investment universe at that time. These portfolios are based on the mean-variance optimization taking account of the top-bottom phenomenon. The constructed portfolios are long-short portfolios, where the sum of positive and negative weights add up to one. We choose this restriction of 1$ long and 1$ short to have a fair comparison with the standard momentum strategy outlined in Section 2.2.1, which is also a long-short strategy with 1$ long and 1$ short exposure.

In order to ease the technical steps of the optimization, we first construct an initial portfolio h_t^* that does not satisfy the constraint that both positive and negative weights add up to one. We find the initial portfolio h_t^* by minimizing the portfolio variance given unit exposure to the expected returns f_t. The optimization problem is formulated as follows

$$\min_{h_t^*} \left\{ h_t^{*\prime} V_t h_t^* \left| h_t^{*\prime} \iota = 0, h_t^{*\prime} f_t = 1 \right. \right\} \tag{2.1}$$

where V_t is the $N_t \times N_t$ covariance matrix and ι is an $N_t \times 1$ vector of ones. In Appendix 2.A we present the derivation of the following solution to this minimization problem

$$\widehat{h_t^*} = \frac{V_t^{-1} f_t - f_t' V_t^{-1} \iota (\iota' V_t^{-1} \iota)^{-1} V_t^{-1} \iota}{f_t' V_t^{-1} f_t - (f_t' V^{-1} \iota)^2 (\iota' V_t^{-1} \iota)^{-1}} = \frac{V_t^{-1} (f_t - C_{2,t} \iota)}{C_{1,t}} \tag{2.2}$$

where $C_{1,t} = f_t' V_t^{-1} f_t - (f_t' V_t^{-1} \iota)^2 (\iota' V_t^{-1} \iota)^{-1}$ and $C_{2,t} = f_t' V_t^{-1} \iota (\iota' V_t^{-1} \iota)^{-1}$ are time-varying scalars.

Equation 2.2 can be written as follows:

$$\widehat{h}_t^* \quad = \quad w_{1,t}\left(V_t^{-1}f_t - w_{2,t}V_t^{-1}\iota\right) \tag{2.3}$$

where $w_{1,t} = \frac{1}{C_{1,t}}$ and $w_{2,t} = C_{2,t}$. As such, the optimized portfolio is a weighted average of scaled and demeaned expected returns f_t where the weights depend on the inverse of the variance-covariance matrix V_t. We note that $V_t^{-1}f_t$ refers to the vector of Sharp ratios of returns and that $V_t^{-1}\iota$ refers to the inverse of the variance, that is, the precision of returns.

We transform the initial portfolios obtained from Equation 2.2 such that the optimized long-short portfolios have 1$ long and 1$ short exposure:

$$\widehat{h}_t = \frac{2\widehat{h}_t^*}{\left|\left|\widehat{h}_t^*\right|\right|_1}, \tag{2.4}$$

We note that this scaling does not affect the zero net-investment constraint for a long-short portfolio and also does not change its ex-ante Sharpe ratio.

An important concern in obtaining the weights using equation (2.2) is to ensure that the covariance matrix V_t is invertible. Standard methods, such as obtaining V_t from the historical sample covariance matrix, may fail to satisfy this condition. This is particularly the case in our example with 36 monthly past returns and a large investment universe with at least 400 and at most over 2000 stocks. See Ledoit and Wolf (2004) and DeMiguel et al. (2009) for details on this invertibility problem. We use the following factor covariance matrix using the Fama and French (1993) equity risk factors which ensures an invertible covariance matrix

$$V_t = \Gamma_t' V_{\mathrm{FF},t} \Gamma_t + \Delta_t \tag{2.5}$$

where Γ_t is the $3 \times N_t$ matrix with the three estimated equity risk factor loadings for each of the N_t stocks in the investment universe for that investment month t, $V_{\mathrm{FF},t}$ is the sample covariance matrix of the equity risk factors and Δ_t is an $N_t \times N_t$ diagonal matrix with the residual variances of the stocks. These residual variances are the variances of the residuals from regressing the stock returns on the equity risk factors. We use the past 36 months for all estimations.

The covariance matrix V_t does not reflect the momentum phenomenon, but momentum will be reflected in the remaining input parameter: the expected returns f_t. In the next section we discuss in which ways these expected returns reflect momentum and its top-bottom phenomenon.

2.3.2 Optimization strategies using the top-bottom phenomenon

In this section we present two mean-variance optimization procedures which differ in the way they take into account momentum's top-bottom phenomenon.

For convenience, we start to describe a straightforward way to apply mean-variance optimization to equity momentum. In the previous section we already discussed the problem set-up including the covariance matrix. It would be straightforward to make use of the past six month average returns after a skip month for ranking the stocks in the momentum strategy as expected returns in the mean-variance optimization:

$$f_{it} = \sum_{t^*=t-7}^{t-2} r_{it^*} \forall i \in I_t. \tag{2.6}$$

We will compare our results with this pure mean-variance optimization strategy, denoted by 'Full-data'.

We take the top-bottom phenomenon into account in two ways: by filtering the expected returns or by reducing the investment universe to only the stocks in the top and bottom decile.

The first method we propose, filtering returns, implies zero expected returns for stocks that are ranked in the intermediate deciles, defines a value of 1 for stocks at the top decile, and a value of -1 for stocks at the bottom decile. Intuitively, this approach smooths extreme returns, and the portfolio constructed with filtered returns does not rely on the returns of the stocks in the intermediate deciles. We denote this optimization strategy as '101' optimization and present details of the filtering procedure below.

The second method, reducing the investment universe, ensures that the optimization algorithm does not select stocks from intermediate deciles. Discarding stocks from these intermediate deciles is expected to improve the realized return of the strategy as Figure 2.2 already showed that long and short positions in these stocks in the intermediate deciles have a zero net return. We note that a disadvantage of reducing the investment universe according to this strategy is the reduction in diversification benefits. The stocks in the top and bottom deciles keep their past six month average returns after a skip month as expected returns. We denote this optimization strategy as 'Top-Bottom-Data' optimization.

The different optimized portfolio strategies can be summarized as follows:

- *Full-data*: This strategy uses mean-variance optimization to allocate all stocks in a zero net investment long-short portfolio using the momentum signals used for ranking the stocks in equation (2.6) as expected returns.

- *Full-101*: The portfolio is optimized over all stocks, but past expected returns are filtered to reflect momentum's top-bottom phenomenon. This optimization results in a minimum variance combination of all stocks. The expected returns of the stocks in the top decile equal one, that of stocks in the bottom decile equal minus one and that of the stocks in the intermediate deciles equal zero. This is also formulated in equation (2.7). The rationale for using expected returns equal to plus and minus one is that this results in minimum variance combinations of these stocks within the top and bottom deciles. This reduces the "error maximization problem" in wrongfully estimated expected returns in mean-variance optimization by have equal expected returns for the stocks in the top and bottom deciles. We use a zero expected return for the stocks in the intermediate deciles to reflect momentum's top-bottom phenomenon that these stocks do not out- or underperform each other; allocating to these stocks would not be beneficial from a return perspective. However, due to the weighted average of returns in equation (2.2), optimized weights of returns within the same decile can differ according to the correlation between stocks.

- *Top-Bottom* data: We reduce the investment universe to only the stocks in the top and bottom decile of momentum's ranking and use these stocks' recent returns, their momentum signals according to Equation 2.6, as expected returns. The portfolio is constructed using mean-variance optimization on the reduced investment universe.

Filtering returns according to momentum's top-bottom phenomenon The methodology to filter past returns is as follows. We base the expected returns f_t of the stocks in a particular month in the investment universe I_t on the standard momentum signals, the six month average realized returns, lagged by one month to deal with short-term reversals as in equation (2.6). We then test if taking into account momentum's top-bottom phenomenon is able to boost the performance. We change the expected returns to plus one for the top decile, minus one for the bottom decile and zero for the intermediate deciles.

The filtered or adjusted expected returns f_t^* become

$$f_{it}^* = \begin{cases} 1 & f_{it} \geq F_t^{-1}(0.1) \\ 0 & F_t^{-1}(0.1) < f_{it} \leq F_t^{-1}(0.9) \\ -1 & f_{it} \leq F_t^{-1}(0.9) \end{cases} \tag{2.7}$$

where F_t is the sample cumulative distribution function of the momentum signals used for investing in month t.

We emphasize that the covariance matrix is not based on filtered returns even when past returns are adjusted according to equations (2.6) and (2.7). The obtained variances, calculated from the Fama and French (1993) equity risk factors, are by construction less volatile than returns. Given a single extreme observation, variances will not change drastically, hence there is not a clear need to 'filter' variances. Strategy *Full-101* benefits from smoothing by filtering extreme returns to -1 and 1, and filtering intermediate returns to 0. Still, the portfolio benefits from diversification since the weight of each stock is determined by filtered returns but non-filtered variances in equation (2.2).

An illustrative example As an illustration, consider the investment universe of 20 stocks. The standard momentum strategy in Section 2.2.1 and the alternative momentum strategies proposed above are as follows

- Strategy *Standard* momentum, presented in Section 2.2.1, takes the 2 best performing stocks and the 2 worst performing stocks according to the average returns in the last 7 months. The return of the strategy is the equal weighted return of these 4 stocks.

- Strategy *Full-data* takes 20 stocks as investment universe. The choice of stocks and the weights of stocks are determined using the mean-variance optimization and the average past returns. The return of this strategy is a mean-variance optimized weighted returns of all 20 stocks. No filtering is applied on returns.

- Strategy *Full-101* takes 20 stocks as investment universe. Past performance of the stocks are evaluated according to 'filtered' returns taking values $-1, 0, 1$, corresponding to bottom, intermediate and top deciles, respectively. The optimized portfolio of this strategy is a mean-variance optimized weighted holding in all 20 stocks. Even when stocks have equal filtered returns, e.g. for stocks with 0 filtered returns in deciles 2 and 3, weights of stocks can be different due to the correlation structure. Hence this strategy takes into account momentum's top-bottom phenomenon, but it does not simply reduce the investment universe to only four stocks.

- Strategy *Top-Bottom* data takes 4 stocks from the top and bottom deciles as investment universe. The choice of stocks and the weights of stocks are determined using the mean-variance optimization and the average past returns. The return of this strategy is a mean-variance optimized weighted returns of 4 stocks. No filtering is applied on returns.

2.3.3 Empirical results

We apply the proposed momentum strategies and the standard momentum strategy to the data described in Section 2.2.1. Table 2.2 presents several results obtained from different portfolio strategies including the annualized mean, volatility, Sharpe-ratio and the crash risk characteristics. For all strategies, the investment universe consists of stocks with at least 36 months from July 1929 until December 2014.

The first three rows of Table 2.2 indicate the standard mean-variance trade-off in portfolio strategies. The standard strategy leads to a relatively high annualized mean, associated with the highest volatility. Strategy *Full-101* has a slightly lower annualized mean, but a considerably lower volatility. The Sharpe-ratio takes account of this mean-variance trade-off. The proposed two strategies, with results listed in the last two columns, lead to a higher Sharpe-ratio than the standard strategy. Hence the return results are improved by using *Full-101* and *Top-Bottom* data, which take into account the top-bottom phenomenon by filtering the expected returns (*Full-101*) or by reducing the investment universe to only the stocks in the top and bottom decile (*Top-Bottom* data).

We next compare strategies *Full-101* and *Top-Bottom* data with the pure mean-variance optimization strategy *Full-data*. The former strategies take momentum's top-bottom phenomenon into account when applying mean-variance optimization. The strategy *Full-data*, on the other hand, is only based on mean-variance optimization. Table 2.2 shows that the proposed strategies *Full-101* and *Top-Bottom* data perform better than the pure

Table 2.2: **Return and crash risk characteristics of standard and optimized momentum.** The optimized momentum strategies differ in two aspects. The *Full* strategies optimize over all available stocks in the investment universe, while the *Top-Bottom* (T-B) data strategy optimizes over just the stocks in the top and bottom decile of momentum's ranking. The *data* strategies use the unfiltered momentum signals as expected returns, while the *101* strategy adjusts the expected returns to plus one for the stocks in the top decile of momentum's ranking, minus one for the stocks in the bottom decile and zero for the remaining stocks. The return and crash risk characteristics cover the period from July 1929 to December 2014.

	Std. Mom.	Full Data	Full 101	T-B Data
Annualized Mean	6.5%	4.2%	6.1%	9.9%
Annualized Volatility	22.2%	10.2%	9.7%	17.1%
Annualized Sharpe-Ratio	29.3%	41.3%	63.6%	57.9%
Largest Monthly Loss	-57.3%	-23.7%	-23.8%	-40.6%
1% Expected Shortfall	-38.1%	-17.7%	-15.1%	-26.2%
Maximum Drawdown	323.6%	73.5%	44.3%	132.9%
Max. Recovery Period	430	202	197	239

mean-variance optimization strategy in terms of mean and Sharpe-ratio. Hence, both taking the top-bottom phenomenon of momentum and mean-variance optimization improves portfolio performance. Taking momentum's top-bottom phenomenon into account is beneficial since in this way the allocation to stocks in deciles 2 to 9 decreases. The *Full-data* optimizes under the assumption that stocks in decile 2 will outperform all other stocks, but the stocks in decile 1. Figure 2.2 has, however, shown that the stocks in decile 2 only outperform stocks in decile 10. As a consequence the *Full-Data* has long-short positions in stocks in deciles 2 to 9 that are assumed to be beneficial, but in practice are not. Both removing the stocks in the intermediate deciles in the *Top-Bottom-Data* strategy and adjusting the expected returns of these stocks in these deciles in the *Full-101* strategy is beneficial. The Sharpe ratios of these latter strategies do not differ that much. The *Full-101* strategy has lower risk because of its diversification benefits from optimizing over the full investment universe rather than only the stocks in the top and bottom deciles and because of forming minimum variance combinations of the stocks in the top and bottom deciles. The higher realized return of the *Top-Bottom* data strategy stems from the fact that it allocates 100 percent to the top and bottom deciles and does not 'suffer' from allocating to stocks in the intermediate deciles that do not outperform each other.

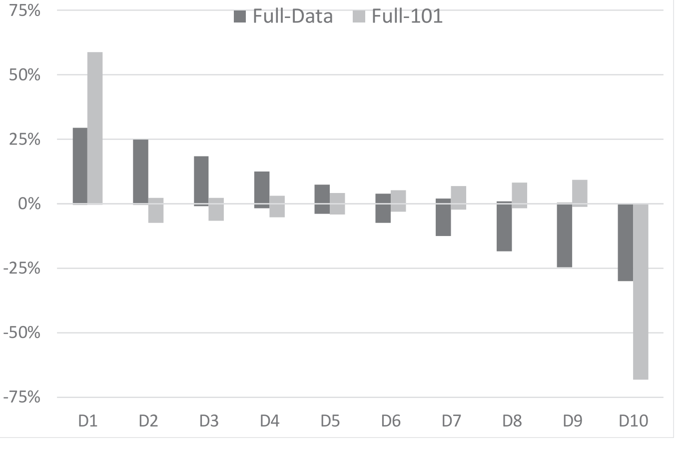
In terms of crash risk characteristics, we conclude that optimization improves the results of the standard momentum strategy substantially, as shown in the bottom panel of Table 2.2. Within the optimized strategies, *Full-101* performs best.

In Figure 2.3, we present the average monthly optimized weights for each decile for two strategies, *Full-data* and *Full-101*, in order to clarify the effect of filtered returns in mean-variance optimization. The *Full-101* strategy is on average long 58.8 percent in the top decile and short 68.2 percent in the bottom decile in comparison to 29.4 percent long in the top decile and 30.0 percent short in the bottom decile for the *Full-Data* strategy with the unadjusted momentum signals as expected returns. As such the *Full-101* strategy implicitly takes into account the top-bottom phenomenon of momentum. The higher Sharpe ratio for *Full-101* compared to *Full-data* in Table 2.2 is explained with this exposure.

An interesting feature of these optimized weights is the slight increase in weights from deciles 2 to 9 for strategy *Full-101*. Adjusted returns for all stocks in deciles 2 to 9 are 0, but the covariance matrix V_t leads to different weights for each stock and thus each decile. Therefore the optimized portfolio with filtered returns benefits from diversification. For example, slightly low weights of stocks in decile 2 potentially result from the positive correlation of these stocks with those in decile 1; mean-variance optimization typically results in offsetting positions in positively correlated stocks with different expected returns.

The increased allocation to the top and bottom deciles of the *Full-101* strategy does not come at the cost of higher risk; the *Full-101* strategy has actually lower crash risk properties than the strategy based on momentum signals as expected returns. We ascribe this to the optimization of the stocks within the top decile: mean-variance optimization loads on the input expected returns and within the top decile the *Full-Data* strategy with unadjusted expected returns heavily loads on the few stocks with extremely high and low returns. These stocks are likely the most risky ones within the top and bottom deciles. The *Full-101* strategy rather selects the less risky stocks within the top and bottom deciles. In general, having equal expected returns, for example by adjusting the expected returns to one, results in the minimum variance portfolio. In our case, adjusting the expected returns to plus one and minus one results in minimum variance combinations

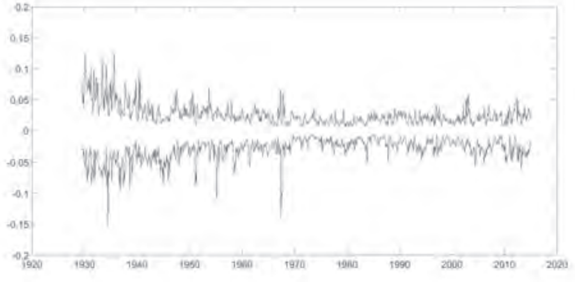
Figure 2.3: **Weights over deciles for optimized momentum.** Average monthly weights when the investment universe contains all stocks for the strategy that use the standard momentum signals as expected returns (Full-Data) and the strategy that uses filtered returns for momentum ranking (Full-101).



of the stocks within the top and bottom deciles[4]. The outperformance of this *Full-101* is in this way in line with the findings of Clarke, de Silva, and Thorley (2006) who also find that the minimum variance portfolio outperforms the market value weighted portfolio.

Mean-variance optimization is known to often result in extreme weights. We deem the minimum and and maximum weights of the *Full-101* equity momentum strategy in Figure 2.4 not to be extreme. We achieve these non-extreme weights by our problem set up that restricts all weights to be between minus and plus 100 percent. This is comparable to the solution Jagannathan and Ma (2003) come up with to prevent mean-variance optimized portfolios to have extreme weights: by adding a no short-selling restriction they ensure that all weights are between 0 and 100 percent. Without this restriction standard fully-invested mean-variance optimized portfolios typically have long positions in part of their stocks and short positions in stocks in the other part; for example all long holdings add up to 500 percent and all short holdings add up to 400 percent.

Figure 2.4: Minimum and maximum weights of *Full-101* optimized momentum strategy over the period from July 1929 to December 2014.



Based on the presented results, we conclude that momentum improves by mean-variance optimization and by accounting for the top-bottom phenomenon of momentum, either by reducing the investment universe or by filtering past returns. According to the Sharpe ratios and crash risk characteristics, strategy *Full-101* is the best strategy from the compared alternatives so far.

[4]Of course, the adjustment of the expected returns to plus and minus one also results in long-short positions in highly correlated stocks with high versus low risk.

2.4 Alternative momentum strategies: residual and constant-risk

In this section we consider two recent alternative momentum strategies and compare their performance with the proposed optimized momentum strategy, the standard momentum strategy, as well as with each other. First, we consider residual momentum proposed by Blitz et al. (2011), where stocks are ranked on the recent residual returns from regressing the stocks' total returns on the Fama and French (1993) equity risk factors. Second, we consider the constant-risk momentum strategy put forward by Barroso and Santa-Clara (2015) where the momentum strategy is leveraged to a strategy with constant risk over time.

In the following tables and figures, we make use of the symbol 'S' for the standard momentum strategy; the symbol 'O' for the *Full-101* optimization strategy; the symbol 'R' for residual momentum and the symbol 'C' for constant-risk momentum.

Residual momentum (R) Residual momentum, proposed by Blitz et al. (2011), ranks stocks based on their recent residual returns from regressing the stock returns on the Fama and French (1993) equity risk factors rather than ranking stocks on their recent total returns. Blitz et al. (2011) find that a strategy that is long in the top decile and short in the bottom decile according to the ranking of these residual returns leads to a higher Sharpe ratio compared to the standard momentum strategy.

The residual momentum strategy we apply first estimates the stocks' equity risk factor exposures over the months $t-T$ to $t-1$ where T equals 36 months. The average residual returns from these regressions over $t-7$ to $t-2$ are then used for residual momentum's ranking. The average residuals returns used for ranking do not include the most recent month $t-1$ for the same argument as for standard momentum: in order to overcome short-term reversals in residual returns. The residual momentum strategy invests in month t in the stocks in the top decile of the ranking on residual returns and short-sells the stocks in the bottom decile of this ranking.

Unlike Blitz et al. (2011), we do not rank the stocks on the risk-adjusted residual returns by normalizing recent average residual returns with the corresponding residual volatilities, i.e. the standard errors of the regressions. We choose to do so to remain closer to the standard momentum strategy in which the total returns are also not divided by the total volatility of the stocks. Dividing the residual, or total, returns used for momentum's ranking would add exposure to the low volatility effect as documented by Blitz and Vliet (2007). Nevertheless, we note that our results are similar when we use average residual returns or these normalized residual returns.

Column 3 in Table 2.3 confirms the results of Blitz et al. (2011) that residual momentum outperforms standard momentum. Residual momentum's Sharpe ratio is 77.9 percent in comparison to 29.3 percent for standard momentum. Residual momentum also exhibits less crash risk than standard momentum. Its largest monthly loss is 40.1 percent versus 57.3 for standard momentum and the maximum drawdown is 63.0 percent in comparison to 323.6 percent for standard momentum. Moreover its maximum recovery period to recover from a drawdown to the level of the previous peak is 84 months, 7 years, while this is more than 430 months, 35 years, for the standard momentum strategy.

Constant-risk momentum (C) Barroso and Santa-Clara (2015) adjust the momentum strategy in a different way. They find that leveraging the momentum strategy to a strategy with constant risk over time outperforms standard momentum. This constant-risk momentum strategy reduces the exposure of standard momentum when momentum's volatility has been high during the recent months. At times when momentum's recent volatility has been low, the strategy increases its exposure. At times of low volatility, the total holdings in both the top and bottom decile add up to more than one, while these holdings add up to less than one when the recent momentum returns have been volatile. However, the constant-risk strategy still attaches equal weights to the stocks in the top and bottom deciles. The holdings in the top and bottom decile of the standard momentum strategy add up to one every month, irrespective of momentum's recent volatility.

The constant-risk momentum strategy we apply estimates the volatility of the standard momentum strategy over the past 6 months and then adjust the total, although still equally weighted, holdings in the top and bottom deciles of momentum's ranking. In this way we differ from Barroso and Santa-Clara (2015) who use daily data to estimate momentum's volatility to adjust the leverage. Nevertheless, we do find that our variant of constant-risk momentum also outperforms standard momentum. The total holdings in the top and bottom are adjusted such that the momentum's recent volatility times these total holdings are constant over time. For a fair comparison with the other momentum strategies that have total holdings in the top and bottom decile equal to one, we multiply the returns of the constant-risk momentum strategy afterwards, such that the constant-risk strategy also has on average total holdings in the top and bottom deciles equal to one.

Results All proposed alternative momentum strategies in the columns 2 to 4 of Table 2.3 outperform standard momentum, S, in terms of their Sharpe ratios which are higher than standard momentum's Sharpe ratio of 29.3 percent. Furthermore, their largest monthly loss is lower, their maximum drawdown is much lower than standard momentum's 323.6 percent and their maximum recovery period from a drawdown is well below the 430 months for standard momentum. In terms of Sharpe ratio, the residual momentum strategy performs best with a Sharpe ratio of 77.9 percent and this strategy does not take more than 84 months (7 years) to recover to the previous peak after a drawdown. The second highest Sharpe ratio is obtained by the constant-risk strategy. In terms of the risk indicators largest monthly loss and maximum drawdown, however, the proposed optimized strategy outperforms the constant-risk and residual momentum strategies.

We conclude that the proposed optimized and alternative strategies outperform standard momentum strategy in all risk and return characteristics, except that strategy O has a slightly lower annualized mean (6.1 percent) than strategy S (6.5 percent). However, a second conclusion is that none of the proposed three strategies outperforms the other two strategies on all risk and return characteristics. In terms of a higher Sharpe ratio residual momentum, R, performs best, while in terms of crash risk, the optimization strategy, O, outperforms the other improvements mostly.

Table 2.3: **Return and crash risk characteristics of four different momentum strategies:** Standard momentum (S), mean-variance optimization of momentum using filtered data, earlier denoted by Full-101 and now as (O), residual momentum (R) and constant-risk momentum (C). The results cover the period July 1929 to December 2014.

	S	O	R	C
Annualized Mean	6.5%	6.1%	10.3%	11.8%
Annualized Volatility	22.2%	9.7%	13.2%	18.0%
Annualized Sharpe-Ratio	29.3%	63.6%	77.9%	65.4%
Largest Monthly Loss	-57.3%	-23.8%	-40.1%	-35.0%
1% Expected Shortfall	-38.1%	-15.1%	-22.7%	-18.0%
Maximum Drawdown	323.6%	44.3%	63%	93.2%
Max. Recovery Period	430	197	84	198

2.5 Combining momentum strategies

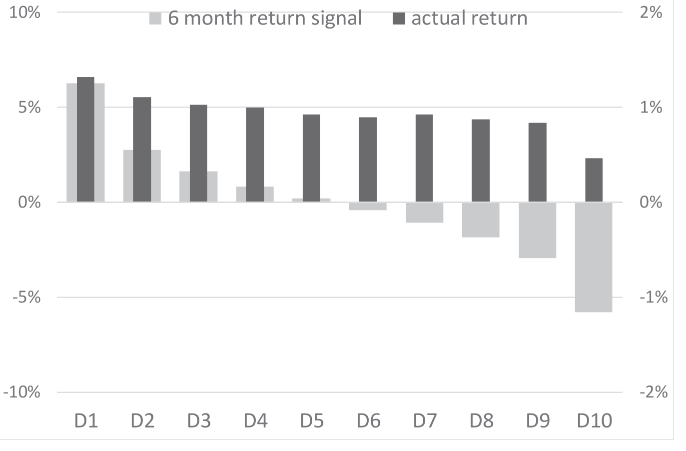
In the previous section we compared the three momentum strategies and found some mixed results on the risk and return characteristics of the optimized (with filtered data), residual and constant-risk momentum strategies. Optimization is better according to largest monthly loss and maximum drawdown, while residual and constant-risk momentum are better in terms of Sharpe ratio. Therefore, it is natural to continue the analysis using several combination strategies: pairwise as well as a joint strategy of using optimized, residual and constant-risk momentum. In the following tables and figures, we make use of an asterisk '*' to denote optimization strategies that use the *Top-Bottom* data instead of the *Full-101* method.

2.5.1 Combination strategies

Residual Optimized Momentum: R-O In order to investigate whether optimization improves residual momentum, we first investigate whether residual momentum also exhibits a top-bottom phenomenon. Figure 2.5 presents the average 6 month return signal for residual momentum and the actual realized return in the investment month per decile. Average returns obtained from residual momentum show a different pattern compared to realized returns. Actual realized returns exhibit a different decreasing pattern between deciles 2 to 9 than the residual momentum signals that by construction show a steady decrease over these deciles. Like standard momentum, residual momentum seems to be most pronounced in the top and bottom deciles of the ranking. This makes optimization strategies that take account of this phenomenon also appropriate for optimizing residual momentum. Similarly as for standard momentum we choose to apply the *Full-101* strategy to optimize residual momentum.

In Table 2.4 risk and return characteristics are presented when we apply mean-variance optimization to residual momentum. This combined strategy is denoted by 'R-O'. These results confirm that optimization does not only work for standard momentum, but also for residual momentum. Although the improvement in the Sharpe ratio and the reduction in the crash risk characteristics are smaller than when optimizing the standard momentum strategy. The results for the *Top-Bottom* data R-O strategy, not reported, are comparable to those of the *Full-101* R-O strategy. The *Full-data* R-O strategy, also not reported here, performs less than these strategies that take the top-bottom phenomenon into account. This is in line with the results when optimizing standard momentum. More details are

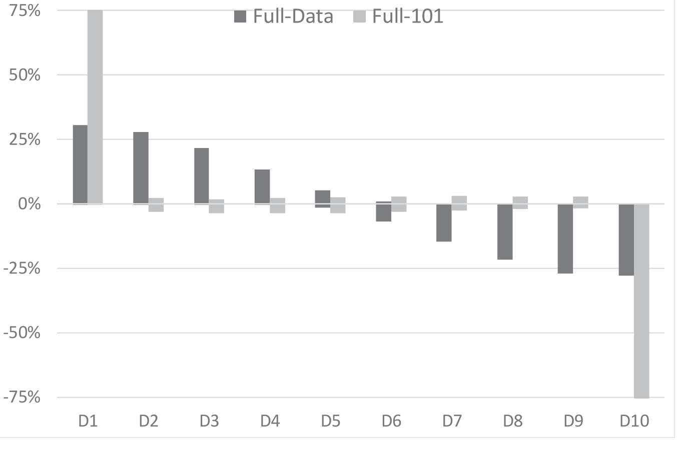
Figure 2.5: **Actual versus lagged residual returns.** Average residual return signals over the six month ranking period (left axis) and actual realized returns (right axis) for each decile from July 1929 to December 2014.



discussed below.

Just like in Section 2.3, we consider in Figure 2.6 the average monthly weights of the optimized strategies, but now for the residual momentum strategies. The pattern of the weights for the optimized residual momentum strategies is similar to the pattern of the weights of the optimized standard momentum strategies in Figure 2.3. The optimized momentum *Full-101* strategy using filtered returns, shown in Figure 2.3, allocates approximately 60 percent of the portfolio to the top and bottom deciles, while residual momentum with filtered returns, shown in Figure 2.6, allocates almost 80 percent of the portfolio to the top and bottom deciles. A possible explanation for this finding is the existence of more pronounced diversification benefits for residual momentum within the top and bottom deciles. The top and bottom deciles of residual momentum are always populated by stocks with both high and low factor loadings, irrespective of the factor returns during momentum's ranking period. The top and bottom deciles of standard momentum are populated by just stocks with high or low factor loadings if the factor returns have been highly positive or strongly negative during momentum's ranking period. This follows from the fact that if a factor return has been highly positive during the ranking period the stocks with large loadings on this factor tend to end up in the top decile of standard momentum as they profit from being largely exposed to a factor with a positive return. Residual momentum ranks stocks on their residual returns and as such stocks with high and low factor loadings are equally likely to end up in the top and bottom deciles, irrespective of the returns on these factors during momentum's ranking period. This phenomenon also relates to the time-varying exposures of the equity momentum strategy that we discuss in Section 2.7.

Figure 2.6: **Weights over deciles for optimized residual momentum.** Average monthly weights when the investment universe contains all stocks for the strategy that use the residual momentum signals as expected returns (Full-Data) and the strategy that uses filtered returns for momentum ranking (Full-101).



The allocation to the deciles in Figure 2.6 for residual momentum also exhibits the pattern over the deciles 2 to 9 as the allocation of optimized standard momentum in Figure 2.3, but to a lesser extent. This is also a consequence of ranking stocks on their residual return rather than on their total return. When ranking stocks on their total returns stocks with high or low factor loadings tend to end up in the top or bottom deciles when the factor return during momentum's ranking period has been highly positive. The average factor loading of stocks in a particular decile in this situation declines from decile 1 to 10. As a consequence stocks in neighboring deciles tend to be more correlated when the factor return has been highly positive or strongly negative during momentum's ranking period. Mean-variance optimization results in offsetting positions in stocks that are highly correlated and that is why we find the increasing pattern in allocation to the stocks in deciles 2 to 9 for standard momentum. Since residual momentum ranks stocks on their residual returns, stocks with high and low factor loadings are equally likely to be in the different deciles and as such neighboring deciles are less correlated. Therefore we find a smaller increasing pattern in the allocation from deciles 2 to 9 than for standard momentum.

Optimized constant-risk momentum: O-C It is possible to apply the mean-variance optimization in a constant-risk setting. Rather than scaling the solution to the subproblem in equation 2.2 to a 1$ long and 1$ short exposure, one may scale the optimal weights such that the long-short portfolio has constant risk over time:

$$\widehat{h}_t = \widehat{h}_t^* \frac{CR}{\sqrt{\widehat{h}_t^{*\prime} V_t \widehat{h}_t^*}} \tag{2.8}$$

where CR is the ex-ante constant-risk the portfolio is scaled to. After constructing all monthly portfolios $\widehat{h}_t$ we scale these weights again such that on average the optimized constant-risk momentum strategy has unit long and unit short exposure. We do this for a fair comparison of the different strategies. Note that this scaling does not affect the Sharpe ratio or the time to recover from a large drawdown, but it does affect the crash risk figures largest loss and maximum drawdown.

Optimized constant-risk momentum at lower costs: O-C* Since mean-variance optimization typically increases the number of transactions we also apply the constant-risk momentum strategy to just the stocks in the top and bottom deciles of momentum's ranking. Given that this *optimized constant-risk at lower costs momentum* strategy is only long and short in the stocks in the top and bottom decile its number of transactions will be lower than when optimizing over all available stocks. It is thus similar to the optimized constant-risk momentum O-C strategy except that the strategy buys and short-sells less stocks every investment month. The strategy is in that sense also similar to the *Top-Bottom* data strategy in Section 2.3; it only differs from this strategy by keeping its ex-ante risk constant over time.

Residual momentum with constant risk: R-C The constant-risk residual momentum strategy is similar to the constant-risk momentum (C) strategy, but the risk adjustment to achieve a constant-risk over time is now applied to the residual R momentum strategy. The realized volatility over the past 6 months of residual momentum is used to leverage and deleverage to a constant risk over time.

Residual momentum optimized with constant risk: R-O-C The strategy with all three improvements combined is the optimized constant-risk residual momentum strategy, denoted by 'R-O-C'. It is similar to the optimized constant-risk momentum O-C strategy but with the adjusted residual returns used as expected returns in the mean-variance optimization.

Residual momentum optimized with constant risk at lower costs: R-O-C* Similarly as for standard momentum we also test the performance of the optimization of constant-risk residual momentum over just the stocks in top and bottom deciles of residual momentum's ranking. Given that this *optimized constant-risk residual momentum at lower costs* strategy is only long and short in the stocks in the top and bottom decile its number of transactions will be lower than when optimizing over all available stocks.

2.5.2 Summary of momentum strategies

In the next sections we compare all ten momentum strategies that we have discussed so far. First, we briefly summarize all ten momentum strategies:

- **S:** The **standard momentum** strategy is long in the stocks ranked in the top decile and short in the stocks in the bottom decile. The stocks are ranked on their total return over the previous six months with a delay of one month to account for short-term reversals.

- **O:** The **optimized momentum** strategy is the *Full-101* strategy and uses mean-variance optimization to construct portfolios that are 1$ long and 1$ short. The expected returns are plus one for the stocks ranked in the top decile, minus one for the stocks ranked in the bottom decile and zero for the remaining stocks.

- **C:** The **constant-risk momentum** strategy leverages and deleverages the standard momentum S strategy such that its ex-ante risk is constant over time. The ex-ante risk is determined as the realized volatility of standard momentum S over the past six months. Afterwards the strategy is scaled such that its exposure is on average 1$ long and 1$ short.

- **O-C:** The **optimized constant-risk momentum** strategy is similar to the standard optimized O strategy, but constructs portfolios that have a constant ex-ante volatility over time rather than portfolios that are 1$ long and 1$ short. Afterwards the strategy is scaled such that its exposure is on average 1$ long and 1$ short.

- **O-C*:** The **optimized constant-risk momentum at lower costs** strategy is similar to the optimized constant-risk momentum O-C strategy, but does not optimize over all available stocks with expected returns adjusted to plus one, minus one and zero. This strategy is rather the *Top-Bottom* data optimization strategy that applies mean-variance optimization to only the stocks in the top and bottom deciles with the total returns over the previous six months as expected returns. The strategy keeps the ex-ante risk constant over time; afterwards the strategy is scaled such that its exposure is on average 1$ long and 1$ short.

- **R:** The **residual momentum** strategy is long in the stocks ranked in the top decile and short in the stocks in the bottom decile. The stocks are ranked on their residual return over the previous six months with a delay of one month to account for short-term reversals. The residual returns stem from regressing the three equity risk factors market, size and value on the past 36 months.

- **R-O:** The **residual optimized momentum** strategy is the *Full-101* residual momentum strategy and uses mean-variance optimization to construct portfolios that are 1$ long and 1$ short. The expected returns are plus one for the stocks ranked on their residual returns in the top decile, minus one for the stocks ranked in the bottom decile and zero for the remaining stocks.

- **R-C:** The **residual momentum with constant-risk** strategy leverages and deleverages the residual momentum R strategy such that its ex-ante risk is constant over time. The ex-ante risk is determined as the realized volatility of residual momentum R over the past six months. Afterwards the strategy is scaled such that its exposure is on average 1$ long and 1$ short.

- **R-O-C: residual optimized constant-risk momentum** strategy is similar to the optimized constant-risk momentum O-C strategy, but uses the ranking on the residual returns instead of the ranking on total returns to determine the expected returns of plus one, minus one and zero used for the mean-variance optimization.

- **R-O-C*:** The **residual optimized constant-risk momentum at lower costs** strategy is similar to the optimized constant-risk momentum O-C* strategy, but uses the residual returns instead of the total returns for the expected returns used in the mean-variance optimization.

2.5.3 Results

Table 2.4 shows the results for return and crash characteristics of the standard strategy, the proposed individual and combined strategies. As stated above, the proposed individual momentum strategies O, C, and R outperform standard momentum on all criteria, except one case, see the results in columns 2-4 compared to the results in column 1. Applying pairwise combination strategies gives improvements on the Sharpe ratios from R-O and R-C and two of their crash risk characteristics.

Combining all proposed alternatives in the R-O-C strategy improves standard momentum's performance and reduces its crash risk substantially. The Sharpe ratio improves from 29.3 percent to 100.3 percent. In addition, crash risk indicators like the largest monthly loss, the maximum drawdown and the time to recover to the level of the previous peak after a drawdown improve from, respectively, 57.3 percent, 323.6 percent and 430 months to 13.8 percent, 35.8 percent and 104 months. The R-O-C* strategy has a Sharpe ratio of 106.6% that is even higher than that of the R-O-C strategy, but does exhibit slightly more crash risk as shown by two criteria. Its higher volatility and the larger crash risk are due to the fact that this strategy only optimizes over the stocks in the top and bottom deciles and thus exhibits less diversification than the R-O-C strategy that optimizes over all available stocks. On the other hand the R-O-C* strategy will likely profit from imposing less transactions than the R-O-C strategy. We will look into transaction costs in the next section. Both the R-O-C and R-O-C* strategies also outperform the alternative momentum strategies, O, R and C in terms of Sharpe ratios and most crash risk characteristics.

2.5.4 Statistical significance of Sharpe ratios

In this section we test whether the Sharpe ratios of the ten different strategies are significantly larger than zero and whether a Sharpe ratio of a specific strategy is significantly higher than that of the other strategies. We do so by using the bootstrap method presented in Appendix 2.C. Table 2.5 presents the p-values of the bootstrap test whether the Sharpe ratio of a strategy is larger than zero. It is directly seen that the Sharpe ratios of all momentum strategies have p-values equal to zero except for standard momentum with a p-value of 1.04%. This indicates that all Sharpe ratios are significantly larger than zero at the 5% confidence level and all but one are significant at the 1% confidence level.

Table 2.4: **Return and crash risk characteristics of different momentum strategies and their combinations.** The return and risk characteristics are as follows: Average annualized returns, annualized volatility, annualized Sharpe ratio, the largest monthly loss (Lowest), the 1% expected shortfall (Shortfall), maximum drawdown (Drawdown) and the maximum period to recover from a drawdown to the level of the prior peak (Recovery). See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.

	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Mean	6.5%	6.1%	11.8%	10.3%	6.5%	12.2%	7.9%	10.1%	8.7%	12.4%
Volatility	22.2%	9.7%	18.0%	13.2%	8.4%	15.8%	9.5%	11.8%	8.7%	11.7%
Sharpe	29.3%	63.6%	65.4%	77.9%	76.9%	77.6%	83.5%	86.0%	100.3%	106.6%
Lowest	-57.3%	-23.8%	-35.0%	-40.1%	-17.3%	-34.0%	-24.0%	-24.6%	-13.8%	-18.7%
Shortfall	-38.1%	-15.1%	-22.7%	-18.0%	-10.8%	-20.9%	-12.9%	-14.8%	-9.8%	-12.0%
Drawdown	323.6%	44.3%	93.2%	63.0%	29.9%	62.0%	70.3%	30.2%	35.8%	35.5%
Recovery	430	197	198	84	82	184	109	126	104	70

Table 2.5: **Statistical significance Sharpe ratios.** The p-value of the bootstrap test whether a momentum strategy's Sharpe ratio is larger than zero. See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.

	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Sharpe	29.3%	63.6%	65.4%	77.9%	76.9%	77.6%	83.5%	86.0%	100.3%	106.6%
p-value	1.04%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Table 2.6 shows the p-values of the bootstrap tests whether the Sharpe ratio of a specific strategy is higher than those of the other strategies. The entries in Table 2.6 read as follows. Consider the case of comparing strategy R versus strategy O, then the p-value of the test that the Sharpe ratio of strategy R is higher than that of strategy O equals 92.64%. Obviously, the test whether the Sharpe ratio of R versus O is lower is then equal to 7.36%. In Table 2.6 this is reflected by the fact that the sum of an entry and the opposite entry along the diagonal add up to one.

