Improving Time Series and Cross-Sectional Momentum Trading Strategies using Stochastic Programming

by

Xiaoshi Guo

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Industrial Engineering

Program of Study Committee:
Sarah M. Ryan, Major Professor
Cameron MacKenzie
K. Jo Min
Farzad Sabzikar
Cindy L Yu

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation/thesis. The Graduate College will ensure this dissertation/thesis is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2021

DEDICATION

I would like to dedicate this thesis to my husband, my parents, and my grandparents for their unconditional love and support.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to take this opportunity to express my gratitude to those who helped me with conducting research to make this work possible.

First and foremost, I would like to thank my advisor, Dr. Sarah M. Ryan for her continuous guidance, patience, and insightful comments in the last four years. Her rigorous attitude towards research, wisdom, and diligence is always a good example for me to follow.

I would like to express my gratefulness to my committee members, Dr. Cameron MacKenzie, Dr. K. Jo Min, Dr. Farzad Sabzikar, and Dr. Cindy L Yu, for their valuable time and feedback on this dissertation. Their suggestions from different domains helped me open my mind and think about my topic from various perspectives. Many thanks to the faculty and staff from the IMSE Department and my friends for bringing me a fulfilled and colorful life at Iowa State University. My sincere thanks also go to authors I consulted, such as Dr. David Woodruff, Dr. Juan Pablo Contreras, Dr. Taylor Arnold, and Dr. John W. Emerson, for their patient guidance.

In the end, I would like to give my deepest thanks to my husband, my parents, and my grandparents for their unconditional love. It is their love and support that motivates me to keep moving forward and encourages me to complete this four-year PhD program.

ABSTRACT

For nearly three decades, empirical studies have shown that momentum trading strategies, which rely on return predictability based on past returns, have produced stable profits across multiple periods, for diverse asset and asset classes, and in various financial markets. These momentum strategies can be categorized as time-series momentum or cross-sectional momentum strategies. This dissertation consists of three papers discussing some our improvements on the time series and the cross-sectional momentum strategies using stochastic portfolio optimization.

To improve on the time series momentum strategy using stochastic programming, in the first paper, we construct a model of market index returns using a variant of the geometric Brownian motion process. This model is reduced to a Gaussian process with time-dependent drift capturing and reflecting the time series momentum and constant volatility. According to this model, we test three schemes for updating time series momentum estimates on a rolling basis for scenario generation. They differ in the frequency, either daily or monthly, to update the momentum estimation and the approach, either simple moving average or exponential moving average, to estimate momentum. The performance of these schemes is evaluated by the mass transportation distance (MTD) rank histograms. By carrying out computational studies on the S&P 500 and the FTSE 100 indices, we find: 1) The approach to estimating time series momentum can greatly influence the quality of the generation procedure while the frequency with which to update the estimate only has trivial effects; 2) All schemes are highly sensitive to the estimated volatility of the index returns; and 3) Among the three processes tested, the algorithm that estimates time series momentum according to an exponential moving average can generate reliable scenarios when the volatility estimation error is small.

Inspired by the findings in our first paper, we modify the model of market index returns by considering additional moving average techniques to estimate time series momentum and using time-dependent volatility to capture the most recent change in market index volatility over time as well as to match the procedure of estimating the time-dependent drift. The scenarios generated

using this model are input into three alternative two-stage stochastic optimization models, time series momentum strategy controlling downside risk (TSMDR), mean-CVaR, and mean-variance, to derive the strategy for continually rebalancing a portfolio of a risk-free asset and a risky asset. By backtesting all three strategies, along with the traditional time series momentum strategy, using S&P 500 index returns and 1-month Treasury bill rates during 2001-2019, we validate our stochastic model of index returns. Furthermore, our newly created TSMDR strategy consistently provides higher excess returns and less investment risk than either the traditional time series momentum strategy or the other strategies derived from solutions for the mean-risk stochastic programs. By decomposing all strategies, we find the outperformance of TSMDR occurs because 1) weighted moving average of recent returns can better forecast the trend of the stock market than the past 12-month excess return, 2) risk-parity strategies have less investment risk whereas mean-risk strategies generally have higher returns, and 3) using CVaR can control investment risk better than using variance can.

In the third paper, we discuss our improvement made on the cross-sectional momentum strategy. The performance of the remarkable cross-sectional momentum strategy has been deteriorated by producing persistent negative returns in 1932 and 2009, phenomena widely known as momentum crashes. The occurrence of momentum crashes is due to the fact that the investors' risk attitudes and the momentum portfolio return for the next period are hard to predict when the stock market is highly volatile and in a recovery phase. To avoid such crashes, we develop a two-stage stochastic mean-CVaR optimization model with time-varying risk aversion to optimize the weight on the momentum portfolio for a trade-off between maximizing the expected return of the momentum portfolios and minimizing the expected loss of momentum investing under the extreme cases. Three volatility-related measures, 1) the ex-ante market volatility, 2) the ratio of the ex-ante market return to the ex-ante market volatility, and 3) the Chicago Board Options Exchange Volatility Index, are utilized to gauge the market condition in the next period and automate the change in risk aversion accordingly. Scenarios of momentum portfolio returns are generated using a combination of heuristic and optimization moment-matching algorithms to capture the statistical properties of momentum portfolio returns in the next period. The reliability of the scenarios is evaluated through the MTD rank histograms. By backtesting using total returns of NYSE, NASDAQ, and AMEX equities, we find our model with time-varying risk parameters automated using ex-ante volatility provides not

only higher returns but also less investment risk compared to a cross-sectional momentum heuristic. Most importantly, our model completely avoids momentum crashes by avoiding drastic loss in all market conditions from January 1926 to December 2020.

CHAPTER 1. INTRODUCTION

1.1 Background

Asset allocation is a trading strategy that aims to balance risk and reward by adjusting the percentage of wealth invested in each asset in an investment portfolio according to an individual's goals, risk tolerance, and investment horizon. To achieve the individual's goal to gain profits and mitigate investment risk, investors need to determine whether to long (buy a security with the anticipation it will increase in value) or short (sell a security first with the plan to buy it later at a lower price) assets from time to time in the investment horizon. In this dissertation, we focus our discussion on two types of momentum trading strategies to help investors make asset allocation decisions more wisely.

Momentum in finance is a phenomenon that assets trending strongly in a certain direction will continue to move in the same direction in the short term. This concept of momentum is based on a similar observation in physics that an object will continue moving at its current velocity until some force causes its speed or direction to change. To calculate momentum of each asset, the rate of change on price movements is measured to help investors obtain a quantitative measure of the strength of the trend for this asset. The momentum effect is considered to be a market anomaly as it contradicts the efficient market hypothesis that past returns cannot determine the current return. Investors employ this return predictability of momentum to construct trading strategies, such as the cross-sectional (CS) momentum strategy (Jegadeesh and Titman, 1993) and the time series (TS) momentum strategy (Moskowitz et al., 2012).

The CS strategy ranks the cross-section of securities according to the three- to 12-month past returns to identify winners (securities that have performed well) and losers (securities that have performed poorly), then buy past winners and short past losers. Jegadeesh and Titman (1993) show that the CS strategy can achieve a compounded excess return of 12.01% per year on average and its Sharpe ratio, which is defined as the difference between the returns of the investment and the risk-free return divided by the standard deviation of the investment return and represents the additional

amount of return that an investor receives per unit of increase in risk (Sharpe, 1994), exceeds that of the US stock market from 1965 to 1989. Subsequently, Jegadeesh and Titman (2001) and Israel and Moskowitz (2013) document the continuing efficacy of US equity momentum portfolios in common stock returns during the years 1927 – 1965 and 1990 – 2012. Furthermore, plentiful research demonstrates that cross-sectional momentum is not just a puzzling phenomenon in the US stock market but a worldwide abnormality. By applying the CS strategy in the European stock market, Rouwenhorst (1998) finds that this strategy works well on both individual country and cross-national basis, which adds new evidence to support momentum strategy. Most importantly, momentum exhibits its existence in assets other than stocks. Country indices (Asness et al., 1997), developed and emerging equity markets (Rouwenhorst, 1998), industry portfolios (Moskowitz and Grinblatt, 1999), currencies (Okunev and White, 2003), exchange traded futures contract (Moskowitz et al., 2012), and bonds (Asness et al., 2012) are examples to illustrate the existence of momentum. The fact that the CS strategy has proved robust across time, countries and asset classes has made Carhart (1997) add cross-sectional momentum as the fourth factor in addition to the well-known Fama and French three-factor asset pricing model (Fama and French, 1993) and led Fama (1998) to observe that momentum remains the premier unexplained anomaly.

Different from the CS strategy, the TS strategy makes the long/short decision according to each security's own past performance during the ranking period rather than its relative rank in the security universe. Moskowitz et al. (2012) and Asness et al. (2012) apply the TS strategy on the data covering the period January 1986 to December 2009 across diverse asset classes, including equity index, currency, commodity, and bond futures. They both find that the TS strategy exhibits its profitability in all asset classes tested, has small loadings on standard risk factors, and performs extraordinarily well in extreme periods.

1.2 Motivation and Problem Statement

Even though both CS and TS strategies are shown to be profitable in various markets, in multiple time periods, and across numerous types of assets and asset classes, there are still additional improvements to be discovered. For example, the remarkable performance of the CS strategy suffers from occasionally large crashes in which past losers stop being losers and outperform past winners in panic states of the volatile recovery market, while the TS strategy exhibits its best performance

in the extreme market but mediocre performance in other cases. The occurrence of these issues is mostly due to two reasons. First, both trading strategies simply use the momentum, either CS or TS, as a trading signal, which may not be an optimal trading signal. In other words, neither of these trading strategies is optimized. Second, they do not explicitly account for any uncertainty of the assets' future returns, and ignoring uncertainty can lead to wrong decisions since every decision in the momentum strategies is made for a single predicted circumstance which may not be realized. These two issues can be effectively addressed by stochastic optimization through its flexibility in dealing with uncertainty.

Stochastic programming provides a framework for modeling optimization problems that involve uncertainty. Stochastic portfolio optimization models have four major components:

- Decision variables represent the weights on different assets in the portfolio;
- The objective function represents a quantity to be maximized or minimized such as future wealth, the returns generated by the portfolio in each period, and/or some measure of investment risk;
- Constraints express limits on acceptable investment behavior and tolerable levels of risk while describing the links between decisions in one period and available funds to invest in the next;
- Probability distributions for uncertain parameters represent the uncertainty of the assets' future performance.

These components are all incorporated in the deterministic optimization, except for the uncertain parameters, which make stochastic programs more attractive for modelling practical financial planning problems. In stochastic programs, the uncertain parameters are flexibly characterized by scenarios, which discretize or approximate the stochastic process governing the underlying uncertainty so that the stochastic program can be solved by accounting for all scenarios and their probabilities of occurrence. Using probabilistic scenarios to represent the continuous-valued process of random parameters allows the model to be freed from rigid assumptions of the distribution of the random parameters, and thus, motivates the creation of various scenario generation approaches (Yu et al., 2003; Kaut and Wallace, 2007; Roman et al., 2010; Gülten and Ruszczyński, 2015). In finance, the uncertain parameters often follow asymmetric (or skewed) heavy-tailed distributions,

which can be well accommodated by scenario-based models (Topaloglou et al., 2008). Besides the flexibility of assumptions for the distribution of random parameters, stochastic programming is well suited for portfolio management for another reason. In real finance world, the investing decisions have to be made before the information about asset future performance is revealed. However, the stock market changes continuously, making it hard for investors to predict what will happen in the next second and thus difficult to respond correctly to the unknown change. Stochastic programming takes advantage of the scenario-based uncertainty to optimize for more than one of the possibilities that could happen in the future, allowing investors to obtain a better trading strategy by thinking from a more comprehensive perspective in a world full of uncertainty.

In general, our goal in this dissertation is to improve on the momentum strategies by stochastic programming. To achieve it, we need to answer the following questions. At the very beginning, we should consider what decision to make, what our objective is, how to measure the investment risk, what constraints to set, and what the uncertain parameters are in the stochastic optimization model so that the momentum trading strategies can be viewed as feasible, but not necessarily optimal, solutions. With these questions answered, how to generate scenarios for the uncertain parameters becomes our next problem. Upon scenario generation, the quality of scenarios should be evaluated to ensure the scenarios will result in a trading strategy which is not only optimal to the stochastic program itself but also in the proximity of the optimal trading strategy in reality. Therefore, how to evaluate the quality of scenarios will be another critical issue to solve. Finally, the selection of performance measures is always important as it provides quantitative criteria to judge whether the optimal strategies derived by our stochastic programming models have improved the performance of the traditional CS and TS strategies from various perspectives.

1.3 Dissertation Structure

Three papers are provided in this dissertation, one in each chapter. Chapters 2 and 3 combined enhance the performance of the TS strategy and Chapter 4 endeavors to avoid the momentum crashes that have been observed in the CS strategy. The remainder of the dissertation is organized as follows.

Chapter 2 (Guo and Ryan, 2020) focuses on the discussion of the scenario generation of the uncertain parameter, market index returns, which will serve as input of the stochastic portfolio op-

timization model, using time-series momentum. In this chapter, we develop a variant of geometric Brownian motion process with a nonstationary mean incorporating time series momentum to estimate the distribution of the future rate of return. Based on this model, three different procedures for generating scenarios on a rolling basis are devised, that differ according to how frequently the momentum parameter is updated and whether it is estimated according to a simple or an exponential moving average of returns. Since these continuous process models for the uncertain parameter make it challenging to optimize directly, Monte Carlo simulation is invoked to discretize the continuous process and generate scenarios for the index rate of return. To investigate which generation procedure can provide reliable scenarios that cannot be distinguished from the corresponding observations, the reliability of scenario sets generated for multiple instances in a historical backtest is assessed by applying a recently developed statistical tool, the mass transportation distance rank histogram. Case studies of the Standard Poor's 500 Index and the Financial Times Stock Exchange 100 Index for two different historical periods are conducted.

Chapter 3 continues the discussion of scenario generation in Chapter 2 and develops multiple trading strategies using stochastic optimization with generated scenarios as input. First, we consider additional moving average techniques to estimate time series momentum, including simple, exponential and weighted moving average, and find that weighted moving average can most accurately capture the trend of stock market. Second, the volatility in the geometric Brownian motion process of index returns is estimated on a rolling basis and no longer a constant, which greatly reduces the volatility estimation error mentioned in Chapter 2. To be consistent with the traditional TS strategy, all investment strategies are constructed to rebalance a portfolio of a risky asset and a risk-free asset that varies between 100% long and 100% short in the risky asset. The first two strategies are derived by two rolling two-stage mean-risk stochastic optimization models, respectively. These two models both decide on a long or short position to hold in the risky asset for the period at the beginning of each period and optimize the trade-off between maximizing the expected return and minimizing the investment risk at the end of each period. The only difference between these two models is the risk measure used to quantify the investment risk. One uses variance, which characterizes investment risk by both upside gain and downside loss. The other one employs conditional value-at-risk (CVaR), a downside risk measure concentrating on the potential loss. Analogous to these two models, the third trading strategy, time series momentum strategy controlling downside

risk (TSMDR), is proposed. TSMDR is a novel hybrid model that follows the time series momentum strategy but controls the CVaR of return instead of variance while setting the risky asset position. When evaluating the results of implementing the rolling two-stage mean-risk stochastic optimization models as well as TSMDR in a historical backtest, we find that TSMDR frequently dominates the traditional TS strategy and the mean-risk optimization strategies in terms of both Sharpe ratio and Sortino ratio.

The goal of Chapter 4 is to improve the performance of another momentum trading strategy, the CS strategy, by avoiding the momentum crashes that have been observed to occur in historical backtests of this strategy. Momentum crashes are due to the uncertainty of the stock market, especially when the market is volatile and recovered from the bear market. In Chapter 4, we develop a stochastic mean-risk optimization model with time-varying risk aversion to avoid momentum crashes from three perspectives: 1) Rather than the widely-used variance (volatility) to quantify the risk of momentum trading, a downside risk measure, CVaR, is exploited to control the expected investment loss under the extreme case, including during the momentum crash period, 2) We extend the traditional mean-CVaR model by allowing the change in the tail probability in CVaR and the risk-aversion coefficient to flexibly adjust the investment strategy according to the constantly-changing market condition, and 3) A combination of heuristic and optimization matching algorithms is applied to generate scenarios that capture the inherent structure, including the first four moments and correlation, of uncertain momentum portfolio returns, whose quality is evaluated using the mass transportation distance rank histogram and verified by a Cramér–von Mises hypothesis test. A backtest using total returns of NYSE, NASDAQ, and AMEX equities from January 1926, to December 2020 is conducted, from which we find that our model successfully avoids momentum crashes by greatly reducing the investment loss and maintaining remarkable performance in almost all market conditions.

Finally, Chapter 5 concludes the dissertation with a summary of research findings and contributions.

1.4 References

Asness, C. S., Frazzini, A., and Pedersen, L. H. (2012). Leverage aversion and risk parity. *Financial Analysts Journal*, 68(1):47–59.

Asness, C. S., Liew, J. M., and Stevens, R. L. (1997). Parallels between the cross-sectional predictability of stock and country returns. *Journal of Portfolio Management*, 23(3):79.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, 49(3):283–306.

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

Gülten, S. and Ruszczyński, A. (2015). Two-stage portfolio optimization with higher-order conditional measures of risk. *Annals of Operations Research*, 229(1):409–427.

Guo, X. and Ryan, S. M. (2020). Reliability assessment of scenarios generated for stock index returns incorporating momentum. *International Journal of Financial & Economics*, 26(3):4013–4031.

Israel, R. and Moskowitz, T. J. (2013). The role of shorting, firm size, and time on market anomalies. *Journal of Financial Economics*, 108(2):275–301.

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

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

Kaut, M. and Wallace, S. W. (2007). Evaluation of scenario-generation methods for stochastic programming. *Pacific Journal of Optimization*, 3(2):257–271.

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

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

Okunev, J. and White, D. (2003). Do momentum-based strategies still work in foreign currency markets? *Journal of Financial and Quantitative Analysis*, 38(2):425–447.

Roman, D., Mitra, G., and Spagnolo, N. (2010). Hidden Markov models for financial optimization problems. *IMA Journal of Management Mathematics*, 21(2):111–129.

Rouwenhorst, K. G. (1998). International momentum strategies. *The Journal of Finance*, 53(1):267–284.

Sharpe, W. F. (1994). The Sharpe ratio. *Journal of Portfolio Management*, 21(1):49–58.

Topaloglou, N., Vladimirou, H., and Zenios, S. A. (2008). Pricing options on scenario trees. *Journal of Banking & Finance*, 32(2):283–298.

Yu, L.-Y., Ji, X.-D., and Wang, S.-Y. (2003). Stochastic programming models in financial optimization: A survey. *AMO - Advanced Modeling and Optimization*, 5(1):1–26.

CHAPTER 2. RELIABILITY ASSESSMENT OF SCENARIOS GENERATED FOR STOCK INDEX RETURNS INCORPORATING MOMENTUM

Modified from a manuscript published in *International Journal of Finance and Economics*

Abstract

Stochastic programming models for portfolio optimization rely on scenario paths for returns derived from stochastic process models. This paper investigates a variant of the geometric Brownian motion process for stock index returns that incorporates index momentum. Based on this model, three different processes for generating scenarios on a rolling basis are devised, that differ according to how frequently the momentum parameter is updated and whether it is estimated according to a simple moving average or an exponentially weighted moving average of returns. The reliability of scenario sets generated for multiple instances are assessed by applying a recently developed statistical tool. Backtesting is conducted in case studies of the Standard & Poor's 500 Index and the Financial Times Stock Exchange 100 Index for two different historical periods. The numerical results show that the frequency with which the expected return is updated does not significantly influence the performance of the scenario generation procedure, whereas how the expected return is calculated affects the autocorrelation and dispersion of generated scenarios drastically. All scenario generation schemes are highly sensitive to the estimated volatility of the index returns. Among the three processes tested, the algorithms that incorporate momentum estimates according to an exponentially weighted moving average can generate reliable scenarios when the volatility estimation error is small.

2.1 Introduction

Portfolio optimization is the problem of dynamically allocating investment funds among a set of assets or asset classes. This problem can be formulated as a mathematical program over a given time horizon by defining three major components of the model. First, decision variables represent the investment choices. Second, an objective function represents a goal such as maximizing future

wealth, the returns generated by the portfolio in each period, and/or some measure of investment risk. Finally, constraints express limits on acceptable investment behavior and tolerable levels of risk while describing the links between decisions in one period and available funds to invest in the next. The uncertainty of future investment performance poses daunting challenges for modeling and computation. In some situations it is possible to formulate and solve a dynamic optimization problem where both time and outcomes are continuous (Koijen et al., 2009). However, it is more common to discretize the time horizon into periods and generate discrete probabilistic scenarios for asset performance (Fulton and Bastian, 2018). In practice, the resulting stochastic programming formulation may be updated and solved repeatedly to rebalance a portfolio in response to changing financial conditions.

To generate scenarios of investment performance over multiple future periods, a common approach is to identify and estimate parameters for a stochastic process model. Such a model may be derived by fitting to historical data (Høyland and Wallace, 2001; Fulton and Bastian, 2018) or be based on assumptions about markets and investor behavior (**?**). Based on the model specified, sample paths for future realizations can be generated by Monte Carlo methods. For stock prices, the classical lognormal or geometric Brownian motion model is appealing due to its simplicity and widespread use. It rests on simple assumptions on investor rationality in efficient markets and requires only two parameters, representing drift and volatility, to be estimated. However, a considerable amount of empirical research on stock market momentum and strategies to exploit it has raised questions concerning these fundamental assumptions (Jegadeesh and Titman, 1993; Conrad and Kaul, 1998; Lewellen, 2002). The various definitions of, and explanations for, momentum unite around the idea that trends in market performance persist over periods of up to a few months. Mathematically, the existence of momentum undermines the plausibility of a constant drift parameter that could be estimated once and used to generate sample paths extending far into the future. It also raises questions about how to estimate volatility in a nonstationary regime.

For an investment manager to place confidence in a stochastic programming model for portfolio optimization, convincing evidence must be provided that the probabilistic scenarios accurately reflect the range of possible financial outcomes over the investment horizon. Frequently, computational resources limit the number of scenarios that solvers can accommodate in reasonable solution times. The stochastic programming literature on scenario generation and reduction is rich in methods

for discretizing continuous distributions and reducing large sets of generated scenarios to smaller representative sets for use in the optimization model. However, unlike in studies of investment strategies where backtesting over historical time periods is routine, it is rare for observational data on actual realizations to be used in testing the performance of scenario generation methods for use in stochastic programming models.

The goals of this paper are to investigate a variant of the geometric Brownian motion model that replaces the constant drift parameter with a time-dependent momentum parameter, and to explore different schemes for updating momentum estimates while generating sample paths. We use a recently developed tool for assessing the reliability of sets of probabilistic scenarios against observational data. Originally motivated by work on verification of ensemble forecasts in meteorology, and successfully applied in the context of energy planning, a graphical and quantitative approach is applied to loosely test a hypothesis that the observed realization over a given period is indistinguishable from the collection of scenarios generated to represent possible realizations over that period. We find that the procedures that update the momentum estimate while emphasizing more recent observations outperform the alternative, as they generate reliable scenarios as long as the error in the estimate of volatility is small.

The remainder of this paper is organized as follows: Section 2.2 briefly reviews the relevant literature. Section 2.3 presents our methodology for generating scenarios of index returns by (1) defining the model of stock index returns and index levels, (2) detailing three alternative processes for scenario generation on a rolling basis, and (3) conducting the scenario assessment. In Section 2.4, the scenario generation method is applied to the Standard and Poor's 500 Index and the Financial Times Stock Exchange 100 Index in two separate time periods, and the results are analyzed. Section 2.5 concludes the paper with a summary of findings and directions for future research.

2.2 Literature Review

When determining how much of an investment fund should be allocated to equities, the market return is one of the primary concerns. Here we focus on the return of a market index, computed as the change in the index level over a period as a proportion of the level at the beginning of the period. Due to the uncertainty of the market, stochastic process models are employed to describe

the index levels and returns. Typically, the form of the stochastic process is fit to historical data and its parameters are calibrated using some prior information (Hochreiter and Pflug, 2007).

Recent empirical research suggests that investors can improve the timing of their transactions by following momentum signals. Momentum can be defined as the tendency of stock prices to keep moving in the same direction for a period after an initial impulse. There are many forms of momentum, such as price momentum, return momentum, and earnings momentum, which are distinguished by the initial impulse used. Price momentum, where the initial impulse is simply a change in the price itself, is the most widely used, and it is stronger than the return momentum (Hong and Satchell, 2012). Index momentum is a type of price momentum, because the stock market index is a weighted average of the prices of several representative stocks or other investment vehicles within a sector of the stock market. It shares all the properties of price momentum. Even though the price momentum seems basic, there is little consensus in the literature about what drives it. Generally speaking, price momentum is attributed to behavioral explanations, which fall into two main categories. The first category addresses a process of gradual adjustment to news. Stock prices initially underreact to news, then adjust as time goes by so that the long-term response becomes appropriate and rational. The second category stresses the overreaction by irrational investors to stories of dubious relevance. When overreaction develops gradually, the stock prices may display momentum for a period of time but will eventually reverse and return to their fundamental values. Attempting to exploit momentum, researchers have developed the relative strength strategy, which consists of buying past winners and selling past losers. Jegadeesh and Titman (1993) show that the past winners have continued to outperform past losers over horizons of 3-12 months, indicating that the return continuation lasts for at most a year. There are three sources for the expected profit of this relative strength strategy. This paper is concerned with only one source of momentum profits: stock returns might be positively autocorrelated (Moskowitz and Grinblatt, 1999), as a firm with a high return today is expected to have high returns in the future.

Admittedly, momentum can play an important role in estimating the future stock returns. Nevertheless, as momentum has various forms and controversial causes, models of returns and stock prices that incorporate momentum are few. Carhart (1997) extends Fama and French's (1993) three-factor model (Fama and French, 1993) by incorporating a one-year momentum factor in the regression function of stock returns. Hong and Satchell (2012) define momentum as the ratio of

today's price to its value n periods ago, and describe the asset price process as a log Ornstein-Uhlenbeck process so that momentum can manifest itself as either trend or autocorrelation. They validate this form of momentum for single asset prices; however, they do not discuss how to apply this formulation of momentum to simulate future asset prices. Koijen et al. (2009) propose a model of returns that accommodates both momentum and mean reversion with two state variables: the weighted average of past returns and the dividend yield.

A stochastic process model for stock index returns can be most accurately formulated with continuous marginal distributions. Nonetheless, it is difficult to solve a stochastic programming problem with continuous distributions for the uncertain parameters. Thus, the continuous distributions at future time points are often discretized in terms of a finite number of scenarios, the process of which is called scenario generation. Various scenario generation methods have been proposed and applied to portfolio optimization problems. Sampling from the continuous distributions is a straightforward approach that is widely used to generate scenarios. In previous studies (Bertsimas and Pachamanova, 2008; Dantzig and Infanger, 1993), the return has been assumed to follow a normal or lognormal distribution. Yu et al. (2003) illustrate three methods for generating asset return scenarios: 1) bootstrapping historical data, 2) statistical modeling with the value at risk approach, and 3) modeling economic factors and asset returns with vector autoregressive models. Bootstrapping is sampling with replacement. By bootstrapping, each scenario is obtained by sampling returns that were observed in the past. The second method is to use the historical data to estimate the volatility and covariance matrix of a multivariate normal distribution and then use Monte Carlo simulation to generate scenarios of returns of different asset classes. The third one is to apply a vector autoregressive model to accommodate the progress of time series while considering the correlation of asset returns. Şakar and Köksalan (2013) estimate the return through a regression equation for the single index model and generate scenarios of index returns from a random walk model. Gulten and Ruszczynski (2015) apply a multivariate generalized orthogonal GARCH model (Gülten and Ruszczyński, 2015; Van der Weide, 2002), which involves the covariance of uncorrelated return ensembles based on both squared values and the covariance matrix of past uncorrelated return ensembles. Chen and Yang (2017) apply heuristic moment matching to approximate the joint distribution of risky asset returns for future periods. Fulton and Bastian (2018) employ Monte Carlo simulation to generate scenarios based on the estimate of sample means and covariance matrices

from a multivariate normal distribution, remove the outlier data based on percentiles, and resample the remaining data to obtain three types of scenarios: positive, negative, or neutral outlook.

Although numerous scenario generation methods have been proposed, no universally superior scenario generation method exists. Therefore, we need some techniques to evaluate the quality of generated scenarios. The verification rank histogram is one such recently-developed technique. The advantage of rank histograms over other scenario quality assessment methods is their ability to provide insights into how well the scenarios compare to the realizations. The verification rank histogram is constructed from the ranks of the observation within each group, consisting of a constant number of scenarios and the corresponding observation. Scalar rank histograms (Hamill and Colucci, 1997) originally were developed to assess univariate ensemble weather forecasts. The minimum spanning tree (MST) rank histogram, which tabulates the frequencies of the rank of the MST length for each scenario, within the group of such lengths that is obtained by replacing an observation for each of its scenario member in turn, is a multivariate extension of the scalar rank histogram (Wilks, 2004). One limitation of the MST histogram is that it is applicable only to equally likely scenarios. While scenarios generated by sampling are equally likely, the application of scenario reduction techniques (Hochreiter and Pflug, 2007)) results in unequally likely scenarios. Sarı et al. (2016) thus develop a new mass transportation distance (MTD) rank histogram, which can be applied to both equally and unequally likely scenarios. It substitutes the MTD (or Wasserstein distance) between the scenario members and the observation for the MST length as a pre-rank function. The principle behind the rank histogram is that the set of ideal scenario values at a given point and its associated observation can be regarded as random samples from the same distribution. Accordingly, the rank histogram of ideal scenarios is supposed to be flat (reflecting a discrete uniform distribution for the ranks), which would indicate that the generation method produces scenarios that well represent the underlying uncertainty. In contrast, overdispersion will present as an upward sloping rank histogram, and underdispersion in the scenarios will show up as a downward sloping histogram (Hamill, 2001). Sarı and Ryan (2018) demonstrate a connection between scenario reliability as assessed by the MTD rank histogram and quality of solutions to the resulting stochastic program instances in the context of short-term power system scheduling.

In summary, our paper focuses on identification, estimation and validation of the scenarios derived from a stochastic process model of stock index returns. In particular, we modify a model

proposed by Koijen et al. (2009) to generate scenarios of stock market returns by Monte Carlo simulation. This model captures momentum as an exponentially weighted moving average of recent past index returns to predict future returns. On the basis of this model, three generation procedures are proposed, and the MTD rank histograms developed by Sarı et al. (2016) are plotted to assess the quality of the generated scenarios. Goodness of fit tests are applied to test a hypothesis of uniformity, which could indicate scenario reliability.

2.3 Methodology

2.3.1 Momentum

Momentum refers to the tendency of stock prices to continue moving in the same direction for several months after an initial impulse. Technical analysts interpret momentum in terms of moving averages. An upward (bullish) trend appears when the current return rises above its moving average, while a downward (bearish) trend emerges when the current return falls below its moving average. Since the return is the price change relative to its current value, momentum can be used as an indicator for predicting future returns.

2.3.2 Model of Stock Index Returns and Levels

In this section, we introduce a model using momentum as the expected index return rate for a nonstationary variant of the geometric Brownian motion process of index levels. This model is built based on the work of Koijen et al. (2009). Their model accommodates both return continuation (momentum) over short horizons and return reversals (mean reversion) over longer horizons. Because our model is intended for use to generate scenarios over short time horizons with frequent parameter updates, the effect of return reversals is incorporated in the periodically renewed observations. For this reason, by keeping the short-run return continuation and neglecting the long-run return reversals, we can simplify the stock return model as

$$r_t = \frac{\Delta S_t}{S_t} = M_t \Delta t + \sigma \Delta Z_t \tag{2.1}$$

where r_t is the return of the index at time t, S_t is the index level at time t, M_t is the expected return over a short time interval, σ is the volatility of the index level, and Z_t is standard Brownian motion with mean 0 and variance t. We use Δ to denote a change over a time increment Δt

(i.e. $\Delta S_t \equiv S_{t+\Delta t} - S_t$). According to the independent increments property of Brownian motion (Hull, 2009), ΔZ_t is normally distributed with mean 0 and standard deviation $\sqrt{\Delta t}$. By linear transformation of the normal distribution, the marginal distribution of stock market return at time t is normal with mean $M_t \Delta t$ and standard deviation $\sigma\sqrt{\Delta t}$:

$$r_t \sim N(M_t \Delta t, \sigma\sqrt{\Delta t}) \tag{2.2}$$

To specify this model for returns, the process of the expected return, M_t, and the volatility, σ, must be estimated.

Similar to Koijen et al. (2009), we approximate the expected return by a moving average model of momentum over the past T periods:

$$M_t = \sum_{i=1}^{T} \frac{e-1}{e^i} \frac{S_{t-i+1} - S_{t-i}}{S_{t-i}} \tag{2.3}$$

This model, reflecting the recent performance of the index levels, can also be viewed as an exponentially weighted moving average. It modifies the so-called performance variable in Koijen et al. (2009) so that the weights of returns used in the estimate sum to one and M_t is an unbiased estimate of the return. The exponentially diminishing weights place emphasis on the most recent return values, while the function of summation is to smooth the effect of random index level fluctuations.

Next, the volatility of the index level, σ, is a statistical measure of investors' uncertainty about market returns. A sector of the stock market with high volatility is riskier, but this risk cuts both ways: Investing in a volatile market will increase the chance of success as much as the risk of failure. Hull (2009) points out that, for the geometric Brownian motion model with constant drift, $\log S_t$ follows a Brownian motion process with a standard deviation of $\sigma\sqrt{t}$. Equivalently, $\log \frac{S_t}{S_{t-1}}$ follows a normal distribution with a standard deviation of σ, which suggests that the volatility can be estimated based on the standard deviation of the log ratios of successive index levels. Let $u_t \equiv \log \frac{S_t}{S_{t-1}}$ be this log ratio at time t, N be the number of historical periods used to calculate σ, and $\bar{u}$ be the mean of the sequence $\{u_1, ..., u_N\}$. Then the past volatility can be computed as

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{k=1}^{N} (u_k - \bar{u})^2}. \tag{2.4}$$

2.3.3 Scenario Generation of Stock Index Returns

To generate scenarios of the stock index returns according to the model in Section 2.3.2, we apply Monte Carlo simulation, described by Hull (2009), by sampling randomness for the process of the stock returns. To update momentum estimates and recalibrate the model with recent data, we repeatedly generate scenario paths over short time horizons on a rolling basis.

Let F be length of the forecast horizon. We use the observations up to time t to generate scenarios of returns for the following F periods. In each period t, J scenarios are generated. Returns over a single period are generated by sampling from the normal distribution with mean equal to the current momentum estimate and standard deviation equal to the estimated volatility. The momentum variable is an exponentially weighted sum of the returns over the past T periods, and the volatility, which is calculated as the standard deviation of the sequence of log ratios of index levels in the historical period, is a constant throughout the whole simulation horizon. A critical issue is that if this forecast horizon is greater than a year, the effect of momentum will gradually diminish while the effect of mean reversion will increase, which can cause a poor performance of the generated scenarios. Therefore, the forecast horizon should be set to less than one year. This leads us to apply the idea of rolling horizon: cut the horizon short, and consider the dynamic problem on a rolling basis. Specifically, at period t, scenarios of returns representing periods $t+1$ to $t+F$ are generated according to the realized index levels from period $t-T+1$ to t. These scenarios along with the associated observation from $t+1$ to $t+F$ compose one instance. To obtain non-overlapping instances, we proceed to period $t+F$ in the next step. The index level in period $t+F$ is observed, and scenarios over periods $t+F+1$ to $t+2F$ are generated based on the historical data from period $t+F-T+1$ to $t+F$. By repeating this procedure, we not only maintain an effective forecast horizon but also generate relatively reliable scenarios with such frequent updates of observational data.

In this study, three different scenario generation processes are proposed: a process with daily momentum updated daily, a process with daily momentum updated monthly, and a process employing expected daily return updated monthly based on a simple moving average. The notation used to present the algorithms is defined as follows.

J	Number of scenarios

T	Number of days to calculate each estimate of momentum
S_d	Observed index level on day d
H	Total number of trading days in the simulation horizon
H'	Total number of trading months in the simulation horizon
F	Forecast horizon (days)
F'	Forecast horizon in months
σ	Volatility computed on a daily basis
t_d	Day d in month t
D	Number of trading days in each month
L	Number of months of data used to estimate the mean values of returns
M_d	Daily momentum on day d
$M^*_{d,j}$	Updated daily momentum on day d in scenario j
μ_d	Expected daily return computed by simple moving average on day d
$r^*_{d,j}$	Generated return on day d in scenario j
$S^*_{d,j}$	Generated index level on day d in scenario j
N	Number of non-overlapping instances

2.3.3.1 Process with Daily Momentum Updated Daily

In this subsection, daily momentum is updated each day upon generating the new index level. For example, on day d, the realized index level, S_d, can be observed. We use it, along with the past T periods' index levels, to compute the daily momentum, M_d. Based on the volatility and this momentum, the return on this day, $r^*_{d,j}$, can be derived for each scenario j. Then, we can forecast the index level for the next day in scenario j, $S^*_{d+1,j}$, according to the projected return, $r^*_{d,j}$. By taking account of this newly generated index level, the momentum in scenario j is updated

to be $M^*_{d+1,j}$. We use this updated momentum to generate the return for the next day, $r^*_{d+1,j}$. By repeating this procedure, each scenario will have its own paths of daily momentum. To avoid overlapping generation, we obtain the observation data only once every F days. This procedure is detailed in Algorithm 1.

Algorithm 1 Process with daily momentum updated daily

Input: J, T, H, F, σ, S_d for $d = 1, \cdots, H$

Output: $r^*_{d,j}$ for $d = T+1, \cdots, T+NF$ and $j = 1, \cdots, J$

begin:

$N \leftarrow \lfloor \frac{H-T}{F} \rfloor$ ▷ compute the number of non-overlapping instances

for $d \leftarrow 1$ to T **do**

for $j \leftarrow 1$ **to** J **do** $S^*_{d,j} \leftarrow S_d$ **end for** ▷ initialize the scenarios of daily index levels up to day T

end for

for $d \leftarrow T+1, T+F+1, T+2F+1, \cdots, T+(N-1)F+1$ **do**

for $j \leftarrow 1$ to J **do**

$S^*_{d,j} \leftarrow S_d$ ▷ initialize all scenarios to the observed value on day d

$M_d \leftarrow \frac{e-1}{e}\frac{S_d - S_{d-1}}{S_{d-1}} + ... + \frac{e-1}{e^T}\frac{S_{d-T+1} - S_{d-T}}{S_{d-T}}$ ▷ estimate momentum on day d

$r^*_{d,j} \leftarrow rnorm(M_d, \sigma)$ ▷ generate scenarios for day d

for $l \leftarrow d$ to $d+F-2$ **do** ▷ generate scenarios for day $d+1$ to $d+F-1$

$S^*_{l+1,j} \leftarrow (1 + r^*_{l,j}) S^*_{l,j}$

$M^*_{l+1,j} \leftarrow \frac{e-1}{e}\frac{S^*_{l+1,j} - S^*_{l,j}}{S^*_{l,j}} + ... + \frac{e-1}{e^T}\frac{S^*_{l-T+2,j} - S^*_{l-T+1,j}}{S^*_{l-T+1,j}}$

$r^*_{l+1,j} \leftarrow rnorm(M^*_{l+1,j}, \sigma)$

end for

end for

return $r^*_{l,j}$, where $l = d$ to $l = d+F-1$ and $j = 1$ to J.

end for

2.3.3.2 Process with Daily Momentum Updated Monthly

In this subsection, the daily momentum within the same month is held to a constant value and updated monthly. To illustrate, we denote the number of trading days in every month as D. The daily momentum in month t, M_{t_d}, is calculated according to the index level of the first trading day in that month, S_{t_1}, to the last trading day in that month, S_{t_D}. Daily returns in month $t+1$ for each scenario j, $r^*_{(t+1)_d,j}$, can be generated based on this momentum value and volatility. With the generated returns, we can forecast the index level for each day in the next month, $S^*_{(t+1)_d,j}$, and derive a new daily momentum in the next month for each scenario, $M^*_{(t+1)_d,j}$. Returns for each day in the next month, $r^*_{(t+2)_d,j}$, can be generated by this newly updated momentum. To acquire non-overlapping instances, we observe the realized data once every F' months. The specific procedure is demonstrated in Algorithm 2.

