Dynamic Investment Strategy with Factor Models under Regime Switches

by

Takahiro Komatsu

Submitted to the Graduate School of Business Sciences
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

at the

University of Tsukuba

January 2016

レジーム・スイッチを考慮した
ファクターモデルによる動的投資戦略

筑波大学審査学位論文 (博士)

2016

小　松　高　広

筑波大学大学院

ビジネス科学研究科 企業科学専攻

Dynamic Investment Strategy with Factor Models under Regime Switches

by

Takahiro Komatsu

Submitted to the Graduate School of Business Sciences
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Abstract

The conventional models to simplify a return generating process of financial assets have not paid sufficient amount of attention to discontinuous behaviors represented by expected returns, volatilities and correlation. Despite of time series occasionally observed in the financial markets, another simplification assumes that an investor is not aware of information on the return generating process beyond one time step ahead. These simplifications mislead the investment strategy to build a suboptimal portfolio for the investor. In this work, we study the effects of a regime switching in the return generating process as well as in key parameters that comprise the investor's utility to maximize. For this purpose, we attempt to propose three layers of optimal portfolios all subject for a regime switching framework in the investment management, building on established literature: factor models for return prediction, dynamic investment over multiple periods, and portfolio optimization under investment constraints in practice.

First, we focus why the regime dependent assumption to the return generating process improves investment efficacy in optimal portfolios. By taking an example of popular equity investment strategies, i.e., sector momentum, we find that the conventional Gaussian assumption continues to work over enormously a long out-of-sample period to keep generating decent investment performance if the regime dependence is introduced to the models. For the utility function to maximize, a regime dependent risk aversion which an investor can be aware further improves the efficacy. As an empirical finding, specific to the sector momentum example, the momentum effect is revealed to be more significant in a sort of tranquil regimes while a reversal in a turbulent regime. We claim that these three findings contribute to the literature from both perspectives of pure practitioners and academia in the behavioral finance.

Second, as a theoretical core of this thesis, we achieve a tractable semi-analytical solution for multiple period investment horizons when the regime dependency is introduced to both of the return generating processes and the mean-variance utility function with an adjustment of the transaction cost. Traditionally, the dynamic investment over multi-period investment horizons makes sense for rigorous reasons to investors who are transaction cost conscious and not myopic but aware of the return generating process further ahead. In addition, the regime dependency enforces the dynamic investments because the discontinuous changes are optimally reflected into the portfolio over multiple periods effectively. The derived semi-analytical solution takes an intuitive form of a linear summation over a current portfolio and a target portfolio combined with a regime dependent weight. The target portfolio is a function of the factors to predict returns to the assets to hold. An empirical example is another popular equity investment strategy, i.e., style rotation, to testify the derived solution. Size and Book-to-Market portfolios predicted by SMB and HML are identified to be regime dependent. The more number in regime and/or the longer the optimization horizon, the better investment efficacy. Although the derived solution is achievable only in a constraint-free space, the tractable form delivers us meaningful insights to deploy it to practical applications. We emphasize that the theoretical contribution with the semi-analytical solution is quite valuable to advanced practitioners and academia in finance as well as operations research.

Finally we extend the derived semi-analytical solution to an optimization problem imposed by a constraint. In an attempt to argument the optimal solution to include not only hedge funds management to the long-only and/or asset allocation practices, we choose a short sales as the constraint to apply the solution. An earlier study reports that a problem formulation is achievable under such investment constraints as a short sales for a widely ranging class to include the linear combination, known as the Linear Rebalancing Rules. As such our finding in the constraint-free space, i.e., the linear combination of a current portfolio and a target portfolio as a function of the factors, theoretically encourages us to extend the Linear Rebalancing Rules to the regime dependent space. We show that the problem formulation and numerical experiments that prove reasonable basis to confirm the optimality of the numerically solved solutions under the framework of the second order cone problem.

We proudly overcome complexities on both of the regime dependency and the investment constraint incorporated into the multi-period optimization. The achievements in this thesis provide with opportunities to build significantly improved portfolios for a wide variety of investors ranging from a high level asset allocation to those in individual asset classes including hedge funds. Throughout the thesis for both of academia and investment professionals as practitioners in finance, we extend the latest horizon in knowledge for the communities.

Chief Thesis Advisor: Dr. Naoki Makimoto
Title: Professor

Thesis Advisor: Dr. Tadahiko Sato
Title: Professor

Thesis Advisor: Dr. Yuji Yamada
Title: Professor

Acknowledgments

I am greatly indebted to my chief thesis advisor Professor Naoki Makimoto for his irreplaceable support and guidance from start to the last minute of completing this thesis. His advices have always played a decisive role at various stages in my graduate study. I also thank my associate thesis advisors - Professor Tadahiko Sato and Professor Yuji Yamada - for their time and encouragements. Dr. Sato's suggestions have substantially improved Chapter 3, 4, and 5. Also I am also grateful for two anonymous referees. In addition, I wish to thank my colleagues and alumni at the Investment Management Division in the Goldman Sachs & Co. for their encouragements. Finally, I would like to dedicate this work to my wife, Yumi, my daughter, Miyabi and my son, So, who have supported and encouraged me throughout this program at the University of Tsukuba.

Contents

List of Tables

List of Figures

List of Symbols

$\boldsymbol{B}$	Transaction cost coefficient matrix
$\boldsymbol{f}$	Returns to factor to predict assets' returns
$I(t)$	Regime at time t
$I_t(t+1)$	Regime as of time $t+1$ predicted at time t
i	Suffix for a parameter in regime i
$\boldsymbol{L}$	Factor loading of assets
$\boldsymbol{P}$	Transition probability matrix
p_{ij}	(i,j) element in the transition probability matrix
R_f	Risk Free Rate
R_m	Market Return
$\boldsymbol{r}$	Returns to assets to invest
T	Time horizon
$\boldsymbol{u}$	Unredictable noise in return prediction to assets
$\boldsymbol{W}$	Covariance matrix of unredictable noise in return prediction to assets
$\boldsymbol{\epsilon}$	Unredictable noise in VAR(1) for returns to factors
λ	Risk aversion coefficient
$\boldsymbol{\mu}$	Drift term in VAR(1) for returns to factors
ρ	Discount rate
$\boldsymbol{\Sigma}$	Covariance matrix of unredictable noise in VAR(1) for returns to factors
$\boldsymbol{\Phi}$	Autoregressive coefficient in VAR(1) for returns to factors

Chapter 1

Introduction

Since Markowitz (1952) having inspired the investment management society to contemplate the exact science, studies on and practices of optimal portfolio have attracted notable efforts among academia and practitioners. Continuous challenges have contributed to evolve beyond the capitalization weight portfolio advocated by Sharpe (1964). Now literature of finance and investment largely converges that shapes of optimal portfolios depend upon information investors are aware and investment utility to maximize.

This thesis analyzes optimal portfolio investments by studying unique extensions to Hamilton (1989) and Merton (1971). First, a model specification to assume regime dependency is empirically examined in a well shared market savvy of industry momentum anomaly in equity markets. A next focus is put on analytical study to derive optimal solutions for infinite multi-period investment horizons to invest in regime dependent assets. Finally, we formulate the linear rebalancing rules to solve a finite horizon solutions under a short sales constraint.

Spearheading the thesis in Chapter 2, literature survey is conducted to relevant areas to subsequent chapters. As a common piece of models to study in the following chapters, we chose factor models to first area of survey. In particular, nature of time series is of emphasized focus. Motivating extensions to the factor models, studies on regime switching models are surveyed as a potential approach to specify discontinuous changes in return generating process of assets and the factors as predictors. Reaching at a last phase, the survey explores wide range of aspects of optimal portfolios, including regime switches and multi-period optimal solutions, in order to narrow down the scope of study.

Chapter 3 discusses in-depth analyses of a sector rotation investment strategy in the US equity under the regime switch framework for monthly data since 1927. It turns out that the momentum as a return predictor is more significant in tranquil regimes, that are largely separable around post-Oil Crisis early in 1980s, than a turbulent regime. Maximizing the mean-variance utility penalized by transaction costs out-of-sample basis since 1976, regime aware optimal portfolios perform decently. Moreover, regime dependent risk aversions further improve investment efficacy. It is revealed that multiple regime approaches reduce chances to mis-specify sector return generating processes that exhibit discontinuous behaviors.

In Chapter 4, we develop a model for dynamic investment strategy where assets' returns are predicted by multiple factors. In a mean-variance framework with factor models under regime switches, we derive a semi-analytic solution for the optimal portfolio with transaction costs. Due to the existence of transaction costs, the optimal portfolio is characterized as a linear combination of current and target portfolios, the latter of which maximizes the value function in the current regime. For some special cases of interest, simplified analytical solutions are also achieved. To see the effect of regime switches, the proposed model is applied to US equity market in which small minus big and high minus low are employed as factors. Investment strategy based on a proposed model demonstrates empirically that the regime switching models exhibit superior performance over the single regime model for all types performance measures evaluated, including realized utility and Sharpe ratio which are of particular interest in practice. Taking a close look at the time series of portfolio returns, the result shows the usefulness of the regime switching model as investors flexibly optimize asset allocations depending on the state of the market.

Chapter 5 arguments tractable solutions for dynamic investments to include complex portfolio set up. Since a seminal work of Merton (1971) academic endeavor and practical application continue to face scarcely encouraging consequences to derive analytical solutions. There also lies the difficulty to accommodate to wide range of models to specify return generating process of assets. This study introduces two complexities, 1) no short sales as one of popular portfolio constraints, and 2) regime switches in return generating process, in attempting to achieve optimal portfolios under multi-period return predictability by multiple common factors. We formulate linear rebalancing rules to solve under the second order cone constraints. A numerical experiment details that the formulated linear rebalancing rule leads to promising results that decent investment performance is reasonably achievable for investors even under the complexity both in the model and the practical investment constraint.

Studies presented in following chapters contribute to three areas of literature of finance and investments. First, Chapter 3 and 4 contribute to literature on equity market anomaly. Letting a momentum factor model specification be regime dependent, traditional wisdom of the momentum factor as a predictor is identified more promisingly in a tranquil state of assets' residual returns while a turbulent state does not necessarily contribute the momentum anomaly. In-depth analyses to understand the conditional efficacy of the anomaly reveal that, as an intuitive background to the improve, an introduction of regime dependency improves normality of residuals in return forecasting model, In out-of-sample basis over nearly forty years in monthly space. It is encouraging that three is sufficient as number of regimes to expect decent performance. The fact that twelve industries comprise the regime aware optimal portfolio encourages practitioners to apply the regime switching models to practically sizable investment problems. In addition to the industry momentum, Anomaly on Size and Value is also presented in weekly space. Similar to the industry momentum, Size and Value anomaly is more significant in tranquil regime than in turbulent regimes in both of two and three regime models.

Second, a theoretical achievement in Chapter 4 contributes to literature of areas of dynamic portfolio choice. Presence of regime dependency gives non trivial impacts upon to derive semi-analytical optimal solutions in discrete time space for an infinite time horizon. With extensions of regime dependency to classical models in the literature, it is revealed that, similar to regime independent models, the derived optimal solution takes a form of a linear combination of portfolio holdings as of the end of previous period and a function of realized returns to common factors as predictors for asset returns. Under a VAR(1) specification for the factors as asset return predictors, the literature is enriched by the fact that introduction of complexity by regime dependency leads to optimal solution as tractable as classical models.

Finally, Chapter 5 contributes to area of portfolio optimizations in finance literature. The notable achievement in the Chapter 4 justifies a validity of extending the linear rebalancing rules to apply the achievement to practical problems imposed by investment constraints as another complexity. A regime dependent formulation of an objective function to maximize adopts the second order cone problem which accommodates to practical investment constraints such as no short sales. Under a finite time horizon on a discrete time grid, numerical experiments demonstrate the efficacy of investments and practical feasibility of numerical computations.

Major contributions of this thesis are summarized as follows:

- Emphasized focuses on factor models that are widely accepted for advantages of dimension reductions and economic interpretation.
- Extensions of the asset return model predicted by the factors and the factor models to regime dependent space.
- Achievements of semi-analytical optimal solutions to derive for multi-period regime dependent investment solutions over infinite horizons under no investment constraint.
- Successful formulation and numerical solutions of a constraint bearing problem for the regime dependent asset return prediction by factors model over finite horizons.
- Superior investment performance measured in popular equity investment strategies as sector rotation and style rotation.

The thesis contributes to literature and investment managers to formally incorporate discontinuous changes in market behavior into investment decisions and manage portfolios for ranging from hedge funds and long only investments including asset allocation with forecasting factors.

Chapter 2

Literature Survey

2.1 Introduction

The main objective of research in portfolio management is to deliver superior investment performance under circumstances both when tail winds and when head winds blow. Markowitz's portfolio selection and Sharpe's CAPM are the major pillars for the investment society to have built well received and formal approaches to deliver efficient portfolios for investors. Because of tractable mathematical foundations to implement the uncovered theories, despite of unrealistic assumptions, the two pillars still play central roles in the most of investment practices.

By virtue of the seminal works, investment decisions have evolved over a half century in the past. Recent experiences in the financial markets, however, are obviously beyond the scope of the works. For instance, a recent decade has witnessed stochastic behaviors of financial instruments in the markets displaying discontinuous changes in key parameters. Although the changes have been specified by statistical models for quite long time, practitioners rather see structural changes behind the discontinuities. Hamilton (1989) opens eyes of not only econometricians but investors to consider regime switching to characterize the changes. This consideration gives meaningful impacts upon a process to convert investment views into investment actions to build optimal portfolios. Prior to literature reviews, a following set of questions arises for finding research paths in the survey:

- The factor models have established irreplaceable popularity due to accountability in the investment decisions and computation reality. How the factor models evolves beyond the static and linear specification? Non Gaussian and/or time variant behaviors in factor models generally identified across asset classes. Is any time series in the time variant behaviors predictable? What sorts of model to specify Non Gaussian and/or time variant behaviors studied?

- What sorts of approaches to find optimal portfolios under the Non Gaussian and/or time variant behaviors? If the time variant nature is predictable then myopic solution is not necessarily optimal but multi-period solutions. Is the the solution analytically derived and tractable? What if realistic constraints applied, e.g., no short sales?

Motivated by theoretical and empirical literature, this section surveys the progress and limitations to date in three related area: factor model, regime switch and portfolio optimization. Special attentions are paid to attempt to derive analytical solutions for optimal portfolios under realistic complexity in financial markets. While the attempt may not necessarily solve all of issues around optimal portfolios in the real markets, insights that analytical solutions tell investors are valuable to discuss investment processes and to understand consequent outcome for further improvements under the governance structure.

This chapter starts by reviewing how typical factors models, e.g., CAPM, have evolved to fill gaps between the theory and reality in Section 2.2. Section 2.3 covers the empirical works of regime switching. Section 2.4 explores approaches to optimal portfolios composed of assets that exhibit fat tail distributions and discusses multi-period optimizations. Section 2.5 provides suggestions for future research.

2.2 Factor Models

Spearheaded by Sharpe (1964) with the Capital Asset Pricing Model (CAPM) under the presence of the utility which an investor is aware of, factor models have shed light on structure and mechanism pricing securities in the financial markets. Fama and French (1992, 1993, 1996) extends the CAPM to add size and value to the market factor followed by Carhart (1997) to add momentum. Based upon no arbitrage opportunity, a development of the Arbitrage Pricing Theory (APT) extends the CAPM to find that a securities price is a linear combination of multiple factors. The BARRA delivers commercially successful factor models for individual assets and for asset allocation. In fixed income area, Nelson and Siegel (1987) specifies the term structure of interest rates with three factors.

Sharpe (1964) pioneers a factor model in the financial literature. Under an equilibrium condition, the Capital Asset Pricing Model assumes a single factor model which draws the Capital Market Line claiming that the expected return is a linear to the risk. Although the theory requires five conditions including those unrealistically achievable in the real world, together with Markowitz (1952), an investment practice of market capitalization weighted portfolio plays a central role in institutional investments around the world.

On CAPM and APT, massive amount of earlier studies is documented from a perspective of time variant behaviors of factors. Jagannathan and Wang (1996) and Lettau and Ludvigson (2001) show that conditional CAPM with a time-varying beta outperforms the unconditional CAPM with a constant beta. Akdeniz, Salih and Caner (2003) proposes the threshold CAPM which assumes nonlinear relationships between asset returns and the market return. Ross (1976) extends the capital pricing theory of Sharpe (1964) to an arbitrage model. Similar to the case with CAPM, Gonzalez-Rivera (1997) studies Ross (1976) and finds the time varying beta.

Basu and Stremme (2007) is one of examples that claim that the beta are allowed to vary as a function of lagged business cycle variables. A conditioning variable of the lagged business cycle appears in a linear predicting model in which both of the market portfolio return and the assets' returns are predicted by lagged business cycle variables shared in the model. As a function of observed macroeconomic and financial variables, modeling beta requires that econometricians are comfortable with the predictors to employ while investors are not necessarily the econometricians in general. Beyond the econometric approaches, economic explanations are enhanced in such other approaches as the agent model if researchers develop economic theories to test.