Column 1 of Table 2.6 shows that the Sharpe ratios of all alternative and combined momentum strategies are significantly higher than the Sharpe ratio of standard momentum. Furthermore, it turns out that the highest Sharpe ratio of momentum strategy R-O-C* is significantly higher than the Sharpe ratios of the other strategies at the 5% significance level; it is also significantly higher at the 1% level for all strategies except for the R-O-C strategy; this R-O-C strategy is similar to the R-O-C* strategy, but uses all

Table 2.6: **Statistical significance Sharpe ratio differences.** The p-values of the bootstrap test whether the Sharpe ratio of the momentum strategy in the particular row is higher than the Sharpe ratio of the strategy in that particular column. See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.

	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
S		100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
O	0.00%		59.51%	92.64%	99.99%	99.50%	99.37%	98.56%	100.00%	100.00%
C	0.00%	40.49%		85.70%	90.15%	93.25%	95.68%	97.61%	99.94%	100.00%
R	0.00%	7.36%	14.30%		47.04%	49.61%	77.62%	89.72%	99.87%	99.99%
O-C	0.00%	0.01%	9.85%	52.96%		56.92%	76.90%	81.21%	99.82%	100.00%
O-C*	0.00%	0.50%	6.75%	50.39%	43.08%		75.75%	80.61%	99.94%	100.00%
R-O	0.00%	0.63%	4.32%	22.38%	23.10%	24.25%		61.65%	100.00%	100.00%
R-C	0.00%	1.44%	2.39%	10.28%	18.79%	19.39%	38.35%		95.70%	99.69%
R-O-C	0.00%	0.00%	0.06%	0.13%	0.18%	0.06%	0.00%	4.30%		97.06%
R-O-C*	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.31%	2.94%	

available stocks. The Sharpe ratio of this strategy itself is also significantly higher than any other strategy at the 5% significance level.

2.6 Transaction costs

So far, our analysis has not taken account of transaction costs, while mean-variance optimization is known to increase the number of transactions compared to $\frac{1}{N}$ strategies. In our case this feature holds particularly true for the strategies where the optimization is over all available stocks rather than only over the stocks in the top and bottom deciles. The constant-risk strategies are also expected to increase the number of transactions, since the leveraging and deleveraging in these strategies requires buying and selling additional amounts of equities. In this section we investigate the effect of transaction costs. We are interested for which level of transaction costs the Sharpe ratios remain positive and whether the Sharpe ratio of a strategy remains higher than those of other strategies for different levels of transaction costs. We also investigate the break-even transaction costs for which the Sharpe ratio bootstrap test remains significantly higher than zero for the 5% significance level.

Number and size of transactions Table 2.7 presents the number and size of the transactions for the ten momentum strategies. Standard momentum, for example, turns out to require on average 4248 transactions per year and the average size of all transactions per year equals 20.3. This 20.3 is relative to being 1$ long and 1$ short and is the consequence of rebalancing the portfolio to the $\frac{1}{N}$ long and short portfolio 12 times a year.

Table 2.7: **Average number and size of transactions plus break-even transaction costs.** The number of transactions in row *Number* presents the average number of transactions per year for the different momentum categories; row *Size* contains the average size of all transactions per year. The table also presents the break-even transaction costs in row *BE SR>0* up to which the Sharpe ratio would remain positive; row *BE p(5%)* presents the break-even transactions up to which the Sharpe ratio would remain higher than zero at the 5% significance level (see Appendix 2.C for the bootstrap method used). See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.

	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Sharpe	29.3%	63.6%	65.4%	77.9%	76.9%	77.6%	83.5%	86.0%	100.3%	106.6%
Number	4248	15497	4249	4345	15497	4227	15497	4346	15497	4322
Size	20.3	21.2	44.5	21.8	21.5	25.1	23.1	39.2	23.8	25.9
BE SR>0	0.32%	0.29%	0.26%	0.47%	0.30%	0.49%	0.34%	0.26%	0.36%	0.48%
BE p(5%)	0.10%	0.21%	0.19%	0.36%	0.23%	0.37%	0.26%	0.20%	0.29%	0.39%

Table 2.7 also shows a distinction between the strategies that only use stocks in the top and bottom deciles and the strategies that optimize over all available stocks. The number of transactions for the optimization strategies over all available stocks have on average 15497 transactions per year, while the strategies that only use the stocks in the top and bottom deciles have on average somewhere around 4300 transactions per year. Although both strategies based on residual and standard momentum are invested in the same number of stocks every month, their numbers of transactions slightly differ. This is due to the fact that the number of stocks entering and leaving the top and bottom deciles differs for these strategies. Given that the strategies based on residual momentum require more transactions, residual momentum strategies' turnover of this type is thus higher.

The size of the transactions for the different strategies is typically between 20$ and 25$ per year, relative to being 1$ long and 1$ short. However, the constant-risk strategies do have much larger transactions with an average total transaction size of 39.2 for residual constant-risk momentum and 44.5 for standard constant-risk momentum. Relative to the transactions for standard and residual momentum the constant-risk strategies typically have larger transactions due to their leveraging and deleveraging in order to keep their risk constant. The constant-risk optimized strategies do not exhibit these larger transactions; this is probably due to the fact that the risk of the optimized strategies is based on 36 months of return history, while the constant-risk strategies themselves only use 6 month of return history to estimate the volatility used for leveraging. This 6 month realized volatility varies substantially over time and as a consequence these strategies require a substantial amount of leveraging and deleveraging every month.

Statistical significance of Sharpe ratios under transaction costs Table 2.7 shows, first, in the row *BE SR>0* that the transaction costs for which the Sharpe ratio would still be positive ranges from 0.26% to 0.49%. As a consequence of their relatively large transactions the Sharpe ratio of the constant-risk momentum strategies only remains positive for transactions costs up to 0.26%. The strategies R, O-C* and R-O-C* would keep a positive Sharpe ratio for transaction costs up to 0.47%, 0.48% and 0.49%, respectively. This is substantially higher than for the other strategies.

The last row of Table 2.7 presents the break-even transaction costs for which the Sharpe ratio of the different strategies remains positive at the 5% level. These break-even costs are obviously lower than those for the Sharpe ratio to remain just positive and range from 0.10% for standard momentum to 0.36%, 0.37% and 0.39% for the strategies R, O-C* and R-O-C*.

Mutual outperformance and transaction costs So far, we have considered the transaction costs in isolation for each momentum strategy and how these costs would affect the Sharpe ratio and the significance of that Sharpe ratio of a specific strategy. We now turn to testing for which transaction costs the different strategies outperform each other. Table 2.8 shows for which ranges of transaction costs a particular strategy outperforms another strategy in terms of Sharpe ratio and for which its own Sharpe ratio is still positive. For example, the entry in column S and row C shows that the Sharpe ratio of the constant-risk momentum strategy C has a higher Sharpe ratio than the standard momentum strategy S as long as transaction costs are between 0.00% and 0.23%. Column C and row S show that standard momentum's Sharpe ratio is higher than that of constant-risk momentum for transaction costs between 0.23% and 0.32%. An empty entry means that the strategy in that row does not have a higher and still positive Sharpe ratio than the strategy in the corresponding column for any, positive, range of transaction costs.

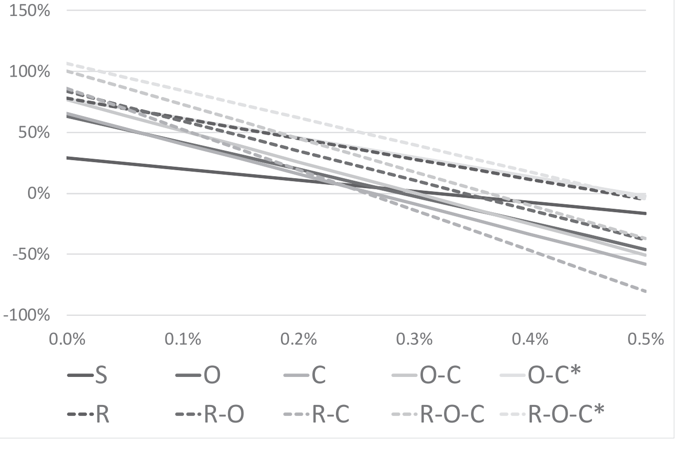
Table 2.8: **Transaction costs and mutual outperformance.** The range of transaction costs for which the strategy in the specific row has a higher, and still positive, Sharpe ratio than the strategy in the specific column. If an entry is empty the strategy in that row does not outperform the strategy in that column for a range of transaction costs for which its Sharper ratio would still be positive. See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.

	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
S		0.27%	0.23%		0.29%			0.24%		
		0.32%	0.32%		0.32%			0.32%		
O	0.00%		0.07%					0.20%		
	0.27%		0.29%					0.29%		
C	0.00%	0.00%						0.24%		
	0.23%	0.07%						0.26%		
R	0.00%	0.00%	0.00%		0.00%	0.00%	0.07%	0.05%	0.20%	
	0.47%	0.47%	0.47%		0.47%	0.06%	0.47%	0.47%	0.47%	
O-C	0.00%	0.00%	0.00%					0.12%		
	0.29%	0.30%	0.30%					0.30%		
O-C*	0.00%	0.00%	0.00%	0.06%	0.00%		0.07%	0.05%	0.20%	0.46%
	0.49%	0.49%	0.49%	0.49%	0.49%		0.49%	0.49%	0.49%	0.49%
R-O	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%		0.03%		
	0.34%	0.34%	0.34%	0.07%	0.34%	0.07%		0.34%		
R-C	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%			
	0.24%	0.20%	0.24%	0.05%	0.12%	0.05%	0.03%			
R-O-C	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%		
	0.36%	0.36%	0.36%	0.20%	0.36%	0.20%	0.36%	0.36%		
R-O-C*	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	
	0.48%	0.48%	0.48%	0.48%	0.48%	0.46%	0.48%	0.48%	0.48%	

The results in Table 2.8 confirm our earlier results: The strategies R, O-C* and R-O-C* outperform the other strategies for a wide range of transaction costs. It is interesting that the constant-risk residual momentum strategy, R-C, also outperforms most other strategies for relatively low transaction costs, but is itself outperformed by any other strategy for relatively larger transaction costs. We ascribe this to its relatively large average size of the transactions due to leveraging and deleveraging to a constant risk over time. Finally, the optimized constant-risk residual momentum at lower costs strategy, R-O-C*, outperforms all other strategies up to the level of its break-even transaction costs of 0.48%, except for the optimized constant-risk momentum at lower costs strategy, O-C*, that has a higher Sharpe ratio for transaction costs between 0.46% and 0.49%.

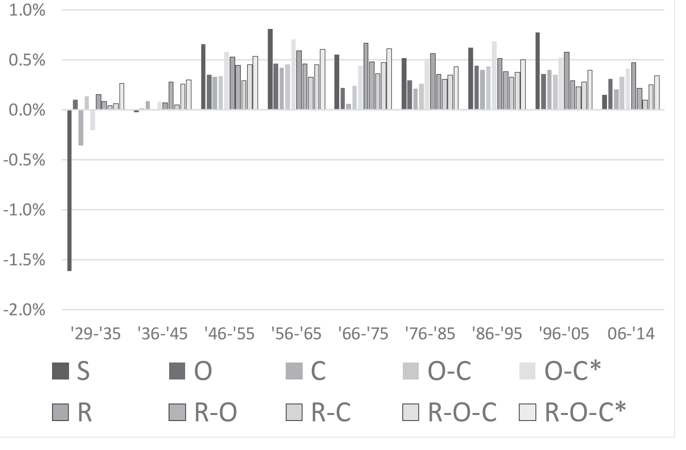
Figure 2.7 shows the Sharpe ratios of the different momentum strategies for different levels of transaction costs. This figure confirms that the three strategies R, O-C* and R-O-C* have the highest Sharpe ratios for most levels of transaction costs. Strategy R-O-C* has the highest Sharpe ratio up to 0.46% and beyond transaction costs of 0.49% the Sharpe ratio of every strategy becomes negative.

Figure 2.7: **Sharpe ratios of momentum strategies for varying transaction costs.** The annual Sharpe ratios of the different strategies over the period from July 1929 to December 2014 for transaction costs ranging from 0.00% to 0.50%. See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.



Break-even transaction costs over time Although transaction costs nowadays might be well below the break-even numbers we have presented in this section, these costs might have been considerably larger in the past. Figure 2.8 shows the break-even transaction costs per decade for the different momentum strategies. For example, the Sharpe ratio of the standard momentum strategy, S, would be larger than zero for transaction costs below 0.78% in the period 1996 to 2005. As a consequence of the relatively poor performance of all strategies until 1945 the break-even transaction costs are also relatively low or negative in this period as holds particularly true for the standard momentum strategy S. Since 1946 the break-even transaction costs are positive for every strategy over all decades. Note that the standard momentum strategy, S, has the highest break-even transaction costs during two of the post-war decades. This means that standard momentum does keep a positive Sharpe ratio for relatively high transaction costs of, say, 0.70%, during these periods, while all the other momentum strategies would then have a negative Sharpe ratio for these transaction costs.

Figure 2.8: **Break-even transaction costs over time.** The break-even transaction costs for different time spans are the transaction costs that would imply a Sharpe ratio of zero during that specific time frame. See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.



Transaction costs[5] are an important aspect of the profitability of the different momentum strategies; it depends on an investor's actual costs if a specific momentum strategy is a, significantly, profitable investment strategy. The analysis over the whole period from July 1929 to December 2014 indicates that most alternative and, in particular, the combined momentum strategies have a higher Sharpe ratio than the standard momentum strategy, also after correcting for transaction costs. However, during a few decades the Sharpe ratio of the standard momentum strategy would be higher than that of any other momentum strategy, but this only holds for relatively high transaction costs.

Our analysis shows that the number and size of the transactions of the combined equity

[5]An investor's transaction costs do not only consist of brokerage fees, but also include the impact of the bid-ask spread and the market impact of actually doing the transactions. Whereas brokerage fees relatively decrease with the size of an investment, the market impact could increase substantially for larger investments.

momentum strategy, R-O-C*, are comparable to those of the standard and residual momentum strategies, S and R, and that this R-O-C* strategy outperforms those strategies with a higher Sharpe ratio for a large range of transaction costs, while it possesses lower crash risk characteristics. Moreover, we find that its Sharpe ratio is also significantly higher than those of the other strategies. Thus, we conclude that this combined strategy R-O-C* is worthwhile for an investor to explore.

2.7 Time-varying risks and returns

The standard equity momentum strategy is known for its time-varying risks and returns. In this section we look into these features and investigate if the optimized momentum strategy, O, the residual and constant-risk momentum strategies, R and C, as well as their combinations are able to reduce momentum's time-varying risks and returns.

First, we look at the strategies' static exposures to the three Fama and French (1993) equity risk factors. Then we test if the strategies also exhibit time-varying exposures to these risk factors as put forward by Grundy and Martin (2001). Finally, we test if, and by how much, the returns on all these momentum strategies vary over the business cycle as Chordia and Shivakumar (2002) find for standard momentum.

Chan, Jegadeesh, and Lakonishok (1996) show that the three Fama and French (1993) equity risk factors do not explain standard momentum's performance. Table 2.9 confirms that momentum's exposures to the equity risk factors do not explain its returns: after correcting for these risk factors a significantly positive alpha of 0.91 percent remains. Nevertheless, standard momentum does negatively load on the three equity risk factors.

Table 2.9 also shows the exposures of the alternative and combined strategies. All alternative and combined strategies are less exposed to the equity risk factors as can be seen from the fact that the monthly volatility explained by the equity risk factors is less than the 2.5 percentage points of the standard momentum strategy. For example, the equity risk factors explain only 0.6 percentage points of the optimized constant-risk residual momentum (R-O-C and R-O-C*) strategies.

Table 2.9: **Regression of momentum returns on three equity risk factors.** The estimated coefficient and t-statistics (in brackets), the standard deviation explained by these factors (Vol. Explained), the total standard deviation (Vol. Total) and the R-squared of the regressions of the returns of the different momentum strategies on the three equity risk factors market, size and value from July 1929 to December 2014. See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.

	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Intercept	0.91%	0.70%	1.18%	0.95%	0.68%	1.25%	0.78%	0.92%	0.81%	1.12%
	(4.90)	(9.15)	(7.48)	(8.10)	(9.64)	(9.40)	(9.40)	(8.68)	(10.47)	(10.68)
Market	-0.22	-0.04	-0.07	-0.11	-0.02	-0.06	-0.09	-0.07	-0.06	-0.07
	(-5.88)	(-2.79)	(-2.37)	(-4.76)	(-1.46)	(-2.38)	(-5.50)	(-3.53)	(-3.65)	(-3.48)
Size	-0.26	-0.24	-0.18	0.01	-0.19	-0.29	-0.05	-0.04	-0.05	-0.05
	(-4.30)	(-9.47)	(-3.44)	(0.21)	(-8.40)	(-6.70)	(-1.86)	(-1.22)	(-2.16)	(-1.43)
Value	-0.43	-0.26	-0.27	-0.08	-0.18	-0.30	-0.12	-0.05	-0.09	-0.07
	(-8.02)	(-11.75)	(-6.02)	(-2.42)	(-9.16)	(-8.00)	(-5.17)	(-1.66)	(-3.88)	(-2.19)
Vol. Explained	2.5%	1.4%	1.4%	0.7%	1.0%	1.7%	0.8%	0.5%	0.6%	0.6%
Vol. Total	6.4%	2.8%	5.2%	3.8%	2.4%	4.5%	2.7%	3.4%	2.5%	3.4%
R-squared	15.1%	23.9%	7.1%	3.7%	16.8%	13.6%	8.4%	2.6%	4.9%	3.0%

2.7.1 Time-varying exposures

Grundy and Martin (2001) find that momentum has time-varying exposures to the Fama and French (1993) equity risk factors. After a large positive return on a specific factor during momentum's ranking period the momentum strategy has a large positive exposure to this factor. This is a consequence of ranking the stocks on their past returns in the momentum strategy. If a factor return has been high during momentum's ranking period, stocks with the highest factor loadings on this factor tend to end up in the top of momentum's ranking as they profit from this high factor return. At the same time stocks with the lowest loadings on this particular factor tend to populate the bottom of the ranking as these stocks profit the least from this high factor return. The momentum strategy then, being long in the stocks with high factor loadings and short in the stocks with low factor loadings, is positively exposed to this particular factor. Equivalently, momentum is negatively exposed to an equity risk factor when the return on that factor has been negative during momentum's ranking period.

Martens and Van Oord (2014) achieve a reduction in momentum's time-varying exposures by directly hedging these exposures. We do not consider hedging in our combined strategies as optimization is supposed to take account of a reduction in the exposures to the equity risk factors as it penalizes these exposures via the factor covariance matrix. Our approach is to determine momentum strategies' time-varying exposures to the equity risk factors by regressing momentum returns on conditional equity risk factors according to the following model:

$$\begin{aligned} r_{\text{Mom }i,t} = \alpha \quad &+ \quad \left(\beta_{\text{down}} D_t^{\text{MKT,down}} + \beta_{\text{flat}} D_t^{\text{MKT,flat}} + \beta_{\text{up}} D_t^{\text{MKT,up}}\right) r_{\text{MKT},t} \\ &+ \quad \left(s_{\text{down}} D_t^{\text{SMB,down}} + s_{\text{flat}} D_t^{\text{SMB,flat}} + s_{\text{up}} D_t^{\text{SMB,up}}\right) r_{\text{SMB},t} \\ &+ \quad \left(h_{\text{down}} D_t^{\text{HML,down}} + h_{\text{flat}} D_t^{\text{HML,flat}} + h_{\text{up}} D_t^{\text{HML,up}}\right) r_{\text{HML},t} + e_t \end{aligned} \tag{2.9}$$

where D_t are the dummies that denote whether the market (MKT), the size factor (SMB) or the book-to-market factor (HML) during the ranking period ($t-7$ to $t-2$) went down by more than one standard deviation relative to the mean (down), or up by more than one standard deviation relative to their mean (up) or stayed within one standard deviation from their mean (flat). As such we estimate three coefficients per equity risk factor: one for the situation where the factor returns have been highly positive during momentum's ranking period, one where they have been strongly negative and one where they are around zero. This results in the estimation of ten coefficients since we also include a constant in the estimation.

The first column in Table 2.10 confirms standard momentum's time-varying exposures: for the market, size and value factor the coefficients after "Up" returns in the ranking period are significantly positive, while they are significantly negative after "Down" returns. The significance of these time-varying exposures of momentum is clear from the R-squared of 46.4 percent of this regression, whereas the unconditional equity risk factors only explain 15.1 percent of standard momentum's returns.

All alternative and combined momentum strategies have less time-varying exposure to the equity risk factors than the standard momentum strategy. The time-varying equity risk factors explain 4.4 percentage points of standard momentum's monthly volatility of 6.4 percent. The time-varying exposures explain less for all alternative and combined strategies. For example, these risk factors only explain 0.9 percentage points of the optimized constant-risk residual momentum, R-O-C, strategy's monthly volatility of 2.5 percent. The decrease in time-varying exposures of the alternative and combined strategies is also clear from the coefficients being closer to zero.

Table 2.10: **Regression of momentum returns on time-varying equity risk factors.** The estimated coefficients and t-statistics (in brackets), the standard deviation explained by these time-varying factors (Vol. Explained), the total standard deviation (Vol. Total) and the R-squared of the regression of the returns of the different momentum strategies on the time-varying equity risk factors market, size and value from July 1929 to December 2014. See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.

		S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Intercept		0.74%	0.69%	1.09%	0.93%	0.67%	1.22%	0.73%	0.89%	0.77%	1.08%
		(4.98)	(9.34)	(7.57)	(8.32)	(9.77)	(9.50)	(9.32)	(8.71)	(10.38)	(10.63)
Market	Up	0.39	0.03	0.24	0.03	0.03	0.03	0.01	0.05	0.03	0.02
		(4.32)	(0.77)	(2.76)	(0.42)	(0.74)	(0.44)	(0.11)	(0.78)	(0.71)	(0.39)
	Flat	0.06	0.01	0.06	-0.01	0.02	0.03	0.02	-0.01	0.02	0.01
		(1.53)	(0.40)	(1.64)	(-0.46)	(1.06)	(0.94)	(0.77)	(-0.28)	(0.94)	(0.36)
	Down	-0.64	-0.11	-0.28	-0.25	-0.08	-0.19	-0.24	-0.17	-0.16	-0.19
		(-13.48)	(-4.86)	(-6.14)	(-7.04)	(-3.58)	(-4.61)	(-9.46)	(-5.31)	(-6.90)	(-6.02)
Size	Up	0.61	-0.08	0.60	0.24	-0.03	0.11	0.10	0.19	0.12	0.22
		(5.80)	(-1.59)	(5.91)	(2.99)	(-0.65)	(1.26)	(1.88)	(2.63)	(2.37)	(3.03)
	Flat	-0.52	-0.36	-0.43	-0.22	-0.30	-0.48	-0.13	-0.20	-0.13	-0.17
		(-8.13)	(-11.28)	(-6.96)	(-4.54)	(-10.21)	(-8.74)	(-3.81)	(-4.48)	(-4.11)	(-3.94)
	Down	-0.41	-0.08	-0.27	0.35	-0.08	-0.20	-0.01	0.13	-0.04	-0.01
		(-4.18)	(-1.65)	(-2.80)	(4.76)	(-1.81)	(-2.38)	(-0.12)	(1.83)	(-0.77)	(-0.15)
Value	Up	0.69	-0.02	0.20	0.21	-0.01	0.08	0.04	0.11	0.04	0.11
		(6.53)	(-0.38)	(1.96)	(2.61)	(-0.23)	(0.90)	(0.74)	(1.49)	(0.69)	(1.55)
	Flat	-0.25	-0.23	-0.13	0.04	-0.17	-0.26	-0.02	0.05	-0.01	-0.01
		(-4.45)	(-8.24)	(-2.38)	(0.86)	(-6.51)	(-5.40)	(-0.61)	(1.19)	(-0.51)	(-0.30)
	Down	-1.02	-0.39	-0.55	-0.39	-0.24	-0.44	-0.30	-0.25	-0.19	-0.12
		(-10.78)	(-8.30)	(-5.91)	(-5.46)	(-5.59)	(-5.37)	(-6.08)	(-3.79)	(-3.95)	(-1.78)
Vol. Explained		4.4%	1.5%	2.5%	1.4%	1.1%	2.1%	1.2%	1.0%	0.9%	1.0%
Vol. Total		6.4%	2.8%	5.2%	3.8%	2.4%	4.5%	2.7%	3.4%	2.5%	3.4%
R-squared		46.4%	30.1%	23.0%	14.0%	21.7%	20.4%	18.5%	9.1%	12.2%	8.7%

The reduction in the time-varying exposures in Table 2.10 for the optimized and residual momentum strategy are perfectly understood from their construction. The residual momentum, R, strategy ranks stocks on their recent residual returns rather than their total returns. As such the equity risk factor loadings of the stocks in the top and bottom of this ranking are independent of the factor returns during momentum's ranking period, whereas the top and bottom of standard momentum's ranking are populated by stocks with either high or low factor exposures if the factor return has been highly positive or strongly negative. Consequently, residual momentum's equity risk factor exposures are close to zero over time. The optimized (O) momentum strategy has, by construction, less time-varying exposures but for different reasons. First and most importantly, the strategy takes into account risk by applying mean-variance optimization: the optimization prefers portfolios with less risk and as such more or less aims at long-short portfolios with zero or small exposures to the equity risk factors. Secondly, although optimized momentum's expected returns are based on the recent total returns that do depend on the factor returns over momentum's ranking period we adjust the expected returns to plus one and minus one for the stocks ranked in the top and bottom of momentum's ranking. As such, after highly positive or strongly negative factor returns during momentum's ranking period all stocks with either high or low factor loadings get an expected return of either plus one or minus one. Using these equal expected returns puts more emphasis on the risks and factor exposures of the stocks in the top and bottom of momentum's ranking when optimizing. For example, when we do not filter the expected returns to plus one, minus one and zero, but use the stocks' recent average returns as expected returns the time-varying factors explain 1.9 percentage points of the monthly volatility (not reported in Table 2.10) versus 1.5 percentage points for the optimization using the filtered expected returns in column 2 of Table 2.10.

The constant-risk momentum (C) strategy has the largest time-varying exposures among all alternative and combined strategies: the conditional model in Equation 2.9 still explains 2.5 percentage points of its monthly volatility. We ascribe this to the fact that this risk-adjusted momentum strategy responds somewhat delayed to changes in momentum's risk profile. Consider the case where factor returns have been relatively stable and positive for some time including during momentum's ranking period. In that case the standard momentum strategy builds up a positive exposure to these risk factors. So far, the risk-adjusted constant-risk strategy has not reduced the exposure to the standard momentum strategy. If in the investment month the factor returns realize strongly negative returns the standard momentum strategy realizes a relatively large loss. The risk-adjusted momentum strategy then also realizes a large loss since it has not been adapted to these built-up large factor exposures. Only after this realized loss, standard momentum's realized volatility increases and this constant-risk momentum strategy reduces its total exposures. In this way the constant-risk momentum strategy only reduces its factor loadings after it has already suffered from these large loadings.

Nevertheless, when adding the constant-risk constraint to the optimized standard and residual momentum strategies (O-C and O-C respectively) the time-varying risk factors explain even less of these strategies' volatilities: 1.1 and 0.9 percentage points of monthly volatility is ascribable to the time-varying exposures versus 1.5 and 1.2 percentage points for the optimized standard momentum (O) strategy in column 2 of Table 2.10 and optimized residual momentum (R-O), respectively.

2.7.2 Variation over the business cycle

The fact that momentum has a positive factor loading after a positive factor return during momentum's ranking period also has the implication that it suffers when the factor return is negative in the investment month. These factor return reversals occur, in particular, at the end of recessions. We establish in this section that standard equity momentum loses on average 2.1 percent per month in the second half of recessions while it earns on average 1.5 percent per month in the second half of expansionary periods when fewer factor return reversals occur. The optimized strategy only loses 0.1 percent per month at the end of recessionary periods while the combined strategies no longer experience losses in these periods. Moreover, the monthly average returns over the business cycle vary less than 1.0 percent for the optimized and combined strategies while this number is 3.6 percent for the standard momentum strategy.

The literature also points to these data features. Chordia and Shivakumar (2002) show that momentum's performance varies over the business cycle. Martens and Van Oord (2014), see Chapter 3, explain that these time-varying exposures play a key role in the time-varying performance over the business cycle. A consequence of the time-varying exposures is that momentum's returns vary over the business cycle. For example, momentum's performance suffers when a factor return has been strongly negative during momentum's ranking period and reverses to a positive return in the investment month; momentum then suffers from its built-up negative exposure to this particular factor. For the market risk factor these factor return reversals from negative to positive particularly occur at the end of recessions when equity markets start to recover from previous large losses, even before the real economy starts to recover. For example, momentum's two largest monthly losses occur in July 1932 and April 2009 when equity markets strongly recover after their largest losses in history.

Table 2.11 shows the average returns during four different stages of the business cycle of the different momentum strategies. These four stages of the business cycle are defined as

Table 2.11: **Return variation of different momentum strategies over the business cycle.** Average monthly returns during four different stages of the business cycle as well as the overall monthly average and the maximum return difference between the different stages of the business cycle. The four stages of the business cycle are based on the peaks and troughs of the US economy as established by the NBER. Stage I contains the months in the first half of expansionary periods and stage II contains the months in the second half of these expansionary periods. Stage III contains the months in the first halfs of recessions and stage IV are the latter halfs of recessions. See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.

		S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Expansion	Stage I	0.2%	0.3%	0.5%	0.6%	0.3%	0.7%	0.4%	0.6%	0.5%	0.7%
	Stage II	1.5%	0.8%	1.7%	1.3%	0.9%	1.6%	1.0%	1.3%	1.1%	1.4%
Recession	Stage III	0.8%	0.9%	1.8%	0.8%	0.8%	1.3%	1.2%	0.7%	1.1%	1.4%
	Stage IV	-2.1%	-0.1%	-0.7%	0.3%	0.0%	-0.3%	0.2%	0.3%	0.2%	0.4%
	Overall	0.5%	0.5%	1.0%	0.9%	0.5%	1.0%	0.7%	0.8%	0.7%	1.0%
	Difference	3.6%	0.9%	2.5%	1.0%	0.9%	1.9%	1.0%	1.0%	0.9%	1.1%

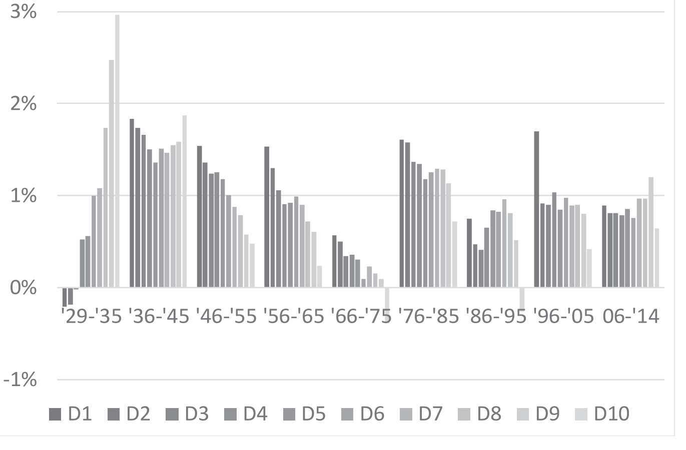
follows: first, the months are classified in expansionary and recessionary periods according to the NBER data on the US economy and secondly, these periods are split in two equal halfs to end up with four stages. The results in Table 2.11 confirm that momentum's returns vary over the business cycle: during the second stage of the business cycle in expansionary periods momentum earns on average 1.5 percent per month, while it loses on average 2.1 percent per month during the second halfs of recessions. That is a return difference of 3.6 percent per month over the business cycle.

The returns on all alternative and combined momentum strategies vary less over the business cycle. The returns on the constant-risk (C) momentum strategy still vary over the business cycle: the difference between its best and worst stages of the business cycle is still 2.5 percent per month, while this difference reduces to 1.0 percent for optimized (O) and residual (R) momentum and to 0.9 percent for the optimized constant-risk momentum (O-C and R-O-C) strategies. This is in line with our findings in the previous section that the constant-risk (C) strategy keeps the largest time-varying exposures, while the optimized constant-risk strategies exhibit the least time-varying exposures.

2.8 Episode analysis

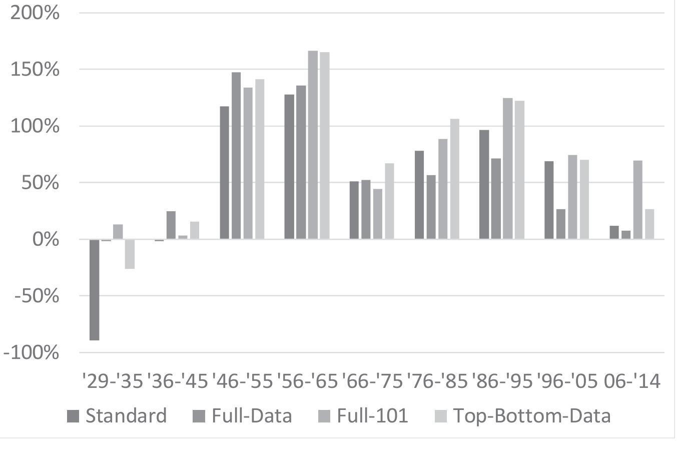
In this section we look at the performance of different strategies over severalsubperiods of our total ninety year period. One motivation to do so stems from the fact that we founda rather constant return pattern for deciles 2 to 9 in Figure 2, while Jegadeesh and Titman (1993) find a decreasing return pattern for these deciles. Figure 2.9 shows that this is a consequence of the different time periods considered in the analyses: Jegadeesh and Titman (1993) use data from 1965 to 1989 in which indeed a decreasing return pattern exists, whereas we use July 1929 to December 2014 that also covers the periods 1929 to 1945 and 1995 to 2014. This might indicate that the performance of the different strategies varies over time.

Figure 2.9: Average returns of different deciles of the equity momentum strategy during nine decades from 1929-2014.



Another motivation for the episode analysis stems from analyzing some first results for the nine decades for the case of the mean-variance optimized standard momentum strategies of Section 2.3. Given that Figure 2.9 already showed that momentum's top-bottom phenomenon varies over time, it is clear that the *Full-Data* strategy will perform better when momentum's top-bottom phenomenon is increasing over all deciles rather than that only stocks in decile 1 outperform and stocks in decile 10 underperform. Figure 2.10 shows the Sharpe ratios of the three optimization strategies compared to standard momentum for different decades and confirms the variation in the performance of the different optimization strategies. The *Full-Data* strategy has indeed a higher Sharpe ratio than standard momentum, particularly between 1946 to 1975 when momentum's top-bottom phenomenon in Figure 2.9 is strictly decreasing. During these years the optimization strategies *Full-101* and *Top-Bottom-Data* that take account of momentum's top and bottom phenomenon over the whole period also perform relatively well in terms of Sharpe ratio.

Figure 2.10: Sharpe ratios over time of different optimization strategies.



These results motivate us to explore the dynamic behavior of the risk and return characteristics of the ten momentum strategies and to focus on particular episodes. We do so in two different ways. We start to analyze the time series behavior of the average returns, volatilities and Sharpe ratios for the ten strategies over the total period of 1929-2014. Next, we specify three particular episodes, namely, (I) Pre-World-War-II period, denoted by Pre-war, where an economic crisis occurred and the preparation for World War II started; (II) Post-World-War-II, denoted by Post-war Restoration, since it may be characterized as an economic restoration period, which lasted until the first oil crisis in the early seventies; (III) Deregulation after the oil crisis, which is the period with economic liberation and several economic shocks with the most recent one in 2008 as the most severe. We investigate whether the return and crash risk characteristics and the possible effects of transaction costs on Sharpe ratios are different during these three episodes.

2.8.1 Dynamic patterns in average returns, volatilities and Sharpe ratios over the period 1929-2014

The time series of the 10-year annualized means of the ten different momentum strategies is shown in Figure 2.11. For the annualized volatilities it is shown in Figure 2.12, while Figure 2.13 contains the derived time series of the annualized Sharpe ratios. We summarize the major results as follows.

Means Strategy S performs poorly in the crisis period of the thirties with a substantial negative mean return. Similarly, strategies C and R perform poorly, albeit less so. The combination strategies R-O-C and R-O-C* do much better in terms of mean return during this period. After World-War II, during the recovery period, all strategies improve substantially in terms of mean return and show a more stable behavior. It is interesting to observe that strategy C does well in quiet periods like the fifties and nineties but does very poorly in turbulent times like the seventies and the recent crisis. This pattern is also typical for several other strategies. However, the combination strategy R-O-C* shows a rather robust behavior with respect to shocks and strategy R is a major component in this combination. It is also interesting to observe that strategy R does best in the recent period.

Volatilities There are three distinct periods of larger volatility and turbulence: the economic crisis of the thirties, after the first oil price shock in 1973, and during the recent crisis starting in 2008. It is seen that the individual strategies S, C and R are negatively affected by the high volatility, although residual momentum (R) only exhibits this high volatility in the beginning. The optimization strategy O copes better with volatility. The combination strategies R-O-C and R-O-C* are also much less affected by shocks and volatility. It is clear that during the relatively quiet periods in the fifties and early sixties as well as in the nineties both strategies R and O positively contribute to the low volatility of the combination strategies R-O-C and R-O-C*.

Sharpe ratios A first observation is that the Sharpe ratios appear as a rather non-stationary series that vary much over time for all strategies with negative values in the thirties for strategies S and C. A remarkable strong restoration process occurs until the mid fifties. Next, a very good period comes along until the first oil price shock and after that again a very good period over the nineties with a recent downturn that started already before the crisis of 2008. Clearly, the observed time series patterns in the volatilities have strongly affected the more stationary behavior of the mean returns.

It is clear that strategy S does not do well in terms of Sharpe ratio over most decades and, in particular, in the three periods of high volatility. Until 1945 most momentum strategies have relatively low or even negative Sharpe ratios. During and after the first oil price shock and during the years that cover the last financial crisis, 2008 to 2014, we also observe relatively low Sharpe ratios for all strategies.

With respect to the group of residual momentum strategies we find that the R-O-C and R-O-C* strategies have a higher Sharpe ratio than residual momentum over the decades running up to the first oil price shock. During the decades thereafter strategy R 'catches up' with a higher Sharpe ratio in recent times. This might be a consequence of the possibly increased comovements of the stock returns over time; in particular, the increase in the comovements of the residual returns. Mean-variance optimization then becomes less beneficial because of the decrease in the diversification benefits. It may require a modeling approach using factor models like in Chapter 4 to support this line of reasoning.

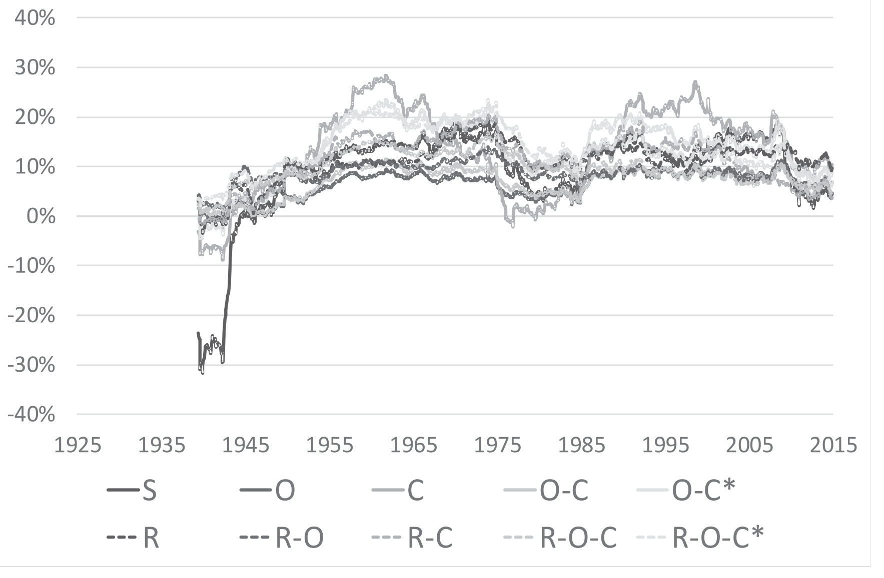
We conclude that the combination strategies show better Sharpe ratios than individual strategies and are robust in turbulent times. In quiet times an individual strategy like strategy R performs well in terms of Sharpe ratio. We note that the combination strategies also showed the best performance in terms of crash risk criteria in our earlier analysis, which confirms our present conclusion.

2.8.2 Three episodes of return and crash risk characteristics for all momentum strategies

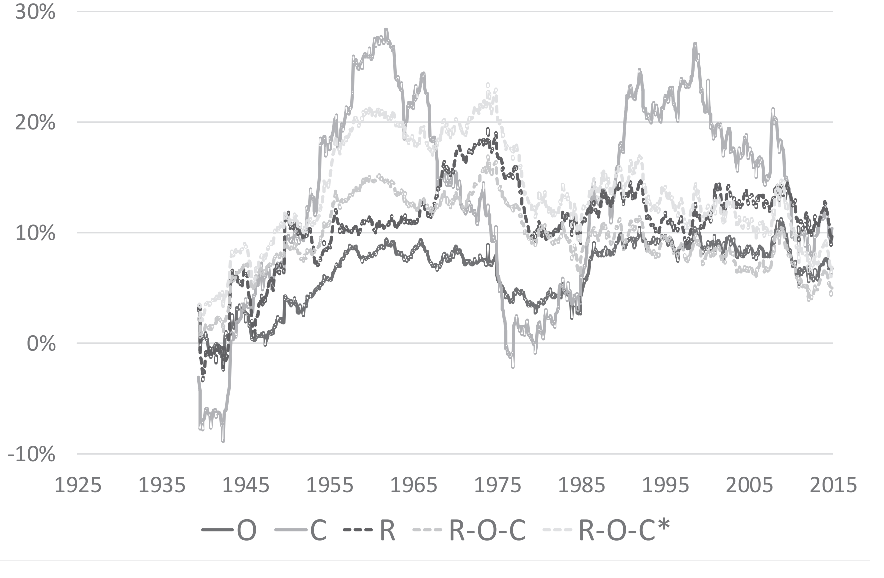
In this section we investigate how the data features on annualized means, volatilities and Sharpe ratios from the previous section translate into summary statistics for the three episodes. These results are reported jointly with the results on the four crash risk characteristics. Table 2.12 presents the main risk and return characteristics for three different subperiods, namely, (I) Pre-war period, (II) Post-war restoration period, which maybe characterized as a restoration period in industrialized countries that lasted until the first oil crisis in the early seventies; (III) Deregulation, post oil shock.