Algorithm 2 Process with daily momentum updated monthly

Input: J, H, F', D, σ, S_{t_d} for $t = 1, \cdots, H'$ and $d = 1, \cdots, D$

Output: $r^*_{t_d,j}$ for $t = 2, \cdots, 1 + NF', d = 1, \cdots, D$ and $j = 1, \cdots, J$

begin:

$H' \leftarrow \lfloor \frac{H}{D} \rfloor$, $N \leftarrow \lfloor \frac{H'-1}{F'} \rfloor$ ▷ compute the number of non-overlapping instances

for $t \leftarrow 1, 1 + F', 1 + 2F', \cdots, 1 + (N-1)F'$ **do**

for $j \leftarrow 1$ to J **do**

$S^*_{(t+1)_1,j} \leftarrow S_{(t+1)_1}$ ▷ initialize all scenarios to the observed value

$r^*_{t_D,j} \leftarrow \frac{S_{(t+1)_1} - S_{t_D}}{S_{t_D}}$

for $d \leftarrow 1$ to D **do**

$M_{t_d} \leftarrow \frac{e-1}{e}\frac{S_{t_D} - S_{t_{D-1}}}{S_{t_{D-1}}} + \frac{e-1}{e^2}\frac{S_{t_{D-1}} - S_{t_{D-2}}}{S_{t_{D-2}}} + ... + \frac{e-1}{e^{D-1}}\frac{S_{t_2} - S_{t_1}}{S_{t_1}}$ ▷ get momentum in month t

$r^*_{(t+1)_d,j} \leftarrow rnorm(M_{t_d}, \sigma)$ ▷ generate scenarios for month $t+1$

end for

for $f \leftarrow t$ to $t + F' - 1$ **do** ▷ generate scenarios for month $t+2$ to $t+F'$

for $d \leftarrow 1$ **to** $D - 1$ **do** $S^*_{(f+1)_{(d+1)},j} \leftarrow (1 + r^*_{(f+1)_d,j})S^*_{(f+1)_d,j}$ **end for**

for $d \leftarrow 1$ to D **do**

$M^*_{(f+1)_d,j} \leftarrow \frac{e-1}{e}\frac{S^*_{(f+1)_D,j} - S^*_{(f+1)_{D-1},j}}{S^*_{(f+1)_{D-1},j}} + ... + \frac{e-1}{e^{D-1}}\frac{S^*_{(f+1)_2,j} - S^*_{(f+1)_1,j}}{S^*_{(f+1)_1,j}}$

$r^*_{(f+2)_d,j} \leftarrow rnorm(M^*_{(f+1)_d,j}, \sigma)$

end for

end for

end for

return $r^*_{f_d,j}$, where f_d is from $(t+1)_1$ to $(t+F')_D$ and j is from 1 to J

end for

2.3.3.3 Process with Expected Daily Return Updated Monthly

To understand whether momentum improves the accuracy of forecast returns, we consider a third generation process using a simple, rather than exponentially weighted, average of recent returns. We replace the momentum estimates used in Subsection 2.3.3.2 with the equally weighted average

of daily returns over the most recent L months to estimate the expected returns. To be more specific, we take month t as an illustration. On each day in month t, the expected return, μ_{t_d}, is computed as the mean of daily returns in the past L months. Then, the return of day d in month $t+1$ for scenario j, $r^*_{(t+1)_d,j}$, can be simulated by sampling from the normal distribution with mean equal to this expected return, μ_{t_d}, and standard deviation equal to the volatility, σ. Note that, besides the approach to calculate expected return, there is another difference between Algorithm 2 and Algorithm 3. In Algorithm 2, the expected return is updated when new stock index levels are generated; while for Algorithm 3, the expected return is always based on the observational data and remains the same throughout the instance horizon.

Algorithm 3 Process with expected daily return updated monthly

Input: J, L, H, F', D, σ, S_{t_d} for $t = 2, \cdots, L+H'$ and $d = 1, \cdots, D$

Output: $r^*_{t_d,j}$ for $t = L+2, \cdots, L+1+NF'$, $d = 1, \cdots, D$ and $j = 1, \cdots, J$

begin:

$H' \leftarrow \lfloor \frac{H}{D} \rfloor$, $N \leftarrow \lfloor \frac{H'-1}{F'} \rfloor$ ▷ compute the number of non-overlapping instances

for $t \leftarrow L+1, L+1+F', L+1+2F', \cdots, L+1+(N-1)F'$ **do**

  for $j \leftarrow 1$ to J **do**

    for $d \leftarrow 1$ to D **do**

      $\mu_{t_d} \leftarrow \frac{1}{DL}\left(\frac{S_{(t+1)_1}-S_{t_D}}{S_{t_D}} + \cdots\cdots + \frac{S_{(t-L+1)_2}-S_{(t-L+1)_1}}{S_{(t-L+1)_1}}\right)$ ▷ get the expected daily return using the past L-month index levels

    end for

    for $f \leftarrow t$ to $t+F'-1$ **do** ▷ generate scenarios of month $t+1$ to $t+F'$

      for $d \leftarrow 1$ **to** D **do** $r^*_{(f+1)_d,j} \leftarrow rnorm(\mu_{t_d}, \sigma)$ **end for**

    end for

  end for

  return $r^*_{f_d,j}$, where f_d is from $(t+1)_1$ to $(t+F')_D$ and j is from 1 to J

end for

2.3.4 Assessment of Scenarios

To evaluate the reliability of a scenario generation method, we seek to quantify the extent to which scenarios and observations are interchangeable. A scenario set is called reliable or calibrated if the relative frequency of occurrence of scenario assigned a probability p tends to be close to p (Hsu and Murphy, 1986). In other words, the more similar the distribution of generated scenarios and the distribution of real observations are, the more reliable the generation method is judged to be. Sarı et al. (2016) develop and test a novel mass transportation distance (MTD) rank histogram to assess whether the scenarios have similar patterns as the corresponding observations. Based on this work, Sarı and Ryan (2016) create a MTDrh package in R.

The mass transportation distance is the minimum total cost of transporting the probability of one distribution to the other. For two discrete distributions $P = \sum_i p_i \delta_{x_i}$ and $Q = \sum_j q_j \delta_{y_j}$, the mass transportation distance can be computed by solving an optimization problem

$$MTD = \min \left\{ \sum_{i,j} c(x_i, y_j)\eta_{i,j} : \eta_{i,j} \geq 0, \sum_i \eta_{i,j} = q_j, \sum_j \eta_{i,j} = p_i \right\}, \tag{2.5}$$

where $c(x_i, y_j)$ is the distance between mass points x_i and y_j. For example, $c(x_i, y_j)$ could be the Euclidean distance. In our case, suppose the number of instances is N, $\boldsymbol{r_d}$ is a vector of the observed (realized) values from day $d+1$ to day $d+F$ (i.e. $\boldsymbol{r_d} = (r_{d+1}, \cdots, r_{d+F})$), and $\boldsymbol{r^*_{d,j}}$ is the generated scenario j for day $d+1$ to day $d+F$ (i.e. $\boldsymbol{r^*_{d,j}} = (r^*_{d+1,j}, \cdots, r^*_{d+F,j})$). Then when calculating the mass transportation distance between the distribution of $\boldsymbol{r_d}$, denoted as P, and the distribution of $\boldsymbol{r^*_{d,j}}$, denoted as Q, the mass transportation minimization problem has a simple solution:

$$MTD_0^d = \sum_{j=1}^{J} c(\boldsymbol{r_d}, \boldsymbol{r^*_{d,j}}) p_j^d, \quad d = 1, \cdots, N, \tag{2.6}$$

where $P = \delta_{\boldsymbol{r_d}}$, $Q = \sum_j p_j^d \delta_{\boldsymbol{r^*_{d,j}}}$, and $c(\boldsymbol{r_d}, \boldsymbol{r^*_{d,j}}) = \sqrt{\sum_{k=d+1}^{d+F} (r_k - r^*_{k,j})^2}$.

Next, we compute MTD_j^d by replacing the scenario j in instance d with the observation during the same time period, and order these MTD values from the largest to the smallest to determine the rank of MTD_0^d in each instance. Finally, the rank histogram plots the frequency with which this rank attains each of its possible values, $1, \ldots, J+1$, over the N instances.

If the scenario sets are reliable, then their distributions should be similar to the distribution of the observation, which means the mass transportation distance from the scenarios to the observation

should not be distinguishable from the distances to each scenario. Moreover, by examining the relationship between the shape of the MTD histogram and the difference between the autocorrelation of scenarios and observations, Sarı et al. (2016) point out that if the MTD rank histogram exhibits uniformity, it indicates that the autocorrelation of scenarios and observations are similar and the observation is indistinguishable from the scenarios. If the histogram displays a U-shape or a downward trend, it means the autocorrelations of the scenarios are less than that of the observation and the scenarios are under-dispersed; if it shows an upward trend, it suggests that the autocorrelation of the scenario is greater than that of the observation and the scenarios are over-dispersed.

Note that even if the histogram exhibits uniformity, it cannot be guaranteed that the ranks indeed follow a uniform distribution because the effect of bias can offset the effect of overdispersion when being displayed in the MTD histogram. In order to prevent misdiagnosis, a bias-correction technique should be utilized. The scenario data are de-biased according to the following formula (Sarı et al., 2016):

$$\tilde{\boldsymbol{r}}^*_{\boldsymbol{d,j}} = \boldsymbol{r}^*_{\boldsymbol{d,j}} - \frac{1}{N}\sum_{d=1}^{N}\left(\frac{1}{J}\sum_{j=1}^{J}\boldsymbol{r}^*_{\boldsymbol{d,j}} - \boldsymbol{r_d}\right), \quad d = 1, \cdots, N, \quad j = 1, \cdots, J. \tag{2.7}$$

Additionally, we are interested in the effects of correlations within scenarios on effective distances between scenarios. Therefore, the variances of the marginal distributions are equalized by applying the Mahalanobis transformation (Wilks, 2004) for both the observation, $\boldsymbol{r_d}$, and the de-biased scenarios, $\tilde{\boldsymbol{r}}^*_{\boldsymbol{d,j}}$, to derive the transformed observation, $\boldsymbol{z_{d,o}}$, and the transformed de-biased scenario, $\boldsymbol{z_{d,j}}$:

$$\boldsymbol{z_{d,o}} = \boldsymbol{S_d^{-1/2}}(\boldsymbol{r_d} - \bar{\tilde{\boldsymbol{r}}}^*_{\boldsymbol{d}}), \quad d = 1, \cdots, N \tag{2.8}$$

$$\boldsymbol{z_{d,j}} = \boldsymbol{S_d^{-1/2}}(\tilde{\boldsymbol{r}}^*_{\boldsymbol{d,j}} - \bar{\tilde{\boldsymbol{r}}}^*_{\boldsymbol{d}}), \quad d = 1, \cdots, N, \quad j = 1, \cdots, J, \tag{2.9}$$

where

$$\bar{\tilde{\boldsymbol{r}}}^*_{\boldsymbol{d}} = \frac{1}{J+1}\left(\boldsymbol{r_d} + \sum_{j=1}^{J}\tilde{\boldsymbol{r}}^*_{\boldsymbol{d,j}}\right), \quad d = 1, \cdots, N \tag{2.10}$$

and

$$\boldsymbol{S_d} = \frac{1}{J}\left[(\boldsymbol{r_d} - \bar{\tilde{\boldsymbol{r}}}^*_{\boldsymbol{d}})(\boldsymbol{r_d} - \bar{\tilde{\boldsymbol{r}}}^*_{\boldsymbol{d}})^T + \sum_{j=1}^{J}(\tilde{\boldsymbol{r}}^*_{\boldsymbol{d,j}} - \bar{\tilde{\boldsymbol{r}}}^*_{\boldsymbol{d}})(\tilde{\boldsymbol{r}}^*_{\boldsymbol{d,j}} - \bar{\tilde{\boldsymbol{r}}}^*_{\boldsymbol{d}})^T\right], \quad d = 1, \cdots, N \tag{2.11}$$

After de-biasing and transforming, we use the Cramér-von Mises goodness-of-fit test to assess the nearness to uniformity. The Cramér–von Mises statistic, W^2, can be computed to measure the distance between the uniform distribution and the distribution of the rank of MTD_0. Let F^* be the uniform distribution with possible outcomes of $\{1, 2, \ldots, J+1\}$ that are equally likely to occur, and F be the distribution of the rank of MTD_0. Then W^2 is computed as (Choulakian et al., 1994)

$$W^2 = \frac{N}{J+1} \sum_{x=1}^{J+1} (F(x) - F^*(x))^2 \tag{2.12}$$

2.4 Computational Results

The index levels for the S&P 500 index and the FTSE 100 index are obtained from the CRSP (www.crsp.com) and Yahoo Finance (https://finance.yahoo.com) databases, respectively. To test the model, we consider two time intervals for simulation. The first runs from July 1, 1999, to December 27, 2006, which follows the trend of up-down, and the second goes from January 3, 2011, to June 29, 2018, which has an upward trend (Figure 2.1). These two data sets each contains a total of 1886 trading days. Therefore, the value of H in Algorithm 1 is 1886, while $H' = 94$ in Algorithms 2 and 3. All the following statistical simulations and analyses are conducted in R.

Figure 2.1: Index levels of S&P 500 and the FTSE 100 from July 1, 1999, to June 29, 2018

In the experiments, $J = 25$ scenarios are generated for each instance. By calculation, the number of trading days in each month, D, is around 20. For simplicity, we assume that $D = 20$ for all months. The forecast horizon, F, in Algorithm 1 is set to be 20 (days) and F' in Algorithms 2 and 3 is set to be 1 (month). The number of days to calculate each daily momentum estimate, T, is 20. As a result, N is 93 in all three algorithms. To match the length of the simulation horizon, we set L in Algorithm 3 to be 94 (trading months). Moreover, because of the financial crisis in 2008, the data between January 3, 2007, and December 31, 2009, are omitted to obtain a relatively valid expected return for the second simulation period. To obtain the candidate values of σ, we compute the daily volatility of roughly every non-overlapping sequence of 1886 trading days based on the historical data back to the very beginning of the S&P 500 index (Table 2.1) and the FTSE 100 index (Table 2.2) for consistency with the length of the simulation horizon. Based on these two tables, we test the maximum volatility, the minimum volatility, and some other values in between. The candidate values of daily volatility are listed in Table 2.3. The three algorithms to generate scenarios are compared on the basis of the performance of their de-biased and transformed MTD rank histogram through the W^2 value of the Cramér–von Mises test. As the scenarios generated by the Monte Carlo simulation are random, each trial is performed thirty times, with common random numbers used across algorithms for each trial, to ensure the comparability of results.

Table 2.1: Historical daily volatility of the S&P 500 returns

Period	**Daily volatility**
01/1963 - 06/1970	0.0060
07/1970 - 12/1977	0.0087
01/1978 - 06/1985	0.0088
07/1986 - 12/1993	0.0110
01/1994 - 06/2001	0.0095
07/2001 - 12/2006	0.0110

Table 2.2: Historical daily volatility of the FTSE 100 returns

Period	**Daily volatility**
01/1984 - 06/1991	0.010
07/1991 - 12/1998	0.008
01/1999 - 06/2006	0.012

Table 2.3: Candidate values of daily volatility

Candidate daily volatility (σ)	Period 1 (07/01/1999-12/26/2006) & Period 2 (01/03/2011-06/29/2018)				
S&P 500	0.0060	0.0070	0.0085	0.0100	0.0110
FTSE 100	0.0080	0.0090	0.0100	0.0110	0.0120

Figures 2.2 - 2.4 show the performance of Algorithms 1 - 3, respectively, using S&P 500 index levels from January 3, 2000, to December 31, 2018. Each panel depicts the results of non-overlapping instances, consisting of the observation and 25 scenarios, each of which is a time series of length 20. Note that, within in each figure, the difference in the estimated value of σ is the only source of changes in the shapes of the histograms. In Figures 2.2 and 2.3, the observation ranks exhibit a downward trend when daily volatility is 0.006 or 0.007, and an upward trend when daily volatility is 0.010 or 0.011. When daily volatility is 0.0085, the histogram for index returns is nearly flat. These results indicate that when estimated daily volatility is low, the autocorrelation of scenarios is less than that of the observation and scenarios are under-dispersed. When estimated daily volatility is high, the autocorrelation of scenarios becomes greater than that of the observation, and scenarios are over-dispersed. When daily volatility is 0.0085, the autocorrelation of scenarios and observations are similar and the observation is indistinguishable from the scenarios. In Figure 2.4, we observe a downward trend for index returns when daily volatility is set to be 0.006, flatness when volatility is 0.007, and an upward trend in other cases. Comparing Algorithms 1 and 2, both the volatility of generated returns and the autocorrelation of observations in each column of figures are controlled to be the same, and the frequency with which momentum estimates are updated causes the difference in the shape of histograms, indicating that this frequency has trivial influence on the autocorrelation or dispersion of scenarios. Comparing Algorithms 2 and 3, we notice that even when the frequency with which the expected return is updated is the same, using exponentially weighted average or simple moving average to calculate can cause the differences in the ensembles of index return scenarios. The rank histograms of the other datasets, including the return of the S&P 500 index in the first simulation period, and the FTSE 100 in both simulation periods, shown in the Appendix, exhibit similar behavior.

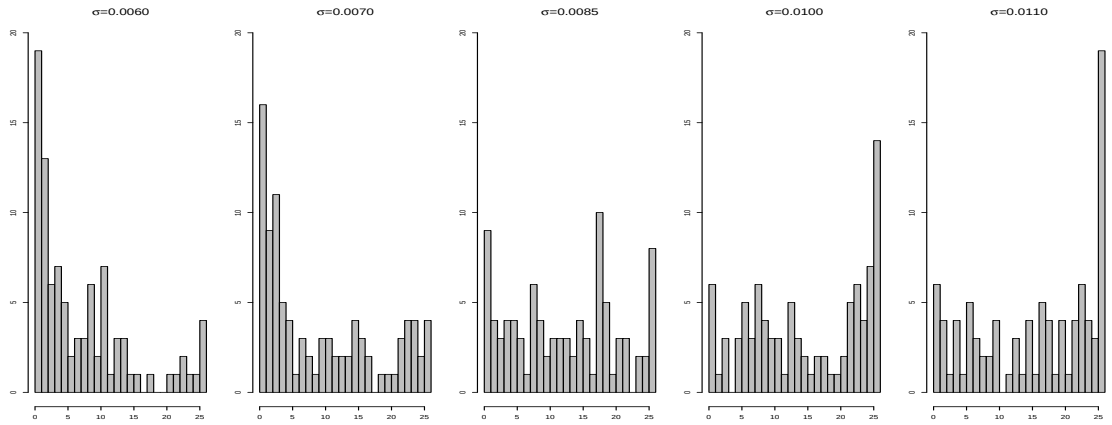


Figure 2.2: Performance of de-biased and transformed returns generated by Algorithm 1 for index return of S&P 500 in the second simulation period.

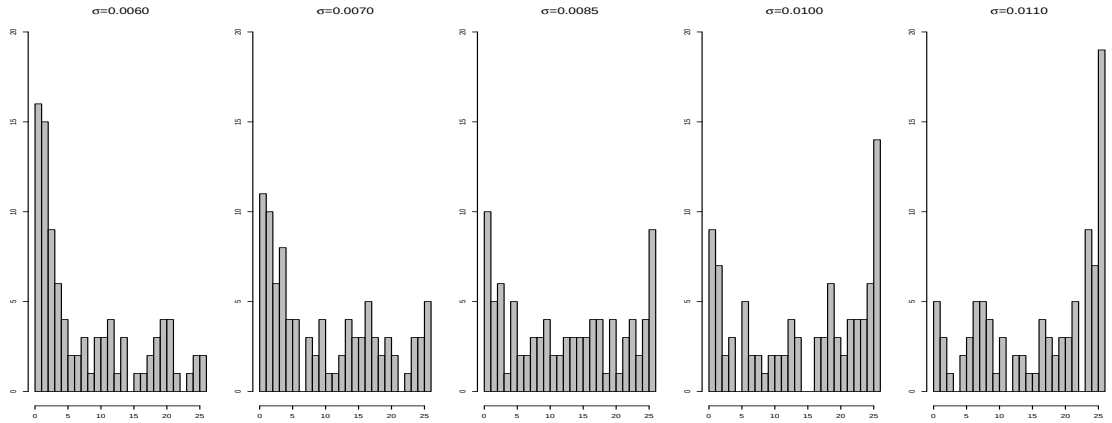


Figure 2.3: Performance of de-biased and transformed returns generated by Algorithm 2 for index return of S&P 500 in the second simulation period.

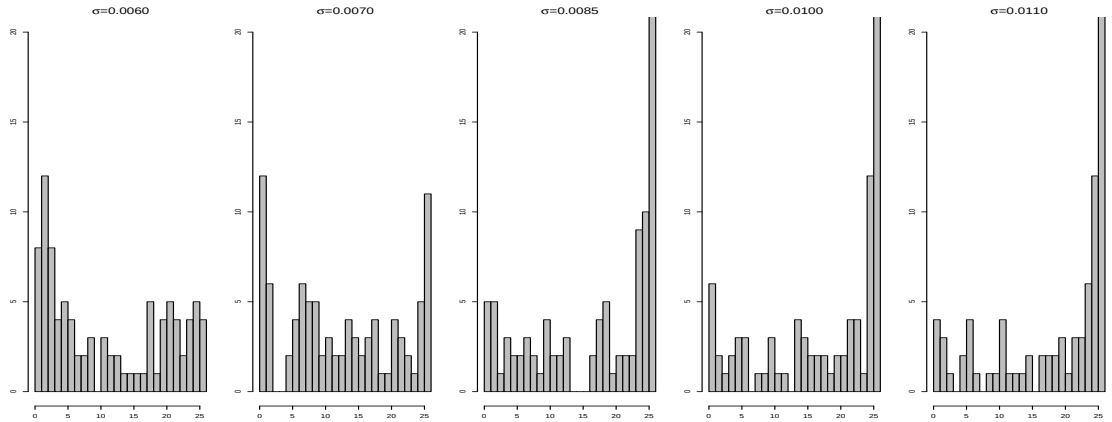


Figure 2.4: Performance of de-biased and transformed returns generated by Algorithm 3 for index return of S&P 500 in the second simulation period.

The uniformity of the de-biased and transformed MTD rank histogram is tested using the Cramér-von Mises (CvM) statistic. According to Choulakian et al. (1994), the critical values of W^2 are 0.871 for the 1% significance level, 0.743 for the 2% significance level, and 0.581 for the 5% significance level in a two-sided hypothesis test. Tables 2.4 and 2.5 display the 98% t-statistic confidence intervals for W^2 values of the ranks of observed returns of the S&P 500 and the FTSE 100. In the table bodies, underlining is used to indicate that we cannot reject the null hypothesis of uniformity at the significance level of 1%, while italic and boldface correspond to the 2% and the 5% significance level, respectively. The value of σ shown in bold in each table heading is the one closest to the observed value for that index and time period. From Tables 2.4 and 2.5, it can be seen that the performance of Algorithms 1 and 2 is similar: we cannot reject the null hypothesis of uniformity at the 5% (or lower) significance level when daily volatility is set close to the observed value of 0.0110 in Period 1 or 0.0090 in Period 2, for either index. For Algorithm 3, we fail to reject the null hypothesis of uniformity at the 5% significance level only when the estimated volatility is lower than the observed value. Scenarios generated by any of these three algorithms can be considered reliable at the 5% significance level as long as the volatility is set appropriately. The appropriate volatility for Algorithm 1 is always the largest among the three algorithms, and the appropriate volatility for Algorithm 3 is always the smallest. If we can estimate the volatility of observation correctly, the performance of Algorithm 1 and 2 will be reliable. However, even if σ is set equal to the true volatility, Algorithm 3 cannot provide reliable scenarios in any of these four settings.

For additional insight into the results, Figure 2.4 shows the time-series plots of de-biased and transformed S&P 500 index returns over a small portion of the simulation horizon, from January 2, 2013, to April 29, 2013, for each algorithm when the daily volatility is set to 0.0085 for scenario generation. In these plots, the thicker line represents the path of observation data, and other dashed lines are for the 25 generated scenario. From Figure 2.4, we can see that the scenarios and observations are intertwined with each other, and it is hard to differentiate between them as the red thicker line is sometimes above most scenarios, sometimes below the majority of scenarios, and sometimes in between the scenarios. From Figure 2.4, it can be observed that the scenarios almost always envelop the observation data before mid February but the relative position of the observation oscillates drastically in the latter portion of the time horizon, which explains why the rank histogram for Algorithm 2 is less flat than that for Algorithm 1 in this setting. Compared

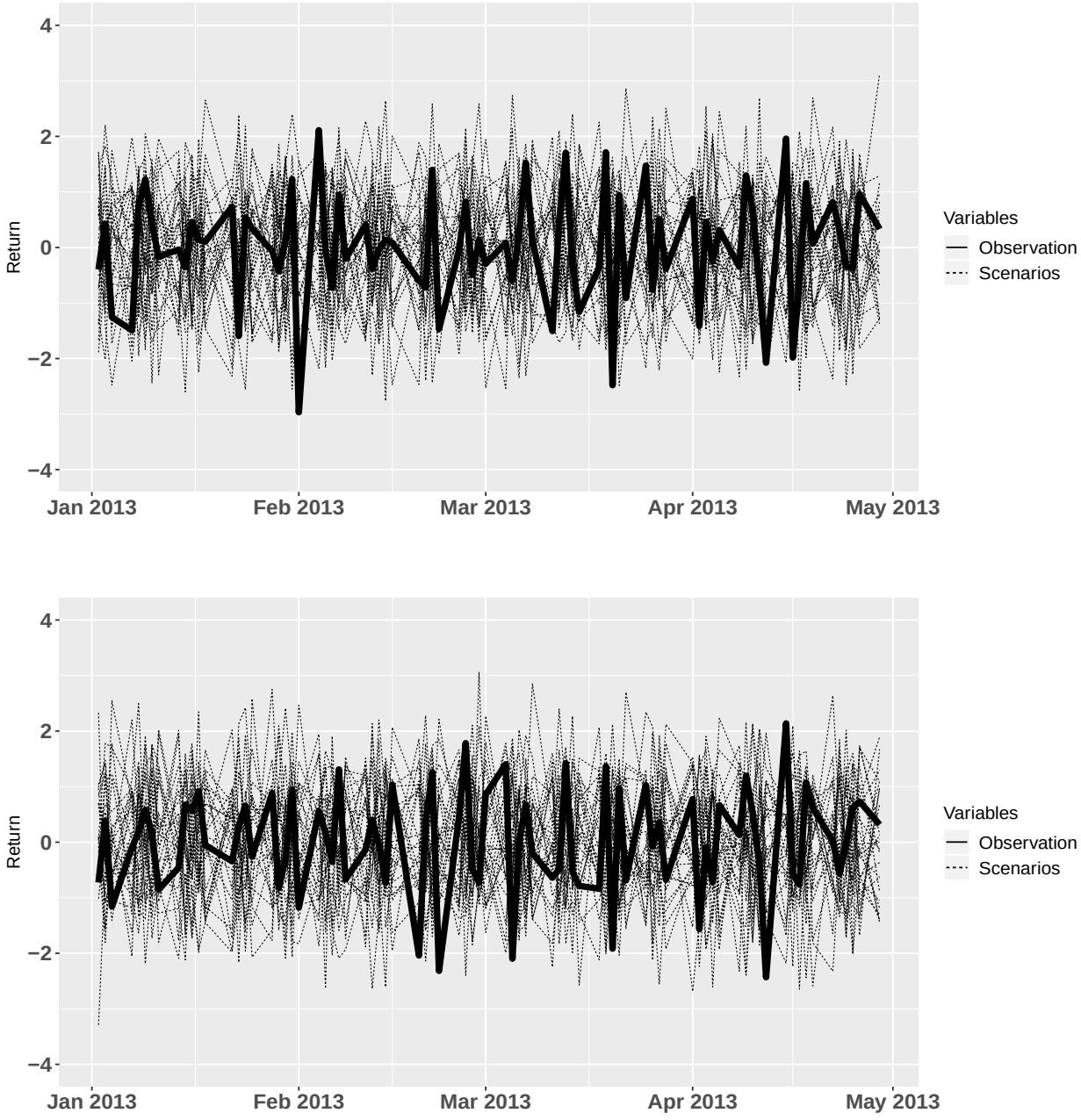
Table 2.4: 98% confidence interval of W^2 for de-biased and transformed S&P 500 index returns. Underline is used to indicate that the null hypothesis of uniformity cannot be rejected at the significance level of 1%, and italic and bold correspond to the 2% and the 5% significance level, respectively.

Period	Period 1				
Daily Volatility	0.0060	0.0070	0.0085	0.0100	**0.0110**
Algorithm 1	[11.05,12.17]	[7.02, 8.06]	[2.57, 3.63]	[1.08, 1.49]	***[0.43, 0.77]***
Algorithm 2	[10.18, 11.49]	[6.13, 7.22]	[2.38, 3.04]	[0.80, 1.10]	***[0.53, 0.76]***
Algorithm 3	[6.14, 7.32]	[2.92, 3.83]	*[0.66, 1.06]*	***[0.53, 0.78]***	[0.88, 1.34]
Period	Period 2				
Daily Volatility	0.0060	0.0070	**0.0085**	0.0100	0.0110
Algorithm 1	[4.65, 5.58]	[1.62, 2.21]	***[0.34, 0.48]***	[0.77, 1.25]	[1.96, 2.58]
Algorithm 2	[4.45, 5.31]	[1.86, 2.46]	***[0.46, 0.72]***	[0.89, 1.31]	[1.70, 2.29]
Algorithm 3	[1.33, 1.82]	***[0.36, 0.56]***	[0.91, 1.57]	[3.47, 4.23]	[5.08, 6.07]

Table 2.5: 98% confidence interval of W^2 for de-biased and transformed FTSE 100 index returns. Underline is used to indicate that the null hypothesis of uniformity cannot be rejected at the significance level of 1%, and italic and bold correspond to the 2% and the 5% significance level, respectively.

Period	Period 1				
Daily Volatility	0.0080	0.0090	0.0100	**0.0110**	0.0120
Algorithm 1	[2.96, 3.76]	[1.42, 1.86]	*[0.68, 0.96]*	***[0.50, 0.75]***	*[0.72, 1.08]*
Algorithm 2	[2.26, 2.90]	[1.02, 1.37]	***[0.57, 0.80]***	*[0.62, 0.92]*	[1.09, 1.52]
Algorithm 3	[0.80, 1.17]	***[0.52, 0.68]***	[0.76, 1.12]	[1.47, 2.09]	[2.67, 3.51]
Period	Period 2				
Daily Volatility	0.0080	**0.0090**	0.0100	0.0110	0.0120
Algorithm 1	[1.08, 1.69]	***[0.38, 0.55]***	***[0.39, 0.63]***	[1.16, 1.57]	[2.02, 2.68]
Algorithm 2	[0.85, 1.29]	***[0.40, 0.61]***	***[0.57, 0.79]***	[1.54, 2.08]	[2.15, 3.05]
Algorithm 3	***[0.23, 0.38]***	[0.91, 1.41]	[1.88, 2.60]	[4.12, 5.19]	[5.30, 6.79]

with the other two time series plots, the observation data in Figure 2.4 spends the least time being either above or below all the scenarios. The scenario data mostly surround it, indicating that the generated scenarios are over-dispersed, which is consistent with the upward slope of the MTD rank histogram for Algorithm 3 in this setting.



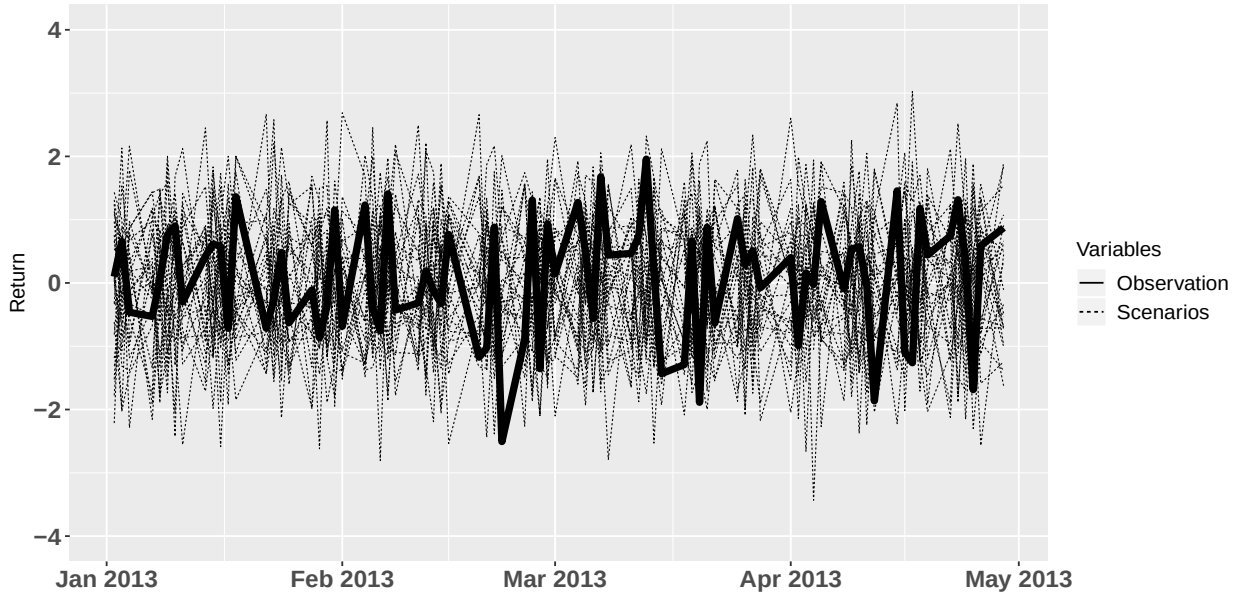


Figure 2.5: Time-series plot: 25 generated scenarios of S&P 500 index returns, represented by the dashed lines, along with the corresponding observations, represented by the thicker solid line, from Feb 2, 2011, to April 28, 2011

2.5 Conclusion

In this study, we propose a model that substitutes short-term momentum for the drift in a geometric Brownian motion process for index levels to produce a nonstationary variant of this widely used model. While the volatility can be estimated by computing the standard deviation of log ratios of index levels when the drift is constant, it is not straightforward to produce a reliable estimate of volatility to use in our momentum model. Therefore, we test different volatility estimates along with various schemes for updating the momentum or drift estimates in rolling horizon procedures for scenario generation. The performance of algorithms with different values of volatility is compared based on the MTD rank histograms.

By carrying out computational studies on the S&P 500 and the FTSE 100 indices in two different time periods, we find that the approach to update the expected index returns has a large effect on the reliability of scenarios generated, while the frequency to update the expected return does not influence the generation significantly. Besides the expectation of index returns, the volatility of

returns also plays an important role in the scenario generation process. All algorithms are able to produce reliable scenarios when the volatility is set appropriately. For Algorithm 1 and 2, as long as the volatility is set close to the true observed volatility, the results will be reliable. For Algorithm 3, the volatility parameter value that produces reliable scenarios is always smaller than the true volatility. Thus, even if we estimate the volatility accurately, Algorithm 3 does not provide reliable scenarios for any combination of stock indices and time periods tested. Therefore, Algorithms 1 and 2 seem to be more reliable in that they provide reliable scenarios as long as the volatility estimation error is small. These results provide indirect evidence for the existence of momentum in index returns, as weighting recent observations more heavily when updating expected return estimates produces better scenarios for future returns.

The lack of uniformity of the rank histograms produced even by the most suitable method tested, if the volatility is wrong, suggests that the model and scenario generation methods could be improved. Some possibilities to be investigated are to incorporate nonstationary volatility as well as momentum in the model, to obtain better estimates of volatility, or to include other factors such as mean reversion. The relationship between scenario reliability and performance of the solutions to corresponding stochastic programs also must be explored thoroughly. In this paper, the forecast horizon of 20 trading days (1 month) has been tested. Comparing the performance of the algorithms with a longer forecast horizon could be another avenue to explore.

2.6 References

Bertsimas, D. and Pachamanova, D. (2008). Robust multiperiod portfolio management in the presence of transaction costs. *Computers & Operations Research*, 35(1):3–17.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82.

Chen, H.-H. and Yang, C.-B. (2017). Multiperiod portfolio investment using stochastic programming with conditional value at risk. *Computers & Operations Research*, 81:305–321.

Choulakian, V., Lockhart, R. A., and Stephens, M. A. (1994). Cramér-von Mises statistics for discrete distributions. *Canadian Journal of Statistics*, 22(1):125–137.

Conrad, J. and Kaul, G. (1998). An anatomy of trading strategies. *The Review of Financial Studies*, 11(3):489–519.

Dantzig, G. B. and Infanger, G. (1993). Multi-stage stochastic linear programs for portfolio optimization. *Annals of Operations Research*, 45(1):59–76.

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

Fulton, L. V. and Bastian, N. D. (2018). Multiperiod stochastic programming portfolio optimization for diversified funds. *International Journal of Finance & Economics*.

Gülten, S. and Ruszczyński, A. (2015). Two-stage portfolio optimization with higher-order conditional measures of risk. *Annals of Operations Research*, 229(1):409–427.

Hamill, T. M. (2001). Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Review*, 129(3):550–560.

Hamill, T. M. and Colucci, S. J. (1997). Verification of Eta–RSM short-range ensemble forecasts. *Monthly Weather Review*, 125(6):1312–1327.

Hochreiter, R. and Pflug, G. C. (2007). Financial scenario generation for stochastic multi-stage decision processes as facility location problems. *Annals of Operations Research*, 152(1):257–272.

Hong, K. J. and Satchell, S. (2012). Defining single asset price momentum in terms of a stochastic process. *Theoretical Economics Letters*, 2(03):274.

Høyland, K. and Wallace, S. W. (2001). Generating scenario trees for multistage decision problems. *Management Science*, 47(2):295–307.

Hsu, W.-r. and Murphy, A. H. (1986). The attributes diagram a geometrical framework for assessing the quality of probability forecasts. *International Journal of Forecasting*, 2(3):285–293.

Hull, J. C. (2009). *Option, Futures and other Derivatives*. Upper Saddle River, NJ: Prentice Hall,.

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

Koijen, R. S. J., Rodríguez, J. C., and Sbuelz, A. (2009). Momentum and mean reversion in strategic asset allocation. *Management Science*, 55(7):1199–1213.

Lewellen, J. (2002). Momentum and autocorrelation in stock returns. *Review of Financial Studies*, 15(2):533–564.

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

Şakar, C. T. and Köksalan, M. (2013). A stochastic programming approach to multicriteria portfolio optimization. *Journal of Global Optimization*, 57(2):299–314.

Sarı, D., Lee, Y., Ryan, S., and Woodruff, D. (2016). Statistical metrics for assessing the quality of wind power scenarios for stochastic unit commitment. *Wind Energy*, 19(5):873–893.