The Kalman filter approach is a second methodology which has been used to provide estimates of conditional risk. In the state space model, the CAPM β is specified in the state model while the observation model defines the CAPM market model. This model recursively estimates the beta series from an initial set of priors, generating a series of conditional alphas and betas in the CAPM market model. Black, Fraser and Power (1992) and Wells (1994) advocate the state space specification of the time varying beta estimated in the Kalman filter. Huang and Hueng (2008) runs the adoptive least square method for the Kalman filter and concludes that it outperforms a constant beta model specified by Pettengill, Sundaram, and Mathur (1995). Adrian and Franzoni (2009) counter-argues a criticism against the time varying beta. In an autoregressive model, the Kalman filter is justified to apply because of low frequency of variation of the time varying beta. The one-period-ahead forecast of the factor loading is a combination of its long-run behavior and the current estimate of the level of risk. Trecroci (2013) documents that Kalman filter gives much better results than OLS estimation of CAPM parameter estimated in rolling windows.

The multivariate GARCH approach to model time-varying beta uses the conditional variance information produced by the GARCH model to construct the conditional beta series. This approach has been utilized in various studies including Braun, Nelson and Sunier (1995), McClain, Humphreys and Boscan (1996), Gonzalez-Rivera (1997), and Brooks, Faff and McKenzie (1998). For example, Braun, Nelson and Sunier (1995) fits a bivariate EGARCH model to monthly US stock returns over the period July 1926 to December 1990. The EGARCH specification enables to formally test for the presence of predictive asymmetry in market and portfolio specific volatility as well as systematic risk or beta. The results indicate that whilst asymmetry is found to be present in the market component of returns, it was weak and/or inconsistent in the non-market components of risk and completely absent in the conditional beta. To understand the latest crisis in the financial markets, Hasnaoui and Fatnassi (2014) applies bivariate BEKK-GARCH model to create the time-varying betas in CAPM for the US 10 industrial sectors for a period of the subprime financial crisis. As an alternative to the GARCH approach to estimate conditional betas, Schwert and Seguin (1990) proposes a single factor model of heteroscedasticity found in stock returns and incorporates it into the market model equation in which its modified form provides estimates of time-varying market risk. Testing for monthly size-ranked US portfolios for over the sample period 1927-1986, it is turned out that the ability of previous studies to validate the CAPM model may be due to their failure to account for the heteroscedasticity in stock returns.

By introducing the Kernel regression for non-parametric estimation, unlike the vast majority of the literature, a nonparametric approach is adopted to test the time variant nature of factors by Ang and Kristensen (2012) in finding time dependency of alpha and beta in the Fama-French factor model. Although much is left for further research to understand what leads the alpha to deviate from zero, this suggests some dynamics drives the alpha and beta over time.

Reeves and Wu (2010) is one of limited examples of research claiming that constant beta estimated with daily returns over the last 12 months generates lower forecast error for quarterly beta forecast than those estimated under autoregressive time series analysis across for stock markets of United States, United Kingdom and Australia.

Some research tests if length of regression window matters or not. Shortening windows is one method as Lewellen and Nagel (2006) to estimate alpha and beta for size, BE/ME and momentum portfolios. A large unconditional CAPM alpha is not reasonably explained by it. Fama and Fench (2006) also tests the rolling window estimates in one year windows and confirms that size, value and momentum still exist.

A couple of earlier studies comprehensively compare several approaches discussed above. First, Brooks, Faff and McKenzie (1998) compares the three of them. In both of in-sample and out-of-sample basis, the Kalman filter works better than two other approaches in case of Australian equity sector returns. Nieto, Orbe and Zarraga (2014) also compares varying beta estimates taken from three different methodologies: least-square estimators including nonparametric weights, GARCH estimators and Kalman filter estimators. Taking the Mexican stock markets where beta dispersion is wide, the results is consistent with Brooks, Faff and McKenzie (1998) that Kalman filter estimators with random coefficients outperform the others in capturing both the time series of market risk and their cross-sectional relation with mean returns.

Chiarella, Dieci and He (2013) summarizes econometric approaches to the time varying beta by classifying earlier studies as follows:

- GARCH and Multivariate GARCH: Engle (1982), Bollerslev (1986, 1990)
- EGARCH: asymmetric and nonlinear effects of beta on conditional volatility of positive and negative shocks: Braun, Nelson and Sunier (1990)
- The random walk model: Fabozzi and Francis (1978) and Collins, Ledolter and Rayburn (1987)
- The mean-reverting model: Bos and Newbold (1984)
- The Markov switching models: Hamilton (1989)
- Ang and Chen (2007) treats betas as endogenous variables that vary slowly and continuously over time and finds that a single-factor model performs substantially better at explaining the book-to-market premium.

Summarizing this Subsection 2.2, Table 2.1 compares examples of model specifications across major approaches to CAPM β estimation.

Table 2.1: Examples of specifications across major CAPM β estimation

CAPM	Specification	β estimate
Unconditional[1]	$R_j - R_f = \alpha + \beta_j(R_m - R_f) + \epsilon_j$	$cov(R_j - R_f, R_m - R_f)/var(R_m - R_f)$
Kalman fiter for State Space[2]	$R_{j,t} - R_{f,t} = \alpha + \beta_{j,t}(R_{m,t} - R_{f,t}) + \xi_{j,t}$ $\xi_{j,t} \sim \mathcal{N}(0, \Omega_{j,t})$ $\beta_{j,t} = \beta_{j,t-1} + \nu_{j,t}$ $\nu_{j,t} \sim \mathcal{N}(0, Q_{j,t})$	$\beta_{j,t\|t-1}$ $+P_{j,t\|t-1} z_{j,t}[(R_{j,t} - R_{f,t}) - z_{j,t}^{\top}\beta_{j,t\|t-1}]$ $/z_{j,t}^{\top}P_{j,t\|t-1}z_{j,t}$ $+h_{j,t}$
Bivariate GARCH[3]	$R_{j,t} = \delta_j + \epsilon_{j,t}$ $R_{m,t} = \delta_m + \epsilon_{m,t}$ $\epsilon_t \mid \psi_{t-1} \sim \mathcal{N}(0, h_t)$ $h_{j,t} = c_{11} + a_{11}\epsilon_{j,t-1}^2 + b_{11}h_{j,t-1}$ $h_{j,m,t} = c_{21} + a_{22}\epsilon_{j,t-1}\epsilon_{m,t-1} + b_{22}h_{j,m,t-1}$ $h_{m,t} = c_{31} + a_{33}\epsilon_{m,t-1}^2 + b_{33}h_{m,t-1}$	$h_{j,m,t}/h_{m,t}$
Regime Switch	$R_{j,t} - R_{f,t} = \alpha_{I_t} + \beta_{j,I_t}(R_{m,t} - R_{f,t}) + \xi_{j,t,I_t}$ I_t is a Markov process on $\{1, \ldots, J\}$.	β_{j,I_t}

[1] See Sharpe (1964).
[2] See Black, Fraser and Power (1992).
[3] See McClaim, Humperes and Boscon (1996).

In sum, a literature survey on the factor models summarizes five representative items in model structure in a following way:

- OLS: When betas vary over time, the standard OLS inference is misspecified and cannot be used to assess the fit of a conditional CAPM. Sensitivity of estimated parameters to length of sampling windows raises a robustness question.
- Exogenous variables as predictors for beta, provided that economically rational and intuitive basis are assured.
- Multivariate GARCH across an asset and the market portfolio might be fatally complicated for advanced optimal portfolio solutions to be derived.
- State-Space model with Kalman filter: Empirical results report occasionally better performance than OLS, exogenous variables as predictors and GARCH. The algorithm may be viable only if behaviors of assets and factors are reasonably stable. An inherited Gaussian assumption may not necessarily be adequate for modeling factors and assets.
- Regime Switch: Not many earlier studies have explored yet. Although decent ability to specify discontinuous changes in factor behaviors presumably because not accountable for reasons and limited interpretation available.

2.3 Regime Switch

Based upon a mathematical cornerstones founded by Baum and Petrie (1966), extensive studied in the statistics and econometrics literature, e.g., Titterington, Smith and Markov (1985) and Hamilton (1994) deliver a Markov mixture of dynamic models that have attracted increasing interests. Since Hamilton (1990) introduces the E-M algorithm to estimate the regime switching models, a wide spectrum of interests is attracted including following domains tightly related to investment practices:

- Time series models
- Factor models
- Optimal portfolio selection

First, two theoretical studies on interaction of the Markov switching with underlying time series analyses include Timmermann (2000) which derives the moments for a range of Markov switching models that produce volatility, skewness and kurtosis as a function of the transition probabilities and parameters of the underlying regime probability densities entering the switching process. Also relationship between volatility clustering, regime switches and structural breaks in time-series models are well connected to autocovariance of time series generated by Markov switching processes. The success of a family of ARCH model in empirical studies is attributable to the fact that the model can generate autocorrelation occasionally seen in financial markets. Similar to this, it is revealed that the Markov switching models can also generate autocorrelation in the squares of a time series. Francq and Zakoïan (2001) considers multivariate ARMA models subject to Markov switching. Stationarity within each regime is theoretically examined. Derived stationarity indicates that local stationarity of the observed process is neither sufficient nor necessary to obtain the global stationarity with a few exception, e.g., VAR(1). Timmermann (2000) and Francq and Zakoïan (2001) give important insights on a selection of underlying models to specify. In order to avoid unwelcome interactions between the regime switching and underlying models, the literature indicates that one should pay careful attentions to assume regime switching over those models to capture the volatility clustering such as GARCH and other time-series models that handle structural breaks.

In the macro economics, Hamilton (1989) is one of seminal works pioneering that the regime switch model is proven to be workable in macro economic variables. A regime dependent autoregressive model reveals that the postwar US real GNP experiences periodic shifts of growth rate from positive to negative. The probabilistic inference on the shifts can be used as an objective criterion to define economic recessions. Kim and Nelson (1999) studies real GDP growth in post war period in the U.S. The regime switch model handles a structural break at an unknown change point. It reveals that a 1st quarter in 1984 is a break point of it. The classical linear model claims that a structural decline in the volatility of U.S. real GDP growth happens in a same point in time as 1984:Q1.

Stimulated by the Hamilton (1989), the regime switching model has been researched in not only economics but finance and investments. In equities and equity style investments, Schaller and Norden (1997) finds that the regime switch model replaces conventional price/dividend ratio to predict stock market return. The ratio rather adds value to the regime switch model once it appears in a model to forecast time variant probabilities of regime transitions. On equity style, Coggi and Manescu (2004) finds that an unconditional Fama-French model is quite poor in some periods when a regime switching model identifies two regime, one is characterized by a very high factor loading on the value risk factor. Intuitively the regime is identified as a financial crisis regime. Although no regime switching is specified, Arshanapalli, Fabozzi and Nelson (2006) shows that market risk and small cap premium behave more like risk factors while value premium does not. A momentum premium exists under different economically distressed scenarios in the tested period. This implies that the behavior of these premium under different macro economic scenarios is different across factors. The study implies potential presence of different mechanism to drive the equity factor returns from those handled in traditional linear models. Ma, MacLean, Xux and Zhao (2011) applies linear factors of MKT, SMB, HML, VIX, YS and CS to forecast US sector ETF returns. Not maximizing mean-variance utility but maximizing returns under constraints of risk factors as a linear programing, a three regime model, i.e., Bull, Bear and Transient, is identified as the best model under the BIC. Measured performance for 150 days without transaction cost shows decent results for the multiple regime solution. Tu (2010) maximizes the mean-variance utility function for investment universe including cash, MKT, SMB, HML and Fama-French 25 portfolios sorted on size and book-to-market.

Combining a data generating process,which is regime dependent, with uncertainty of an asset pricing model in portfolio decisions, results reveal the economic importance of regime switching under model uncertainty. The regime dependence is assumed to an expected return and volatility as well as correlation. Bekaert and Harvey (1995) allows a degree of integration of local markets into a global market to be time varying. A regime switching model to specify the time varying nature uncovers that coupling or de-coupling of each country to the global market is reasonably explained by the regimes.

Moving on to the fixed income space, Ang and Bekaert (2002a) documents a regime switching model for interest rates for US, UK and Germany. Incorporating the international short-rate and the term spread information, the regime switching models work better in forecasting the interest rates and fit to sample moments better than conventional models. Allowing the transition probabilities to be a function of exogenous variables, it finds the regimes in interest rates correspond reasonably well to business cycles at least in the US. With economical reasons to believe that the interest rate is subject for regime switches, e.g., economic expansion and recession, Bansal and Zhou (2002) extends the term-structure model proposed by Cox, Ingersoll and Ross (1985) and an affine specification by Dai and Singleton (2000). While the CIR and the multifactor version of it are sharply rejected to be accountable for the violation of the expectation hypothesis, a two factor regime switch model survives the tests.

In currencies, Engle (1994) reports that the regime switching model does not perform better than the random-walk model in foresting currencies in out-of-sample basis although workable in in-sample basis. The tested model simply specifies the exchange rate returns with expected return and volatility. A promising observation is concluded that the regime switching model seems to add value for a large swing in directions of exchange rates.

2.4 Optimal Portfolios

The mean-variance optimal portfolio is optimal for an investor who assumes the quadratic utility and markets follow to normal distributions. Optimizations under this problem falls under the quadratic problem in the Operations Research. However, recent experiences in financial markets are harmful for investors who implement the mean-variance approaches into practice. The investors should be reminded that the quadratic utility truncates higher order terms in the Taylor expansion of the exponential utility. If markets move extremely large, the quadratic utility may not reasonably represent a real shape of investor's utility any more. As such, the mean-variance optimal portfolio discounts so called tail risks that appear often in turbulent states in financial markets and/or crisis in real economy. More than a half century ago, Mandelbrot (1963) discusses the fat tail and applies Levy distribution and Cauchy distribution to specify the data generation process.

That said, for an investor's perspective, it is less and less practical than the Markowitz solution to solve an optimal portfolio solution under expected utility that entails the higher order terms. Followings are major examples that explore optimal solutions that do not require normal distributions for assets' returns and the rest in this sub-section surveys some of them:

- Independent Component Analysis (ICA)
- Non parametric approach
- CVaR (Expected Shortfall) and CDaR
- Extreme Value Theory and Copula
- Regime Switches

2.4.1 Independent Component Analysis (ICA)

Madan and Yen (2008) solves an optimal solution accepting higher moments of assets' returns in the portfolio in a parametric way. The optimal portfolio is solved under the exponential utility by taking advantage of the Independent Component Analysis (ICA) to orthogonalize assets without truncating 3rd or 4th moments. Handling the non-Gaussian multivariate time series, the ICA decomposes time series into statistically independent components. An analytical solution is derived under an assumption that the decomposed components follow the Variance-Gamma process. The derived optimal portfolio exhibits higher Sharpe ratios than those solved under an assumption of normal distributions.

Jondeau and Rockinger (2006) calculates numerically optimal several solutions across 2nd, 3rd and 4th moments to truncate the terms in the Taylor expansion for the exponential utility. Those portfolios considering higher moments perform better than that truncated at the 2nd moment especially when the market exhibits non-normality.

2.4.2 CVaR and CDaR

VaR fails to measure tail risks if a shape of distributions does not follow a normal distribution. CVaR sometimes referred as Expected Shortfall works to the case of non-Gaussian distributions. The CVaR makes a lot sense not only when financial crisis arises but such non-linear financial instruments as derivatives and credit even in a tranquil environments.

Rockafellar and Uryasev (2000, 2002) introduces an optimization algorithm minimizing Conditional Value-at-Risk (CVaR) rather than minimizing Value-at-Risk (VaR). The algorithm is practically useful because actual problem to solve is reduced to a linear programming.

Concerned with significant serial correlation toward negative direction of returns, drawdown is well recognized amongst investors. Originating from the CVaR as one of risk family, Chekhlov, Uryasev and Zabarankin (2003) proposes a metric named conditional drawdown-at-risk (CDaR) to capture it into portfolio optimizations. CDaR is defined as the mean of the worst $(1 - \alpha) \times 100\%$ drawdowns. An optimization problem to maximize subject to the CDaR can also be reduced to a linear programming similar to Rockafellar and Uryasev (2000, 2002) for CVaR.

2.4.3 Full-Scale Optimizations

An innovation to portfolio construction called full-scale optimization accommodates to any kinds of utility functions and return distributions due to a numerical search algorithm as a non-parametric approach.