Episode I: Pre-war In this episode all momentum strategies have their worst performance as can be seen from the lowest, and even negative, Sharpe ratios as well as the

Figure 2.11: **Average returns over time.** The annualized means of the ten different momentum strategies over ten year moving windows between June 1939 to December 2014. See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.

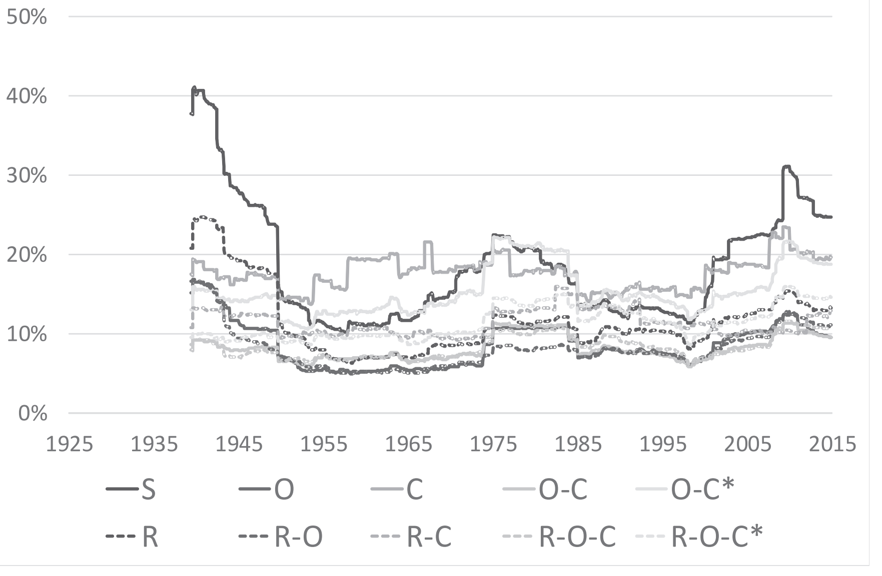


(a) All strategies

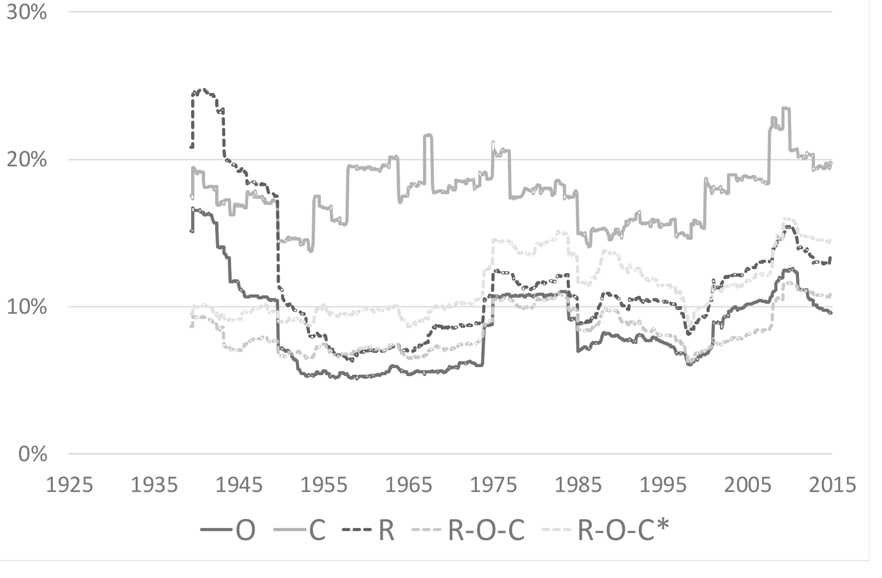


(b) Subset of five strategies

Figure 2.12: **Volatilities over time.** The annualized volatilities of the ten different momentum strategies over ten year moving windows between June 1939 to December 2014. See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.

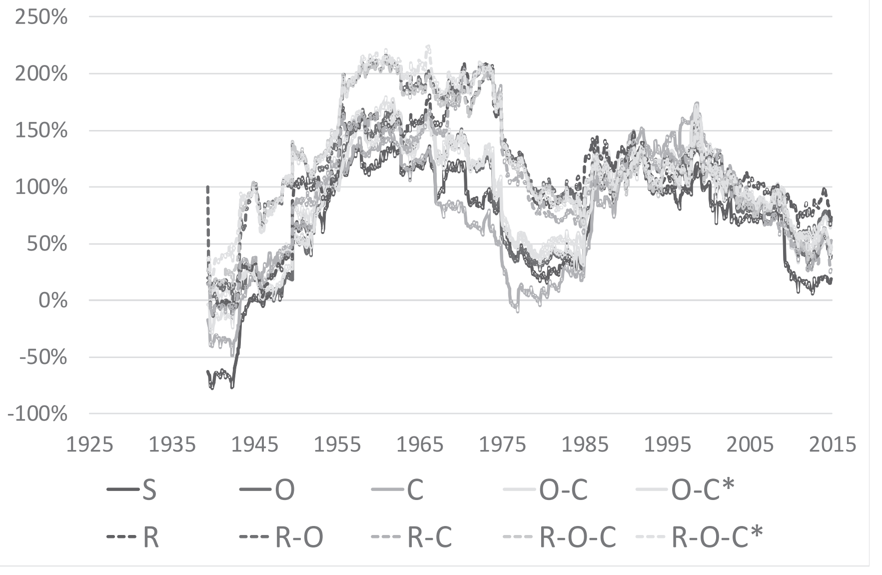


(a) All strategies

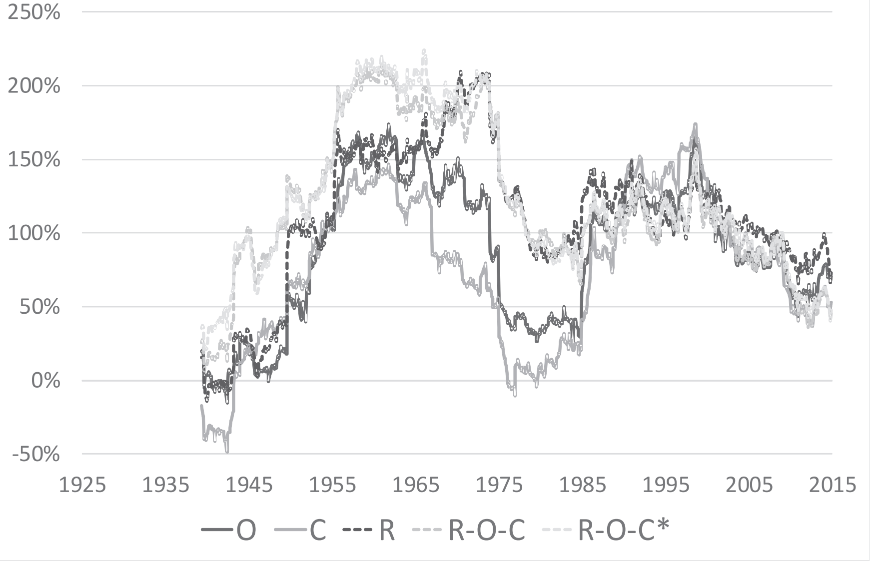


(b) Subset of five strategies

Figure 2.13: **Sharpe ratios over time.** The annualized Sharpe ratios of the ten different momentum strategies over ten year moving windows between June 1939 December 2014. See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.



(a) All strategies



(b) Subset of five strategies

Table 2.12: **Episode analysis.** Annualized mean, volatility and Sharpe ratio as well as lowest monthly return, realized 1% shortfall, maximum drawdown and maximum recovery period over the whole period and three different subperiods.

Whole period: July 1929 - December 2014

	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Mean	6.5%	6.1%	11.8%	10.3%	6.5%	12.2%	7.9%	10.1%	8.7%	12.4%
Volatility	22.2%	9.7%	18.0%	13.2%	8.4%	15.8%	9.5%	11.8%	8.7%	11.7%
Sharpe	29.3%	63.6%	65.4%	77.9%	76.9%	77.6%	83.5%	86.0%	100.3%	106.6%
Lowest	-57.3%	-23.8%	-35.0%	-40.1%	-17.3%	-34.0%	-24.0%	-24.6%	-13.8%	-18.7%
Shortfall	-38.1%	-15.1%	-22.7%	-18.0%	-10.8%	-20.9%	-12.9%	-14.8%	-9.8%	-12.0%
Drawdown	323.6%	44.3%	93.2%	63.0%	29.9%	62.0%	70.3%	30.2%	35.8%	35.5%
Recovery	430	197	198	84	82	184	109	126	104	70

Pre-war: July 1929 - December 1945

	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Mean	-15.2%	1.6%	-0.9%	3.1%	1.2%	0.4%	5.5%	1.6%	3.9%	5.4%
Volatility	34.8%	14.2%	17.8%	21.0%	8.4%	14.4%	14.2%	12.1%	8.4%	9.7%
Sharpe	-43.8%	11.3%	-4.9%	14.8%	14.6%	2.6%	38.6%	13.0%	46.4%	56.2%
Lowest	-54.8%	-23.8%	-29.0%	-40.1%	-16.2%	-30.1%	-24.0%	-24.6%	-13.2%	-10.6%
Shortfall	-54.8%	-23.8%	-29.0%	-40.1%	-16.2%	-30.1%	-24.0%	-24.6%	-13.2%	-10.6%
Drawdown	323.6%	44.3%	93.2%	63.0%	18.6%	61.5%	70.3%	30.2%	35.8%	21.6%
Recovery	186	156	157	84	82	156	109	126	104	59

Post-war restoration: January 1946 - December 1972

	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Mean	14.0%	7.4%	17.1%	13.1%	9.1%	18.5%	10.4%	15.6%	13.0%	18.5%
Volatility	14.3%	5.8%	17.9%	8.0%	7.1%	13.8%	5.9%	9.9%	7.3%	9.9%
Sharpe	97.3%	128.1%	95.7%	164.0%	126.9%	133.9%	176.7%	158.6%	179.6%	187.2%
Lowest	-19.3%	-6.1%	-35.0%	-5.6%	-7.7%	-17.0%	-4.9%	-8.1%	-5.5%	-8.4%
Shortfall	-16.9%	-5.4%	-20.9%	-5.2%	-6.6%	-14.0%	-4.1%	-7.4%	-5.3%	-7.3%
Drawdown	25.5%	8.6%	35.9%	8.2%	8.8%	17.0%	9.1%	11.8%	8.8%	10.6%
Recovery	29	21	29	10	21	9	13	10	13	11

Deregulation, post oil crisis: January 1973 - December 2014

	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Mean	9.9%	7.0%	13.1%	11.2%	6.8%	12.7%	7.2%	9.8%	7.7%	11.1%
Volatility	19.8%	9.5%	18.0%	11.9%	9.1%	17.2%	9.2%	12.6%	9.5%	13.2%
Sharpe	50.0%	73.8%	72.7%	93.7%	74.5%	73.6%	78.9%	77.4%	80.8%	84.3%
Lowest	-57.3%	-18.4%	-29.5%	-17.8%	-17.3%	-34.0%	-12.4%	-23.8%	-13.8%	-18.7%
Shortfall	-33.4%	-13.0%	-21.5%	-12.8%	-12.4%	-23.9%	-10.3%	-16.4%	-10.2%	-14.3%
Drawdown	93.4%	56.8%	92.4%	41.4%	53.9%	62.0%	42.4%	43.0%	40.6%	40.7%
Recovery	112	117	117	58	116	78	81	58	71	70

largest crash risk characteristics. The strategies that optimize residual momentum R-O, R-O-C and R-O-C* have the highest Sharpe ratios, while the constant-risk optimized residual momentum R-O-C and R-O-C* strategies also exhibit the least crash risk.

Episode II: Post-war, restoration This episode is characterized by the best performance for all strategies in terms of Sharpe ratios as well as crash risks. The higher average returns and lower volatilities result in the highest Sharpe ratios among the three episodes. The relative performance in terms of Sharpe ratios, which strategies outperform other strategies, is similar to that for the whole period. All strategies outperform standard momentum S. Residual momentum R performs best among the three proposed alternative momentum strategies and it is a major component in the combined strategies R-O-C and R-O-C* that themselves perform best among all strategies.

Episode III: Deregulation, post oil crisis All momentum strategies still have positive Sharpe ratios, but lower than during the episode leading up to the oil crisis. All proposed individual and combined momentum strategies still outperform the standard momentum strategy in terms of higher Sharpe ratios and lower crash risks. The performance of all proposed individual and combined momentum strategies is more or less similar as over the whole period. It is noteworthy that the constant-risk standard momentum strategies C and O-C* exhibit relatively large crash risk. Also the residual momentum strategy R stands out: its Sharpe ratio of 93.7% over this episode is the highest among all strategies, while its crash risk is close to the lowest crash risk among the strategies. This is different from the previous episodes and the full period analysis. As stated before in Section 2.8.1, the superior performance of the residual momentum strategy R might be caused by the increased correlation of the residual returns of the stocks, since increased correlations would reduce the diversification benefits from optimization.

The conclusion from this empirical analysis is that over different periods different strategies perform well. More specifically, it is seen that combination strategies work well in long periods of turbulent times, while it is noteworthy that the residual momentum strategy does well in more quiet times and in recent decades.

2.8.3 Statistical significance of Sharpe ratios over three episodes

In this section we explore possible variations in the statistical significance of the Sharpe ratios over the three episodes. We consider both the significance of the Sharpe ratios for each of the different strategies and, next, the significance of the differences in the Sharpe ratios between different strategies.

Table 2.13 shows the statistical significance of the Sharpe ratios of the different strategies for the whole period and the three episodes. The strategies' worst performance in terms of the lowest and even negative Sharpe ratios in the pre-war episode are also reflected in Sharpe ratios that are not significantly larger than zero. In the pre-war episode only the Sharpe ratios of the optimized constant-risk residual momentum (R-O-C and R-O-C*) strategies are significantly larger than zero at the 5% significance level. During the post-war restoration and deregulation episodes the Sharpe ratios of all momentum strategies are significantly larger than zero with the p-values of all strategies equalling 0.0%. This does not come as a surprise given the relatively high Sharpe ratios during these episodes.

Table 2.13: **Episode analysis of statistical significance of Sharpe ratios.**

Whole period: July 1929 - December 2014										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Sharpe	29.3%	63.6%	65.4%	77.9%	76.9%	77.6%	83.5%	86.0%	100.3%	106.6%
p-value	1.04%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Pre-war: July 1929 - December 1945										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
SR	-43.8%	11.3%	-4.9%	14.8%	14.6%	2.6%	38.6%	13.0%	46.4%	56.2%
p-value	95.52%	30.27%	58.36%	29.99%	23.91%	46.01%	7.71%	31.31%	4.91%	1.40%
Post-war restoration: January 1946 - December 1972										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
SR	97.3%	128.1%	95.7%	164.0%	126.9%	133.9%	176.7%	158.6%	179.6%	187.2%
p-value	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Deregulation, post oil crisis: January 1973 - December 2014										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
SR	50.0%	73.8%	72.7%	93.7%	74.5%	73.6%	78.9%	77.4%	80.8%	84.3%
p-value	0.09%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Table 2.14 shows the statistical significance of the variation in the Sharpe ratios over the three episodes. A first observation is that all strategies have a significantly higher Sharpe ratio than standard momentum in any of the three different episodes at the 5% significance level, except for the constant-risk momentum strategy C during the restoration period. The relatively high and significantly positive Sharpe ratio of the strategy R-O-C* during

the pre-war episode has also resulted in a significantly higher Sharpe ratio than the other strategies at the 5% level, except when it is compared with the other two optimized residual momentum strategies R-O and R-O-C. In the post-war restoration episode we also find that the Sharpe ratio of this strategy is significantly higher than any other momentum strategy except for the other two optimized residual momentum strategies.

During the deregulation and post oil crisis episode we find hardly any significantly higher Sharpe ratios among the different strategies, apart from the significantly lower Sharpe ratios for the standard momentum strategy S. This is a consequence of the level of the Sharpe ratios being close to each other in this episode. The highest Sharpe ratio of the residual momentum strategy R is not significantly higher than the Sharpe ratios of the other strategies at the 5% level, except with respect to the constant-risk residual momentum strategy R-C.

Table 2.14: **Episode analysis of statistical significance of Sharpe ratio differences.**

Whole period: July 1929 - December 2014										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
S		100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
O	0.00%		59.51%	92.64%	99.99%	99.50%	99.37%	98.56%	100.00%	100.00%
C	0.00%	40.49%		85.70%	90.15%	93.25%	95.68%	97.61%	99.94%	100.00%
R	0.00%	7.36%	14.30%		47.04%	49.61%	77.62%	89.72%	99.87%	99.99%
O-C	0.00%	0.01%	9.85%	52.96%		56.92%	76.90%	81.21%	99.82%	100.00%
O-C*	0.00%	0.50%	6.75%	50.39%	43.08%		75.75%	80.61%	99.94%	100.00%
R-O	0.00%	0.63%	4.32%	22.38%	23.10%	24.25%		61.65%	100.00%	100.00%
R-C	0.00%	1.44%	2.39%	10.28%	18.79%	19.39%	38.35%		95.70%	99.69%
R-O-C	0.00%	0.00%	0.06%	0.13%	0.18%	0.06%	0.00%	4.30%		97.06%
R-O-C*	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.31%	2.94%	
Pre-war: July 1929 - December 1945										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
S		99.74%	99.95%	98.50%	99.65%	99.59%	100.00%	98.61%	100.00%	100.00%
O	0.26%		20.68%	56.03%	69.26%	27.99%	91.25%	53.92%	95.80%	98.25%
C	0.05%	79.32%		79.66%	82.62%	65.29%	98.30%	78.58%	99.29%	99.84%
R	1.50%	43.97%	20.34%		49.25%	26.70%	88.75%	43.21%	96.08%	99.58%
O-C	0.35%	30.74%	17.38%	50.75%		16.57%	86.80%	48.30%	94.33%	98.34%
O-C*	0.41%	72.01%	34.71%	73.30%	83.43%		97.05%	71.22%	99.56%	99.99%
R-O	0.00%	8.75%	1.70%	11.25%	13.20%	2.95%		10.13%	85.56%	89.04%
R-C	1.39%	46.08%	21.42%	56.79%	51.70%	28.78%	89.87%		97.02%	99.87%
R-O-C	0.00%	4.20%	0.71%	3.92%	5.67%	0.44%	14.44%	2.98%		81.80%
R-O-C*	0.00%	1.75%	0.16%	0.42%	1.66%	0.01%	10.96%	0.13%	18.20%	
Post-war restoration: January 1946 - December 1972										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
S		98.75%	46.07%	100.00%	99.00%	99.98%	100.00%	99.96%	100.00%	100.00%
O	1.25%		3.25%	96.57%	39.19%	75.01%	99.92%	93.85%	99.95%	100.00%
C	53.93%	96.75%		99.87%	96.90%	99.57%	100.00%	99.95%	100.00%	100.00%
R	0.00%	3.43%	0.13%		2.50%	4.08%	87.35%	33.23%	90.17%	98.44%
O-C	1.00%	60.81%	3.10%	97.50%		81.99%	99.94%	94.53%	99.98%	100.00%
O-C*	0.02%	24.99%	0.43%	95.92%	18.01%		99.94%	91.17%	99.98%	100.00%
R-O	0.00%	0.08%	0.00%	12.65%	0.06%	0.06%		13.23%	78.89%	94.78%
R-C	0.04%	6.15%	0.05%	66.77%	5.47%	8.83%	86.77%		91.60%	97.96%
R-O-C	0.00%	0.05%	0.00%	9.83%	0.02%	0.02%	21.11%	8.40%		91.40%
R-O-C*	0.00%	0.00%	0.00%	1.56%	0.00%	0.00%	5.22%	2.04%	8.60%	
Deregulation, post oil crisis: January 1973 - December 2014										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
S		99.27%	98.69%	100.00%	99.52%	99.81%	99.88%	97.71%	99.80%	99.92%
O	0.73%		47.98%	94.29%	59.03%	48.37%	70.64%	61.65%	75.85%	86.50%
C	1.31%	52.02%		90.03%	54.16%	51.04%	65.89%	61.87%	70.51%	77.82%
R	0.00%	5.71%	9.97%		6.90%	5.11%	6.18%	3.95%	9.64%	13.50%
O-C	0.48%	40.97%	45.84%	93.10%		42.60%	68.10%	58.95%	74.66%	85.98%
O-C*	0.19%	51.63%	48.96%	94.89%	57.40%		73.56%	61.45%	79.40%	90.64%
R-O	0.12%	29.36%	34.11%	93.82%	31.90%	26.44%		45.34%	72.77%	87.20%
R-C	2.29%	38.35%	38.13%	96.05%	41.05%	38.55%	54.66%		61.74%	74.00%
R-O-C	0.20%	24.15%	29.49%	90.36%	25.34%	20.60%	27.23%	38.26%		80.05%
R-O-C*	0.08%	13.50%	22.18%	86.50%	14.02%	9.36%	12.80%	26.00%	19.95%	

2.8.4 Transaction costs over three episodes

In this section we start to investigate the variation in the number and size of the transactions over the three episodes. Next, we consider break-even transaction costs and the transaction costs for which the Sharpe ratios of the strategies remain significantly larger than zero at the 5% level. We then look at the level of the transaction costs for which the different strategies outperform the other strategies in terms of a higher Sharpe ratio.

Table 2.15 shows the number and size of the transactions for the different strategies. Given that the number of stocks increases over time, the number of transactions for the different strategies also increases over the three different episodes. However, the pattern of the number of transactions over the different strategies does not change over the three episodes. The size of transactions does change over the different episodes. In particular, the strategies that involve the constant-risk setting have relatively smaller transactions during the pre-war episode than in the other two episodes. This is due to the relatively high volatility of the standard S and residual R strategies during this episode; these constant-risk momentum strategies deleverage in this episode to get a constant-risk over the full period and as such attach relatively small weights to all stocks in this episode. Relatively smaller weights obviously also result in relatively smaller transactions. The other strategies that do not have an ex-ante constant-risk setting over the whole period have a relatively stable transaction size over the different episodes. This stems from the fact that the standard S and residual R strategies as well as their optimized counterparts O and R-O have a 1$ long and 1$ short exposure in every investment month; the weights attached to the different stocks are of a similar magnitude in all episodes.

Given that the performance of the different strategies, in terms of Sharpe ratios, varies over the three different episodes we also find that the break-even transaction costs vary over these episodes. The break-even transaction costs in Table 2.15 are high during the episodes that all momentum strategies perform well and low, and even negative for some strategies, during the pre-war episode. This also holds for the break-even transaction costs for which the strategies still have a significantly positive Sharpe ratio at the 5% significance level. Only the optimized constant-risk residual momentum strategies R-O-C and R-O-C* keep a significantly positive Sharpe ratio for positive transaction costs.

Negative transaction costs indicate that an investor would have to get paid for doing transactions in order to end up with a significantly positive Sharpe ratio. In the pre-war episode transaction costs of up to 0.35% of the investment size are required in order to have the best strategy R-O-C* keep its positive Sharpe ratio; it is questionable if transaction costs were that low at that time. During the post-war restoration episode all momentum strategies keep a, significantly, positive Sharpe ratio for relatively high transaction costs; for transaction costs up to 0.33% all strategies keep a positive Sharpe ratio and for standard momentum this is even the case for transaction costs up to 0.71%. Finally, the lower Sharpe ratios during the deregulation post oil crisis episode are also reflected in lower break-even transaction costs.

Table 2.15: **Episode analysis of transaction costs.**

Whole period: July 1929 - December 2014										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Sharpe	29.3%	63.6%	65.4%	77.9%	76.9%	77.6%	83.5%	86.0%	100.3%	106.6%
Number	4248	15497	4249	4345	15497	4227	15497	4346	15497	4322
Size	20.3	21.2	44.5	21.8	21.5	25.1	23.1	39.2	23.8	25.9
BE SR>0	0.32%	0.29%	0.26%	0.47%	0.30%	0.49%	0.34%	0.26%	0.36%	0.48%
BE p(5%)	0.10%	0.21%	0.19%	0.36%	0.23%	0.37%	0.26%	0.20%	0.29%	0.39%
Pre-war: July 1929 - December 1945										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Sharpe	-43.8%	11.3%	-4.9%	14.8%	14.6%	2.6%	38.6%	13.0%	46.4%	56.2%
Number	1977	7084	1982	1988	7084	1952	7084	1993	7084	1960
Size	22.4	22.2	27.3	23.0	15.3	16.5	23.6	25.1	15.3	15.4
BE SR>0	-0.68%	0.07%	-0.03%	0.14%	0.08%	0.02%	0.23%	0.06%	0.25%	0.35%
BE p(5%)	-1.48%	-0.17%	-0.30%	-0.29%	-0.11%	-0.35%	-0.04%	-0.16%	0.00%	0.09%
Post-war restoration: January 1946 - December 1972										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Sharpe	97.3%	128.1%	95.7%	164.0%	126.9%	133.9%	176.7%	158.6%	179.6%	187.2%
Number	3765	13679	3765	3864	13679	3736	13679	3864	13679	3834
Size	19.7	20.3	51.1	21.7	25.3	30.0	22.9	47.7	28.4	31.8
BE SR>0	0.71%	0.37%	0.33%	0.60%	0.36%	0.61%	0.46%	0.33%	0.46%	0.58%
BE p(5%)	0.53%	0.28%	0.23%	0.48%	0.28%	0.50%	0.38%	0.26%	0.39%	0.49%
Deregulation, post oil crisis: January 1973 - December 2014										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
Sharpe	50.0%	73.8%	72.7%	93.7%	74.5%	73.6%	78.9%	77.4%	80.8%	84.3%
Number	5429	19892	5430	5557	19892	5415	19892	5558	19892	5541
Size	19.9	21.3	46.7	21.5	21.4	25.1	23.0	39.0	23.9	26.2
BE SR>0	0.50%	0.33%	0.28%	0.52%	0.32%	0.50%	0.31%	0.25%	0.32%	0.43%
BE p(5%)	0.24%	0.22%	0.18%	0.38%	0.21%	0.33%	0.21%	0.17%	0.21%	0.29%

Table 2.16 now presents the ranges of transaction costs for which a particular momentum strategy outperforms the other momentum strategies over the three different episodes. We find that during the pre-war episode the optimized constant-risk residual momentum strategies R-O-C and R-O-C* outperform all other strategies for a range of transaction costs for which these strategies keep a positive Sharpe ratio themselves, with the strategy with lower costs R-O-C* performing best. This is mainly a consequence of the other strategies having a relatively low Sharpe ratio during this episode. During the post-war restoration episode we also find that the R-O-C* strategy outperforms all other strategies for moderate transaction costs, up to a level around 0.50%. Still, the standard S, residual R and optimized constant-risk at lower costs O-C* strategies outperform this strategy for relatively high transaction costs above 0.50%, up to 0.71% for standard momentum. In this episode the average size of the transactions of the R-O-C* strategy is relatively large and as such these strategies with a relatively smaller size of the transactions are able to outperform for relatively high transaction costs. During the deregulation post oil crisis episode residual momentum R has a higher Sharpe ratio than any other momentum strategy for transaction costs up to 0.52%. This is in line with its Sharpe ratio being the highest in this episode, although not significantly higher than most other strategies.

Table 2.16: **Episode analysis of mutual outperformance for ranges of transaction costs.**

Whole period: July 1929 - December 2014										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
S		0.27% 0.32%	0.23% 0.32%		0.29% 0.32%			0.24% 0.32%		
O	0.00% 0.27%		0.07% 0.29%					0.20% 0.29%		
C	0.00% 0.23%	0.00% 0.07%						0.24% 0.26%		
R	0.00% 0.47%	0.00% 0.47%	0.00% 0.47%		0.00% 0.47%	0.00% 0.06%	0.07% 0.47%	0.05% 0.47%	0.20% 0.47%	
O-C	0.00% 0.29%	0.00% 0.30%	0.00% 0.30%					0.12% 0.30%		
O-C*	0.00% 0.49%	0.00% 0.49%	0.00% 0.49%	0.06% 0.49%	0.00% 0.49%		0.07% 0.49%	0.05% 0.49%	0.20% 0.49%	0.46% 0.49%
R-O	0.00% 0.34%	0.00% 0.34%	0.00% 0.34%	0.00% 0.07%	0.00% 0.34%	0.00% 0.07%		0.03% 0.34%		
R-C	0.00% 0.24%	0.00% 0.20%	0.00% 0.24%	0.00% 0.05%	0.00% 0.12%	0.00% 0.05%	0.00% 0.03%			
R-O-C	0.00% 0.36%	0.00% 0.36%	0.00% 0.36%	0.00% 0.20%	0.00% 0.36%	0.00% 0.20%	0.00% 0.36%	0.00% 0.36%		
R-O-C*	0.00% 0.48%	0.00% 0.48%	0.00% 0.48%	0.00% 0.48%	0.00% 0.48%	0.00% 0.46%	0.00% 0.48%	0.00% 0.48%	0.00% 0.48%	
Pre-war: July 1929 - December 1945										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
S										
O	0.00% 0.07%		0.00% 0.07%			0.00% 0.07%		0.03% 0.07%		
C										
R	0.00% 0.14%	0.00% 0.14%	0.00% 0.14%		0.00% 0.14%	0.00% 0.14%		0.00% 0.14%		
O-C	0.00% 0.08%	0.00% 0.08%	0.00% 0.08%			0.00% 0.08%		0.00% 0.08%		
O-C*	0.00% 0.02%		0.00% 0.02%							
R-O	0.00% 0.23%	0.00% 0.23%	0.00% 0.23%	0.00% 0.23%	0.00% 0.23%	0.00% 0.23%		0.00% 0.23%		
R-C	0.00% 0.06%	0.00% 0.03%	0.00% 0.06%			0.00% 0.06%				
R-O-C	0.00% 0.25%	0.00% 0.25%	0.00% 0.25%	0.00% 0.25%	0.00% 0.25%	0.00% 0.25%	0.00% 0.25%	0.00% 0.25%		
R-O-C*	0.00% 0.35%	0.00% 0.35%	0.00% 0.35%	0.00% 0.35%	0.00% 0.35%	0.00% 0.35%	0.00% 0.35%	0.00% 0.35%	0.00% 0.35%	

Continued on next page

Table 2.16 – *Continued from previous page*

Post-war restoration: January 1946 - December 1972										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
S		0.14%	0.00%	0.49%	0.14%	0.45%	0.32%	0.18%	0.32%	0.49%
		0.71%	0.71%	0.71%	0.71%	0.71%	0.71%	0.71%	0.71%	0.71%
O	0.00%		0.00%		0.00%			0.23%		
	0.14%		0.37%		0.37%			0.37%		
C								0.32%		
								0.33%		
R	0.00%	0.00%	0.00%		0.00%	0.00%	0.11%	0.00%	0.13%	0.46%
	0.49%	0.60%	0.60%		0.60%	0.56%	0.60%	0.60%	0.60%	0.60%
O-C	0.00%		0.00%					0.24%		
	0.14%		0.36%					0.36%		
O-C*	0.00%	0.00%	0.00%	0.56%	0.00%		0.25%	0.09%	0.26%	0.51%
	0.45%	0.61%	0.61%	0.61%	0.61%		0.61%	0.61%	0.61%	0.61%
R-O	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%		0.00%		
	0.32%	0.46%	0.46%	0.11%	0.46%	0.25%		0.46%		
R-C	0.00%	0.00%	0.00%		0.00%	0.00%				
	0.18%	0.23%	0.32%		0.24%	0.09%				
R-O-C	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%		
	0.32%	0.46%	0.46%	0.13%	0.46%	0.26%	0.46%	0.46%		
R-O-C*	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	
	0.49%	0.58%	0.58%	0.46%	0.58%	0.51%	0.58%	0.58%	0.58%	
Deregulation, post oil crisis: January 1973 - December 2014										
	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
S		0.19%	0.14%		0.18%		0.19%	0.13%	0.20%	0.35%
		0.50%	0.50%		0.50%		0.50%	0.50%	0.50%	0.50%
O	0.00%		0.00%		0.07%	0.00%	0.19%	0.04%	0.24%	
	0.19%		0.33%		0.33%	0.00%	0.33%	0.33%	0.33%	
C	0.00%							0.10%		
	0.14%							0.28%		
R	0.00%	0.00%	0.00%		0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	0.52%	0.52%	0.52%		0.52%	0.52%	0.52%	0.52%	0.52%	0.52%
O-C	0.00%	0.00%	0.00%			0.00%	0.27%	0.04%		
	0.18%	0.07%	0.32%			0.01%	0.32%	0.32%		
O-C*	0.00%	0.00%	0.00%		0.01%		0.05%	0.02%	0.07%	0.21%
	0.50%	0.50%	0.50%		0.50%		0.50%	0.50%	0.50%	0.50%
R-O	0.00%	0.00%	0.00%		0.00%	0.00%		0.00%		
	0.19%	0.19%	0.31%		0.27%	0.05%		0.31%		
R-C	0.00%	0.00%	0.00%		0.00%	0.00%				
	0.13%	0.04%	0.10%		0.04%	0.02%				
R-O-C	0.00%	0.00%	0.00%		0.00%	0.00%	0.00%	0.00%		
	0.20%	0.24%	0.32%		0.32%	0.07%	0.32%	0.32%		
R-O-C*	0.00%	0.00%	0.00%		0.00%	0.00%	0.00%	0.00%	0.00%	
	0.35%	0.43%	0.43%		0.43%	0.21%	0.43%	0.43%	0.43%	

2.9 Conclusions

In this chapter we investigate the working of several equity momentum strategies and compare their performances in terms of return and crash risk characteristics. We use data that contain all monthly NYSE and AMEX common equity returns from July 1929 until December 2014 from CRSP. Our conclusions are listed in two groups. The first group is related to the whole period and the second one to three episodes that together form the full period from July 1929 to December 2014.

For the whole period we find that basic mean-variance optimization, in particular, when taking account of momentum's top-bottom phenomenon, improves standard momentum's risk and return characteristics and reduces momentum's time-varying exposures. This top-bottom phenomenon is the fact that the stocks in the top of momentum's ranking outperform and that stocks in the bottom of the ranking underperform. The other alternative momentum strategies, residual momentum by Blitz et al. (2011) and constant-risk momentum by Barroso and Santa-Clara (2015) also outperform standard momentum. No individual strategy is superior over the whole period in all characteristics. Mean-variance optimization is best in crash risk reduction, while residual momentum is best in Sharpe ratio improvement.

Combining all three alternative strategies in a joint strategy further improves momentum's performance and reduces its crash risk substantially. The Sharpe ratio improves from 29.3% to 106.6%, while crash risk indicators like the largest monthly loss, the maximum drawdown and the time to recover to the level of the previous peak after a drawdown improve from respectively 57.3%, 323.6% and 430 months for standard momentum to 18.7%, 35.5% and 70 months. The joint strategy is also able to outperform residual and constant-risk momentum. The obtained results are statistically significant.

Although transaction costs play a key role in the performance of the momentum strategies, the outperformance of standard momentum by other strategies is independent of the transaction size. Combining the three individual strategies improves momentum's performance for a similar number of transactions while only slightly increasing the average size of the transactions. For transaction costs up to 0.46% we find that this joint momentum strategy outperforms any other momentum strategy, including residual and constant-risk momentum.

A second group of conclusions refers to how the results vary over time and in particular episodes. A major concern with respect to the performance of equity momentum strategies is their sensitivity to special data features as turbulence and shocks that may occur in particular episodes.

We find that all three alternative momentum strategies, as well as the combined strategies, reduce the time-variation in momentum's exposures and returns. This also implies that the returns become more stable over different stages of the business cycle.

Episode analysis shows that the performance of the different strategies varies substantially over time. Nevertheless, the performance of our proposed optimized and combined momentum strategies over standard momentum is robust for the choice of the episode, also when moderate transaction costs are taken into account.

Episode analysis also shows reasonably stationary behavior of mean returns for all strategies after the economic crisis of the thirties and turbulent behavior of the volatilities in three subperiods: the crisis of the thirties, the oil price shock of 1973 and the recent crisis since 2008. This strongly affects the dynamic pattern of the Sharpe ratios over the three episodes. It leads to the conclusion that combined strategies work fine over the whole period with very good performance during turbulent periods, whereas residual momentum appears to be a good strategy in relatively quiet times. These results are statistically significant and not sensitive for a reasonable range of transaction costs.

In general, we conclude that the proposed combination of equity momentum strategies may provide practitioners with an attractive strategy with good performance in terms of return and crash risk over a long period of time. Our second general conclusion is that different strategies may perform better in different episodes depending upon the amount and level of turbulence and volatility. The results presented in this chapter also indicate that momentum's improved performance is still a challenge for the concept of market efficiency and this remains a topic for further academic research.

Some topics for further research Given the possible sensitivity of our results for the chosen data, one may explore different data sets than the one we considered. That is, data for different countries, like emerging markets, and data for particular sectors of the economy.

Searching for a causal explanation of the improved performance will require more detailed and complex modeling.

Exploring the uncertainty in the results and connect this inferential uncertainty with forecast and policy uncertainty and effectiveness in a coherent way remains a major research topic. A very first methodological attempt is presented in Chapter 4.

Given that different strategies perform well in different time periods, exploring a weighted time-varying combination of strategies where the weights depend on the past performance of a strategy might be beneficial. For some first results on this topic we refer to Casarin, Grassi, Ravazzolo, and Van Dijk (2015).

All these topics will require more data, better and more complex modeling and more intense computational procedures where possibly parallel computing is needed. Connections between this financial econometric topic and the field of data science should be fruitfully pursued.

2.A Analytical Solution to the Mean-Variance Optimization Problem

Every investment month we solve the following problem to form the optimized momentum portfolios.

$$\min_{h^*} \left\{ h^{*\prime} V h^* \left| h^{*\prime} \iota = 0, h^{*\prime} f = 1 \right. \right\} \tag{2.10}$$

We use the Lagrange function

$$L\left(h^*, \lambda_1, \lambda_2\right) = h^{*\prime} V h^* - \lambda_1 h^{*\prime} \iota - \lambda_2 \left(h^{*\prime} f - 1\right) \tag{2.11}$$

and set its derivatives to zero

$$\frac{\delta L}{\delta h^*} = 2Vh^* - \lambda_1 \iota - \lambda_2 f = 0 \tag{2.12}$$

$$\frac{\delta L}{\delta \lambda_1} = h^{*\prime} \iota = 0 \tag{2.13}$$

$$\frac{\delta L}{\delta \lambda_2} = h^{*\prime} f - 1 = 0 \tag{2.14}$$

The optimal solution for h^* becomes

$$h^* = \frac{1}{2} V^{-1} \left(\lambda_1 \iota + \lambda_2 f\right) \tag{2.15}$$

We then solve for λ_1

$$\begin{aligned} h^{*\prime} \iota &= \frac{1}{2} V^{-1} \left(\lambda_1 \iota + \lambda_2 f\right)' \iota = \frac{1}{2} \left(\lambda_1 \iota' V^{-1} \iota + \lambda_2 f' V^{-1} \iota\right) = 0 \\ \lambda_1 &= \frac{-\lambda_2 f' V^{-1} \iota}{\iota' V^{-1} \iota} \end{aligned} \tag{2.16}$$

and for λ_2

$$\begin{aligned} h^{*\prime} f - 1 &= \left(\frac{1}{2} V^{-1} \left(\lambda_1 \iota + \lambda_2 f\right)\right)' f - 1 = \frac{1}{2} \left(\lambda_1 \iota' V^{-1} f + \lambda_2 f' V^{-1} f\right) - 1 = 0 \\ \lambda_2 &= \frac{2 - \lambda_1 \iota' V^{-1} f}{f' V^{-1} f} \end{aligned} \tag{2.17}$$

which results in the following solutions for λ_1

$$\begin{aligned}
\lambda_1 &= \frac{-\lambda_2 f'V^{-1}\iota}{\iota'V^{-1}\iota} = \frac{-\left[\frac{2-\lambda_1\iota'V^{-1}f}{f'V^{-1}f}\right] f'V^{-1}\iota}{\iota'V^{-1}\iota} = \frac{\lambda_1\iota'V^{-1}ff'V^{-1}\iota - 2f'V^{-1}\iota}{f'V^{-1}f\iota'V^{-1}\iota} \\
\lambda_1 &= -2\frac{f'V^{-1}\iota}{f'V^{-1}f\iota'V^{-1}\iota - \iota'V^{-1}ff'V^{-1}\iota}
\end{aligned} \tag{2.18}$$

and λ_2

$$\begin{aligned}
\lambda_2 &= \frac{2-\lambda_1\iota'V^{-1}f}{f'V^{-1}f} = \frac{2-\left[-2\frac{f'V^{-1}\iota}{f'V^{-1}f\iota'V^{-1}\iota-\iota'V^{-1}ff'V^{-1}\iota}\right]\iota'V^{-1}f}{f'V^{-1}f} \\
\lambda_2 &= 2\frac{1+\frac{f'V^{-1}\iota\iota'V^{-1}f}{f'V^{-1}f\iota'V^{-1}\iota-\iota'V^{-1}ff'V^{-1}\iota}}{f'V^{-1}f}
\end{aligned} \tag{2.19}$$

These Lagrange multipliers λ_1 and λ_2 then yield the following optimal solution

$$\begin{aligned}
h^* &= \frac{1}{2}V^{-1}(\lambda_1\iota + \lambda_2 f) \\
&= \frac{1}{2}V^{-1}\left(-2\left[\frac{f'V^{-1}\iota}{f'V^{-1}f\iota'V^{-1}\iota - \iota'V^{-1}ff'V^{-1}\iota}\right]\iota\right. \\
&\quad \left. +2\left[\frac{1+\frac{f'V^{-1}\iota\iota'V^{-1}f}{f'V^{-1}f\iota'V^{-1}\iota-\iota'V^{-1}ff'V^{-1}\iota}}{f'V^{-1}f}\right] f\right) \\
&= \frac{V^{-1}f}{f'V^{-1}f} + \frac{f'V^{-1}\iota\iota'V^{-1}f}{f'V^{-1}f\iota'V^{-1}\iota - \iota'V^{-1}ff'V^{-1}\iota} \times \frac{V^{-1}f}{f'V^{-1}f} \\
&\quad - \frac{f'V^{-1}\iota}{f'V^{-1}f\iota'V^{-1}\iota - \iota'V^{-1}ff'V^{-1}\iota}V^{-1}\iota \\
&= \frac{f'V^{-1}f\iota'V^{-1}\iota V^{-1}f - f'V^{-1}\iota f'V^{-1}fV^{-1}\iota}{f'V^{-1}f\left(f'V^{-1}f\iota'V^{-1}\iota - (f'V^{-1}\iota)^2\right)} \\
&= \frac{\iota'V^{-1}\iota V^{-1}f - f'V^{-1}\iota V^{-1}\iota}{f'V^{-1}f\iota'V^{-1}\iota - (f'V^{-1}\iota)^2} \\
&= \frac{V^{-1}f - f'V^{-1}\iota(\iota'V^{-1}\iota)^{-1}V^{-1}\iota}{f'V^{-1}f - (f'V^{-1}\iota)^2(\iota'V^{-1}\iota)^{-1}} \\
&= \frac{V^{-1}(f - C_1\iota)}{C_2}
\end{aligned} \tag{2.20}$$

where $C_1 = f'V^{-1}\iota(\iota'V^{-1}\iota)^{-1}$ and $C_2 = f'V^{-1}f - (f'V^{-1}\iota)^2(\iota'V^{-1}\iota)^{-1}$. I.e. the optimized portfolio is a weighted average of demeaned returns f, and the weights are determined according to inverse of the variance-covariance matrix V.