Sarı, D. and Ryan, S. M. (2016). MTDrh: Mass transportation distance rank histogram. https://cran.r-project.org/web/packages/MTDrh/index.html.

Sarı, D. and Ryan, S. M. (2018). Statistical reliability of wind power scenarios and stochastic unit commitment cost. *Energy Systems*, 9(4):873–898.

Van der Weide, R. (2002). GO-GARCH: a multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics*, 17(5):549–564.

Wilks, D. S. (2004). The minimum spanning tree histogram as a verification tool for multidimensional ensemble forecasts. *Monthly Weather Review*, 132(6):1329–1340.

Yu, L.-Y., Ji, X.-D., and Wang, S.-Y. (2003). Stochastic programming models in financial optimization: A survey. *AMO - Advanced Modeling and Optimization*, 5(1):1–26.

2.7 Appendix

Figures 2.6-2.14 show the performance of the generated scenarios using Algorithms 1, 2 and 3 for the S&P 500 in the first period, and the FTSE 100 in both periods.

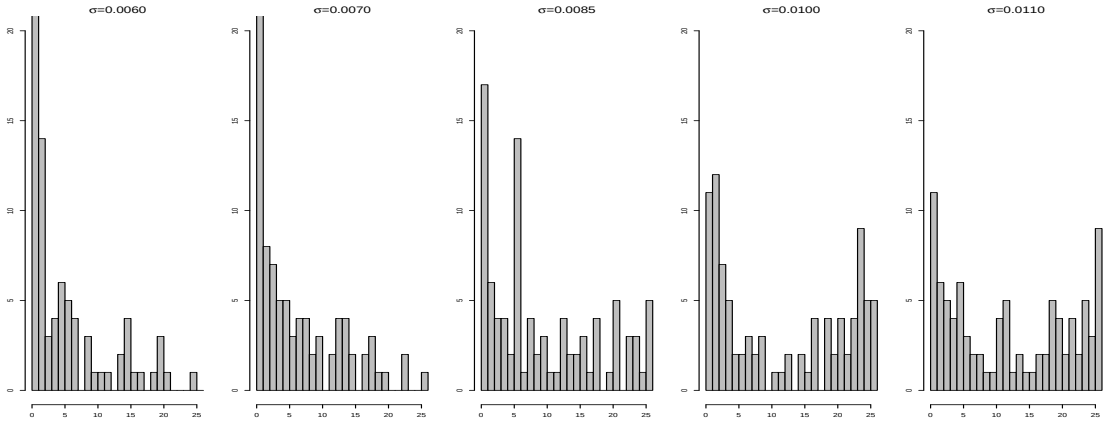


Figure 2.6: Performance of de-biased and transformed returns generated by Algorithm 1 for index return of the S&P 500 in the first simulation period.

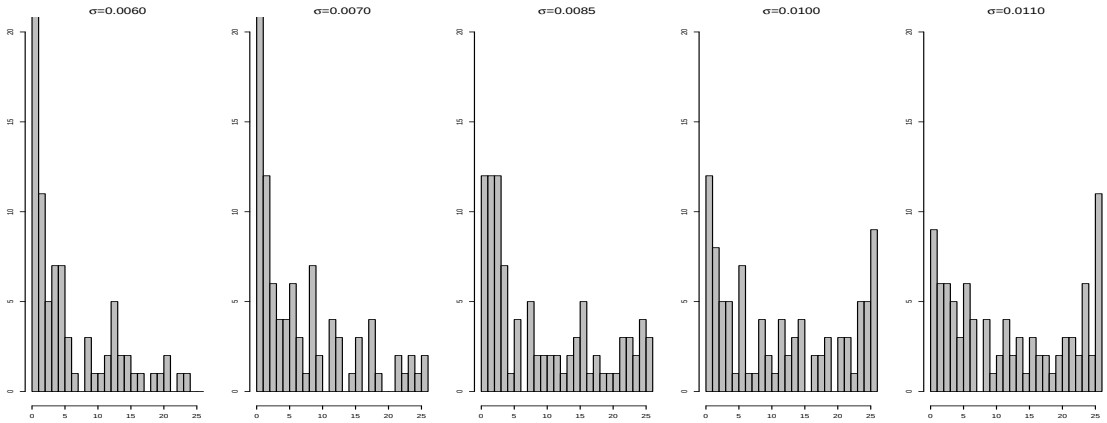


Figure 2.7: Performance of de-biased and transformed returns generated by Algorithm 2 for index return of the S&P 500 in the first simulation period.

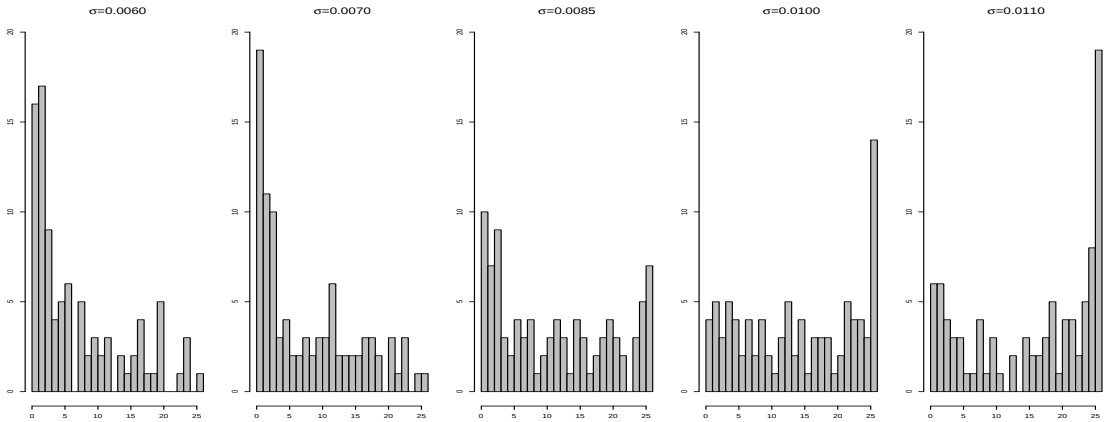


Figure 2.8: Performance of de-biased and transformed returns generated by Algorithm 3 for index return of the S&P 500 in the first simulation period.

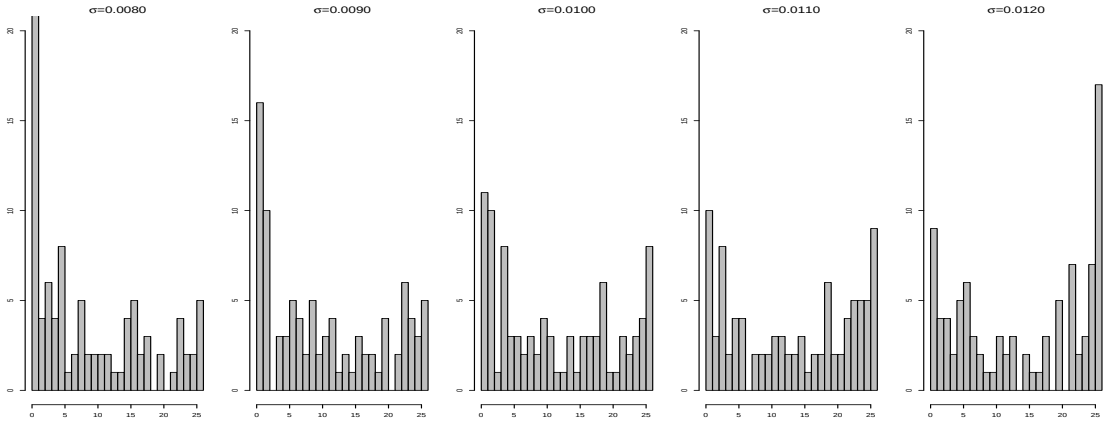


Figure 2.9: Performance of de-biased and transformed returns generated by Algorithm 1 for index return of the FTSE 100 in the first simulation period.

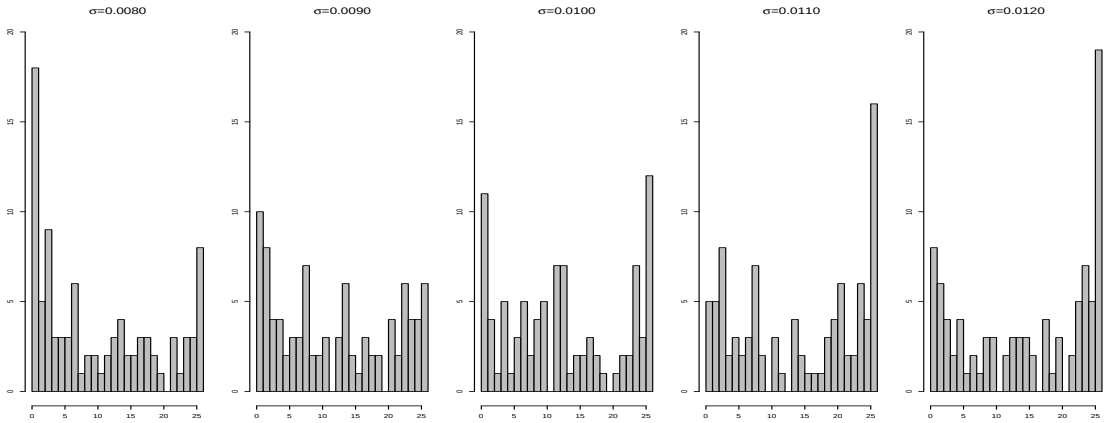


Figure 2.10: Performance of de-biased and transformed returns generated by Algorithm 2 for index return of the FTSE 100 in the first simulation period.

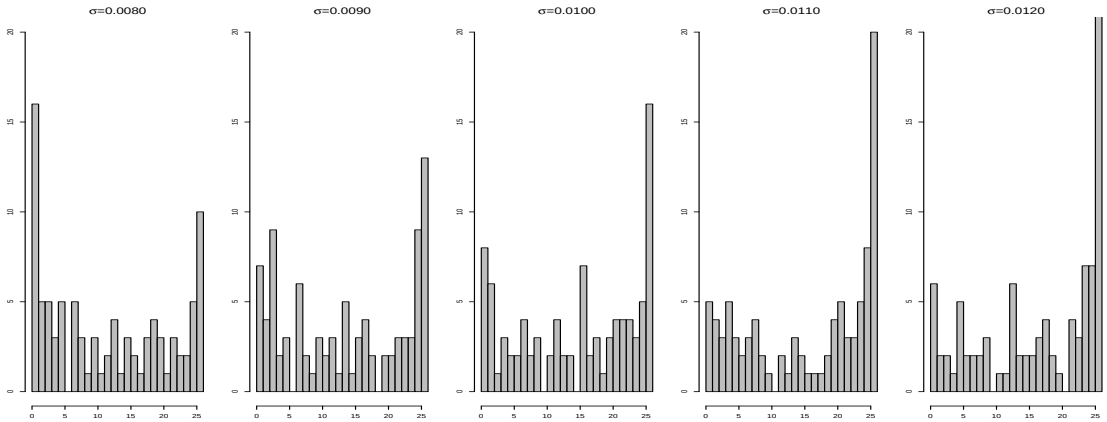


Figure 2.11: Performance of de-biased and transformed returns generated by Algorithm 3 for index return of the FTSE 100 in the first simulation period.

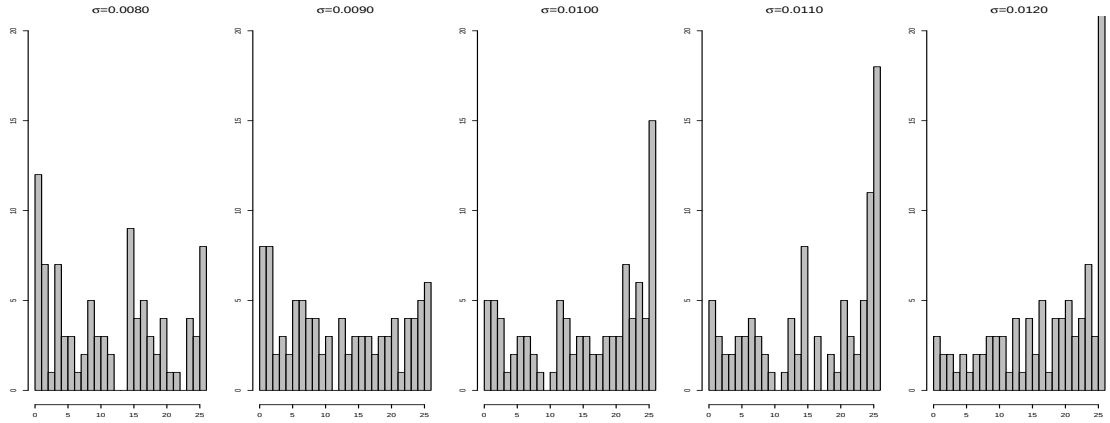


Figure 2.12: Performance of de-biased and transformed returns generated by Algorithm 1 for index return of the FTSE 100 in the second simulation period.

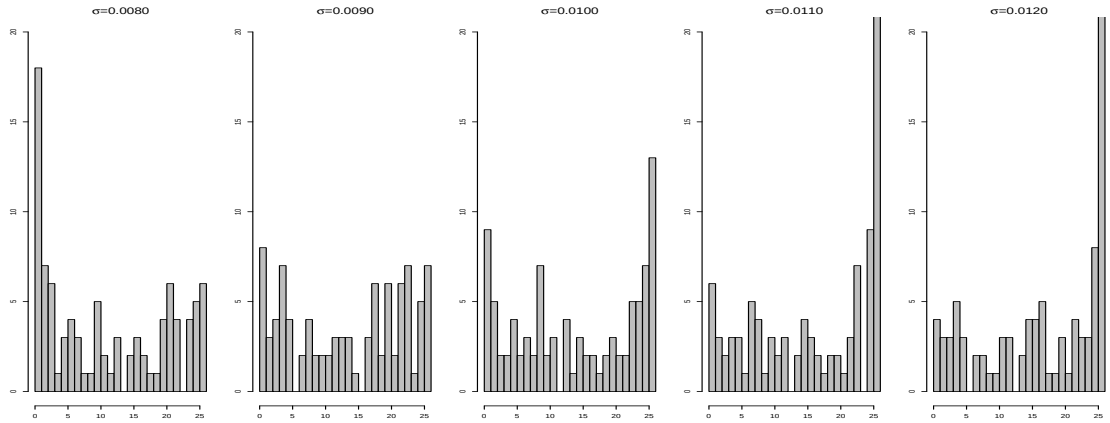


Figure 2.13: Performance of de-biased and transformed returns generated by Algorithm 2 for index return of the FTSE 100 in the second simulation period.

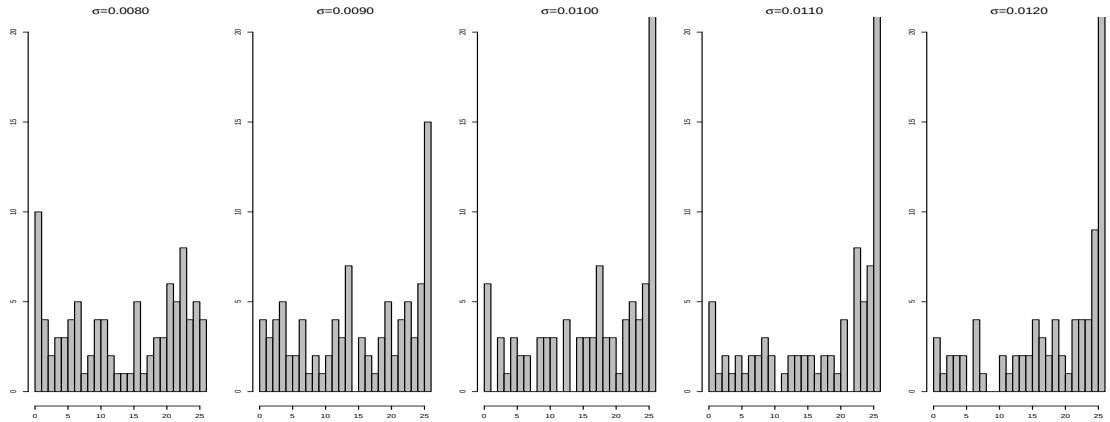


Figure 2.14: Performance of de-biased and transformed returns generated by Algorithm 3 for index return of the FTSE 100 in the second simulation period.

CHAPTER 3. PORTFOLIO REBALANCING BASED ON TIME SERIES MOMENTUM AND DOWNSIDE RISK

A manuscript under second review by *IMA Journal of Management Mathematics*

Abstract

To examine the familiar tradeoff between risk and return in financial investments, we use a rolling two-stage stochastic program to compare mean-risk optimization models with time series momentum strategies. In a backtest of allocating investment between a market index and a risk-free asset, we generate scenarios of future return according to a momentum-based stochastic process model. A new hybrid approach, time series momentum strategy controlling downside risk (TSMDR), frequently dominates traditional approaches by generating trading signals according to a modified momentum measure while setting the risky asset position to control the CVaR of return. For insight into the outperformance of TSMDR, we decompose each strategy into two aspects, the trading signal and the asset allocation model that determines the risky asset position. We find that 1) weighted moving average can better capture the trend of the stock market than time series momentum computed as past 12-month excess return, 2) mean-risk strategies generally provide better returns whereas risk parity strategies have less investment risk, and 3) controlling CVaR limits the investment risk better than controlling variance does.

3.1 Introduction

In their quest to achieve high rates of return with tolerable levels of risk, many investors rebalance their portfolios at regular intervals. Periodic rebalancing results in a dynamic asset allocation where the allocation of wealth to riskier assets can vary over time in response to changing market conditions. While broad market indicators and fundamental assessments of individual assets' prospects for growth could be considered, many asset allocation approaches rely simply on the use of previous return data to either (i) estimate parameters of stochastic models for returns to be used in optimization models or (ii) detect signals that may indicate the directions in which returns will move before the next rebalancing point.

Momentum is a term employed to describe one set of such signals that have been studied empirically and analytically over the past few decades. In contradiction to the efficient-market hypothesis, that future returns are independent of their past values, considerable evidence has accumulated in support of momentum. Empirical studies suggest the existence of two types of momentum in the short term. Time series momentum describes a phenomenon in which an asset that has been trending strongly in a certain direction will continue to move in the same direction in the short term, while cross-sectional momentum identifies the tendency for assets that have recently outperformed (or underperformed) their peers to continue to do so for a while longer. Both of these forms of momentum have been incorporated in stochastic models of asset returns (Koijen et al., 2009; He et al., 2018) and in so-called momentum strategies for asset allocation (Jegadeesh and Titman, 1993; Moskowitz et al., 2012).

At the same time as momentum strategies have emerged, many investors have adopted a passive approach to reducing risk by investing in diversified index funds rather than individual assets. In the US stock markets, the total amount invested in index funds recently overtook the amount in actively managed funds (**?**). An investor who relies on indexing to diversify their investments in risky assets is not concerned with cross-sectional momentum. But they still must consider, and periodically reevaluate, how much wealth to hold in the index fund as opposed to a risk-free alternative. The goal of this paper is to analyze and compare optimization-based and signal-based approaches for dynamically allocating assets between a risky asset (an index fund) and a risk-free asset. Both the expected future return used in the optimization models and the trading signals employed in the rule-based strategies are based on continually updated estimates of time series momentum in the index fund returns.

As a foundation for comparison, we formulate a generic mean-risk optimization problem for periodic rebalancing, of which a version of the Markowitz Mean-Variance model is a special case. We also consider Mean-CVaR optimization to employ a coherent measure of downside risk. We show that the optimal solutions to these two mean-risk formulations differ in their structure. The well-studied time series momentum strategy, TSMOM, is recast as a feasible solution to this mean-risk optimization problem. We also develop a new hybrid strategy, termed time series momentum controlling downside risk (TSMDR), as another feasible solution that combines a time series momentum trading signal with a position size based on CVaR rather than volatility. Although the

CVaR can be evaluated in closed form for normally-distributed returns, we compute it using discrete probabilistic scenarios to allow flexibility in the estimated return distribution. Also by doing so, we are able to identify methods and time windows for estimating the mean and volatility parameters that result in sets of scenarios that are evaluated as statistically reliable in a historical backtest.

We empirically study the performance of the four strategies in a historical backtest of monthly rebalancing, considering the Standard & Poor's 500 index as the risky asset and the one-month US Treasury bill as the risk-free asset. Using historical data from 2001 through 2019, we compare the results of all strategies according to multiple performance metrics. Our newly created TSMDR strategy significantly outperforms the others in terms of the Sortino ratio and the Sharpe ratio, while dominating the others in a Pareto sense for all but the highest risk tolerances. For certain values of its adjustable risk parameters, it also would have achieved the highest cumulative return over nearly all periods in the backtest. Lastly, we analyze all four trading strategies from three angles: the trading signal, the type of asset allocation model, and the risk measure. The analysis suggests that the outperformance of TSMDR is due to 1) using weighted moving average, which can better forecast the trend of the stock market than time series momentum, 2) using risk parity, which helps reduce investment risk and increase risk-adjusted returns, and 3) using CVaR to control the downside investment risk while not limiting the upside return.

The paper proceeds as follows. Section 3.2 sets the scholarly context for this work. Section 3.3 introduces the mean-risk optimization model and analyzes its mean-variance and mean-CVaR variants. It presents the two time series momentum strategies, TSMOM and TSMDR, as heuristic solutions to this model. In Section 3.4, we introduce the model for the risky asset rate of return, the scenario generation method, and procedures to assess the reliability of generated scenarios and test the stability of optimal solutions. Section 3.5 describes the performance metrics used to evaluate the results of applying the investment strategies. Section 3.6 presents the numerical results of the evaluation and comparison of the investment strategies. Section 3.7 finally concludes.

3.2 Literature Review

Investors employ the return predictability of momentum to construct trading strategies by going long in recently outperforming securities and short in recently underperforming ones. The profitability of momentum strategies across different markets and time periods (De Bondt and Thaler, 1985;

Jegadeesh and Titman, 1993; He and Li, 2015; Goyal and Jegadeesh, 2018) suggests that investors can improve the timing of their transactions by following momentum signals. In general, these momentum strategies can be categorized as cross-sectional (Jegadeesh and Titman, 1993) or time series (Moskowitz et al., 2012). They differ in whether the strategy constructs portfolios based on a security's historical performance relative to other securities or only to itself. Time series momentum strategies focus on the return predictability of a single risky asset while cross-sectional momentum strategies focus on the relative returns of securities in the cross-section.

Moskowitz et al. (2012) report that the time series momentum strategy (TSMOM) is significantly more profitable than the cross-sectional momentum strategy across diverse asset classes, including equity indices, bonds, currencies and commodities, especially during the extreme periods. In this strategy, labeled TSMOM(k, h), the trading signal is the total excess return over a lookback period of k months. Based on this signal, traders hold a long or short position for h months. The size of the position in TSMOM(k, h) is determined by volatility scaling, which can be seen as a risk parity approach (Kim et al., 2016), in which asset allocation is determined according to relative risk contribution to control the investment risk (Asness et al., 2012).

Despite the remarkable performance of TSMOM, it does not explicitly model the assets' future returns. In this paper, we adopt the flexibility of stochastic portfolio optimization in the representation of the assets' future returns and build three alternative rebalancing strategies, including a hybrid that improves on TSMOM with respect to the trading signal, the risk measure and the approach to balancing the trade-off between risk and return.

First, the trading signal in TSMOM is the sign of the past k-month excess returns. Alternatively, we can estimate the distribution of the future rate of return according to a stochastic process model and construct the trading signal based on the sign of the difference between the expectation of the estimated distribution of the asset future return rates and the risk-free rate.

Second, TSMOM uses volatility of excess returns as the risk measure. However, much theory and experience indicate that volatility (as the square root of variance) is not an ideal risk measure, because it treats the chance of success and the risk of failure equivalently, which does not accurately reflect most investors' attitudes towards risk (see Kolm et al. (2014) for an overview of portfolio optimization developments over the past 60 years). Therefore, various downside risk measures have been proposed (Roy, 1952; Nawrocki, 1999; Artzner et al., 1999). One of these, the value-

at-risk (VaR) with probability α, corresponds to an upper estimate of losses which is exceeded with probability α. However, since VaR does not consider information beyond this upper estimate, it cannot distinguish among different extents of losses beyond the VaR (Rockafellar and Uryasev, 2002). What is worse, Artzner et al. (1999) prove that VaR is not a coherent risk measure, as it does not satisfy the subadditivity property which is consistent with the investment principle that diversification can reduce risk. Rockafellar and Uryasev (2000) propose an improved risk measure, conditional value-at-risk (CVaR), which is coherent (Pflug, 2000). CVaR is defined as the expected loss beyond the VaR, and for a discrete distribution of loss it can be computed by solving a linear program (Rockafellar and Uryasev, 2002). Hence, we consider CVaR as an additional risk measure.

Third, TSMOM employs the risk parity approach to allocate assets. The risk parity approach uses solely the measure of risk to determine the proportion of wealth to be allocated in each asset and does not require the estimation of expected returns, which significantly reduces the estimation error. Nevertheless, this approach has very limited power when helping investors seek for high returns. Accordingly, we also consider mean-risk optimization, where the objective is to minimize a convex combination of the expectation of negative investment return and the risk of investment loss.

In this paper, two stochastic portfolio optimization models are formulated along with the hybrid strategy, TSMDR. They all use the sign of the difference between the expectation of the estimated distribution of the asset future return rates and the risk-free rate as the trading signal but vary in the risk measures and asset allocation approaches: TSMDR uses CVaR and a risk parity approach; Mean-Variance uses variance and a mean-risk tradeoff; and Mean-CVaR uses CVaR and a mean-risk tradeoff.

The advantage of stochastic programming models is to incorporate more of the distribution of the random parameter, the future rate of return of the risky asset. The success of TSMOM suggests that this random parameter can be modeled by the past returns of the same asset, which is partially validated by Koijen et al. (2009). Our model of the asset rate of return is inspired by that of Koijen et al. (2009), which characterizes both short-run momentum and long-run mean reversion, but different in three aspects. First, we incorporate the effect of mean reversion in the periodically renewed observations rather than regarding it as a state variable. Second, Koijen et al. (2009) use exponential moving average to estimate momentum. However, by comparing different moving

average techniques in terms of the percentage of positive excess returns (prediction accuracy) and cumulative absolute log returns, we find that weighted moving average, where weights on more remote past observations decrease in arithmetical progression, can better capture the trend of the stock market. There is no consensus on how to quantify the trend of a single asset's performance (e.g., Moskowitz et al. (2012) compute time series momentum based on rates of return differently from Koijen et al. (2009), **?** use simple moving averages of prices to detect the market trend, and **?** analyze moving averages of prices in terms of the equivalent weights on return rates). In this paper we adopt the time series momentum definition of Moskowitz et al. (2012), as (total) excess return over a fixed time period in the past. Third, Koijen et al. (2009) assume that the volatility of risky asset returns is constant. To respond to recent changes, we allow volatility to be non-stationary and estimated on a rolling basis. Consequently, our model of the risky asset's future rate of return reduces to a normal distribution with non-stationary mean, estimated by weighted moving average of recent historical rates of return, and a time-dependent volatility.

As it is usually intractable to solve a stochastic program with continuous distributions for the uncertain parameters, these distributions are often discretized or approximated by a finite number of probabilistic scenarios. Plentiful methods for scenario generation (Yu et al., 2003; Kaut and Wallace, 2007; Roman et al., 2010; Gülten and Ruszczyński, 2015) have been proposed and applied to portfolio optimization problems. Among them, the Monte Carlo method, which simulates the distribution by random sampling, is the most straightforward approach that has been used in many studies (Dantzig and Infanger, 1993; Bertsimas and Pachamanova, 2008; **?**; Guo and Ryan, 2020).

After generating scenarios, the next question arising is how well they approximate the distribution. The verification rank histogram is a recently-developed technique for evaluating scenario quality that compares sets of scenarios that would have been generated in the past with the corresponding observations. Mass transportation distance (MTD) rank histograms, developed by Sarı et al. (2016), use the distance between scenario sets and the corresponding observations measured by MTD, also known as the Wasserstein distance, which accounts for the scenario probabilities. The principle behind the rank histogram is that an ideal set of scenarios for a given problem instance and its associated observation could be regarded as a random sample from the same distribution. If the observation is indistinguishable from a scenario, then the rank of the observation-to-scenarios distance, among distances of each single scenario to the rest (including the observation) is uniformly

distributed. Accordingly, the quality of a scenario generation method can be measured by the flatness of the MTD rank histogram over multiple instances. In this paper, we set a parameter value for the scenario generation method based on a hypothesis test of uniform MTD ranks.

However, King and Wallace (2012) argue that simply evaluating the scenarios themselves without reference to the optimization model may not be sufficient to test the quality of scenarios, because the quality of a set of scenarios ultimately is determined by the resulting solution to the optimization problem. In addition, the randomness in the scenario generation procedure may cause inconsistency in the optimal solutions found when using different sets of scenarios, and thus, a loss of confidence in the true optimality of a solution obtained with any particular scenario set. Therefore, we further check the stability of the optimal solutions (Kaut and Wallace, 2007). The optimal objective value is frequently the focus of stability analysis because there may be a large number of potentially optimal solutions, and the value of the objective function presumably is what the decision maker cares about. However, in this paper, we are mainly interested in the optimal solutions themselves, as they are what guide investors in how to allocate assets at each time point. We demonstrate simple structures for the solutions of both the Mean-Variance and Mean-CVaR models that allow the stability analysis to focus on the similarity among the optimal solutions. Mindful of the "fragility" of CVaR observed by Lim et al. (2011), we generate large samples of scenarios to overcome it.

This paper is closely related to He et al. (2018), who modify the asset price model of Koijen et al. (2009) and derive the optimal investment strategy. Like Koijen et al. (2009) and He et al. (2018), we consider investments in only two assets, a single risky asset and a risk-free asset, to avoid the ambiguity of time series momentum strategies applied to multiple risky assets in that the proportion of wealth allocated each risky asset is unclear and unconstrained. This paper, however, differs from He et al. (2018) in several ways. First, He et al. (2018) build a continuous-time asset price model incorporating both short-term time series momentum and long-term mean reversion components. In the context of regular rebalancing, we focus on the effect of time series momentum and build a short-term model of risky asset returns using time series momentum alone. Furthermore, like Koijen et al. (2009), He et al. (2018) assume the volatility of the risky asset to be constant. We estimate the nonstationary volatility as the standard deviation of the difference between the risky asset returns and the momentum estimated at each time point using recent historical data. We verify the return model with non-stationary time series momentum component and time-varying

volatility and methods for estimating its parameters using the MTD rank histograms. Finally, He et al. (2018) formulate a long-term optimization model to maximize the constant relative risk aversion (CRRA) utility of future wealth. In the special case of log utility, the resulting optimal time-varying proportion of wealth to invest in the risky asset is the ratio of the nonstationary excess return estimate to the fixed estimate of the variance of risky asset returns. We observe a similarity to the optimal solution for our Mean-Variance model, which provides insight into the high risk tolerance implicit in the log utility function.

To the best of our knowledge, this paper is the first to use the flexible representation of uncertain parameters in stochastic programming to improve on TSMOM from three perspectives: trading signal, risk measure, and the approach to balancing risk and return.

3.3 Rebalancing strategies

Given a risky asset and a risk-free alternative, we consider four strategies to dynamically rebalance a portfolio that varies between 100% long and 100% short in the risky asset. At the beginning of each period, we decide on a long or short position to hold in the risky asset for the period. We define w_t as the proportion of current wealth to invest in the risky asset at the beginning of period t. If $w_t > 0$ (< 0), the investor longs (shorts) the risky asset and $|w_t|$ is the size of the position taken. The generic problem, to be solved repeatedly for each t, is to determine w_t by balancing the trade-off between maximizing excess returns and minimizing the investment risk, where the rate of return of the risky asset over period t, R_t, is uncertain and the rate of return of the risk-free asset, f_t, is known at the beginning of the period. We denote the beginning of period t as time t and its end as time $t + 1$. The information release schedule and the time point to make a decision of each period are illustrated in Figure 4.2, where r_t is the realized value of R_t.

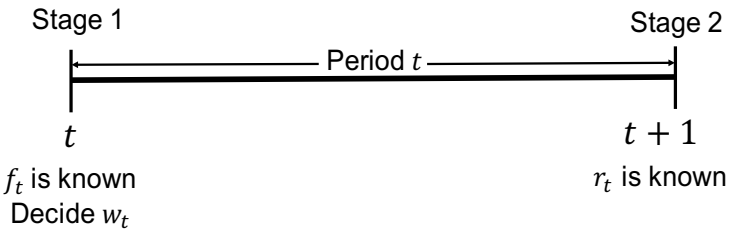


Figure 3.1: Timeline of period t

At each time t, we can either 1) estimate the distribution of R_t according to a stochastic process model and determine the asset allocation by solving a stochastic program, or 2) simply generate a trading signal (sign of w_t) and position size for the risky asset according to a time series momentum strategy. Both approaches are grounded on excess return rates of the risky asset before time t. Our goal is to compare optimal solutions to a two-stage stochastic programming formulation, using various risk measures (including no risk measure) and weights assigned to them, with variants of the time series momentum strategy according to different ways of generating the trading signal and scaling the position. For simplicity to focus on the comparison of strategies, we omit any considerations of transaction costs, management fees, taxes or any other investment costs.

3.3.1 Mean-Risk Stochastic Optimization Model

The purpose of the mean-risk stochastic optimization model is to find an allocation of wealth between the two assets that maximizes the expected excess return while minimizing the investment risk. Define a risk measure, ρ, and a risk-aversion parameter $\lambda \in [0, 1]$. Then the generic mean-risk optimization model to be solved at time t is

$$\min_{w_t} (1-\lambda)\mathrm{E}[-(R_t - f_t)w_t] + \lambda\rho[-(R_t - f_t)w_t] \tag{3.1a}$$

$$\text{s.t.} \quad -1 \leq w_t \leq 1 \tag{3.1b}$$

Here, $(R_t - f_t)w_t$ is the excess return to be realized at the end of period t. Given monetary wealth, D_t, at the beginning of period t, if $w_t \geq 0$ we long a monetary value $w_t D_t$ of the risky asset while investing $(1 - w_t)D_t$ in the risk-free asset. Then, at the end of period t, the wealth is $D_{t+1} = w_t D_t(1 + R_t) + (1 - w_t)D_t(1 + f_t)$ and the rate of return is $\frac{D_{t+1}-D_t}{D_t} = w_t(R_t - f_t) + f_t$. Therefore, the excess return rate is $w_t(R_t - f_t)$. Similarly, if $w_t < 0$, we borrow $w_t D_t$ worth of the risky asset and sell it at the beginning of period t to be repurchased at time $t + 1$. Therefore, we will have $(1 - w_t)D_t > D_t$ invested in the risk-free asset during period t. The expression for D_{t+1} is identical to the case where $w_t \geq 0$, and hence, the excess return equals $w_t(R_t - f_t)$ in either case. Constraint (4.15b) limits the absolute position size of the risky asset to be no larger than 1.

If the risk measure is coherent, a simple structure for optimal solutions is found.

Proposition 1. *For a coherent risk measure, ρ, one of the three values 1, 0, or -1 is optimal for model (4.15a) - (4.15b).*

Proof. The linearity of expectation and coherence properties of ρ allow the objective to be simplified as:

$$
\begin{aligned}
&\min_{-1\le w_t\le 1}(1-\lambda)\mathrm{E}[-(R_t-f_t)w_t]+\lambda\rho[-(R_t-f_t)w_t] \\
&=\min_{-1\le w_t\le 1}(1-\lambda)f_tw_t-(1-\lambda)\mathrm{E}[R_t]w_t+\lambda\rho[-(R_t-f_t)w_t] \text{ (Linearity of E)} \\
&=\min_{-1\le w_t\le 1}(1-\lambda)f_tw_t-(1-\lambda)\mathrm{E}[R_t]w_t+\lambda f_tw_t+\lambda\rho[-R_tw_t] \text{ (Translation invariance of } \rho) \\
&=\min_{-1\le w_t\le 1}\begin{cases}\Big(f_t-(1-\lambda)\mathrm{E}[R_t]+\lambda\rho[-R_t]\Big)w_t, & \text{if } w_t\ge 0\\ \Big(f_t-(1-\lambda)\mathrm{E}[R_t]-\lambda\rho[R_t]\Big)w_t, & \text{if } w_t<0\end{cases} \text{ (Positive homogeneity of } \rho)
\end{aligned}
$$

Denote $c_t^1(\lambda)\equiv(\lambda-1)\mathrm{E}[R_t]+f_t+\lambda\rho[-R_t]$ and $c_t^2(\lambda)\equiv(\lambda-1)\mathrm{E}[R_t]+f_t-\lambda\rho[R_t]$. By inspection, it is optimal to weigh the risky asset as follows:

$$
w_t^*=\begin{cases}-1, & \text{if } c_t^2(\lambda)>0 \text{ and } c_t^1(\lambda)>-c_t^2(\lambda)\\ 0, & \text{if } c_t^1(\lambda)\ge 0 \text{ and } c_t^2(\lambda)\le 0\\ 1, & \text{if } c_t^1(\lambda)<0 \text{ and } c_t^2(\lambda)<-c_t^1(\lambda)\end{cases}
$$

□

3.3.1.1 Mean-Variance Model

The Mean-Variance model, a variant of the one proposed by Markowitz (1952), results from setting ρ in (4.15a) to be variance. Model (4.15a)-(4.15b) simplifies to

$$
\min_{w_t}\ -(1-\lambda)(\mathrm{E}[R_t]-f_t)w_t+\lambda\mathrm{Var}[R_t]w_t^2 \tag{3.2a}
$$

$$
\text{s.t.}\quad -1\le w_t\le 1 \tag{3.2b}
$$

Because variance is not a coherent risk measure, Proposition 1 does not apply. But optimal solutions for model (3.2a) - (3.2b) can be found in terms of $c_t^3(\lambda)\equiv\frac{(1-\lambda)(\mathrm{E}[R_t]-f_t)}{2\lambda\mathrm{Var}[R_t]}$ as follows:

$$w_t^* = \begin{cases} -1, & \text{if } c_t^3(\lambda) < -1 \\ c_t^3(\lambda), & \text{if } c_t^3(\lambda) \in [-1, 1] \\ 1, & \text{if } c_t^3(\lambda) > 1 \end{cases}$$

Note that, in contrast to coherent ρ, optimal solutions to the Mean-Variance model may take absolute values between 0 and 1.

Additionally, when the objective function is to maximize the constant relative risk aversion (CRRA) utility function and the return rate of the risky asset has a constant drift and volatility over time, **?** proves that the optimal proportion in the risky asset is $c_t^4(\gamma) = \frac{(\mathrm{E}[R_t]-f_t)}{\gamma \mathrm{Var}[R_t]}$, where γ is the constant coefficient of relative risk aversion. He et al. (2018) find a similar result for $\gamma = 1$ (log utility function), in the absence of mean reversion, but with nonstationary drift for the risky asset return rate to capture the time series momentum effect. To compare our result with He et al. (2018), we can set $\lambda = \frac{1}{3}$ by equating $c_t^4(\gamma)$ with $c_t^3(\frac{\gamma}{2+\gamma})$, .

3.3.1.2 Mean-CVaR Model

Although variance (or standard deviation, also called volatility) is a traditional risk metric, it may not reflect an investor's actual perspective on the risk. Over the past three decades, researchers in finance, economics and psychology have noted that individuals view return dispersion asymmetrically; that is, losses weigh more heavily than gains (Harlow, 1991). The equal treatment of gains and losses in the variance calculation does not characterize the investors' attitudes towards risk accurately. Therefore, various downside risk measures have been proposed (Roy, 1952; Nawrocki, 1999; Artzner et al., 1999).