From a practitioner's perspective, shaping an optimal portfolio of hedge funds to compose a fund of hedge funds is a real problem to assume non-Gaussian nature of return generating process. To this end, Cremers, Kritzman and Page (2005) studies hedge funds to compare the mean-variance optimal portfolio with the full-scaled optimized portfolio for such non traditional utility such as bilinear utility functions or S-shaped value functions. The mean-variance optimal portfolio works reasonably for investors under the power utility and not workable for investors under the bilinear utility functions or S-shaped value functions.

Estimation errors in investment models can occasionally hurt investment performance for optimal portfolios. Adler and Kritzman (2007) employs a bootstrapping procedure to compare the estimation error of the combined approximation with an estimation error of the mean-variance analysis. And concludes that the full-scale optimization works in out-of-sample space.

Hagströmer and Binner (2009) seriously examines computation burden inevitable for the full scale optimization once number of assets grows. By applying the Differential Evolution proposed by Storn and Price (1997) and further studied by Price, Stone and Lampinen (2005), an optimal portfolio composed of 97 stocks performs decently in out-of-sample basis for several utility functions including the S-shaped and a kinked power.

2.4.4 Extreme Value Theory (EVT) and Copulas

Longin (1996) reports that US stock index belongs to a region of attraction of the Frechêt distribution. Poon, Rockinger and Tawn (2003) finds that tail indices for the Generalized Pareto Distribution (GPD) are all positive to representative stock indices in US, UK, Germany, France and Japan. Those evidences applying the Extreme Value Theory imply an promising outlook to develop an portfolio optimization framework.

Sheikh and Qiao (2010) approaches to non-Gaussian characteristics of data generation process focusing on serial correlation, fat tails and correlation breakdown. The fat tails are estimated with the GPD for marginal distributions and copulas for dependence. Under the Monte Carlo simulation, optimal portfolios are obtained under the CVaR (ES) algorithm.

Assuming the EVT for equity returns to countries, Longin and Solnik (2001) formally establishes the statistical significance of this asymmetric correlation phenomenon, while standard models of time varying volatility (such as GARCH models) fail to capture salient behaviors of international equity returns. Estimating marginal distribution of return series under the GPD, the Monte Carlo simulation method applies to a known distribution for which the tail index is calculated in each time series for different levels of threshold. The optimal value of the tail index is identified on a basis of the mean square error (MSE) criterion. Dependence structure is estimated under a multivariate logistic function. Applying the EVT, a choice of threshold level plays a crucial role.

As surveyed above, the EVT sometimes coupled with copulas is a developing approach to portfolio optimizations in the investment management. As one of key parameters to define distribution shapes specified by the EVT, the tail index is sensitive to where a threshold is identified. Danielson and de Vries (1998) proposes an algorithm to choose a region of exceedance.

A combination of the EVT with copula has occasionally been studied for financial markets. For example, Patton (2004) focuses on two asymmetries, i.e., distribution in each of individual stocks and dependence amongst individual stocks. Assuming time-varying moments up to a 4th order, the copula is introduced to estimate models of the time-varying dependence structure which accepts a different dependence during bear markets from bull markets. Because a double-integral defining the expected utility of wealth does not have a closed-form solution, an optimal portfolio for CRRA investors is obtained throughout the Monte Carlo method on forecast copula.

Chollete, Heinen and Valdesogo (2009) proposes an important choice of copulas under regime dependent framework especially in risk management of international portfolio measured in value at risk. The dependence is modeled with the Gaussian and one canonical vine copula regime. The canonical vines are constructed from bivariate conditional copulas and provide a flexible way of characterizing dependence in multivariate problems. Empirical tests apply to G5 and Latin and American regions. Major findings include that, first, models with canonical vines generally dominate alternative dependence structures. Second, since it modifies the Value-at-Risk (VaR) of international portfolios and produces a better out-of-sample, the choice of copula is found to be important for risk management.

2.4.5 Regime Switch Model

As discussed earlier in the Subsection 2.4, several approaches to handle non-Gaussian distributions of assets' returns in the portfolio have already reached to present optimal portfolio solutions. When examining the non-Gaussian distributions, it is occasionally turned out to be mixture of normal distributions. In this case, it is plausible that the mixture follows to a time series which governs a generating process of the mixture. Unlike an insurance area, the finance area tends to identify time series in data generating process of assets to invest. The regime switching process is the time series which follows the Markov property so that it is expected to specify the discontinuous behaviors of the markets.

Over a recent decade, number of studies on optimal portfolios under regime switching has been increasing. For example, Ang and Bekaert (2002b) constructs and numerically solves a regime switching model for international equity markets and for the US domestic market timing across cash, 10-yr treasury and equity. Number of regime is chosen to be 2 for both models. Ignoring the regimes could cost under a presence of cash in asset allocation problems in both of in-sample and out-of-sample space.

On equity style investments, Ammann and Verhofen (2006) proposes a model to let the Carhart (1997) four-factor model assume regime switching structure. As is almost all other research documenting, different mean returns, volatilities and correlation are found across separable regimes. Value style is beneficial in a regime in the high-variance regime, momentum investing in the low-variance Regime. Beyond the regime estimation, out-of-sample backtest is conducted to conclude the regime-switching model seems to be weak compared to the i.i.d. model.

Ma, MacLean, Xu and Zhao (2011) forecasts sector returns in the US equity market. A model to forecast sector return employs three factors as MKT, SMB and HML and three more as VIX, YS and DS. Intercepts, factor loading and residuals are regime dependent. An optimal portfolio is defined to maximize a linear combination of intercepts across the sectors. A three regime model is chosen in an information criterion and the optimal portfolio generates higher return in the three regime model than a single regime model.

Ang and Bekaert (2002b) models the behavior with regime switch models, solves it numerically in the dynamic programming and reports the ignoring the regimes could costs under a presence of cash in asset allocation problems. Ang and Bekaert (2004) extends Ang and Bekaert (2002b) to solve the dynamic portfolio choice problem. The model is classified dynamic in the sense that an investor is assumed to be exposed to time-varying investment opportunity set modeled using a regime-switching process, i.e., correlation and volatilities increasing in turbulent states. A dynamic programming solves for optimal solutions under the regime dependent CRRA utility to maximize end-of-period wealth in absence of transaction costs. It concludes that the international diversification is still valuable with regime changes and currency hedging imparts further benefit. The costs of ignoring the regimes are small for all-equity portfolios but become expensive when a risk-free asset can be held. The proposed model is uniquely of value because a multi-period optimal solution is numerically computed.

Guidolin and Timmermann (2004) studies strategic asset allocation and consumption choice. Four regimes are separated and asset allocation across them is significantly different each other across bonds, stocks, large-cap and small-cap stocks as well as cash. Maximizing the power utility, no analytical solution but optimal portfolios are derived by a backward solution of the joint consumption and asset allocation problem under regime switching which employs Monte Carlo simulations.

Guidolin and Timmermann (2006) models a regime dependent VAR(1) model for t-bill, bonds and stocks in the US to maximize a power utility over the future. Regime is assumed unobservable and Monte Carlo method applies to simulate future paths to generate for calculating the utility. Four regime is chosen in-sample basis. One of major finding in an empirical example highlights an ability of the framework to show different asset allocation across different investment horizons, e.g., stock allocations are found to be monotonically increasing as the investment horizon gets longer in only one of the four regimes.

Guidolin and Timmermann (2008a) documents a regime dependent style investments to equity market, SMB and HML. Both of SMB and HML are turned out to be regime dependent. Assuming an investor to maximize a power utility, the Monte Carlo method finds the optimal solution. Since regime switches generate predictability in future investment opportunities, derived optimal investment solutions show horizon effects and hedging demands. Although the approach does not give any tractable nature of derived solution, sensible characteristics of regime dependent optimal solutions is clearly addressed.

Guidolin and Timmermann (2008b) applies a regime switch to a four-moment ICAPM where returns on the market portfolio depend not on variance but skewness and kurtosis that are time varying. Intercepts and risk premium in the model are regime dependent. Beside to the ICAPM, the model includes predictor variables that follow VAR(1) sharing the regime with the ICAPM. A major finding is that co-skewness and co-kurtosis risk have economically intuitive signs, i.e., investors dislike risk in the form of higher volatility or fatter tails but like positively skewed return distributions. Furthermore, the co-skewness and co-kurtosis risk premium appear to be important in economic terms, as they are of the same order of magnitude as the covariance risk premium. An optimal solution is derived for a CARA utility to maximize. As studied in Guidolin and Timmermann (2006), the Monte Carlo method applies to simulate future paths for out-of-sample tests that show better results than in single regime model.

2.4.6 Multi-period Optimizations

As a second item in the set of questions mentioned at the beginning of this chapter, it is conceivable to identify other portfolios that are expected to have better investment efficacy than that solved under the myopic problem if an investor knows predictable behaviors of assets' returns for further futures. Section 2.2 finds that factors occasionally display time variant behaviors. Section 2.3 tells that financial time series follows to the regime switching. Section 2.4 gives examples of optimal portfolios that hold assets exhibiting non-Gaussian nature. Especially, Subsection 2.4.5 reveals that some of earlier studies find optimal solutions to hold regime dependent assets. Some studies discuss multi-period optimal portfolios that are numerically obtained by the Monte Carlo method. In this section, a domain of the multi-period optimization is focused in order to identify the latest front of analytical work for a closed form solution. Largely, two major approaches are available in the multi-period optimizations:

- Stochastic programming
- Stochastic control

Besides to the two formal approaches, a couple of more ways are available to solve this sort of problems. One example is to solve a sequence of myopic problem and other is a stochastic simulation applying decision rules under the Monte Carlo method.

The stochastic programming is a powerful approach to calculate optimal solutions for highly complicated and practical problems that are not possible to be solved for analytical solutions. A recent evolution in computational cost reduction is also encouraging to apply the stochastic programming.

On the other hand, the stochastic control theory provides a framework to solve the multi-period problem in analytical ways. The solution is highly tractable to understand how the solutions behave given changes in key parameters that compose the solution. The stochastic control theory gives tremendous contributions to Liu (2004) for optimal consumption and investment choice and Grinold (2007) discussing sensitivity of analytical solutions to length of life of factors. Grinold (2007) points out vintage of information can be found in portfolios under a presence of transaction costs. Prior to Grinold (2007), Grinold (1997) finds importance of the information horizon of a signal or of an investment strategy under a presence of the transaction costs. Stimulated by Grinold (1997, 2007), Sneddon (2008) solves mean-variance optimal portfolio solutions under quadratic t-cost penalties investing assets with across different decay speed in forecasted return signals modeled in AR(1). The solutions reveal that the optimal portfolio should trade more aggressively high decay assets than low decay assets.

Gârleanu and Pedersen (2013) solves a closed form solution for a multifactor portfolio in which the multifactor follows to VaR(1) with a transaction cost to bear in a quadratic form. In the stochastic control problem, the predicted trajectory of the market impact to securities prices are well reflected to derived solutions. Especially for the high frequency trading model, the market impact is occasionally expected not to reset to before the previous trades. As such the integration of predicted market impact is highly valuable if the prediction is reasonable. The solution takes a form of a linear combination of a current portfolio and an aim portfolio which is a combination of the Markowitz portfolio and an expected optimal target portfolio in the future. If a factor decays more slowly then the optimal portfolio weighs more to those assets exposed to the factor. The more expensive transaction cost, the more slowly approach to the aim portfolio.

Applying the solutions to the investment practices, it is occasional to impose investment constraints. For example, no short sales in portfolio holdings is a popular constraint in long-only and asset allocation problems. To this end, Li, Zhou and Lim (2002) and Cui, Gao, Li and Li (2014) solve multiperiod optimal solutions for the mean-variance investment utility under the no short sales constraint. None of them assumes transaction costs to bear in the optimization problem. As Grinold (2007), Sneddon (2008) and Gârleanu and Pedersen (2013) agree, the presence of transaction costs dominates the shape of optimal solutions, yet the no short sales constraint is successfully incorporated in the multi-period optimal solutions.

In order to relax the tightly constrained problems to have solved analytically, Collin-Dufresne, Daniel, Moallemi and Sağlam (2014) finds analytical solutions in a linear-quadratic optimization problem reduced from a highly non-linear dynamic optimization problem if the investment is classified into the linearity generating strategies. The strategy is as a strategy for which the dollar position in each security is a weighted average of current and lagged exposures. Moallemi and Sağlam (2013) summarizes a broad class of dynamic portfolio optimization problems that accept complex models of return predictability, transaction costs, trading constraints, and risk considerations. For those classified into a class of linear rebalancing rules, numerical procedures enable to calculate optimal solutions. Yet neither Collin-Dufresne, Daniel, Moallemi and Sağlam (2014) nor Moallemi and Sağlam (2013) incorporates the regime switch process which governs to flip amongst underlying models.

Wrapping up Section 2.4, Table 2.2 compares data generating processes, objective functions and optimal solutions for some of the optimization problems.

Table 2.2: Examples of optimization problems

Problem	DGPs	Objectives	Solution $\boldsymbol{x}_t^*$
Markowitz[1]	$\boldsymbol{r}_t = \boldsymbol{\mu} + \boldsymbol{\epsilon}_t$ $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma})$	$\max \left\{ \boldsymbol{x}^\top \boldsymbol{\mu} - \frac{\lambda}{2} \boldsymbol{x}^\top \boldsymbol{\Sigma} \boldsymbol{x} \right\}$	$(\lambda \boldsymbol{\Sigma})^{-1} \boldsymbol{\mu}$
Dynamic M-V policy under quadratic t-cost[2]	$\Delta \boldsymbol{f}_t = -\Phi \boldsymbol{f}_{t-1} + \boldsymbol{\epsilon}_t$ $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \boldsymbol{\Omega})$ $\boldsymbol{r}_t = \boldsymbol{B} \boldsymbol{f}_{t-1} + \boldsymbol{u}_t$ $\boldsymbol{u}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma})$	$\max_{\{\boldsymbol{x}_0, \boldsymbol{x}_1, \dots\}}$ $\boldsymbol{E}_0[\sum_t (1-\rho)^{t+1} (\boldsymbol{x}_t^\top \boldsymbol{r}_{t+1} - \frac{\gamma}{2} \boldsymbol{x}_t^\top \Sigma \boldsymbol{x}_t)$ $-\frac{(1-\rho)^t}{2} \Delta \boldsymbol{x}_t^\top \Lambda \Delta \boldsymbol{x}_t]$	$\boldsymbol{x}_{t-1} + \Lambda^{-1} \boldsymbol{A}_{xx} (aim_t - \boldsymbol{x}_{t-1})$ where $aim_t = \boldsymbol{A}_{xx}^{-1} \boldsymbol{A}_{xf} \boldsymbol{f}_t$ $\boldsymbol{A}_{xx} = (\rho\gamma \Lambda^{\frac{1}{2}} \Sigma \Lambda^{\frac{1}{2}}$ $+\frac{1}{4}(\rho^2 \Lambda^2 + 2\rho\gamma \Lambda^{\frac{1}{2}} \Sigma \Lambda^{\frac{1}{2}} + \gamma^2 \Lambda^{\frac{1}{2}} \Sigma \Lambda^{-1} \Sigma \Lambda^{\frac{1}{2}}))^{\frac{1}{2}}$ $-\frac{1}{2}(\rho\Lambda + \gamma\Sigma)$ $vec(\boldsymbol{A}_{xf}) = \rho(\boldsymbol{I} - \rho(\boldsymbol{I} - \boldsymbol{\Phi})^\top \otimes (\boldsymbol{I} - \boldsymbol{A}_{xx} \Lambda^{-1}))^{-1}$ $vec((\boldsymbol{I} - \boldsymbol{A}_{xx} \Lambda^{-1}) \boldsymbol{B})$
Optimal Linear Rebalance[3]	Same as above except for $\boldsymbol{r}_t = \boldsymbol{\mu} + \boldsymbol{B} \boldsymbol{f}_{t-1} + \boldsymbol{u}_t$	$\max_{\{\boldsymbol{E}, \boldsymbol{c}\}}$ $E[\sum_{t=1}^T ((\boldsymbol{d}_t + \sum_{s=1}^t \boldsymbol{J}_{s,t} \boldsymbol{f}_s)^\top \boldsymbol{B} \boldsymbol{f}_t$ $-\frac{1}{2}(\boldsymbol{c}_t + \sum_{s=1}^t \boldsymbol{E}_{s,t} \boldsymbol{f}_s)^\top \Lambda$ $(\boldsymbol{c}_t + \sum_{s=1}^t \boldsymbol{E}_{s,t} \boldsymbol{f}_s))]$ where trades $\boldsymbol{u}_t$ and positions $\boldsymbol{x}_t$ $\boldsymbol{u}_t = \boldsymbol{c}_t + \sum_{s=1}^t \boldsymbol{E}_{s,t} \boldsymbol{f}_s$ $\boldsymbol{x}_t = \boldsymbol{d}_t + \sum_{s=1}^t \boldsymbol{J}_{s,t} \boldsymbol{f}_s$ subject to $\boldsymbol{d}_t = \boldsymbol{x}_0 + \sum_{i=1}^t \boldsymbol{c}_i, \quad 1 \le t \le T$ $\boldsymbol{J}_{s,t} = \sum_{i=s}^t \boldsymbol{E}_{s,i}, \quad 1 \le s \le t \le T$ $P(\boldsymbol{d}_t + \sum_{s=1}^t \boldsymbol{J}_{s,t} \boldsymbol{f}_s) \le \eta, \quad 1 \le t \le T$ $P(\boldsymbol{c}_t + \sum_{s=1}^t \boldsymbol{E}_{s,t} \boldsymbol{f}_s) \le \eta, \quad 1 \le t \le T$ $d_T = \boldsymbol{0}$ $J_{t,T} = \boldsymbol{0}$	Quadratic Program

[1] See Markowitz (1952).
[2] See Gârleanu and Pedersen (2013).
[3] See Moallemi and Sağlam (2013).