2.B Momentum strategies with different percentiles

Throughout this chapter all strategies were based on dividing stocks over the top and bottom deciles of the ranking on total or residual returns. In this appendix we consider the dependence of the performance of the strategies on this choice for deciles by considering other percentiles.

Table 2.B.1 presents the Sharpe-ratios of the momentum strategies if we use percentiles other than deciles (tenth percentile). The 10 percent row in this table corresponds to the Sharpe-ratios for the compared strategies in Table 2.3. For example, the standard momentum (S) strategy would have a Sharpe ratio of 37.8% if only the stocks in the top and bottom fifth percentile are used for the strategy. The increasing Sharpe ratios for smaller percentiles of the standard momentum strategy reflects momentum's top-bottom phenomenon that is discussed in Section 2.2.2: using the stocks in the top and bottom 20% of momentum's ranking results in a Sharpe ratio of 18.7% versus the Sharpe ratio of 29.3% when using deciles.

Table 2.B.1 also shows that the differences between the Sharpe ratios when using deciles generally remains if we use other percentiles. The conclusions in this chapter thus do not seem to depend on the percentile chosen. However, we note that Table 2.B.1 indicates that most strategies have the highest Sharpe ratio around the fifth percentile; the performance of these strategies could thus be improved by choosing this smaller percentile.

Table 2.B.1: **Sharpe ratios for different percentiles.** The annualized Sharpe ratio of the momentum strategies for different percentiles used to divide the stocks over the top and bottom deciles. See Section 2.5.2 for a brief description of the strategies corresponding to the different abbreviations.

	S	O	C	R	O-C	O-C*	R-O	R-C	R-O-C	R-O-C*
1%	43.1%	60.3%	55.2%	59.8%	73.5%	62.8%	68.3%	62.4%	78.2%	73.9%
2%	42.5%	64.1%	56.6%	72.3%	77.3%	69.4%	80.1%	82.5%	91.1%	93.4%
3%	36.4%	64.6%	60.0%	80.4%	81.5%	71.9%	82.3%	94.0%	94.6%	99.0%
4%	38.2%	62.6%	61.3%	86.6%	81.5%	75.1%	90.1%	95.2%	102.8%	106.5%
5%	37.8%	67.0%	63.9%	80.5%	86.0%	80.6%	86.8%	90.5%	101.6%	107.5%
6%	37.2%	68.8%	68.6%	78.4%	83.6%	79.7%	83.5%	86.8%	99.8%	109.6%
7%	34.8%	69.9%	67.4%	79.3%	83.6%	79.7%	85.3%	84.4%	101.5%	109.6%
8%	32.3%	64.9%	64.8%	77.5%	79.4%	78.9%	84.7%	85.2%	101.6%	109.9%
9%	30.0%	64.1%	64.6%	78.8%	77.7%	78.4%	85.1%	85.6%	101.8%	109.9%
10%	29.3%	63.6%	65.4%	77.9%	76.9%	77.6%	83.5%	86.0%	100.3%	106.6%
11%	26.7%	59.0%	65.7%	76.8%	72.6%	77.4%	80.8%	80.3%	97.6%	105.1%
12%	26.0%	58.8%	64.3%	75.7%	71.1%	76.0%	78.0%	81.5%	96.0%	103.5%
13%	25.1%	58.3%	65.1%	76.3%	70.6%	75.4%	78.0%	85.4%	95.3%	103.1%
14%	23.5%	54.4%	63.1%	72.4%	66.3%	74.1%	75.9%	82.8%	90.9%	99.9%
15%	22.2%	50.5%	61.8%	70.0%	62.5%	72.4%	70.7%	80.2%	86.5%	96.2%
16%	22.2%	49.9%	62.2%	66.2%	61.9%	71.2%	67.0%	77.4%	84.2%	96.0%
17%	21.1%	48.1%	61.5%	66.3%	60.2%	71.0%	66.6%	77.0%	82.4%	93.5%
18%	19.9%	44.2%	59.9%	68.2%	57.4%	70.0%	65.7%	76.1%	81.9%	93.4%
19%	19.0%	43.3%	61.0%	67.1%	56.8%	70.3%	63.9%	76.3%	80.4%	93.6%
20%	18.8%	42.0%	59.4%	66.2%	55.4%	69.2%	64.2%	74.8%	80.2%	93.4%

2.C Bootstrapping Sharpe ratios

In this section we present a bootstrap method to test whether a Sharpe ratio is significantly larger than zero as well as a similar bootstrap method to test whether a Sharpe ratio of a strategy is significantly higher than that of another strategy. We do so by applying the circular bootstrap method by Politis and Romano (1992). Following Ledoit and Wolf (2008), we do not use the Jobson and Korkie (1982) test with the Ledoit and Wolf (2003) adjustment for differences between the Sharpe-ratios as it is not valid for non-normal time series. Momentum returns are typically not normally distributed given their relatively extreme drawdowns. Given these return characteristics the standard bootstrap approach of sampling single observations in every bootstrap sample by Efron (1979) is also inappropriate, because the bootstrap methodology should take into account autocorrelations in the first and second moments. Among others, Barroso and Santa-Clara (2015) show that momentum returns have persistence in volatility.

Bootstrap method Several bootstrapping methods exist to take into account persistence in the first and second order moments of the data. Equivalent to the circular bootstrap are the moving block bootstraps by Liu and Singh (1992) and Kunsch (1989), apart from their small sample behavior. The circular bootstrap solves that observations in the beginning or at the end of a time series end up less frequently in the bootstrapped samples by having the last observations also in a continuing block with the first observations.[6]

Unlike Ledoit and Wolf (2008) we do not perform a two-sided test on the difference between two Sharpe ratios, but simply test how many of the bootstrapped Sharpe ratio differences are higher than that of the actual observed difference between the Sharpe ratios of the two strategies. The p-value of this test if a strategy's Sharpe ratio is higher than that of another strategy is one minus the portion of bootstrapped samples for which their bootstrapped difference is larger than their actual observed difference in Sharpe ratio. It is key to sample the Sharpe ratios under H_0 by demeaning the return series such that the Sharpe ratio of the series becomes zero. Testing if a Sharpe ratio is different from zero is equivalent; rather than comparing the actual difference between two ratios with the bootstrapped differences, the actual Sharpe ratio is compared to the bootstrapped Sharpe ratios. This method follows Ledoit and Wolf (2008) in that it does not sample from the distribution under H_0, but rather inverts a confidence interval; Ledoit and Wolf (2008) refer to paragraph 1.8 in Politis, P.Romano, and Wolf (1999) to do so.

The bootstrap algorithm we apply to find the p-value of the Sharpe ratio test of return series 1 r_1 being larger than zero or being higher than the Sharpe ratio of return series 2 r_2 is as follows:

1. Demean the return series r_1 such that its Sharpe ratio becomes zero in order to sample Sharpe ratios under H_0. If testing the difference between two Sharpe ratios also demean the return series r_2 such that its Sharpe ratio becomes zero.

[6] Politisand and P.Romano (1994) have come up with a circular stationary bootstrap with random block length to preserve the persistence in the bootstrapped samples. We do not see the advantages of this stationary bootstrap over the non-random block length bootstraps and choose the regular circular bootstrap, like Ledoit and Wolf (2008) use for their hypothesis testing on the difference between two Sharpe ratios.

2. Set $i = 0$, where i will count the number of times the bootstrapped Sharpe ratio of series 1 is larger than the Sharpe ratio for the actually observed data, or the number of times the difference between the bootstrapped Sharpe ratios is larger than the difference between the Sharpe ratios for the actually observed data.

3. Determine the optimal block length size b, by either the algorithm in Ledoit and Wolf (2008) or the corrected Politis and White (2004) optimal block length size in Patton, Politis, and White (2009).

4. Sample $l = \lceil \frac{T}{b} \rceil$ blocks of b random integers between 1 and T, where T equals the length of the original return series. Following the circular bootstrap by Politis and Romano (1992) a block that starts at the end of the series continues at the start of the series until its length equals the block size b. The last sampled block l might have more observations than required to have the series filled to length T; in that case the last observations of the last block l are thrown away.

5. Construct a single index series x that contains the subsequently sampled integers of the l blocks in the previous step.

6. Construct the adjusted return series r_1^* and that has at time t the sampled observations $r_1^* = r_1(x_t)$; if applicable also construct r_2^* in a similar way $r_2^* = r_2(x_t)$.

7. Calculate the Sharpe ratios of the constructed series r_1^* and, if applicable, r_2^*. If the Sharpe ratio of the series r_1^* is larger than the Sharpe ratio for the actually observed data or if the difference between the bootstrapped Sharpe ratios is larger than the difference between the Sharpe ratios for the actually observed data, then $i = i + 1$, else $i = i$.

8. Repeat steps 3 to 6 M times

9. The p-value of the bootstrap test then equals i, the number of times the bootstrapped Sharpe ratio of series 1 has been larger than the Sharpe ratio for the actually observed data, or the number of times the difference between the bootstrapped Sharpe ratios is larger than the difference between the Sharpe ratios for the actually observed data, divided by the number of bootstrap samples M.

Bootstrap settings In this study we construct 10000 bootstrap samples from the original observations; a Sharpe ratio is thus significantly higher at the 5 percent level if 9500 of the bootstrapped Sharpe ratio differences are smaller than the actual difference.

We choose a block length of 6 months as this is the optimal block length when we apply the algorithm by Ledoit and Wolf (2008) for block lengths of 1, 3, 6, 10 and 15 months. As such our methodology takes into account the optimal block length put forward by Politis and White (2004)[7] and corrects for the persistency in the first and second order moments of the raw and optimized momentum returns.

Bootstrap test and transaction costs Not only the statistical significance is of interest, but also if this significance would hold after transaction costs. We have therefore more or less reversed the bootstrap test on the Sharpe ratios being larger than zero. Given all bootstrapped Sharpe ratios for the original test on being different from zero we have solved for the transaction costs that would lower the actual Sharpe ratio by an amount such that a specific percentage, equal to the significance level, of the bootstrapped Sharpe ratio would be higher. We have presented the results of this test for a significance level of 5% in Table 2.7 in Section 2.6.

[7]Patton et al. (2009) have later corrected the formulas for the optimal block lengths

Chapter 3

Hedging the Time-Varying Exposures of Equity Momentum*

3.1 Introduction

Momentum is a strategy that ranks stocks on past performance, buys the past winners and sells the past losers[1]. Kothari and Shanken (1992) and Grundy and Martin (2001) show that the resulting momentum returns have time-varying exposures to the market and all three Fama and French (1992) and Fama and French (1993) equity risk factors respectively. In this study we show that an ex-ante feasible hedging strategy accounting for these time-varying exposures outperforms the standard momentum strategy. The hedged returns are larger, less risky, more stable over time and vary less over market and economic conditions. The results weaken the crash risk explanation of momentum returns in Daniel and Moskowitz (2013) and Daniel et al. (2012) as the average of the twelve worst one-month losses is reduced by 60 percent. We show that the way of hedging is important. Taking into account the momentum's conditional factor exposures outperforms both hedging based on the momentum's unconditional factor exposures and hedging based on the individual loadings of the constituents of the momentum strategy. We also find that it is crucial to account for all three equity risk factors, not just the market which has been the focus of recent studies.

*This chapter is based on Martens and Van Oord (2014).

[1]See Jegadeesh and Titman (1993) and Jegadeesh and Titman (2001), among others.

3.1.1 Time variation in the returns and exposures of the momentum strategy

The average factor exposures of momentum to the equity risk factors are all negative. For example, the full-sample market beta is -0.25. These factor exposures, however, vary substantially over time. If the market returns are positive (negative) during the ranking period, the market beta of the zero-investment momentum strategy is 0.45 (-0.62). When subsequently the market reverses in the investment period the strategy will lose due to its adverse loading on the market factor. Moreover, if all three equity risk factors move in the opposite direction in the investment period compared to the ranking period, momentum on average loses 3.6 percent per month. On the contrary, momentum earns 3.5 percent when the factors move in the same direction. These time-varying exposures of momentum also result in variation of the returns over the business cycle: during the second half of recessions momentum loses 2.1 percent per month on average, while it earns 1.5 percent per month in the second half of expansions.

The reason for the time variation in the exposures is as follows: Consider the simplified case that stock returns are governed by their market beta, the market return and a residual return. If during the ranking period the market return is positive, the momentum strategy tends to buy high beta stocks and sell low beta stocks. The resulting positive beta exposure then positively contributes to the momentum return if the market return is positive again in the investment period. In contrast, if during the investment period the market return is negative, this positive beta exposure negatively contributes to the return. In short, when the market factor moves in the same direction in both ranking and investment period, it is beneficial for the momentum return. If, however, the market factor moves in the opposite direction in the investment period, it is detrimental for the momentum returns. The same arguments hold for the size and book-to-market factors.

3.1.2 Hedging the time-varying exposures of momentum

We investigate three different ways to hedge the time-varying risk exposures of momentum. First, we use the common approach of regressing historical momentum returns for a specific window on the three Fama and French (1993) equity risk factors. This mostly corrects for the long-term negative factor loadings of momentum. This is a rather naive strategy as it assumes the current momentum portfolio has the same factor loadings as the past momentum portfolios, while the constituents of the strategy change rapidly over time. As a consequence, this hedging method still has strong time-varying factor exposures. Therefore we consider two hedging methods that do take into account the current factor exposures of momentum. First, we use individual historical betas of the constituents of the momentum portfolio to estimate the exposures. Second, we capture the time-varying patterns in the momentum exposures by regressing momentum returns on the equity risk factors conditional on the factor returns in the ranking period of the momentum strategy, as proposed by Grundy and Martin (2001).

Hedging based on either individual stock betas or the conditional factor model reduces the volatility of the momentum strategy by about 25 percent. When all three equity risk factor returns show a reversal from the ranking to the investment period, the hedge based on individual betas gains 0.6 percent, and the hedge based on conditional momentum loadings loses -0.4 percent, compared to -3.6 percent for the unhedged momentum strategy.

The hedge based on individual betas has an annual Sharpe ratio of 0.47 and the conditionally hedged returns have a Sharpe ratio of 0.71 compared to a Sharpe ratio of 0.36 for unhedged momentum. These Sharpe ratios further improve to 0.63 and 0.75, respectively, if we only take offsetting positions when the estimated factor exposures are negative. These asymmetric hedges profit from the average positive factor returns and reduce the overall negative loadings on the equity risk factors.

3.1.3 Crash risk as an explanation for momentum profits

Daniel and Moskowitz (2013) and Daniel et al. (2012) show that the momentum strategy has many large monthly losses. As such they hypothesize that the momentum profits are a compensation for crash risk. For our data we find that the average loss equals 36.8 percent for the twelve months where momentum loses more than 20 percent. The conditional hedge reduces the losses in these months to an average loss of 14.8 percent. We achieve this by understanding the time-varying exposures of momentum and only using ex-ante available information. Hence these results substantially weaken the crash risk explanation of momentum profits.

3.1.4 Stability of hedged momentum returns over the business cycle

Hedging also stabilizes the momentum returns over the different stages of the business cycle. Similar to Chordia and Shivakumar (2002) we find that momentum returns are on average higher periods of expansion (0.90 percent) than in periods of recession (-0.50 percent). For the conditionally hedged returns, however, we find average monthly returns of 1.13 percent and 0.46 percent in periods of expansions and recessions, respectively. Especially in the second half of recession periods, raw momentum suffers with an average return of -2.09 percent per month. We find that this is mainly due to reversals in the market, size and value-growth factors. Hedging reduces the adverse conditional loadings on the equity risk factors and thereby improves performance in the latter stages of recessions, from -2.09 percent to 0.05 percent per month.

Hedging also improves momentum returns after down markets. Cooper et al. (2004) find substantial momentum losses after the 3-year market return is negative. The conditional hedge improves momentum returns in down markets from -1.78 percent per month for the unhedged strategy to 0.27 percent. Hence we attribute the differences between down and up markets partly to time-varying risk exposures rather than only to overreaction theories as Cooper et al. (2004) do.

3.1.5 Systematic under- and overestimation historical betas

The common approach to estimate a portfolio's factor exposures by the historical betas of its constituents leads to systematic over- and underestimation of the momentum exposures[2]. After positive factor returns in the formation period momentum selects stocks with high loadings on the winning factors. It turns out that these loadings are overestimated. Hence the ex-ante hedge will ex-post turn out to be an over-hedge and results in negative loadings on the equity risk factors. Following negative factor returns the exposures of stocks with high loadings on a particular factor are underestimated, while momentum is short in these stocks. It follows that the ex-ante hedge ratios will ex-post turn out to be insufficient, leaving negative loadings on the equity risk factors. Thus both after positive and after negative factor returns hedging based on individual stock betas

[2]Boguth, Carlson, Fisher, and Simutin (2011) find for ex-post betas of the momentum strategy a negative correlation between beta biases and market returns. They, however, focus on estimating unbiased alphas, not hedging.

will leave on average negative exposures to the factors.

The remainder of this chapter is organized as follows. Section 3.2 describes the data. Section 3.3 explains how we measure the ex-post exposures of momentum to the equity risk factors and how the hedging methods work. Section 3.4 presents the key results. Section 3.5 describes the systematic under- and overestimation of the momentum exposures when using the commonly used individual historical estimated betas. Finally, Section 3.6 will conclude.

3.2 Data

Total monthly returns of all NYSE and AMEX common stocks in the CRSP database are used[3], covering the period 1926 to 2011. Stocks that have a stock price lower than \$1 at the beginning of the investment period are excluded as also done by Amihud (2002) and Pastor and Stambaugh (2003) for \$5 stocks. Following Grundy and Martin (2001) we use a six-month ranking period, wait for one month to avoid return reversals[4], and then go long in the 10 percent winners and short in the 10 percent losers for one month assigning equal weight to each individual stock. We rule out those stocks that do not exist in the whole six-month ranking period. With these criteria the first ranking period is from January 1926 to June 1926 to invest in August 1926, and there are 44 stocks in each decile at that time. The final ranking period is from April 2011 to October 2011 to invest in December 2011, and by then there are 146 stocks in each decile.

Table 3.1 column 1 shows the sample characteristics of the standard momentum strategy from February 1932 to December 2011. We leave out the data before February 1932 as for the hedging strategies we require a burn-in period to estimate the first individual factor loadings and the conditional model presented in section 3.3[5]. The average return to the momentum strategy is 0.67 percent per month with a volatility of 6.5 percent per month. The worst month is April 2009 with a return of -60.5 percent, and the best month is

[3]Common AMEX and NYSE stocks with an exchange code equal to 1 or 2 and a share code equal to 10 or 11 during the investment month are eligible for the momentum strategy.

[4]See e.g. Jegadeesh and Titman (1993), Narasimhan Jegadeesh (1995), Lehmann (1990) and Lo and MacKinlay (1990).

[5]Data on the Fama and French (1993) equity risk factors start in July 1926. Our conditional regressions need these equity risk factor returns from $t-7$ to $t-2$ besides the skip month $t-1$. Subsequently we require at least 60 months for the estimation of the conditional regression presented in Section 3.3.

March 1938 with a return of 23.0 percent. The final rows of Table 3.1 also show that momentum on average has significant negative loadings on the market (-0.25), size (-0.25) and HML (-0.42) factors.

Table 3.1: **Sample statistics momentum strategy.** Monthly sample statistics of investing according to the regular momentum strategy (Unhedged), i.e. sort stocks based on their performance from $t-7$ to $t-2$, skip month $t-1$, and invest at time t in the top 10% of the ranked stocks and short the bottom 10%. The next three columns show the same results for the hedged momentum returns, where hedging takes place by taking positions in the equity risk factor portfolios offsetting predicted exposures to these equity risk factors. These predicted factor exposures are based on using past unconditional loadings on the three equity risk factors, over the most recent 120 months (Hedge 1); using the average of the individual loadings of the constituents of the momentum strategy using a maximum estimation window of 120 (Hedge 2); or using past conditional loadings using Equation 3.1 (Hedge 3) estimated over the most recent 120 months. The final four rows show the results of regressing the (hedged) momentum returns on the equity risk factors over the full sample. ***, ** and * denote significance at the 1%, 5% and 10% significance levels, respectively. The sample period covers Febuary 1932 to December 2011 (959months).

	Unhedged	Hedge 1	Hedge 2	Hedge 3
Average return	0.67%	1.08%	0.68%	1.02%
Standard deviation	6.5%	6.1%	5.0%	5.0%
Minimum	-60.5%	-56.5%	-42.9%	-42.3%
Maximum	2.03%	22.1%	17.0%	23.6%
Alpha	1.10%***	1.06%***	1.04%***	1.07%***
Market beta	-0.25***	-0.06*	-0.08***	-0.02
Size beta	-0.25***	0.05	-0.44***	-0.10**
HML beta	-0.42***	0.14***	-0.42***	0.00
R-squared	15.5%	0.8%	23.2%	0.6%

3.3 Methods

Let $r_{i,t}$ be the return on stock i in month t. At the end of each month $t-1$ all NYSE and AMEX stocks are ranked on the basis of their cumulative returns over the formation months $t-7$ to $t-2$. Note that month $t-1$ itself is known as the skip month, and is explicitly not used for the ranking to avoid mean-reversion in returns. Subsequently equally weighted deciles are formed. The return differential between the winners' portfolio (top ranked 10 percent stocks) and the losers' portfolio is denoted by $r_{WL,t}$. This is the momentum return in month t.

3.3.1 Ex-post evaluation of the time-varying risks of the momentum strategy

Grundy and Martin (2001) estimate the conditional model in Equation 3.1, since regressing the momentum returns on the equity risk factors ignores the time-variation in the factor exposures of momentum:

$$\begin{aligned} r_{WL,t} = \alpha \quad &+ \quad \left(\beta_{\text{down}} D_t^{\text{MKT,down}} + \beta_{\text{flat}} D_t^{\text{MKT,flat}} + \beta_{\text{up}} D_t^{\text{MKT,up}}\right) r_{\text{MKT},t} \\ &+ \quad \left(s_{\text{down}} D_t^{\text{SMB,down}} + s_{\text{flat}} D_t^{\text{SMB,flat}} + s_{\text{up}} D_t^{\text{SMB,up}}\right) r_{\text{SMB},t} \\ &+ \quad \left(h_{\text{down}} D_t^{\text{HML,down}} + h_{\text{flat}} D_t^{\text{HML,flat}} + h_{\text{up}} D_t^{\text{HML,up}}\right) r_{\text{HML},t} + e_t \end{aligned} \tag{3.1}$$

where D_t are the dummies to denote whether the market (MKT), the size factor (SMB) or the book-to-market factor (HML) during the ranking period ($t-7$ to $t-2$) went down by more than one standard deviation relative to the mean (down), up by more than one standard deviation relative to their mean (up) or stayed within one standard deviation from their mean (flat).

Consider for example the exposure to the market factor. If the market is up during the ranking period only $D_t^{\text{MKT,up}}$,up equals one. In that case the winners (losers) portfolio will have more high (low) beta stocks and momentum will have a positive beta exposure. Hence we expect β_{up} to be larger than zero. For similar reasons we expect β_{down} to be smaller than zero. Obviously similar arguments apply to the SMB and HML exposures. It is these time-varying factor exposures that we would like to reduce with hedging. Hence we evaluate the adequacy of hedging strategies by checking if the estimated conditional betas in Equation 3.1 get closer to zero for the hedged return series. Using adequately hedged momentum returns $r_{WL,t+1}^{\text{Hedge } j}$ on the left-hand-side of Equation 3.1 should reduce the regression R-squared of this conditional model.

3.3.2 Hedging strategies for the standard momentum returns

If we were to know the momentum factor exposures at the start of the investment month we could hedge the momentum return by taking opposite positions in the market, the size and the value-growth factor[6]. We look at three different ways to hedge returns in this way. They differ from each other in the way they estimate the factor exposures of momentum. The hedged momentum returns become:

$$r_{WL,t+1}^{\text{Hedge } i} = r_{WL,t+1} - \widehat{\beta_{W-L,t}^{\text{Hedge } j}} r_{\text{MKT},t+1} - \widehat{s_{W-L,t}^{\text{Hedge } j}} r_{\text{SMB},t+1} - \widehat{h_{W-L,t}^{\text{Hedge } j}} r_{\text{HML},t+1} \tag{3.2}$$

Using unconditional loadings of momentum returns (hedge 1)

The simplest approach is to regress the momentum returns in months tT to $t1$ on the unconditional factor returns to find the loadings for month t,

$$r_{WL,t^*} = \alpha_t + \beta_t r_{\text{MKT},t^*} - s_t r_{\text{SMB},t^*} - h_t r_{\text{HML},t^*} + e_t \qquad t^* = t-T, ..., t-1 \tag{3.3}$$

The estimated betas can be used directly in Equation 3.2 to create a hedged version of momentum.

Using individual factor loadings of the current constituents (hedge 2)

The composition of the momentum portfolio changes over time, while we are interested in the loadings of the current portfolio. Hence a more logical approach to determine a portfolio's current exposures is to first compute the historical factor loadings of the individual stocks that make up the current portfolio and then compute the factor loadings of the portfolio. Hence each month t we estimate for each stock in the momentum portfolio

$$r_{i,t^*} = \alpha_t + \beta_{i,t} r_{\text{MKT},t^*} - s_{i,t} r_{\text{SMB},t^*} - h_{i,t} r_{\text{HML},t^*} + e_{i,t} \qquad t^* = t-T, ..., t-1 \tag{3.4}$$

for months tT to $t1$ (or less history when the stock does not exist for T months). Then we compute the factor loadings of the momentum portfolio for month t as follows,

$$\widehat{\beta_t^{\text{Hedge 2}}} = \sum_{i=1}^{N_t} w_{i,t} \widehat{\beta_{i,t}} \quad \widehat{s_t^{\text{Hedge 2}}} = \sum_{i=1}^{N_t} w_{i,t} \widehat{s_{i,t}} \quad \widehat{h_t^{\text{Hedge 2}}} = \sum_{i=1}^{N_t} w_{i,t} \widehat{h_{i,t}} \tag{3.5}$$

where $w_{i,t}$ is the weight of stock i in the momentum portfolio in month t. In our case with equal weighting for the top and bottom decile the weights will be $10/N_t$ or $-10/N_t$

[6]We assume it is possible to do this. Daniel and Titman (1997) point to the efficacy of the approach of trading offsetting portfolios of long-short positions that mimic the Fama and French (1993) factors. In practice it could be difficult and costly to create short positions. As future research it will be interesting to test hedging using recently launched futures on value, growth and small cap indices, including transaction costs.

respectively for the constituents of these deciles with N_t the number of stocks at time t; stocks in deciles two to nine get a weight of zero.

Using conditional loadings of momentum returns (hedge 3)

Rather than estimating the individual exposures of the current constituents of the momentum strategy we directly take offsetting positions according to the current conditional loadings. We only use in-sample data that is known prior to the investment month t to estimate these conditional loadings. For that purpose we estimate Equation 3.1 for the past T months up to t to determine the conditional loadings. Next we use the factor returns in the formation period to establish for each factor whether we are in the up, down or flat state and use the corresponding estimate for the offsetting positions. The advantage of this approach over the individual hedge is that we do not have to assume that individual stocks have constant loadings for a longer historical period. The disadvantage is that we assume to know the functional form and the parameters of the conditional model[7].

3.4 Results

3.4.1 Sample characteristics of hedged momentum returns

Table 3.1 shows the sample characteristics of momentum and the three versions of hedged momentum. We use T equal to 120 months for all hedge methods. In Section 3.5 we show results for smaller T. First, as expected hedging reduces the volatility. The hedge based on individual stock betas and the conditional hedge both have a standard deviation of 5.0 percent per annum, compared to 6.5 percent for unhedged momentum. The reduction in risk is also visible in the minimum return, with the worst monthly loss reduced from 60.5 percent to less than 43 percent. In addition the hedge based on the conditional model in Equation 3.1 provides a higher average monthly return of 1.02 percent, compared to 0.67 percent for both unhedged momentum and for hedged momentum based on individual loadings. The unconditional hedge also increases the return to 1.09 percent but reduces the volatility only to 6.1 percent and the max loss is still -56.5 percent. We will see in the next section that the unconditional hedge is not effective in reducing the time-varying

[7] Jr. and Prinsky (2007) and Blitz et al. (2011) follow a completely different approach that aims at avoiding the conditional factor loadings of momentum altogether. In particular the residual momentum strategy looks at the past residual returns to form decile portfolios, based on the residuals from Equation 3.4. Here we take the positions of the total return momentum strategy as given

exposures to the risk factors.

The final rows of Table 3.1 explain the return differences of the hedged momentum strategies. Both the unconditional and conditional hedge on average increase the loadings on the market, size and HML, three factors known to have on average a positive return. For our sample the market excess return is 3.9 percent per annum, and the returns for size and HML are respectively 2.9 and 4.6 percent per annum. In contrast the hedge based on individual stock loadings still has negative loadings on all three factors. We investigate this in more detail in Section 3.5.

3.4.2 Ex-post time-varying exposures

We started the hedging exercise with the objective to reduce the time-varying exposures of momentum. These exposures are shown in column 1 of Table 2. The results are comparable to those on page 47 in Grundy and Martin (2001). In particular we see that momentum has a negative market beta (0.62) when the market return in the ranking period is negative and a positive beta (0.45) after the market return is positive during the ranking period. Similarly the momentum strategy has higher (lower) loadings to the size and HML factors when these factor returns have been positive (negative) during the ranking period. The conditional model explains 51.3 percent of the variation in the momentum returns. This is in contrast with the unconditional equity risk factor model that explains only 15.5 percent, see column 1 final row in Table 3.1.

All three hedging methods reduce the impact of the time-varying exposures to the equity risk factors. The conditional hedge is the best method to reduce the time-varying exposures. After hedging the regression R-squared is reduced to just 5.2

Column 2 in Table 3.1 shows the ex-post conditional factor exposures of the hedged momentum returns, where the hedge parameters have been computed using the unconditional model in Equation 3.3. The unconditional hedge is the least effective in reducing the time-varying risk exposures of momentum. Effectively this method most of the time only adds the market, size and HML because the average loading of momentum to these factors is negative (see Table 3.1 column 1). This is also visible from column 2 in Table 3.3 where for all three factors the loadings are higher in all three states. For example the exposure to the market, after the market went up, has increased from 0.45 to 0.66. And the exposure, after the market went down, has increased from 0.62 to 0.38. Both

Table 3.1: **Ex-post risk exposures of (hedged) momentum strategies.** This table shows the estimation results of the conditional equity three risk factor model in Equation 3.1, estimated for February 1932 to December 2011 (959 monthly observations). The first column shows the results for the original momentum strategy that ranks stocks based on their returns in the past 6 months, skips one month and then invests for one month. The next three columns show the same results for the hedged momentum returns, where hedging takes place by taking positions in the equity risk factor portfolios offsetting predicted exposures to these risk factors. These predicted factor exposures are based on using past unconditional loadings on the three equity risk factors, over the most recent 120 months (column 2); using the average of the individual loadings of the constituents of the momentum strategy using a maximum estimation window of 120 (column 3); or using past conditional loadings using Equation 3.1 (final column) estimated over the most recent 120 months.

		Unhedged	Hedge 1	Hedge 2	Hedge 3
Intercept		0.0095	0.0092	0.0103	0.0101
		(6.42)	(5.79)	(7.42)	(6.35)
MKT	Up	0.445	0.661	-0.075	0.108
		(5.03)	(6.94)	(-0.90)	(1.13)
	Flat	0.067	0.195	0.058	0.075
		(1.64)	(4.42)	(1.5)	(1.7)
	Down	-0.621	-0.384	-0.243	-0.094
		(-12.43)	(-7.13)	(-5.17)	(-1.74)
SMB	Up	0.537	0.837	-0.462	0.212
		(5.47)	(7.92)	(-5.02)	(2.01)
	Flat	-0.58	-0.254	-0.645	-0.24
		(-9.14)	(-3.72)	(-10.83)	(-3.51)
	Down	-0.422	-0.17	-0.023	-0.181
		(-4.62)	(-1.73)	(-0.27)	(-1.83)
HML	Up	0.687	1.249	-0.208	0.003
		(6.66)	(11.24)	(-2.14)	(0.03)
	Flat	-0.129	0.376	-0.344	0.211
		(-2.18)	(5.87)	(-6.16)	(3.29)
	Down	-1.048	-0.468	-0.534	-0.201
		(-13.19)	(-5.47)	(-7.16)	(-2.35)
R-squared		51.3%	37.8%	27.9%	5.2%

increases are not far away from the average (negative) 0.25 loading on the market for momentum (see Table 3.1 Column 1). Hence the unconditional hedge is merely adding betas to offset the average negative betas, but it is not providing a solution for the time-varying exposures.

Column 3 in Table 3.1 shows the results for the hedge based on individual stock loadings, see Equations 3.4 and 3.5. The results show that this hedging approach partially succeeds in reducing the time-varying risk exposures. For example the down market beta is still 0.24, and the down HML beta is still 0.53. Also in some instances the hedge ratio has been overestimated, as e.g. the up market beta is 0.08 and the up size beta is even 0.46. The regression R-squared of the conditional model is still 27.9 percent. Hence this hedging approach improves over not hedging and the unconditional hedge, but there is room for further improvement. Hence the hedge based on individual loadings does not work that well, and we believe the main cause is estimation error. The estimated loadings of the constituents of the momentum portfolio are, for example, biased upward after positive factor returns during the formation period. We investigate this in more detail in Section 3.5 where we also motivate the choice for the 120 month window.

Column 4 in Table 3.1 shows the ex-post conditional factor exposures of the hedged momentum returns, where the hedge parameters have been computed by applying the conditional regression in Equation 3.1 to the past 120 momentum returns prior to investing. The results show that this approach succeeds to a large extent in reducing the time-varying risk exposures. All market betas are no longer significantly different from zero at the 5 percent significance level. For the HML factor the loading in the up state is no longer significant at the 5 percent level. This hedging method is the least successful for the size factor, with the up state beta lowered to a still significant 0.21, a flat beta of 0.24 and the down size beta still 0.18. The regression R-squared has been substantially reduced from 51 percent to just 5 percent.

We therefore conclude on this criterion that the conditional hedge has the best hedging performance than the hedge based on individual stock betas. Apparently imposing a prior structure on the time-varying factor loadings improves their estimation substantially.

3.4.3 Momentum returns over factor return reversals

Table 3.2 shows the results when sorting the raw and hedged momentum returns based on the factor returns during the ranking and investment period. This exercise can be viewed as the non-parametric alternative to the regression-based analysis in the previous section to analyze the consequences of the time-varying exposures to the factor loadings. For each of the three equity risk factors we consider the case where the factor return moves in the same direction in ranking and investment period (same) or in the opposite direction (reversal). This gives 8 possibilities. The raw momentum strategy earns on average 3.5 percent per month when all factor returns have the same sign in the ranking and investment period, whereas this is -3.6 percent when all factor returns have opposite signs in the ranking and investment period, a difference of 7.1 percent.

Table 3.2: **Momentum returns conditional upon factor returns.** This table shows the average monthly (hedged) momentum returns in 8 different states for February 1932December 2011 (959 monthly observations). If the factor returns have the same sign in the ranking period ($t-7$ to $t-2$) and investment period (t) it is labeled Same. If the factor moves in the opposite direction in ranking and investment period it is labeled Reversal. For the three equity risk factors we thus get 8 possibilities. The second column (Unhedged return momentum) shows the results for the original momentum strategy that ranks stocks based on their returns in the past 6 months, skips one month and then invests for one month. The next three columns show the same results for the hedged momentum returns, where hedging takes place by taking positions in the equity risk factor portfolios offsetting predicted exposures to these equity risk factors. These predicted factor exposures are based on using past unconditional loadings on the three equity risk factors, over the most recent 120 months (column 3); using the average of the individual loadings of the constituents of the momentum strategy using a maximum estimation window of 120 (column 4); or using past conditional loadings using Equation 3.1 (final column) estimated over the most recent 120 months.

MKT	SMB	HML	Obs	Unhedged	Hedge 1	Hedge 2	Hedge 3
Same	Same	Same	188	3.5%	4.1%	0.6%	2.0%
Same	Same	Reversal	120	1.7%	2.5%	0.9%	1.4%
Same	Reversal	Same	119	1.6%	2.2%	0.6%	1.4%
Same	Reversal	Reversal	117	-0.3%	-0.1%	1.0%	0.6%
Reversal	Same	Same	96	1.8%	2.4%	0.5%	1.3%
Reversal	Same	Reversal	78	-1.0%	-1.2%	0.5%	0.1%
Reversal	Reversal	Same	116	-0.2%	0.2%	0.7%	1.1%
Reversal	Reversal	Reversal	125	-3.6%	-3.5%	0.6%	-0.4%
		All	959	0.67%	1.08%	0.68%	1.02%

Hedge 2 based on individual stock betas reduces the impact of the time-varying factor exposures. The range is reduced to 0.5 percent (best 1.0 percent, worst 0.5 percent). The conditional hedge 3 is worse on this criterion with the range being reduced to 2.4 percent (best 2.0 percent, worst -0.4 percent). The unconditional hedge 1 is clearly the worst. The range between the best case at 4.1 percent and the worst case at -3.5 percent is 7.6

percent, even larger than that of unhedged momentum. This illustrates once again that the unconditional hedge is not effective in estimating the factor exposures of momentum accurately at each moment in time.

3.4.4 Momentum crashes

Daniel and Moskowitz (2013) and Daniel et al. (2012) show that momentum crashes occur following market declines, when market volatility is high, and contemporaneous with market rebounds. We show that this is largely due to the time-varying exposures in the factor loadings of momentum. When the market crashes, low beta stocks will lose less than high beta stocks. Hence momentum will load on low beta stocks and short high beta stocks, leading to a negative beta for the strategy. If subsequently the market rebounds, the negative beta exposure leads to a loss.

Hedging this negative beta should reduce these losses; the better the ex-ante hedge the larger this loss reduction. Table 3.3 shows that we manage to reduce the momentum crashes quite substantially. Following Daniel and Moskowitz (2013) we look at 12 drawdowns, months in which WML loses 20 percent or more. The average loss of WML in our case is -36.8 percent.

This reduces to -23.1 percent after the unconditional hedge, because in the market down state the beta is -0.38 instead of -0.62 (Table 3.1 row labeled MKT down), reducing the loss due to a market rebound. The hedge based on individual stock betas reduces the average loss in the worst unhedged momentum return months to -24.2 percent. It also has a less negative market beta in market down states. Finally the conditional hedge manages to bring down the average loss to just -14.8 percent, a reduction of 60 percent compared to the -36.8 percent for unhedged momentum. At the same time this hedging method provides an even better average monthly return, up from 0.67 percent to 1.02 percent. Hence this makes a crash risk based explanation of momentum less likely[8].

[8]Of course after hedging other months than the original twelve months could have greater losses. For the conditional hedge the new twelve worst months have an average loss of 21.9 percent, substantially lower than the worst twelve months of unhedged momentum at -36.8 percent. Also these losses will not necessarily take place after market downturns and subsequent recoveries (see also 3.4.5 and 3.4.6), or when market volatility is high.

Table 3.3: **Twelve largest monthly WML losses and these losses after hedging.** The table reports all momentum (unhedged column 2) losses in excess of 20 percent in a single month from February 1932 to December 2011. The final three columns show the losses in the same months for the hedged momentum returns, where hedging takes place by taking positions in the equity risk factor portfolios offsetting predicted exposures to these equity risk factors. These predicted factor exposures are based on using past unconditional loadings on the three equity risk factors, over the most recent 120 months (column 3); using the average of the individual loadings of the constituents of the momentum strategy using a maximum estimation window of 120 (column 4); or using past conditional loadings using Equation 3.1 (final column) estimated over the most recent 120 months.