Conditional value-at-risk (CVaR), which quantifies the expected loss under the extreme cases, is one type of downside risk measure that has the added benefit of being coherent (Pflug, 2000). Rockafellar and Uryasev (2000) introduce the CVaR of a continuous loss random variable, Z, at a given probability, α, to protect against outcomes worse than the value-at-risk (VaR), η, as

$$\mathrm{CVaR}_\alpha[Z] = \mathrm{E}[Z|Z \geq \eta] \tag{3.3}$$

Here, η can be interpreted as the maximum loss associated with the specified probability of α, and is equivalent to the α-quantile of the loss distribution.

Later, Rockafellar and Uryasev (2002) define CVaR for a general distribution by transforming the computation of CVaR into an optimization problem:

$$\mathrm{CVaR}_\alpha[Z] = \min_{\eta \in R} \left\{ \eta + \frac{1}{1-\alpha} \mathrm{E}[(Z-\eta)_+] \right\} \tag{3.4}$$

Accordingly, when ρ represents CVaR in (4.15a), the model using probability α considers the values in the lower $(1-\alpha)$-probability tail of the excess return distribution:

$$\min_{w_t, \eta_{t+1}} \ (1-\lambda)\mathrm{E}[-(R_t - f_t)w_t] + \lambda \left\{ \eta_{t+1} + \frac{1}{1-\alpha} \mathrm{E}[(-(R_t - f_t)w_t - \eta_{t+1})_+] \right\} \tag{3.5a}$$

$$\text{s.t.} \quad -1 \leq w_t \leq 1 \tag{3.5b}$$

where η_{t+1} represents the α-quantile of the negative excess return distribution realized at time $t+1$.

If the distribution of Z is discrete, (3.5a) can be reformulated as a linear program. Here, by approximating the distribution of R_t with J scenarios, $r_{t,1}, \ldots, r_{t,J}$, where $r_{t,j}$ occurs with probability p_j and $\sum_{j=1}^{J} p_j = 1$, we can rewrite model (3.5a)-(3.5b) as

$$\min_{w_t, \eta_{t+1}, \{\nu_{t+1,j}\}_{j=1}^{j=J}} \ \lambda \eta_{t+1} + \sum_{j=1}^{J} p_j \left[-(1-\lambda)(r_{t,j} - f_t)w_t + \frac{\lambda}{1-\alpha} \nu_{t+1,j} \right] \tag{3.6a}$$

$$\text{s.t.} \quad -1 \leq w_t \leq 1 \tag{3.6b}$$

$$\nu_{t+1,j} \geq -(r_{t,j} - f_t)w_t - \eta_{t+1}, \ j = 1, \ldots, J \tag{3.6c}$$

$$\nu_{t+1,j} \geq 0, \ j = 1, \ldots, J \tag{3.6d}$$

Constraints (3.6c) and (3.6d) combine to compute the deviation of excess returns below this value in scenario j as a non-negative simple recourse variable: $\nu_{t+1,j} = [-(r_{t,j} - f_t)w_t - \eta_{t+1}]_+$. Proposition 1 has the following corollary for the Mean-CVaR model.

Corollary 1. *Suppose the risk measure, ρ, in Proposition 1 is* CVaR_α. *Define* $d^+_{\alpha,t} \equiv \mathrm{CVaR}_\alpha[R_t] - \mathrm{E}[R_t] \geq 0$ *and* $d^-_{\alpha,t} \equiv \mathrm{CVaR}_\alpha[-R_t] + \mathrm{E}[R_t] \geq 0$. *Then in model (4.15a) - (4.15b) it is optimal to weigh the risky asset as:*

$$w_t^* = \begin{cases} -1, & \text{if } \mathrm{E}[R_t] - f_t < -\lambda d^+_{\alpha,t} \\ 0, & \text{if } -\lambda d^+_{\alpha,t} \leq \mathrm{E}[R_t] - f_t \leq \lambda d^-_{\alpha,t} \\ 1, & \text{if } \lambda d^-_{\alpha,t} < \mathrm{E}[R_t] - f_t \end{cases}$$

Proof. Let $q_t(\lambda) = (\lambda - 1)\mathrm{E}[R_t] + f_t$. Then the conclusion of the proof of Proposition 1 can be restated as

$$w_t^* = \begin{cases} -1, & \text{if } q_t(\lambda) > \max\{\lambda\rho[R_t], \frac{\lambda}{2}(\rho[R_t] - \rho[-R_t])\} \\ 0, & \text{if } -\lambda\rho[-R_t] \leq q_t(\lambda) \leq \lambda\rho[R_t] \\ 1, & \text{if } q_t(\lambda) < \min\{-\lambda\rho[-R_t], \frac{\lambda}{2}(\rho[R_t] - \rho[-R_t])\} \end{cases}$$

Specifying ρ as CVaR_α, we have $\rho[R_t] = \mathrm{E}[R_t] + d_{\alpha,t}^+$ and $\rho[-R_t] = -\mathrm{E}[R_t] + d_{\alpha,t}^-$. Then:

$$\begin{aligned} \rho[R_t] - \rho[-R_t] &= 2\mathrm{E}[R_t] + d_{\alpha,t}^+ - d_{\alpha,t}^- \\ \max\{\lambda\rho[R_t], \frac{\lambda}{2}(\rho[R_t] - \rho[-R_t])\} &= \lambda\rho[R_t] \\ \min\{-\lambda\rho[-R_t], \frac{\lambda}{2}(\rho[R_t] - \rho[-R_t])\} &= -\lambda\rho[-R_t] \end{aligned}$$

Hence, it is optimal to set

$$w_t^* = \begin{cases} -1, & \text{if } q_t(\lambda) > \lambda\rho[R_t] \\ 0, & \text{if } -\lambda\rho[-R_t] \leq q_t(\lambda) \leq \lambda\rho[R_t] \\ 1, & \text{if } q_t(\lambda) < -\lambda\rho[-R_t]. \end{cases}$$

The result follows from substitution. □

Note that, unlike the formulation to minimize risk subject to a constraint on expected return as analyzed by Rockafellar and Uryasev (2000), where the optimal solutions for variance and CVaR are identical for normally distributed returns, in our formulation to minimize a convex combination of mean and risk, the structure of optimal solutions differs according to the risk measure. Specifically, if R_t is normal distribution, $d_{\alpha,t}^+ = \mathrm{d}_{\alpha,t}^- = k(\alpha)\sigma_t$ (Rockafellar and Uryasev, 2000). In our Mean-Variance model, the continuous w_t^* depends on the expected excess return in relation to variance, while in the Mean-CVaR model the discrete w_t^* depends on expected excess return in relation to volatility.

3.3.2 Variants of Time Series Momentum Strategy

In this paper we consider a time series momentum strategy as a heuristic approach to solving the mean-risk optimization model. In this approach, the sign of past excess returns is taken as the

trading signal to determine whether w_t is positive or negative, and the position size is chosen to manage the risk.

3.3.2.1 Time Series Momentum Strategy (TSMOM)

Moskowitz et al. (2012) studied time series momentum strategies with various values of the lookback interval (number of months over which to estimate the trading signal) and holding interval (number of months to hold a position). Here, to avoid complications of overlapping holding intervals, we consider only a special case in which the holding interval is equal to the time period for our optimization problem. To illustrate TSMOM$(k, 1)$ as proposed by Moskowitz et al. (2012), we first define the trading signal in period t as (Moskowitz et al., 2012)

$$a_t \equiv sign\left(\Pi_{i=1}^{k}(r_{t-i}+1) - \Pi_{i=1}^{k}(f_{t-i}+1)\right) \tag{3.7}$$

which is the sign of excess returns over the past k periods.

The position size of the risky asset in period t is set to be inversely proportional to the asset's ex-ante volatility, $\frac{C}{\widetilde{\sigma}_t}$. This volatility scaling is a form of risk parity approach (Kim et al., 2016) as it helps to produce a time series with relatively stable volatility so that the strategy is not dominated by a few volatile periods. Here, as in He et al. (2018), we set

$$\widetilde{\sigma}_t^{\,2} = N\sum_{i=0}^{\infty}(1-\delta)\delta^i(r_{t-1-i} - f_{t-1-i} - \overline{r_t})^2 \tag{3.8}$$

where the factor, N, the number of periods in a year, scales the variance to be annual. According to Moskowitz et al. (2012), the parameter δ is chosen so that the center of mass of the weights $(\frac{\delta}{1-\delta})$ is 60 trading days. Likewise, $\bar{r}_t$ is the exponential moving average of excess return:

$$\overline{r_t} = \sum_{i=0}^{\infty}(1-\delta)\delta^i(r_{t-1-i} - f_{t-1-i}) \tag{3.9}$$

Based on Hurst et al. (2013) and He et al. (2018), the scale, C, is an annualized volatility of the realized rate of return of the risky asset over the simulation horizon, calculated analogously to Equation (3.8). Kim et al. (2016) show that the choice of the value of C will affect the profitability of TSMOM. In our numerical study, we compare the performance of various quartiles of $\widetilde{\sigma}_t$ in a separate training dataset as the values of C. This allows C to represent the annualized volatility through all periods without data leakage. To ensure the amount of investment in the risky asset

cannot exceed the total wealth, we constrain the absolute value of the position size of the risky asset to be no larger than 1. Therefore, the position size of period t is given by

$$p_t = \min\left\{1, \left|\frac{C}{\widetilde{\sigma}_t}\right|\right\}. \tag{3.10}$$

Consequently, the prescribed allocation in the risky asset at the beginning of period t is

$$w_t^* = a_t p_t. \tag{3.11}$$

3.3.2.2 Time Series Momentum Controlling Downside Risk Strategy (TSMDR)

TSMOM$(k, 1)$ uses volatility as the risk measure, which has the same drawback as the Mean-Variance model in that it regards downside risk and upside reward of the investment equally. Motivated by Mean-CVaR, we consider CVaR as the risk measure to compute the position size in TSMOM$(k, 1)$ to focus on downside risk.

The new trading signal of period t is the sign of the difference between the expectation of the distribution of the rate of return of the risky asset and the risk-free rate of period t:

$$b_t \equiv sign(\mathrm{E}[R_t] - f_t) \tag{3.12}$$

Investment risk in period t is measured by the conditional value at risk of the excess return of the risky investment:

$$c_t \equiv \mathrm{CVaR}_\alpha[-(R_t - f_t)] = f_t + \mathrm{CVaR}_\alpha[-R_t] \tag{3.13}$$

To compute c_t, we approximate R_t by J scenarios, $r_{t,1}, \ldots, r_{t,J}$, where $r_{t,j}$ occurs with probability p_j and $\sum_{j=1}^{J} p_j = 1$; thus,

$$c_t = f_t + \beta_t + \frac{\sum_{i=1}^{J} p_j[-r_{t,j} - \beta_t]_+}{1-\alpha} \tag{3.14}$$

where β_t is the α-quantile of this discrete approximation of $-R_t$.

Similar to TSMOM, the position size of period t is set to be inversely proportional to the CVaR estimated for period t, $\frac{C^*}{c_t}$, to equalize the risk contribution through time. In the numerical study we experiment with setting the scale, C^*, equal to various quartiles of c_t in the training dataset. The position size for period t is

$$q_t = \min\left\{1, \left|\frac{C^*}{c_t}\right|\right\}. \tag{3.15}$$

Consequently, the prescribed allocation in the risky asset at the beginning of period t is

$$w_t^* = b_t q_t. \tag{3.16}$$

3.4 Scenarios of the Risky Asset Rate of Return

While the Mean-Variance model and the TSMOM strategy rely on point estimates of the mean and volatility, the Mean-CVaR and TSMDR approaches require a representation of the distribution of the risky asset's rate of return. A major challenge to implementing these investment strategies is to model the uncertainty of the rate of return in the period ahead. Because a stochastic process model for an uncertain continuous parameter can be most accurately formulated with continuous marginal distributions, we first apply a nonstationary Brownian motion process to represent the change in the return rate over time. This model implies a normal distribution for R_t. Although the CVaR could be computed in closed form using the estimated nonstationary mean and volatility, we approximate the continuous distribution at each time t by employing Monte Carlo simulation to generate a large number of rate of return scenarios. To produce trustworthy solutions, the quality of the scenarios produced by identifying, specifying, and approximating this rate of return distribution must be assessed carefully. We employ a rank histogram technique in a backtest to check how well the scenarios match past realizations and choose the value of a key parameter in the scenario generation process. Then, to ensure that the set of scenarios is large enough, we conduct stability analysis on optimal solutions obtained from scenario sets of different sizes.

3.4.1 Model of the Risky Asset Rate of Return

Our model of the rate of return of the risky asset is motivated by the work of Koijen et al. (2009), which accommodates both return continuation (i.e., momentum) over short horizons and return reversals (mean reversion) over longer horizons. Because our model is intended for use to generate scenarios over short time horizons with frequent parameter updates, the effect of return reversals is incorporated in the periodically renewed observations. Moreover, the Koijen et al. (2009) model has a fixed market volatility for all periods. Rather than attempting to estimate a constant volatility that would be valid over all periods, we use a time-dependent volatility, estimated according to the standard deviation of return rate on a rolling basis, to capture the most recent change in volatility over time as well as match the procedure of estimating the expected rate of return.

By neglecting the long-run return reversals and utilizing a nonstationary volatility, we modify the Koijen et al. (2009) model of the rate of return for the risky asset in discrete time as

$$R_t = M_t + \sigma_t(Z_{t+1} - Z_t), \tag{3.17}$$

where the drift parameter, M_t, is the expected rate of return over a short time interval, σ_t is the volatility of the rate of return, and Z_t is a standard Brownian motion process (with mean 0 and variance t). According to the independent increments property of Brownian motion (Hull, 2009), $Z_{t+1} - Z_t$ is normally distributed with mean 0 and standard deviation 1. Equivalently,

$$R_t \sim N(M_t, \sigma_t) \tag{3.18}$$

The continuation of equity returns over short horizons indicates that the recent rate of return has predictive power for the future rate of return (Koijen et al., 2009; Moskowitz et al., 2012). Based on this assumption, Koijen et al. (2009) use an exponential moving average of the historical return rates to estimate M_t. However, in our computational study of the S&P 500 index as the risky asset, we found that using a weighted moving average, where the weights on past observations decrease in arithmetic progression, can better predict the sign of future returns (see Section 3.6.3). Therefore, we use the rates of return over the past T periods to estimate the expected rate of return at time t as

$$M_t = \left[\frac{(1+T)T}{2}\right]^{-1} \sum_{i=1}^{T} (T - i + 1) r_{t-i} \tag{3.19}$$

The decaying weights, $T - i + 1$, place emphasis on the most recent rates of return, while the summation smooths the effect of fluctuations of the rate of return. The weights are normalized to sum to one so that M_t is an unbiased estimate.

Volatility is a statistical measure of investors' uncertainty about market returns. In our model, the correlation between the past and future return rates is captured by this volatility estimate. From (3.18), it follows that

$$(R_t - M_t) \sim N(0, \sigma_t) \tag{3.20}$$

Let $u_t \equiv r_t - M_t$, G be the number of historical periods used to calculate σ_t, and $\bar{u_t}$ be the mean of the sequence $\{u_{t-G}, ..., u_{t-1}\}$. We can then estimate σ_t by

$$\sigma_t \approx \sqrt{\frac{1}{G-1} \sum_{k=t-G}^{t-1} (u_k - \bar{u_t})^2}. \tag{3.21}$$

In summary, Table 3.1 compares and contrasts the four rebalancing strategies. In the Mean-CVaR and TSMDR approaches, the CVaR of the return is estimated using Equation 3.4 and scenarios generated according to the procedure described next.

Mean-Variance $\mathrm{w}_t^* = \begin{cases} -1, & \text{if } M_t - f_t < -\frac{2\lambda\sigma_t^2}{1-\lambda} \\ 1, & \text{if } M_t - f_t > \frac{2\lambda\sigma_t^2}{1-\lambda} \\ \frac{(1-\lambda)(M_t - f_t)}{2\lambda\sigma_t^2}, & \text{otherwise} \end{cases}$	TSMOM $\mathrm{w}_t^* = a_t \min\left\{1, \left\lvert\frac{C}{\widetilde{\sigma}_t}\right\rvert\right\}$ $a_t = \text{sign}\left(\Pi_{i=1}^{k}(r_{t-i}+1) - \Pi_{i=1}^{k}(f_{t-i}+1)\right)$ $\overline{r_t} = \sum_{i=0}^{\infty}(1-\delta)\delta^i(r_{t-1-i} - f_{t-1-i})$ $\widetilde{\sigma}_t = \sqrt{N\sum_{i=0}^{\infty}(1-\delta)\delta^i(r_{t-1-i} - f_{t-1-i} - \overline{r_t})^2}$
Mean-CVaR $\mathrm{w}_t^* = \begin{cases} -1, & \text{if } M_t - f_t < -\lambda(\mathrm{CVaR}_\alpha[R_t] - M_t) \\ 1, & \text{if } M_t - f_t > \lambda(\mathrm{CVaR}_\alpha[-R_t] + M_t) \\ 0, & \text{otherwise} \end{cases}$	TSMDR $\mathrm{w}_t^* = b_t \min\left\{1, \left\lvert\frac{C^*}{c_t}\right\rvert\right\}$ $b_t = \text{sign}(\mathrm{M}_t - f_t)$ $\mathrm{c}_t = f_t + \mathrm{CVaR}_\alpha[-R_t]$

Table 3.1: Summary of rebalancing strategies

3.4.2 Scenario Generation Procedure

To generate rate of return scenarios for the risky asset according to the model in Section 3.4.1, we apply Monte Carlo simulation, as described by Hull (2009) and Guo and Ryan (2020). To update the drift and volatility estimates and recalibrate the model with recent data, we repeatedly generate scenarios for the next period on a rolling basis. For instance, at the end of period $t-1$, the realized rate of return, r_{t-1}, is released. We use it, along with the past T periods' rates of return, to compute the expected rate of return in period t, M_t, by (4.13). The volatility estimate for period t, σ_t, is based on the standard deviation of the difference in the rate of return and its mean value over the past G periods based on (4.14). Using M_t and σ_t, the rate of return in period t, $r_{t,j}$, can be randomly generated for each scenario j according to (3.18). Note that since each estimate of M_t requires T realized rates of return, an estimate of M_t is not available until period $T+1$. Similarly, we obtain our first estimate of volatility, σ_1, after computing $M_{-G+1}, M_{-G+2}, \ldots, M_0$. By repeating this procedure, J scenarios for each period $t = 1, \ldots, H$ can be generated.

The procedure for generating rate of return scenarios with rolling parameter updates is detailed in Algorithm 4. The sequence $\{r_t, t = -(T+G)+1, \ldots, H-1\}$ represents observed values of the return rate, whose values up to time $t-1$ are known at time t. The collection $\{r_{t,j}, j = 1, \ldots, J\}$ represents the set of scenarios generated for R_t.

Algorithm 4 Procedure to generate scenarios of the risky asset rate of return

Input: J, T, G, H, r_t for $t = -(T+G)+1, \ldots, H-1$

Output: $r_{t,j}$ for $t = 1, \ldots, H$ and $j = 1, \ldots, J$

begin:

for $t \leftarrow -G+1$ to 0 **do** ▷ initialize the estimation of drift

$M_t \leftarrow T\left[\frac{(1+T)T}{2}\right]^{-1} r_{t-1} + \cdots + 1\left[\frac{(1+T)T}{2}\right]^{-1} r_{t-T}$

end for

for $t \leftarrow 1, 2, \ldots, H$ **do**

$M_t \leftarrow T\left[\frac{(1+T)T}{2}\right]^{-1} r_{t-1} + \cdots + 1\left[\frac{(1+T)T}{2}\right]^{-1} r_{t-T}$ ▷ estimate drift in period t

$\sigma_t \leftarrow$ `standard deviation of` $\{r_{t-G} - M_{t-G}, \ldots, r_{t-1} - M_{t-1}\}$ ▷ estimate volatility in period t

for $j \leftarrow 1$ to J **do** ▷ generate scenarios for period t

$r_{t,j} \leftarrow \text{rnorm}(M_t, \sigma_t)$ ▷ generate a random observation of the standard normal distribution

end for

return $r_{t,j}$, $j = 1, \ldots, J$.

end for

Algorithm 4 involves three parameters that can affect the quality of scenarios generated and the resulting solution (or position scaling in the TSMDR strategy). In our numerical study the number, T, of past return rates used to estimate the drift, is chosen to match the corresponding value in the TSMOM strategy. The value of G, the number of past return rates used to estimate volatility, is selected according to a reliability study. The number of scenarios, J, is determined by stability analysis of optimal solutions.

3.4.3 Reliability of Scenarios

To evaluate the reliability of a scenario generation method, we seek to quantify the extent to which scenarios and observations are interchangeable. A scenario set is called reliable or calibrated if the relative frequency of occurrence of a scenario assigned a probability p tends to be close to p (Hsu and Murphy, 1986). In other words, the more similar the distribution of generated scenarios and the distribution of real observations are, the more reliable the generation method is judged to be. Sarı et al. (2016) develop and test a novel mass transportation distance (MTD) rank histogram, implemented as an R package by Sarı and Ryan (2016), to assess whether the scenarios have similar patterns as the corresponding observations.

Sarı et al. (2016) point out that if the MTD rank histogram exhibits uniformity, it indicates that the observation is indistinguishable from the scenarios, and thus, the generated scenarios are reliable. We use a goodness-of-fit test to assess the nearness to uniformity of the MTD histogram. Elmore (2005) explains that the Cramér–von Mises (CvM) family of statistics is better at detecting the ordered departures from the null distribution than the χ^2 test. In addition, it retains considerable power for relatively small samples. Therefore, we choose the W^2 statistic (Choulakian et al., 1994; Arnold and Emerson, 2011), one of the CvM statistics, to measure the distance between the uniform distribution and the empirical distribution of the MTD ranks. If the value of W^2 is smaller than the critical value, the generation procedure is judged to be reliable. According to this criterion, we can select the best value for G from among all candidates when T and J are fixed.

3.4.4 Stability of the Optimal Solutions

King and Wallace (2012) argue that the quality of generated scenarios is ultimately determined by solutions to the optimization problem. We determine the sample size according to the stability of the optimal solutions. In many studies, the stability of the optimal objective value is examined rather than the optimal solution, because there may be a large number of potentially optimal solutions and the value of the objective function presumably reflects the concerns of the decision maker. However, we are mainly interested in the optimal solution obtained from a set of scenarios, to facilitate comparison of the stochastic optimization models with the heuristic approaches. Furthermore, the simple structure of solutions to our two optimization models makes it easy to check their stability. Recall that the Mean-Variance model in Section 3.3.1.1 does not involve random scenarios, because

$\mathrm{E}[R_t] = M_t$ and $\mathrm{Var}[R_t] = \sigma_t^2$ are both deterministic parameters at the time we solve the model for period t. The existence of only three possible values for the optimal solution to the Mean-CVaR model in Section 3.3.1.2 allows us to easily compare the optimal solutions found using different scenario sets.

When checking the stability of the optimal solution in the Mean-CVaR model, there is a trade-off between the computational efficiency and the scenario quality. A large sample size, J, helps avoid the fragility of CVaR (Lim et al., 2011) as well as guarantee that the optimality gap between the objective value of the true optimization model (with continuous distribution of the rate of return of the risky asset) and the approximate optimization model (with discrete rate of return scenarios) is small enough. But a small value of J will save computation time. Therefore, we initialize with a sufficiently large J value to balance this trade-off. For a candidate value of J, we conduct 30 replications of the generation procedure using different random seeds. By solving the problem at each time t with each scenario set generated at that time, we obtain 30 sequences of optimal solutions. Next, we check the similarity of these 30 sequences pairwise. Similarity is measured according to the number of times the same element appears in the same position of the sequence divided by the length of the sequence. In this way, we compute 29×15 values of the similarity, the average of which measures the stability of the optimal solutions across the 30 replications. We increase J until the value of stability passes a predetermined threshold.

3.5 Performance Metrics

We evaluate the out-of-sample performance of the different investment strategies according to multiple metrics in historical backtesting. For each strategy, we simulate the process of assigning weights w_t^* in periods $1, \ldots, H$, where w_t^* is computed or optimized according to the risky asset return rates observed before time t.

The Sharpe ratio of excess returns to the standard deviation of portfolio returns is a widely-used risk-adjusted performance measure (Sharpe, 1966, 1994). In our case, the annualized Sharpe ratio is computed as

$$SR = \sqrt{N} \frac{\frac{1}{H} \sum_{t=1}^{H} w_t^* (r_t - f_t)}{\sqrt{\frac{1}{H-1} \sum_{t=1}^{H} \left[w_t^* (r_t - f_t) - \bar{B} \right]^2}} \tag{3.22}$$

where $\bar{B} = \frac{1}{H} \sum_{t=1}^{H} \left[w_t^* (r_t - f_t) \right]$ and N is the number of periods in a year.

However, the Sharpe ratio is often criticized on the grounds that it penalizes both the downside risk and the upside return potential equally. The Sortino ratio (Sortino and Price, 1994; Eling, 2008), whose risk metric focuses on returns that fall below a minimum threshold or minimum acceptable return (r_{MAR}), corrects this deficiency. The formula for the annualized Sortino ratio is

$$ST = \sqrt{N}\frac{\frac{1}{H}\sum_{t=1}^{H}[w_t^* r_t + (1-w_t^*)f_t] - r_{MAR}}{\sqrt{\frac{1}{H}\sum_{t=1}^{H}\min\{0, w_t^* r_t + (1-w_t^*)f_t - r_{MAR}\}^2}} \tag{3.23}$$

To be aligned with the numerator of the Sharpe ratio, we set $r_{MAR} \equiv \frac{1}{H}\sum_{t=1}^{H} f_t$.

Maximum drawdown (Young, 1991) is an indicator of downside risk over a specified time period, representing the maximum loss an investor can suffer in the portfolio by buying at the highest point and selling at the lowest. A drawdown is defined as any time when the cumulative return dips below the maximum cumulative return. Drawdowns are measured as a percentage of that maximum cumulative return. Using the simulated sequence of periodic return rates generated by implementing any strategy, we calculate this quantity using the `maxDrawdown` function in the PerformanceAnalytics R package (Peterson et al., 2014).

Motivated by Tse (2015) and He et al. (2018), we also investigate the relationship between the state of market and the portfolio returns by plotting the time series cumulative absolute returns in log scale. The cumulative absolute log return up to period t is

$$CA_t = \sum_{i=1}^{t} log[1 + w_i^* r_i + (1-w_i^*)f_i] \tag{3.24}$$

3.6 Numerical Results

In this study, we consider the S&P 500 index as the risky asset and the one-month US Treasury bill (T-bill) as the risk-free investment. The monthly rates of return of both assets are obtained from the CRSP database (`www.crsp.com`).

The time interval of the simulation covers January 1, 2001, to December 31, 2019, containing a total of $H = 228$ trading months. In the experiments, J is initialized to be 10,000 for each instance. By constructing portfolios based on $k = 12$ in TSMOM$(k, 1)$, referred to simply as TSMOM, Moskowitz et al. (2012) find that this strategy performs the best among all the momentum strategies with k varying from 1 to 48 months. Thus, we fix $k = 12$ and focus only on the discussion of TSMOM. Accordingly, T in Algorithm 1 is set to 12. To illustrate results for multiple risk

attitudes, we consider $\alpha \in \{75\%, 90\%, 99\%\}$, $\lambda \in \{0, 0.1, 1/3, 0.5, 0.7, 0.9\}$ for Mean-Variance and $\lambda \in \{0, 0.02, 0.04, 0.06, 0.08, 0.1\}$ for Mean-CVaR. Specifically, risk-neutral investors are represented by $\lambda = 0$. We choose relatively small values of λ for Mean-CVaR because most risky periods, such as the extreme bear markets, historically have persisted over several months, even years, which has been taken into account in the construction of M_t. For Mean-Variance, along with evenly-spaced values, we test $\lambda = 1/3$ to compare with the solution found by He et al. (2018) for maximizing log utility in the long run. Besides $\lambda = 0$ in the mean-risk optimization model, we consider another risk-neutral investment strategy, replacing (3.10) by $p_t \equiv 1$. We call this strategy Unscaled TSMOM since it ignores volatility scaling and simply uses the trading signal to determine the position without considering risk. All the statistical simulations and analyses are conducted in R.

3.6.1 Parameter Estimation and Stability Analysis

To avoid data leakage while identifying suitable values for the risk-related parameters G, C and C^*, we use data from January, 1982, to December, 2000, as the training dataset of the same length as the simulation horizon.

The value of G is selected from among $\{2, 4, 6, 8, 10, 12\}$ according to scenario reliability based on $J = 10000$. We use a common set of 30 random seeds across different values of G to conduct 30 replications of the generation procedure for each candidate value of G. Values of G that can offer reliable scenarios can be found on the basis of their MTD rank histograms according to the observed values of the W^2 statistic. The results for each candidate value of G are shown in Table 3.2, along with critical values provided by Choulakian et al. (1994). Because a larger significance level increases the power of the test and we prefer to guard against falsely failing to reject the null hypothesis of uniformity, we choose $G = 10$ or 12 based on the 0.25 significance level. Furthermore, $G = 10$ yields a smaller mean, standard deviation and maximum value of W^2. Therefore, we fix $G = 10$ in the following analysis.

To further validate our model of index returns with $T = 12$ and $G = 10$, the Q-Q plots for the standardized actual returns, $\frac{r_t - M_t}{\sigma_t}$, in the training and test periods against the standard normal distribution are displayed in Figure 3.2. In both plots, the points fall approximately along the reference line. Based on the Shapiro-Wilk normality test (**?**), the p-values for the training and test data are 0.3077 and 0.0455, respectively. The Q-Q plots, along with the Shapiro-Wilk test, make

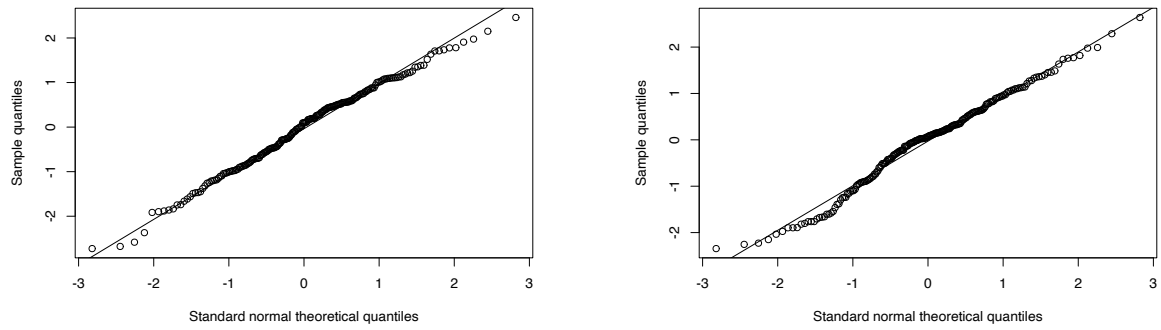
it evident that our model of index returns with nonstationary mean estimated by past 12-month returns and time-varying volatility estimated by past 10-month returns can well represent the actual index returns. In addition, the empirical density of the actual return of the S&P index during the simulation horizon is shown in Figure 3.3. For comparison, the empirical density of all generated scenarios for all time periods using $G = 10$ is shown in Figure 3.4. The statistics of the combined scenarios are close to the statistics of the actual returns and both densities are asymmetric and heavy-tailed. Performance comparisons for $G = 6, 8$ and 12 are similar, as shown in the Electronic Companion.

G	2	4	6	8	10	12
Mean of W^2	3.92	0.39	0.13	0.09	0.07	0.09
Standard deviation of W^2	0.29	0.06	0.05	0.04	0.03	0.04
Maximum W^2	4.40	0.51	0.27	0.25	0.13	0.20
% rejecting H0 at significance level 0.01 (0.743)	100.00	0.00	0.00	0.00	0.00	0.00
% rejecting H0 at significance level 0.025 (0.581)	100.00	0.00	0.00	0.00	0.00	0.00
% rejecting H0 at significance level 0.05 (0.461)	100.00	13.30	0.00	0.00	0.00	0.00
% rejecting H0 at significance level 0.10 (0.347)	100.00	70.00	0.00	0.00	0.00	0.00
% rejecting H0 at significance level 0.15 (0.284)	100.00	96.70	0.00	0.00	0.00	0.00
% rejecting H0 at significance level 0.25 (0.209)	100.00	100.00	3.30	3.30	0.00	0.00

Table 3.2: Uniformity test for G using the training dataset. The critical value for each significance level is shown in parentheses.

The median value ($Q2$) of $\widetilde{\sigma}_t$ is 0.109. To check its sensitivity, we also consider the first and third quartiles and the 90$^{\text{th}}$ percentile ($Q1$, $Q3$ and 90%) of $\widetilde{\sigma}_t$, hence, $C \in \{0.087, 0.109, 0.135, 0.166\}$ in TSMOM. In the same way, we obtain $Q1, Q2, Q3$ and 90$^{\text{th}}$ percentile of C^*, which constitutes the candidate set for C^* in TSMDR: when $\alpha = 75\%$, $C^* \in \{0.032, 0.045, 0.061, 0.075\}$; when $\alpha = 90\%$, $C^* \in \{0.049, 0.066, 0.085, 0.107\}$; when $\alpha = 99\%$, $C^* \in \{0.080, 0.107, 0.130, 0.166\}$.

For the stability analysis, we find that stable optimal solutions are obtained when the number of scenarios, J, is increased to 20,000. The stability of optimal solutions in the Mean-CVaR model for different values of α and λ is displayed in Table 3.3 based on a 98% threshold. The solutions for



(a) Q-Q plot for the training data

(b) Q-Q plot for the test data

Figure 3.2: Q-Q plots. These plots show the quantiles of empirical distribution of the standardized actual returns in the training and test periods against the standard normal distribution.

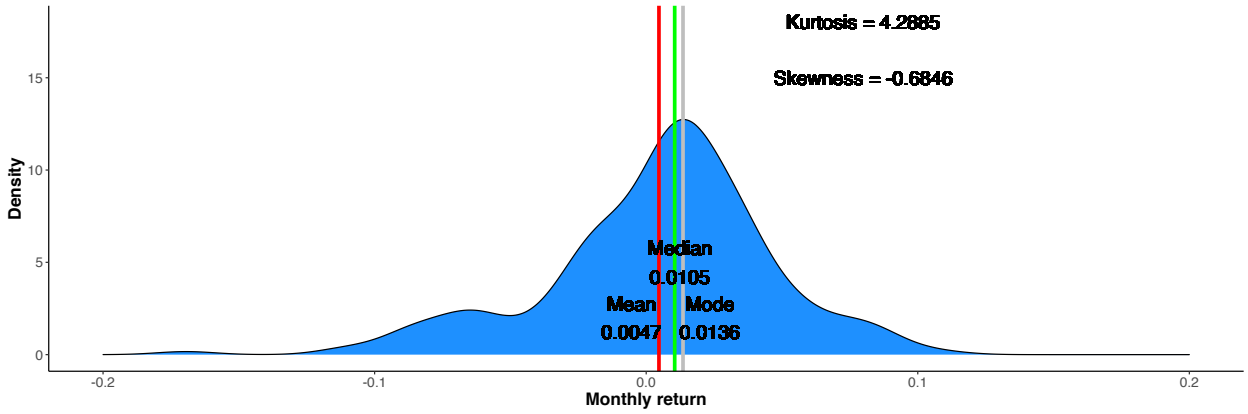


Figure 3.3: Empirical density plot of actual monthly return during January, 2001 to December, 2019

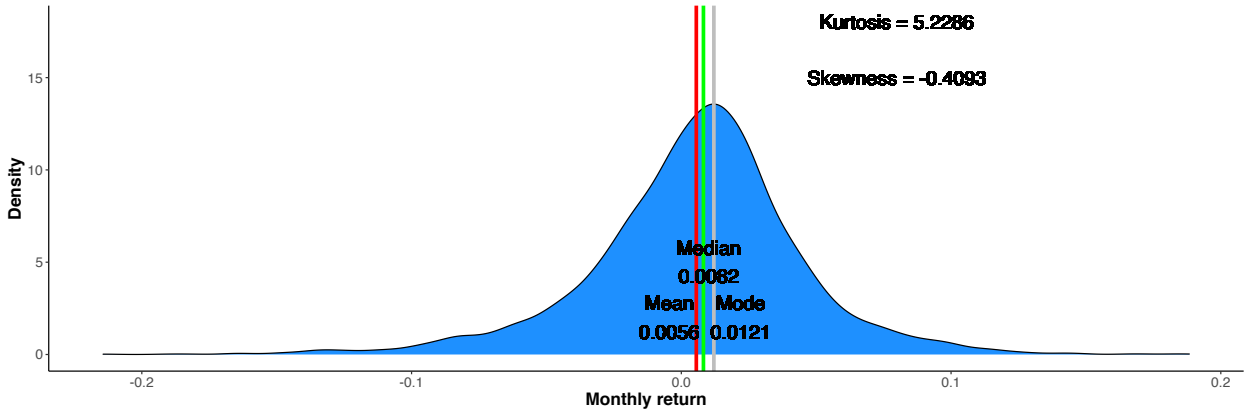


Figure 3.4: Empirical density plot of the generated scenarios of monthly return using $G = 10$ during January, 2001 to December, 2019

all combinations of α and λ are at least 98.37% similar, implying that at most four of the $H = 228$ instances have different optimal solutions.

λ \ α	75%	90%	99%
0	99.33%	99.35%	99.37%
0.02	98.85%	98.95%	98.37%
0.04	98.76%	99.19%	98.75%
0.06	98.68%	98.60%	99.45%
0.08	99.70%	98.58%	98.93%
0.10	98.43%	98.53%	98.60%

Table 3.3: Results of stability test for 20000 scenarios

3.6.2 Evaluation and Comparison of Investment Strategies

Figure 3.5 illustrates the combinations of the annualized volatility and excess return achieved by different strategies and values of risk parameters. The points for TSMDR occupy the top left corner, indicating that it provides a combination of low volatility and high excess return, which is typically desired by investors. The volatility and the excess return of TSMOM increases as the position scale, C, increases; however, the rate of increase of the volatility is greater than that of

the excess return. When comparing between the TSMOM and TSMDR strategies with the scale set at the same quartile of the respective risk distribution, in all pairwise comparisons TSMDR has a better return and less volatility. The excess return and the volatility of portfolio returns obtained using the Mean-CVaR model vary considerably. A relatively large value for the risk-aversion parameter, λ, often provides a small excess return and a moderate volatility. Additionally, a larger value of the probability α (smaller tail probability) will result in lower volatility in general. For Mean-Variance, when λ is 0.1, it produces the largest return and the largest volatility among all risk-aversion strategies. Thus, investors with high risk tolerance may prefer the Mean-Variance model with a small risk-aversion parameter. On the other hand, when λ is 0.9, the Mean-Variance model has the smallest return and the smallest volatility among all risk-aversion strategies. In addition, compared to the Mean-CVaR model, the Mean-Variance model is less sensitive to changes in λ. For the same value of λ tested, the Mean-CVaR model results in much lower volatility than the Mean-Variance model, which suggests that CVaR can better control the investment risk than variance when the distribution of returns is not symmetric. The volatilities of returns resulting from the two Risk-Neutral strategies are similar, but the optimization model with $\lambda = 0$ produces a higher return than the Unscaled TSMOM strategy. This result implies that using the weighted moving average of the past 12-month index rates of return in excess of 1-month T-bill rate as the trading signal performs better than using past 12-month excess return.

Table 3.4 summarizes the annualized performance metrics of the various strategies sorted by descending Sortino ratio. The three largest values of the Sharpe ratio and Sortino ratio are produced by our newly created TSMDR when C^* is set at its first quartile. The TSMDR strategy produces a larger Sharpe ratio than the other strategies for all values of α, λ, and risk-scaling parameters tested, including the optimal strategy derived by the log utility function in He et al. (2018). For (α, λ) combinations (0.99, 0.08), Mean-CVaR produces the worst results in terms of the Sortino ratio. However, when λ of Mean-CVaR and Mean-Variance takes a relatively large value, they achieve the smallest maximum drawdown values. TSMDR also has low values of maximum drawdown when C^* is at $Q1$. As for the scale parameter in TSMDR and TSMOM, setting Q1 always results in the largest Sortino ratio and Sharpe ratio, all else equal.