2.5 Discussion on Future Issues

This section has reviewed three domains where we could potentially identify key breakthrough for the investment efficacy for optimal portfolios with factor models. First, vast majority of literature agrees that common factors in the markets are time varying. The time variant nature addresses a couple of important issues in deriving optimal portfolios.

1. Predictability in common factor loading, dispersions and dependence of residuals
2. Non-Gaussian shapes of distributions of the key parameters

For the first item, if the system to consider is dynamic, a myopic optimal solution is not optimal but a multi-period optimal portfolio delivers an optimal solution by summing up expected investment utility over the future over investment horizons ahead of the investors. Unlike insurance area, financial data often exhibits time series nature.

On the second item mentioned above as the Non-Gaussian shapes of distributions of the key parameters, earlier studies tend to agree that the state space specification works with the Kalman filter algorithm better than other approaches. On the other hand, across all major asset classes, recent studies document the regime switch structure is identified. Factors are also found to exhibit the regime switching. Rapid and discontinuous changes in key parameter behaviors are expected to be modeled to supersede conventional models. Beyond those documented in the literature, there still are significant room to study the regime switching model in a diversified investment portfolio problem.

Third, the conventional Markowitz model solves an optimal portfolio if investors live under the quadratic utility and the data generation process for asset's return follows normal distributions. Recent experiences in turbulent behaviors of the markets accelerate growth of number of research in the literature. To this end, although a combination of the extreme value theory and copula is emerging as one of promising approaches, it is potentially true that much room exists to stick to normal distributions if an actual distribution is deemed to be a mixture of the normal distributions. On the other hand, an analytical solution is derived for a multi-period problem if the factor model is specified in the VAR(1) model. The solution solved under the stochastic control theory is tractable enough to discuss if augmented to include the regime switching model.

In sum, tackling with both of the time series in common factors and non Gaussian distributions in both of asset returns and factor behaviors, amongst potential approaches, the regime switching model is worth to apply to a multi-period optimization problem to solve because of growing amount of regime switching models for investment management and as yet not solved under multi-period optimization problems.

Now issues for future research on the optimal portfolios are summarized in following three directions:

- A first interest focuses on how multiple regime model potentially improves efficiency of optimal portfolios. To answer to this research question, a myopic problem for a factor model is to examine.

- Second, provided that time series behaviors of factor loading, a multi-period extension of the myopic optimal solution is to study for analytical solutions to derive.

- Third, as an investment constraint, non negativity as a constraint is imposed in deriving the multi-regime and multi-period optimal solution. The non negativity is a practical constraint typically in an asset allocation problem. Due to the constraint, an analytical solution is hardly achievable but a problem for an numerical solution is formulated and solved.

Figure 2.1 exhibits how each of following three sections works for the three research issues and illustrates major contributions to the literature.

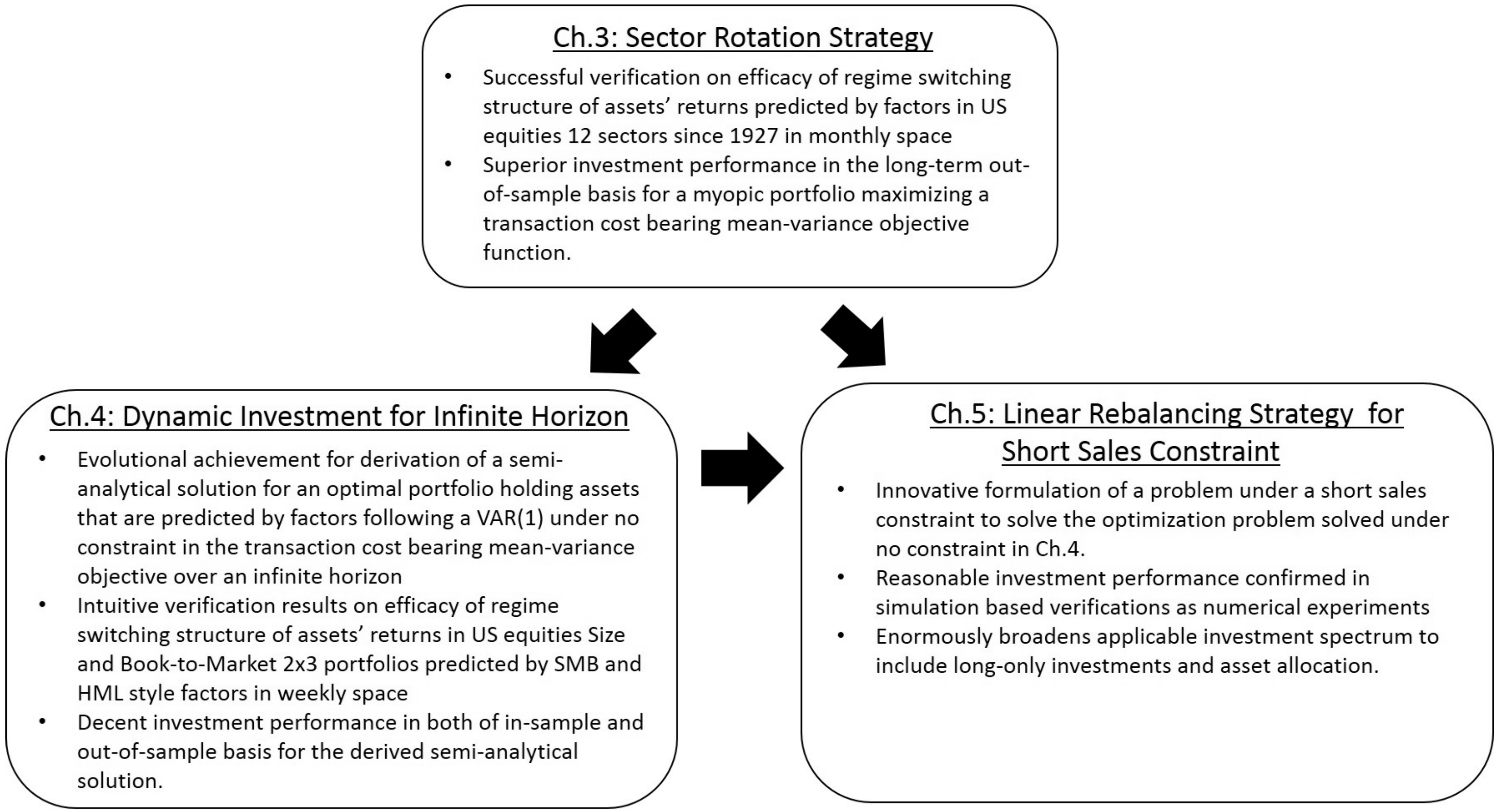


Figure 2.1: Contribution summary of following three sections

Chapter 3

Sector Rotation Strategy

3.1 Introduction

Recent innovations of investment portfolios augment sources of risk premiums from traditional equity centric risk allocation in the CAPM to include such exotic beta as credit, emerging, inflation and small capitalization equities as well as alternative risk premium. Some of the alternative risk premium comes from long-lived anomaly in the markets. An investor justifies the risk premium to allocate some of permissible total risk because of less correlation to major risks in financial markets. Low correlation to major risks in the financial markets motivates investors to budget risks to the exotic beta and alternative risk premium to monetize. Over reasonable length of periods, the alternative risk premium is expected to bring in portfolios unique source of return generation pieces.

Over decades financial markets have exhibited drastic changes in return generating processes that deviate from long-term expectations of investors. The drastic changes have appeared at least four times in recent decades; 1) late 1970s and early 1980s that are known as a lost decade for equities and rates, 2) technology bubble growth and bursting around 2000, 3) late 2000s global crisis in economy and in the financial markets triggered by subprime loan and 4) sovereign crisis in developed and emerging market in early 2010s. The drastic changes challenge investors to re-examine if the risk premiums deserves major seats in diversified portfolios.

This chapter addresses potential issues of the sector rotation strategy in the US equity if it deserves one of strategy buckets in diversified portfolios partly because consistent abilities and stable investment efficacy are still the case in the recent markets. This chapter focuses on the sector rotation strategy because earlier studies find that momentum anomaly drives the strategy profitable for investors. The momentum anomaly is known as one of the alternative risk premiums identified in equity, rates and currency markets quite sometime. Asness, Liew and Stevens (1997) argues that momentum anomaly is robust to major countries in equity markets. Carhart (1997) concludes mutual fund performance in the US does not reflect managers' investment skills but common factors including momentum and transaction costs as well as fund expenses. Chan, Hameed and Tong (2000) echoes to earlier studies on momentum effect in international equity market and finds volume increase in past period can further add efficacy of momentum strategy. Also Asness, Moskowitz and Pedersen (2013) documents interactions between value and momentum across diverse global markets and asset classes and discusses that common risks, e.g., funding liquidity, to global markets drives the interactions. The findings challenge to traditional behavioral models to explain value and momentum anomaly and inspire to explore innovative structure behind the momentum. Especially in the equity markets, as one of well known financial anomaly, the equity sector rotation has been sometimes focused to apply momentum strategies. There is growing amount of evidence that sector returns are related to past profile of own returns, such as, momentum in the US and other markets in the world. Jagadeesh and Titman (1993) claims one year momentum is attributed not to systematic factors but to delayed price reactions to firm specific information. Moskowitz and Grinblatt (1999) finds that an industry component of stock returns accounts for much of individual stock momentum anomaly.

Apart from the momentum in the equity markets, the value effect is of major focus amongst academia and practitioners for quite long time. Lakonishok, Shleifer and Vishny (1994) stands in the behavioral model to understand value effect as investors incorrectly extrapolate past earnings growth rates, i.e., investors are overly optimistic (pessimistic) to firms done well (worse) in the past. Fama and French (1996) claims to higher systematic risks associated for value stocks as representing the risk based root causes. In the long-lasting value discussion, Guirguis, Theodore and Suen (2012) gives a new insight to the discussion by Lakonishok, Shleifer and Vishny (1994) and Fama and French (1996) by applying regime switching structure for value anomaly. It finds that prediction power of the value anomaly by earnings yield dispersion is dependent on risk regime of value index.

Similar to the value discussions, academia puts inexhaustible endeavor on risk based or behavioral model to understand background of momentum. Grifin, Ji and Martin (2003) approaches to the momentum effect in the point of view of macro economic risk and finds that momentum is driven by country specific risks and weakly co-moves across countries. Inflation, term spread and changes in industrial production are not useful to explain momentum efficacy. In the international space, Chordia and Shivakumar (2002) finds that momentum profits are positive across all tested macroeconomic environments and are not explained by macro economic risks. Rouwenhorst (1998) also shows that the momentum return based on past medium term performance is found in international 12 countries and indicates that the effect is not attributed to such conventional risks as market risks or size risks.

As mentioned at the beginning, while the academic discussions continue, practitioners may want make sure if the momentum is a reliable sources of risk premium. Daniel and Moskowitz (2013) discussing negative skews in momentum strategy suggests that crash in momentum occurs in the panic states and reports that a hike in equity volatility is a potential predictor for momentum returns in individual stocks. The key findings by Daniel and Moskowitz (2013) motivates to study the regime switch in the momentum strategy since the crash in financial time series sometimes exhibits nature of regime switching.

In a recent decade, increasing number of research articles have documented optimal portfolio decisions over regime switching return generating processes. Ang and Bekaert

(2002a, 2004) solves for the optimal asset allocation using Gaussian quadrature methods for international markets and reports that ignoring the regimes could cost under a presence of cash in asset allocation problems. Ammann and Verhofen (2006) extends the Carhart (1997) four-factor model to regime dependent ways and reports the extended model delivers profitable results. Guidolin and Timmermann (2004, 2008a) introduces an optimal portfolio allocation using Monte-Carlo methods for approximating expected utility under the set up that regime is not observable and handled as latent states. Liu, Xu and Zhao (2010) applies a three regime model to the Fama-French model together with three macroeconomic factors to predict returns to sector ETF in the equity market. They confirm improved fits of the model by introducing regime switches. Seidl (2012) discusses portfolio optimization where the mean-variance utility is weighted by regime probabilities. Comparing with the mean-variance model by Markowitz (1952), the proposed model applied to stocks, bonds, hedge funds, commodities and real estates performs much better than the classical mean-variance model under no transaction cost.

This chapter contributes to literature in several ways. First, over the long term history since 1927 toward 2013 in monthly basis, it is revealed that the sector momentum is regime dependent and changes over two regimes and three regimes. In multiple regimes, cross-sectional sector dispersion estimated in variance-covariance distinguishes a turbulent risk regime from the others. Momentum is more significant in the tranquil regime than the turbulent regime when some sectors behave reversal. Configuring the three regime structure in the proposing sector return forecasting model for the 86 year period, it is noteworthy that regime shifts in the monetary and the fiscal policies in early 1980s are the second most notable events. Second, out-of-sample portfolio optimization practices from 1976 toward 2013 deliver encouraging investment performance. We show statistical evidence of normality test to support decent investment performance for the three regime model. Third, a regime dependent risk aversion coefficient, which makes sense to risk averse investors, plays crucial roles to further improve investment efficacy with shallow maximum drawdown.

The outline of this chapter is as follows. In Section 3.2, we define momentum factors and a model to predict returns to sector under regime switches. Section 3.3 describes data and estimated models over 86 years to find a reasonable basis to expect decent performance in multiple regime models. Section 3.4 shows the optimal investment solutions and reports out-of-sample investment performance of the optimal portfolios for 37 years. Finally, we conclude the chapter in Section 3.5.

3.2 Model

In this section, we first describe a return forecasting factor and a regime switch model to employ the factor for sector return predictions.

3.2.1 Return Forecasting Factor

In general, momentum is known as market anomaly inducing past winners tend to outperform past losers in future periods. Since no single past winner/loser continues to outperform/underperform others forever, winners to hold more and losers to hold less in the investment strategy change over time. As such investment practitioners sometime refer this strategy as a sort of a rotation across assets to hold in the portfolio. A number of earlier studies find the momentum as common behavior to major financial assets across the globe as introduced in the previous section. A momentum factor is, at certain point in time, a metric to quantify how significantly each asset outperforms or underperforms each other in the past. Jagadeesh (1990) and Lehmann (1990) shows that short-term reversals, e.g., the contrarian strategies on stocks in the previous week or month. That is bid-ask spread sensitive and cancels a trend formed in past 12 months. Our goal is to forecast a sector return one month ahead and to build optimal portfolios to generate positive absolute returns in risk and transaction cost adjusted basis. As such, to define the factor for the one month forecast purpose, our study here excludes returns in a recent one month and accumulates past 11 monthly returns starting from a monthly return from 12 months ago to 11 months ago up until a monthly return from 2 months ago to 1 month ago. This is known as a PAST(2,12).