Rank	Month	Unhedged	Hedge 1	Hedge 2	Hedge 3
1	Apr-09	-60.5%	-56.5%	-42.9%	-42.3%
2	Sep-39	-57.7%	-38.6%	-37.9%	-41.2%
3	Jul-32	-50.9%	-24.0%	-27.2%	1.6%
4	Apr-33	-44.6%	-25.1%	-4.3%	14.4%
5	Aug-32	-35.8%	9.6%	-41.9%	-5.2%
6	Jan-01	-35.4%	-37.9%	-33.3%	-29.6%
7	Jan-75	-31.4%	-24.3%	-26.0%	-13%
8	May-33	-28.9%	-5.4%	-11.5%	-20.3%
9	Nov-02	-27.9%	-26.9%	-22.7%	-23.5%
10	Jan-34	-25.6%	-8.2%	-14.7%	-3.0%
11	Sep-70	-21.6%	-24.3%	-3.2%	-4.5%
12	Jan-74	-20.8%	-15.3%	-24.4%	-11.5%
	Average	-36.8%	-23.1%	-24.2%	-14.8%

3.4.5 Stability of momentum returns over the business cycle

Momentum returns vary over the different stages of the business cycle; hedging the time-varying exposures of momentum reduces this variation. The National Bureau of Economic Research (NBER) defines expansions and recessions for the US economy, including specific dates for troughs and peaks. Several studies (see e.g. MichaelDeStefano (2004)) use this information to split the business cycle into four stages: Stage 1 starts in the month after the trough (early expansion) with stage 2 ending at the peak (end of expansion). Stage 3 starts in the month after the peak (first part of recession period) and stage 4 ends with the trough (end of the recession). Table 3.4 shows the average (hedged) momentum returns during the various stages of the business cycle.

Momentum performs below par in stage 1 (early expansion) with an average return of

Table 3.4: **(Hedged) momentum returns over the business cycle.** This table shows the average monthly (hedged) momentum returns from February 1932 to December 2011 divided over 4 stages of the business cycle between the peaks and troughs of the US economy as established by the NBER. Stage 1 starts in the month after the trough with stage 2 ending in the month of the peak, with months being equally divided between stages 1 and 2. Similarly stage 3 starts in the month after the peak with stage 4 ending in the month with the trough, with months being equally divided between stages 3 and 4. Hence stages 1 and 2 reflect the different stages of the expansion, whereas stages 3 and 4 are the different stages of the recession. The second column (Unhedged) shows the results for the original momentum strategy that ranks stocks based on their returns in the past 6 months, skips one month and then invests for one month. The next three columns show the same results for the hedged momentum returns, where hedging takes place by taking positions in the equity risk factor portfolios offsetting predicted exposures to these equity risk factors. These predicted factor exposures are based on using past unconditional loadings on the three equity risk factors, over the most recent 120 months (column 3); using the average of the individual loadings of the constituents of the momentum strategy using a maximum estimation window of 120 (column 4); or using past conditional loadings using Equation 3.1 (final column) estimated over the most recent 120 months.

		Obs	Unhedged	Hedge 1	Hedge 2	Hedge 3
Expansion	Stage 1	418	0.32%	0.75%	0.58%	0.83%
	Stage 2	384	1.53%	2.02%	1.09%	1.46%
Recession	Stage 3	73	1.33%	1.11%	0.74%	0.94%
	Stage 4	84	-2.09%	-1.55%	-0.77%	0.05%
	All	959	0.67%	1.08%	0.68%	1.02%

0.32 percent. In the second part of the expansion (stage 2) momentum performs better with an average monthly return of 1.53 percent. The first part of recession periods (stage 3) is also good for momentum with an average return of 1.33 percent per month. In contrast, in the latter stage of the recessions momentum is generating losses with an average monthly return of -2.09 percent. Overall this means that momentum on average loses during recession periods (on average -0.50 percent per month), as also reported by Chordia and Shivakumar (2002). Our explanation for this is that there are more market return reversals in stage 4 that are bad for momentum and these market returns in reversal months are usually large; stage 4 typically contains the last part of recessions where market returns are positive (and large) and lead the recovery in the coming period of expansion. Hence momentum returns are unstable over the business cycle. Both the hedge based on individual betas (hedge 2) and the conditional hedge (hedge 3) improve the stability of returns over the business cycle, with the conditional hedge even posting a tiny positive return of 0.05 percent during recession. This last result again provides evidence against a crash risk explanation for momentum profits. The hedge based on the unconditional loadings of momentum is less good in stabilizing the returns over the business cycle, still posting a loss of 1.55 percent in stage 4.

3.4.6 Stability of momentum returns after up and down markets

Cooper et al. (2004) show that momentum profits depend on the state of the market. In particular momentum is profitable after the market has been up in the past years and loses money when the market has been down in the past years. For our sample period we also find that when the market return has been positive in the past 3 years momentum makes 1.23 percent per month, whereas momentum loses 1.78 percent per month when the market return has been negative in the past 3 years, see Table 3.5.Cooper et al. (2004) say that this is consistent with the overreaction models of Daniel, Hirshleifer, and Subrahmanyam (1998) and Hong and Stein (1999).

Results are different when hedging momentum returns for time-varying exposures to the equity risk factors. Again focusing on the most successful method, the conditional hedge, we see that after up markets the average return is 1.19 percent, whereas after down markets it is 0.27 percent, in sharp contrast with unhedged momentum returns. Hence we attribute Cooper et al. (2004) results partly to time-varying risk exposures rather than only to overreaction models. The small positive returns at 0.27 percent after 3-year market losses also contradict the crash risk explanation of momentum returns after large

Table 3.5: **(Hedged) momentum returns after up and down markets.** This table shows the average monthly (hedged) momentum returns from February 1932 to December 2011 split into up and down markets. The market is up if its return is positive in the past 36 months. The market is down if its return is negative in the past 36 months. The second column (Unhedged) shows the results for the original momentum strategy that ranks stocks based on their returns in the past 6 months, skips one month and then invests for one month. The next three columns show the same results for the hedged momentum returns, where hedging takes place by taking positions in the equity risk factor portfolios offsetting predicted exposures to these equity risk factors. These predicted factor exposures are based on using past unconditional loadings on the three equity risk factors, over the most recent 120 months (column 3); using the average of the individual loadings of the constituents of the momentum strategy using a maximum estimation window of 120 (column 4); or using past conditional loadings using Equation 3.1 (final column) estimated over the most recent 120 months.

	Obs	Unhedged	Hedge 1	Hedge 2	Hedge 3
Up	782	1.23%	1.53%	0.99%	1.19%
Down	177	-1.78%	-0.90%	-0.72%	0.27%
All	959	0.67%	1.08%	0.68%	1.02%

market downturns.

The unconditional hedge improves both the Up and Down states. Most of the time it simply adds market beta, which is beneficial both in the Up and in the Down state, because irrespective of these states the market more often goes up than down. Hence the unconditional hedge profits from its average positive loading on the market factor.

3.4.7 Only hedging negative loadings on the equity risk factors

The momentum strategy loads positively on the equity risk factors when the factor returns are positive during the formation period, and negatively when the returns are negative during the formation period. From a return perspective it should actually be beneficial to have a positive exposure on average, because the equity risk factors each have positive returns on average. On the other hand we have seen that loading negatively on factors after they posted a negative return during the formation period, will be detrimental for returns. Hence hedging should improve momentum returns in those cases.

We apply an asymmetric hedge to illustrate this point: We only take offsetting positions

Table 3.6: **Only hedging negative loadings on equity risk factors.** This table shows the average monthly (hedged) momentum returns, standard deviation and annualized Sharpe ratios from February 1932 to December 2011 when hedging the exposures all the time or when hedging only negative estimated loadings on the equity risk factors. The first column (Unhedged) shows the results for the original momentum strategy that ranks stocks based on their returns in the past 6 months, skips one month and then invests for one month. The next three columns show the same results for the hedged momentum returns, where hedging takes place by taking positions in the equity risk factor portfolios offsetting predicted exposures to these equity risk factors. These predicted factor exposures are based on using past unconditional loadings on the three equity risk factors, over the most recent 120 months (column 2); using the average of the individual loadings of the constituents of the momentum strategy using a maximum estimation window of 120 (column 3); or using past conditional loadings using Equation 3.1 (final column) estimated over the most recent 120 months.

		Unhedged	Hedge 1	Hedge 2	Hedge 3
	Return	0.67%	1.08%	0.68%	1.02%
Always hedging	St. dev	6.5%	6.1%	5.0%	5.0%
	Sharpe ratio p.a.	0.36	0.61	0.47	0.71
	Return		1.1%	0.93%	1.14%
Only hedging negative loadings	St. dev		6.1%	5.1%	5.3%
	Sharpe ratio p.a.		0.63	0.63	0.75

when the estimated factor exposures are negative. The results are shown in Table 3.6. As expected for the two hedging methods most accurately measuring the time-varying factor exposures of momentum hedge 2 and hedge 3 only hedging negative loadings on the equity risk factors yields a higher return and also a higher risk than hedging both positive and negative exposures. The Sharpe ratio when only hedging negative exposures improves.

On an annualized basis (times the square root of 12) the Sharpe ratio for the unhedged momentum strategy is 0.36, and this increases to 0.47 for hedge 2 and 0.71 for hedge 3. Only hedging negative loadings further increases the Sharpe ratio to 0.63 for hedge 2 and 0.75 for hedge 3. Hence understanding the conditional risk exposures of the original momentum strategy with a Sharpe ratio of 0.36 provides a (partially) hedged momentum strategy that can (more than) double the Sharpe ratio to (0.71) 0.75. The results for the unconditional hedge do not change a lot. This hedge estimates negative loadings for all factors most of the time given the overall negative loading of momentum on these factors (see Table 3.1). Hence, this hedge already frequently takes positive offsetting positions.

3.5 Estimation windows and biases

In Section 3.4 we have used an estimation window of 120 months for all hedges. We do this because we find worse hedging results for shorter estimation windows. In this section we show results for these shorter estimation windows, and also compare ex-ante estimated betas with ex-post realized betas of the momentum constituents. In Section 3.5.3 we look at the importance of including size and HML when hedging.

3.5.1 Estimation windows

Shorter estimation windows reduce the risk reduction of the hedging strategies. Table 3.1 shows the average monthly returns, standard deviations and Sharpe ratios for shorter estimation windows of 36 and 60 months. For all three hedges we see that the longer the estimation window, the lower the remaining risk. For the conditional hedge 3 a short estimation window even leads to hedged momentum having a higher standard deviation than unhedged momentum. For this conditional hedge we find near-singularity issues when estimating 10 parameters (a constant plus three betas for each factor return) on just 36 months of momentum return history. This results in unstable estimated hedge ratios and as a consequence an increase of the variance rather than a reduction. Hence it is important to have a sufficiently large estimation window for the hedge ratios to reduce the risk of the momentum strategy.

The final column of Table 3.1 shows the regression R-squared for Equation 3.1 applied to unhedged and hedged momentum. This number corresponds to the explanatory power of the time-varying risk exposures. Interestingly for the unconditional hedge we do find a lower R-squared for the regression in Equation 3.1 when using shorter estimation windows. By using shorter estimation windows the unconditional hedge will shift from simply removing the average negative exposures to the equity risk factors to getting closer to the actual exposures of momentum. As a result the R-squared drops from 37.8 to 32.5 percent.

In contrast the time-varying exposure to the equity risk factors decreases for longer estimation windows for hedge 2 and hedge 3, especially for the conditional hedge. Although the conditional hedge is much better in reducing the conditional risks of the momentum strategy for the shorter estimation windows the reduction in R-squared decreases. For the 36-month estimation window the R-squared is still 17.9 percent, whereas it drops to 5.2

Table 3.1: **Performance and exposures for different estimation windows.** This table shows the monthly returns, standard deviation and annualized Sharpe ratio of momentum (Unhedged) and hedged momentum where the hedge is based on unconditional momentum betas, individual stock betas or conditional momentum betas from Equation 3.1 based on estimation windows of the past 36, 60 or 120 months. The final column shows the R-squared from Equation 3.1 for unhedged momentum and the various hedges. The sample period covers February 1932 to December 2011 (959 months).

	Window	Return	St. dev.	Sharpe ratio	R-squared
Unhedged		0.67%	6.5%	0.36	51.3%
	36	1.03%	6.3%	0.56	32.5%
Hedge 1	60	1.09%	6.2%	0.60	35.3%
	120	1.08%	6.1%	0.61	37.8%
	36	0.55%	5.6%	0.34	29.9%
Hedge 2	60	0.62%	5.2%	0.42	28.1%
	120	0.68%	5.0%	0.47	27.9%
	36	1.00%	7.1%	0.49	17.9%
Hedge 3	60	0.91%	5.9%	0.54	10.5%
	120	1.02%	5.0%	0.71	5.2%

percent when using the 120-month estimation window. This indicates for the conditional hedge it is important to have sufficient data to estimate the long-run conditional exposures of momentum.

3.5.2 Estimation error in individual stock betas

The hedge based on individual stock betas has a systematic bias: the hedge overestimates momentum's factor exposures after positive factor returns and underestimates momentum's factor returns after negative factor returns. The results in Table 3.1 in Subsection 3.4.2 already suggest this systematic bias since the positive exposures after positive factor returns become negative after hedging, while the negative exposures after negative factor returns in the ranking period remain negative. This phenomenon is further illustrated in Table 3.2 by the comparison of the estimated ex-post realized momentum betas with the ex-ante hedge ratios of the hedge methods after up and down factor returns in the ranking period.

Table 3.2 shows the betas estimated for the individual hedge and the realized betas of the momentum strategy after the market, size or value-growth factors posted a large

Table 3.2: **Bias in individual stock betas.** The table shows the ex-post and ex-ante betas of the hedged momentum strategy (WML) based on individual stock betas. WML is the top 10 percent minus bottom 10 percent portfolio based on stock returns from $t-7$ to $t-2$ to invest in month t. The ex-post betas use the period t to $t+11$. The ex-ante betas use the 36, 60 or 120 months preceding month t. All betas are conditional upon the market MKT, SMB or HML having a return larger than one standard deviation above their mean (up) or smaller than one standard deviation below their mean (down) in the ranking period from $t-7$ to $t-2$. The sample period runs from February 1932 to January 2011 (948 months).

	Ex-post betas	Ex-ante betas		
Window	12	36	60	120
MKT up	**0.43**	**0.54**	**0.47**	**0.41**
Top	1.24	1.33	1.30	1.27
Bottom	0.81	0.79	0.84	0.86
MKT down	**-0.51**	**-0.49**	**-0.41**	**-0.35**
Top	0.87	0.81	0.86	0.88
Bottom	1.38	1.30	1.27	1.23
SMB up	**0.69**	**1.47**	**1.23**	**1.09**
Top	1.36	1.89	1.74	1.65
Bottom	0.67	0.42	0.50	0.56
SMB down	**-1.01**	**-1.03**	**-0.92**	**-0.76**
Top	0.51	0.48	0.55	0.64
Bottom	1.52	1.51	1.47	1.40
HML up	**0.44**	**1.33**	**1.11**	**0.88**
Top	0.54	1.13	0.96	0.81
Bottom	0.09	-0.20	-0.15	-0.07
HML down	**-1.05**	**-1.09**	**-0.91**	**-0.70**
Top	-0.10	-0.24	-0.15	-0.04
Bottom	0.95	0.85	0.76	0.66

positive (up) or negative (down) return in the formation period. The realized betas are approximated by the 12 month beta of the constituents of the momentum strategy over months t to $t+11$. For example if the market return exceeds its average return plus one standard deviation (MKT up) during the formation period of WML ($t-7$ to $t-2$) we see that based on individual stock betas estimated over t to $t+11$ the WML portfolio has an (ex-post) beta of 0.43. For hedge 2 we use the individual betas estimated over $t-T$ to $t-1$ and find that these (ex-ante) betas are 0.54, 0.47 and 0.41 for the estimation windows of 36, 60 and 120 months.

Momentum's market exposures are thus overestimated after positive factor returns in the ranking period. This overestimation is also present for momentum's exposures to the size and value factor. These overestimated factor loadings after positive factor returns in the formation period result in the change from a positive loading on the factors for unhedged momentum in these situations to a negative exposure after hedging (as found in Table 3.1).

When looking at the overestimation of momentum's of factor loading after positive factor loadings for the individual hedge in more detail we find that this overestimation decreases for larger estimation windows. Table 3.2 also shows that the overestimation is mostly caused by an overestimation of the high beta stocks after positive factor returns in the top decile rather than an underestimation of the stocks in the bottom decile.

After down factor returns similar, but opposite results: the estimated factor loadings of the individual hedge underestimate the realized betas for the coming 12 months. These underestimated factor loadings then fail to fully hedge the negative loading of the unhedged momentum strategy after negative factor returns leaving a negative exposure to the three factors. Like the overestimation after positive factor returns the underestimation after negative factor returns is mainly caused by the underestimation of the high beta stocks. However, this underestimation increases with larger estimation windows contrary to the decreasing overestimation after positive factor returns,

It is well-known that individual stock betas contain estimation errors and a common solution is to apply Bayesian shrinkage to the individual estimates. This solution, however, does not help when hedging the time-varying exposures of momentum. Shrinkage would help to reduce the overestimation of the high betas in the top decile after up factor returns, but would also increase the underestimation of the high betas in the bottom decile after down factor returns. These two opposite effects result in a hedge performance similar to the performance of the hedge with the original individual loadings.

Overall we find that using 120 months of history is best in reducing the bias in the estimates of the momentum exposures. These results also explain why in Table 8 we found that the hedge based on individual betas improves when moving from a 36-month estimation window to the 120-month estimation window. This is an important finding given that practitioners and academics often make use of the 36-month and 60-month window, respectively.

We believe that these results are related to the nature of the momentum strategy. Assume the true stock market beta is 1. Suppose we look back 36 months and the market goes up in each month and in addition the (true) residual return of this stock also goes up in each month. In that case the market beta estimated over this period will be higher than 1, as it is impossible to distinguish between systematic and stock specific returns. After a positive market return in the ranking period the top portfolio is not only populated by high beta stocks but also populated by stocks that posted high residual returns. Our example stock would definitely be in the top 10 percent portfolio. Increasing the estimation window decreases the possibility that residual returns are predominantly in the same direction as the market return, and hence the bias is reduced. Obviously all of the above also holds for the size and value factor.

3.5.3 The importance of hedging for all three equity risk factors

Recent studies on momentum only focus on the market exposures. We, however, find that it is important to also take into account the size and HML exposures of momentum to fully understand and hedge its exposures. In this section we provide a number of results for the conditional hedge based on only the market factor, without repeating all the tables in detail.

First, the conditional hedge based on only the market leaves on average negative exposures to size and HML of -0.26. Also the worst loss is 52.9 percent compared to 42.3 percent when hedging all three equity risk factors. The Sharpe ratio drops from 0.71 to 0.46.

Second, the conditional regression in Table 3.1 applied to the conditional hedge based on just the market gives an R-squared of 29.4 percent, up from 5.2 percent when hedging based on all three factors. Hence, only hedging out the market exposure leaves the time-varying exposures to the size and HML factors.

Third, if we repeat the non-parametric analysis for the market factor we are left with just two cases: The market moves in the same direction during the formation and investment period, or it moves in the opposite direction. Unhedged momentum gains 1.9 percent per month when the market continues in the same direction, and loses 0.9 percent when the market reverses. Hence ignoring reversals in HML and size does not reveal the much wider gap of a gain of 3.5 percent per month when all factors continue in the same direction, and a loss of 3.6 percent when all factors have a reversal.

Fourth, for the momentum crashes in Table 3.3 we find that the conditional hedge based on just the market factor leads to an average loss of 24.8 percent, up from 14.8 percent when also hedging for size and HML.

Fifth, in stage 4 of the business cycle (see Table 3.4) the conditional hedge earns 0.05 percent per month. The conditional hedge based on the market factor only leads to a loss of -0.77 percent per month. Similarly, the conditional hedge using only the market loses 1.00 percent after a negative 3-year market return, compared to a gain of 0.27 percent reported in Table 3.5 when hedging all three factors.

All results and conclusions improve when considering all three equity risk factors. We therefore recommend not only focusing on just the market factor when investigating the momentum strategy.

3.6 Conclusion

Selecting the past winners and selling the past losers result in time-varying exposures to the Fama and French (1992), Fama and French (1993) equity risk factors. In particular the exposures are higher after the factor returns went up during the ranking period. These exposures lead to losses on the momentum strategy when factors move in the opposite direction in the ranking and investment period.

We show that specifically accounting for the conditional pattern in the time-variation of the exposures provides the best hedge. Hedging decreases the volatility of momentum by about 25 percent and does not suffer from negative returns at turning points of the factor returns as raw momentum does. Furthermore hedged momentum has more stable returns over the business cycle, in contrast to raw momentum that suffers losses in the latter part of recessions.

We also show that hedged momentum returns no longer suffer losses after the stock market has shown a negative return in the recent years. This puts a different light on the findings of Cooper et al. (2004) who claim that the negative returns after down markets and the positive returns after up markets for the raw momentum strategy support the overreaction theories of Daniel et al. (1998) and Hong and Stein (1999). Also the crash risk explanation of momentum by Daniel and Moskowitz (2013) and Daniel et al. (2012) seems less likely as the ex-ante available hedged momentum reduces the loss of the 12 worst months from -36.8 to -14.8 percent.

For investors we have good news as well as an important warning; the latter is also relevant for risk-management. Hedging momentum using conditional factor loadings doubles the annualized Sharpe ratio from 0.36 to 0.71, but the common approach to hedge using individual stock betas leads to systematic biases. Momentum's ex-post factor exposures are generally overestimated by the individual betas after positive factor returns and underestimated after negative factor returns. This is mainly caused by the over- and underestimation of the betas of high beta stocks after positive and negative factor returns respectively. The overestimation of the betas of high beta stocks after positive factor returns decreases when longer estimation windows are used, while the underestimation of the betas of high beta stocks decreases when shorter estimation windows are used.

Hence using these individual betas leads to ex-post negative factor exposures where one ex-ante expects to have zero exposures. Longer windows of 120 months rather than the common 36 and 60 month windows reduce the estimation bias.

Chapter 4

Bayesian Factor Modeling with Industry Momentum Strategies*

4.1 Introduction

Financial and economic relations vary over time. One of these time varying relations is the increase in the correlations between equities during market downturns. During equity market crashes all stocks lose and diversification does not help much in reducing these losses. In this chapter we analyze this time variation in the comovements of equity returns with a Bayesian latent factor model. The model allows the number of latent factors to vary over time; that is, during equity market downturns fewer latent factors are assumed to be able to explain the same amount of variation in equity returns. We apply this Bayesian latent factor model to evaluate risk and return properties of a residual industry momentum strategy, which is a combination of residual momentum by Blitz et al. (2011) and industry momentum by Moskowitz and Grinblatt (1999) and compare the results with those obtained from using a Bayesian factor model with the three Fama and French (1993) equity risk factors as explanatory variables.

The advantage of using a Bayesian approach in this context is that the uncertainty with respect to the time varying latent factors is taken into account in a coherent way for forecasting and decision analysis. For the theoretical model, we make use of a standard static latent factor model where the factors are defined as latent variables. Our aim is to let the information in the *likelihood* dominate the prior information and we therefore use flat and weakly informative priors. We show that, under flat or conjugate priors,

*This chapter is based on Baştürk, Grassi, et al. (2016).

the conditional posteriors of the static factors and of the factor loadings are well-known distributions. The marginal posteriors, however, do not have this property. The marginal posteriors of the factors potentially have fat tails, but these densities are integrable due to our 'regularization' of the likelihood information through weakly informative normal prior distributions of factors and their loadings. Given our analytical results some basic Markov Chain Monte Carlo methods can be used to numerically evaluate results.

In our Bayesian latent factor model we address the following well-known issue: The number of factors in a static factor model cannot be easily treated as a model parameter due to the identification assumptions of the model. As a consequence the number of factors are often fixed prior to the analysis. This difficulty also presents itself if one wants to allow for time variation in the number of factors by extending the standard factor model. Neglecting this possibility of time variation may adversely affect model forecasts. We present a method that accounts for time variation in the number of factors through a moving window estimation. For each estimation, the optimal number of factors are chosen according to the predictive likelihoods of models with a different number of factors. We use the predictive likelihood approach instead of the marginal likelihood approach since the former ensures that adequate priors are taken using the prior data sample. This also refrains from the (Bartlett, 1957) paradox occurring under diffuse priors; see (Baştürk, Çakmaklı, Ceyhan, & Van Dijk, 2014) for a comprehensive discussion.

An important issue in predictive likelihood comparisons is the size of the training sample, particularly as a ratio of the full sample size. We show the impact of the training sample size using simulated data and find that the correct number of factors are obtained in almost all simulated data sets, regardless of the training sample size.

One focus of this chapter is the derivation of a Bayesian time varying latent factor model using the concept of predictive likelihoods. A second focus is to apply this modeling approach to residual industry momentum; a regular application of residual industry momentum is found in Hühn (2015). Standard momentum strategies, see Jegadeesh and Titman (1993), are long-short equity strategies that rank stocks on their recent total returns, skip a short period and then buy the stocks in the top of this ranking and short-sell the stocks in the bottom of this ranking. Residual momentum is different in the sense that it ranks stocks on their recent residual returns from regressing the total stock returns on the three Fama and French (1993) equity risk factors rather than ranking the stocks on their recent total returns. Industry momentum differs from standard momentum by ranking industry portfolios rather than individual stocks. The main reason that we apply all procedures using industry portfolios rather than individual stocks is that this reduces the computational workload. For a large number of individual stocks the computational workload of our procedures increases substantially.

The residual industry momentum strategy that we apply differs from the approach by Blitz et al. (2011) and Hühn (2015). Rather than using the static Fama and French (1993) equity risk factors to compute residual returns we use a latent factor model in which the number of factors varies over time. In our strategy the time varying latent factors are not explicitly linked to the equity risk factors market, size and value. We aim to let the basic data features 'speak' for themselves. We compare the results of our Bayesian time-varying latent factor modeling approach with the results from using a Bayesian factor model that takes account of the three Fama and French (1993) equity risk factors as explanatory variables.

Using US industrial portfolios over the period 1980-2015, our empirical results show that the Bayesian time varying latent factor model applied to residual industry momentum outperforms in terms of several return and risk characteristics a Bayesian factor model with standard equity risk factors in turbulent times, in particular during the crisis that started in 2008. We also find that the optimal number of latent factors varies substantially over time and that the number of optimal factors indeed decreases when the equity markets experience large losses.

An other empirical result is that in quiet times like in the nineties, the Fama and French (1993) equity risk factor model has very good performance features. Thus a major conclusion of our empirical analysis is that a time-varying combination of model structures and momentum strategies may give overall superior performance.

The contents of this chapter is organized as follows. In Section 4.2, the standard static latent factor model is presented. Section 4.3 and Section 4.4 contain our Bayesian posterior analysis of the latent factor and the implied simulation technique is discussed. In Section 4.5 the Bayesian factor model using the three Fama and French (1993) equity risk factors is presented. The predictive likelihood approach to determine the optimal number of factors is given in Section 4.6. In Section 4.7 empirical results are presented.

Section 4.8 states conclusions and suggestions for further research.

We note that some of the theoretical results presented in this chapter are part of a larger research project about Bayesian analysis of reduced rank econometric models. For some proofs we refer, therefore, to Zellner, Ando, Baştürk, Hoogerheide, and Van Dijk (2014) and Baştürk, Hoogerheide, Kleijn, and Van Dijk (2015).

4.2 A static factor model

Consider a time series of observations on p economic variables during the period $t = 1, ..., T$. A static factor model may be defined as:

$$y_t' = f_t'\Lambda + \varepsilon_t', \tag{4.1}$$

$$f_t' = 0 + u_t', \tag{4.2}$$

where y_t' is a $1 \times p$ vector of observations on the economic variables of interest, f_t' is the $1 \times r$ vector of the unobserved random factors, and Λ is a $r \times p$ matrix of unknown random factor loadings. We assume that $p >> r$, that is the information in the p economic variables of interest can be compressed to a much lower number of r unobserved random factors. Deterministic factors, individual fixed effects and trends can be added to the factor structure. For convenience of the analysis, we refrain from these latter extensions of the factor model. For a background on factor models we refer to Lawley and Maxwell (1971) and Anderson (1984).

The following distributional assumptions are made for the model in (4.1) and (4.2):

$$u_t \sim NID(0, I_r), \; \text{cor}(u_t, \varepsilon_s) = 0, \; \forall s, t, \tag{4.3}$$

$$\varepsilon_t \sim NID(0, \Sigma), \; \Sigma = \begin{pmatrix} \sigma_{11}^2 & 0 & \dots & 0 \\ 0 & \sigma_{22}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{pp}^2 \end{pmatrix}, \tag{4.4}$$

where NID stands for indepently and normally distributed and the diagonal covariance matrix assumption with respect to the disturbances ε_t implies that all cross-sectional correlation is captured by the factors f_t.

The model introduced in (4.1) and (4.2) can be specified in matrix notation as follows:

$$\begin{aligned} Y &= F\Lambda + E, &(4.5)\\ F &= 0 + U, &(4.6) \end{aligned}$$

where Y is the $T \times p$ matrix of observations, F is the $T \times r$ matrix of factors, Λ is the $r \times p$ matrix of factor loadings, E is the $T \times p$ matrix of disturbances and U is the $T \times r$ matrix of disturbances. In addition, $\text{cor}(U, E) = 0$, $E \sim MN(0, \Sigma, I_T)$ and $U \sim MN(0, I_p, I_T)$. In this notation, $MN(X, \Omega, \Phi)$ denotes the matricvariate normal distribution with mean M and scale parameters Ω, Φ, and I_k is the $k \times k$ identity matrix.

The identification problem of determining factors F and their loadings Λ in this model stems from the following equality:

$$F\Lambda = FRR^{-1}\Lambda,$$

for any $r \times r$ invertible matrix R, which has r^2 free parameters. Hence at least r^2 restrictions are needed for the model to be identified. In addition, matrix R can as well be an $r \times r$ invertible rotation matrix, i.e. rotated factors and loadings provide the same likelihood unless further restrictions are set Λ.

Several identification strategies can be followed using alternative restrictions on F and Λ. We consider two types of linear normalization:

1. Linear normalization of Λ using a lower-triangular matrix, as in Geweke and Zhou (1996) and Lopes and West (2004):

$$\Lambda = \begin{pmatrix} \Lambda_1^{(r\times r)} & \Lambda_2^{(r\times(p-r))} \end{pmatrix},\ \Lambda_1 = \begin{pmatrix} \lambda_{11} & 0 & \dots & 0 \\ \lambda_{21} & \lambda_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{r1} & \lambda_{r2} & \dots & \lambda_{rr} \end{pmatrix} \tag{4.7}$$

   where Λ_2 is unrestricted.

2. Linear normalization of Λ with identity restrictions, as in Bai and Ng (2013):

$$\Lambda = \begin{pmatrix} I^{(r\times r)} & \Lambda_2^{(r\times(p-r))} \end{pmatrix}, \tag{4.8}$$

where Λ_2 is unrestricted.

We note that different normalizations exist in the literature. For a detailed discussion of identification restrictions for the static factor model, see Bai and Ng (2013). We also refer to Baştürk et al. (2015) for an extended discussion on possible identification restrictions for a similar reduced rank model, namely the cointegration model. We follow the present identification restrictions for ease of tractability in the posterior and predictive analysis.

4.3 Conditional Posterior densities of F, Λ, and Σ

In this section we derive conditional posterior densities of Σ, the factors F and their loadings Λ. We analyze the existence of posterior moments for flat as well as normal prior specifications in the next section. These results are important for the construction of simulation algorithms for the numerical evaluation of the densities.

We start with the interpretation of the model in (4.5) and (4.6) as a model that includes data augmentation with respect to the factors. That is, the factors are interpreted as latent variables on which there is specified a normal distribution. The likelihood function using data augmentation is as follows

$$\begin{aligned} p(Y \mid \Lambda, \Sigma) &= \int p(Y|F, \Lambda, \Sigma) p(F \mid \Lambda, \Sigma) dF \\ &\propto \int |\Sigma|^{-\frac{T}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}(F'F)\right) \\ &\quad \times \exp\left(-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}(Y - F\Lambda)'(Y - F\Lambda)\right)\right) dF \qquad (4.9) \end{aligned}$$

We consider the following flat prior for Λ, Σ:

$$p(\Lambda, \Sigma) \propto |\Sigma|^{-h/2}, h > 1, \qquad (4.10)$$

and we choose $h = p + 1$ which implies that the information in the likelihood is the same from a Bayesian or a Frequentist perspective at this stage of the analysis.

An alternative interpretation of prior and likelihood is to specify only the model given in equation (4.1) or equation (4.5) and derive a likelihood function of the parameters F, Λ, Σ

given the data Y:

$$p(F, \Lambda, \Sigma|Y) = (2\pi)^{rT/2}|\Sigma|^{-\frac{T}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}(Y - F\Lambda)'(Y - F\Lambda)\right)\right) \tag{4.11}$$

where this likelihood is similar to (4.9) except that the normal prior on F is deleted and no integration with respect to F takes place. Then we assume as prior:

$$p(F, \Lambda, \Sigma) \propto \exp\left(-\frac{1}{2}\operatorname{tr}(F'F)\right)|\Sigma|^{-h/2}, h = p + 1. \tag{4.12}$$

It is obvious that this Bayesian model yields the same posterior as the one with the augmented likelihood and a flat prior on F. We will use both interpretations in subsequent analysis.

Apart from the prior specifications on F we make use of prior specifications for Λ that are similar. That is, for the prior of Λ both a flat and a normal prior will be used. This allows us to investigate the prior effects on posterior moments in several cases by combining flat and normal priors on F and on Λ and both on F and Λ.

We start with the joint posterior density for F, Λ, Σ. Using the likelihood in (4.11) and the prior in (4.12), we obtain the following density:

$$\begin{aligned} p(F, \Lambda, \Sigma \mid Y) \quad \propto \quad & |\Sigma|^{-\frac{T+p+1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}(F'F)\right) \\ & \times \exp\left(-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}(Y - F\Lambda)'(Y - F\Lambda)\right)\right). \end{aligned} \tag{4.13}$$

We note that the same result would be obtained by using the model with data augmentation in (4.5) and (4.6), where the integration with respect to F is not yet performed.

Conditional posterior of $\Sigma|F, \Lambda$ From (4.13), the conditional posterior of $\Sigma|F, \Lambda, Y$ is:

$$p(\Sigma \mid F, \Lambda, Y) \propto |\Sigma|^{-\frac{T+p+1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}(Y - F\Lambda)'(Y - F\Lambda)\right)\right) \tag{4.14}$$

$$= |\Sigma|^{-\frac{T+p+1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left((Y - F\Lambda)'(Y - F\Lambda)\Sigma^{-1}\right)\right), \tag{4.15}$$

$$\propto \operatorname{IW}\left(\Sigma|(Y - F\Lambda)'(Y - F\Lambda), T\right) \tag{4.16}$$

where $\text{IW}(\Sigma|\Omega, T)$ denotes the inverted Wishart density with a $p \times p$ positive definite symmetric scale matrix Ω and T degrees of freedom. For the factor model, this density exists when the inner product $\Omega = (Y - F\Lambda)'(Y - F\Lambda)$ has full rank p and $T > p - 1$. We assume that these conditions, which depend on the data Y and the sample size, hold.

Conditional posterior of $F|\Lambda, \Sigma$ To derive the conditional posterior of $F|\Lambda, \Sigma$, we complete the squares on F in (4.13):

$$\begin{aligned} p(F, \Lambda, \Sigma \mid Y) \propto & |\Sigma|^{-\frac{T+p+1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}(F'F)\right) \\ & \times \exp\left(-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}(Y - F\Lambda)'(Y - F\Lambda)\right)\right) \qquad (4.17) \\ = & |\Sigma|^{-\frac{T+p+1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left(FF' + (Y - F\Lambda)\Sigma^{-1}(Y - F\Lambda)'\right)\right) \qquad (4.18) \\ = & |\Sigma|^{-\frac{T+p+1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left(Y\left(\Sigma^{-1} - \Sigma^{-1}\Lambda'(I + \Lambda\Sigma^{-1}\Lambda')^{-1}\Lambda\Sigma^{-1}\right)Y'\right)\right) \\ & \times \exp\left(-\frac{1}{2}\operatorname{tr}\left((F - \hat{F})\left(I + \Lambda\Sigma^{-1}\Lambda'\right)(F - \hat{F})'\right)\right) \qquad (4.19) \end{aligned}$$

where $\hat{F} = Y\Sigma^{-1}\Lambda'(I_r + \Lambda\Sigma^{-1}\Lambda')^{-1}$.

From (4.19), the conditional posterior of $F|\Lambda, \Sigma$ is a matricvariate normal density, where the latter is indicated by the symbol mn:

$$p(F \mid \Lambda, \Sigma, Y) \propto MN\left(F; Y\Sigma^{-1}\Lambda'(I_r + \Lambda\Sigma^{-1}\Lambda')^{-1}, I_T, \left(I + \Lambda\Sigma^{-1}\Lambda'\right)^{-1}\right), \qquad (4.20)$$

where the first scale parameter I_T stems from the assumption on the distribution of the disturbances $\varepsilon_t \sim NID(0, \Sigma)$. Using properties from the matricvariate normal distribution, one can derive in a straightforward way that the marginal posterior density of each column f_t with $t = 1, ..., T$ is also normal.

$$p(f_t \mid \Lambda, \Sigma, y_t) \propto N\left(f_t; (I_r + \Lambda\Sigma^{-1}\Lambda')^{-1}\Lambda\Sigma^{-1}y_t, \left(I_r + \Lambda\Sigma^{-1}\Lambda'\right)^{-1}\right), \qquad (4.21)$$

where the symbol N indicates the normal density function.

Conditional posterior of $\Lambda|F, \Sigma$ For convenience we present results for the case of linear normalization with identity restrictions as given in (4.8). The case of triangular restrictions gives similar results but the derivations are more involved. We note that Λ_1

was defined as an identity matrix: $\Lambda = \begin{pmatrix} \Lambda_1^{(r\times r)} & \Lambda_2^{(r\times(p-r))} \end{pmatrix} = \begin{pmatrix} I^{(r\times r)} & \Lambda_2^{(r\times(p-r))} \end{pmatrix}$, where Λ_2 is unrestricted. This restriction implies that the posterior density is defined over the matrix of parameters, Λ_2, since Λ_1 is now fixed. Thus the conditional posterior of Λ_2 from (4.13) is:

$$p(\Lambda_2 \mid F, \Sigma, Y) \propto |\Sigma|^{-\frac{T+p+1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}(Y - F\Lambda)'(Y - F\Lambda)\right)\right). \tag{4.22}$$

We next complete the squares on Λ in (4.22):

$$\begin{aligned} p(\Lambda_2 \mid F, \Sigma, Y) &\propto |\Sigma|^{-\frac{T+p+1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}(Y - F\Lambda)'(Y - F\Lambda)\right)\right) && (4.23)\\ &= |\Sigma|^{-\frac{T+p+1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}Y'M_FY\right)\right) \\ &\quad \times \exp\left((-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}(\Lambda - \hat{\Lambda})'F'F(\Lambda - \hat{\Lambda})\right)\right) && (4.24)\\ &\propto \exp\left(-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}(\Lambda - \hat{\Lambda})'F'F(\Lambda - \hat{\Lambda})\right)\right), && (4.25) \end{aligned}$$

where $Y'M_FY = Y'Y - Y'F(F'F)^{-1}F'Y$ and $\hat{\Lambda} = (F'F)^{-1}F'Y$. The right-hand side of (4.25) is for an unrestricted Λ proportional to a matricvariate normal density with location $\hat{\Lambda}$, and scale matrices Σ and $(F'F)^{-1}$. We restrict $|F'F|$ to be nonsingular for all values of F except events with measure zero. We comment on this in the next section.

From (4.25), it is also seen that the conditional posterior $\Lambda_2|\Lambda_1 = I$ has a conditional matricvariate normal density with the following parameters:

$$p(\Lambda_2 \mid \Lambda_1, F, \Sigma, Y) \propto MN(\Lambda_2; \hat{\Lambda}_2, \Sigma_{22}, P_{2|1}) \tag{4.26}$$

where the parameters of the matricvariate normal density are defined using the following equations and matrix partitionings:

$$\begin{aligned}
\hat{\Lambda} &= (F'F)^{-1}F'Y = \begin{pmatrix} \hat{\Lambda}_1^{(r\times r)} & \hat{\Lambda}_2^{r\times(p-r)} \end{pmatrix} \\
\Sigma &= \begin{pmatrix} \Sigma_{11}^{(r\times r)} & \Sigma_{12}^{(r\times(p-r))} \\ \Sigma_{21}^{((p-r)\times r)} & \Sigma_{22}^{(p-r)\times(p-r)} \end{pmatrix} \\
P &= F'F \\
P_{2|1} &= P + (I - \hat{\Lambda}_1)'\Sigma_{11.2}(I - \hat{\Lambda}_1) \\
\Sigma_{11.2} &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}.
\end{aligned}$$

We refer to Zellner (1971), appendix B.5 and Bauwens, Lubrano, and Richard (2000), appendix A.2 for the details of this property. See also Baştürk et al. (2015) for an extension to the conditionals of a matricvariate t density.

Under the fixed parameter assumption $\Lambda_1 = I$, (4.26) is adjusted such that the conditional posterior density is as follows:

$$p(\Lambda_2 \mid F, \Sigma, Y) \quad \propto \quad MN(\Lambda_2; \hat{\Lambda}_2, \Sigma_{22}, P_{2|1}) \tag{4.27}$$

and this density is proper if Σ is positive definite and $P = F'F$ is positive definite for all values of F. We investigate this latter condition in the next section.