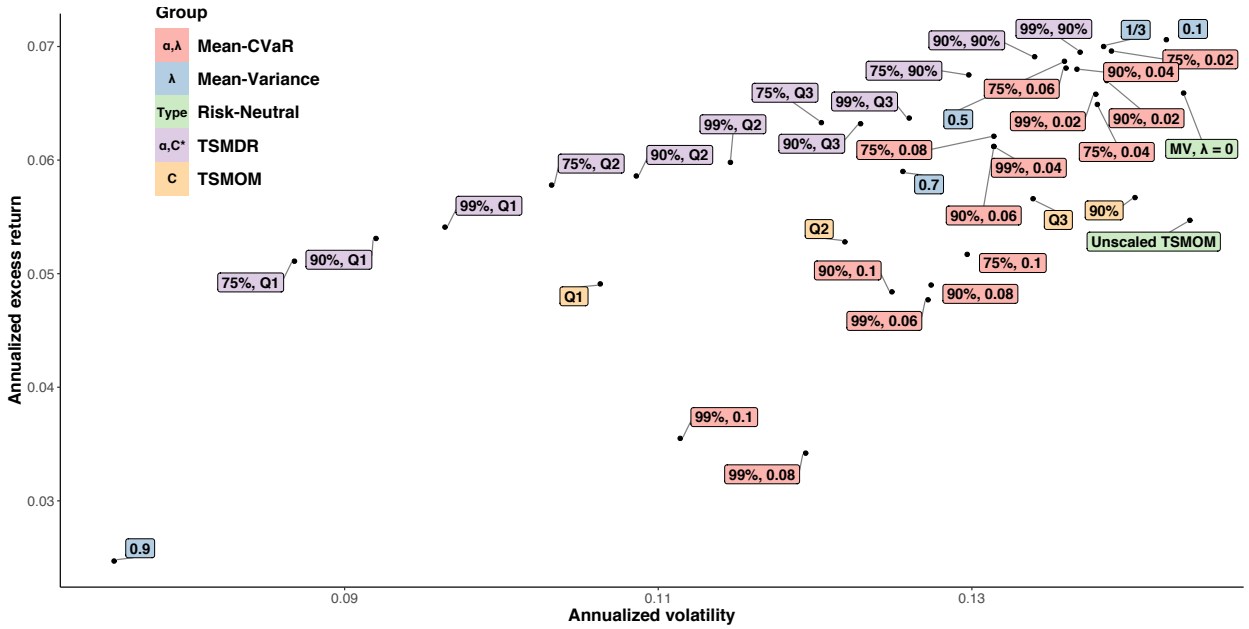


Figure 3.5: Scatter plot of annualized excess returns and volatility for different strategies for $T = 12, G = 10$ and $J = 20000$

Table 3.4: Annualized performance metrics of different strategies sorted by the descending Sortino ratio

Group	Parameter value	Sortino ratio (%)	Sharpe ratio (%)	Maximum drawdown (%)
TSMDR	$\alpha = 75\%$, C* at Q1	88.23	58.83	22.24
TSMDR	$\alpha = 90\%$, C* at Q1	86.88	57.70	23.23
TSMDR	$\alpha = 99\%$, C* at Q1	84.63	56.14	23.85
TSMDR	$\alpha = 75\%$, C* at Q2	84.10	56.04	25.37
TSMDR	$\alpha = 90\%$, C* at Q2	80.92	53.97	26.93
TSMDR	$\alpha = 90\%$, C* at 90%	78.93	51.58	32.32
TSMDR	$\alpha = 75\%$, C* at 90%	78.83	51.99	31.47
TSMDR	$\alpha = 75\%$, C* at Q3	78.79	52.57	29.9
TSMDR	$\alpha = 99\%$, C* at Q2	78.38	52.18	28.95
Mean-Variance	$\lambda = 0.5$	77.98	50.53	36.66
TSMDR	$\alpha = 99\%$, C* at 90%	77.88	50.79	33.84
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.06$	77.87	50.07	30.30

Mean-Variance	$\lambda = 1/3$	77.70	50.53	39.40
TSMDR	$\alpha = 90\%$, C* at Q3	77.19	51.40	30.88
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.02$	77.16	50.12	38.53
Mean-CVaR	$\alpha = 90\%$, λ= 0.04	77.16	49.77	33.61
TSMDR	$\alpha = 99\%$, C* at Q3	76.31	50.54	32.21
Mean-Variance	$\lambda = 0.1$	75.78	49.57	40.25
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.08$	74.43	47.26	22.22
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.02$	74.24	48.35	38.53
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.06$	73.35	46.58	22.22
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.04$	73.35	46.58	22.22
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.02$	73.27	47.74	38.53
Mean-Variance	$\lambda = 0.7$	73.04	46.98	30.85
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.04$	72.16	47.03	38.53
Mean-Variance	$\lambda = 0$	70.60	45.93	40.85
TSMOM	C at Q1	70.04	46.14	27.54
TSMOM	C at Q2	65.90	43.35	30.96
TSMOM	C at Q3	65.10	42.25	31.62
TSMOM	C at 90%	62.13	40.41	32.92
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.1$	62.09	39.87	29.99
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.1$	60.28	38.72	30.52
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.08$	59.65	38.5	25.99
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.06$	58.15	37.54	25.99
TSMOM	Unscaled TSMOM	57.94	38.05	37.99
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.1$	50.55	31.88	22.12
Mean-Variance	$\lambda = 0.9$	48.15	32.84	20.89
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.08$	44.19	28.63	32.47

Figure 3.9 plots the cumulative absolute log returns of the TSMDR, Mean-CVaR, Mean-Variance, TSMOM, and Risk-neutral to compare the cumulative performance of different approaches through

time. Although α is relevant to TSMDR and Mean-CVaR only, the plots of the other three strategies are repeated in the three panels to facilitate comparison. When comparing among different strategies for each value of α, Mean-Variance with $\lambda = 0.1$ and TSMDR with C^* set at the 90% are the best during most of the simulation horizon. All strategies, especially the Mean-Variance models, would have performed well during the 2007-2008 financial crisis. When comparing the same strategy with different parameters, we observe that a large scale value in TSMOM or TSMDR and a small λ value in Mean-CVaR or Mean-Variance almost always provide a large cumulative return. Between the risk-neutral approaches, simply maximizing expected returns is better than unscaled TSMOM during most of the periods. Comparing across the three panels, we see that Mean-CVaR is much more sensitive to the change in α values than TSMDR is.

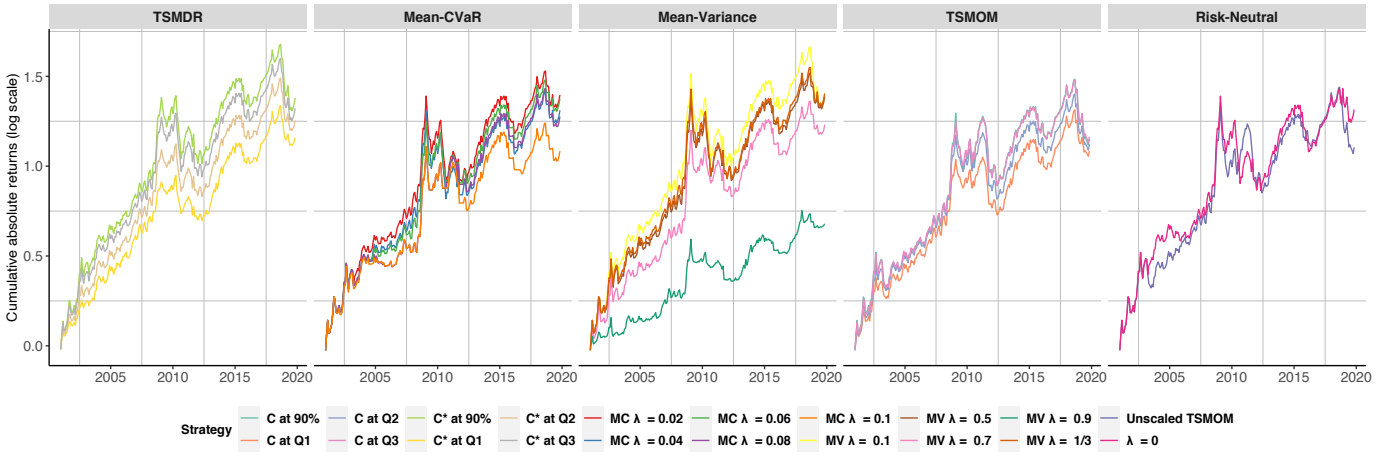


Figure 3.6: Cumulative absolute log returns when $\alpha = 0.75$

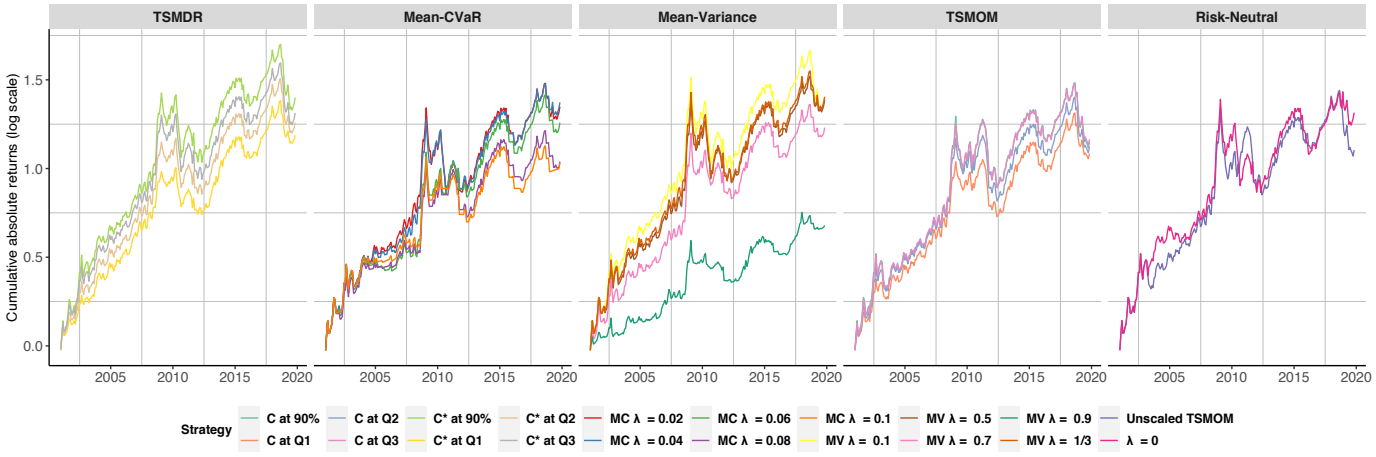


Figure 3.7: Cumulative absolute log returns when $\alpha = 0.90$

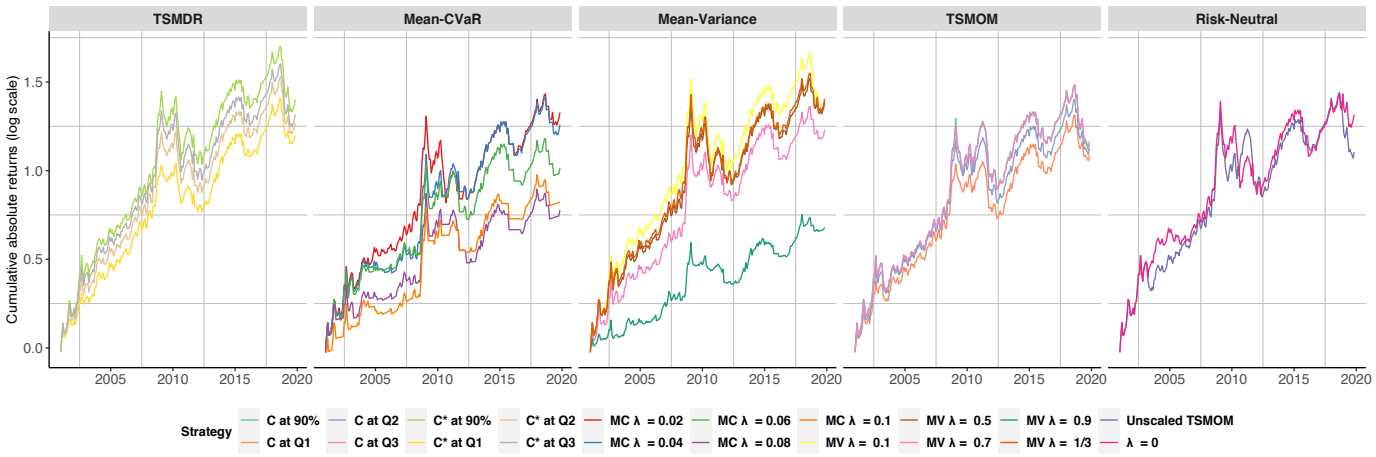


Figure 3.8: Cumulative absolute log returns when $\alpha = 0.99$

Figure 3.9: Cumulative absolute log returns of TSMDR, Mean-CVaR, Mean-Variance, Risk-neutral, and TSMOM. Only TSMDR and Mean-CVaR are different in three panels.

3.6.3 Insights from Investment Strategy Decomposition

In general, we can conclude that TSMDR is a promising strategy as it has the largest Sortino ratio, Sharpe ratio and cumulative returns. To investigate why TSMDR consistently outperforms, we decompose all strategies into three aspects, the trading signal, the asset allocation model used to determine the position size, and the risk measure, as described in Table 3.5.

Investment strategy	Trading signal	Model type	Risk
Mean-Variance	Weighted moving average	Mean-risk	Variance
Mean-CVaR	Weighted moving average	Mean-risk	CVaR
TSMOM	Time series momentum	Risk parity	Volatility
TSMDR	Weighted moving average	Risk parity	CVaR

Table 3.5: Decomposition of investment strategies

To analyze the trading signal, we apply several typical moving average techniques, including simple, weighted, and exponential moving average, to the past excess returns to compute the trading signal and compare them with the TSMOM trading signal which is the sign of the past 12-month excess return. If the trading signal is positive, the investment position in the S&P 500 index is 1. Otherwise, it is -1. This comparison excludes the effect of the investors' attitudes towards risk and uses the trading signal as the only element to determine the quality of the strategy.

Table 3.6 shows the performance of different trading signals in terms of the percentage of positive and negative excess returns during the 228 months from January, 2001, to December, 2019. These values indicate that the weighted moving average produces the most accurate predictions of the trend of the S&P index, and exponential moving average is the worst.

Trading signal	Percentage of positive excess returns (Prediction accuracy)
Weighted moving average	60.4%
Simple moving average	59.9%
Time series momentum	57.7%
Exponential moving average	55.5%

Table 3.6: Percentage of positive excess returns during January, 2001 to December, 2019 when using different trading signals

Figure 3.10 plots the sequence of trading signals generated by different moving average techniques and time series momentum, along with the sign of the actual excess return. It can be observed that weighted moving average, time series momentum, and simple moving average produce similar signals. However, exponential moving average performs like a lagged realized excess return, which helps explain the results in Table 3.6 that exponential moving average has less accuracy in predicting the sign of the excess return of the next month. In addition, the top three trading signals appear to be less volatile and produce less turnover, whereas exponential moving average is very sensitive to any change in the stock market, even small ones, which results in frequent transactions. When using exponential moving average, we put 37% of the total weight on the most recent index rate of return and only 0.0006% weight on the rate of return observed 12 months previously (the corresponding weights for the weighted moving average are 15% and 1%, respectively).

Figure 3.11 plots the cumulative absolute log return of portfolios built using different trading signals, and includes the cumulative returns of the realized S&P 500 index, which can be treated as the cumulative return of a buy-and-hold (or passive long) strategy, as a comparison. From Figure 3.11, we find that all trading signals produce their best performance under extreme conditions. We conjecture that this is because most extreme bear or bull markets historically have happened over several months or years, with the result that using past returns as the trading signal takes short (long) positions as markets begin to decline (improve) in prolonged bear (bull) markets. The returns of simple moving average and time series momentum are overlapped during most periods.

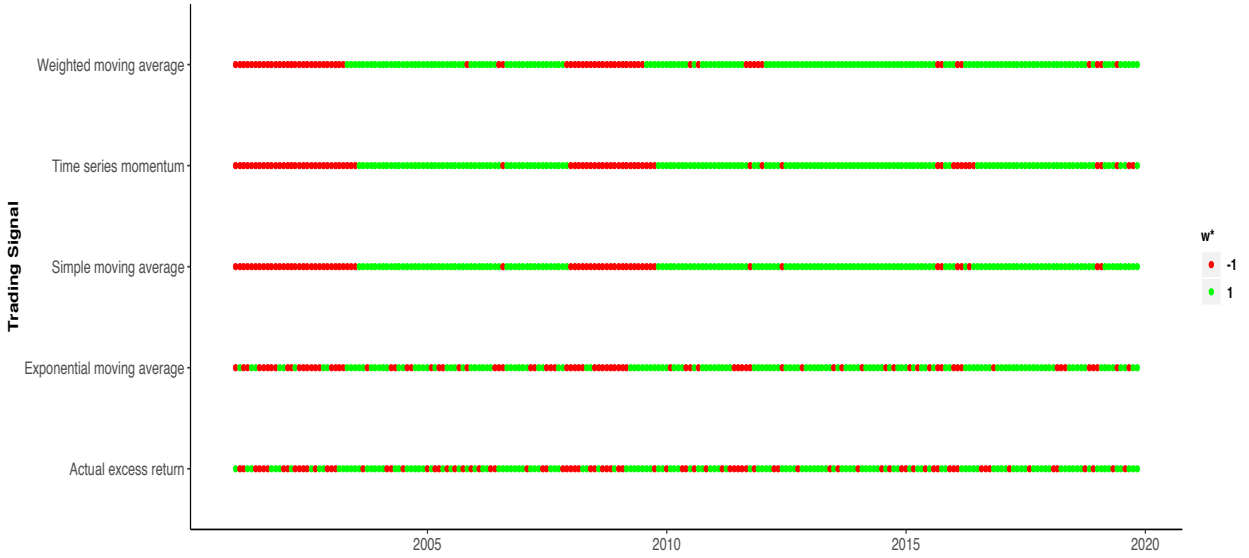


Figure 3.10: Sequence of trading signals, along with the sign of the actual excess return

The return from following weighted moving average remains above the returns from all other trading signals during most of the time horizon.

In short, using weighted moving average as the trading signal can provide the best accuracy of predicting the sign of excess returns among all tested trading signals, shown in Table 3.6, as well as a higher cumulative absolute return than time series momentum, exhibited in Figure 3.11. Recall that TSMOM uses the sign of time series momentum as the trading signal while TSMDR employs weighted moving average to obtain the trading signal. The trading signal helps explain why TSMDR outperforms TSMOM.

A good trading signal is not the only factor that determines performance. For example, both Mean-CVaR and TSMDR use weighted moving average as a signal and CVaR as the risk measure, but TSMDR achieves better Sortino and Sharpe ratios, suggesting that the asset allocation model to determine the position size also plays an important role. From Figure 3.5, we observe that the mean-risk strategies generally have a better return whereas risk parity strategies have less risk. These results were to be expected because mean-risk models balance the trade-off between maximizing the expected return and minimizing the risk but risk parity models simply determine the allocation by equalizing the risk without attempting to maximize the expected return.

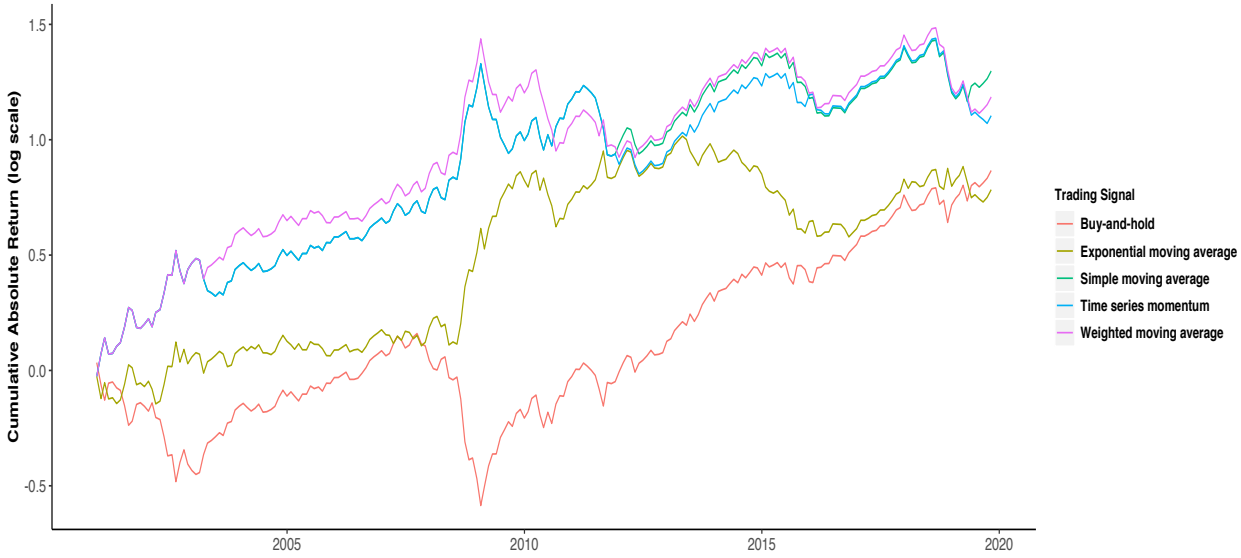


Figure 3.11: Cumulative absolute log return when using different trading signals. The blue line, representing time series momentum, and the green line, representing simple moving average, are only separate during December, 2011 and December, 2015, and August, 2019 to the end.

Finally, by comparing Mean-Variance and Mean-CVaR, which differ only in the risk measure, we find that CVaR can better control the investment risk than variance, as Mean-Variance models have much worse performance in both volatility and maximum drawdown when using the same risk-aversion parameter value (see Figure 3.5 and Table 3.4).

3.7 Conclusion

To find an optimal dynamic allocation between a risky asset and a risk-free investment, two mean-risk stochastic optimization models and two heuristics based on the return predictability of the time series momentum are compared. To estimate the future return rates of the risky asset, we develop a modified geometric Brownian motion process with nonstationary drift to accommodate the time series momentum and time-dependent volatility to capture the most recent change in volatility. The optimal solutions to the Mean-Variance and Mean-CVaR models differ in their structure, even for normally distributed returns. In the heuristic strategies, different estimates of time series momentum are used to determine the trading signals. In both optimization and heuristic strategies, two risk measures are considered: variance (volatility), which treats both upside reward and downside risk as risk, and CVaR, which quantifies downside risk only. By backtesting all

strategies using stock index returns and 1-month T-bill rates during 2001-2019, we validate the stochastic model of index returns. We observe that conditional normal distributions with time-varying mean and volatility produce a simulated unconditional return distribution with skewness and kurtosis similar to the empirical distribution.

Our newly created TSMDR strategy, a hybrid of time series momentum heuristic with a stochastic programming approach to risk estimation, significantly outperforms the others in terms of the Sharpe ratio and the Sortino ratio. TSMDR can not only provide a relatively high excess return but also maintain a very low investment risk, as preferred by many investors. When the risk-aversion parameter, λ, is small, Mean-Variance optimization can provide the largest excess returns among all strategies, which is favored by investors with high risk tolerance. Although the investment risk of Mean-CVaR optimization is large, it results in a larger Sharpe ratio and the Sortino ratio than TSMOM when the tail probability, $1 - \alpha$, of CVaR is set to a moderate value. The insights obtained from considering trading signal and position size separately are that 1) using weighted moving average of returns can better capture the trend of the stock market than time series momentum, 2) mean-risk strategies generally provide better returns whereas risk parity strategies have less investment risk, and 3) CVaR can better control the investment risk than variance.

This paper can be extended in various directions. One possible extension could be to conduct a similar comparison between the cross-sectional momentum strategy and multi-asset portfolio optimization. In addition, we considered only basic moving average techniques to compute the trading signal in this paper. Because the trading signal is essential to the performance of the strategy, it is worthwhile to test additional moving average, or even moving median, techniques. Finally, besides CVaR and variance (or volatility), there are many other risk measures, such as entropic risk measures and expected conditional risk measures for time consistency over multiple stages. Employing different types of risk measures to see how they would affect the performance of an investment strategy could be another avenue to explore.

3.8 References

Arnold, T. B. and Emerson, J. W. (2011). Nonparametric goodness-of-fit tests for discrete null distributions. *R Journal*, 3(2).

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

Asness, C. S., Frazzini, A., and Pedersen, L. H. (2012). Leverage aversion and risk parity. *Financial Analysts Journal*, 68(1):47–59.

Bertsimas, D. and Pachamanova, D. (2008). Robust multiperiod portfolio management in the presence of transaction costs. *Computers & Operations Research*, 35(1):3–17.

Choulakian, V., Lockhart, R. A., and Stephens, M. A. (1994). Cramér-von Mises statistics for discrete distributions. *Canadian Journal of Statistics*, 22(1):125–137.

Dantzig, G. B. and Infanger, G. (1993). Multi-stage stochastic linear programs for portfolio optimization. *Annals of Operations Research*, 45(1):59–76.

De Bondt, W. F. and Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance*, 40(3):793–805.

Eling, M. (2008). Does the measure matter in the mutual fund industry? *Financial Analysts Journal*, 64(3):54–66.

Elmore, K. L. (2005). Alternatives to the chi-square test for evaluating rank histograms from ensemble forecasts. *Weather and Forecasting*, 20(5):789–795.

Goyal, A. and Jegadeesh, N. (2018). Cross-sectional and time-series tests of return predictability: What is the difference? *The Review of Financial Studies*, 31(5):1784–1824.

Gülten, S. and Ruszczyński, A. (2015). Two-stage portfolio optimization with higher-order conditional measures of risk. *Annals of Operations Research*, 229(1):409–427.

Guo, X. and Ryan, S. M. (2020). Reliability assessment of scenarios generated for stock index returns incorporating momentum. *International Journal of Financial & Economics*, 26(3):4013–4031.

Harlow, W. V. (1991). Asset allocation in a downside-risk framework. *Financial Analysts Journal*, 47(5):28–40.

He, X.-Z. and Li, K. (2015). Profitability of time series momentum. *Journal of Banking & Finance*, 53:140–157.

He, X.-Z., Li, K., and Li, Y. (2018). Asset allocation with time series momentum and reversal. *Journal of Economic Dynamics and Control*, 91:441–457.

Hsu, W.-r. and Murphy, A. H. (1986). The attributes diagram a geometrical framework for assessing the quality of probability forecasts. *International Journal of Forecasting*, 2(3):285–293.

Hull, J. C. (2009). *Option, Futures and other Derivatives*. Upper Saddle River, NJ: Prentice Hall,.

Hurst, B., Ooi, Y. H., and Pedersen, L. H. (2013). Demystifying managed futures. *Journal of Investment Management*, 11(3):42–58.

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

Kaut, M. and Wallace, S. W. (2007). Evaluation of scenario-generation methods for stochastic programming. *Pacific Journal of Optimization*, 3(2):257–271.

Kim, A. Y., Tse, Y., and Wald, J. K. (2016). Time series momentum and volatility scaling. *Journal of Financial Markets*, 30:103–124.

King, A. J. and Wallace, S. W. (2012). *Modeling with stochastic programming*. Springer Science & Business Media.

Koijen, R. S. J., Rodríguez, J. C., and Sbuelz, A. (2009). Momentum and mean reversion in strategic asset allocation. *Management Science*, 55(7):1199–1213.

Kolm, P. N., Tütüncü, R., and Fabozzi, F. J. (2014). 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, 234(2):356–371.

Lim, A. E., Shanthikumar, J. G., and Vahn, G.-Y. (2011). Conditional value-at-risk in portfolio optimization: Coherent but fragile. *Operations Research Letters*, 39(3):163–171.

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

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

Nawrocki, D. N. (1999). A brief history of downside risk measures. *The Journal of Investing*, 8(3):9–25.

Peterson, B. G., Carl, P., Boudt, K., Bennett, R., Ulrich, J., and Zivot, E. (2014). PerformanceAnalytics: Econometric tools for performance and risk analysis. *R Package Version*, 1(3541):107.

Pflug, G. C. (2000). Some remarks on the value-at-risk and the conditional value-at-risk. In Uryasev, S., editor, *Probabilistic Constrained Optimization: Methodology and Applications*, pages 272–281. Springer.

Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–42.

Rockafellar, R. T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471.

Roman, D., Mitra, G., and Spagnolo, N. (2010). Hidden Markov models for financial optimization problems. *IMA Journal of Management Mathematics*, 21(2):111–129.

Roy, A. D. (1952). Safety first and the holding of assets. *Econometrica: Journal of the Econometric Society*, pages 431–449.

Sarı, D., Lee, Y., Ryan, S., and Woodruff, D. (2016). Statistical metrics for assessing the quality of wind power scenarios for stochastic unit commitment. *Wind Energy*, 19(5):873–893.

Sarı, D. and Ryan, S. M. (2016). MTDrh: Mass transportation distance rank histogram. https://cran.r-project.org/web/packages/MTDrh/index.html.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, 39(1):119–138.

Sharpe, W. F. (1994). The Sharpe ratio. *Journal of Portfolio Management*, 21(1):49–58.

Sortino, F. A. and Price, L. N. (1994). Performance measurement in a downside risk framework. *The Journal of Investing*, 3(3):59–64.

Tse, Y. (2015). Momentum strategies with stock index exchange-traded funds. *The North American Journal of Economics and Finance*, 33:134–148.

Young, T. W. (1991). Calmar ratio: A smoother tool. *Futures*, 20(1):40.

Yu, L.-Y., Ji, X.-D., and Wang, S.-Y. (2003). Stochastic programming models in financial optimization: A survey. *AMO - Advanced Modeling and Optimization*, 5(1):1–26.

3.9 Appendix: Electronic Companion

This electronic companion provides performance results corresponding to those in Section 5.2 for alternative values of G, the number of past periods used to estimate volatility in the scenario generation procedure. Qualitatively, the results for $G = 6$, 8 or 12 are consistent with those reported in the paper for $G = 10$.

Table 3.7: Annualized performance metrics of different strategies sorted by the descending Sortino ratio when $G = 6$

Group	Paramter value	Sortino ratio (%)	Sharpe ratio (%)	Maximum drawdown (%)

TSMDR	$\alpha = 75\%$, C* at Q2	88.92	59.19	25.59
TSMDR	$\alpha = 75\%$, C* at Q1	88.28	59.18	23.07
TSMDR	$\alpha = 90\%$, C* at Q1	85.78	57.60	23.95
TSMDR	$\alpha = 90\%$, C* at Q2	84.83	56.85	27.27
TSMDR	$\alpha = 99\%$, C* at Q1	83.04	55.93	24.88
TSMDR	$\alpha = 75\%$, C* at Q3	82.38	54.76	29.31
TSMDR	$\alpha = 99\%$, C* at Q2	82.04	55.15	28.20
TSMDR	$\alpha = 90\%$, C* at Q3	81.30	53.80	29.87
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.02$	80.84	52.28	38.53
Mean-Variance	$\lambda = 0.5$	80.02	51.61	34.83
Mean-Variance	$\lambda = 1/3$	79.93	51.96	38.60
TSMDR	$\alpha = 99\%$, C* at Q3	79.81	52.75	30.33
TSMDR	$\alpha = 75\%$, C* at 90%	78.43	51.76	30.96
Mean-Variance	$\lambda = 0.1$	78.38	50.98	40.40
TSMDR	$\alpha = 90\%$, C* at 90%	77.96	51.26	31.48
TSMDR	$\alpha = 99\%$, C* at 90%	77.58	50.88	32.51
Mean-Variance	$\lambda = 0.7$	75.49	48.37	28.41
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.04$	74.93	48.73	38.53
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.02$	74.93	48.73	38.53
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.06$	74.01	47.73	31.68
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.02$	73.81	48.07	38.53
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.06$	73.56	46.72	22.47
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.04$	73.47	46.66	22.47
Mean-CVaR	$\alpha = 75\%$, λ= 0.08	72.02	45.77	22.47
TSMOM	C at Q1	70.81	46.64	26.72
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.04$	70.62	45.93	35.46
Risk-Neutral	$\lambda = 0$	70.60	45.93	40.85
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.1$	67.48	43.13	21.46
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.08$	66.36	42.46	21.46
TSMOM	C at Q2	65.92	43.41	30.82

TSMOM	C at Q3	65.80	42.70	31.52
TSMOM	C at 90%	62.25	40.50	32.59
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.1$	59.04	37.94	30.36
Risk-Neutral	Unscaled TSMOM	57.94	38.05	37.99
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.06$	53.77	34.88	30.52
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.08$	44.41	28.88	32.47
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.1$	43.02	27.59	28.69
Mean-Variance	$\lambda = 0.9$	38.63	28.21	24.47

Table 3.8: Annualized performance metrics of different strategies sorted by the descending Sortino ratio when $G = 8$

Group	Parameter value	Sortino ratio (%)	Sharpe ratio (%)	Maximum drawdown (%)
TSMDR	$\alpha = 75\%$, C* at Q1	89.04	59.51	21.98
TSMDR	$\alpha = 75\%$, C* at Q2	86.87	57.98	25.88
TSMDR	$\alpha = 90\%$, C* at Q1	85.72	57.48	23.27
TSMDR	$\alpha = 90\%$, C* at Q2	84.19	56.29	27.25
TSMDR	$\alpha = 99\%$, C* at Q1	82.89	55.65	24.53
TSMDR	$\alpha = 99\%$, C* at Q2	81.17	54.37	28.36
TSMDR	$\alpha = 75\%$, C* at Q3	81.16	53.87	30.54
TSMDR	$\alpha = 90\%$, C* at Q3	80.78	53.26	31.12
TSMDR	$\alpha = 75\%$, C* at 90%	80.25	52.53	31.34
TSMDR	$\alpha = 99\%$, C* at Q3	80.16	52.53	31.86
TSMDR	$\alpha = 90\%$, C* at 90%	79.76	52.07	32.05
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.08$	78.01	49.46	21.47
TSMDR	$\alpha = 99\%$, C* at 90%	77.97	50.85	33.79
Mean-Variance	$\lambda = 0.5$	77.96	50.47	36.45
Mean-Variance	$\lambda = 0.7$	77.46	49.42	29.87
Mean-Variance	$\lambda = 1/3$	77.00	50.17	39.24
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.06$	76.76	48.68	21.47

Mean-Variance	$\lambda = 0.1$	75.55	49.51	40.31
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.04$	75.31	47.78	21.46
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.02$	74.76	48.68	38.53
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.02$	74.24	48.35	38.53
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.04$	72.13	47.00	38.53
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.02$	72.13	47.00	38.53
TSMOM	C at Q1	70.71	46.58	26.84
Risk-Neutral	$\lambda = 0$	70.60	45.93	40.85
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.06$	69.78	45.39	35.46
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.04$	69.50	45.32	38.53
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.1$	68.01	43.49	21.46
TSMOM	C at Q2	66.10	43.58	30.64
TSMOM	C at Q3	65.75	42.66	31.53
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.08$	63.22	40.71	25.99
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.06$	63.02	40.47	30.52
TSMOM	C at 90%	62.25	40.50	32.58
Risk-Neutral	Unscaled TSMOM	57.94	38.05	37.99
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.1$	54.79	35.29	30.52
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.1$	48.13	30.45	21.54
Mean-Variance	$\lambda = 0.9$	46.68	32.52	24.87
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.08$	44.11	28.76	32.47

Table 3.9: Annualized performance metrics of different strategies sorted by the descending Sortino ratio when $G = 12$

Group	Parameter value	Sortino ratio(%)	Sharpe ratio(%)	Maximum drawdown(%)
TSMDR	$\alpha = 75\%$, C* at Q1	88.06	58.87	20.71
TSMDR	$\alpha = 75\%$, C* at 90%	88.06	58.87	20.71
TSMDR	$\alpha = 90\%$, C* at Q1	86.32	57.49	22.09
TSMDR	$\alpha = 99\%$, C* at Q1	84.14	55.96	23.98

TSMDR	$\alpha = 75\%$, C* at Q2	80.09	54.00	26.22
TSMDR	$\alpha = 90\%$, C* at 90%	79.42	51.86	32.62
TSMDR	$\alpha = 75\%$, C* at Q3	78.80	52.66	29.39
TSMDR	$\alpha = 99\%$, C* at 90%	78.76	51.25	34.12
TSMDR	$\alpha = 90\%$, C* at Q3	77.62	51.62	31.01
Mean-Variance	$\lambda = 0.5$	77.61	50.31	36.86
TSMDR	$\alpha = 90\%$, C* at Q2	77.55	52.23	27.64
TSMDR	$\alpha = 99\%$, C* at Q3	76.89	50.79	33.00
Mean-Variance	$\lambda = 1/3$	76.01	49.58	40.44
TSMDR	$\alpha = 99\%$, C* at Q2	75.69	50.71	30.79
Mean-Variance	$\lambda = 0.1$	74.80	49.02	40.25
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.02$	74.24	48.35	38.53
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.02$	74.13	48.30	40.55
Mean-Variance	$\lambda = 0.7$	71.62	46.24	31.31
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.04$	71.36	46.49	38.53
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.08$	70.96	45.33	25.01
TSMOM	C at Q1	70.90	46.70	26.61
Risk-Neutral	$\lambda = 0$	70.60	45.93	40.85
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.04$	69.93	45.58	38.53
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.02$	69.93	45.58	38.53
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.06$	69.57	45.36	38.53
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.06$	68.85	44.03	27.43
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.04$	68.85	44.03	27.43
Mean-CVaR	$\alpha = 75\%$, $\lambda = 0.1$	66.45	42.36	23.47
TSMOM	C at Q2	66.17	43.63	30.58
TSMOM	C at Q3	65.91	42.78	31.50
TSMOM	C at 90%	62.25	40.50	32.59
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.06$	61.66	39.69	25.99
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.08$	59.65	38.50	25.99
Mean-CVaR	$\alpha = 90\%$, $\lambda = 0.1$	59.27	38.16	26.34

Risk-Neutral	Unscaled TSMOM	57.94	38.05	37.99
Mean-Variance	$\lambda = 0.9$	52.59	34.52	21.87
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.08$	38.97	25.57	31.03
Mean-CVaR	$\alpha = 99\%$, $\lambda = 0.1$	37.53	24.07	27.69