On the PAST(2,12) time frame, the momentum factor is defined as a cumulative relative return to the each sector excess of a market average in the market risk adjusted basis. Let $r_k(t)$, $R_f(t)$ and $R_m(t)$ respectively denote a return of sector k, a risk free rate and a market return. For each time t, we use a set of data $\{r_k(s), R_m(s), R_f(s);\ s = t-11, \dots, t-1\}$ of the past 11 months to run a CAPM analysis

$$r_k(s) - R_f(s) = \alpha_k(t) + \gamma_k(t)\{R_m(s) - R_f(s)\} + \epsilon_k(s), \quad s = t-1, \dots, t-11 \tag{3.1}$$

for estimating $\alpha_k(t), \gamma_k(t)$ and $\epsilon_k(s)$. Based on the estimated $\alpha_k(t)$ and $\epsilon_k(s)$, we define a factor $f_k(t)$ for sector k at t by

$$f_k(t) = \left\{ \prod_{s=t-11}^{t-1} \{1 + \alpha_k(t) + \epsilon_k(s)\} \right\}^{1/11} - 1. \tag{3.2}$$

3.2.2 Regime Switch Model

Our model for forecasting sector returns is a multivariate factor based CAPM with regime dependent coefficients. We are motivated to apply multivariate framework by a couple of recent studies. Menzly and Ozbas (2006) finds strong cross momentum through the supply chain upstream-downstream among industries. Cohen and Frazzini (2008) reports an evidence that firms in customer-supplier relationships generate customer momentum which brings positive returns to buy long firms' customers recorded better returns in stock prices and sell short firms'customers did worse returns. Accordingly, we use vector $\boldsymbol{f}(t) = (f_1(t), \dots, f_k(t), \dots, f_N(t))^\top$ ($\top$ denotes transpose) of all N factors to forecast one time step ahead return vector $\boldsymbol{r}(t+1) = (r_1(t+1), \ldots, r_k(t+1), \ldots, r_N(t+1))^\top$ by

$$\boldsymbol{r}(t+1) - R_f(t+1)\mathbf{1} - \{R_m(t+1) - R_f(t+1)\}\,\boldsymbol{\beta}_i = \boldsymbol{L}_i\boldsymbol{f}(t) + \boldsymbol{u}_i(t+1) \tag{3.3}$$

where $\mathbf{1} = (1, \ldots, 1)^\top$ is a column vector of 1's. In order to represent discontinuous state changes of the market, we introduce the Markov switching regime process that drives random fluctuation of the coefficients over time. The subscript i in (3.3) represents the regime at $t+1$ which can randomly take one on $\{1, ..., J\}$ when number of regimes is J. In (3.3), both CAPM beta $\boldsymbol{\beta}_i = (\beta_{i,1}, \ldots, \beta_{i,k}, \ldots, \beta_{i,N})^\top$ and $N \times N$ factor loading matrix $\boldsymbol{L}_i$ are dependent on regime i. The second term $\boldsymbol{u}_i(t+1) = (u_{i,1}(t+1), \ldots, u_{i,k}(t+1), \ldots, u_{i,N}(t+1))^\top$ in (3.3) represents an unpredictable noise vector that follows a multivariate normal distribution and satisfies $E(u_{i,k}(t+1)) = 0$ for all regime i and sector k. We suppose that the covariance matrix of $\boldsymbol{u}_i(t+1)$ is also regime-dependent and is denoted by $\boldsymbol{W}_i = V(\boldsymbol{u}_i(t+1))$. The first term $\boldsymbol{L}_i\boldsymbol{f}(t)$ denotes the expected excess return known to the investor at time t. We remark that positive elements in the factor loading matrix $\boldsymbol{L}_i$ correspond to the positive momentum while negative elements indicate the return reversal.

3.3 Data and Model Estimation

Over 86 year long data in monthly frequency, we estimate the regime switching structure of a momentum factor to predict US equity sector returns. As a tool for model estimation, we use the R as a statical analysis package running in the Intel(R) Core(TM) i7-4960X CPU 3.60GHz 6 Cores 12 Threads under 64bit operating system with 8G byte memory.

3.3.1 Data

In this subsection, we discuss analysis on regime switches over meaningfully long period of time. For this purpose, visiting to the Kenneth French Data Library, we retrieve monthly series of research data covering from July 1926 to June 2013 for approximately 87 years. See Table 3.13 in Subsection 3.6.1 for a list of the sectors. In the rest of this chapter, we refer the short names in Table 3.13. The number of sectors to apply the model is 12 that is reasonably and practically large to form a well diversified portfolio. These 12 sector portfolios are composed of all NYSE, AMEX and NASDAQ stocks that hold individual stocks in market value weighted basis. We choose the value weighted series out of the US Research Returns Data rather than equal weight series. The value weight series is more realistic than the equal weight because of liquidity reasons. We use a one-month Treasury bill rate for $R_f(t)$ and market return $R_m(t)$ given by the value-weighted return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ. These are also taken from the Kenneth French Data Library. Table 3.1 gives a summary statistics of the returns to the sectors and the market. Figure 3.1 and Figure 3.2 depict cumulative profiles of returns to the sectors and the market, respectively. As investment instruments to implement the sector rotation strategy, such sector ETF in SPDR, iShares and Vanguard are appropriate examples providing tradable market liquidity for sizable assets to manage and short sale capabilities. Many of them have been listed to the exchange over a decade. Number of sectors in the listed sector ETF is more than 12 today. Table 3.2 and Figure 3.3 display a summary statistics and cumulative profiles of the factors of the sectors defined in (3.2), respectively.

Table 3.1: Summary Statistics of returns to 12 sectors and the market

	min.	Q1	median	mean	Q3	max.	std
NoDur	−.2461	−.0139	.0110	.0098	.0365	.3439	.0466
Durbl	−.3482	−.0275	.0104	.0110	.0489	.7987	.0781
Manuf	−.2883	−.0211	.0148	.0103	.0447	.6015	.0679
Enrgy	−.2600	−.0232	.0090	.0106	.0453	.3347	.0601
Chems	−.3162	−.0201	.0113	.0102	.0433	.4885	.0581
BusEq	−.3463	−.0295	.0113	.0109	.0515	.5868	.0762
Telcm	−.2156	−.0135	.0094	.0086	.0320	.2819	.0463
Utils	−.3285	−.0164	.0107	.0088	.0360	.4285	.0559
Shops	−.3022	−.0213	.0114	.0101	.0413	.4225	.0591
Hlth	−.3408	−.0194	.0109	.0108	.0408	.3713	.0565
Money	−.3959	−.0211	.0118	.0101	.0448	.5978	.0689
Other	−.3122	−.0231	.0101	.0084	.0431	.5856	.0666
Market	−.1255	−.0107	.0048	.0069	.0211	.3411	.0353

Monthly figures in July 1926 to June 2013

3.3.2 Model Estimation

(3.3) is rearranged in the state space representation, for an example of number of regime $J = 3$, of which an observation model is given as

$$\boldsymbol{r}(t+1)-R_f(t+1)\mathbf{1}_N = \begin{cases} \{R_m(t+1)-R_f(t+1)\}\boldsymbol{\beta}_1 + \boldsymbol{L}_1\boldsymbol{f}(t) + \boldsymbol{u}_1(t+1), & I(t+1)=1 \\ \{R_m(t+1)-R_f(t+1)\}\boldsymbol{\beta}_2 + \boldsymbol{L}_2\boldsymbol{f}(t) + \boldsymbol{u}_2(t+1), & I(t+1)=2 \\ \{R_m(t+1)-R_f(t+1)\}\boldsymbol{\beta}_3 + \boldsymbol{L}_3\boldsymbol{f}(t) + \boldsymbol{u}_3(t+1), & I(t+1)=3 \end{cases} \tag{3.4}$$

where $\boldsymbol{u}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{W}_i), i = 1, 2, 3$. And a system model is

$$\boldsymbol{q}_{t+1} = \boldsymbol{P}^\top \boldsymbol{q}_t \tag{3.5}$$

or

$$\begin{bmatrix} q_{1,t+1} \\ q_{2,t+1} \\ q_{3,t+1} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \\ p_{13} & p_{23} & p_{33} \end{bmatrix} \begin{bmatrix} q_{1,t} \\ q_{2,t} \\ q_{3,t} \end{bmatrix}$$

where $P\big(I(t) = i\big) = q_{i,t}$ and $q_{1,t} + q_{2,t} + q_{3,t} = 1$.

We apply a filtering in identifying a regime $\{I(t)\}$. The filtering algorithm is given in Subsection 3.6.2. In the entire period for an analysis on regime dependency across

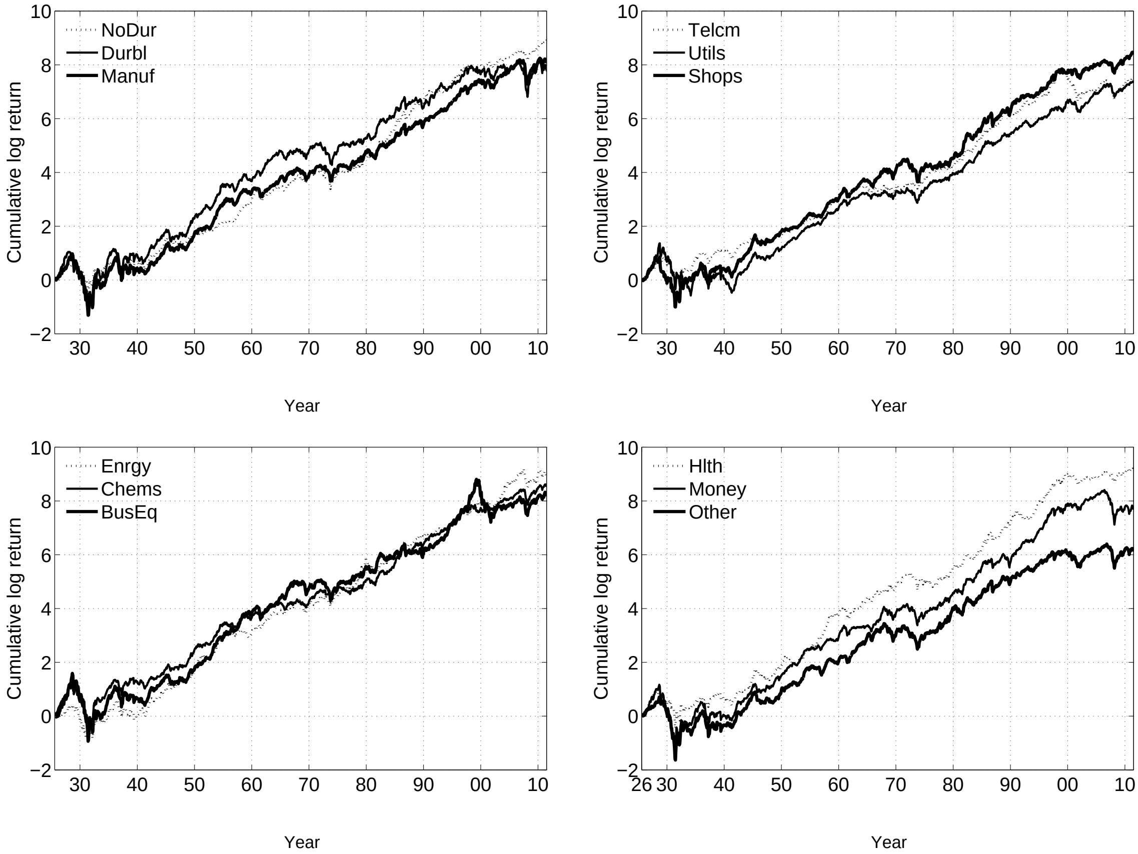


Figure 3.1: Cumulative performance of 12 sectors

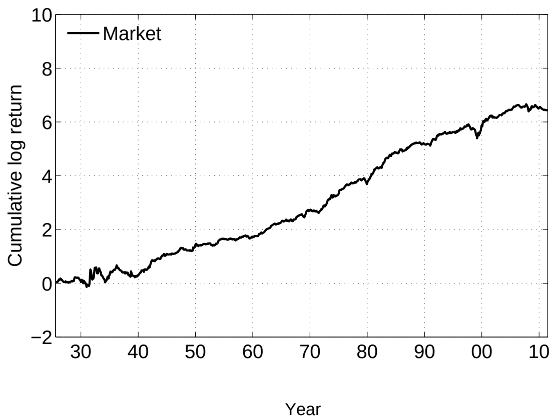


Figure 3.2: Cumulative performance of the market

Table 3.2: Summary Statistics of factors of 12 sectors

	min.	Q1	median	mean	Q3	max.	std
NoDur	−.0290	−.0028	.0018	.0018	.0063	.0338	.0077
Durbl	−.0441	−.0084	−.0011	−.0008	.0070	.0472	.0123
Manuf	−.0250	−.0049	−.0009	−.0007	.0036	.0224	.0063
Enrgy	−.0492	−.0056	.0015	.0017	.0092	.0391	.0113
Chems	−.0343	−.0047	.0002	.0002	.0055	.0280	.0075
BusEq	−.0457	−.0072	−.0013	−.0002	.0065	.0416	.0104
Telcm	−.0352	−.0038	.0019	.0016	.0074	.0452	.0099
Utils	−.0668	−.0046	.0023	.0014	.0086	.0760	.0119
Shops	−.0367	−.0044	.0014	.0013	.0077	.0323	.0094
Hlth	−.0319	−.0066	.0017	.0015	.0093	.0351	.0115
Money	−.0303	−.0044	.0010	.0010	.0062	.0357	.0094
Other	−.0548	−.0070	−.0018	−.0022	.0028	.0310	.0080

Monthly figures in June 1927 to June 2013

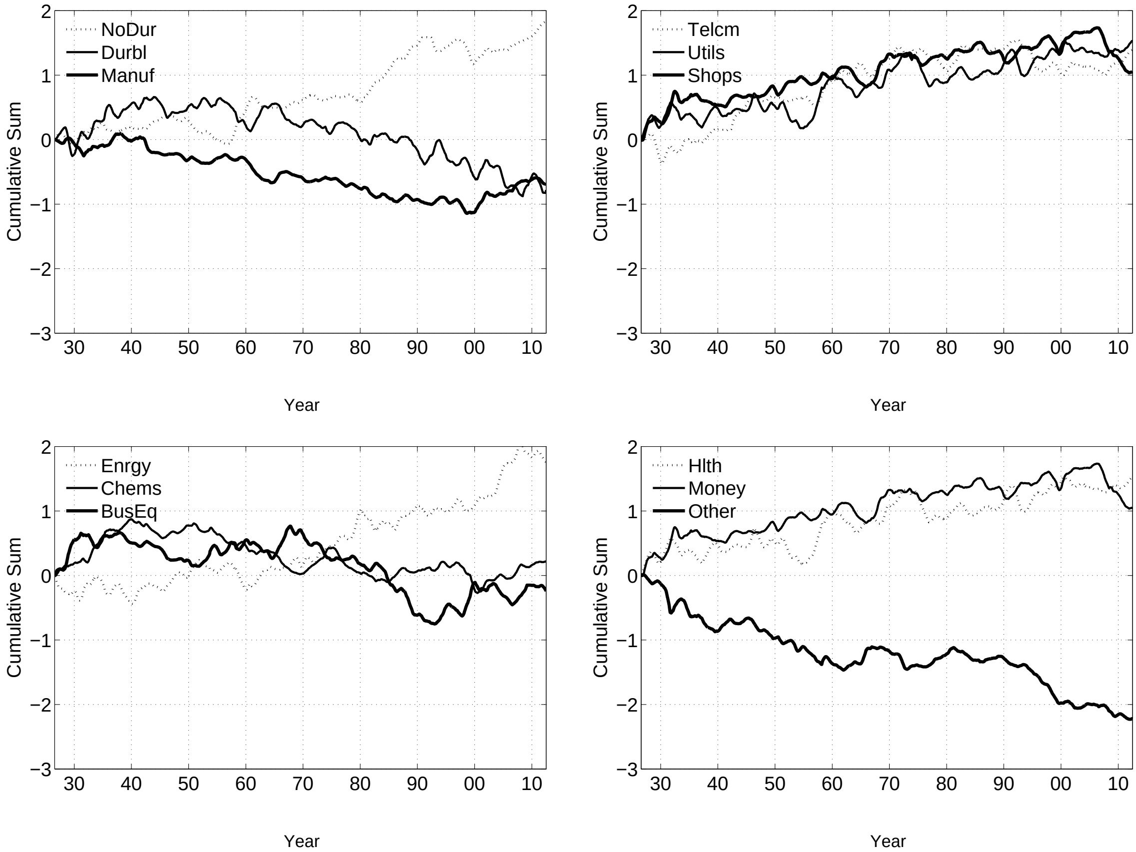


Figure 3.3: Cumulative profiles of 12 factors

12 sectors, the model parameters in (3.4) and (3.5) are estimated based on 1044 month long data set which provides with 1032 month long 12 factor series throughout (3.1) and (3.2). The quasi-Newton algorithm identifies the parameter set in the framework of the maximum likelihood estimation under the assumption that the residuals in (3.4) follow a regime dependent multivariate normal distribution. We adopt the filtering likelihood in the maximum likelihood estimation. Concerning the number of regimes, we estimate two and three regime models respectively while single regime model is also studied for comparison. The Akaike's Information Criterion of the two and three regime models are 57391.52 and 56441.69 that are much better than 59927.07 of the single regime model. As a labeling convention over observed regimes, we label higher numbers as names of regimes for more turbulent regimes. Specifically, the larger the sum of the variances of all sectors is, the higher the number of regime. For notational simplicity, we denote Regime i in the J regime model by Regime i/J. For instance, Regime 2/3 indicates Regime 2 in the three regime model.