We conclude that the conditional posterior distributions, with their corresponding density functions, of each of the parameters F, Λ, Σ have known analytical forms that are given as matricvariate normal and inverse Wishart distributions, and thus are integrable. It appears, therefore, that a straightforward simulation algorithm, such as the Gibbs sampler, can be used for posterior inference. However, this depends on whether the joint and marginal posterior distributions exist with corresponding proper densities. We will investigate this in the next section.

A second conclusion can also be derived. It follows rather directly from the analysis of this section that a uniform prior on F or a normal prior on Λ_2 or normal priors on both parameters will yield the same results. That is, the conditional distributions of F, Λ, Σ are all members of known classes of distributions only their parameters differ. Again, Gibbs sampling appears possible but depends on the existence of the joint and marginal distributions, which we investigate next.

4.4 Marginal posterior densities of F and Λ and existence conditions for posterior moments

We start to derive the joint marginal posterior of (F, Λ) and then derive the individual marginal posteriors of F and Λ by making use of the decomposition of a joint density into a conditional density and a marginal density and by using properties of the multivariate and matricvariate normal and t distributions.

Using an inverse Wishart integration step on Σ in (4.13), the joint marginal density of F and Λ is given by:

$$p(F, \Lambda \mid Y) \propto \exp\left(-\frac{1}{2}\operatorname{tr}(F'F)\right)\left|(Y - F\Lambda)'(Y - F\Lambda)\right|^{-T/2}, \tag{4.28}$$

under the condition that $(Y - F\Lambda)'(Y - F\Lambda)$ is a PDS matrix for all values of Λ and F and the sample size T was already restricted to $T > p - 1$.

4.4.1 Marginal posterior of F

As mentioned above, we make use of the decomposition of $p(F, \Lambda \mid Y)$ into the conditional density $p(\Lambda|F, Y)$ and the marginal $p(F \mid Y)$. The steps of this derivation follow from Baştürk et al. (2015), with slight changes in the notation.

We first rewrite (4.28) by completing the squares on F as follows:

$$p(F, \Lambda \mid Y) \propto \exp\left(-\frac{1}{2}\operatorname{tr}(F'F)\right)\left|Y'M_F Y + \left(\Lambda - \hat{\Lambda}\right)' F'F\left(\Lambda - \hat{\Lambda}\right)\right|^{-T/2} \tag{4.29}$$

where we use the earlier definitions of $\hat{\Lambda}$ and $Y'M_F Y$ and the earlier condition on $F'F$.

Using the definition of the matricvariate t density the joint posterior density in (4.29) is:

$$p\left(F,\Lambda|Y\right)\propto\exp\left(-\frac{1}{2}\,\mathrm{tr}(P)\right)\left|Q+\left(\Lambda-\hat{\Lambda}\right)'P\left(\Lambda-\hat{\Lambda}\right)\right|^{-T/2}$$
$$\propto\exp\left(-\frac{1}{2}\,\mathrm{tr}(P)\right)|P|^{-\frac{p}{2}}|Q|^{-\frac{T-r}{2}}p_{Mt}\left(\Lambda|\hat{\Lambda},P,Q,T\right) \qquad (4.30)$$

where we use the earlier definition $P=F'F$ and $Q=Y'M_FY$, which do not depend on Λ, and $p_{Mt}\left(\Lambda|\hat{\Lambda},P,Q,T\right)$ denotes the matricvariate t density function with location $\hat{\Lambda}$, scale parameters P,Q, and degrees of freedom T.

In case of no normalization for Λ, the matricvariate t density in (4.30) can, given F, be decomposed as follows:

$$p\left(\Lambda\mid F,Y\right)=p\left(\Lambda_2\mid\Lambda_1,F,Y\right)\times p\left(\Lambda_1\mid F,Y\right) \qquad (4.31)$$

and the conditional density of $\Lambda_2|\Lambda_1$ and the marginal density Λ_1, both conditional upon (F,Y), are given as:

$$p\left(\Lambda_2\mid\Lambda_1,F,Y\right)=p_{Mt}\left(\Lambda_2|M_{\Lambda_2|\Lambda_1},P_{\Lambda_2|\Lambda_1},Q_{22},T\right) \qquad (4.32)$$
$$p\left(\Lambda_1\mid F,Y\right)=p_{Mt}\left(\Lambda_1|M_{\Lambda_1},P,Q_{11.2},T-p+r\right) \qquad (4.33)$$

where

$$M_{\Lambda_2|\Lambda_1}=\hat{\Lambda}_2-(\Lambda_1-\hat{\Lambda}_1)Q_{12}Q_{22}^{-1},\;\; M_{\Lambda_1}=\hat{\Lambda}_1 \qquad (4.34)$$
$$P_{\Lambda_2|\Lambda_1}=P+(\Lambda_1-\hat{\Lambda}_1)Q_{11.2}(\Lambda_1-\hat{\Lambda}_1)' \qquad (4.35)$$
$$Q_{11.2}=Q_{11}-Q_{12}Q_{22}^{-1}Q_{21} \qquad (4.36)$$
$$\hat{\Lambda}=(F'F)^{-1}F'Y=\begin{pmatrix}\hat{\Lambda}_1^{(r\times r)} & \hat{\Lambda}_2^{(r\times(p-r))}\end{pmatrix} \qquad (4.37)$$
$$Y'M_FY=Q=\begin{pmatrix}Q_{11}^{(r\times r)} & Q_{12}^{(r\times(p-r))}\\ Q_{21}^{((p-r)\times r)} & Q_{22}^{(p-r)\times(p-r))}\end{pmatrix} \qquad (4.38)$$

see Dickey (1967), Zellner (1971), Bauwens et al. (2000) and Baştürk et al. (2015) for details.

Inserting (4.31)–(4.33) in (4.30) we obtain:

$$\begin{aligned} p\left(F,\Lambda|Y\right) \propto \exp\left(-\frac{1}{2}\operatorname{tr}(P)\right)|P|^{-\frac{p}{2}}|Q|^{-\frac{T-r}{2}} \\ \times\, p_{Mt}\left(\Lambda_2; M_{\Lambda_2|\Lambda_1}, P_{\Lambda_2|\Lambda_1}, Q_{22}, T\right) \\ \times\, p_{Mt}\left(\Lambda_1; M_{\Lambda_1}, P, Q_{11.2}, T-p+r\right). \end{aligned} \tag{4.39}$$

From (4.39), it is clear that the conditional density of $\Lambda_2|\Lambda_1, F, Y$ is a matricvariate t density, given all values of Λ_1:

$$p\left(\Lambda_2|\Lambda_1, F, Y\right) = p_{Mt}\left(\Lambda_2; M_{\Lambda_2|\Lambda_1}, P, Q_{22}, T\right), \tag{4.40}$$

and under the restriction $\Lambda_1 = I$, this density is simply evaluated at $\Lambda_1 = I$.

Marginal posterior density of F: We next derive the marginal posterior density of F under the linear restriction on Λ. This density is obtained using (4.39) under the restriction $\Lambda_1 = I$:

$$\begin{aligned} p\left(F,\Lambda_2|\Lambda_1 = I, Y\right) \propto \exp\left(-\frac{1}{2}\operatorname{tr}(P)\right)|P|^{-\frac{p}{2}}|Q|^{-\frac{T-r}{2}} \\ \times\, p_{Mt}\left(\Lambda_2; M_{\Lambda_2|\Lambda_1=I}, P, Q_{22}, T\right) \\ \times\, p_{Mt}\left(\Lambda_1 = I; M_{\Lambda_1}, P, Q_{11.2}, T-p+r\right) \end{aligned} \tag{4.41}$$

where this density is now defined for F, Λ_2, since Λ_1 is fixed.

From (4.41) we have:

$$p(F \mid Y) \propto \int p(F, \Lambda_2 \mid Y)\, d\Lambda_2 \tag{4.42}$$

$$\begin{aligned} &= \exp\left(-\frac{1}{2}\operatorname{tr}(P)\right) |P|^{-\frac{p}{2}} |Q|^{-\frac{T-r}{2}} \\ &\quad \times p_{Mt}\left(\Lambda_1 = I; M_{\Lambda_1}, P, Q_{11.2}, T-p+r\right) \\ &\quad \times \int p_{Mt}\left(\Lambda_2; M_{\Lambda_2|\Lambda_1}, P, Q_{22}, T\right) d\Lambda_2 \end{aligned} \tag{4.43}$$

$$\begin{aligned} &= \exp\left(-\frac{1}{2}\operatorname{tr}(P)\right) |P|^{-\frac{p}{2}} |Q|^{-\frac{T-r}{2}} \\ &\quad \times p_{Mt}\left(\Lambda_1 = I; M_{\Lambda_1}, P, Q_{11.2}, T-p+r\right) \end{aligned} \tag{4.44}$$

$$\begin{aligned} &\propto \exp\left(-\frac{1}{2}\operatorname{tr}(P)\right) |P|^{-\frac{p}{2}} |Q|^{-\frac{T-r}{2}} |P|^{-\frac{T-p+r-r}{2}} |Q_{11.2}|^{-\frac{r}{2}} \\ &\quad \times \left|P^{-1} + (I - M_{\Lambda_1}) Q_{11.2}^{-1} (I - M_{\Lambda_1})'\right|^{-\frac{T-p+r}{2}}. \end{aligned} \tag{4.45}$$

We next simplify (4.45) using the determinant equality $|Q_{11.2}| = |Q_{11} - Q_{12}Q_{22}^{-1}Q_{21}|$:

$$\begin{aligned} p(F \mid Y) &\propto \exp\left(-\frac{1}{2}\operatorname{tr}(P)\right) |P|^{-\frac{T}{2}} |Q|^{-\frac{T}{2}} |Q_{22}|^{\frac{r}{2}} \\ &\quad \times \left|P^{-1} + (I - M_{\Lambda_1}) Q_{11.2}^{-1} (I - M_{\Lambda_1})'\right|^{-\frac{T-p+r}{2}}. \end{aligned} \tag{4.46}$$

Note that the marginal posterior in (4.46) is in line with the marginal posterior of the adjustment parameters of a cointegration model given in Baştürk et al. (2015), with different definitions of the positive definite symmetric matrices P and Q. These authors further analyze the existence conditions of the factors and show that for any positive definite symmetric matrix Q, the third and the fourth factors, $|Q|^{-\frac{T}{2}}|Q_{22}|^{\frac{r}{2}}$ in (4.46) are bounded from above and below. Therefore the existence conditions of the density in (4.46) do not depend on these terms.

Thus, the existence of the marginal distribution of F, with density in (4.46), depends on the product of the first, second and fifth factor:

$$\begin{aligned} &\exp\left(-\frac{1}{2}\operatorname{tr}(F'F)\right) |F'F|^{-\frac{T}{2}} \left|F'F^{-1} + (I - M_{\Lambda_1}) Q_{11.2}^{-1} (I - M_{\Lambda_1})'\right|^{-\frac{T-p+r}{2}} \\ &\leq \exp\left(-\frac{1}{2}\operatorname{tr}(F'F)\right) |F'F|^{-\frac{p-r}{2}}. \end{aligned} \tag{4.47}$$

We refer to Baştürk et al. (2015) where a proof is presented that $|F'F|^{-(p-r)/2}$ is integrable in a finite area around the event $F = 0$ but that the function does not tend fast enough to zero when the arguments tend to infinity in order to be integrable on the whole real space of F. The argument is that $|F'F|^{-(p-r)/2}$ can be interpreted as indicating the limiting tail behavior of a matricvariate t density of the matrix F with a degrees of freedom restriction of $r = 0$ which implies non-regular tail behavior of the marginal posterior of F. However, since the exponential factor tends to zero much faster and in a regular way, this first factor is responsible for integrability of the marginal posterior of F.

For expository purpose, we present here the result for the case of $p = 2$ and $r = 1$, where $F = f$ is a $T \times 1$ vector of latent observations on one factor with typical element f_t. In this case the following holds for the product of the first factor and the bounding function in (4.46):

$$\exp\left(-\frac{1}{2}\operatorname{tr}(F'F)\right) = \exp\left(-\frac{1}{2}\sum_{t=1}^{T} f_t^2\right) \tag{4.48}$$

$$|F'F|^{-(p-r)/2} \leq \left(\sum_{t=1}^{T} f_t^2\right)^{-1/2} \tag{4.49}$$

where it is immediately clear that the exponential factor in (4.48) tends to zero much faster than the third factor in (4.49).

For an illustration, consider the limit $\sum_{t=1}^{T} f_t^2 \to \infty$. The multiplication of the first and the bounding function in the limit is:

$$\begin{aligned} &\lim_{\sum_{t=1}^{T} f_t^2 \to \infty} \exp\left(-\frac{1}{2}\operatorname{tr}(F'F)\right) |(F'F)|^{-(p-r)} \\ &\leq \lim_{\sum_{t=1}^{T} f_t^2 \to \infty} \exp\left(-\frac{1}{2}\sum_{t=1}^{T} f_t^2\right)\left(\sum_{t=1}^{T} f_t^2\right)^{-1/2} = 0, \end{aligned} \tag{4.50}$$

and the left hand side of (4.50) is non-negative by definition, hence we have:

$$\lim_{\sum_{t=1}^{T} f_t^2 \to \infty} \exp\left(-\frac{1}{2}\operatorname{tr}(F'F)\right) |F'F|^{-(p-r)/2} = 0. \tag{4.51}$$

From (4.46) and (4.51), we conclude that the marginal posterior of F under flat priors has an asymptote when one or more of the factors approach infinity. Still, the density is integrable due to the first factor in (4.46), which comes from the normal prior on F. Whether the prior for Λ_2 is uniform or normal is not essential for these results. We note again that all conditionals are integrable both under a normal and a uniform prior.

The results of this subsection are important for the construction of a Markov Chain Monte Carlo method. Mechanical application of the Gibbs sampling method under a uniform prior on F would lead to the situation where the sampler gets stuck at an absorbing state. This situation of proper conditional densities but improper marginal densities is also been analyzed by Zellner et al. (2014) and by Baştürk, Stefano, Hoogerheide, Opschoor, and Van Dijk (2015) for another reduced rank model, i.e. the instrumental variables model, and by Hobert and Casella (1998) for the hierarchical linear mixed model.

4.4.2 Marginal posterior of Λ_2

In this subsection, we explore features of the marginal posterior of Λ_2. For the general case of a full covariance matrix Σ and flat priors on Λ_2, this density is improper. However, imposing a diagonal covariance matrix on Σ and making use of regularization (normal) priors for Λ_2 make the posterior of Λ_2 proper and an easy simulation method is available for computing the density. We start to consider a flat prior instead of (4.10). The joint posterior in this case is similar to (4.19) except for the factor $|\Sigma|^{-\frac{p+1}{2}}$.

We integrate out F in the joint posterior (4.19), using the conditional posterior of F in (4.20). We further use the identification restriction, hence only partition Λ_2 of Λ is stochastic:

$$\begin{aligned}
p(\Lambda_2, \Sigma \mid Y) \propto & \int |\Sigma|^{-\frac{T}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left(Y\left(\Sigma^{-1} - \Sigma^{-1}\Lambda'(I + \Lambda\Sigma^{-1}\Lambda')^{-1}\Lambda\Sigma^{-1}\right)Y'\right)\right) \\
& \times \exp\left(-\frac{1}{2}\operatorname{tr}\left((F - \hat{F})\left(I + \Lambda\Sigma^{-1}\Lambda'\right)(F - \hat{F})'\right)\right) dF \qquad (4.52) \\
\propto & |\Sigma|^{-\frac{T}{2}} |I + \Lambda\Sigma^{-1}\Lambda'|^{-T/2} \\
& \times \exp\left(-\frac{1}{2}\operatorname{tr}\left(Y\left(\Sigma^{-1} - \Sigma^{-1}\Lambda'(I + \Lambda\Sigma^{-1}\Lambda')^{-1}\Lambda\Sigma^{-1}\right)Y'\right)\right) \qquad (4.53) \\
\propto & |\Sigma|^{-\frac{T}{2}} |(I + \Lambda\Sigma^{-1}\Lambda'|^{-T/2} \exp\left(-\frac{1}{2}\operatorname{tr}\left(Y\left(\Sigma + \Lambda\Lambda'\right)^{-1}Y'\right)\right) \qquad (4.54)
\end{aligned}$$

where the last line uses Woodbury's identity, see, e.g., Seber (1984).

Next, we use the matrix determinant properties to obtain:

$$p(\Lambda_2, \Sigma \mid Y) \propto |\Sigma|^{-\frac{T}{2}} \left|I + \Lambda\Sigma^{-1}\Lambda'\right|^{-T/2} \exp\left(-\frac{1}{2}\,\mathrm{tr}\left(Y\left(\Sigma + \Lambda\Lambda'\right)^{-1} Y'\right)\right) \tag{4.55}$$

$$= |\Sigma|^{-\frac{T}{2}} |\Sigma|^{\frac{T}{2}} \left|\Sigma + \Lambda\Lambda'\right|^{-T/2} \exp\left(-\frac{1}{2}\,\mathrm{tr}\left(Y\left(\Sigma + \Lambda\Lambda'\right)^{-1} Y'\right)\right) \tag{4.56}$$

$$= \left|\Sigma + \Lambda\Lambda'\right|^{-T/2} \exp\left(-\frac{1}{2}\,\mathrm{tr}\left(Y\left(\Sigma + \Lambda\Lambda'\right)^{-1} Y'\right)\right), \tag{4.57}$$

The result is a density function that is proportional to an inverse Wishart density in the sum $\Sigma + \Lambda\Lambda'$. For details on the required matrix determinant properties as well as for the definitions of matricvariate densities, we refer to the appendix of Baştürk et al. (2015).

From (4.57) it is clear that only the sum $\Sigma + \Lambda\Lambda'$ is identified under a flat prior or a normal prior on F for the special case of $p = 1$ and $r = 1$ and for the general case where Σ is a full matrix. For example, the conditional posterior of $\Sigma|\Lambda_2, Y$ in (4.57) is a 'shifted' inverse Wishart density with a shift parameter of $\Lambda\Lambda'$. A similar conclusion holds for the posterior of $\Lambda_2|\Sigma, Y$, but these parameters are not jointly identified. This non-identification of the joint posterior of (Λ_2, Σ) implies that the marginal of Λ_2 is also not identified in the sense that the marginal posterior will be flat. Thus, posterior moments do not exists and a uniform prior on F amplifies this problem. However, given that $\Sigma = D$ a diagonal matrix, the off-diagonal zero restrictions are sufficient for identification of the Σ and Λ_2 parameters when the number of variables p is greater or equal than two. Further, when normal priors are adopted for F as well as for Λ, proper posteriors for all model parameters are ensured. However, no analytical results exist like in the case of the lower left hand entry of Table 1.

As a summary, we present results in a diagram in Table 4.1. Given that conditional distributions and higher order moments of Λ_2 and F exist, Gibbs sampling is possible for cases I–IV, but only for case IV one obtains meaningful numerical results. The important point is that a normal prior on F and a normal prior on Λ_2 are sufficient to ensure existence of the distributions and their higher order moments.

Table 4.1: **Prior choice and posterior existence**

		Λ_2	
	Prior	Flat	Normal
F	Flat	I No moments (F, Λ_2)	III No moments for (F, Λ_2)
	Normal	II No moments for (F, Λ_2) Marginal moment for F	IV Moments for F and Λ_2, no analytical results

4.4.3 Posterior sampler

Given the conditional densities in section 4.3, a straightforward sampler is the Gibbs sampler. Gibbs sampler iteratively obtaining draws from:

$$p(\Sigma \mid \Lambda, F, Y), \;\; p(\Lambda \mid F, \Sigma, Y), \;\; p(F \mid \Lambda, \Sigma, Y),$$

where the first conditional is a density with an inverse Wishart distribution, and the latter are densities with matricvariate Normal distribution function.

We follow Lopes and West (2004) in specifying proper but vague priors for factors, factor loadings as well as the variance parameters. The unrestricted elements of the factor loading matrix Λ, have independent priors with $\lambda_{ij} \sim N(0, 1)$ for $i \neq j$. The diagonal elements of the loading matrix λ_{ii} for $i = 1, \ldots, r$ have a standard normal distribution truncated above from 0. For the variance parameter, $\sigma_{ii}^2, i = 1, \ldots, p$, we use inverse gamma priors with a shape parameter of $\nu = 5$, and a scale parameter of $\eta = 0.1$. The prior means and variances for the variance parameters based on these values are similar to those of Lopes and West (2004), where a slightly different parametrization is used for the inverse Gamma distribution. We note that the results are not sensitive to small changes in parameters ν and η.

4.5 A Bayesian Factor Model using the Fama-French risk factors

In this section we present a brief analysis of a Bayesian Factor Model using the three Fama and French (1993) equity risk factors as explanatory variables. Our analysis so far was based on an unobserved factor structure. Given that the Fama-French risk factors are "observed" (constructed) we are back in the multivariate regression model and the factor model in (4.5) simplifies to a multivariate linear regression model, where

$$Y = X\Lambda + E, \tag{4.58}$$

where Y is the $T \times p$ matrix of observations, X is the $T \times r$ matrix of observable explanatory variables, E is the $T \times p$ matrix of disturbances and $E \sim MN(0, \Sigma, I_T)$ with a positive definite symmetric matrix Σ. We use the notation X for known factors in order to clarify the distinction between this model and the unobserved factor models in earlier sections.

For the three risk factors, X in equation (4.58) is a $T \times 4$ matrix of a constant in the first column and three observed factors proposed by (Fama & French, 1992) and (Fama & French, 1993). These factors relate to the overall market return (β), firm size (small minus big capitalization), and book to market ratio (high minus low book-to-market ratio).

Using flat priors, the following equations provide the Gibbs sampling steps for parameters Λ and Σ:

$$\Lambda|\Sigma, X, Y \sim MN(\hat{\Lambda}, \Sigma, (X'X)^{-1}) \tag{4.59}$$

$$\Sigma|X, Y \sim IW\left(\left(Y - X\hat{\Lambda}\right)'\left(Y - X\hat{\Lambda}\right), v\right), \tag{4.60}$$

where $\hat{\Lambda} = (X'X')^{-1}X'Y$ is the OLS estimate of Λ and $v = T + p - r + 1$. We refer to (Anderson, 1958) and (Zellner, 1971) for the derivations of these distributions.

In addition, the marginal posterior of the parameters Λ in equation (4.58) is a matric-variate student t density with mean equal to the least squares value:

$$E(\Lambda|X, Y) = \hat{\Lambda} = (X'X)^{-1}X'Y. \tag{4.61}$$

under the condition that $T + p - r > 1$, see (Zellner, 1971) for details.

Given these results, we present two procedures for the computation of the implied momentum strategies. The first is an exact analysis that uses Gibbs sampling and the second one is an approximate analysis using the posterior mean of the residuals. The latter approach is mathematically equivalent to the regression approach but the interpretation of the results is different. It is based on a posterior analysis using one data set and not a frequentist interpretation of the results.

The first approach uses the Gibbs sampling steps in equations (4.59) and (4.60) to obtain draws from the model parameters. Given posterior draws $\Lambda^{(m)}$ for $m = 1, \ldots, M$, the investment strategy at time $t^\star$ is based on the expected "residual returns", equivalently, "excess returns", in the last L periods calculated by

$$\hat{E}_{t^\star} = \frac{1}{ML} \sum_{m=1}^{M} \sum_{l=1}^{L} \left(Y_{t^\star - l} - X_{t^\star - l} \Lambda^{(m)} \right) \tag{4.62}$$

where Y_t and X_t are row-vectors of dependent variable and the explanatory variables at time t. The investment decision is then made based on the ranking of the values of $\hat{E}_{t^\star}$.

The second approach, i.e. the approximate analysis we present, uses the property in equation (4.61) to obtain the expected "residual returns":

$$\hat{E}_{t^\star} = \frac{1}{L} \sum_{l=1}^{L} E(E_{t^\star - l} | Y, X) = \frac{1}{L} \sum_{l=1}^{L} \left(Y_{t^\star - l} - X_{t^\star - l} E(\Lambda | Y, X) \right), \tag{4.63}$$

i.e. the expected values of the posterior residuals equivalent to least squares residuals. Thus the frequentist columns have a Bayesian interpretation under prior ignorance. Similar to the full Bayesian analysis, the investment decision is then made based on the sorted values of $\hat{E}_{t^\star}$.

4.6 Predictive analysis to determine the number of factors

An important decision for the factor model is the number of factors, r, in the above setting. The purpose in this section is to propose a 'predictive likelihood' approach to assess the number of factors in the factor model.

Predictive likelihood in order to determine a varying number of factors The first step in such a predictive approach is to assess the number of factors, as in standard Bayesian model comparison. The gain is the use of 'prior data points' to obtain model probabilities (predictive model probabilities). In this approach, part of the data are used to obtain 'priors', hence model selection or model weights are arguably less sensitive to the prior. In the extreme case of diffuse priors, part of the data selected as the 'training sample' regulates the diffuse priors. Further, predictive likelihoods evaluated at different times (using different training samples) will provide time-varying model probabilities. Any decision making based on these model weights, such as a momentum strategy, will therefore be time-varying as well.

A predictive likelihood approach to obtain model probabilities Let M_r denote the factor model with $r << p$ factors for the factor model in (4.5) and (4.6). A predictive likelihood for model M_r is computed by splitting the dataset as follows:

$$Y = \begin{pmatrix} Y_{t_0:t_1} \\ Y_{t_1+1:t_2} \end{pmatrix} = \begin{pmatrix} Y^\star \\ \tilde{Y} \end{pmatrix} \tag{4.64}$$

where observations from t_0 to t_1 are defined as the 'training sample' and observations from $t_1 + 1$ to t_2 are defined as the 'hold-out sample'.

The predictive likelihood for the hold-out sample is then defined as:

$$p(\tilde{Y}|Y^\star, M_r) = \frac{p(\tilde{Y}, Y^\star|M_r)}{p(Y^\star|M_r)} = \frac{p(Y|M_r)}{p(Y^\star|M_r)} \tag{4.65}$$

Important notes:

1. The choice of the training and hold-out samples is important theoretically, and the results may be sensitive to this choice.

2. The size of the hold-out sample is also important, especially for decision making. Consider a time-varying momentum strategy based on rolling windows. If the hold-out sample is very small, the decision of the momentum strategy will be heavily affected by the training sample and not the hold-out sample (which is the most recent and potentiality informative part of the sample). On the contrary, if the hold-out sample is very large, information from this sample and the training sample are potentially outdated, and momentum strategy will only have slight changes over

time.

Equal training and hold-out sample sizes seem to be an intuitive choice. This way both the ‘data based prior’ and the ‘posterior’ have similar effects on the momentum strategy. Naturally, robustness checks should be performed to see the effect of this training and hold out sample size selection.

Estimating predictive model probabilities A simple method to estimate model probabilities is the harmonic mean estimator. Several alternatives exist, see e.g. Ardia, Baştürk, Hoogerheide, and Van Dijk (2012). The harmonic mean estimator has the advantage that it is easily estimated using $n = 1, ..., N$ parameter draws from the posterior distribution of parameters of model M_r, $\theta_r = \{\Lambda_r, F_r, \Sigma\}$.

Calculation of the predictive likelihoods and predictive model probabilities using the harmonic mean estimator is outlined as follows: Define the set of models, M_r for $r \in 1, \ldots, p-1$.

1. For each model obtain draws from the posterior, obtain N draws of parameters θ_r from the two posterior distributions for $n = 1, \ldots, N$:
   (a) $\theta_r^{(f,n)}$ from $p(\theta_r|Y, M_r)$, draws from the posterior distribution of model M_r given all data points (including training and hold-out sample)
   (b) $\theta_r^{(\star,n)}$ from $p(\theta_r|Y^\star, M_r)$, draws from the posterior distribution of model M_r given the training sample only.
2. Using the harmonic mean estimator, calculate (approximate) two marginal likelihoods for each model m, first for the whole sample, second for the training sample using the exact likelihood:

$$p(Y \mid F, \Lambda, \Sigma) \quad = \quad (2\pi)^{-\frac{Tp}{2}} |\Sigma|^{-\frac{T}{2}} \exp\left(-\frac{1}{2}\,\mathrm{tr}\left(\Sigma^{-1}(Y - F\Lambda)'(Y - F\Lambda)\right)\right) \tag{4.66}$$

   (a) Full sample marginal likelihood

$$p(Y|M_r)^{-1} = \int_{\theta_r} p(Y|\theta_r, M_r)^{-1} p(\theta_r|M_r) d\theta_r \approx \frac{1}{N}\sum_{n=1}^{N} p(Y|\theta_r^{f,n}, M_r)^{-1}$$

   where the likelihood is given in (4.66).

(b) Training sample marginal likelihood

$$p(Y^\star|M_r)^{-1} = \int_{\theta_r} p(Y^\star|\theta_r, M_r)^{-1} p(\theta_r|M_r) d\theta_r \approx \frac{1}{N}\sum_{n=1}^{N} p(Y^\star|\theta_r^{\star,n}, M_r)^{-1}$$

where the likelihood is given in (4.66).

3. Calculate predictive likelihoods for each model M_r using (4.65) and the approximated marginal likelihoods in step (2):

$$p(\tilde{Y}|Y^\star, M_r) = \frac{p(Y|M_r)}{p(Y^\star|M_r)} \approx \frac{\sum_{n=1}^{N} p(Y^\star|\theta_r^{\star,n}, M_r)^{-1}}{\sum_{n=1}^{N} p(Y|\theta_r^{f,n}, M_r)^{-1}} \tag{4.67}$$

4. From the predictive likelihoods for each model in step (3), compute model probabilities for M_r for $r \in 1, \ldots, p-1$:

$$p(M_r|Y) = \frac{p(\tilde{Y}|Y^\star, M_r) \times p(M_r)}{\sum_{r'=1}^{p-1} p(\tilde{Y}|Y^\star, M_{r'}) \times p(M_{r'})},$$

where $p(M_r)$ is the prior model probability. An uninformative prior, such as $p(M_r) = 1/(p-1)$ is intuitive to use in this setting.

Based on the predictive likelihood calculation from a rolling window of predictive likelihoods, an optimal model $M^\star$ for the momentum strategy can be chosen. Based on this model, the momentum strategy can be applied. This will be pursued in the next section.

4.6.1 Simulated data experiments

For illustrative purposes we apply the predictive likelihood approach for the factor model to simulated data. We consider a set simulated datasets with $T = 100$ and $T = 250$ observations. In order to see the effect of parameters p and r on the predictive likelihood methodology, we consider $r = 1, 2$ common factors for $p = 2, 4, 10, 20$ data series. For each simulation experiment, we apply the predictive likelihood approach with different sizes of training samples, consisting of 5%, 10%, 20% and 50% of observations. In addition, we replicate each simulation experiment 100 times to decrease the effect of simulation noise. For all simulations, factors are generated from independent standard normal distributions.

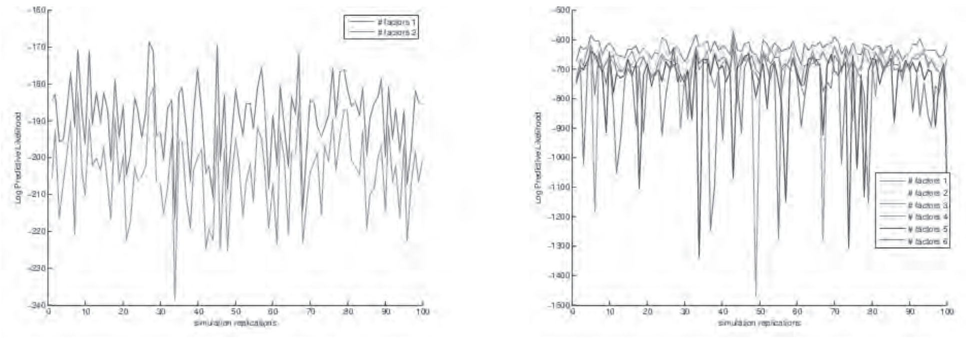
Table 4.1 presents the posterior probabilities from all simulation experiments, where we report the posterior model probabilities for different number of factors averaged over 100 simulation experiments for each simulation setting. Posterior results are based on 4000 posterior draws, where the first 2000 draws are burn-in draws.

Table 4.1: **Posterior probabilities.** Average posterior probabilities from 100 simulation replications with T observations, p variables and r factors. Highest probabilities are indicated by **boldface** table entries.

5% training sample								
T	p	r	pr($r=1$)	pr($r=2$)	pr($r=3$)	pr($r=4$)	pr($r=5$)	pr($r=6$)
100	2	1	**1.00**	0.00	-	-	-	-
100	10	1	**0.96**	0.02	0.00	0.02	0.00	0.00
250	4	2	0.00	**1.00**	0.00	0.00	-	-
250	20	2	0.00	**1.00**	0.00	0.00	0.00	0.00
10% training sample								
T	p	r	pr($r=1$)	pr($r=2$)	pr($r=3$)	pr($r=4$)	pr($r=5$)	pr($r=6$)
100	2	1	**1.00**	0.00	-	-	-	-
100	10	1	**0.89**	0.07	0.00	0.03	0.01	0.00
250	4	2	0.00	**1.00**	0.00	0.00	-	-
250	20	2	0.00	**1.00**	0.00	0.00	0.00	0.00
20% training sample								
T	p	r	pr($r=1$)	pr($r=2$)	pr($r=3$)	pr($r=4$)	pr($r=5$)	pr($r=6$)
100	2	1	**1.00**	0.00	-	-	-	-
100	10	1	**0.77**	0.14	0.05	0.02	0.01	0.00
250	4	2	0.01	**0.98**	0.01	0.00	-	-
250	20	2	0.00	**0.95**	0.01	0.00	0.03	0.01
50% training sample								
T	p	r	pr($r=1$)	pr($r=2$)	pr($r=3$)	pr($r=4$)	pr($r=5$)	pr($r=6$)
100	2	1	**1.00**	0.00	-	-	-	-
100	10	1	**0.34**	0.13	0.02	0.08	0.18	0.24
250	4	2	0.00	**0.95**	0.05	0.00	-	-
250	20	2	0.00	**0.88**	0.02	0.02	0.01	0.07

The results in Table 4.1 indicate that the highest probabilities (indicated by boldface entries) for each simulation experiment, indicated in rows, correspond to the true number of factors. In most simulation studies, the posterior probability is very close to 1 for the correct model specification. Hence the predictive likelihoods provide a clear choice of models. Comparing the bottom panel of Table 4.1 with the other panels, we conclude that the predictive likelihood approach with a smaller training sample than 50% provides more clear indications of the correct number of factors, with posterior probabilities being closer to 1 compared to the same simulation setting but a larger training sample (50%).

Figure 4.1: Log-predictive likelihoods for different number of factors for two sets of simulated data with $T = 100$ observations. The left panel corresponds to $p = 2$ series and $r = 1$ common factor. The right panel corresponds to $p = 10$ series and $r = 1$ common factor. Each simulation experiment is repeated 100 times, as shown in the x-axes. Predictive likelihoods are calculated using 10 percent of the sample as the training sample.



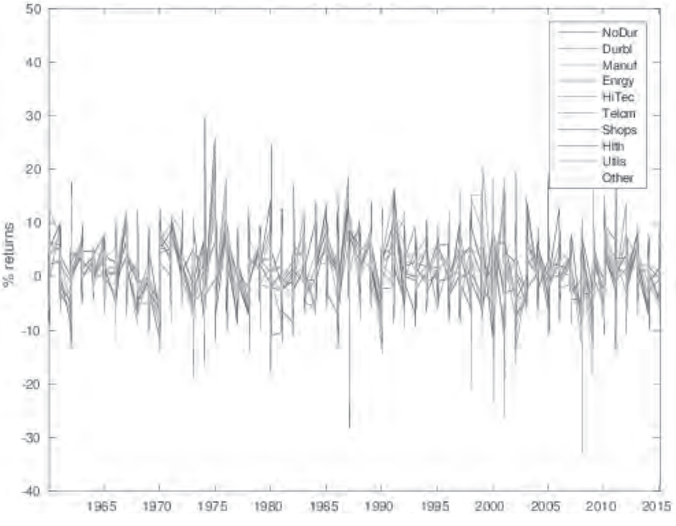
Thus, the length of the training sample should not be chosen too large compared to the total length of the sample and a sensitivity analysis with respect to the length of the training sample will give more confidence in the results.

Figure 4.1 presents the details of predictive likelihoods for two sets of simulated data and for each simulation. These data correspond to $T = 100, p = 2, r = 1$ on the left panel of Figure 4.1 and $T = 100, p = 10, r = 1$ on the right panel of Figure 4.1. For both simulation specifications, the correct number of factors $r = 1$, shown by the red lines in the figure, has the highest posterior probability in almost all simulation replications. We therefore conclude that the predictive likelihood approach accurately detects the number of factors, even with a small sample size.

4.7 Application to ten industry portfolios

In the previous sections a Bayesian approach is outlined to choose the optimal number of time-varying latent factors using the concept of predictive likelihood. In this section we empirically evaluate the impact of modeling dynamic versus static factors as well as latent factors versus Fama and French (1993) factors on the performance of a residual industry momentum strategy. We use return data on ten industry portfolios between 1960M7 and 2015M6. The dataset, shown in Figure 4.1, includes monthly returns from each industry

Figure 4.1: Monthly percentage returns for ten industry portfolios.



portfolio. The 10 industries are labeled as 'non-durables', 'durables', 'manufacturing', 'energy', 'hi-tech', 'telecom', 'shops', 'health', 'utilities' and the final category 'others' [1].

For convenience, we start to present results for a standard momentum strategy applied to the ten industry portfolios. This strategy ranks the industries on their recent total performance, skips a short period of time to overcome short-term return reversals and subsequently buys the industry portfolio ranked first and short-sells the industry portfolio ranked tenth. Jegadeesh and Titman (1993) find that this strategy is profitable when applied to individual stocks and Chan et al. (1996) show that these standard momentum returns cannot be fully explained by the Fama and French (1993) equity risk factors. Moskowitz and Grinblatt (1999) find that the returns to this standard momentum strategy applied to individual stocks are mainly ascribable to momentum in industry portfolios. The empirical results of this standard industry momentum strategy are presented in the top panel of Table 4.1b.

[1] These returns data are retrieved on 24.10.2015 from `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/10_Industry_Portfolios_CSV.zip`. See `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_10_ind_port.html` for definitions. The Fama and French (1993) equity risk factors are also retrieved from this data library, see `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

Standard momentum's risk and return vary over time as a consequence of the time-varying exposures to the Fama and French (1993) equity risk factors, as documented by Grundy and Martin (2001). (Hühn, 2015) confirms this time variation in the exposures for industry momentum. Recently, Blitz et al. (2011) have come up with residual momentum. This strategy ranks stocks on their recent residual returns from the regression of the stock returns on the Fama and French (1993) equity risk factors. By ranking the stocks on their residual returns rather than their total returns Blitz et al. (2011) are able to reduce these time-varying exposures to the equity risk factors.

The purpose of our Bayesian modeling approach is twofold. We investigate the effect on momentum performance of dynamic factor modeling versus static modeling by assessing the effect of moving window estimation instead of estimation with a fixed sample together with assessing gains from using a time-varying number of latent factors compared to a fixed number of latent factors. Second, we compare latent factors versus Fama and French (1993) equity risk factors.

The proposed estimation procedure using a moving window is an additional contribution of our Bayesian estimation of a factor model. In order to assess the effect of this analysis, we estimate two sets of models, namely with a fixed and a moving window estimation sample and evaluate model parameters at each decision time. We apply this approach to the model with latent and the model with Fama and French (1993) factors. Results are presented for the fixed sample case at the top panel of Table 3 and for the moving estimation sample at the bottom panel.

In order to allow for changes in the number of factors, we apply the moving window estimation method as follows. Within one estimation period, the static factor model is estimated for $r = 2, \ldots, 5$ factors, and the optimal number of factors is chosen based on the predictive likelihoods. The moving window estimation is based on a sample of $T = 240$ observations. We consider two cases for the predictive likelihood calculation with a 'small training sample' and a 'large training sample', consisting of 10% and 20% of the full moving window sample, respectively.

Every year in July we rank the ten industry portfolios on the residual returns in the most recent 12 months. These residuals are constructed as follows:

- Sample 4000 matrices of factors F, matrices of factor loadings Λ and matrices Σ for different numbers of factors $r = 2, ..., 5$ using the conditional posterior densities in Section 4.3.

- Use the matrices from the model with the number of factors r with the highest predictive likelihood according to the method described in Section 4.6. Use the first 1000 sampled matrices for the burn-in period and use the last 3000 sampled matrices to construct the average of the current in-sample forecasts of the industry returns $\widehat{Y} = F\Lambda$.

- The residual returns are then $\widehat{E} = Y - \widehat{Y}$.

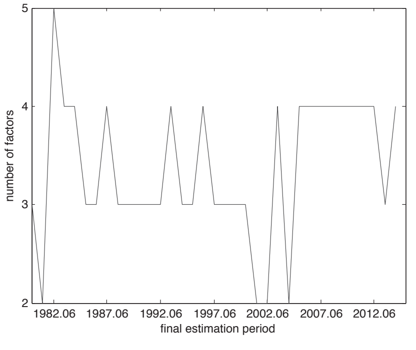
The ten industry portfolios are then sorted on their mean residuals in the last twelve months, where the mean residuals are calculated using the model and the posterior draws using equation (4.62) with $L = 12$ observations. Our residual industry momentum strategy is then long in the industry with the highest residual return and short in the industry with the lowest residual return and holds this position for 12 months. The first investment months is July 1980, since we require 240 months since 1960 for the first estimation of the model parameters.

4.7.1 Optimal number of factors based on moving window estimates and predictive likelihoods

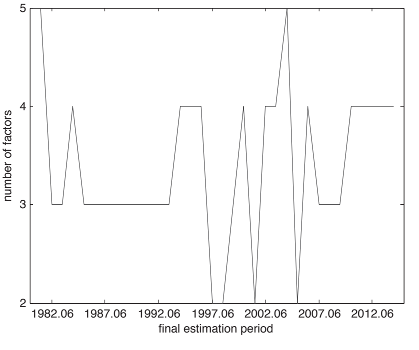
We first analyze the evolution of the number of factors through the moving windows, for two training sample percentages. In order to ensure the existence of posterior densities, a normal prior is used for F and Λ. As stated in Section 4.4, we follow the identification restrictions together with the priors as given in Lopes and West (2004). A sensitivity analysis with respect to the choice of the prior parameters did not affect the results.