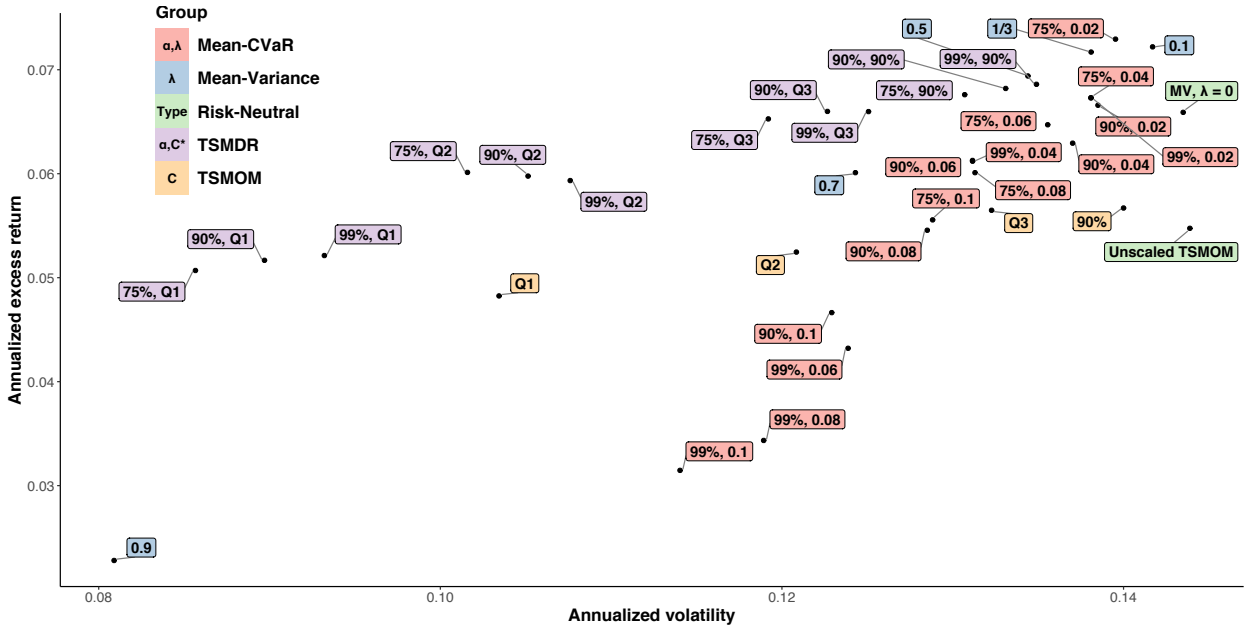


Figure 3.12: Scatter plot of annualized excess returns and volatility for different strategies for $T = 12, G = 6$ and $J = 20000$

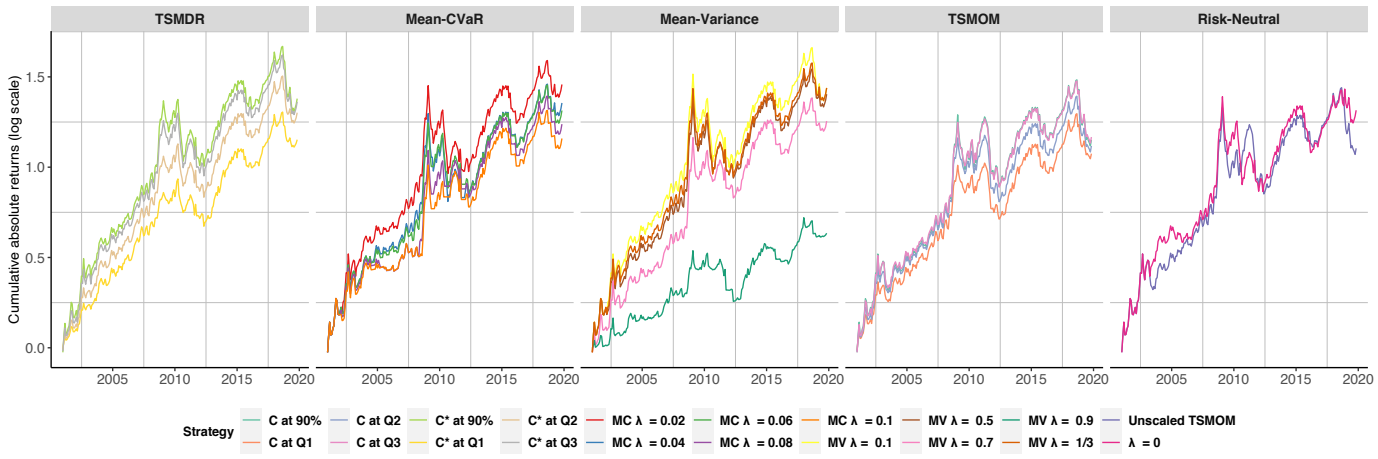


Figure 3.13: Cumulative absolute log returns when $\alpha = 0.75$

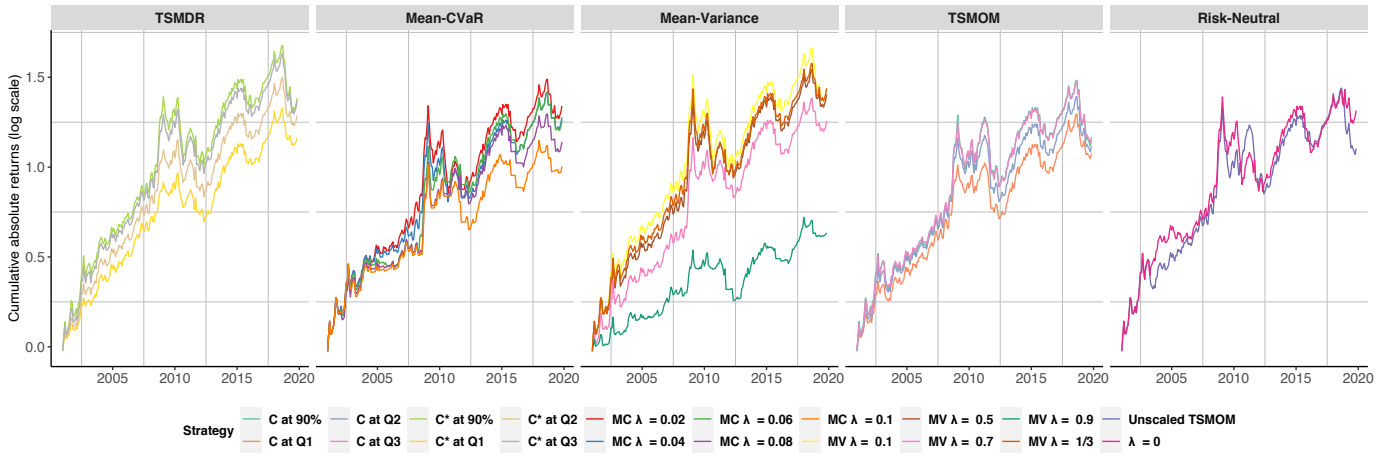


Figure 3.14: Cumulative absolute log returns when $\alpha = 0.90$

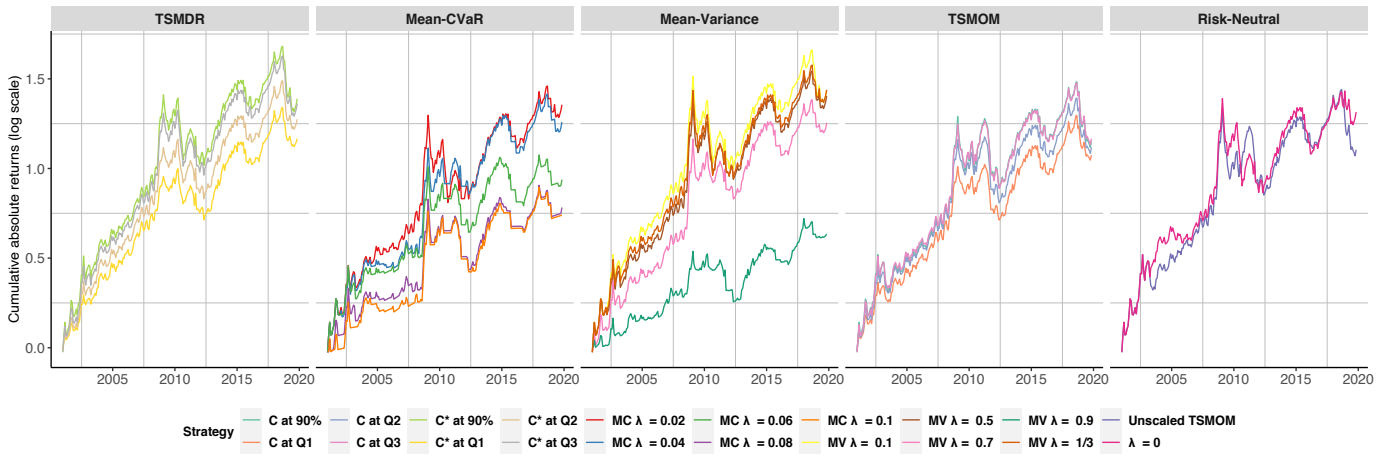


Figure 3.15: Cumulative absolute log returns when $\alpha = 0.99$

Figure 3.16: Cumulative absolute log returns of TSMDR, Mean-CVaR, Mean-Variance, Risk-Neutral, and TSMOM when $G = 6$. Only TSMDR and Mean-CVaR are different in three panels.

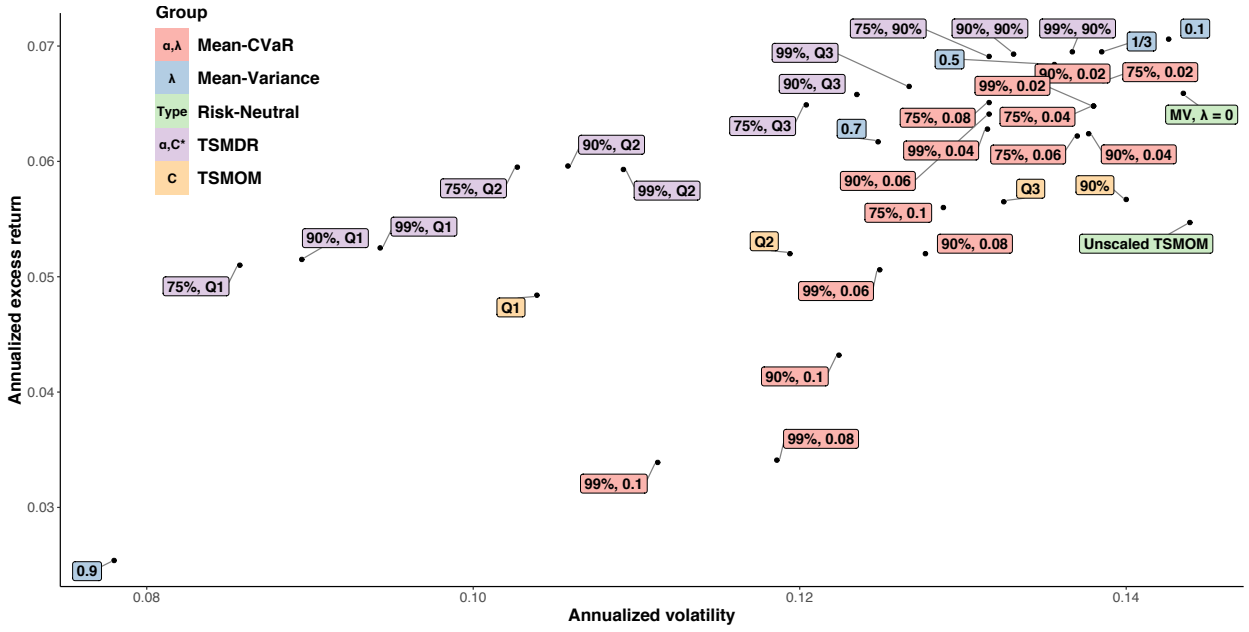


Figure 3.17: Scatter plot of annualized excess returns and volatility for different strategies for $T = 12, G = 8$ and $J = 20000$

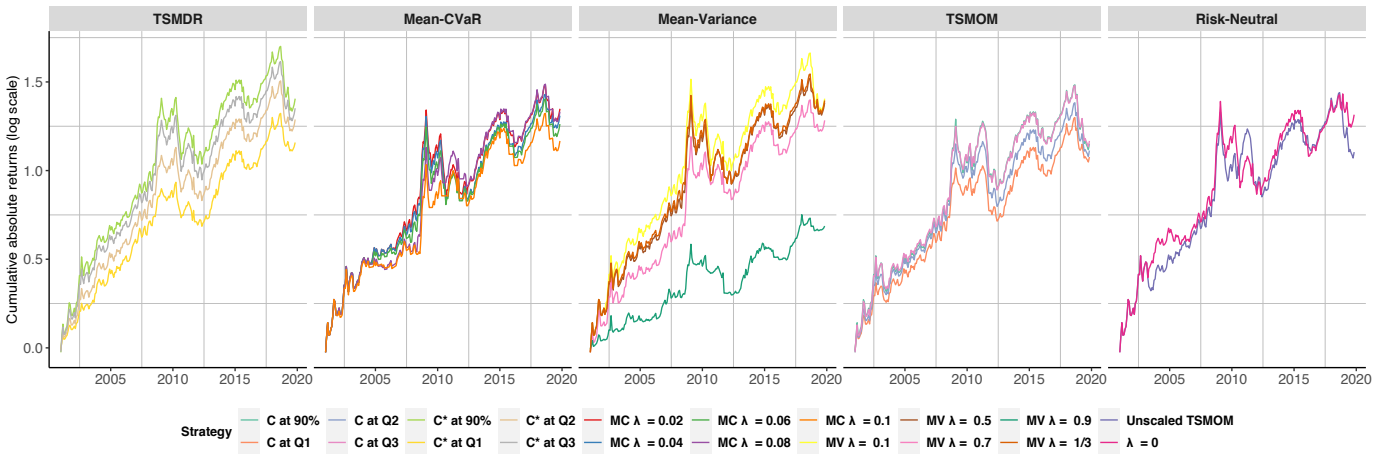


Figure 3.18: Cumulative absolute log returns when $\alpha = 0.75$

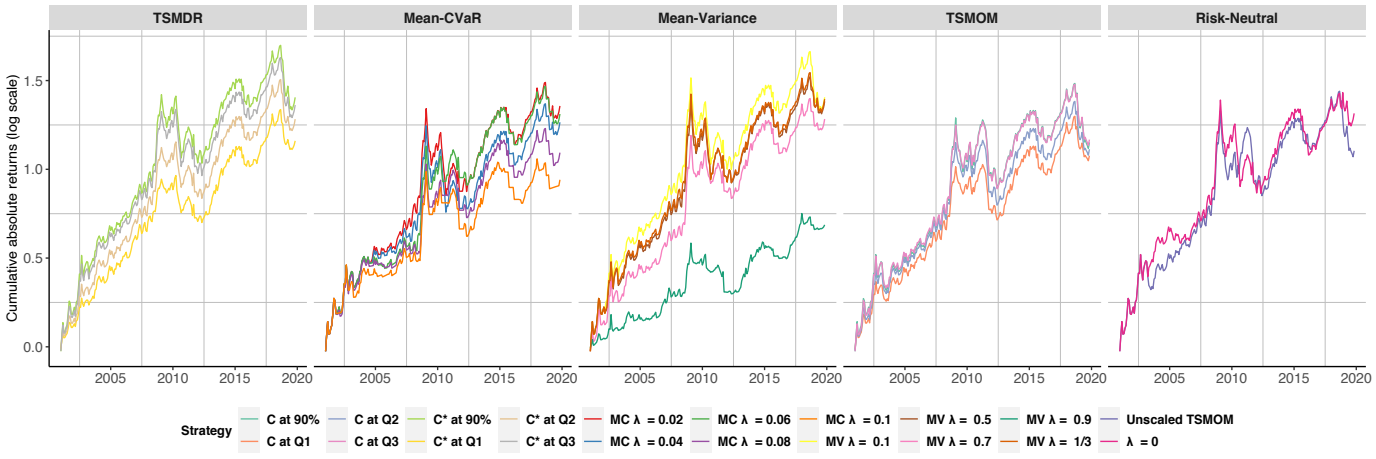


Figure 3.19: Cumulative absolute log returns when $\alpha = 0.90$

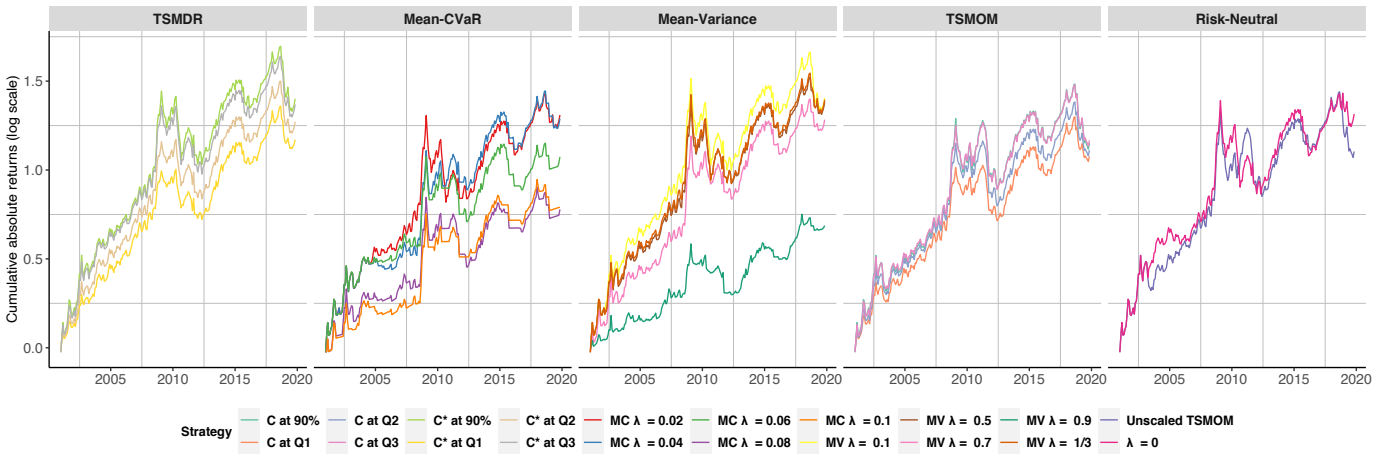


Figure 3.20: Cumulative absolute log returns when $\alpha = 0.99$

Figure 3.21: Cumulative absolute log returns of TSMDR, Mean-CVaR, Mean-Variance, Risk-Neutral, and TSMOM when $G = 8$. Only TSMDR and Mean-CVaR are different in three panels.

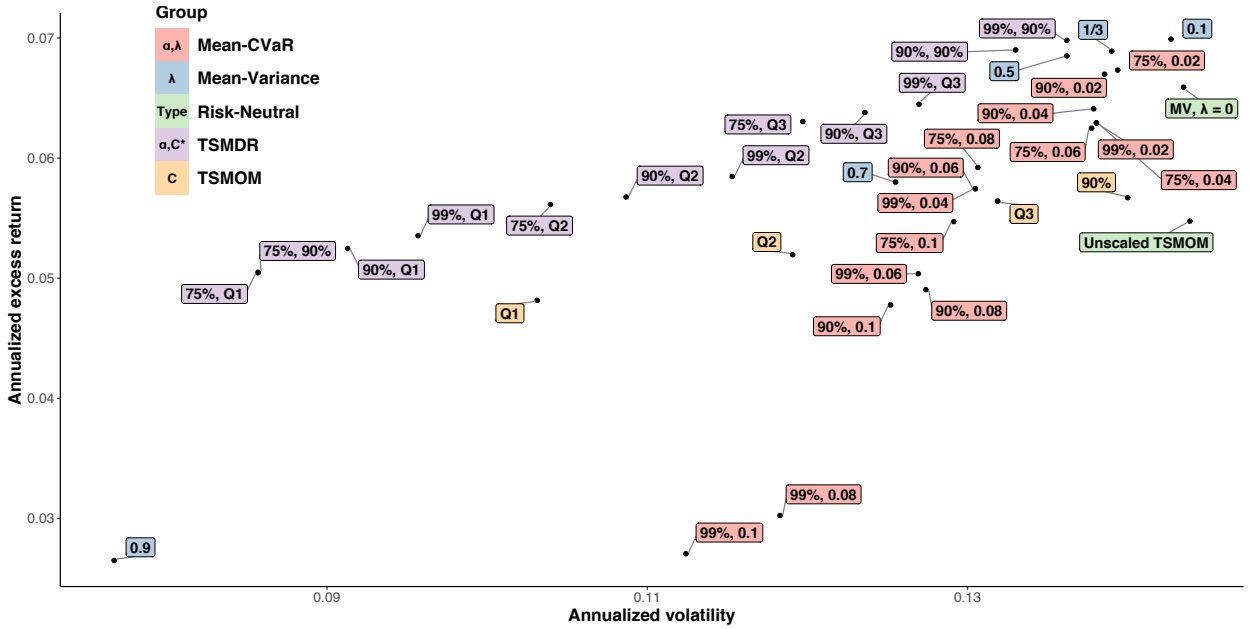


Figure 3.22: Scatter plot of annualized excess returns and volatility for different strategies for $T = 12, G = 12$ and $J = 20000$

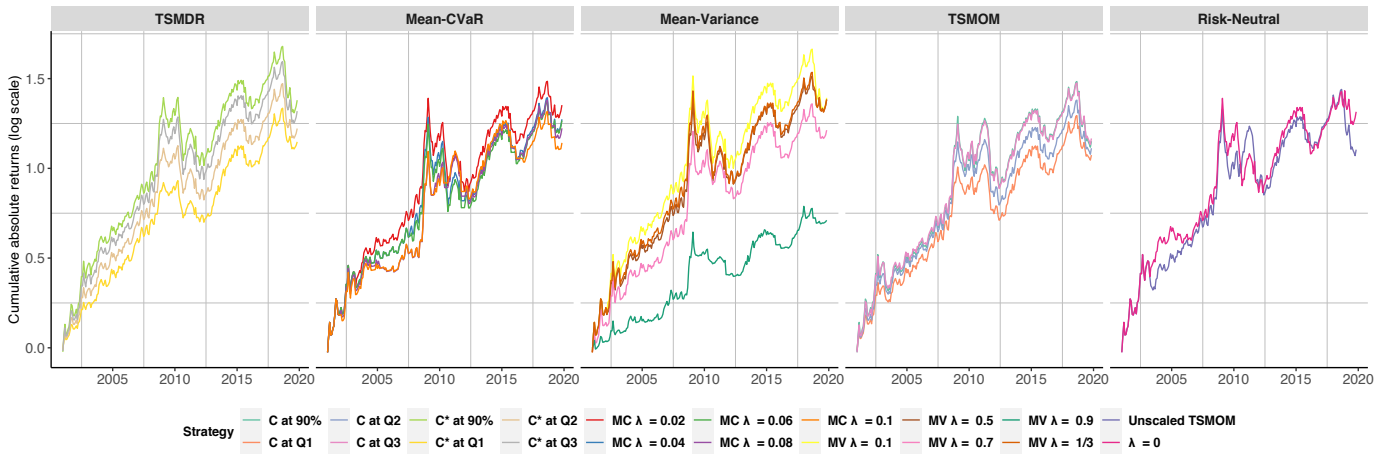


Figure 3.23: Cumulative absolute log returns when $\alpha = 0.75$

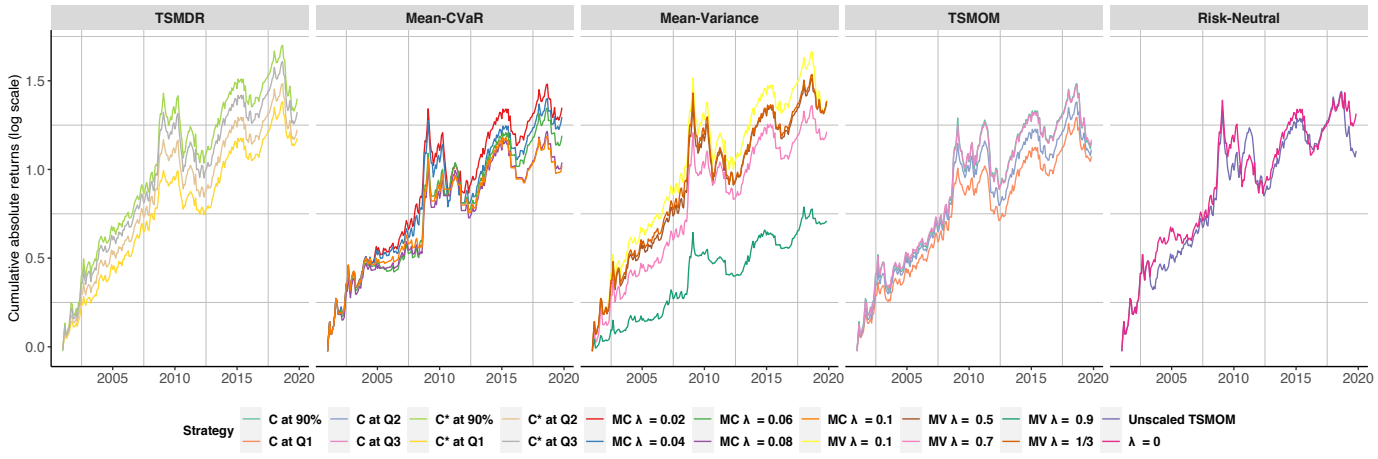


Figure 3.24: Cumulative absolute log returns when $\alpha = 0.90$

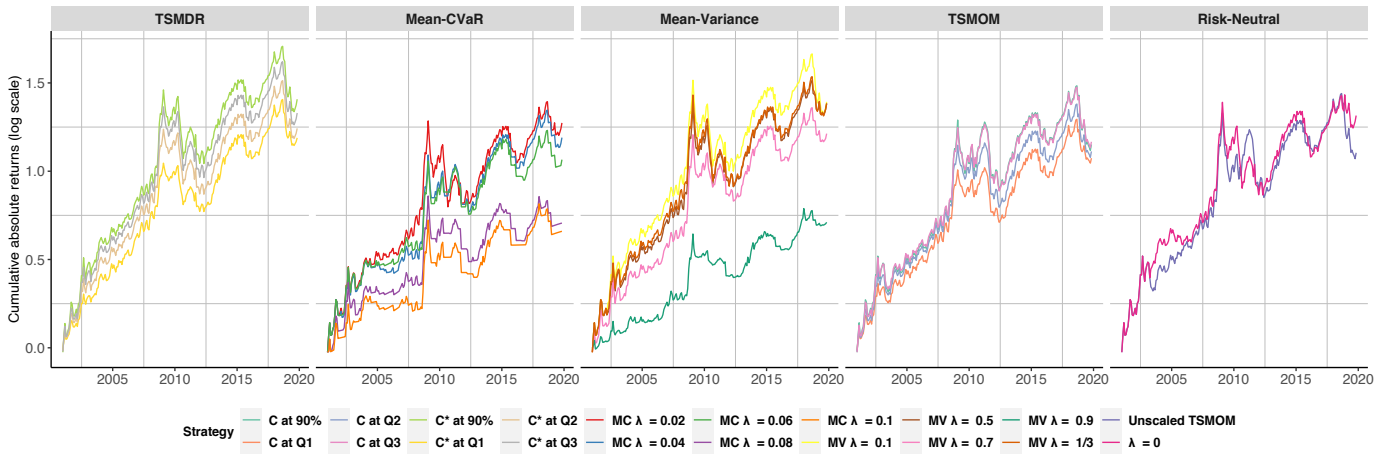


Figure 3.25: Cumulative absolute log returns when $\alpha = 0.99$

Figure 3.26: Cumulative absolute log returns of TSMDR, Mean-CVaR, Mean-Variance, Risk-Neutral, and TSMOM when $G = 12$. Only TSMDR and Mean-CVaR are different in three panels.

CHAPTER 4. AVOIDING MOMENTUM CRASHES USING STOCHASTIC MEAN-CVAR OPTIMIZATION WITH TIME-VARYING RISK AVERSION

A manuscript to be submitted to *The Engineering Economist*

Abstract

Despite its exceptional performance across multiple periods and various asset classes, the cross-sectional momentum strategy for financial asset allocation infrequently produces drastically negative returns, making the strategy risky for investors. In this paper, we develop a stochastic mean-risk optimization model with time-varying risk aversion to avoid momentum crashes in three ways: 1) A downside risk measure, CVaR, is used to control the expected investment loss under the extreme case, 2) The tail probability used to compute CVaR and the risk-aversion coefficient in the objective function are time-dependent to adjust the investment strategy according to the changing market conditions, and 3) A hybrid moment-matching algorithm is applied to generate scenarios that capture the statistical properties, including the first four moments and correlation, of uncertain momentum portfolio returns. Scenario quality is evaluated according to uniformity of the mass transportation distance rank histogram. By backtesting from January, 1926, to December, 2020, using a broad universe of equities, we find that portfolios reblanced according to our model not only provide higher returns but also reduce investment risk compared to a cross-sectional momentum heuristic. More importantly, there is no drastic decline in cumulative returns during the whole simulation horizon, implying that the momentum crashes are successfully avoided.

4.1 Introduction

Momentum in finance is an empirical observation that trends in asset returns persist over the short term. The existence of momentum is a market anomaly as it contradicts the efficient market hypothesis that past returns cannot determine the current return. Investors exploit this return predictability of momentum to construct trading strategies that take a long position (buy in anticipation of an increase in value) in upward trending assets and a short position (sell first with

the plan to buy later at a lower price) in downward trending assets with the expectation that their momentum will extend the trends.

The cross-sectional momentum trading strategy, that forms portfolios by cross-sectional ranking, is the most typical and frequently-used momentum trading strategy. It originates from the work of Jegadeesh and Titman (1993), who rank the cross-section of stocks according to the three- to twelve-month past returns to identify winners (stocks that have performed well) and losers (stocks that have performed poorly), buy past winners and short past losers, and hold these positions for up to twelve months. They find that over the 1965 to 1989 period, past losers continued to perform badly and past winners continued to perform well during these holding periods, indicating that cross-sectional momentum exists and no return reversals occur in the short term. In this paper, we refer to portfolios composed of past winners and/or past losers as momentum portfolios.

Evidence supporting cross-sectional momentum is plentiful. Jegadeesh and Titman (1993) find that the momentum strategy can gain as much as 1.49% per month in the US stock market. Furthermore, momentum is not only an anomaly in the US stock market, but also observed in European equities, emerging markets, country indices, industry portfolios, currencies, commodities, and across asset classes (Grinblatt and Titman, 1993; Asness et al., 1997; Okunev and White, 2003; Grobys et al., 2018). Compared to the factors (market, value, and size) in the classic Fama and French three-factor asset pricing model, cross-sectional momentum offers the highest Sharpe ratio (Barroso and Santa-Clara, 2015).

However, the remarkable performance of momentum strategies comes with large crashes occasionally. Daniel and Moskowitz (2016) show that the cross-sectional momentum strategy would have produced infrequent and persistent negative returns in the US stock market in 1932 and 2009, where the return derived by the strategy would have been -91.59% over two months and -73.42% over three months, respectively. Recovering from the negative impact of these crashes would have taken decades. Daniel and Moskowitz (2016) find that these two crashes both happened when the market rebounded following a severe bear market. They conclude that the reason behind this phenomenon is the asymmetric behavior of past winner and loser portfolios: the return of past losers grows rapidly when the market is volatile and the market return increases after a bear market, making the past losers behave like a call option on the market; while past winner portfolios do not show any similar behaviors during the same period. This asymmetric behavior causes a reversal in

the relationship between past losers and past winners in the recovery phase and, thus, leads to a huge loss for the momentum strategy that shorts the past losers.

The occurrence of the momentum crashes is due mostly to the uncertainty of the stock market, especially concerning the time when the bear market ends and past losers stop being losers. In this paper, to improve on the performance of the cross-sectional momentum strategy and avoid momentum crashes, we explore alternative ways to allocate investment between a risk-free asset and the two momentum portfolios: 1) control the risk of momentum portfolios, 2) exploit time-varying risk aversion to adjust the momentum trading strategy according to the stock market states, and 3) simulate momentum portfolio returns in a way that comprehensively accounts for the possible returns in the next period.

To manage the risk of the investment strategies, the risk measure needs to be selected. Barroso and Santa-Clara (2015) and Daniel and Moskowitz (2016) both quantify the risk of momentum portfolios in terms of volatility (variance). Inspired by the risk parity approaches (Asness et al., 2012), Barroso and Santa-Clara (2015) scale the winner-minus-loser (WML) portfolio return by its realized volatility of past six-month returns to control the investment risk of the cross-sectional momentum strategy. Daniel and Moskowitz (2016) maximize the Sharpe ratio of the WML portfolio, which turns to be a variant of the traditional Markowitz mean-variance model (Markowitz, 1952). Volatility and variance quantify risk in terms of both upside gains and downside losses, and thus, cannot distinguish the direction of investment movement. However, investors worry about the loss only. In addition, researchers have found that higher moments, such as skewness and kurtosis, are non-negligible in asset allocation when the asset returns are not normally distributed. Therefore, a downside risk measure, capable of being expressed as a function of skewness and kurtosis for non-normal investment returns and assessing the extent of losses only (Xiong and Idzorek, 2011), is more suitable to reduce the risk of momentum investing. By far, value at risk (VaR) and conditional value-at-risk (CVaR) are the two most commonly-used downside risk measures in finance (Kim et al., 2012; Banihashemi and Navidi, 2017; Sharifi and Kwon, 2018; Guo et al., 2019; Kim and Park, 2020; Benati and Conde, 2021). VaR estimates the maximum loss that could occur over a given period with a specified confidence level. CVaR is defined as the conditional expectation of portfolio losses beyond a prespecified percentile of the distribution of the loss (Rockafellar and Uryasev, 2000). Although both VaR and CVaR are superior to variance, there are two reasons for the preference in

portfolio and risk management of CVaR relative to VaR. First, CVaR takes into account not only the probability but also the size of the loss (Rockafellar and Uryasev, 2000). Second, Pflug (2000) proves that CVaR satisfies the properties for a coherent risk measure while VaR fails due to the violation of the sub-additivity property. More recently, Ahmadi-Javid (2012) proposes a new risk measure, entropic value-at-risk (EVaR), which uses the Chernoff inequality to create an upper bound for the CVaR and is proven to be coherent. Compared to CVaR, EVaR has stronger monotonicity and is more computationally efficient when solving large-scale optimization problems due to the number of variables and constraints of the EVaR formulation being independent of the sample size (Ahmadi-Javid and Fallah-Tafti, 2019). However, EVaR is more risk-averse at the same confidence level than CVaR; that is, the incorporation of EVaR will result in allocating more resources to avoid risk, which may not be favored by investors who want to allocate the least amount of resources. This property makes EVaR less attractive than CVaR.

Having selected CVaR as the risk measure, a mean-CVaR stochastic program with time-varying risk aversion is developed to balance the trade-off between investment risk and reward. Stochastic programming provides a modeling framework for decision making under uncertainty. It is well suited for portfolio management as investment decisions have to be made before the return information is revealed. Numerous applications of stochastic programming to portfolio management have been studied in the past decades (Topaloglou et al., 2011; Lim et al., 2011; Moazeni et al., 2016; Cui et al., 2020; Guo and Ryan, 2020a; Bergk et al., 2021). Motivated by the regime-switching model (Bae et al., 2014; Li et al., 2021), in this paper, we incorporate time-varying risk aversion[1] to select portfolio weights that dynamically balance the trade-off between expectation and CVaR of the return of the constructed portfolio according to various stock market conditions. As momentum crashes happen when the market is volatile, we explore three volatility-related measures, namely 1) estimated market volatility, 2) ratio of estimated market return to estimated market volatility, and 3) the Chicago Board Options Exchange (CBOE) Volatility Index (VIX), to determine the tail probability in CVaR and the risk-aversion parameter in the mean-CVaR model.

[1]We use the term risk aversion to be consistent with its use in Li et al. (2021), where risk aversion gauges investors' reactions rather than their attitudes to the risk. For instance, in our usage, an investor who reduces the weight on a risky asset during a bear market is behaving in a more risk-averse way than an investor who maintains the same position, even if their attitudes towards risk may be the same.

One of the biggest advantages of stochastic programming is its flexibility in modeling uncertainty. In a stochastic program, the uncertainty of the input parameters is represented in terms of discretely distributed scenarios. The various approaches to generating scenarios in stochastic programs can be categorized into two classes depending on whether the distribution of the uncertain parameters is specified or not. When the distribution is specified, Monte Carlo simulation can be applied to generate scenarios by repeated random sampling from the specified distribution (Ripley, 1987; Yu et al., 2003; Sharifi et al., 2016; Guo and Ryan, 2020b). On the other hand, when the joint distribution of the uncertain parameters is not specified, bootstrapping, which samples the historical data randomly with replacement, is a straightforward way to generate scenarios (Zou et al., 2019; Barro et al., 2019; Thomann, 2021). Although bootstrapping is easy to implement, it is unable to capture some inherent structure of stock returns. Høyland and Wallace (2001) propose a moment-matching method that consists of sequentially solving an optimization problem with the objective to minimize the distance between the statistical properties of scenarios and specified target values estimated on the basis of historical data. Høyland et al. (2003) propose a heuristic procedure to approximately implement the moment-matching method for the first four marginal moments and correlations. To approximate the unknown distribution of momentum portfolio returns, including the large kurtosis indicated by occasional momentum crashes, we employ a combination of heuristic and optimization moment-matching algorithms to generate scenarios of future returns of the momentum portfolios.

Upon generating scenarios, the next question arising in a backtesting environment is how well they match the observed data. The verification rank histogram, a recently-developed technique that compares sets of scenarios that would have been generated in the past with the corresponding observations, can be applied to evaluate the quality of scenarios. The mass transportation distance (MTD) rank histogram (Sarı et al., 2016) is one such verification rank histogram. The principle behind the rank histogram is that the set of reliable scenarios generated at a given time point and its corresponding observation can be regarded as random samples drawn from the same distribution. It evaluates the reliability of the scenarios according to the uniformity of the distribution of the rank of the scenarios-to-observation distance among corresponding distances from each scenario to the rest, augmented by the observation. Therefore, the rank histogram of reliable scenarios is supposed to be flat.

Finally, we empirically analyze the performance of the mean-CVaR solutions with time-varying risk aversion in a historical backtest of monthly rebalancing. The stock universe consists of all stocks listed on NYSE, AMEX, and NASDAQ with the returns of the common shares. Using historical data from 1926 through 2020, we find that weights derived from our mean-CVaR model with time-varying risk aversion defined using volatility of market returns are robust to the number of months used to estimate market index return mean and volatility. For all such numbers of months tested, the resulting portfolio achieves higher excess returns, Sharpe ratio, and upside potential ratio, as well as lower maximum drawdown, than a heuristic solution corresponding to the cross-sectional momentum strategy. In addition, no drastic decline in cumulative returns occurs during the whole simulation horizon. An interesting finding is that no momentum crash would have occurred in 2020 when using either our mean-CVaR model or the cross-sectional momentum heuristic despite the steep decline in the stock market. During the two momentum crash periods, the cumulative returns achieved by rebalancing according to the mean-CVaR solutions held steady around the level of one, indicating that no huge loss is incurred when using this mean-CVaR model and, thus, the momentum crashes would have been avoided completely.

This paper's contribution is a new trading strategy based on cross-sectional momentum, that successfully avoids the momentum crashes and is profitable in various states of the stock market. In this strategy, the investment risk is quantified by downside loss, the risk aversion is time-dependent to flexibly adjust the investment strategy according to the changing market conditions, and the uncertain momentum portfolio returns are well-represented by reliable scenarios that comprehensively account for the possibilities of the possible winner and loser returns that may occur in the next period. To the best of our knowledge, this is the first study to address the issue of momentum crashes using stochastic optimization.

The remainder of this paper is organized as follows. Section 4.2 lists the notations used throughout the whole paper, describes the construction of past winner and loser portfolios in the cross-sectional momentum strategy, and explains the reason to use returns of these two portfolios as random input parameters in Section 4.3, where a two-stage stochastic mean-CVaR optimization model with time-varying risk aversion is introduced. Section 4.4 illustrates a hybrid of heuristic and optimization moment-matching algorithms to generate scenarios of past winner and past loser portfolio returns and evaluates the reliability of the generation procedure. Section 4.5 describes

the metrics used to evaluate the performance of the investment strategies. Section 4.6 presents the numerical experiments and analyzes the results. Section 4.7 finally concludes.

4.2 Notation and Preliminaries

This section lists the notations used in the whole paper, describes the construction of past winner and loser portfolios in the cross-sectional momentum strategy, and explains the reason to use returns of these two portfolios as input random parameters.