First of all, Table 3.3 summarizes estimated variance of residuals $\boldsymbol{u}_i$ in (3.4). Since a total return to each sector is subtracted by $\{R_m(t) - R_f(t)\}\,\boldsymbol{\beta}_i$ in (3.4), residuals covariance among $\boldsymbol{u}_i$ is more close to zero than variance. We therefore let only the variance appear in the Table 3.3. The most notable finding is that variance in Regime 2/2 is several times higher than those in Regime 1/1. This is also the case for Regime 3/3. It is natural to understand that these regimes represent a sort of turbulent state of cross sectional return variabilities across sectors. Other two regimes Regime 1/3 and Regime 2/3 in the three regime model look tranquil. In addition to this, "Durbl" and "Enrgy" become significantly more volatile in Regime 2/3 than in Regime 1/3. We revisit to this finding later in this chapter. Next, Table 3.4 reports the factor loading $\boldsymbol{L}_i$ in (3.4). Moskowitz and Grinblatt

Table 3.3: Diagonal elements of $\boldsymbol{W}_i$ (Full Sample:1927/07-2013/06)

	Single Regime	Two Regimes		Three Regimes		
($\times 10^{-4}$)	Regime 1/1	Regime 1/2	Regime 2/2	Regime 1/3	Regime 2/3	Regime 3/3
NoDur	4.7	3.0	8.5	2.4	4.3	9.2
Durbl	15.2	7.6	35.5	6.4	11.4	37.2
Manuf	4.1	2.0	10.3	1.7	2.9	11.3
Enrgy	14.7	9.6	30.4	8.5	13.1	28.8
Chems	6.0	3.1	14.8	2.9	3.6	16.4
BusEq	10.0	6.9	18.2	6.2	7.8	19.6
Telcm	8.7	6.0	17.1	5.4	6.0	18.1
Utils	13.2	6.9	32.0	6.5	7.0	34.8
Shops	7.2	4.9	12.9	5.0	4.7	13.7
Hlth	11.2	7.4	20.9	7.7	6.7	23.4
Money	7.3	3.8	16.9	3.8	4.0	17.2
Other	6.5	3.6	14.8	4.0	2.5	16.7

Monthly figures

(1999) argues the momentum effect to sectors. We find that Regime 1/1 in the single regime model let 5 out of 12 sector behave significant positive momentum while 3 sector reversal although nothing significant. On the other hand, number of significant positive momentum increases from 5 in Regime 1/1 to 6 in a tranquil Regime 1/2 in the two regime model. Also note that no sector shows negative number at all. At the turbulent Regime 2/2, reversal is found in 10 sectors of which 4 sectors are negatively significant. This finding implies that sector momentum is regime dependent. The momentum is more significant in the tranquil regime than in the turbulent one when the momentum turns into reversal. In the three regime model, the tranquil Regime 1/3 and Regime 2/3 are fraught with significant positive momentum in 5 and 3 sectors, respectively. Only "Other" is a common positive momentum sector to both regimes, and no significant negative number is found in any sector. In the turbulent Regime 3/3, 4 sectors are significantly negative and nothing is significantly positive. Regime 1/3 and Regime 2/3 are relatively closer to Regime 1/2 and Regime 3/3 is close to Regime 2/2. For a third model parameter, Table

Table 3.4: Diagonal elements of $\boldsymbol{L}_i$ (Full sample:1927/07-2013/06)

	Single Regime	Two Regimes		Three Regimes		
	Regime 1/1	Regime 1/2	Regime 2/2	Regime 1/3	Regime 2/3	Regime 3/3
NoDur	0.38*** (3.38)	0.28*** (2.48)	0.07 (0.19)	0.24** (1.70)	−0.06 (−0.24)	0.26 (0.61)
Durbl	0.21** (1.69)	0.05 (0.45)	0.01 (0.02)	−0.05 (−0.33)	0.39** (1.71)	0.32 (0.67)
Manuf	0.10 (0.78)	0.25** (1.98)	−0.86** (−2.14)	0.54*** (2.97)	−0.37 (−1.27)	−0.93** (−2.07)
Enrgy	−0.02 (−0.14)	0.27** (1.75)	−0.83** (−1.84)	−0.01 (−0.06)	−0.29 (−0.85)	−1.47*** (−2.70)
Chems	0.15 (1.27)	0.13 (1.16)	−0.09 (−0.23)	0.12 (0.85)	0.08 (0.31)	−0.05 (−0.12)
BusEq	0.03 (0.26)	0.03 (0.26)	−0.32 (−0.81)	−0.12 (−0.72)	−0.22 (−0.61)	−0.30 (−0.67)
Telcm	0.26** (2.272)	0.29*** (2.42)	−0.10 (−0.27)	0.36** (2.14)	0.11 (0.57)	0.09 (0.21)
Utils	0.04 (0.33)	0.05 (0.43)	−0.02 (−0.06)	−0.01 (−0.04)	0.17 (0.75)	−0.20 (−0.58)
Shops	−0.08 (−0.70)	0.03 (0.21)	−0.45* (−1.40)	−0.13 (−0.75)	−0.15 (−0.62)	−0.47 (−1.30)
Hlth	0.25** (2.17)	0.31*** (2.73)	−0.48 (−1.17)	0.42*** (3.00)	−0.33 (−1.24)	−0.53 (−1.18)
Money	0.14* (1.41)	0.14* (1.45)	−0.13 (−0.49)	0.05 (0.40)	1.74*** (5.78)	−0.47* (−1.60)
Other	−0.14 (−1.17)	0.14 (1.15)	−1.12*** (−2.99)	0.23* (1.46)	−0.40** (−1.70)	−1.15*** (−2.68)

*, **, *** Significant respectively at 10%, 5%, 1% levels. t-statistics in parenthesis.

3.5 tabulates the beta to the market average $\boldsymbol{\beta}_i$ in (3.4). We see low beta sectors and high beta sectors across estimated three models. Two sectors "Telcm" and "Utils" distinguish between Regime 1/3 and Regime 2/3, i.e., "Telcm" increases from 0.57 in Regime 1/3 to 0.90 in Regime 2/3 while "Utils" decreases from 0.74 to 0.53. These figures are not within those in the two regime model. This is one of plausible reasons to split Regime 2/3 from Regime 1/3. Summarizing above, the two regime model identifies clearly opposite two

Table 3.5: Market beta $\boldsymbol{\beta}_i$ (Full sample:1927/07-2013/06)

	Single Regime	Two Regimes		Three Regimes		
	Regime 1/1	Regime 1/2	Regime 2/2	Regime 1/3	Regime 2/3	Regime 3/3
NoDur	0.76	0.87	0.69	0.85	0.86	0.70
Durbl	1.24	1.12	1.32	1.12	1.29	1.31
Manuf	1.20	1.15	1.24	1.16	1.15	1.23
Enrgy	0.86	0.88	0.84	0.92	0.82	0.83
Chems	0.98	1.00	0.97	1.01	0.95	0.97
BusEq	1.28	1.19	1.33	1.18	1.18	1.34
Telcm	0.65	0.63	0.66	0.57	0.90	0.65
Utils	0.78	0.68	0.85	0.74	0.53	0.85
Shops	0.97	1.03	0.94	1.02	0.99	0.96
Hlth	0.84	0.97	0.79	0.95	0.93	0.80
Money	1.16	1.07	1.24	1.06	1.10	1.24
Other	1.13	1.14	1.13	1.17	1.04	1.13

regimes, i.e., Regime 1/2, is less cross-sectionally volatile and more positive momentum than Regime 2/2. Three regime model identifies Regime 3/3 as similar to Regime 2/2, Regime 1/3 is close to Regime 1/2 and Regime 2/3 resides in between.

For further discussions, we will move onto transition probability matrices and profiles of smoothed probabilities. Transition probabilities between regimes in the two regime model are

$$\boldsymbol{P} = \begin{bmatrix} .937 & .063 \\ .218 & .782 \end{bmatrix}. \tag{3.6}$$

On average, Regime 1/2 continues $1/(1 - .937) = 15.8$ months and Regime 2/2 continues only $1/(1 - .782) = 1.3$ month. In the three regime model,

$$\boldsymbol{P} = \begin{bmatrix} .935 & .012 & .054 \\ .015 & .958 & .027 \\ .157 & .028 & .815 \end{bmatrix}. \tag{3.7}$$

Similar to the two regime model, the turbulent Regime 3/3 is short lived relative to other tranquil Regime 1/3 and Regime 2/3.

Figure 3.4 shows the time series of the smoothed probabilities in the two regime model and similarly Figure 3.5 in the three regime model. Over 86 years, Figures 3.4 and 3.5 both depict estimated regime probabilities in most of the entire period are close to 0 or 1. The binary behavior of the observed profiles is a supportive evidence to assume regime switching structure in the studied sector returns. This is also important for applying it to regime aware investment decisions. The two regime model observes the turbulent Regime 2/2 for roughly 5 years around 1930s for the Great Depression followed by the aftermath until 1940.

Next periods at Regime 2/2 take place from 1998 until 2003 for a growth and meltdown of the Technology Bubble and the Lehman Shock from 2008 to 2009 as well as some other short-lived ones. The profiles in the three regime model show two notable differences form

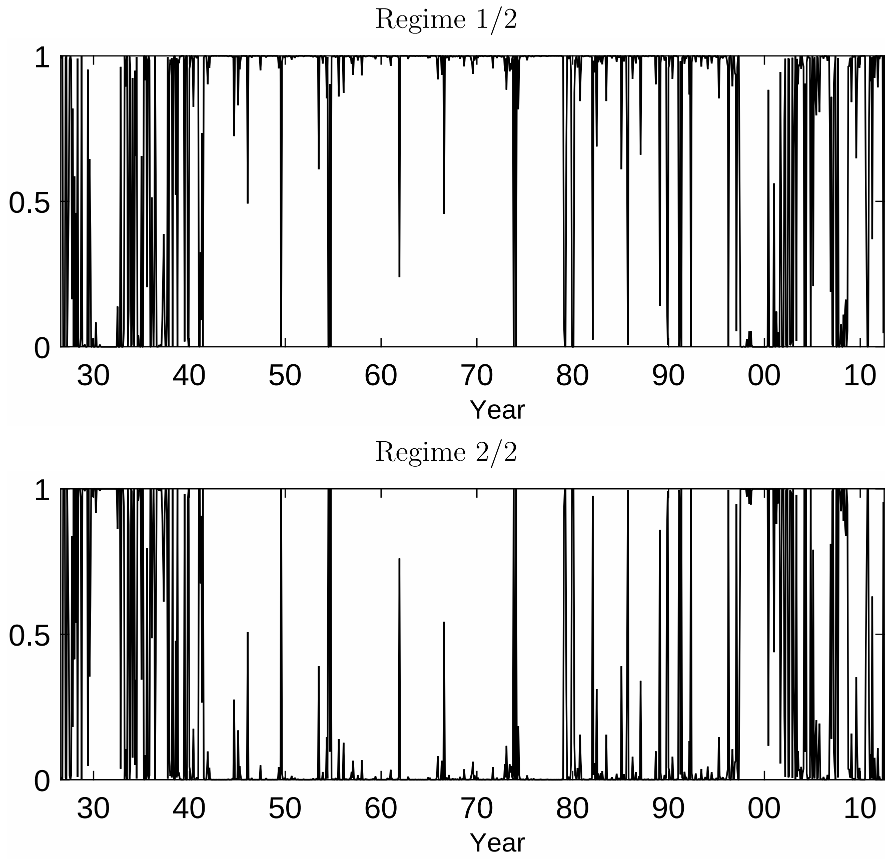


Figure 3.4: Smoothed probabilities in the two regime model (Full sample:1927/07-2013/06)

those in the two regime model. First, Regime 1/3 and Regime 2/3 are decomposed into two halves, split at early in 1980s, of Regime 1/2. Second, Regime 3/3 identifies the turbulent periods more clearly than Regime 2/2. For a possible argument on the split at early 1980s, remember "Durbl" and "Enrgy" displaying higher cross sectional variance in Regime 2/3 than in Regime 1/3 as well as changes in the market beta for "Telcm" and "Utils". In the early 1980s, the US interest rates in both ends of the Treasury curve peaked out around 15% after a long lasted upward trend followed by a secular downturn in a recent quarter century. Behind the change in the trend, the US economy experienced stagflation triggered by the oil crisis. Among number of seminal work in the economics on this major the economic epoch around 1980s, durables and energy are two sectors that are often focused. For example, Bernanke (1983) proposes a theoretical basis to explain why price changes in energy induce consumers postponing purchases of consumer durables that are irreversible. Hamilton (1988), Lee and Ni (2002) as well as Eraker, Shaliastovich and Wang (2012) also pick up the two sectors. "Durbl" and "Enrgy" were inherently sensitive to the inflationary environment in 1970s. Throughout the Reaganomics and the tightening policy led by Volcker at the Federal Reserve Bank, the US economy got out of the inflationary state. The fiscal and monetary regime shift might have driven the switch from Regime 1/3 to Regime 2/3 in our model. Around the same time, "Utils" turned from an outperformer into an underperformer while "Telcm" from an underperformer into an outperformer, changing the market beta as observed in this study. Over the 86 year period, both of two and three regime models commonly distinguish a turbulent state in sector returns. The three regime model uniquely identifies two different tranquil regimes that the two regime model does not resolve. It is noteworthy for the 86 year period that regime shifts in the monetary and the fiscal policies in early 1980s are the second most notable event to many famous economical and market turmoils in configuring the three regime structure for the proposing sector return forecasting model. Observed regime probabilities provide us with intuitive insights to understand the significant difference in the investment efficacy. Revisiting to smoothed probabilities of the two and three regime models in Figures 3.4 and 3.5 respectively, one can notice that the two regime model stays in Regime 1/2 in 1990s until 1998 while the three regime model already switched from Regime 1/3 to Regime 2/3 in the middle of 1980s. It is presumed that the two regime model loses predicting ability due to the changes in regimes. From the late in 1990s to early in 2000s, both of the two and the three regime models agree that the market stays in the turbulent state, represented by Regime 2/2 in the two regime model and Regime 3/3 in the three regime model. During this period, the technology bubble grows and bursts widening inter-sectorial variabilities of return significantly. Once the technology sector meltdown fades away, the two regime model comes back to Regime 1/2 and the three regime returns to Regime 2/3. As reviewed, those two regimes are identified to have notable differences in the factor loading $\boldsymbol{L}_i$ and $\boldsymbol{W}_i$. Remember that the factor loading $\boldsymbol{L}_1$ in the two regime model claims that 6 sectors enjoy positive own momentum and none in own reversal. And $\boldsymbol{L}_2$ in the three regime model supports 3 sector in own momentum and none in own reversal.

3.4 Investment Performance of Sector Rotation Strategy

In this section, we discuss investment efficacy of optimal portfolios for the sector rotation strategy. An out-of-sample period for evaluating portfolio performances starts at July 1976 which is 601^{th} month since the beginning of entire data period. Toward the end of the entire period at June 2013, 444 months (37 years) are available for out-of-sample examination. During the out-of-sample period, all of model parameters are re-estimated every 3 months. For those months in between the re-estimations, we keep taking over previously estimated parameters until next re-estimation is conducted. As an example, Figure 3.6 exhibits historical evolutions of three parameters $\boldsymbol{\beta}_m$, $\boldsymbol{L}$, and $\boldsymbol{W}$ in the three regime model for NoDur sector as one of 12 sectors. The profiles show potential evidence