Figure 4.2 presents the number of factors with the highest predictive likelihood for each estimation sample under two training sample choices, where 10% and 20% of the total sample is used as training samples. Figure 4.2 shows that the model with $r = 3$ factors is the most frequently chosen model using both training samples. Despite this high frequency, the optimal number of factors according to predictive likelihoods changes substantially over time. The obtained number of factors in the left and right panels of Figure 4.2 are different, in particular, at the end of the sample period. On the one hand, this difference indicates that the training sample choice should be made with care in order

Figure 4.2: Optimal number of factors for 10 industry portfolios. The figure presents the number of factors with highest predictive likelihood at each estimation sample for two training sample choices. X-axes correspond to the final month in the estimation sample.



10% training sample



20% training sample

to find the appropriate number of factors. On the other hand, this variation in the number of factors may influence the gains from a trading strategy, like momentum. For the latter reason, we next report the gains from trading strategies using both training samples.

One would expect that the number of optimal factors varies with the performance of equity markets, in particular fewer factors are present in the model during market declines. We have two major market declines in our sample: 2000-2002 and 2008 where equity markets lost 56 and 38 percent respectively. We indeed find that during the equity market losses in 2000 to 2002 the optimal number of factors was 2. For the 2008 crash we do not find that a smaller number of optimal factors. This may be due to our small sample. The effect of the small sample size on our results is not easy to determine and needs to be explored in further research.

4.7.2 Performance of residual industry momentum strategy

In this section we report the performance of a residual industry momentum strategy which uses the proposed latent factor model with changing number of factors over time. In addition, we compare these results with the standard industry momentum strategy and a set of alternative residual industry momentum strategies. The alternative strategies we consider aim to separate the gains from the three contributions of the proposed latent factor model: defining latent factors instead of fixed, Fama and French (1993), factors, allowing for time-varying number of factors and model estimation using moving estimation samples.

The first three strategies we consider do not use an updated estimation sample to obtain model parameters. These strategies are shown on the top-panel of Table 4.1b. The 'standard strategy' is descriptive and by definition does not depend on an underlying model or an estimation sample. The 'F-F (approximate)' and 'F-F (exact)' strategies denote the factor model with known factors, where the residuals are calculated using (4.62) and (4.63), respectively. The known factors in these models are the three equity risk factors of Fama and French (1993), as explained in Section 4.5. Parameters of each model are estimated only prior to the first investment period, July 1980, over the sample from July 1960 to June 1980.

In addition, we consider a set of residual momentum strategies which are based on moving estimation windows to obtain model parameters and residuals. The results of these strategies, together with the proposed latent factor model with two different training samples are shown in the bottom panel of Table 4.1b. Specifically, we report the results of residual strategies based on the factor models of F-F (approximate) and F-F (exact) using moving estimation windows at each portfolio decision time. The comparison of these results with their fixed estimation sample counterparts highlights the effect of moving window estimation on returns from the residual momentum strategy. Finally, we report the results from a latent factor model, estimated with moving windows but with a fixed number of 3 latent factors. The comparison of this model and the proposed latent factor model provides the gains from allowing the number of factors to vary over time. Finally, the comparison between the latent factor model with three factors and both F-F models provides the gains from latent factor modeling.

In Table 4.1b, we report the following risk and return measures for the returns of each strategy: mean return, volatility, Sharpe ratio, largest loss encountered, maximum drawdown, all measured in percentages, and maximum recovery period measured in months. We stress that these values are based on the realized returns of each investment strategy.

The standard industry momentum strategy does not yield positive average returns; its average annual return is minus 1.1 percent. This empirical result is not in line with the typical positive returns on industry momentum as found by Moskowitz and Grinblatt (1999). The difference stems from, among other things, from the fixed estimation sample, from different sample periods, from using 10 rather than 20 industry portfolios, selecting 1 instead of 3 winners and loser industries, and from the fact that we use 12 monthly returns to rank the industries and the fact that we update the strategy once per year rather than every month. We have chosen to do so, because additional industry portfolios, a longer horizon and more frequent updates of the strategy would substantially increase the computational workload of our model. More generally, the results for the top and bottom panel show the importance of taking account of dynamic factors: Moving window results are all superior to results using a fixed window, except for the length of the recovery period. For the dynamic case the 20 percent Bayesian Factor Model scores better in five of six criteria compared to residual industry momentum strategies using the Fama and French (1993) risk factors.

We conclude that a Bayesian latent factor model with a time-varying number of factors is

Figure 4.3: Mean returns over time

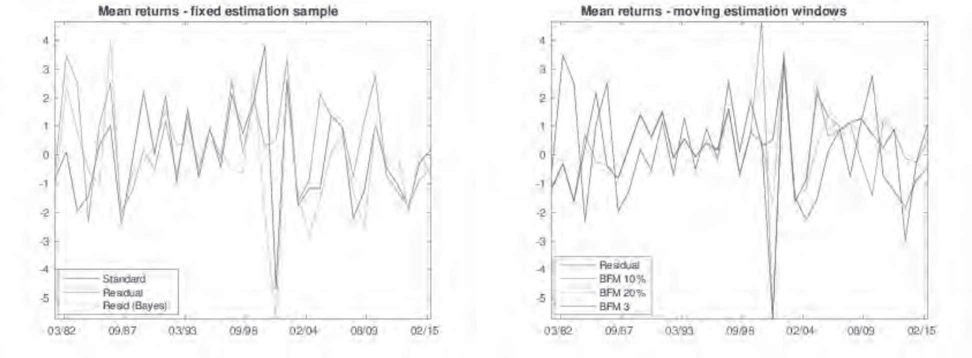


Figure 4.4: Volatility of returns over time

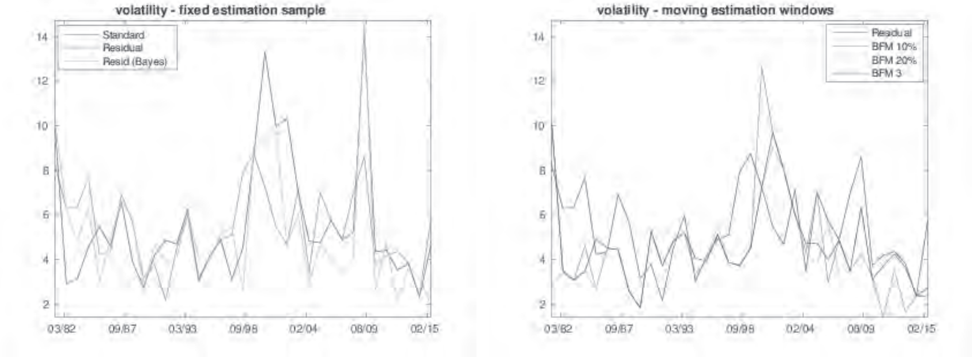


Figure 4.5: Sharpe ratios of returns over time

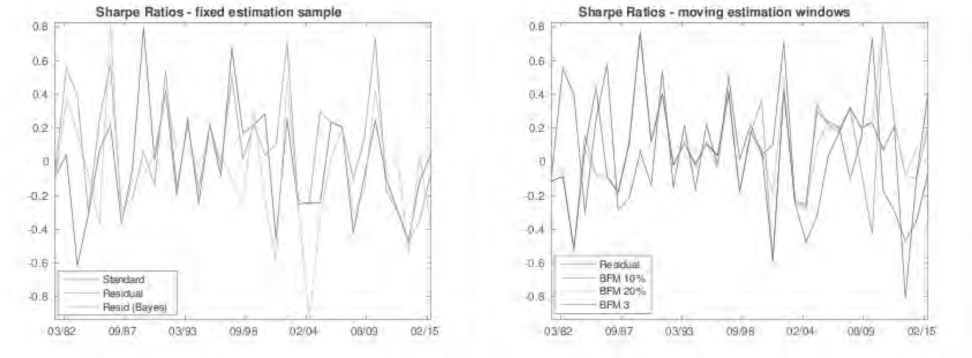


Figure 4.6: Cumulative returns over time

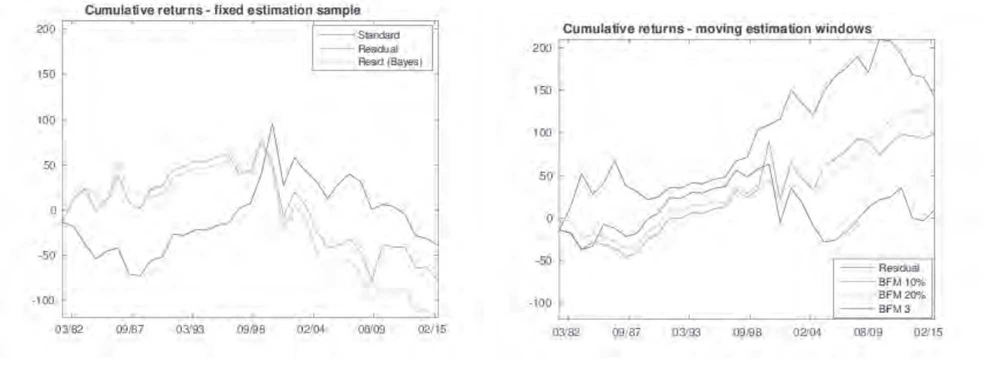


Figure 4.7: Mean returns for two strategies

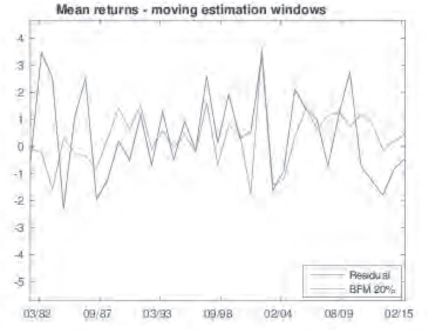


Figure 4.8: Volatility returns for two strategies

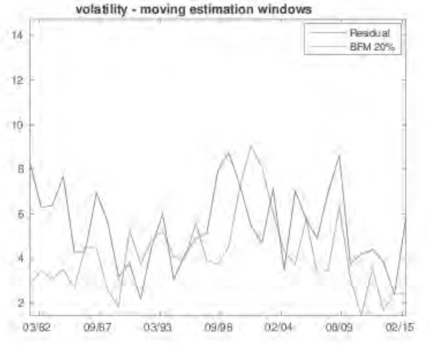


Figure 4.9: Sharpe ratios of returns for two strategies

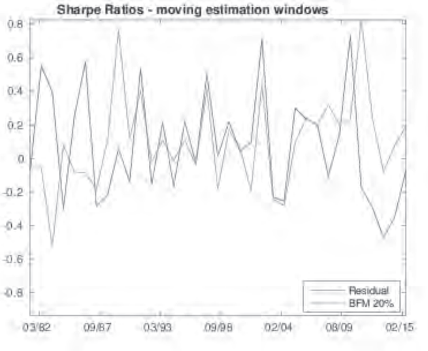


Table 4.1: **Risk and return characteristics of momentum strategies based on different model specifications.**

(a) *Results for strategies with fixed estimation sample*

	Standard	F-F (approx.)	F-F (exact)
mean	-1.13	-2.23	-3.40
volatility	21.54	21.11	21.41
Sharpe ratio	-0.05	-0.11	-0.16
Largest loss	-68.52	-58.06	-68.52
Max. drawdown	135.3	154.88	191.31
Max. recovery	15	16	16

(b) *Results for strategies with moving window estimation*

	F-F	F-F	Latent Factor Models		
	(approx.)	(exact)	10%	20%	3 factors
mean	4.09	3.81	2.82	3.75	0.28
volatility	18.91	20.69	20.16	12.56	19.22
Sharpe ratio	0.22	0.18	0.14	0.30	0.01
Largest loss	-29.05	-46.04	-68.52	-20.95	-68.52
Max. drawdown	66.55	108.12	68.52	35.19	91.01
Max. recovery	11	18	9	9	15

able to outperform the traditional three factor equity risk model in this particular setting for several criteria but not for all. We also emphasize that the Bayesian model with a training sample of 20 percent outperforms in four criteria the model with 10 percent training sample. This may be due to the more complex data structure than the one encountered in the simulation study.

Given the overall results showing that dynamic modeling is important for the performance of residual momentum strategies, we also explore the behaviour over time. Results are shown in Figures 3-9. Figures 3-5 show that in the period considered, namely from July 1980 until June 2015, there are some clear subperiods. In the nineties most procedures did well. The occurrence of the two crashes around 2000 and 2008 have an important influence on the return and volatility and therefore the Sharpe ratio. Apparently, the Bayesian time-varying latent factor model adjusts relatively better to big shocks than the model with the Fama and French (1993) equity risk factors. This result is even more pronounced in the right hand panel of Figure 6. Given that the Bayesian time-varying latent factor model and the Fama and French (1993) equity risk factor model are the best performing strategies, we compare their performance on risk and returns in Figures 7-9. It seems a fair conclusion that in quiet periods the Fama and French (1993) equity risk factor model is good to use and that in more turbulent times the Bayesian time-varying latent factor model performs better. This is particularly shown in Figures 7-9 for the final years of our data. Thus a second conclusion is that a time-varying combination of model structures may give overall superior performance. To explore this topic is outside the scope of the present chapter and left for further research.

4.8 Conclusions

We have specified in this chapter a Bayesian latent factor model with weak prior information that leads to proper posterior distributions of factors and their loadings. Using the predictive likelihood concept a straightforward approach is constructed to account for time variation in the number of factors.

We have applied the Bayesian time-varying latent factor model to residual industry momentum for the period 1980 to 2015 and find that our application outperforms a residual industry momentum strategy based on residual returns from the standard Fama and French (1993) equity risk factors in terms of several risk and return characteristics. Our application outperforms in particular in turbulent times, for example, during the crisis that started in 2008. A possible explanation is that a Bayesian latent factor model adjust more quickly to big shocks than a model with the Fama and French (1993) equity risk factors. We find that the optimal number of latent factors varies substantially over time and that the number of optimal factors indeed decreases when the equity markets experience large losses.

It is also seen that in quiet times like in the nineties, the Fama and French (1993) equity risk factor model is good to use. Thus a major conclusion is that a time-varying combination of model structures and momentum strategies may give overall superior performance.

There is considerable room for extensions of the approach presented. We mention here only two. First, a more formal model with time-varying factors that require filter methods from time series analysis may be very useful; we refer to Billio, Casarin, Ravazzolo, and van Dijk (2013) for more details. See also Casarin, Grassi, Ravazzolo, and Van Dijk (2013) and for a forecasting application see Aastveit, Ravazzolo, and Van Dijk (2014). Secondly, a large panel of individual financial assets, instead of a small number of industry portfolios, that are clustered in smaller batches may yield additional good information and better results. This requires intense and probably parallel computing. Some new approaches are explored in Casarin et al. (2015) for large data sets. These extensions would be of interest to practitioners as well as researchers.

Summary and Conclusions

Summary

This section briefly summarizes in which way we have investigated momentum in this thesis. In Chapter 2 we alter the momentum strategy to improve its performance, while in Chapter 3 we leave the strategy as is, but aim at improving its performance by hedging. In Chapter 4 we develop a Bayesian latent factor model and apply this model to momentum.

In Chapter 2 we apply mean-variance optimization to the equity momentum strategy and compare its performance to other alterations of the momentum strategy. Next to comparing these strategies we test if combining these alterations, including our mean-variance optimization, is able to further improve momentum's performance. We evaluate whether the optimized and other altered momentum strategies as well as their combinations reduce momentum's crash risk and its time-varying risks and returns.

In Chapter 3 we hedge equity momentum's time-varying exposures to the three equity risk factors. We determine the hedge coefficients in two different ways. First, we use the estimated factor loadings of each stock in the momentum strategy at that time to determine momentum's exposures. Secondly, we use a conditional factor model to estimate momentum's exposures using momentum's recent returns. We test whether these hedging strategies reduce momentum's time-varying risks and returns as well as its crash risk. Finally, we investigate whether the bias in the estimated factor loadings used for hedging varies over time.

In Chapter 4 we use a Bayesian latent factor model to investigate the time variation in the comovements of stocks. In particular, we investigate how the optimal number of latent factors varies over time. We determine this optimal number of factors by comparing the predictive likelihoods for models with different numbers of latent factors. Subsequently, we apply the model in a residual industry momentum strategy.

Conclusions

In this section we summarize the main conclusions of this thesis. In the first two essays we use a range of strategies to improve equity momentum's performance and reduce its time-varying risks and returns as well as its crash risk. By doing so, we enhance momentum's challenge for market efficiency as we weaken the crash risk explanation of the positive average returns on the momentum strategy. In the third essay, we find that the comovements of stocks varies over time and that an application to residual industry momentum performs particularly well in turbulent times.

In the first two chapters of this thesis we do not only aim to improve momentum's performance, but also implicitly and explicitly address momentum's time-varying risks. The equity momentum strategy has by construction time-varying risks; when a risk factor return has been negative during momentum's ranking period the strategy is negatively exposed to this particular risk factor and vice versa when the risk factor return has been positive. If the risk factor return then reverses from negative during the ranking period to positive in the investment period, the strategy suffers from its negative exposure to this risk factor. These time-varying exposures do not only cause momentum's time-varying risks, but also cause momentum's crash risk. For example, momentum's largest losses occurred in 1932 and 2009 when the equity markets strongly recovered after their largest losses over the last 100 years. Momentum suffered in these particular years from its negative exposures to the recovering equity markets; the momentum strategy had built up these negative exposures as a consequence of the preceding losses on the equity markets.

In Chapter 3 we conclude that hedging the time-varying exposures of the momentum strategy indeed reduces these exposures and as such also reduces momentum's crash risk. Hedging using factor loadings estimated from a conditional factor model that captures momentum's time-varying exposures turns out to perform better than the more common hedging strategy that uses the estimated factor loadings of the individual stocks. This underperformance of this common hedging strategy based on factor loadings of the individual stocks stems from a systematic bias in the estimation of these factor loadings. This systematic bias also varies with the factor returns during momentum's ranking period.

In Chapter 2 we find that applying standard mean-variance optimization to the equity momentum strategy results in a significant higher Sharpe ratio and reduces momentum's time-varying exposures and crash risk. Although mean-variance optimization is known for its so-called 'error maximization' that results in extreme portfolio weights and bad out-of-sample performance, applying mean-variance optimization does work for equity momentum. We ascribe this to the fact that mean-variance optimization mainly suffers from errors in the estimation of the expected returns, while momentum is quite right about the relative expected returns. We find that the mean-variance optimization benefits from taking into account momentum's top-bottom phenomenon. This is the phenomenon that only the stocks in the top and bottom of momentum's ranking exhibit the momentum effect and out- and underperform, while the stocks in the middle of momentum's ranking do not.

When comparing mean-variance optimization with other improvements of the equity momentum strategy, like residual and constant-risk momentum, we find that residual momentum performs best in terms of the highest Sharpe ratio. The mean-variance optimized momentum strategy turns out to have the least crash risk. Combining mean-variance optimized, constant-risk and residual momentum enables to further improve the Sharpe ratio and further reduce crash risk. This improvement does not come at the cost of a substantial increase in the transaction costs; for moderate transaction costs for which the Sharpe ratios of standard and residual momentum are positive, the Sharpe ratio of the combined strategy is always higher.

Our findings in Chapter 2 hold for the full sample running from July 1929 to December 2014. When looking at different subsamples during this period we find that that the best performing models are different in the various subsamples. Since the oil crisis and the start of the deregulation in 1973 the residual momentum strategy performs best in terms of a higher Sharpe ratio and less crash risk; however, its Sharpe ratio is not significantly higher than those of the combined strategies. With respect to transaction costs, we find that in this subsample mean-variance optimized and combined strategies outperform standard equity momentum up to moderate transaction costs; for relatively high transaction costs standard momentum outperforms the optimized and combined momentum strategies.

In Chapter 4 we find that the comovements of stocks vary over time. This follows from our finding that the optimal number of latent factors in our Bayesian latent factor model varies over time. Applying our Bayesian latent factor model to residual industry momentum turns out to improve the Sharpe ratio and reduce crash risk, in comparison to both standard and residual industry momentum. In this chapter we also find that the best performing model differs from time to time. In relatively quiet periods standard and residual industry momentum perform well, while in turbulent times our Bayesian latent factor model performs better. The flexibility to allow the number of latent factors to vary over time seems valuable in this particular application.

Nederlandse Samenvatting (Summary in Dutch)

Inleiding en motivatie

Aandelen momentum is het fenomeen dat aandelen die recent beter dan andere aandelen hebben gepresteerd, ook de komende maanden weer beter dan andere aandelen zullen presteren. Net zo zullen aandelen die relatief minder hebben gepresteerd, de komende maanden ook relatief minder goed presteren. Jegadeesh and Titman (1993) hebben laten zien dat een strategie die de aandelen koopt die recent relatief goed hebben gepresteerd en de aandelen (ongedekt) verkoopt die relatief slecht hebben gepresteerd gemiddeld een positief rendement oplevert, zonder dat een netto investering vereist is. Academici hebben moeite om dit positieve rendement te verklaren als een vergoeding voor risico en op die manier past momentum niet binnen de efficiënte markthypothese. Nobelprijswinnaar Eugene Fama noemt momentum zelfs de grootste uitdaging van de efficiënte markthypothese.

Momentum's positieve rendementen gaan overigens gepaard met een substantieel risico op een crash. Slechts in enkele maanden kunnen jaren aan positieve rendementen verloren gaan. Naast dit crash risico variëren de rendementen en risico's van de momentum strategie door de tijd, in het bijzonder over de verschillende fases van de conjunctuurcyclus. Deze aspecten maken momentum onaantrekkelijk voor beleggers en zijn voor sommigen het risico waar momentum's gemiddelde positieve rendementen een vergoeding voor ormen; op die manier zou momentum ook binnen de efficiënte markthypothese passen.

In dit proefschrift onderzoeken we verschillende methodes en modellen die het crash risico en de variatie in de rendementen en risico's van de momentum strategie weten te reduceren. Deze strategieën weten eveneens het voor risico gecorrigeerde rendement te verbeteren. Onze resultaten zijn dus interessant voor beleggers en academici; voor laatstgenoemden omdat het reduceren van de risico's de verklaring van momentum als vergoeding voor deze risico's binnen de efficiënte markthypothese verzwakt.

Samenvatting

We onderzoeken aandelen momentum op verschillende manieren in dit proefschrift . In Hoofdstuk 2 vergelijken en combineren we verschillende aanpassingen aan de momentum strategie en in Hoofdstuk 3 laten we de strategie onaangepast, maar dekken we de risico's van de strategie naderhand af. Beide hoofdstukken hebben als doel momentum's crash risico en tijdsvariërende risico's en rendementen te reduceren en tegelijkertijd het positieve rendement te behouden of zelfs te verbeteren. In Hoofdstuk 4 onderzoeken we de tijdsvariatie in de samenhang van aandelenrendementen met een Bayesiaans latent factor model. Dit model passen we vervolgens ook toe in een momentum strategie die gebaseerd is op aandelenportefeuilles van tien verschillende sectoren. Alle hoofdstukken gaan over variatie in rendementen en risico's in het algemeen en voor momentum in het bijzonder.

In Hoofdstuk 3 laten we zien dat het reduceren van momentum's tijdsvariërende blootstellingen aan risicofactoren ook het crash risico en de variatie van de rendementen over de conjunctuurcyclus reduceert, omdat dit crash risico en de variatie van de rendementen juist een gevolg zijn van deze tijdsvariërende blootstellingen. Hoewel de blootstellingen van momentum aan de Fama and French (1993) aandelen risicofactoren gemiddeld ongeveer nul zijn, variëren deze blootstellingen met de rendementen van deze risicofactoren over de recente maanden. Als het rendement op een van deze risicofactoren recent hoog is geweest, dan heeft de momentum strategie een positieve blootstelling aan deze risicofactor, en vice versa. Na bijvoorbeeld een crash van de aandelenmarkten heeft de momentum strategie bijvoorbeeld een negatieve blootstelling aan de aandelenmarkten; als de aandelenmarkten vervolgens herstellen dan verliest momentum aan deze negatieve blootstelling. Momentum's grootste verliezen hebben zich dan ook voorgedaan in juli 1932 en maart 2009 nadat de aandelenmarkten zich sterk herstelden van de twee grootste crashes van de aandelenmarkten in de afgelopen honderd jaar. In het algemeen ondervindt de momentum strategie dus nadeel van omkeringen in het rendement op de risicofactoren; deze omkeringen, zoals na de twee grootste crashes van de aandelenmarkten, vinden veelal plaats (aan het einde) van recessies als aandelenmarkten zich al weer herstellen van hun verliezen, vooruitlopend op herstel in de reële economie. Kortom, het afdekken van deze tijdsvariërende blootstellingen reduceert het crash risico en de variatie van momentum's rendementen over de conjunctuurcyclus.

De standaard aandelen momentum strategie is een relatief simpele strategie die gelijke gewichten toekent aan de aandelen die ten opzichte van andere aandelen recent relatief goed en slecht hebben gepresteerd. Het ligt voor de hand om binnen deze aandelen de aandelen met een lager risico of een hoger verwacht rendement een hoger gewicht toe te kennen, zoals Markowitz (1952) al heeft voorgesteld voor aandelenportefeuilles in het algemeen. Michaud (1989) beschrijft een dergelijke optimalisatiemethode als 'foutenmaximalisatie' en DeMiguel et al. (2009) laten zien dat aandelenportefeuilles met gelijke gewichten geoptimaliseerde portefeuilles juist verslaan. Desondanks blijkt in Hoofdstuk 2 dat optimalisatie van de aandelen momentum strategie wel degelijk de strategie met gelijke gewichten verslaat. Daarbij is het van belang rekening te houden met momentum's 'top-bottom' fenomeen dat niet altijd alle aandelen met een hoger rendement de aandelen met een lager rendement verslaan in de komende maanden; dit geldt voornamelijk alleen voor de aandelen die recent relatief het best hebben gepresteerd. Voor aandelen die recent slecht hebben gepresteerd geldt hetzelfde; alleen aandelen die recent relatief het slechts hebben gepresteerd zullen in de komende periode ook weer slecht presteren. We laten in dit hoofdstuk ook zien dat optimalisatie de tijdsvariërende blootstellingen van momentum reduceert, omdat de optimalisatie al lagere gewichten toekent aan de risicovolle aandelen met hoge blootstellingen. Op die manier reduceert optimalisatie van momentum ook het crash risico en de variatie van de rendementen over de conjunctuurcyclus.

Er bestaan andere aanpassingen aan de momentum strategie die ook tot verbeteringen hebben geleid, zoals residu momentum van Blitz et al. (2011) en constant risico momentum van Barroso and Santa-Clara (2015). Residu momentum sorteert de aandelen op de residuën van de regressie van de rendementen op de Fama and French (1993) risicofactoren in plaats van op de totale rendementen zoals bij standaard momentum. Deze strategie koopt vervolgens de aandelen met de hoogste recente residuële rendementen en verkoopt (ongedekt) de aandelen met de laagste recente residuële rendementen. Constant risico momentum past de blootstelling aan de standaard momentum strategie van tijd tot tijd aan: als de rendementen van momentum recent volatiel zijn geweest dan verlaagt deze strategie de blootstelling aan momentum en vice versa als de rendementen recent minder volatiel zijn geweest. In Hoofdstuk 2 combineren we deze aanpassingen inclusief optimalisatie tot verschillende momentum strategieën. Een strategie die optimalisatie, residu momentum en constant risico momentum combineert leidt tot significant hogere, voor risico gecorrigeerde, rendementen en reduceert het crash risico. Hoewel optimalisatie over het algemeen het aantal transacties sterk doet toenemen, laat deze combinatie strategie het aantal transacties ten opzichte van standaard momentum niet substantieel toenemen. Deze gecombineerde strategie presteert ook beter dan standaard en constant risico momentum in verschillende subperiodes. Sinds 1973 weet residu momentum overigens beter te presteren dan de gecombineerde strategie, hoewel deze verbetering niet significant is.

In dit proefschrift vormt tijdsvariatie een rode draad en in Hoofdstuk 4 benaderen we tijdsvariatie vanuit het perspectief van de samenhang van aandelenrendementen. Door middel van een Bayesiaans latent factor model laten we zien dat het aantal latente factoren dat deze samenhang beschrijft varieert door de tijd. Met dit model adresseren we het probleem in de gebruikelijke modellen dat het aantal latente factoren vast staat; met behulp van een analyse van de voorspelverdelingen zijn we in staat het optimale aantal latente factoren te selecteren. Vervolgens passen we dit model toe in een residu momentum strategie op basis van aandelenportefeuilles van verschillende sectoren. Het Bayesiaans latent factor model is in staat om hogere, voor risico gecorrigeerde, rendementen te realiseren en het crash risico te reduceren ten opzichte van standaard en standaard residu momentum. De prestatie van deze verschillende sectormomentum strategieën varieert eveneens door de tijd. In de relatief rustige periodes in de jaren '90 presteert de standaard residu momentum strategie relatief goed, terwijl in de meer turbulente periodes sinds 2008 de toepassing van het Bayesiaans latent factor model beter presteert; de flexibiliteit om het aantal latente factoren te laten variëren door de tijd blijkt dus waardevol in deze sectormomentum toepassing.

Conclusies

Het reduceren van momentum's crash risico en de variatie van de rendementen en risico's door de tijd in Hoofdstuk 2 en Hoofdstuk 3 verzwakt de verklaring dat momentum een vergoeding voor deze risico's is. Om die reden blijft aandelen momentum een uitdaging vormen voor de efficiënte markthypothese. De samenhang van aandelenrendementen varieert door de tijd en in Hoofdstuk 4 blijkt een Bayesiaans latent factor model deze variatie te beschrijven.

De optimalisatie van momentum en de combinatie van verschillende momentum strategieën in Hoofdstuk 2 weet niet alleen het crash risico te reduceren, maar verbetert ook, statistisch significant, het voor risico gecorrigeerde rendement. De gecombineerde strategie heeft een hoger voor risico gecorrigeerd rendement en minder crash risico dan de afzonderlijke verbeteringen van momentum; deze verbetering gaat niet gepaard met substantieel hogere transactiekosten. Deze conclusies gelden voor de volledige periode van ons onderzoek van juli 1929 tot en met december 2014. Als we uitsluitend naar de periode sinds de oliecrisis in 1973 kijken dan presteert de afzonderlijke residu momentum strategie het best op de genoemde aspecten, maar dit verschil is niet statistisch significant. Afzonderlijke optimalisatie of combinaties presteren wel beter dan standaard momentum over deze periode.

Het direct afdekken van de tijdsvariërende blootstellingen van de standaard momentum strategie in Hoofdstuk 3 reduceert eveneens het crash risico van de momentum strategie en verbetert het voor risico gecorrigeerde rendement. We concluderen dat de gebruikelijke afdekkingsstrategie via de historische blootstellingen van individuele aandelen minder goed presteert dan een afdekkingsstrategie op basis van een conditioneel model op basis van de historische afhankelijkheid van momentum's blootstellingen aan de aandelenrisicofactoren. We schrijven dit toe aan een systematische afwijking van de geschatte historische blootstellingen die eveneens varieert met de recente rendementen op de risicofactoren.

Hoofdstuk 4 laat zien dat het optimale aantal latente factoren in een Bayesiaans latent factor model varieert door de tijd en dit bevestigt dat de samenhang tussen aandelenrendementen eveneens varieert door de tijd. Een toepassing van dit model in een sectormomentum strategie heeft hogere voor risico gecorrigeerde rendementen en lager crash risico dan toepassingen van standaard en residu sectormomentum. Hoewel de traditionele, meer statische, toepassingen van standaard en residu momentum relatief goed presteren in de relatief rustige jaren '90, heeft het meer flexibele Bayesiaans factor model in de turbulente tijden sinds de financiële crisis in 2008 juist toegevoegde waarde.

References

Aastveit, K. A., Ravazzolo, F., & Van Dijk, H. K. (2014). *Combined density nowcasting in an uncertain economic environment* (Tech. Rep. Nos. 14–152/III). Tinbergen Institute. (forthcoming in Journal of Business and Economic Statistics)

Amihud, Y. (2002, January). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, *5*(1).

Anderson, T. (1958). *Multivariate statistical analysis*. New York: John Wiley and Sons.

Anderson, T. (1984). *An introduction to multivariate statistical analysis*. New York: John Wiley & Sons.

Ardia, D., Baştürk, N., Hoogerheide, L. F., & Van Dijk, H. K. (2012). A comparative study of Monte Carlo methods for efficient evaluation of marginal likelihoods. *Computational Statistics and Data Analysis*, *56*(3), 398–414.

Baştürk, N., Grassi, S., Hoogerheide, L. F., Van Dijk, H. K., & Van Oord, A. (2016). *A Bayesian factor model with industry momentum strategies.* (Unpublished manuscript)

Baştürk, N., Hoogerheide, L., Kleijn, R., & Van Dijk, H. K. (2015). *Prior ignorance, likelihood shape and posterior existence in a cointegration model.* (Unpublished manuscript)

Baştürk, N., Hoogerheide, L. F., Van Dijk, H. K., & Van Oord, A. (2016). *Equity momentum strategies at work.* (Unpublished manuscript)

Bai, J., & Ng, S. (2013). Principal components estimation and identification of static factors. *Journal of Econometrics*, *176*(1), 18–29.

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

Bartlett, M. S. (1957). Comment on 'a statistical paradox' by D.V. Lindley. *Biometrika*, *44*(1–2), 533–534.

Baştürk, N., Çakmaklı, C., Ceyhan, S. P., & Van Dijk, H. K. (2014). On the rise of Bayesian econometrics after Cowles Foundation monographs 10, 14. *Œconomia*, *4*, 381-447.

Baştürk, N., Stefano, G., Hoogerheide, L. F., Opschoor, A., & Van Dijk, H. K. (2015). The R-package MitISEM: Efficient and robust simulation procedures for Bayesian inference. *Journal of Statistical Software*. Retrieved from `http://www.tinbergen.nl/discussionpaper/?paper=2466` (forthcoming)

Bauwens, L., Lubrano, M., & Richard, J.-F. (2000). *Bayesian inference in dynamic econometric models*. Oxford University Press.

Billio, M., Casarin, R., Ravazzolo, F., & van Dijk, H. K. (2013). Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics*, *177*(2), 213 - 232. doi: http://dx.doi.org/10.1016/j.jeconom.2013.04.009

Blitz, D., Huij, J., & Martens, M. (2011). Residual momentum. *Journal of Empirical Finance*, *18*, 506–521.

Blitz, D., & Vliet, P. V. (2007, Fall). The volatility effct: Lower risk without lower return. *Journal of Portfolio Management*, *34*(1), 102–113.

Boguth, O., Carlson, M., Fisher, A., & Simutin, M. (2011). Conditional risk and performance evaluation: Volatility timing, overconditioning, and new estimates of momentum alphas. *Journal of Financial Economics*, *102*(2), 363 - 389.

Casarin, R., Grassi, S., Ravazzolo, F., & Van Dijk, H. K. (2013). *Parallel sequential Monte Carlo for efficient density combination: The DeCo MATLAB toolbox* (Tech. Rep. Nos. 13–055/III). Tinbergen Institute. (forthcoming in Journal of Statistical Software)

Casarin, R., Grassi, S., Ravazzolo, F., & Van Dijk, H. K. (2015). *Dynamic predictive density combinations for large data sets in economics and finance* (Tech. Rep. Nos. 15–084/III). Tinbergen Institute.

Chan, L. K. C., Jegadeesh, N., & Lakonishok, J. (1996, December). Momentum strategies. *The Journal of Finance*, *51*(5), 1681–1713.

Chordia, T., & Shivakumar, L. (2002, April). Momentum, business cycle and time-varying expected returns. *The Journal of Finance*, *57*(2), 985 - 1019.

Clarke, R., de Silva, H., & Thorley, S. (2006, Fall). Minimum-variance portfolios in the u.s. equity market. *Journal of Portfolio Management*, *33*(1), 10–24,4.

Cooper, M. J., Jr., R. C. G., & Hameed, A. (2004, June). Market states and momentum. *The Journal of Finance*, *59*(3), 1345–1365.

Daniel, K. D., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *The Journal of Finance*, *53*(6), 1839-1885.

Daniel, K. D., Jagannathan, R., & Kim, S. (2012, June). *Tail risk in momentum strategy returns* (Working Paper No. 18169). National Bureau of Economic Research.

Daniel, K. D., & Moskowitz, T. J. (2013, September). Momentum crashes. *Swiss Finance Institute Research Paper and Columbia Business School Research Paper*.

Daniel, K. D., & Titman, S. (1997, March). Evidence of the characteristics of cross sectional variation in stock returns. *The Journal of Finance*, *52*(1).

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, *22*(5), 1915–1953.

Dickey, J. M. (1967). Matricvariate generalizations of the multivariate t distribution and the inverted multivariate t distribution. *The Annals of Mathematical Statistics*, *38*(2), 511–518.

Efron, B. (1979, January). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, *7*(1), 1–26.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, *47*(2), 427-465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Geweke, J., & Zhou, G. (1996). Measuring the pricing error of the arbitrage pricing theory. *Review of Financial Studies*, *9*(2), 557–587.

Grundy, B. D., & Martin, J. S. (2001, Spring). Understanding the nature of the risks and the source of rewards to momentum investing. *The Review of Financial Studies*, *14*(1), 29–78.

Hobert, J. P., & Casella, G. (1998). Functional compatibility, Markov chains, and Gibbs sampling with improper posteriors. *Journal of Computational and Graphical Statistics*, *7*(1), 42–60.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance*, *54*(6), 2143-2184.

Hühn, H. L. (2015, October). Industry momentum: The role of time-varying factor exposures and market conditions.

Jagannathan, R., & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, *58*(4), 1651-1683.

Jegadeesh, N., & Titman, S. (1993, March). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65–91.

Jegadeesh, N., & Titman, S. (2001, April). Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of Finance*, *56*(2), 699–720.

Jobson, J., & Korkie, B. (1982). Potential performance and tests of portfolio efficiency. *Journal of Financial Economics*, *10*(4), 433–466.

Jr., R. C. G., & Prinsky, C. A. (2007). Momentum, reversal, and the trading behaviors of institutions. *Journal of Financial Markets*, *10*(1), 48 - 75.

Kothari, S., & Shanken, J. (1992). Stock return variation and expected dividends. *Journal of Financial Economics*, *31*(2), 177 - 210.

Kunsch, H. R. (1989, September). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, *17*(3), 1217–1241.

Lawley, D. N., & Maxwell, A. E. (1971). Factor analysis as a statistical method.

Ledoit, O., & Wolf, M. (2003, June). Performance hypothesis testing with the sharpe ratio. *Finance Letters*, *1*(1), 21–23.

Ledoit, O., & Wolf, M. (2004, Fall). Honey, i shrunk the sample covariance matrix. *Journal of Portfolio Management*, *31*(1).

Ledoit, O., & Wolf, M. (2008, December). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, *15*(5), 850–859.

Lehmann, B. N. (1990). Fads, martingales, and market efficiency. *The Quarterly Journal of Economics*, *105*(1), 1-28.

Liu, R. Y., & Singh, K. (1992). Exploring the limits of bootstrap. In R. LePage & L. Billard (Eds.), (pp. 225–248). John Wiley.

Lo, A. W., & MacKinlay, C. (1990). When are contrarion profits due to stock market overreaction. *The Review of Financial Studies*, *3*(2), 175–205.

Lopes, H. F., & West, M. (2004). Bayesian model assessment in factor analysis. *Statistica Sinica*, *14*(1), 41–68.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91.

Martens, M., & Van Oord, A. (2014). Hedging the time-varying risk exposures of momentum returns. *Journal of Empirical Finance*, *28*, 78 - 89.

MichaelDeStefano. (2004). Stock returns and the business cycle. *Financial Review*, *39*(4), 527–547.

Michaud, R. O. (1989). The markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal*, *45*(1), 31–42.

Moskowitz, T. J., & Grinblatt, M. (1999, August). Do industries explain momentum. *The Journal of Finance*, *54*(4), 1249–1290.

Narasimhan Jegadeesh, S. T. (1995). Overreaction, delayed reaction, and contrarian profits. *The Review of Financial Studies*, *8*(4), 973-993.

Pastor, L., & Stambaugh, R. F. (2003, June). Liquidity risk and expected stock returns. *Journal of Political Economy*, *111*(3), 642–685.

Patton, A., Politis, D. N., & White, H. (2009). Correction to automatic block-length selection for the dependent bootstrap by d. politis and h. white. *Econometric Reviews*, *28*(4), 372–375.

Politis, D. N., P.Romano, J., & Wolf, M. (1999). *Subsampling*. Springer.

Politis, D. N., & Romano, J. P. (1992). Exploring the limits of bootstrap. In R. LePage & L. Billard (Eds.), (pp. 263–270). John Wiley.

Politis, D. N., & White, H. (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, *23*(1), 53–70.

Politisand, D. N., & P.Romano, J. (1994, December). The stationary bootstrap. *Journal of the American Statistical Association*, *89*(428), 1303–1313.

Seber, G. (1984). *Multivariate observations*. New York: John Willey & Sons.

Van Oord, A. (2015, September). Optimization of equity momentum: (how) does it work? *Working Paper*. Retrieved from `http://ssrn.com/abstract=2653680`

Zellner, A. (1971). *An introduction to Bayesian inference in econometrics*. John Wiley and Sons, New York. (reprinted in 1996)

Zellner, A., Ando, T., Baştürk, N., Hoogerheide, L. F., & Van Dijk, H. K. (2014). Bayesian analysis of instrumental variable models: Acceptance-rejection within Direct Monte Carlo. *Econometric Reviews*, *33*, 3–35.