4.2.1 Notation

Notations used throughout this paper are defined as follows:
Sets and cardinalities:

B	Set of candidate stocks in the stock market with cardinality $\|B\|$
K	Number of look-back periods for identifying past winner and loser portfolios
T	Number of periods in the simulation horizon
W_t	Set of stocks in the past winner portfolio for period t, $W_t \subset B$, $t = 1, \ldots, T$
L_t	Set of stocks in the past loser portfolio for period t, $L_t \subset B$, $t = 1, \ldots, T$
J	Number of scenarios generated for each period t, $t = 1, \ldots, T$
M	Set of all specified statistical properties for scenario generation
H	Number of periods used to estimate statistics in M
P	Number of periods used to estimate mean of market returns for adjusting risk parameters
G	Number of periods used to estimate volatility of market returns for adjusting risk parameters
N	Number of periods in a year

Deterministic input data:

S_t^b	Observed price of stock $b \in B$ at time $t = 1, \ldots, T-1$
O_t^b	Number of shares outstanding of stock $b \in B$ at time $t = 1, \ldots, T-1$
r_t^b	Observed rate of return for stock $b \in B$ in period $t = 1, \ldots, T-1$
r_t^W	Observed rate of return for the past winner portfolio in period $t = 1, \ldots, T-1$
r_t^L	Observed rate of return for the past loser portfolio in period $t = 1, \ldots, T-1$
r^t	Realized rate of return for the constructed portfolio in period t, which is an output in Section 4.3 and an input in Section 4.5, $t = 1, \ldots, T$
r_t^f	Risk-free rate of return in period $t = 1, \ldots, T$
i_t	Market index return rate in period $t = 1, \ldots, T-1$
$M_{VALm,t}$	Specified value of statistical property $m \in M$ for period $t = 1, \ldots, T$
v_m	Weight assigned to statistical property $m \in M$

Parameters for the portfolio optimization model for period t, $t = 1, \ldots, T$:

R_t^W	Random variable of the rate of return for the past winner portfolio in period t with scenarios $r_{t,j}^W, j = 1, \cdots, J$
R_t^L	Random variable of the rate of return for the past loser portfolio in period t with scenarios $r_{t,j}^L, j = 1, \cdots, J$
m_t	Market return estimated for period t
σ_t	Market volatility estimated for period t
A_t	Rank of σ_t among $\sigma_k, k = 1, \cdots, t$, sorted in increasing order
B_t	Rank of $\frac{m_t}{\sigma_t}$ among $\frac{m_k}{\sigma_k}, k = 1, \cdots, t$, sorted in decreasing order
C_t	Rank of VIX at time t among values up to t sorted in increasing order

λ_t	Risk-aversion parameter for period t, $\lambda_t \in [0, 1]$
α_t	Probability specified for CVaR to pertain to 100(1-α_t)% worst outcomes for period t, $\alpha_t \in [0, 1]$

Decision variables in scenario generation procedure for period t, $t = 1, \ldots, T$:

$p_{t,j}$	Probability of the occurrence of scenario j, $j = 1, \ldots, J$
$r^W_{t,j}$	Return of the past winner portfolio in scenario j, $j = 1, \ldots, J$
$r^L_{t,j}$	Return of the past loser portfolio in scenario j, $j = 1, \ldots, J$

Decision variables in the portfolio optimization model for period t, $t = 1, \ldots, T$:

w^W_t	Weight on the past winner portfolio
w^L_t	Weight on the past loser portfolio
w^f_t	Weight on the risk-free asset
η_t	α-quantile of the negative return rate of the constructed portfolio according to the scenarios generated for period t
$\nu_{t,j}$	Deviation of the return rate of the constructed portfolio below η_t in scenario j as a non-negative simple recourse variable, $j = 1, \ldots, J$
w^{W*}_t	Optimal weight on the past winner portfolio

4.2.2 Past Winner and Past Loser Portfolios

In this paper, the past winner and loser portfolios, instead of the individual stocks, are used as the basis of the optimization model for two reasons. First, Daniel and Moskowitz (2016) conclude that momentum crashes occur because the winner-minus-loser (WML) portfolio behaves like shorting a call option on the market during the rebound from a bear market. However, this optionality is asymmetric as it is mainly associated with the past loser portfolios. The second reason is the concern of the computational cost when generating scenarios of returns and using them as inputs to the stochastic optimization model. Compared to generating scenarios of all stocks that constitute

the past loser and winner portfolios at each time, generating scenarios for only past winner and loser returns is much faster to implement.

We rebalance our portfolio every period and the holding period is the interval between two rebalancing time points. Denote the beginning of period t as time $t-1$ and its end as time t. To form the past winner and loser portfolios used in the cross-sectional momentum strategy, proposed by Jegadeesh and Titman (1993), at time $t-1$ we rank stocks based on their cumulative returns from K periods before to one period before the formation date (i.e. the $t-K$ to $t-2$-period return), $\Pi_{i=t-K}^{i=t-2}(1+r_i^b)-1$, in the ascending order. The one-period gap here is used to avoid short-term reversal, some of the bid-ask spread, price pressure, and lagged reaction effects mentioned in Lehmann (1990) and Jegadeesh and Titman (1993). Based on these rankings, the past winner portfolio for period t is constructed by value weighting the bottom decile stocks, denoted as W_t, and the past loser portfolio for period t is composed by value weighting the top decile stocks, represented by L_t. To obtain the value weight of stock b in W_t or L_t, the stock price, S_{t-1}^b, and the number of shares outstanding, O_{t-1}^b, of stock b at time $t-1$ are used. More specifically, for $b \in W_t$, the value weight in period t is

$$v_t^b = \frac{S_{t-1}^b O_{t-1}^b}{\sum_{a \in W_t} S_{t-1}^a O_{t-1}^a} \tag{4.1}$$

Similarly, the value weight of $b \in L_t$ in period t is

$$v_t^b = \frac{S_{t-1}^b O_{t-1}^b}{\sum_{a \in L_t} S_{t-1}^a O_{t-1}^a} \tag{4.2}$$

Consequently, the return rates of past winner and loser portfolios in period t, r_t^W and r_t^L, are the value-weighted average of returns in period t of the stocks composing the winner and loser portfolios in the look-back period, respectively:

$$r_t^W \equiv \sum_{b \in W_t} v_t^b r_t^b \tag{4.3}$$

and

$$r_t^L \equiv \sum_{b \in L_t} v_t^b r_t^b \tag{4.4}$$

The construction of the cross-sectional momentum strategy is illustrated in Figure 4.1. This strategy has weight one on the past winner portfolio, negative one on the past loser portfolio, and zero on the risk-free asset. To avoid redundancy, we refer to the past winner and past loser portfolios simply as winner and loser portfolios, respectively, for the rest of the paper, unless otherwise noted.

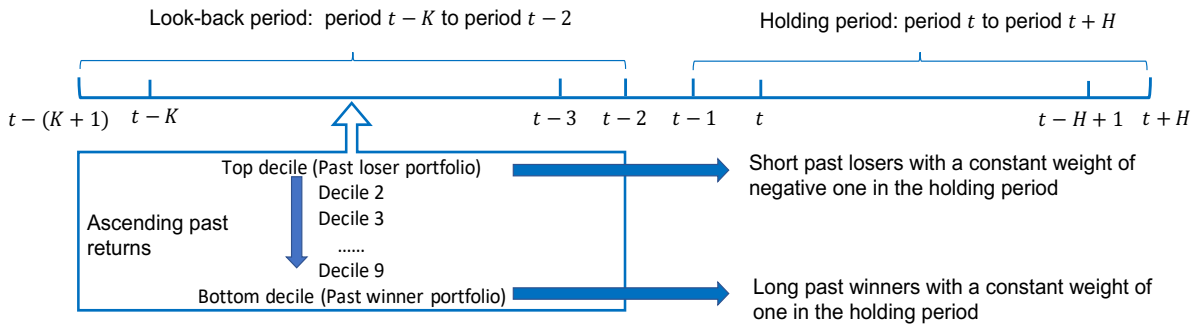


Figure 4.1: Illustration of cross-sectional momentum strategy

4.3 Stochastic Mean-CVaR Optimization Model with Time-Varying Risk Aversion

In this section, we introduce a two-stage stochastic optimization model to dynamically allocate funds between the cross-sectional momentum portfolio and a risk-free asset. At time $t-1$, investors must determine the proportion of wealth, w_t^W, w_t^L, and w_t^f, to invest in the winner and loser portfolios and the risk-free asset, respectively, to maximize the total return and minimize the investment risk of the constructed portfolio in period t. However, the actual returns for the winners and the losers, r_t^W and r_t^L, in which we invest at time $t-1$ will not be known until time t. This model is solved on a rolling basis. More specifically, we use the observation of the risk-free rate in period t and the scenarios for the cross-sectional momentum portfolio return at time t to rebalance the investment at time $t-1$. After solving the model at time $t-1$, the returns of the momentum portfolios in period t t will be realized and stocks in loser and winner portfolios will be updated, according to which we can obtain new scenarios for the return rates of momentum portfolios in period $t+1$. The new scenarios, along with the updated observation of the risk-free rate in period $t+1$, are input to the optimization model to optimally rebalance at time t. This rebalancing process will be continued through period T. The decision point and information release schedule of each period are illustrated in Figure 4.2.

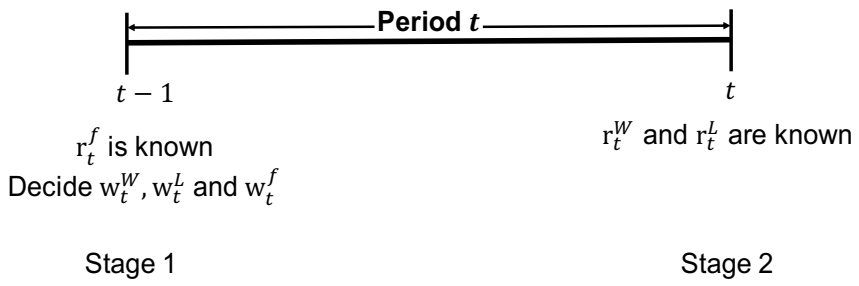


Figure 4.2: Timeline of period t

4.3.1 Mean-CVaR Optimization Model

To balance the trade-off between return and risk, Daniel and Moskowitz (2016) derive a rebalancing strategy to maximize Sharpe ratio over the long term subject to a constraint on variance. However, the variance constraint considers gains and losses equally, which does not characterize the investors' attitudes towards risk accurately when the distribution of the investment loss is not symmetric. As an alternative to variance, we consider a downside risk measure, conditional value-at-risk (CVaR) in each period. The CVaR of a loss random variable, Z, at a given probability, α, is defined as the conditional expectation of the loss of the portfolio exceeding or equal to value at risk, $\eta(\alpha)$:

$$\text{CVaR}_\alpha[Z] = \text{E}[Z|Z \geq \eta(\alpha)] \tag{4.5}$$

where $\eta(\alpha) \equiv \min\{t|P(Z \leq t) \geq \alpha, \alpha \in [0,1]\}$ is the α-quantile of Z. Rockafellar and Uryasev (2002) further demonstrate that CVaR can be computed as follows:

$$\text{CVaR}_\alpha[Z] = \eta(\alpha) + \frac{1}{1-\alpha}\text{E}[(Z-\eta(\alpha))_+] \tag{4.6}$$

where $(Z-\eta(\alpha))_+ \equiv \max(Z-\eta(\alpha), 0)$. By controlling CVaR, we can reduce the expected investment loss during the infrequent but devastating momentum crash periods and, thus, mitigate momentum crashes. In our model, Z represents the negative return of the optimized portfolio realized at time

t. Therefore, the objective at time $t-1$, to maximize the expected total portfolio return while minimizing the investment risk measured by CVaR, can be formulated as

$$\begin{aligned}\min_{w_t^W, w_t^L, w_t^f, \eta_t} \; & (1-\lambda_t)\mathrm{E}\left[-\left(R_t^W w_t^W + R_t^L w_t^L + r_t^f w_t^f\right)\right] + \\ & \lambda_t\left(\eta_t + \frac{1}{1-\alpha_t}\mathrm{E}\left[\left(-\left(R_t^W w_t^W + R_t^L w_t^L + r_t^f w_t^f\right) - \eta_t\right)_+\right]\right)\end{aligned} \tag{4.7}$$

where R_t^W and R_t^L are the random variables of the winner and loser returns, respectively, at time t. The risk-aversion coefficient λ_t represents how risk-averse the investors want the investment strategy to be: The larger the λ_t, the more conservative the strategy. Traditional mean-CVaR models (Lim et al., 2011; Chen et al., 2012; Cui et al., 2020)) use constant tail probability, $1-\alpha$, and risk-aversion parameter, λ. However, in reality, investors' attitude towards risk fluctuates all the time due to the rapidly-shifting financial market. Hence, we extend the traditional mean-CVaR models by allowing α and λ to change over time.

Furthermore, when the joint distribution of R_t^W and R_t^L is approximated by a finite number of probabilistic scenarios, Model (4.7) is transformed into a linear program (Rockafellar and Uryasev, 2000). Upon approximating the joint distribution of R_t^W and R_t^L by J scenarios $\{(r_{t,j}^W, r_{t,j}^L)\}_{j=1}^{j=J}$ with the joint probabilities $\{p_{t,j}\}_{j=1}^{j=J}$, the optimization model for period t is the linear program:

$$\min_{w_t^W, w_t^L, w_t^f, \eta_t, \{\nu_{t,j}\}_{j=1}^J} (1-\lambda_t)\left[-\sum_{j=1}^J p_{t,j}(r_{t,j}^W w_t^W + r_{t,j}^L w_t^L) - r_t^f w_t^f\right] + \lambda_t\left(\eta_t + \frac{1}{1-\alpha_t}\sum_{j=1}^J p_{t,j}\nu_{t,j}\right) \tag{4.8}$$

$$\text{s.t.} \quad -1 \le w_t^W \le 1, \tag{4.9a}$$

$$-1 \le w_t^L \le 1, \tag{4.9b}$$

$$w_t^W + w_t^L = 0, \tag{4.9c}$$

$$w_t^W + w_t^L + w_t^f = 1 \tag{4.9d}$$

$$\nu_{t,j} \ge -p_{t,j}(r_{t,j}^W w_t^W + r_{t,j}^L w_t^L) - r_t^f w_t^f - \eta_t,\ j = 1, \ldots, J \tag{4.9e}$$

$$\nu_{t,j} \ge 0,\ j = 1, \ldots, J \tag{4.9f}$$

As shown in (4.8), the objective function takes a convex combination of mean and CVaR to achieve the goal of maximizing the expected overall return while minimizing CVaR at time t. Note that, the cross-sectional momentum strategy (Jegadeesh and Titman, 1993) is to long winners and short losers, which can be expressed as $w_t^W \equiv 1$ and $w_t^L \equiv -1$ for $\forall t \in \{1, \ldots, T\}$. To be consistent with the original momentum strategy, we allow short sale for the loser portfolios, represented by (4.9b). However, the setting of (4.9a) and (4.9b) is not exactly the same as the original momentum strategy because these two constraints not only allow us to follow the momentum strategy of shorting losers and longing winners but also grant the permission to long losers and short winners. Furthermore, the original momentum strategy for the investment in the winners and losers is zero-cost, which can be expressed as the summation of the weight of the loser and the weight of the winner to be zero, as presented in (4.9c). By forcing the investment in the momentum portfolios to be zero-cost and restricting the weights on the losers and winners to the interval $[-1, 1]$, (4.9a)-(4.9c) include the original momentum strategy coupled with the risk-free asset as a feasible solution to this optimization model. Constraint (4.9d) enforces the total budget to invest in the winner and loser portfolios and the risk-free asset to equal the current wealth, which means the weights of all investment sum up to one. Lastly, constraints (4.9e) and (4.9f) combine to compute the deviation of the weighted returns of the winner, the loser, and the risk-free asset below this value in scenario j as a non-negative simple recourse variable: $\nu_{t,j} = \left(-p_{t,j}(r_{t,j}^W w_t^W + r_{t,j}^L w_t^L) - r_t^f w_t^f - \eta_t\right)_+$.

For each period t, the model (4.8) - (4.9f) can be reduced to the following form, suppressing the subscript t for clarity.

$$\min_{w^W, \eta, \{\nu_j\}_{j=1}^J} (1-\lambda)\left[-\sum_{j=1}^{J} p_j(r_j^W - r_j^L)w^W - r^f\right] + \lambda\left(\eta + \frac{1}{1-\alpha}\sum_{j=1}^{J} p_j \nu_j\right) \tag{4.10}$$

$$\text{s.t.} \quad -1 \leq w^W \leq 1, \tag{4.11a}$$

$$\nu_j \geq -p_j(r_j^W - r_j^L)w^W - r^f - \eta,\ j = 1, \ldots, J \tag{4.11b}$$

$$\nu_j \geq 0,\ j = 1, \ldots, J \tag{4.11c}$$

A simple structure for the optimal solution to the model (4.10) - (4.11c) can be derived in a manner similar to Proposition 1 in Guo and Ryan (2020a). Since CVaR is a coherent risk measure,

the risk-free rate at each time is a deterministic constant and expectation is linear, we find the optimal weight to allocate to the past winner portfolios as follows.

$$w^{W*} = \begin{cases} -1, & \text{if } c^2(\lambda,\alpha) > 0 \text{ and } c^1(\lambda,\alpha) > -c^2(\lambda,\alpha) \\ 0, & \text{if } c^1(\lambda,\alpha) \geq 0 \text{ and } c^2(\lambda,\alpha) \leq 0 \\ 1, & \text{if } c^1(\lambda,\alpha) < 0 \text{ and } c^2(\lambda,\alpha) < -c^1(\lambda,\alpha) \end{cases}$$

where $c^1(\lambda,\alpha) = (\lambda-1)\sum_{j=1}^{J} p_j(r_j^w - r_j^l) + \lambda\left(\eta + \frac{1}{1-\alpha}\left[-\sum_{j=1}^{J} p_j(r_j^w - r_j^l) - \eta\right]_+\right)$, and $c^2(\lambda,\alpha) = (\lambda-1)\sum_{j=1}^{J} p_j(r_j^w - r_j^l) - \lambda\left(\eta + \frac{1}{1-\alpha}\left[\sum_{j=1}^{J} p_j(r_j^w - r_j^l) - \eta\right]_+\right)$.

Another way to interpret this result is to use Corollary 1 in Guo and Ryan (2020a), where w^{W*} depends on the relationship between the expectation and the downside risk of the difference between winner and loser returns:

$$w^{W*} = \begin{cases} -1, & \text{if } (1-\lambda)\mathrm{E}[R^W - R^L] < -\lambda\mathrm{CVaR}[R^W - R^L] \\ 0, & \text{if } -\lambda\mathrm{CVaR}[R^W - R^L] \leq (1-\lambda)\mathrm{E}[R^W - R^L] \leq \lambda\mathrm{CVaR}[-(R^W - R^L)] \\ 1, & \text{if } (1-\lambda)\mathrm{E}[R^W - R^L] > \lambda\mathrm{CVaR}[-(R^W - R^L)] \end{cases}$$

According to this result, we should long winners and short losers, as in the cross-sectional momentum strategy, when the expected return of winners minus losers exceeds a positive multiple of the downside risk of losers minus winners; reverse the strategy to short winners and long losers when the expected winner-minus-loser return is less than a negative multiple of its downside risk; and invest only in the risk-free asset if the expected winner-minus-loser return falls between these extremes. Using the scenarios, $\mathrm{E}[R^W - R^L]$ is approximated by $\sum_{j=1}^{J} p_j(r_j^W - r_j^L)$ and $\mathrm{CVaR}[-(R^W - R^L)]$ is approximated by $\left(\eta + \frac{1}{1-\alpha}\left[-\sum_{j=1}^{J} p_j(r_j^W - r_j^L) - \eta\right]_+\right)$. Note that, according to the representation of CVaR in Pflug and Pichler (2016), when $\alpha_t = 1$, $\mathrm{CVaR}[-(R_t^W - R_t^L)] =: \operatorname{ess\,sup}[-(R_t^W - R_t^L)] \approx \max_{j=1,\cdots,J}\{r_{t,j}^L - r_{t,j}^W\}$.

4.3.2 Time-Varying Tail Probability and Risk-Aversion Parameter

Because momentum crashes happen in periods of high market volatility, we consider three volatility-related measures: 1) the volatility of market index returns, 2) the ratio of estimated market return to market volatility, and 3) the CBOE Volatility Index (VIX), to define the time-varying probability, α_t, and the risk-aversion parameter, λ_t.

We base our volatility estimate on a recent observation of market index returns. Guo and Ryan (2020a) find that S&P 500 stock market index returns can be well characterized by a normal distribution with a time-dependent mean estimated by the weighted moving average of the past P-period return and volatility estimated using the difference between the past G-period returns and the time-dependent mean. We apply the same approach to estimate the market return and volatility. Let i_t be the realized market return, which is the value-weighted return of all stocks listed in B, σ_t be the volatility of the stock market in period t, $u_t \equiv i_t - m_t$, and $\bar{u}_t \equiv \frac{1}{G}\sum_{k=1}^{k=G} u_{t-k}$. Then we have

$$i_t = \sum_{b \in B} \frac{S_t^b O_t^b}{\sum_{a \in B} S_t^a O_t^a} r_t^b, \tag{4.12}$$

$$m_t = \left[\frac{(1+P)P}{2}\right]^{-1} \sum_{k=1}^{P} (P-k+1) i_{t-k}, \tag{4.13}$$

and

$$\sigma_t = \sqrt{\frac{1}{G-1} \sum_{k=t-G}^{t-1} (u_k - \bar{u}_t)^2}. \tag{4.14}$$

When the stock market is highly volatile, the stock price becomes less predictable, and thus, investors should be more cautious and conservative to invest in the risky stock market. One famous example is the large increase in risk aversion after the global financial crisis in 2008 (Guiso et al., 2018), which is during the second momentum crash period. On the contrary, when the market is less volatile, the stock price is more predictable, and rational investors are presumed to have confidence in stock investment. Following this logic, we define α_t and λ_t in an expanding horizon to avoid look-ahead bias by the following steps:

1. Compute σ_k using (4.12)-(4.14) for $k = 1, \cdots, t$
2. Sort these t values in an increasing order and derive the rank of σ_t, denoted as A_t
3. Set $\alpha_t = \frac{A_t}{t}$ and $\lambda_t = \frac{A_t}{t}$.

When A_t is large, we put more weight on CVaR and more focus on the extreme worst cases (smaller tails).

In addition to the standard deviation of market returns, α_t and λ_t can be defined according to the ratio of the estimated market return to market volatility, which is a reward-risk ratio of the

stock market reminiscent of the Sharpe ratio. When this ratio is high, it indicates that the market exhibits a bullish trend and a profitable investment in the market is very likely. Thus, a higher reward-risk ratio should lead to lower risk aversion. To use this ratio to derive α_t and λ_t, we need to

1. Calculate $\frac{m_k}{\sigma_k}$ using (4.12)-(4.14) for $k = 1, \cdots, t$
2. Sort these t values in a decreasing order and derive the rank of $\frac{m_t}{\sigma_t}$, denoted as B_t
3. Set $\alpha_t = \frac{B_t}{t}$ and $\lambda_t = \frac{B_t}{t}$.

The last factor considered to automate the change in α_t and λ_t is VIX, which estimates the expected volatility by aggregating the weighted prices of S&P 500 index put and call options over a wide range of strike prices and has been commonly used as a proxy for market uncertainty (Whaley, 2009; Li et al., 2021). Li et al. (2021) demonstrate that VIX is a strong predictor for stock market regime shifts: A lower VIX implies a higher probability of a bull market regime with high returns and low volatility in the next period regardless of current market conditions. This observation indicates that lower VIX should be associated with less risk aversion. Hence, we take the following steps to specify α_t and λ_t.

1. Obtain VIX at each time $k = 1, \cdots, t$
2. Sort these t values in an increasing order and derive the rank of VIX at time t, denoted as C_t
3. Set $\alpha_t = \frac{C_t}{t}$ and $\lambda_t = \frac{C_t}{t}$.

4.4 Scenario Generation and Evaluation

In deterministic-equivalent stochastic programs, the uncertain parameters are characterized by scenarios, which discretize or approximate the stochastic process governing the underlying uncertainty so that the optimization model can be solvable. There are various approaches to generate scenarios but no universal choice for a scenario generation method. The selection of the scenario generation method depends on the features of the uncertain parameters, which are the returns of winner and loser portfolios in this paper. To generate scenarios for the winner and loser portfolio returns,

we employ a hybrid of heuristic and optimization moment-matching algorithms. After completing scenario generation, a backtesting procedure to evaluate the quality of scenarios is applied.

4.4.1 Optimization and Heuristic Moment Matching

Momentum crashes occur during times of high uncertainty about the returns of winner and loser portfolios, especially at some turning points, such as when the bear market ends and past losers stop being losers. Generating scenarios that are capable of capturing some inherent structure of the returns of these two momentum portfolios could be helpful to avoid momentum crashes. Although the joint distribution of these two uncertain parameters is unknown, the observation of occasional momentum crashes gives us a hint that this joint distribution has large kurtosis while the regular profitability of the cross-sectional momentum strategy indicates that winner and loser returns are dependent. These observations make the inclusion of kurtosis and correlation necessary in scenario generation, and motivate the use of the non-parametric moment-matching method, proposed by Høyland and Wallace (2001). By matching the first four marginal moments (mean, variance, skewness, and kurtosis) and correlation of scenarios to the corresponding sample estimates as closely as possible, this moment-matching optimization method generates scenarios for winner and loser returns without specifying any distributions for them.

Specifically, the moment-matching optimization problem for generating scenario sets for winner and loser portfolios at time t is:

$$\min_{\{r^W_{t,j}, r^L_{t,j}, p_{t,j}\}_{j=1}^{j=J}} \sum_{m \in M} v_m \left[f_m \left((r^W_{t,1}, r^L_{t,1}, p_{t,1}), \ldots, (r^W_{t,J}, r^L_{t,J}, p_{t,J}) \right) - M_{VALm,t} \right]^2 \tag{4.15a}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} p_{t,j} = 1 \tag{4.15b}$$

$$p_{t,j} \geq 0, j = 1, \ldots, J \tag{4.15c}$$

Scenarios of winner and loser portfolio returns, $r^W_{t,j}$ and $r^L_{t,j}$, and their corresponding probabilities, $p_{t,j}$, are constructed by matching the statistical properties of the approximating distribution, $f_m \left((r^W_{t,1}, r^L_{t,1}, p_{t,1}), \ldots, (r^W_{t,J}, r^L_{t,J}, p_{t,J}) \right)$, to the prespecified statistical properties, $M_{VALm,t}$. Here, $M_{VALm,t}$ is estimated using the historical data of winner and loser returns and v_m is a user-specified weight for statistical property m. The function to compute the statistic m of the input scenarios is denoted by f_m. For example, for winner portfolios, if $m = 1$ denotes mean, then $f_1 \equiv \sum_{j=1}^{J} p_{t,j} r^W_{t,j}$.

Similarly, if $m = 2$ represents variance, then f_2 for winner portfolios expands to $\sum_{j=1}^{J} p_{t,j}(r_{t,j}^{W})^2 - f_1^2$. To summarize, the objective (4.15a) minimizes the Euclidean distance between the statistical properties of the constructed distribution and the specifications, subject to constraint (4.15b) ensuring the probability of scenarios to sum up to one and constraint (4.15c) defining the probabilities to be nonnegative.

In this model, the success of the scenario generation relies heavily on the input parameters, $M_{VALm,t}$. Realistic estimates of $M_{VALm,t}$ should produce high-quality scenarios. To estimate the first four moments and the correlation of past winner and loser returns, we assume the returns of past winners/losers portfolio can be estimated by the past H-period returns of the stocks which constitute winners/losers in the last period. This assumption is supported by the cross-sectional momentum observation that stocks constituting winners at time $t-1$ will still be winners at time t. Considering the finding in Jegadeesh and Titman (1993) that cross-sectional momentum exists only when the look-back period is between 3 – 12 months, H should be chosen within the range of 3 and 12 months. Accordingly, the estimates of the first four moments for past winner portfolio returns are

$$M_{VAL1,t} = \frac{1}{H}\left(\sum_{b\in W_t} v_{t-1}^{b} r_{t-H}^{b} + \cdots + \sum_{b\in W_t} v_{t-1}^{b} r_{t-1}^{b}\right) \tag{4.16}$$

$$M_{VAL2,t} = \frac{1}{H}\sum_{i=1}^{H}\left(\sum_{b\in W_t} v_{t-1}^{b} r_{t-i}^{b} - M_{VAL1,t}\right)^2 \tag{4.17}$$

$$M_{VAL3,t} = \frac{\frac{1}{H}\sum_{i=1}^{H}\left(\sum_{b\in W_t} v_{t-1}^{b} r_{t-i}^{b} - M_{VAL1,t}\right)^3}{M_{VAL2,t}^{3/2}} \tag{4.18}$$

$$M_{VAL4,t} = \frac{\frac{1}{H}\sum_{i=1}^{H}\left(\sum_{b\in W_t} v_{t-1}^{b} r_{t-i}^{b} - M_{VAL1,t}\right)^4}{M_{VAL2,t}^{2}} \tag{4.19}$$

By replacing W_t in (4.16)-(4.19) by L_t, we can obtain the first four moments for past loser portfolio returns. The estimated correlation for past winner and loser portfolio returns is

$$M_{VAL5,t} = \frac{\sum_{i=1}^{H}\left(\sum_{b\in W_t} v_{t-1}^{b} r_{t-i}^{b} - M_{VAL1,t}\right)\sum_{i=1}^{H}\left(\sum_{b\in L_t} v_{t-1}^{b} r_{t-i}^{b} - M_{VAL1,t}\right)}{\sqrt{\sum_{i=1}^{H}\left(\sum_{b\in W_t} v_{t-1}^{b} r_{t-i}^{b} - M_{VAL1,t}\right)^2 \sum_{i=1}^{H}\left(\sum_{b\in L_t} v_{t-1}^{b} r_{t-i}^{b} - M_{VAL1,t}\right)^2}} \tag{4.20}$$

We have tested the use of simple moving average, weighted moving average, and exponential moving average, to estimate mean, and the naive approach, exponential weighted moving average, and

various shrinkage estimators, including the single-index shrinkage estimator (Ledoit and Wolf, 2003), the constant-correlation shrinkage estimator (Ledoit and Wolf, 2004a), the one-parameter shrinkage estimator (Ledoit and Wolf, 2004b), and the diagonal shrinkage estimator (Ledoit and Wolf, 2004b), to balance the trade-off between overfitting and underfitting of the covariance matrix estimation. In addition, the use of the composite winner or loser returns, agnostic to the changing composition of these portfolios, to estimate $M_{VALm,t}$ has been investigated (Guo and Ryan, 2021). We find that only the combination of using the simple moving average to estimate $M_{VAL1,t}$ and the naive approach to estimate $M_{VAL2,t}$ and $M_{VAL5,t}$, which are illustrated by (4.16), (4.17) and (4.20), respectively, avoids momentum crashes and provides robust performance in all market conditions. Therefore, in this paper, we only introduce this combination and document its performance.

To select the number of scenarios, J, Høyland and Wallace (2001) propose a rule of thumb, that the degree of freedom in (4.15a)-(4.15c) should be no smaller than the number of specifications, to avoid over and underspecifications while using the optimization moment-matching algorithm. In our problem, we have $3J$ decision variables, $\left\{r_{t,j}^W, r_{t,j}^L, p_{t,j}\right\}_{j=1}^{j=J}$, each of which contributes one degree of freedom. Constraint (4.15b) eliminates one degree of freedom. Hence, the degrees of freedom of (4.15a)-(4.15c) are $3J-1$. The number of specifications is 9 (2 times 4 marginal moments plus 1 correlation). The minimum J that allows for matching 9 statistical specifications is 4, which means at least 4 scenarios need to be generated at each rebalancing point.

Although the cardinality of scenario sets is controllable, solving this highly nonlinear program may end up with non-globally optimal solutions. To avoid such occurrences, a heuristic moment-matching algorithm, proposed by Høyland et al. (2003), is applied to obtain the initial values for the decision variables in (4.15a)-(4.15c). In this heuristic algorithm, scenarios of past winner and loser returns are generated in two phases. In the input phase, based on the prespecified statistical properties, $M_{VALm,t}$, the transformed moments can be calculated using the Cholesky decomposition. In the output phase, scenarios of past winner and loser returns are first generated individually to match their corresponding transformed marginal moments by cubic transformation, neglecting the pre-specified correlation. Then Cholesky decomposition, matrix transformation, and linear transformation are applied to achieve the correct correlations without changing the marginal moments.

4.4.2 Reliability of Scenarios

Upon generating scenario, an evaluation of their quality is needed. This step is critical for two reasons. First, this reliability test assists with selecting values of the parameters used to generate scenarios. Second, the quality of scenarios directly determines the quality of the final investment decision. In this paper, we apply mass transportation distance (MTD) rank histograms, developed and implemented as an R package by Sarı et al. (2016), to assess the reliability of the generated scenarios, which is defined as the similarity of distribution of the generation and observation (Hsu and Murphy, 1986). Sarı et al. (2016) points out that the presence of uniformity of the rank histograms is an indicator of the reliability of scenarios. More specifically, if the rank histogram is flat, it means the generated scenarios are reliable. The degree of the flatness (or uniformity) can be evaluated by the Cramér-von Mises W^2 statistic (Choulakian et al., 1994), which measures the distance between the discrete uniform distribution and the empirical distribution of the MTD ranks and can be calculated using the R package `dgof` (Arnold and Emerson, 2011).

The procedure for scenario generation, reliability evaluation and stochastic portfolio optimization is outlined in Figure 4.3. To avoid data leakage, ideally we should use the early portion of the dataset as the training data set and the rest as the test data set, conduct scenario evaluation and parameter estimation using the training data, and apply the fixed parameter values to the test data. However, considering that the first momentum crash happened in 1932, there would not be sufficient training data to identify a suitable value of H while including this momentum crash in the test data set.

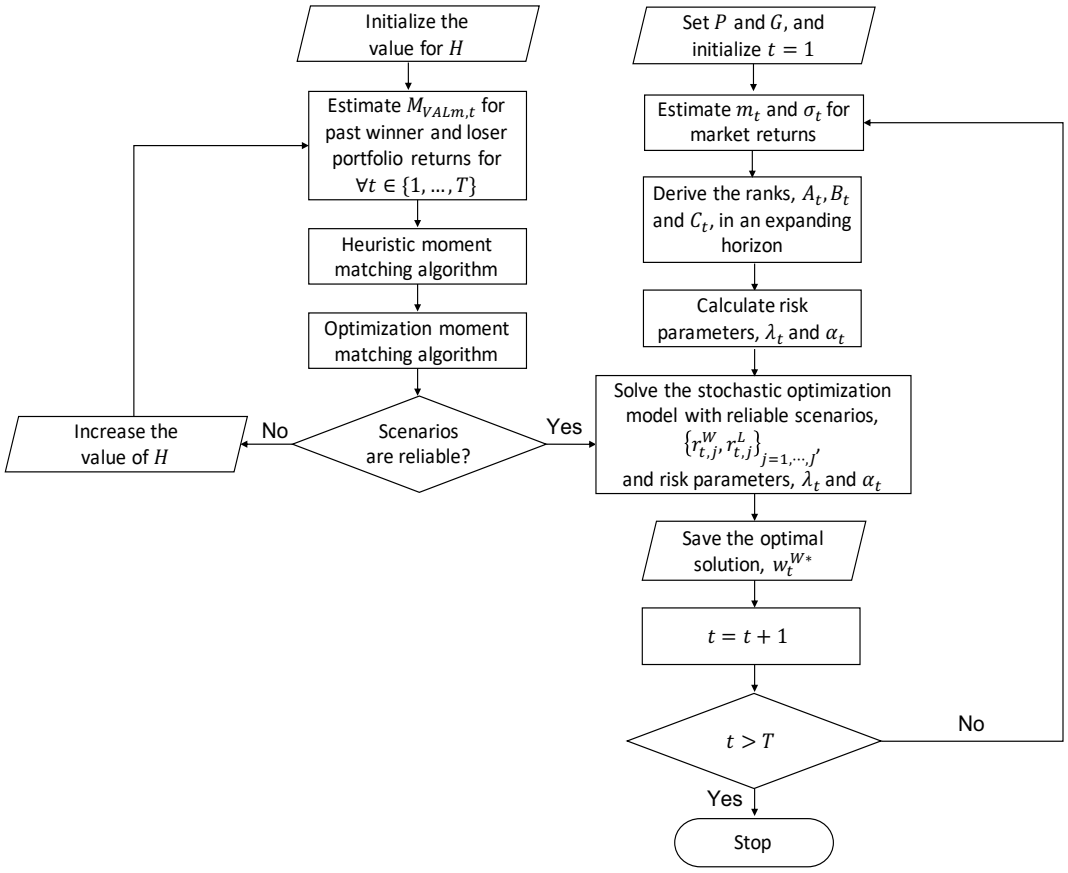


Figure 4.3: Flowchart for scenario generation, evaluation, and portfolio optimization

4.5 Performance Evaluation

Suitable performance measures can provide straightforward quantitative criteria to evaluate the trading strategy derived by the stochastic program, the mean-variance model in Daniel and Moskowitz (2016), and the cross-sectional momentum heuristic ($w_t^W \equiv 1$, $w_t^L \equiv -1$, and $w_t^f \equiv 1$), especially but not exclusively in market downturns.

Daniel and Moskowitz (2016) avoid momentum crashes using a variant of the Markowitz mean-variance model. They build an optimization model with the objective to maximize the Sharpe ratio of the portfolios, constructed by the winner minus loser portfolio and a risk-free asset, for the whole simulation horizon. This optimization model is equivalent to maximizing the long-term expected excess return of the winner minus loser portfolio subject to a constraint on return variance over the simulation horizon. They derive that the optimal weight on the past winner portfolio is

$w_t^{W*} = \frac{\mu_t}{2\gamma\delta_t^2}$, where in period t, μ_t is the expectation of the difference between the past winner and past loser returns, in excess of the risk-free rate, and δ_t^2 is the variance of the difference between the past winner and past loser returns. The parameter γ is chosen so that the in-sample annualized volatility, $\sqrt{\frac{N}{4\gamma^2 T}\sum_{t=1}^{t=T}\frac{\mu_t^2}{\delta_t^2}}$, is 19%. Daniel and Moskowitz (2016) use a time series regression model to estimate μ_t and a GJR-GARCH model to forecast δ_t^2. As their strategy does not impose a limit on the proportion of wealth invested in the past winner and loser portfolios, the net short position of the past loser portfolios in their strategy has reached a maximum of 5.37 in November 1952, which would incur a much higher transaction cost than the cross-sectional momentum heuristic and our mean-CVaR model. To allow a valid comparison between strategies, we let μ_t be the difference between $M_{VALmean,t}$ for past winners and $M_{VALmean,t}$ for past losers, subtracting r_t^f, δ_t^2 be the variance of the difference between past winner and past loser returns in the past H periods, and impose the constraint $-1 \leq w_t^{W*} \leq 1$.

To evaluate these three types of strategies, multiple annualized performance measures, including the excess return, maximum drawdown, Sharpe ratio, upside potential ratio, and time series cumulative returns are considered.

A backtesting simulation is conducted to obtain these performance measures. Let w_t^{W*} be the optimal weight for the winner portfolio in period t, and r_t^W be the observed return for the winner portfolio and r_t^L be the observed return for the loser portfolio in period t. Then we can derive the realized return for the constructed portfolio in period t, r_t, for $t = 1, \ldots, T$:

$$r_t = w_t^{W*}(r_t^W - r_t^L) + r_t^f \tag{4.21}$$

Annualized excess returns and annualized volatility are two basic measures of the overall gain and risk. They are calculated as

$$EX = \frac{N}{T}\sum_{t=1}^{T}(r_t - r_t^f) \tag{4.22}$$

and

$$VO = \sqrt{\frac{N}{T-1}\sum_{t=1}^{T}\left(r_t - \frac{\sum_{i=1}^{T} r_i}{T}\right)^2} \tag{4.23}$$

respectively.