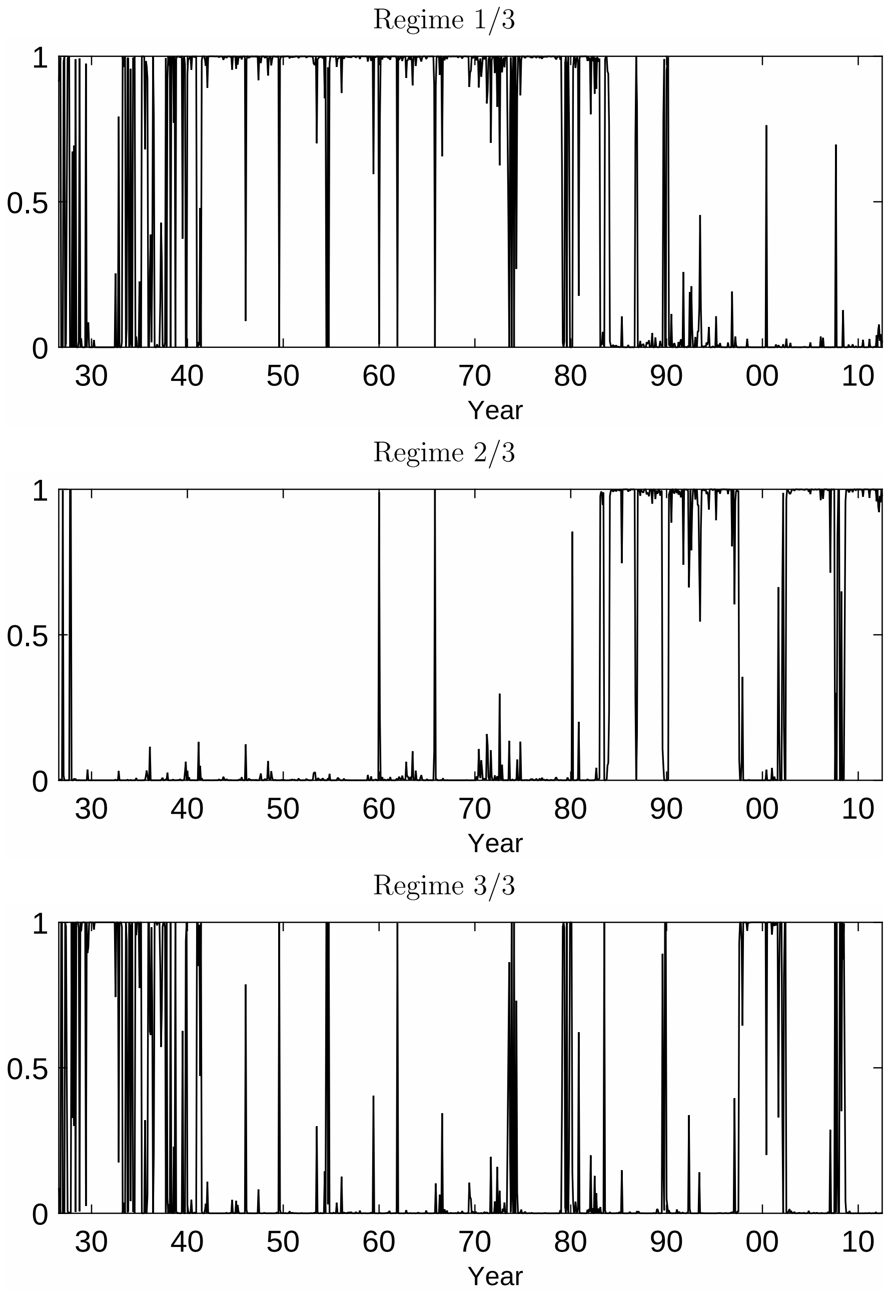


Figure 3.5: Smoothed probabilities in the three regime model (Full sample:1927/07-2013/06)

of time variant nature of the parameters even letting the model being regime dependent. This implies that the best model parameters including optimal number of regimes can depend on sampling periods. The choice of the filtering likelihood in the model estimation may be a cause of the time variant nature. Admitting room for better model specifications than we employ, in the scope of our study, we choose to specify the model to be regime dependent and model parameters are not explicitly time variant. In order to accommodate to the potential time variant nature, we keep re-estimating the model periodically in the out-of-sample period for portfolio optimizations. This is one of practical approaches to aim at decent investment performance by keep tuning model parameters to the latest. As a tool for building optimal portfolios and measurements of investment performance, we use the MATLAB running in the Intel(R) Core(TM) i7-4960X CPU 3.60GHz 6 Cores 12 Threads under 64bit operating system with 8G byte memory.

Parallel to the quarterly model estimations, a monthly exercise for regime observation is maintained by importing realized monthly returns to 12 sectors, market average and risk free rate. In what follows, we use the same labeling convention over the observed regimes as explained in Section 3.3.

3.4.1 Mean-variance Optimization with Transaction Costs

For portfolio optimization, we employ a mean-variance utility function with a quadratic transaction cost under the regime switching market environment. At time t, an investor first observes factor $\boldsymbol{f}(t)$ and estimates filtered probabilities of the current regime i_t, and then rebalances previous portfolio $\boldsymbol{x}(t-1) = (x_1(t-1), \ldots, x_N(t-1))^\top$ to get new portfolio $\boldsymbol{x}(t) = (x_1(t), \ldots, x_N(t))^\top$ where $x_j(t)$ denotes amount invested to sector j. From (3.3), the excess return $y(t+1)$ of the portfolio $\boldsymbol{x}(t)$ between t and $t+1$ is given by

$$y(t+1) = \boldsymbol{x}^\top(t)\{\boldsymbol{L}_i\boldsymbol{f}(t) + \boldsymbol{u}_i(t+1)\} \tag{3.8}$$

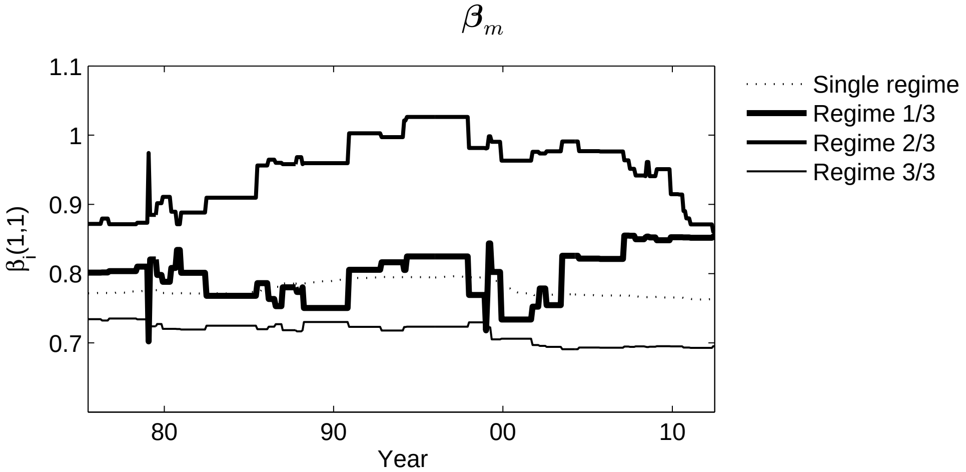
where i represents regime at $t+1$ which is uncertain at t. An investor is supposed to choose $\boldsymbol{x}(t)$ so as to maximize the mean-variance utility penalized by transaction cost:

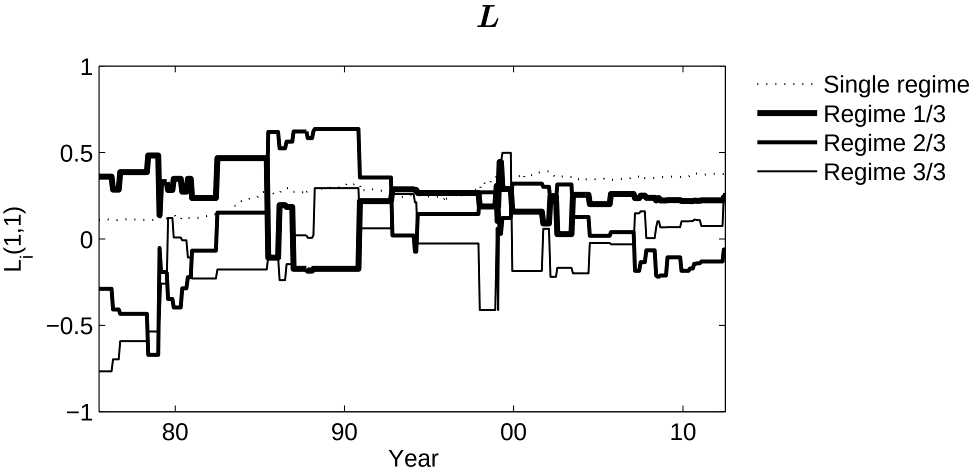
$$E_t[y(t+1)] - \frac{\lambda}{2}V_t[y(t+1)] - \frac{1}{2}\Delta\boldsymbol{x}^\top(t)E_t[\boldsymbol{B}_{I(t+1)}]\Delta\boldsymbol{x}(t) \tag{3.9}$$

where $\Delta\boldsymbol{x}(t) = \boldsymbol{x}(t) - \boldsymbol{x}(t-1)$ and E_t and V_t respectively expresses mean and variance conditioned on the information available at t. The first two terms in (3.9) represent ordinary mean-variance utility with a risk aversion coefficient λ. The third term denotes a quadratic cost function where $\boldsymbol{B}_i$ is an $N \times N$ positive definite matrix. Under no constraint bot allowing short positions, the first order optimality condition gives the optimal portfolio as follows:

$$\boldsymbol{x}^*(t) = [\lambda V_t[\boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t)+\boldsymbol{u}_{I(t+1)}(t+1)]+E_t[\boldsymbol{B}_{I(t+1)}]]^{-1}[E_t[\boldsymbol{L}_{I(t+1)}]\boldsymbol{f}(t)+E_t[\boldsymbol{B}_{I(t+1)}]\boldsymbol{x}(t-1)]. \tag{3.10}$$

Note that $I(t+1)$ an a regime at time $t+1$, which is not available yet at time t, appears in (3.8),(3.9) and (3.10) to derive an optimal solution $\boldsymbol{x}^*(t)$. E_t and V_t defined in Subsection





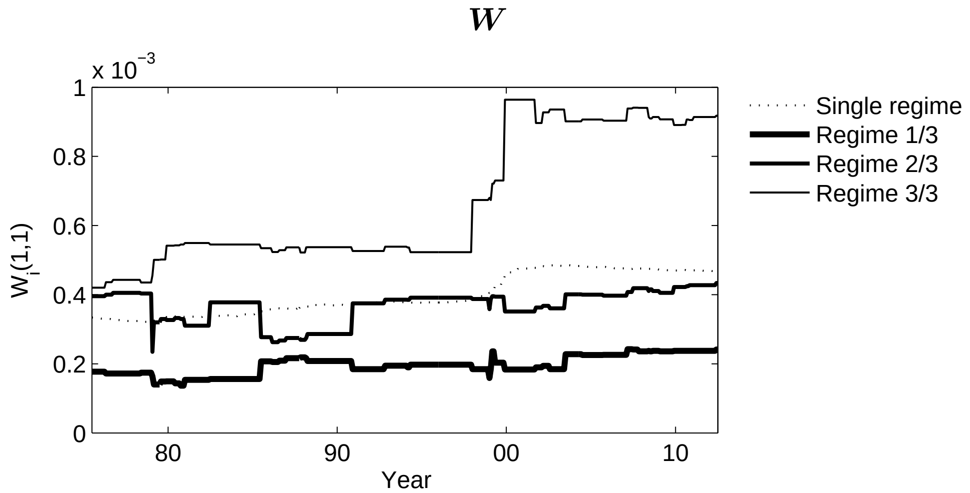


Figure 3.6: Profiles of parameters for NoDur sector: $\boldsymbol{\beta}_m$, factor loading to NoDur in $\boldsymbol{L}$, and variance in $\boldsymbol{W}$ under three regime model and single regime model (Out-of-sample:1976/07-2013/06).

3.6.2 calculate a mean and a variance using both of the estimated filtered probability $I(t)$ for each regime in the model as of time t and the transition matrix $\boldsymbol{P}$. Taking advantage of the Markov property, we predict variance of $\boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t)+\boldsymbol{u}_{I(t+1)}(t+1)$ and expectations to $\boldsymbol{L}_{I(t+1)}$ and $\boldsymbol{B}_{I(t+1)}$ one time step ahead $t+1$ for the optimal solution as of time t. See Subsection 3.6.2 for the derivation of (3.10) and the explicit expressions of the conditional expectations and variance.

On the transaction cost matrix $\boldsymbol{B}_i$ for calculating optimal portfolio, it is in generally difficult to observe realized transaction costs. Here we apply some of earlier studies such as Grinold and Kahn (1999), suggesting the transaction cost proportional to the variability of the security. As reported in Table 3.3, estimated variance of excess returns to each sector of beta adjusted market average exhibits significant regime dependency. In our empirical analysis, we set $\boldsymbol{B} = 0.01\boldsymbol{I}$ for the single regime model where $\boldsymbol{I}$ is an $N \times N$ identity matrix, $\boldsymbol{B}_1 = 0.007\boldsymbol{I}$ and $\boldsymbol{B}_2 = 0.015\boldsymbol{I}$ for the two regime model, and $\boldsymbol{B}_1 = \boldsymbol{B}_2 = 0.007\boldsymbol{I}$ and $\boldsymbol{B}_3 = 0.015\boldsymbol{I}$ for the three regime model.

Evaluating investment performance of the optimal portfolios, the transaction costs also matter because realized transaction costs do not necessarily coincide with what are assumed. In order to be mindful of inevitable uncertainty of the transaction costs, we examine three types of realized transaction costs in performance comparison. The first type assumes transaction costs $\Delta\boldsymbol{x}^{*\top}(t)\boldsymbol{B}_i\Delta\boldsymbol{x}^{*\top}/2$ as in the third term of (3.9) with $\boldsymbol{B}_i$'s given above. The second one assumes variable transaction costs depending upon a linear function of a realizing total return to each sector as $(0.0025 + 0.1\,|\boldsymbol{r}(t)|)^\top|\Delta\boldsymbol{x}^*(t)|$, meaning 25 basis points plus 10% of an absolute return to each of sector as a variable component. The third one combines the first two types the transaction costs of which are $(0.0025+0.1|\boldsymbol{r}(t)|)\Delta\boldsymbol{x}^{*\top}(t)\Delta\boldsymbol{x}^*(t)/2$. The first, second and third types are referred as Perfect foresight, Linear and Quadratic, respectively in Table 3.6. We take into account other inevitable costs for neither short sale nor fund expenses in ETF.

In the mean-variance framework, an investor needs to choose a risk aversion coefficient. In our experiments in Subsections 3.4.2 and 3.4.3, we fix $\lambda = 1$ for all models. A choice of the common and consistent λ to all regimes and over time is chosen in order to penalize neither the single nor multiple regime models for preserving fairness in the intended comparisons. Motivated by the observations in Subsection 3.4.3 that forecasting ability of sector returns shows strong regime dependency, we generalize a λ to be regime dependent in Subsection 3.4.4.

3.4.2 Performance Comparisons across Number of Regimes

In this subsection, we compare investment performances across number of regimes as well as types of transaction costs. Table 3.6 summarizes returns, volatilities and Sharpe ratio achieved by the optimal portfolio for the single, two and three regime models in the out-of-sample period when the initial position of the portfolio equals $\mathbf{0}$. As mentioned earlier in Subsection 3.4.1, the quadratic form is assumed to penalize mean-variance utility to maximize. In evaluating investment performance of the optimized portfolios, three types of realized transaction costs are applied to deduct from gross results. As a control experiment, we show those performance before netting transaction costs at the upper rows in Table 3.6. Regardless of the realized transaction costs, the more number in regimes, the higher returns and volatilities in dollar. As to uncertainty of the transaction costs, we confirm a robustness of the superior investment efficacy with regime dependent models. Because of accelerating improvement in the returns while decelerating increase in volatilities along with increase of regimes, Sharpe ratios improve as the number of regimes increases. For the entire period, Subsection 3.3.2 reports that AIC improves as number of regimes increases. The optimal portfolio performance summarized in Table 3.6 for the out-of-sample period also improves as number of regimes increases. In general, no exact relationship is available between degree of model fitness and optimal investment performance applying the model. That is obviously because of a difference in objective function to maximize. The fact the more number of regimes the better investment performance suggests that the both objectives may share a certain level of similarity each other.

Table 3.6: Optimal portfolio performance (Out-of-sample:1976/07-2013/06)

Transaction cost	Model	Dollar return	Dollar volatilities	Sharpe ratio
Before netting transaction costs	Single regime	0.44	1.81	0.24
	Two regimes	0.97	2.30	0.42
	Three regimes	1.74	2.32	0.75
Perfect foresight	Single regime	0.34	1.81	0.19
	Two regimes	0.73	2.30	0.32
	Three regimes	1.40	2.32	0.60
Linear	Single regime	0.17	1.81	0.09
	Two regimes	0.57	2.30	0.25
	Three regimes	1.22	2.32	0.53
Quadratic	Single regime	0.36	1.81	0.20
	Two regimes	0.81	2.30	0.35
	Three regimes	1.48	2.31	0.64

Annualized monthly figures.

Understanding why significant differences in the investment efficacy are realized, Figure 3.7 depicts historical profiles of the out-of-sample investment performance for single, two and three regime models when the realized transaction costs take the form of Quadratic in Table 3.6. The single regime model performs consistently poorly throughout the out-of-sample period. The two regime model performs as decent as the three regime model up until early 1990s and immediately underperforms until late 1990s for several years. From 1998 to 2001 the two regime model catches up with the three regime model and again it stops positive performance toward the end of the tested period. On the other hand, the three regime model continues to perform consistently throughout the entire period.

Observed regime probabilities provide us with intuitive insights to understand the significant difference in the investment efficacy. Revisiting to smoothed probabilities of the two and three regime models in Figures 3.4 and 3.5 respectively, one can notice that the two regime model stays in Regime 1/2 in 1990s until 1998 while the three regime model already switched from Regime 1/3 to Regime 2/3 in the middle of 1980s. It is presumed that the two regime model loses predicting ability due to the changes in regimes. From the late in 1990s to early in 2000s, both of the two and three regime models agree that the market stays in the turbulent state, represented by Regime 2/2 in the two regime model and Regime 3/3 in the three regime model. During this period, the technology bubble grows and bursts widening inter-sectorial variabilities of return significantly. Once the technology sector meltdown fades away, the two regime model comes back to Regime 1/2 and the three regime returns to Regime 2/3. As reviewed in Section 3.3, those two models are identified to have notable differences in the factor loading $\boldsymbol{L}_i$ and $\boldsymbol{W}_i$. Remember that the factor loading $\boldsymbol{L}_1$ in the two regime model claims 6 sectors enjoy positive own momentum and none in own reversal. And $\boldsymbol{L}_2$ in the three regime model supports 3 sector in own momentum and none in own reversal.