About the Author

Arco van Oord (1984) graduated in Quantitative Finance at the Erasmus University Rotterdam in 2006. In this year he also started as a Ph.D. student at the Erasmus Research Institute of Management on 'Active Portfolio Selection with Uncertainty in the Returns and Covariances'. Throughout the years the focus of his research shifted to the application of equity momentum. This has resulted in the publication of Chapter 3 in the *Journal of Empirical Finance*. Arco has presented his research at various international conferences. He has supervised several master students during their financial case studies and when writing their master thesis and has lectured on portfolio management in the master Quantitative Finance at the Erasmus University Rotterdam.

In 2010 Arco joined the expert center on Risk and Asset Liability Management of De Nederlandsche Bank as a supervisor specialist on pension funds and insurance companies. In 2015 he moved to the Supervision Policy Department on Insurance Companies as a policy advisor. During his time at De Nederlandsche Bank Arco has also performed research on the interest rate risk management of pension funds as well as Dutch pension funds' investments costs.

Portfolio

Ph.D. courses

Course	Institution	Grade	ECTS
Real Option Theory	ERIM	8.1	5
The Theory of Corporate Finance	ERIM	9	5
English	ERIM		4
Presentation Skills	ERIM		2
Publishing Strategy	ERIM		2
Continuous Optimization	LNMB	10	6
Computational Econometrics	Tinbergen	9	3
Bayesian Econometrics	ESE	9	5
Quantitative Methods in International Finance and Macroeconomics	ESE	9	5
ESRC Financial Econometrics PhD Training Course	ESRC, Exeter		2
Recent Development of Nonparametric Methods in Financial Econometrics	CASE, Humboldt, Berlin		3

Teaching

- **Portfolio Management (2006-2008):** Lectures as well as assignments. This course is part of the specialization Quantitative Finance within the master Econometrics and Management Science.

- **Financial Case Studies (2007-2009):** Supervision of groups of students working on a financial case in their master Econometrics and Management Science, also in the specialization Quantitative Finance.

  - Forecasting models for oil and energy commodity prices

- Multivariate GARCH models for portfolio Value-at-Risk: Model choice or model combination?
- Predicting Style Factors, Optimal factors for stock selection
- Predicting Stock Returns using Business Cycle Indicators

- **Master thesis (2007-2009):** Supervising students writing their thesis to graduate from the master Econometrics and Management Science in the specialization Quantitative Finance.
  - Subjective views and stress-tests for optimization of credit risk of loan portfolio *(Liudvika Liubeckiene)*
  - Conservative Credits *(Arja de Jager)*
  - Capital Structure Arbitrage: Can Debt-Equity Trading Be Profitable? *(Raisa Velthuis)*
  - Portfolio Construction: Linear versus Quadratic Optimization *(Andries Blom)*

Conferences

The author has presented his research at the following conferences

- **CFE 2008:** 2nd International Conference on Computational and Financial Econometrics, 19-21 June 2008, Neuchatel, Switzerland
- **CFE 2009:** 3rd International Conference on Computational and Financial Econometrics, 28-31 October 2009, Limassol, Cyprus

Other

Other research by the author includes the following working papers

- van Oord, A & Lin, H (2005). Modelling inter- and intraday payment flows. *DNB Working Paper 74*.
- Broeders, D, van Oord, A & Rijsbergen, D (2015). Scale economies in pension fund investments: A dissection of investment costs across asset classes. *DNB Working Paper 474*.

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT (ERIM)

ERIM PH.D. SERIES RESEARCH IN MANAGEMENT

The ERIM PhD Series contains PhD dissertations in the field of Research in Management defended at Erasmus University Rotterdam and supervised by senior researchers affiliated to the Erasmus Research Institute of Management (ERIM). All dissertations in the ERIM PhD Series are available in full text through the ERIM Electronic Series Portal: http://repub.eur.nl/pub. ERIM is the joint research institute of the Rotterdam School of Management (RSM) and the Erasmus School of Economics at the Erasmus University Rotterdam (EUR).

DISSERTATIONS LAST FIVE YEARS

Abbink, E.J., *Crew Management in Passenger Rail Transport*, Promotor(s): Prof.dr. L.G. Kroon & Prof.dr. A.P.M. Wagelmans, EPS-2014-325-LIS, http://repub.eur.nl/pub/76927

Acar, O.A., *Crowdsourcing for Innovation: Unpacking Motivational, Knowledge and Relational Mechanisms of Innovative Behavior in Crowdsourcing Platforms*, Promotor(s): Prof.dr.ir. J.C.M. van den Ende, EPS-2014-321-LIS, http://repub.eur.nl/pub/76076

Akin Ates, M., *Purchasing and Supply Management at the Purchase Category Level: strategy, structure and performance*, Promotor(s): Prof.dr. J.Y.F. Wynstra & Dr. E.M. van Raaij, EPS-2014-300-LIS, http://repub.eur.nl/pub/50283

Akpinar, E., *Consumer Information Sharing*, Promotor(s): Prof.dr.ir. A. Smidts, EPS-2013-297-MKT, http://repub.eur.nl/pub/50140

Alexander, L., *People, Politics, and Innovation: A Process Perspective*, Promotor(s): Prof.dr. H.G. Barkema & Prof.dr. D.L. van Knippenberg, EPS-2014-331-S&E, http://repub.eur.nl/pub/77209

Almeida e Santos Nogueira, R.J. de, *Conditional Density Models Integrating Fuzzy and Probabilistic Representations of Uncertainty*, Promotor(s): Prof.dr.ir. U. Kaymak & Prof.dr. J.M.C. Sousa, EPS-2014-310-LIS, http://repub.eur.nl/pub/51560

Bannouh, K., *Measuring and Forecasting Financial Market Volatility using High-frequency Data*, Promotor(s): Prof.dr. D.J.C. van Dijk, EPS-2013-273-F&A, http://repub.eur.nl/pub/38240

Ben-Menahem, S.M., *Strategic Timing and Proactiveness of Organizations*, Promotor(s): Prof.dr. H.W. Volberda & Prof.dr.ing. F.A.J. van den Bosch, EPS-2013-278-S&E, http://repub.eur.nl/pub/39128

Benning, T.M., *A Consumer Perspective on Flexibility in Health Care: Priority Access Pricing and Customized Care*, Promotor(s): Prof.dr.ir. B.G.C. Dellaert, EPS-2011-241-MKT, http://repub.eur.nl/pub/23670

Benschop, N, *Biases in Project Escalation: Names, frames & construal levels*, Promotors: Prof.dr. K.I.M. Rhode, Prof.dr. H.R. Commandeur, Prof.dr. M.Keil & Dr. A.L.P. Nuijten, EPS-2015-375-S&E, hdl.handle.net/1765/79408

Berg, W.E. van den, *Understanding Salesforce Behavior using Genetic Association Studies*, Promotor(s): Prof.dr. W.J.M.I. Verbeke, EPS-2014-311-MKT, http://repub.eur.nl/pub/51440

Betancourt, N.E., *Typical Atypicality: Formal and Informal Institutional Conformity, Deviance, and Dynamics*, Promotor(s): Prof.dr. B. Krug, EPS-2012-262-ORG, http://repub.eur.nl/pub/32345

Bliek, R. de, *Empirical Studies on the Economic Impact of Trust*, Promotor(s): Prof.dr. J. Veenman & Prof.dr. Ph.H.B.F. Franses, EPS-2015-324-ORG, http://repub.eur.nl/pub/78159

Blitz, D.C., *Benchmarking Benchmarks*, Promotor(s): Prof.dr. A.G.Z. Kemna & Prof.dr. W.F.C. Verschoor, EPS-2011-225-F&A, http://repub.eur.nl/pub/22624

Boons, M., *Working Together Alone in the Online Crowd: The Effects of Social Motivationsand Individual Knowledge Backgrounds on the Participation and Performance of Members of Online Crowdsourcing Platforms*, Promotor(s): Prof.dr. H.G. Barkema & Dr. D.A. Stam, EPS-2014-306-S&E, http://repub.eur.nl/pub/50711

Brazys, J., *Aggregated Marcoeconomic News and Price Discovery*, Promotor(s): Prof.dr. W.F.C. Verschoor, EPS-2015-351-F&A, http://repub.eur.nl/pub/78243

Burger, M.J., *Structure and Cooptition in Urban Networks*, Promotor(s): Prof.dr. G.A. van der Knaap & Prof.dr. H.R. Commandeur, EPS-2011-243-ORG, http://repub.eur.nl/pub/26178

Byington, E., *Exploring Coworker Relationships: Antecedents and Dimensions of Interpersonal Fit,Coworker Satisfaction, and Relational Models*, Promotor(s): Prof.dr. D.L. van Knippenberg, EPS-2013-292-ORG, http://repub.eur.nl/pub/41508

Camacho, N.M., *Health and Marketing: Essays on Physician and Patient Decision-Making*, Promotor(s): Prof.dr. S. Stremersch, EPS-2011-237-MKT, http://repub.eur.nl/pub/23604

Cancurtaran, P., *Essays on Accelerated Product Development*, Promotor(s): Prof.dr. F. Langerak & Prof.dr.ir. G.H. van Bruggen, EPS-2014-317-MKT, http://repub.eur.nl/pub/76074

Caron, E.A.M., *Explanation of Exceptional Values in Multi-dimensional Business Databases*, Promotor(s): Prof.dr.ir. H.A.M. Daniels & Prof.dr. G.W.J. Hendrikse, EPS-2013-296-LIS, http://repub.eur.nl/pub/50005

Carvalho, L. de, *Knowledge Locations in Cities: Emergence and Development Dynamics*, Promotor(s): Prof.dr. L. Berg, EPS-2013-274-S&E, http://repub.eur.nl/pub/38449

Cox, R.H.G.M., *To Own, To Finance, and To Insure - Residential Real Estate Revealed*, Promotor(s): Prof.dr. D. Brounen, EPS-2013-290-F&A, http://repub.eur.nl/pub/40964

Deichmann, D., *Idea Management: Perspectives from Leadership, Learning, and Network Theory*, Promotor(s): Prof.dr.ir. J.C.M. van den Ende, EPS-2012-255-ORG, http://repub.eur.nl/pub/31174

Deng, W., *Social Capital and Diversification of Cooperatives*, Promotor(s): Prof.dr. G.W.J. Hendrikse, EPS-2015-341-ORG, http://repub.eur.nl/pub/77449

Desmet, P.T.M., *In Money we Trust? Trust Repair and the Psychology of Financial Compensations*, Promotor(s): Prof.dr. D. de Cremer, EPS-2011-232-ORG, http://repub.eur.nl/pub/23268

Dollevoet, T.A.B., *Delay Management and Dispatching in Railways*, Promotor(s): Prof.dr. A.P.M. Wagelmans, EPS-2013-272-LIS, http://repub.eur.nl/pub/38241

Doorn, S. van, *Managing Entrepreneurial Orientation*, Promotor(s): Prof.dr. J.J.P. Jansen, Prof.dr.ing. F.A.J. van den Bosch, & Prof.dr. H.W. Volberda, EPS-2012-258-STR, http://repub.eur.nl/pub/32166

Douwens-Zonneveld, M.G., *Animal Spirits and Extreme Confidence: No Guts, No Glory?* Promotor(s): Prof.dr. W.F.C. Verschoor, EPS-2012-257-F&A, http://repub.eur.nl/pub/31914

Duca, E., *The Impact of Investor Demand on Security Offerings*, Promotor(s): Prof.dr. A. de Jong, EPS-2011-240-F&A, http://repub.eur.nl/pub/26041

Duyvesteyn, J.G. *Empirical Studies on Sovereign Fixed Income Markets,* Promotor(s): Prof.dr P.Verwijmeren & Prof.dr. M.P.E. Martens, EPS-2015-361-F&A, hdl.handle.net/1765/79033

Duursema, H., *Strategic Leadership: Moving Beyond the Leader-Follower Dyad*, Promotor(s): Prof.dr. R.J.M. van Tulder, EPS-2013-279-ORG, http://repub.eur.nl/pub/39129

Eck, N.J. van, *Methodological Advances in Bibliometric Mapping of Science*, Promotor(s): Prof.dr.ir. R. Dekker, EPS-2011-247-LIS, http://repub.eur.nl/pub/26509

Elemes, A, *Studies on Determinants and Consequences of Financial Reporting Quality,* Promotor: Prof.dr. E.Peek, EPS-2015-354-F&A, http://hdl.handle.net/1765/79037

Ellen, S. ter, *Measurement, Dynamics, and Implications of Heterogeneous Beliefs in Financial Markets*, Promotor(s): Prof.dr. W.F.C. Verschoor, EPS-2015-343-F&A, http://repub.eur.nl/pub/78191

Eskenazi, P.I., *The Accountable Animal*, Promotor(s): Prof.dr. F.G.H. Hartmann, EPS-2015-355-F&A, http://repub.eur.nl/pub/78300

Essen, M. van, *An Institution-Based View of Ownership*, Promotor(s): Prof.dr. J. van Oosterhout & Prof.dr. G.M.H. Mertens, EPS-2011-226-ORG, http://repub.eur.nl/pub/22643

Evangelidis, I., *Preference Construction under Prominence*, Promotor(s): Prof.dr. S.M.J. van Osselaer, EPS-2015-340-MKT, http://repub.eur.nl/pub/78202

Faber, N., *Structuring Warehouse Management*, Promotor(s): Prof.dr. MB.M. de Koster, Prof.dr. Ale Smidts, EPS-2015-336-LIS, http://repub.eur.nl/pub/78603

Fernald, K., *The Waves of Biotechnological Innovation in Medicine: Interfirm Cooperation Effects and a Venture Capital Perspective,* Promotor(s): Prof.dr. E.Claassen, Prof.dr. H.P.G.Pennings & Prof.dr. H.R. Commandeur, EPS-2015-371-S&E, http://hdl.handle.net/1765/79120

Fourne, S.P., *Managing Organizational Tensions: A Multi-Level Perspective on Exploration, Exploitation and Ambidexterity*, Promotor(s): Prof.dr. J.J.P. Jansen & Prof.dr. S.J. Magala, EPS-2014-318-S&E, http://repub.eur.nl/pub/76075

Gharehgozli, A.H., *Developing New Methods for Efficient Container Stacking Operations*, Promotor(s): Prof.dr.ir. M.B.M. de Koster, EPS-2012-269-LIS, http://repub.eur.nl/pub/37779

Gils, S. van, *Morality in Interactions: On the Display of Moral Behavior by Leaders and Employees*, Promotor(s): Prof.dr. D.L. van Knippenberg, EPS-2012-270-ORG, http://repub.eur.nl/pub/38027

Ginkel-Bieshaar, M.N.G. van, *The Impact of Abstract versus Concrete Product Communications on Consumer Decision-making Processes*, Promotor(s): Prof.dr.ir. B.G.C. Dellaert, EPS-2012-256-MKT, http://repub.eur.nl/pub/31913

Gkougkousi, X., *Empirical Studies in Financial Accounting*, Promotor(s): Prof.dr. G.M.H. Mertens & Prof.dr. E. Peek, EPS-2012-264-F&A, http://repub.eur.nl/pub/37170

Glorie, K.M., *Clearing Barter Exchange Markets: Kidney Exchange and Beyond*, Promotor(s): Prof.dr. A.P.M. Wagelmans & Prof.dr. J.J. van de Klundert, EPS-2014-329-LIS, http://repub.eur.nl/pub/77183

Hekimoglu, M., *Spare Parts Management of Aging Capital Products,* Promotor: Prof.dr.ir. R. Dekker, EPS-2015-368-LIS, http://hdl.handle.net/1765/79092

Heij, C.V., *Innovating beyond Technology. Studies on how management innovation, co-creation and business model innovation contribute to firm's (innovation) performance*, Promotor(s): Prof.dr.ing. F.A.J. van den Bosch & Prof.dr. H.W. Volberda, EPS-2012-370-STR, http://repub.eur.nl/pub/78651

Heyde Fernandes, D. von der, *The Functions and Dysfunctions of Reminders*, Promotor(s): Prof.dr. S.M.J. van Osselaer, EPS-2013-295-MKT, http://repub.eur.nl/pub/41514

Heyden, M.L.M., *Essays on Upper Echelons & Strategic Renewal: A Multilevel Contingency Approach*, Promotor(s): Prof.dr.ing. F.A.J. van den Bosch & Prof.dr. H.W. Volberda, EPS-2012-259-STR, http://repub.eur.nl/pub/32167

Hoever, I.J., *Diversity and Creativity*, Promotor(s): Prof.dr. D.L. van Knippenberg, EPS-2012-267-ORG, http://repub.eur.nl/pub/37392

Hogenboom, A.C., *Sentiment Analysis of Text Guided by Semantics and Structure,* Promotor(s):Prof.dr.ir. U.Kaymak & Prof.dr. F.M.G. de Jong, EPS-2015-369-LIS, http://hdl.handle.net/1765/79034

Hogenboom, F.P., *Automated Detection of Financial Events in News Text*, Promotor(s): Prof.dr.ir. U. Kaymak & Prof.dr. F.M.G. de Jong, EPS-2014-326-LIS, http://repub.eur.nl/pub/77237

Hollen, R.M.A., *Exploratory Studies into Strategies to Enhance Innovation-Driven International Competitiveness in a Port Context: Toward Ambidextrous Ports,* Promotor(s) Prof.dr.ing. F.A.J. Van Den Bosch & Prof.dr. H.W.Volberda, EPS-2015-372-S&E, hdl.handle.net/1765/78881

Hoogendoorn, B., *Social Entrepreneurship in the Modern Economy: Warm Glow, Cold Feet*, Promotor(s): Prof.dr. H.P.G. Pennings & Prof.dr. A.R. Thurik, EPS-2011-246-STR, http://repub.eur.nl/pub/26447

Hoogervorst, N., *On The Psychology of Displaying Ethical Leadership: A Behavioral Ethics Approach*, Promotor(s): Prof.dr. D. de Cremer & Dr. M. van Dijke, EPS-2011-244-ORG, http://repub.eur.nl/pub/26228

Hout, D.H. van, *Measuring Meaningful Differences: Sensory Testing Based Decision Making in an Industrial Context; Applications of Signal Detection Theory and Thurstonian Modelling*, Promotor(s): Prof.dr. P.J.F. Groenen & Prof.dr. G.B. Dijksterhuis, EPS-2014-304-MKT, http://repub.eur.nl/pub/50387

Houwelingen, G.G. van, *Something To Rely On*, Promotor(s): Prof.dr. D. de Cremer & Prof.dr. M.H. van Dijke, EPS-2014-335-ORG, http://repub.eur.nl/pub/77320

Hurk, E. van der, *Passengers, Information, and Disruptions*, Promotor(s): Prof.dr. L.G. Kroon & Prof.mr.dr. P.H.M. Vervest, EPS-2015-345-LIS, http://repub.eur.nl/pub/78275

Hytonen, K.A., *Context Effects in Valuation, Judgment and Choice: A Neuroscientific Approach*, Promotor(s): Prof.dr.ir. A. Smidts, EPS-2011-252-MKT, http://repub.eur.nl/pub/30668

Iseger, P. den, *Fourier and Laplace Transform Inversion with Applications in Finance*, Promotor(s): Prof.dr.ir. R. Dekker, EPS-2014-322-LIS, http://repub.eur.nl/pub/76954

Jaarsveld, W.L. van, *Maintenance Centered Service Parts Inventory Control*, Promotor(s): Prof.dr.ir. R. Dekker, EPS-2013-288-LIS, http://repub.eur.nl/pub/39933

Jalil, M.N., *Customer Information Driven After Sales Service Management: Lessons from Spare Parts Logistics*, Promotor(s): Prof.dr. L.G. Kroon, EPS-2011-222-LIS, http://repub.eur.nl/pub/22156

Kappe, E.R., *The Effectiveness of Pharmaceutical Marketing*, Promotor(s): Prof.dr. S. Stremersch, EPS-2011-239-MKT, http://repub.eur.nl/pub/23610

Karreman, B., *Financial Services and Emerging Markets*, Promotor(s): Prof.dr. G.A. van der Knaap & Prof.dr. H.P.G. Pennings, EPS-2011-223-ORG, http://repub.eur.nl/pub/22280

Khanagha, S., *Dynamic Capabilities for Managing Emerging Technologies*, Promotor(s): Prof.dr. H.W. Volberda, EPS-2014-339-S&E, http://repub.eur.nl/pub/77319

Kil, J., *Acquisitions Through a Behavioral and Real Options Lens*, Promotor(s): Prof.dr. H.T.J. Smit, EPS-2013-298-F&A, http://repub.eur.nl/pub/50142

Klooster, E. van 't, *Travel to Learn: the Influence of Cultural Distance on Competence Development in Educational Travel*, Promotor(s): Prof.dr. F.M. Go & Prof.dr. P.J. van Baalen, EPS-2014-312-MKT, http://repub.eur.nl/pub/51462

Koendjbiharie, S.R., *The Information-Based View on Business Network Performance: Revealing the Performance of Interorganizational Networks*, Promotor(s): Prof.dr.ir. H.W.G.M. van Heck & Prof.mr.dr. P.H.M. Vervest, EPS-2014-315-LIS, http://repub.eur.nl/pub/51751

Koning, M., *The Financial Reporting Environment: The Role of the Media, Regulators and Auditors*, Promotor(s): Prof.dr. G.M.H. Mertens & Prof.dr. P.G.J. Roosenboom, EPS-2014-330-F&A, http://repub.eur.nl/pub/77154

Konter, D.J., *Crossing Borders with HRM: An Inquiry of the Influence of Contextual Differences in the Adoption and Effectiveness of HRM*, Promotor(s): Prof.dr. J. Paauwe & Dr. L.H. Hoeksema, EPS-2014-305-ORG, http://repub.eur.nl/pub/50388

Korkmaz, E., *Bridging Models and Business: Understanding Heterogeneity in Hidden Drivers of Customer Purchase Behavior*, Promotor(s): Prof.dr. S.L. van de Velde & Prof.dr. D. Fok, EPS-2014-316-LIS, http://repub.eur.nl/pub/76008

Kroezen, J.J., *The Renewal of Mature Industries: An Examination of the Revival of the Dutch Beer Brewing Industry*, Promotor(s): Prof.dr. P.P.M.A.R. Heugens, EPS-2014-333-S&E, http://repub.eur.nl/pub/77042

Kysucky, V., *Access to Finance in a Cros-Country Context*, Promotor(s): Prof.dr. L. Norden, EPS-2015-350-F&A, http://repub.eur.nl/pub/78225

Lam, K.Y., *Reliability and Rankings*, Promotor(s): Prof.dr. Ph.H.B.F. Franses, EPS-2011-230-MKT, http://repub.eur.nl/pub/22977

Lander, M.W., *Profits or Professionalism? On Designing Professional Service Firms*, Promotor(s): Prof.dr. J. van Oosterhout & Prof.dr. P.P.M.A.R. Heugens, EPS-2012-253-ORG, http://repub.eur.nl/pub/30682

Langhe, B. de, *Contingencies: Learning Numerical and Emotional Associations in an Uncertain World*, Promotor(s): Prof.dr.ir. B. Wierenga & Prof.dr. S.M.J. van Osselaer, EPS-2011-236-MKT, http://repub.eur.nl/pub/23504

Legault-Tremblay, P.O., Corporate Governance During Market Transition: Heterogeneous responses to Institution Tensions in China, Promotor: Prof.dr. B. Krug, EPS-2015-362-ORG, http://repub.eur.nl/pub/78649

Lenoir, A.S. *Are You Talking to Me? Addressing Consumers in a Globalised World,* Promotor(s) Prof.dr. S. Puntoni & Prof.dr. S.M.J. van Osselaer, EPS-2015-363-MKT, , http://hdl.handle.net/1765/79036

Leunissen, J.M., *All Apologies: On the Willingness of Perpetrators to Apologize*, Promotor(s): Prof.dr. D. de Cremer & Dr. M. van Dijke, EPS-2014-301-ORG, http://repub.eur.nl/pub/50318

Li, D., *Supply Chain Contracting for After-sales Service and Product Support*, Promotor(s): Prof.dr.ir. M.B.M. de Koster, EPS-2015-347-LIS, http://repub.eur.nl/pub/78526

Li, Z., *Irrationality: What, Why and How*, Promotor(s): Prof.dr. H. Bleichrodt, Prof.dr. P.P. Wakker, & Prof.dr. K.I.M. Rohde, EPS-2014-338-MKT, http://repub.eur.nl/pub/77205

Liang, Q.X., *Governance, CEO Identity, and Quality Provision of Farmer Cooperatives*, Promotor(s): Prof.dr. G.W.J. Hendrikse, EPS-2013-281-ORG, http://repub.eur.nl/pub/39253

Liket, K., *Why 'Doing Good' is not Good Enough: Essays on Social Impact Measurement*, Promotor(s): Prof.dr. H.R. Commandeur & Dr. K.E.H. Maas, EPS-2014-307-STR, http://repub.eur.nl/pub/51130

Loos, M.J.H.M. van der, *Molecular Genetics and Hormones: New Frontiers in Entrepreneurship Research*, Promotor(s): Prof.dr. A.R. Thurik, Prof.dr. P.J.F. Groenen & Prof.dr. A. Hofman, EPS-2013-287-S&E, http://repub.eur.nl/pub/40081

Lovric, M., *Behavioral Finance and Agent-Based Artificial Markets*, Promotor(s): Prof.dr. J. Spronk & Prof.dr.ir. U. Kaymak, EPS-2011-229-F&A, http://repub.eur.nl/pub/22814

Lu, Y., *Data-Driven Decision Making in Auction Markets*, Promotor(s): Prof.dr.ir. H.W.G.M. van Heck & Prof.dr. W. Ketter, EPS-2014-314-LIS, http://repub.eur.nl/pub/51543

Manders, B., *Implementation and Impact of ISO 9001*, Promotor(s): Prof.dr. K. Blind, EPS-2014-337-LIS, http://repub.eur.nl/pub/77412

Markwat, T.D., *Extreme Dependence in Asset Markets Around the Globe*, Promotor(s): Prof.dr. D.J.C. van Dijk, EPS-2011-227-F&A, http://repub.eur.nl/pub/22744

Mees, H., *Changing Fortunes: How China's Boom Caused the Financial Crisis*, Promotor(s): Prof.dr. Ph.H.B.F. Franses, EPS-2012-266-MKT, http://repub.eur.nl/pub/34930

Mell, J.N., *Connecting Minds: On The Role of Metaknowledge in Knowledge Coordination*, Promotor: Prof.dr.D.L. van Knippenberg, EPS-2015-359-ORG, http://hdl.handle.net/1765/78951

Meuer, J., *Configurations of Inter-firm Relations in Management Innovation: A Study in China's Biopharmaceutical Industry*, Promotor(s): Prof.dr. B. Krug, EPS-2011-228-ORG, http://repub.eur.nl/pub/22745

Micheli, M.R., *Business Model Innovation: A Journey across Managers' Attention and Inter-Organizational Networks*, Promotor(s): Prof.dr. J.J.P. Jansen, EPS-2015-344-S&E, http://repub.eur.nl/pub/78241

Mihalache, O.R., *Stimulating Firm Innovativeness: Probing the Interrelations between Managerial and Organizational Determinants*, Promotor(s): Prof.dr. J.J.P. Jansen, Prof.dr.ing. F.A.J. van den Bosch, & Prof.dr. H.W. Volberda, EPS-2012-260-S&E, http://repub.eur.nl/pub/32343

Milea, V., *News Analytics for Financial Decision Support*, Promotor(s): Prof.dr.ir. U. Kaymak, EPS-2013-275-LIS, http://repub.eur.nl/pub/38673

Naumovska, I., *Socially Situated Financial Markets: A Neo-Behavioral Perspective on Firms, Investors and Practices*, Promotor(s): Prof.dr. P.P.M.A.R. Heugens & Prof.dr. A. de Jong, EPS-2014-319-S&E, http://repub.eur.nl/pub/76084

Nielsen, L.K., *Rolling Stock Rescheduling in Passenger Railways: Applications in short term planning and in disruption management*, Promotor(s): Prof.dr. L.G. Kroon, EPS-2011-224-LIS, http://repub.eur.nl/pub/22444

Nuijten, A.L.P., *Deaf Effect for Risk Warnings: A Causal Examination applied to Information Systems Projects*, Promotor(s): Prof.dr. G.J. van der Pijl, Prof.dr. H.R. Commandeur & Prof.dr. M. Keil, EPS-2012-263-S&E, http://repub.eur.nl/pub/34928

Osadchiy, S.E., *The Dynamics of Formal Organization: Essays on bureaucracy and formal rules*, Promotor(s): Prof.dr. P.P.M.A.R. Heugens, EPS-2011-231-ORG, http://repub.eur.nl/pub/23250

Ozdemir, M.N., *Project-level Governance, Monetary Incentives, and Performance in Strategic R&D Alliances*, Promotor(s): Prof.dr.ir. J.C.M. van den Ende, EPS-2011-235-LIS, http://repub.eur.nl/pub/23550

Peers, Y., *Econometric Advances in Diffusion Models*, Promotor(s): Prof.dr. Ph.H.B.F. Franses, EPS-2011-251-MKT, http://repub.eur.nl/pub/30586

Peters, M., *Machine Learning Algorithms for Smart Electricity Markets*, Promotor(s): Prof.dr. W. Ketter, EPS-2014-332-LIS, http://repub.eur.nl/pub/77413

Porck, J., *No Team is an Island: An Integrative View of Strategic Consensus between Groups*, Promotor(s): Prof.dr. P.J.F. Groenen & Prof.dr. D.L. van Knippenberg, EPS-2013-299-ORG, http://repub.eur.nl/pub/50141

Porras Prado, M., *The Long and Short Side of Real Estate, Real Estate Stocks, and Equity*, Promotor(s): Prof.dr. M.J.C.M. Verbeek, EPS-2012-254-F&A, http://repub.eur.nl/pub/30848

Poruthiyil, P.V., *Steering Through: How organizations negotiate permanent uncertainty and unresolvable choices*, Promotor(s): Prof.dr. P.P.M.A.R. Heugens & Prof.dr. S.J. Magala, EPS-2011-245-ORG, http://repub.eur.nl/pub/26392

Pourakbar, M., *End-of-Life Inventory Decisions of Service Parts*, Promotor(s): Prof.dr.ir. R. Dekker, EPS-2011-249-LIS, http://repub.eur.nl/pub/30584

Pronker, E.S., *Innovation Paradox in Vaccine Target Selection*, Promotor(s): Prof.dr. H.J.H.M. Claassen & Prof.dr. H.R. Commandeur, EPS-2013-282-S&E, http://repub.eur.nl/pub/39654

Protzner, S. *Mind the gap between demand and supply: A behavioral perspective on demand forecasting,* Promotor(s): Prof.dr. S.L. van de Velde & Dr. L. Rook, EPS-2015-364-LIS, *http*://repub.eur.nl/pub/79355

Pruijssers, J.K., *An Organizational Perspective on Auditor Conduct*, Promotor(s): Prof.dr. J. van Oosterhout & Prof.dr. P.P.M.A.R. Heugens, EPS-2015-342-S&E, http://repub.eur.nl/pub/78192

Retel Helmrich, M.J., *Green Lot-Sizing*, Promotor(s): Prof.dr. A.P.M. Wagelmans, EPS-2013-291-LIS, http://repub.eur.nl/pub/41330

Rietveld, N., *Essays on the Intersection of Economics and Biology*, Promotor(s): Prof.dr. A.R. Thurik, Prof.dr. Ph.D. Koellinger, Prof.dr. P.J.F. Groenen, & Prof.dr. A. Hofman, EPS-2014-320-S&E, http://repub.eur.nl/pub/76907

Rijsenbilt, J.A., *CEO Narcissism: Measurement and Impact*, Promotor(s): Prof.dr. A.G.Z. Kemna & Prof.dr. H.R. Commandeur, EPS-2011-238-STR, http://repub.eur.nl/pub/23554

Rösch, D. *Market Efficiency and Liquidity,* Promotor: Prof.dr. M.A. van Dijk, EPS-2015-353-F&A, http://hdl.handle.net/1765/79121

Roza-van Vuren, M.W., *The Relationship between Offshoring Strategies and Firm Performance: Impact of innovation, absorptive capacity and firm size*, Promotor(s): Prof.dr. H.W. Volberda & Prof.dr.ing. F.A.J. van den Bosch, EPS-2011-214-STR, http://repub.eur.nl/pub/22155

Rubbaniy, G., *Investment Behaviour of Institutional Investors*, Promotor(s): Prof.dr. W.F.C. Verschoor, EPS-2013-284-F&A, http://repub.eur.nl/pub/40068

Schoonees, P. *Methods for Modelling Response Styles*, Promotor: Prof.dr P.J.F. Groenen, EPS-2015-348-MKT, http://repub.eur.nl/pub/79327

Shahzad, K., *Credit Rating Agencies, Financial Regulations and the Capital Markets*, Promotor(s): Prof.dr. G.M.H. Mertens, EPS-2013-283-F&A, http://repub.eur.nl/pub/39655

Sousa, M.J.C. de, *Servant Leadership to the Test: New Perspectives and Insights*, Promotor(s): Prof.dr. D.L. van Knippenberg & Dr. D. van Dierendonck, EPS-2014-313-ORG, http://repub.eur.nl/pub/51537

Spliet, R., *Vehicle Routing with Uncertain Demand*, Promotor(s): Prof.dr.ir. R. Dekker, EPS-2013-293-LIS, http://repub.eur.nl/pub/41513

Staadt, J.L., *Leading Public Housing Organisation in a Problematic Situation: A Critical Soft Systems Methodology Approach*, Promotor(s): Prof.dr. S.J. Magala, EPS-2014-308-ORG, http://repub.eur.nl/pub/50712

Stallen, M., *Social Context Effects on Decision-Making: A Neurobiological Approach*, Promotor(s): Prof.dr.ir. A. Smidts, EPS-2013-285-MKT, http://repub.eur.nl/pub/39931

Tarakci, M., *Behavioral Strategy: Strategic Consensus, Power and Networks*, Promotor(s): Prof.dr. D.L. van Knippenberg & Prof.dr. P.J.F. Groenen, EPS-2013-280-ORG, http://repub.eur.nl/pub/39130

Teixeira de Vasconcelos, M., *Agency Costs, Firm Value, and Corporate Investment*, Promotor(s): Prof.dr. P.G.J. Roosenboom, EPS-2012-265-F&A, http://repub.eur.nl/pub/37265

Troster, C., *Nationality Heterogeneity and Interpersonal Relationships at Work*, Promotor(s): Prof.dr. D.L. van Knippenberg, EPS-2011-233-ORG, http://repub.eur.nl/pub/23298

Tsekouras, D., *No Pain No Gain: The Beneficial Role of Consumer Effort in Decision-Making*, Promotor(s): Prof.dr.ir. B.G.C. Dellaert, EPS-2012-268-MKT, http://repub.eur.nl/pub/37542

Tuijl, E. van, *Upgrading across Organisational and Geographical Configurations*, Promotor(s): Prof.dr. L. van den Berg, EPS-2015-349-S&E, http://repub.eur.nl/pub/78224

Tuncdogan, A., *Decision Making and Behavioral Strategy: The Role of Regulatory Focus in Corporate Innovation Processes*, Promotor(s): Prof.dr.ing. F.A.J. van den Bosch, Prof.dr. H.W. Volberda, & Prof.dr. T.J.M. Mom, EPS-2014-334-S&E, http://repub.eur.nl/pub/76978

Uijl, S. den, *The Emergence of De-facto Standards*, Promotor(s): Prof.dr. K. Blind, EPS-2014-328-LIS, http://repub.eur.nl/pub/77382

Vagias, D., *Liquidity, Investors and International Capital Markets*, Promotor(s): Prof.dr. M.A. van Dijk, EPS-2013-294-F&A, http://repub.eur.nl/pub/41511

Veelenturf, L.P., *Disruption Management in Passenger Railways: Models for Timetable, Rolling Stock and Crew Rescheduling*, Promotor(s): Prof.dr. L.G. Kroon, EPS-2014-327-LIS, http://repub.eur.nl/pub/77155

Venus, M., *Demystifying Visionary Leadership: In search of the essence of effective vision communication*, Promotor(s): Prof.dr. D.L. van Knippenberg, EPS-2013-289-ORG, http://repub.eur.nl/pub/40079

Vermeer, W., *Propagation in Networks:The impact of information processing at the actor level on system-wide propagation dynamics,* Promotor: Prof.mr.dr. P.H.M.Vervest, EPS-2015-373-LIS, http://repub.eur.nl/pub/79325

Visser, V.A., *Leader Affect and Leadership Effectiveness: How leader affective displays influence follower outcomes*, Promotor(s): Prof.dr. D.L. van Knippenberg, EPS-2013-286-ORG, http://repub.eur.nl/pub/40076

Vlam, A.J., *Customer First? The Relationship between Advisors and Consumers of Financial Products*, Promotor(s): Prof.dr. Ph.H.B.F. Franses, EPS-2011-250-MKT, http://repub.eur.nl/pub/30585

Waltman, L., *Computational and Game-Theoretic Approaches for Modeling Bounded Rationality*, Promotor(s): Prof.dr.ir. R. Dekker & Prof.dr.ir. U. Kaymak, EPS-2011-248-LIS, http://repub.eur.nl/pub/26564

Wang, T., *Essays in Banking and Corporate Finance*, Promotor(s): Prof.dr. L. Norden & Prof.dr. P.G.J. Roosenboom, EPS-2015-352-F&A, http://repub.eur.nl/pub/78301

Wang, Y., *Information Content of Mutual Fund Portfolio Disclosure*, Promotor(s): Prof.dr. M.J.C.M. Verbeek, EPS-2011-242-F&A, http://repub.eur.nl/pub/26066

Wang, Y., *Corporate Reputation Management: Reaching Out to Financial Stakeholders*, Promotor(s): Prof.dr. C.B.M. van Riel, EPS-2013-271-ORG, http://repub.eur.nl/pub/38675

Weenen, T.C., *On the Origin and Development of the Medical Nutrition Industry*, Promotor(s): Prof.dr. H.R. Commandeur & Prof.dr. H.J.H.M. Claassen, EPS-2014-309-S&E, http://repub.eur.nl/pub/51134

Wolfswinkel, M., *Corporate Governance, Firm Risk and Shareholder Value*, Promotor(s): Prof.dr. A. de Jong, EPS-2013-277-F&A, http://repub.eur.nl/pub/39127

Yang, S., *Information Aggregation Efficiency of Prediction Markets*, Promotor(s): Prof.dr.ir. H.W.G.M. van Heck, EPS-2014-323-LIS, http://repub.eur.nl/pub/77184

Zaerpour, N., *Efficient Management of Compact Storage Systems*, Promotor(s): Prof.dr.ir. M.B.M. de Koster, EPS-2013-276-LIS, http://repub.eur.nl/pub/38766

Zhang, D., *Essays in Executive Compensation*, Promotor(s): Prof.dr. I. Dittmann, EPS-2012-261-F&A, http://repub.eur.nl/pub/32344

Zwan, P.W. van der, *The Entrepreneurial Process: An International Analysis of Entry and Exit*, Promotor(s): Prof.dr. A.R. Thurik & Prof.dr. P.J.F. Groenen, EPS-2011-234-ORG, http://repub.eur.nl/pub/23422

ESSAYS ON MOMENTUM STRATEGIES IN FINANCE

This thesis discusses several aspects and possible improvements of equity momentum strategies in finance. Equity momentum is the phenomenon that stocks that have recently outperformed continue to outperform, while underperformers will continue to underperform. Equity strategies that exploit this phenomenon by buying the recent outperformers and short-selling the recent underperformers have proven to be profitable for investors. In his Nobel prize lecture in 2013 Eugene Fama referred to this performance of the momentum strategy as being the biggest challenge for the efficient market hypothesis.

Nevertheless, equity momentum is also known for its crash risk, wiping out years of average positive returns in just a few months, and the fact that its risk and returns vary over time. In this thesis different hedging strategies are applied to reduce momentum's crash risk and time varying exposures without reducing its positive average returns. Furthermore, different recent improvements of momentum are combined in a mean-variance optimization set-up. Optimization also reduces momentum's crash risk and its time varying exposures. Moreover it improves momentum's Sharpe ratio for moderate transaction costs.

Finally, this thesis addresses momentum's time varying risks and returns in a different way. A Bayesian latent factor model where the number of latent factors is allowed to vary over time is derived. Using the predictive likelihood approach this model is then applied to a residual industry momentum strategy. In turbulent times, like the crisis that started in 2008, the Bayesian latent factor model performs well in terms of risk and return characteristics.

ERiM

The Erasmus Research Institute of Management (ERIM) is the Research School (Onderzoekschool) in the field of management of the Erasmus University Rotterdam. The founding participants of ERIM are the Rotterdam School of Management (RSM), and the Erasmus School of Economics (ESE). ERIM was founded in 1999 and is officially accredited by the Royal Netherlands Academy of Arts and Sciences (KNAW). The research undertaken by ERIM is focused on the management of the firm in its environment, its intra- and interfirm relations, and its business processes in their interdependent connections.

The objective of ERIM is to carry out first rate research in management, and to offer an advanced doctoral programme in Research in Management. Within ERIM, over three hundred senior researchers and PhD candidates are active in the different research programmes. From a variety of academic backgrounds and expertises, the ERIM community is united in striving for excellence and working at the forefront of creating new business knowledge.

Erasmus Research Institute of Management - ERiM

Design & layout: B&T Ontwerp en advies (www.b-en-t.nl) Print: Haveka (www.haveka.nl)

ERIM PhD Series
Research in Management

Erasmus Research Institute of Management - ERiM
Rotterdam School of Management (RSM)
Erasmus School of Economics (ESE)
Erasmus University Rotterdam (EUR)
P.O. Box 1738, 3000 DR Rotterdam,
The Netherlands

Tel. +31 10 408 11 82
Fax +31 10 408 96 40
E-mail info@erim.eur.nl
Internet www.erim.eur.nl