However, volatility suffers from the weakness of treating upside potential and downside risk equally. This implies that using volatility as risk guidance may abandon a seemingly high risk

strategy which actually has high returns. Maximum drawdown (Young, 1991), the maximum observed loss from a peak to a trough of a portfolio before a new peak is attained, can overcome the drawback of volatility by only studying the downside risk and better detect the occurrence of momentum crashes. We calculate it using the `maxDrawdown` function in the PerformanceAnalytics R package (Peterson et al., 2014).

All the performance measures mentioned above focus on either risk or return but not both. They cannot tell investors whether the risk is worth the reward. The Sharpe ratio (Sharpe, 1966) is one of the most widely-used risk-adjusted metrics which takes account of both investment's profit and the degree of risk that is taken to achieve it. Annualized Sharpe ratio can be calculated as

$$SR = \frac{EX}{VO}. \tag{4.24}$$

Similar to the disadvantage of using volatility as the risk measure, Sharpe ratio is not able to differentiate between the upside reward and the downside risk. Compared to the Sharpe ratio, the upside potential ratio (UP ratio), proposed by Sortino and Van Der Meer (1991), is a more appropriate risk-adjusted performance measure, which measures the upside performance of portfolio returns for each unit of downside risk. Annualized UP ratio can be calculated as

$$UP = \frac{\frac{N}{T}\sum_{t=1}^{T}\max(0, r_t - r_t^f)}{\sqrt{\frac{N}{T}\sum_{t=1}^{T}\left(\max(0, r_t^f - r_t)\right)^2}} \tag{4.25}$$

Last but not least, to investigate whether our proposed mean-CVaR optimization model with time-varying risk aversions can avoid momentum crashes and make profits in all market conditions, the time series cumulative absolute return is plotted. The cumulative absolute return up to time t is

$$CA_t = \Pi_{i=1}^{i=t}(1 + r_i). \tag{4.26}$$

4.6 Numerical Results

Our data source is the Center for Research in Security Prices (`www.crsp.com`), where all monthly returns and prices of stocks listed on NYSE, AMEX, and NASDAQ with the returns of the common shares (with CRSP sharecode of 10 or 11) and VIX can be found. Following Daniel and Moskowitz (2016), the look-back period $K = 12$, and the holding period $H = 1$. The simulation horizon for

the scenario generation process covers the end of each month from January 31, 1927, to December 31, 2020. VIX is available only from Feb 1, 1990. The number of generated scenarios, J, is set to be 10 for each instance. It is larger than the minimum J so that overspecifications can be avoided and differences in the value of α_t are meaningful, but small enough to avoid underspecification. The user-specified weights, v_m, for all statistics are set to be $1/9$. The values of pre-specified statistical properties, $M_{VALm,t}$, are calculated using the preceding 3, 6, 9 and 12-month returns on a rolling basis, meaning $H \in \{3, 6, 9, 12\}$. The number of months used to estimate the mean of index returns, P, and the number of months used to estimate the market volatility, G, are both selected from {3, 6, 9, 12}. All the statistical simulations and analyses are conducted in R.

Table 4.1 displays the value of the W^2 statistic and its p-value in the reliability test for three types of scenarios generated using different numbers of months to calculate each $M_{VALm,t}$. This table indicates that using past 12-month data to calculate $M_{VALm,t}$ provides the smallest value of W^2 for the scenarios of the winner and loser returns, and thus, the best reliability for these two types of scenarios. By comparing the value of W^2 for winners and losers, we can see that the W^2 value of winners is smaller than that of losers for all tested values, which is consistent with the observation in Daniel and Moskowitz (2016) that the trend of loser portfolio returns is harder to predict and the momentum crashes are mostly due to such unforeseeable behavior of loser portfolios. Figure 4.4 displays the empirical density plots of returns for the winner and loser returns during the whole simulation horizon. The large excess kurtosis value of the loser returns also verifies that the loser has extreme returns occasionally, making its scenario generation less reliable. In terms of the reliability for the joint scenarios, using the past 9-month returns to calculate $M_{VALm,t}$ shows the best performance and is the only one passing the reliability test. By examination of this table, we find using $H = 9$ provides not only the best reliability for joint scenarios but also a decent performance for the scenarios of the winner and loser returns.

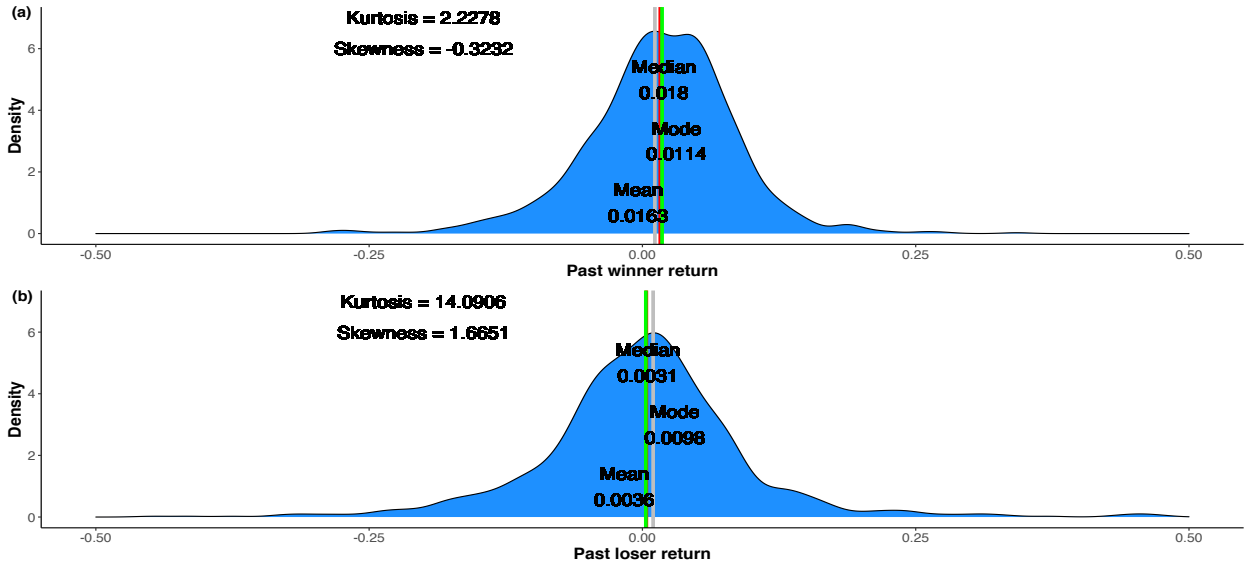


Figure 4.4: Empirical density plot of returns for the past winner and loser portfolios with mean, median, mode, skewness, and excess kurtosis shown during the whole simulation horizon.

H = Number of months used to calculate each $M_{VALm,t}$	3	6	9	12
Joint scenarios	5.5357	2.5632	0.2980	1.3193
	(0.0000)	(0.0000)	(0.1371)	(0.0003)
Scenarios of past winner returns	4.3766	1.2742	0.1418	0.0888
	(0.0000)	(0.0003)	(0.4124)	(0.6330)
Scenarios of past loser returns	7.6213	2.2046	0.6347	0.1464
	(0.0000)	(0.0002)	(0.0186)	(0.3979)

Table 4.1: W^2 statistic for the reliability of scenarios generated using different number of months to calculate each $M_{VALm,t}$. The p-values computed by the `dgof` package are shown in parentheses. Note that the simulation rather than the default approximation had to be used for at least one value in the table.

Figure 4.5 displays the generated scenarios for the joint winner and loser, winner marginal, and loser marginal returns when $H = 9$. All three subplots appear nearly flat, which further confirms that using past 9-month data to estimate $M_{VALm,t}$ in this hybrid moment-matching algorithm

produces scenarios that well represent the underlying uncertainty. Most importantly, the flatness of these plots promotes confidence that, upon employing these reliable scenarios in the stochastic portfolio optimization model on a rolling basis, asset allocation decisions derived from the solutions may avoid momentum crashes. Therefore, $H = 9$ is fixed for the following analysis.

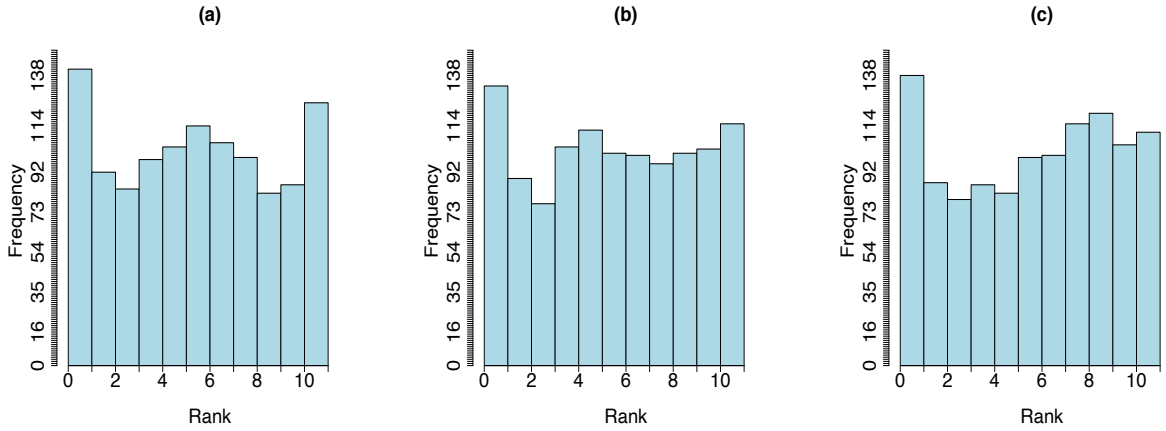


Figure 4.5: MTD rank histogram for (a) joint past winner and loser scenarios, (b) scenarios of past winner returns, and (c) scenarios of past loser returns using past 9-month data to calculate $M_{VALm,t}$ on a rolling basis and scenario sets of cardinality 10.

Table 4.2 presents a performance comparison among the baseline strategy (cross-sectional momentum heuristic), the mean-variance optimization model, and mean-CVaR optimization models with time-varying risk aversion defined by estimated market volatility in terms of annualized excess return, maximum drawdown, Sharpe ratio, and UP ratio. Different numbers of months to estimate the mean ($P \in \{3, 6, 9, 12\}$) and to estimate the volatility of market index returns ($G \in \{3, 6, 9, 12\}$) are considered for adjusting α_t and λ_t. This table indicates that the performance of the mean-CVaR models dominates the cross-sectional momentum heuristic and the mean-variance model. No matter which candidate value of P and G is selected, mean-CVaR always provides a higher excess return, Sharpe ratio, and UP ratio, as well as a smaller maximum drawdown, than the cross-sectional momentum heuristic or the mean-variance model. When $P = 6$ and $G = 3$, mean-CVaR offers the highest excess return, Sharpe ratio, and UP ratio, as well as the smallest maximum drawdown. Furthermore, the difference in the performance of all mean-CVaR models is trivial, which indicates

that the mean-CVaR model with volatility-adjusted risk aversion is robust to the values of P and G.

Table 4.3 shows the performance for mean-CVaR with time-varying risk aversions defined using the reward-risk ratio. From Table 4.3, we can see that using the reward-risk ratio to automate risk aversions underperforms the mean-variance model and the cross-sectional momentum heuristic for all tested numbers of months used to estimate the market return and volatility. The main reason for this poor performance is that estimating mean is much more difficult than estimating volatility due to the strong blurring effect of the mean. Luenberger (1997) illustrates that, when using a short period of historical data to estimate the mean, the standard deviation of the mean estimate is large, suggesting that the mean estimation is not reliable. For a reliable estimation of the mean, we need around 156 years' worth of monthly data, as shown in Luenberger (1997). However, the mean values of market returns are not likely to be constant over such a long time interval. The conflict between the number of periods used to estimate mean and data dependency makes it impossible to obtain an accurate estimate of mean using historical data.

Table 4.4 displays the comparison among the mean-CVaR model with time-varying risk aversion defined using VIX, the mean-variance, and the cross-sectional momentum heuristic. Because the VIX originated in February, 1990, Table 4.4 only shows the results after Feb 1, 1990, meaning that the performance of both strategies during the first momentum crash is not included in this table. From this table, it can be observed that mean-CVaR with risk parameters defined using VIX outperforms the cross-sectional momentum heuristic and mean-variance in terms of Sharpe ratio and UP ratio. What is more, the small maximum drawdown indicates that the second momentum crash is likely to be circumvented. However, the annualized excess return of the mean-CVaR strategy is less than that of the cross-sectional momentum heuristic. From all these observations, we can conclude that using VIX to automate risk parameters controls the investment loss at the cost of being less profitable than using volatility.

Annualized performance	Excess return	Maximum drawdown	Sharpe ratio	UP ratio
$P = 6$, $G = 3$	0.1669	0.5118	0.7757	2.5176
$P = 12$, $G = 3$	0.1657	0.5118	0.7684	2.5026
$P = 9$, $G = 3$	0.1650	0.5118	0.7637	2.4886
$P = 3$, $G = 3$	0.1635	0.5118	0.7585	2.4847
$P = 3$, $G = 9$	0.1643	0.5118	0.7754	2.4610
$P = 6$, $G = 12$	0.1642	0.5118	0.7805	2.4505
$P = 9$, $G = 9$	0.1606	0.5118	0.7626	2.4488
$P = 6$, $G =9$	0.1610	0.5118	0.7649	2.4486
$P = 9$, $G = 12$	0.1606	0.5248	0.7586	2.4278
$P = 12$, $G = 9$	0.1568	0.5118	0.7454	2.4213
$P = 12$, $G = 12$	0.1592	0.5248	0.7520	2.4195
$P = 6$, $G = 6$	0.1615	0.6325	0.7596	2.4142
$P = 3$, $G = 6$	0.1591	0.6325	0.7522	2.3959
$P = 9$, $G = 6$	0.1570	0.6325	0.7387	2.3898
$P = 3$, $G = 12$	0.1601	0.5118	0.7492	2.3872
$P = 12$, $G = 6$	0.1554	0.6325	0.7332	2.3806
Mean-variance	0.1109	0.5118	0.6349	2.1496
Cross-sectional momentum heuristic	0.1512	0.9674	0.5076	1.8855

Table 4.2: Annualized performance comparison among the baseline strategy (cross-sectional momentum heuristic), the mean-variance optimization model, and the mean-CVaR optimization models with time-varying risk aversions defined by estimated market volatility when $H = 9$. This table is sorted by the descending UP ratio.

Annualized performance	Excess return	Maximum drawdown	Sharpe ratio	UP ratio
Mean-variance	0.1109	0.5118	0.6349	2.1496
Cross-sectional momentum heuristic	0.1512	0.9674	0.5076	1.8855
$P = 12$, $G = 12$	0.1163	0.8609	0.4953	1.8696
$P = 12$, $G = 6$	0.1173	0.8197	0.4956	1.8628
$P = 12$, $G = 9$	0.1177	0.8284	0.5000	1.8601
$P = 9$, $G = 9$	0.1094	0.9165	0.4539	1.8141
$P = 9$, $G = 12$	0.1104	0.9053	0.4572	1.8140
$P = 12$, $G = 3$	0.1087	0.8819	0.4547	1.8124
$P = 9$, $G = 6$	0.1098	0.9132	0.4562	1.8119
$P = 9$, $G = 3$	0.1098	0.9755	0.4304	1.6625
$P = 6$, $G = 3$	0.1024	0.9840	0.3990	1.6303
$P = 6$, $G = 6$	0.0944	0.9867	0.3575	1.5574
$P = 6$, $G = 12$	0.0912	0.9880	0.3425	1.5468
$P = 6$, $G =$9	0.0903	0.9894	0.3398	1.5408
$P = 3$, $G = 9$	0.0774	0.9909	0.2855	1.4707
$P = 3$, $G = 3$	0.0764	0.9916	0.2821	1.4689
$P = 3$, $G = 6$	0.0741	0.9938	0.2713	1.4618
$P = 3$, $G = 12$	0.0708	0.9938	0.2597	1.4457

Table 4.3: Annualized performance comparison among the baseline strategy (cross-sectional momentum heuristic), the mean-variance optimization model, and the mean-CVaR optimization models with time-varying risk aversions defined by the reward-risk ratio when $H = 9$. This table is sorted by the descending UP ratio.

Annualized performance	Excess return	Maximum drawdown	Sharpe ratio	UP ratio
Mean-CVaR	0.1224	0.5118	0.4703	1.9306
Cross-sectional momentum heuristic	0.1362	0.8403	0.3899	1.8830
Mean-variance	0.0808	0.5118	0.4081	1.7498

Table 4.4: Annualized performance comparison among the baseline strategy (cross-sectional momentum heuristic), the mean-variance optimization model, and the mean-CVaR optimization models with time-varying risk aversions defined by VIX when $H = 9$. This table is sorted by the descending UP ratio. Since VIX is available after Feb 1, 1990, this table only shows the performance comparison after Feb 1, 1990, and does not account for the first momentum crash.

To check whether our mean-CVaR model with time-dependent risk aversion defined using estimated market volatility is profitable in various market conditions and avoids two momentum crashes, Figure 4.6 plots the time-series cumulative returns of the worst-performing ($P = 6, G = 3$) and the best-performing ($P = 12, G = 6$) mean-CVaR models during the whole simulation horizon in subplot (a), the first momentum crash in (b), and the second momentum crash in (c). Both mean-CVaR models consistently produce higher cumulative returns than either the cross-sectional momentum strategy or the mean-variance model. One notable exception occurs during the Great Depression period (August 1929 – March 1933), but the slopes of mean-CVaR returns are still positive, implying mean-CVaR still would have made a profit during the Great Depression. What is more, in both mean-CVaR models, there is no drastic decline during 1926-2020. Note that Daniel and Moskowitz report much higher returns for the mean-variance strategy but they do not enforce the weight constraint for winner and loser portfolios. Very interestingly, we notice that despite the steep decline in the stock market due to the COVID-19 in 2020, there is no momentum crash in 2020 for any strategy tested, which might be explained by government stimulus spending. For the two momentum crash periods, different from the dramatic sustained drawdown in the cross-sectional momentum heuristic, the cumulative returns of both mean-CVaR models are almost continuously above one, which means no great loss will be incurred during these two periods. Hence, our mean-CVaR model with time-dependent risk aversion successfully avoids momentum crashes. Note that the flat segments of the cumulative reward curves for the mean-CVaR optimization occur when it

places zero weight on the winner-minus-loser portfolio. No optimal weights of -1 are observed in this backtest.

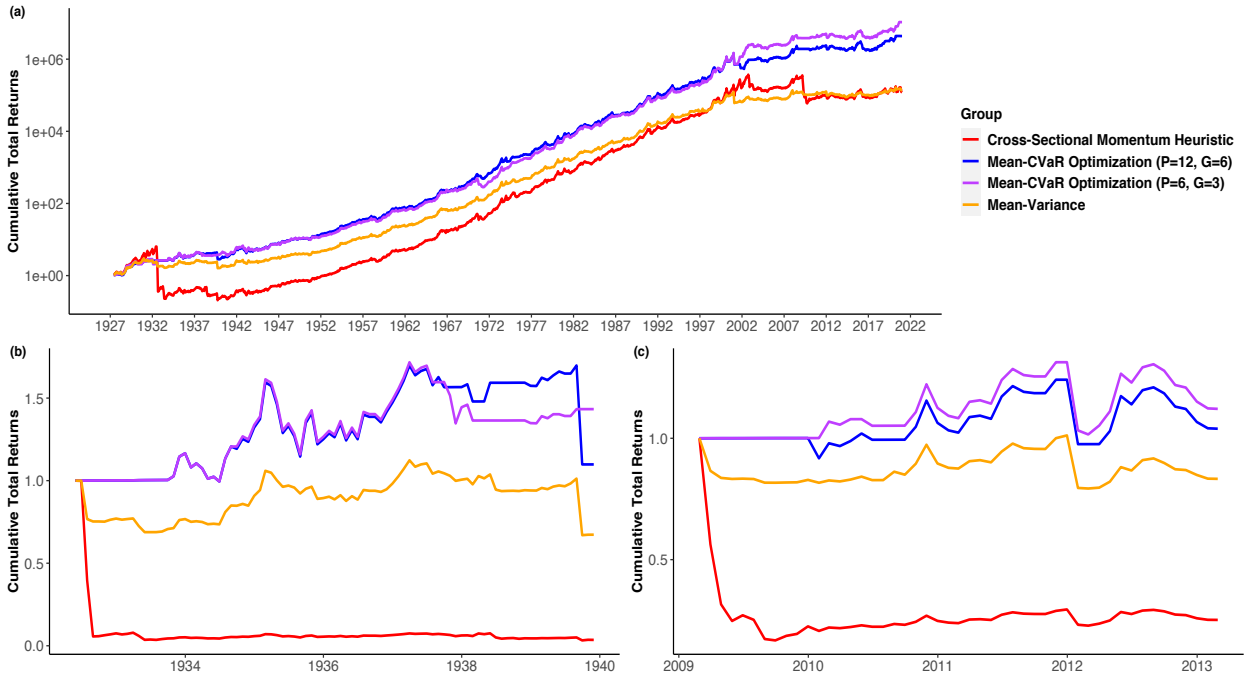


Figure 4.6: Time series cumulative returns of the cross-sectional momentum strategy, the mean-variance optimization model, and the mean-CVaR optimization model with time-varying risk aversions defined using the estimated market volatility during (a) the whole simulation horizon, (b) June 1932 - December 1939, the first momentum crash, following the Great Depression, and (c) March 2009 - March 2013, the second momentum crash, following the 2008-2009 financial crisis, when $H = 9$.

4.7 Conclusion

The cross-sectional momentum strategy exhibits strong and consistent performance across many diverse asset classes and periods. However, the severe and unpredictable momentum crashes make this strategy unfavorable by many investors. In this paper, we develop a stochastic mean-CVaR optimization model with time-varying risk aversion, which successfully avoids momentum crashes. Furthermore, the simple structure of the optimal trading solution to our mean-CVaR model is an added value for its usage in the industry due to the efficient computational effort.

In this strategy, we deal with the uncertainty of the stock market that causes momentum crashes by a combination of three solutions. First, as momentum crashes are infrequent but costly, and thus can be regarded as extreme events, we utilize a downside risk measure, CVaR, to control the expected loss under the extreme case. Second, momentum crashes happen in panic states when the stock market is volatile. Therefore, we adopt time-dependent risk aversion, calculated based on estimated market volatility, which outperforms the estimated reward-risk ratio or volatility index, to flexibly and automatically adjust the investment strategy according to the changing market condition. Third, the return of momentum portfolios (past losers and past winners) is uncertain when making investment decisions. However, the existence of momentum crashes is a signal that the distribution of momentum portfolio returns has large kurtosis and the remarkable performance of the cross-sectional momentum strategy indicates momentum portfolios are highly correlated. To capture the inherent structure of the momentum portfolio returns, we apply a hybrid of heuristic and optimization moment-matching algorithms and generate scenarios that well represent the underlying uncertainty. The reliability of the generated scenarios is evaluated based upon Cramér-von Mises hypothesis testing for the uniformity of the MTD rank histogram.

By carrying out computational studies over the last 94 years on common shares of all stocks listed on NYSE, AMEX, and NASDAQ, we find that reliable scenarios for the joint returns of the winners and losers, the winners only, and the losers only are generated when using nine months' worth of past data to estimate the parameters in this scenario generation approach. Upon employing these reliable scenarios in the mean-CVaR model with time-dependent risk aversion parameters defined using estimated market volatility, the rebalanced portfolio shows performance that is robust to the number of periods for estimating market volatility and achieves higher excess returns, Sharpe ratio, and upside potential ratio, as well as lower maximum drawdown, than either the cross-sectional momentum heuristic or the mean-variance model. We observe no drastic decline in the time series cumulative returns throughout the whole simulation horizon. Particularly during the two momentum crash periods, the cumulative returns of mean-CVaR are maintained around the level of one, indicating no huge loss is incurred when using this mean-CVaR model and the momentum crashes are avoided.

For future research, we plan to explore various approaches to automate risk aversion. For example, empirical research shows that clients' risk aversion changes in response to shocks to client

demographics (Barsky et al., 1997; Guiso and Paiella, 2008) and market returns and economic conditions (Bucciol and Miniaci, 2018; Guiso et al., 2018). Therefore, in addition to market volatility, reward-risk ratio, and VIX, we can consider other signals to automate the change in risk aversion settings. Moreover, robo-advisors, which collect information from clients about their financial situation, risk attitudes, and future goals through an online survey and then use the data to offer advice and automatically invest client assets, have emerged as an alternative to traditional financial advisors due to its high portfolio personalization, affordable portfolio management, and full automation for portfolio construction and rebalancing (Jung et al., 2018; D'Acunto et al., 2019; D'Acunto and Rossi, 2020). The current approach to incorporate risk aversion in robo-advisors involves human-machine interaction. Extending robo-advising algorithms to automatically adjust risk aversion appears promising.

4.8 References

Ahmadi-Javid, A. (2012). Entropic value-at-risk: A new coherent risk measure. *Journal of Optimization Theory and Applications*, 155(3):1105–1123.

Ahmadi-Javid, A. and Fallah-Tafti, M. (2019). Portfolio optimization with entropic value-at-risk. *European Journal of Operational Research*, 279(1):225–241.

Arnold, T. B. and Emerson, J. W. (2011). Nonparametric goodness-of-fit tests for discrete null distributions. *R Journal*, 3(2).

Asness, C. S., Frazzini, A., and Pedersen, L. H. (2012). Leverage aversion and risk parity. *Financial Analysts Journal*, 68(1):47–59.

Asness, C. S., Liew, J. M., and Stevens, R. L. (1997). Parallels between the cross-sectional predictability of stock and country returns. *Journal of Portfolio Management*, 23(3):79.

Bae, G. I., Kim, W. C., and Mulvey, J. M. (2014). Dynamic asset allocation for varied financial markets under regime switching framework. *European Journal of Operational Research*, 234(2):450–458.

Banihashemi, S. and Navidi, S. (2017). Portfolio performance evaluation in mean-CVaR framework: A comparison with non-parametric methods value at risk in Mean-VaR analysis. *Operations Research Perspectives*, 4:21–28.

Barro, D., Canestrelli, E., and Consigli, G. (2019). Volatility versus downside risk: performance protection in dynamic portfolio strategies. *Computational Management Science*, 16(3):433–479.

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

Barsky, R. B., Juster, F. T., Kimball, M. S., and Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics*, 112(2):537–579.

Benati, S. and Conde, E. (2021). A relative robust approach on expected returns with bounded CVaR for portfolio selection. *European Journal of Operational Research*. doi:10.1016/j.ejor.2021.04.038.

Bergk, K., Brandtner, M., and Kürsten, W. (2021). Portfolio selection with tail nonlinearly transformed risk measures—a comparison with mean-CVaR analysis. *Quantitative Finance*, pages 1–15.

Bucciol, A. and Miniaci, R. (2018). Financial risk propensity, business cycles and perceived risk exposure. *Oxford Bulletin of Economics and Statistics*, 80(1):160–183.

Chen, A. H., Fabozzi, F. J., and Huang, D. (2012). Portfolio revision under mean-variance and mean-CVaR with transaction costs. *Review of Quantitative Finance and Accounting*, 39(4):509–526.

Choulakian, V., Lockhart, R. A., and Stephens, M. A. (1994). Cramér-von Mises statistics for discrete distributions. *Canadian Journal of Statistics*, 22(1):125–137.

Cui, T., Bai, R., Ding, S., Parkes, A. J., Qu, R., He, F., and Li, J. (2020). A hybrid combinatorial approach to a two-stage stochastic portfolio optimization model with uncertain asset prices. *Soft Computing*, 24(4):2809–2831.

D'Acunto, F. and Rossi, A. G. (2020). Robo-advising. *Available at SSRN 3545554*.

Daniel, K. and Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, 122(2):221–247.

D'Acunto, F., Prabhala, N., and Rossi, A. G. (2019). The promises and pitfalls of robo-advising. *The Review of Financial Studies*, 32(5):1983–2020.

Grinblatt, M. and Titman, S. (1993). Performance measurement without benchmarks: An examination of mutual fund returns. *Journal of Business*, pages 47–68.

Grobys, K., Ruotsalainen, J., and Äijö, J. (2018). Risk-managed industry momentum and momentum crashes. *Quantitative Finance*, 18(10):1715–1733.

Guiso, L. and Paiella, M. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic Association*, 6(6):1109–1150.

Guiso, L., Sapienza, P., and Zingales, L. (2018). Time varying risk aversion. *Journal of Financial Economics*, 128(3):403–421.

Guo, X., Chan, R. H., Wong, W.-K., and Zhu, L. (2019). Mean–variance, mean–VaR, and mean–CVaR models for portfolio selection with background risk. *Risk Management*, 21(2):73–98.

Guo, X. and Ryan, S. (2020a). A time series momentum hybrid that controls downside risk. *Available at SSRN 3748126*.

Guo, X. and Ryan, S. (2021). Scenario generation for asset returns in a cross-sectional momentum strategy. *Proceedings of the 2021 IISE Annual Conference*.

Guo, X. and Ryan, S. M. (2020b). Reliability assessment of scenarios generated for stock index returns incorporating momentum. *International Journal of Financial & Economics*, 26(3):4013–4031.

Høyland, K., Kaut, M., and Wallace, S. W. (2003). A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24(2-3):169–185.

Høyland, K. and Wallace, S. W. (2001). Generating scenario trees for multistage decision problems. *Management Science*, 47(2):295–307.

Hsu, W.-r. and Murphy, A. H. (1986). The attributes diagram a geometrical framework for assessing the quality of probability forecasts. *International Journal of Forecasting*, 2(3):285–293.

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

Jung, D., Dorner, V., Weinhardt, C., and Pusmaz, H. (2018). Designing a robo-advisor for risk-averse, low-budget consumers. *Electronic Markets*, 28(3):367–380.

Kim, K. and Park, C. S. (2020). Pricing real options based on linear loss functions and conditional value at risk. *The Engineering Economist*, pages 1–24.

Kim, Y. S., Giacometti, R., Rachev, S. T., Fabozzi, F. J., and Mignacca, D. (2012). Measuring financial risk and portfolio optimization with a non-Gaussian multivariate model. *Annals of Operations Research*, 201(1):325–343.

Ledoit, O. and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10(5):603–621.

Ledoit, O. and Wolf, M. (2004a). Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management*, 30(4):110–119.

Ledoit, O. and Wolf, M. (2004b). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411.

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

Li, H., Wu, C., and Zhou, C. (2021). Time-varying risk aversion and dynamic portfolio allocation. *Operations Research*.

Lim, A. E., Shanthikumar, J. G., and Vahn, G.-Y. (2011). Conditional value-at-risk in portfolio optimization: Coherent but fragile. *Operations Research Letters*, 39(3):163–171.

Luenberger, D. G. (1997). Investment science. *Oxford University Press*.

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

Moazeni, S., Coleman, T. F., and Li, Y. (2016). Smoothing and parametric rules for stochastic mean-CVaR optimal execution strategy. *Annals of Operations Research*, 237(1-2):99–120.

Okunev, J. and White, D. (2003). Do momentum-based strategies still work in foreign currency markets? *Journal of Financial and Quantitative Analysis*, 38(2):425–447.

Peterson, B. G., Carl, P., Boudt, K., Bennett, R., Ulrich, J., and Zivot, E. (2014). PerformanceAnalytics: Econometric tools for performance and risk analysis. *R Package Version*, 1(3541):107.

Pflug, G. C. (2000). Some remarks on the value-at-risk and the conditional value-at-risk. In Uryasev, S., editor, *Probabilistic Constrained Optimization: Methodology and Applications*, pages 272–281. Springer.

Pflug, G. C. and Pichler, A. (2016). Time-inconsistent multistage stochastic programs: Martingale bounds. *European Journal of Operational Research*, 249(1):155–163.

Ripley, B. D. (1987). *Stochastic simulation*. John Wiley & Sons.

Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–42.

Rockafellar, R. T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471.

Sarı, D., Lee, Y., Ryan, S., and Woodruff, D. (2016). Statistical metrics for assessing the quality of wind power scenarios for stochastic unit commitment. *Wind Energy*, 19(5):873–893.

Sharifi, M. and Kwon, R. H. (2018). Performance-based contract design under cost uncertainty: A scenario-based bilevel programming approach. *The Engineering Economist*, 63(4):291–318.

Sharifi, M., Kwon, R. H., and Jardine, A. K. (2016). Valuation of performance-based contracts for capital equipment: A stochastic programming approach. *The Engineering Economist*, 61(1):1–22.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, 39(1):119–138.

Sortino, F. A. and Van Der Meer, R. (1991). Downside risk. *Journal of Portfolio Management*, 17(4):27.

Thomann, A. (2021). Multi-asset scenario building for trend-following trading strategies. *Annals of Operations Research*, 299(1-2):293–315.

Topaloglou, N., Vladimirou, H., and Zenios, S. A. (2011). Optimizing international portfolios with options and forwards. *Journal of Banking & Finance*, 35(12):3188–3201.

Whaley, R. E. (2009). Understanding the VIX. *The Journal of Portfolio Management*, 35(3):98–105.

Xiong, J. X. and Idzorek, T. M. (2011). The impact of skewness and fat tails on the asset allocation decision. *Financial Analysts Journal*, 67(2):23–35.

Young, T. W. (1991). Calmar ratio: A smoother tool. *Futures*, 20(1):40.

Yu, L.-Y., Ji, X.-D., and Wang, S.-Y. (2003). Stochastic programming models in financial optimization: A survey. *AMO - Advanced Modeling and Optimization*, 5(1):1–26.

Zou, J., Ahmed, S., and Sun, X. A. (2019). Stochastic dual dynamic integer programming. *Mathematical Programming*, 175(1-2):461–502.

CHAPTER 5. GENERAL CONCLUSION

Momentum is a prevalent market anomaly that can be used to identify the return trends of assets believed to persist into the future. Given the return predictability of momentum, TS and CS momentum trading strategies are constructed to guide an investor's selection of an investment portfolio. These two strategies are shown to be profitable across time, countries and numerous types of assets and asset classes. However, they both suffer from the weaknesses that uncertainty in stock market is ignored and the trading strategies are not optimized, both of which can be remedied by stochastic portfolio optimization. Therefore, in this dissertation, we aim at improving these momentum trading strategies using stochastic programming. The followings are the contributions of our work.

The first two papers combined focus on the possible improvement for the TS strategy. In the first paper, we propose three procedures to generate scenarios of the uncertain parameter, market index returns, which serve as input of the stochastic optimization model in the second paper. The main contributions in this paper were to:

- Develop a nonstationary variant of the geometric Brownian motion process incorporating an updated estimate of time series momentum to model the dynamics of index returns;
- Establish three scenario generation procedures, that differ in how frequently the time series momentum parameter is updated and whether it is estimated according to a simple moving average or an exponential moving average of returns;
- Assess the quality of scenario generation procedures by MTD rank histograms;
- Backtest in the S&P 500 and the FTSE 100 indices in two different time periods, and observe that
  - The reliability of scenarios is largely affected by the moving average technique used but not influenced by the frequency with which to update the parameters;

  - All procedures are able to produce reliable scenarios when the volatility is set appropriately;
  - The procedure using exponential moving average to estimate momentum appears to be more reliable in that it can provide reliable scenarios as long as the volatility estimation error is small.

The second paper continues the discussion of scenario generation in the first paper and extends it in several directions:

- Consider more moving average techniques to estimate time series momentum;
- Reduce the volatility estimation error mentioned in the first paper using a time-varying volatility, which is estimated on a rolling basis, in the geometric Brownian motion process of index returns;
- Construct two mean-risk stochastic optimization models, one of which uses variance as the risk measure and the other of which uses CVaR, solved on a rolling basis to sequentially obtain trading strategies that can balance the trade-off between maximizing the expected return and minimizing the investment risk;
- Develop a novel hybrid strategy, TSMDR, which combines the idea of risk parity from the TS strategy and controlling risk by CVaR, and find that TSMDR significantly outperforms the original TS strategy in terms of the Sharpe ratio and the Sortino ratio;
- Decompose all trading strategies into the trading signal and the position size to explore the components that can construct a quality trading strategy, and conclude that
  - Using weighted moving average can better capture the trend of the stock market than time series momentum;
  - Mean-risk strategies generally provide better returns whereas the risk parity model used in the TS and TSMDR strategies has less investment risk;
  - CVaR can better control the investment risk than variance.

The last paper endeavors to avoid the momentum crashes in the CS strategy. Momentum crashes caused by the uncertainty of the stock market in the recovery phase are the most challenging issue to

be solved when using the CS strategy. They occur when past losers in the recovery phase outperform past winners, and thus, to short past losers and long past winners will cause a great loss. By building a stochastic mean-CVaR optimization model with time-dependent risk parameters, we improve the CS strategy from the following perspectives.

- Use past winner and loser portfolios rather than individual stocks as two single units, in order to save computational resources and represent past winners and losers' asymmetric behaviors;
- Build a two-stage mean-CVaR stochastic program to balance the trade-off between maximizing expected investment returns and minimizing the investment loss;
- Extend the traditional mean-CVaR model by automating the change in the tail probability in CVaR and the risk-aversion coefficient according to three volatility-related measures: 1) estimated market volatility, 2) reward-risk ratio, and 3) volatility index, to flexibly adjust the momentum trading strategy according to the updated market condition;
- Conduct scenario generation for past winner and loser portfolio returns by an optimization moment-matching algorithm, where the initial values are derived from a heuristic moment-matching to avoid local optimal solutions, and evaluate the scenario reliability through MTD rank histograms;
- Develop a new momentum trading strategy that provides higher returns and less investment risk than the cross-sectional momentum heuristic and completely avoids momentum crashes.

For future research, there are multiple directions worth investigating. First, to be directly comparable with the benchmark strategy (the TS or the CS strategy), the transaction cost, such as the market impact cost, and tax effects are ignored in all three papers. To make our models more realistic and implementable, the incorporation of the trading costs can be considered. Second, in the third paper, we adopt a combination of heuristic and optimization moment-matching algorithms to generate scenarios. This combination requires solving a highly nonlinear program, which is not computationally efficient. Ponomareva et al. (2015) propose a moment-matching algorithm that can produce probabilistic scenarios that match exactly the specified mean, the covariance matrix, and the average of the marginal skewness and kurtosis without employing any optimization and considerably reduce the computational cost. Contreras et al. (2018) further improve this algorithm

by removing the spurious scenarios with negative probabilities. However, we find limitations when implementing their approaches in that neither algorithm is able to generate any scenarios when all entries of the specified covariance matrix are smaller than one. Thus, it would be very useful if we can follow their steps and push further. Lastly, due to the rapid development of technology, traditional wealth management has faced a big challenge in the form of robo-advisors. Typical robo-advisors collect risk profiles from clients and future goals through a survey and provide automated, customized, algorithm-driven financial planning services with very little cost and almost zero human supervision. Robo-advising has received a wealth of attention in the industry. Various financial institutions and banks, such as Betterment, Vanguard, and Wealthfront, have entered this space. In our third paper, we develop a model which can automate portfolio construction, rebalancing, and the change in risk aversion. It would be more interesting if we can acquire information about clients' risk attitudes and build a customized model that can optimally balance risk preferences from clients and the suggestions in our model.

5.1 References

Contreras, J. P., Bosch, P., and Herrera, M. (2018). Comment on "an algorithm for moment-matching scenario generation with application to financial portfolio optimization". *European Journal of Operational Research*, 269(3):1180–1184.

Ponomareva, K., Roman, D., and Date, P. (2015). An algorithm for moment-matching scenario generation with application to financial portfolio optimisation. *European Journal of Operational Research*, 240(3):678–687.