To quantitatively investigate the superior performance of the three regime model, Table 3.7 summarizes a cross tabulation of the number of months classified into each regime in the two and three regime models when the regime of each month is identified as that with the highest filtered probability. We observe the two models agree to classify the state of each month into tranquil or turbulent in that only 8 out of total 443 months fall in Regime 1/2 and Regime 3/3 and only 1 month falls in Regime 2/2 and Regime 1/3. On the other hand, Regime 2/3 splits into Regime 1/2 for 241 months and into Regime 2/2 for 48 months. In terms of the observed series of the regimes, this means that the three regime model differs from the two regime model in Regime 2/3 belonging to both of Regime 1/2 and Regime 2/2. Because of a tighter limitation of the number of regimes, the two regime model does not distinguish Regime 2/3 but attempts to replicate it by a combination of Regime 1/2 and Regime 2/2. The cross tabulation applies to realized gross returns and Quadratic transaction costs generated by the two and three regime models are summarized in Tables 3.8 and 3.9, respectively. In a row for Regime 2/3 in Table 3.8, it is notable that the realized returns by the two regime model are worse than those by the three regime model, especially in Regime 2/2. We make sure in Table 3.9 that the two regime model pays less transaction costs than the three regime model. This implies that the return forecasting ability matters rather than the transaction costs to discuss the performance difference. In sum, we confirm that major reason why the three regime model performs better than the two regime model is that the increasing number of regimes in the market is overflown from the two regime model. The three regime model continues to capture the regimes over 86 years in the expanding model estimation windows.

In order to see how three models accommodate to structural changes captured throughout the market beta, factor loading and unpredictable noise in the multivariate framework, we show how model fitness differs across number of regimes in terms of Akaike's Information Criterion (AIC). Figure 3.8 plots a scatter AIC of a two regime model versus to a three regime both subtracted by AIC of a single regime model. Negative numbers on both axis mean that both of the two and the three regime models better fit than the single regime model. The plots distribute in a lower region than a 45° line with no exception. As such the more number of regimes, the better fitness. Awareness of regimes assumed in (3.4) and

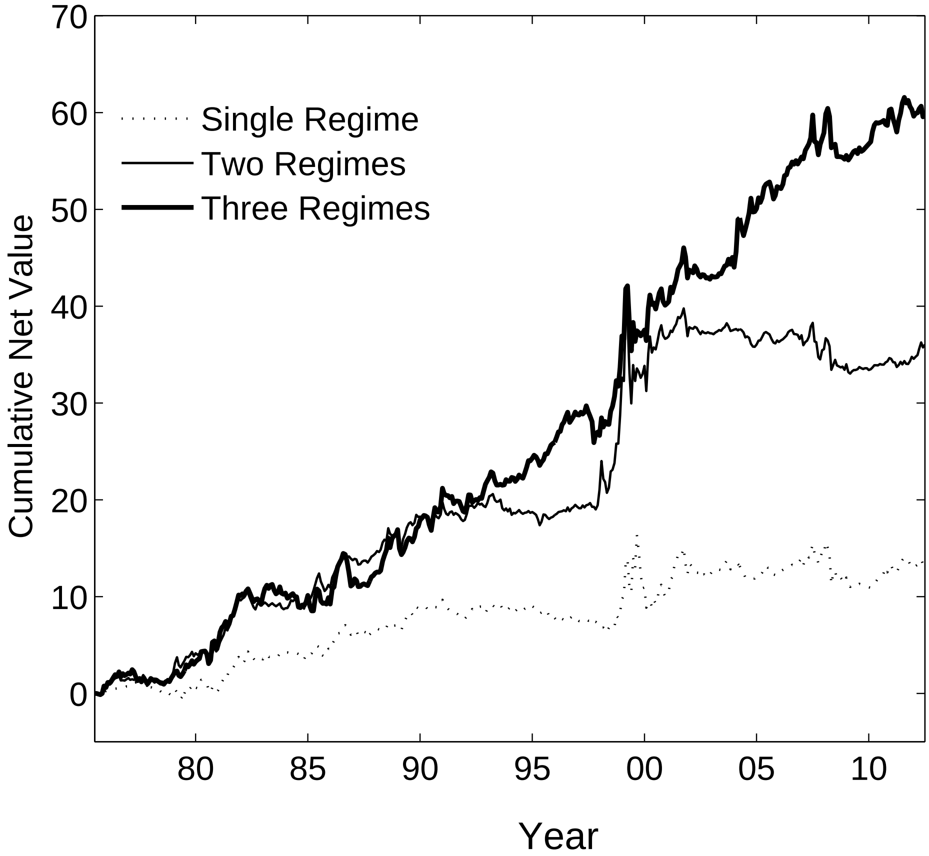


Figure 3.7: Cumulative net value of the optimal portfolios for realized Quadratic transaction costs (Out-of-sample:1976/07-2013/06).

Table 3.7: Cross tabulation (1): Number of months (Out-of-sample:1976/6-2013/6)

	Regime 1/2	Regime 2/2
Regime 1/3	58	1
Regime 2/3	241	48
Regime 3/3	8	87

Regime defined as that with the highest filtered probabilities.

Table 3.8: Cross tabulation (2) : Mean realized gross returns (Out-of-sample:1976/6-2013/6)

	by Two regime model		by Three regime model	
	Regime 1/2	Regime 2/2	Regime 1/2	Regime 2/2
Regime 1/3	0.106	0.501	0.078	0.286
Regime 2/3	0.093	0.003	0.186	0.173
Regime 3/3	0.129	0.170	0.349	0.155

Table 3.9: Cross tabulation (3) : Mean realized transaction costs (Out-of-sample:1976/6-2013/6)

	by Two regime model		by Three regime model	
	Regime 1/2	Regime 2/2	Regime 1/2	Regime 2/2
Regime 1/3	0.0115	0.0155	0.0231	0.0349
Regime 2/3	0.0149	0.0196	0.0276	0.0302
Regime 3/3	0.0272	0.0404	0.0339	0.0529

Figures in this table are from the case with Quadratic transaction costs in Table 3.6.

(3.5) adds value to the sector rotation strategy in the out-of-sample period.

3.4.3 Forecasting Ability across Regimes

In the model shown in (3.4) and (3.5) to estimate for predicting multivariate sector returns, we assume the unpredictable noises $\boldsymbol{u}_i(t)$ follows a normal distribution. If actual residuals do not satisfy this assumption, estimated parameter are inevitably biased and sector return forecasts are consequently misled. Table 3.10 summarizes the Jarque-Bera test to see if $\boldsymbol{u}_i(t)$ follows a normal distribution in the out-of-sample examination. Worth to mention for the single regime model is that none of 12 sectors follows a normal distribution. It is not surprising because number of earlier studies report that majority of financial time series exhibit fat tails that annoy investors especially for left tails. The two regime model significantly reduces non-Gaussian property with 8 sectors and 7 sectors following normal distribution in Regime 1/2 and Regime 2/2, respectively. As expected, the three regime model further mitigates chance of non-normality, especially in Regime 3/3. The more number of regimes, the less chance for biased model estimation. This is one of reasons why the three regime model performs the best among three models. Going beyond the normal distributions, a great deal of effort in finance research has been going to overcome obviously serious issues of the complicated behavior of financial returns. For example, Chollete, Heinen and Valdesogo (2009) proposes importance choice of copulas under regime dependent framework especially in risk management of international portfolio measured in value at risk. As long as our test results suggest, the normal distributions have much room

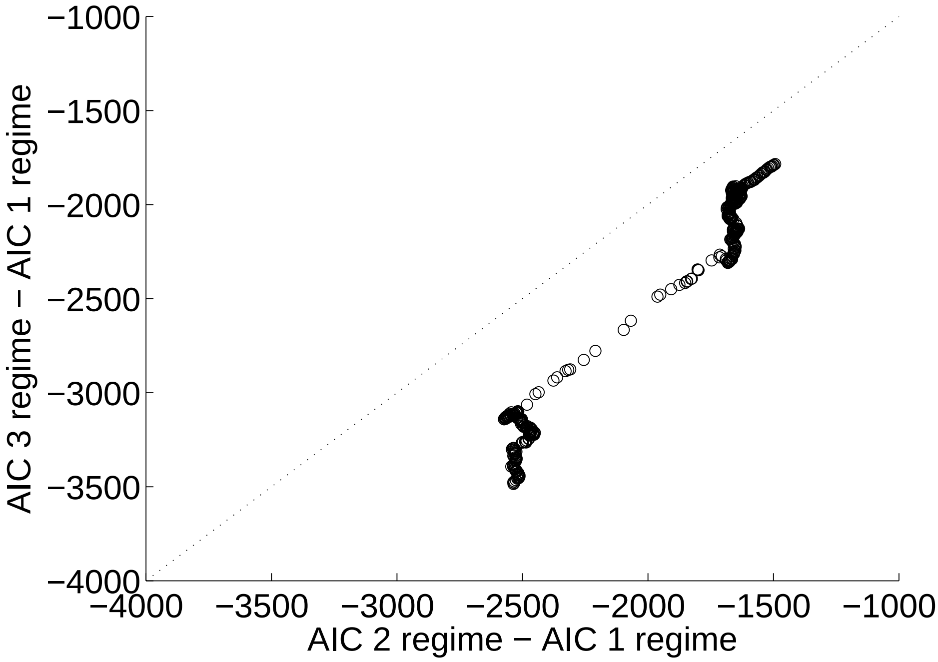


Figure 3.8: AIC comparisons over three regime and two regime both in differences from the single regime (Out-of-sample:1976/06-2013/06)

to add value if incorporated into the regime switching time series structure. A primary source of the added value to the sector rotation strategy owes to the awareness of regimes assumed in (3.4) and (3.5) to prevent from investment performance degradation due to a potential model misspecification.

Table 3.10: Jarque-Bera test (Out-of-sample:1976/6-2013/6)

Single regime	Two regimes		Three regimes		
Regime 1/1	Regime 1/2	Regime 2/2	Regime 1/3	Regime 2/3	Regime 3/3
0	8	7	8	9	10

Out of all 12 sectors, this table shows the number of sectors for which the null hypothesis that the residuals are normally distributed is not rejected at 5% significance level.

Next, a goodness-of-fit is of interest to understand if any difference in forecasting ability of sector returns across regimes. The root mean squared errors (RMSE) of return forecasts are tabulated in Table 3.11 to compare across 12 sectors and periods each of regimes dominates as well as the entire out-of-sample period in the three models. Without any exception in multiple regime models, the regimes for turbulent states entail higher forecasting errors than those for tranquil states. This is not only obviously because of inherently high risk nature of the regimes for turbulent state but also potentially because of possible estimation errors in return variabilities. Also noteworthy that the more number of regimes the RMSE tends to be lower.

Table 3.11: RMSE (Out-of-sample:1976/6-2013/6)

	Single regime	Two regimes			Three regimes			
	Regime 1/1	Regime 1/2	Regime 2/2	Overall	Regime 1/3	Regime 2/3	Regime 3/3	Overall
NoDur	.0257	.0199	.0342	.0252	.0158	.0207	.0380	.0250
Durbl	.0416	.0304	.0599	.0417	.0260	.0331	.0632	.0407
Manuf	.0215	.0155	.0318	.0218	.0136	.0169	.0360	.0221
Enrgy	.0437	.0348	.0592	.0438	.0302	.0386	.0609	.0435
Chems	.0260	.0177	.0383	.0258	.0150	.0193	.0431	.0259
BusEq	.0353	.0286	.0457	.0347	.0264	.0289	.0516	.0348
Telcm	.0336	.0279	.0453	.0342	.0235	.0275	.0519	.0338
Utils	.0378	.0283	.0539	.0380	.0241	.0280	.0607	.0371
Shops	.0267	.0234	.0319	.0263	.0227	.0240	.0347	.0265
Hlth	.0314	.0253	.0412	.0311	.0289	.0259	.0436	.0310
Money	.0284	.0189	.0414	.0278	.0169	.0218	.0463	.0285
Other	.0191	.0172	.0232	.0193	.0143	.0175	.0273	.0197

3.4.4 Regime Dependent Risk Aversion

In Subsection 3.4.3, we review why multiple regime models perform absolutely better than a conventional single regime model. We also find intuitive sources for a difference in forecasting power across regimes in multiple regime models, i.e., forecasting errors annoy the turbulent regimes. Under the mean-variance utility, this finding is not necessarily a bad news at all if an investor is risk averse. In this subsection, we assume that the investor becomes more risk averse if a turbulent regime is predicted. As in the other parameters, the risk aversion coefficient λ_i is now regime dependent where i denotes one step ahead regime forecast. Under the set-up, the optimal portfolio (3.10) is modified by replacing λ with the conditional expectation of the risk aversion coefficient $\sum_j \widehat{q}_j(t+1)\lambda_j$ where $\widehat{q}_j(t+1)$ represents estimated probability that the regime is j at $t+1$. See Subsection 3.6.3 for calculation of $\widehat{q}_j(t+1)$.

Table 3.12 compares investment performances when the investor chooses higher risk aversion coefficients for Regime 2/2 in the two regime model and for Regime 3/3 in the three regime model while leaving the coefficients for tranquil regimes as 1. Among all in each of the multiple regime models, the net Sharpe ratios attain the highest when the investor chooses $\lambda_2 = 10$ in the two regime model and $\lambda_3 = 3$ and 5 in the three regime model for the turbulent regimes. Also note that depth of the maximum draw down for the three regime model are mitigated to \$3.32 and \$3.12 for $\lambda_3 = 3$ and 5, respectively, vis-a-vis \$5.25 for $\lambda_3 = 1$.

Figure 3.9 picks up and draws one of the experimentations reported in Table 3.12 when the investor chooses 5 for the turbulent regimes, i.e., λ_2 in the two regime and λ_3 in the three regime model. Comparing with Figure 3.7 when the investor always chooses 1, the regime dependent risk aversion delivers to a risk averse investor more efficient investment results than who is indifferent to regimes. In a recent decade when many of investors and investment managers suffered from the credit crisis, the regime dependent risk aversion delivers more consistent result than shown in Figure 3.7 to the investor who is aware of the regimes for the sector rotation strategy.

Table 3.12: Optimal portfolio performance under regime dependent risk aversion λ (Out-of-sample:1976/07-2013/06)

Model	λ	Gross dollar return ($)	Net dollar return ($)	Dollar vols ($)	Gross Sharpe	Net Sharpe	Maximum depth of dollar draw down ($) (when)	Maximum length of dollar draw down (months) (starts:ends)
Single regime	1	0.44	0.37	1.81	0.24	0.20	7.49 (01/01)	154 (00/08:Present)
Two regimes	1/1	0.97	0.81	2.30	0.42	0.35	8.00 (00/05)	129 (02/09:Present)
λ_1/λ_2	1/1.5	0.93	0.79	1.99	0.47	0.39	6.51 (00/05)	129 (02/09:Present)
	1/2	0.88	0.74	1.76	0.50	0.42	5.41 (00/05)	129 (02/09:Present)
	1/3	0.78	0.66	1.44	0.54	0.46	3.93 (00/05)	129 (02/09:Present)
	1/5	0.64	0.53	1.09	0.59	0.48	2.42 (00/05)	129 (02/09:Present)
	1/10	0.48	0.37	0.74	0.64	0.50	1.74 (10/02)	106(04/08:Present)
	1/50	0.20	0.14	0.31	0.67	0.46	0.82 (10/02)	99 (05/03:Present)
	1/100	0.13	0.09	0.20	0.65	0.45	0.48 (10/02)	71 (07/07:Present)
Three regimes	1/1/1	1.74	1.48	2.32	0.75	0.64	5.25 (00/05)	32 (09/02:11/10)
$\lambda_1/\lambda_2/\lambda_3$	1/1/1.5	1.67	1.42	2.06	0.82	0.69	4.07 (00/05)	29 (09/02:11/07)
	1/1/2	1.61	1.36	1.88	0.85	0.72	3.46 (08/09)	35 (08/06:11/05)
	1/1/3	1.49	1.25	1.67	0.90	0.75	3.32 (08/09)	32 (08/06:11/02)
	1/1/5	1.32	1.08	1.45	0.92	0.75	3.12 (08/10)	32 (08/06:11/02)
	1/1/10	1.10	0.86	1.22	0.91	0.71	2.92 (08/10)	32 (08/06:11/02)
	1/1/50	0.67	0.47	0.77	0.89	0.61	2.04 (98/10)	77 (98/05:04/10)
	1/1/100	0.49	0.33	0.59	0.86	0.56	1.89 (98/10)	80 (98/05:04/10)

Annualized monthly figures.

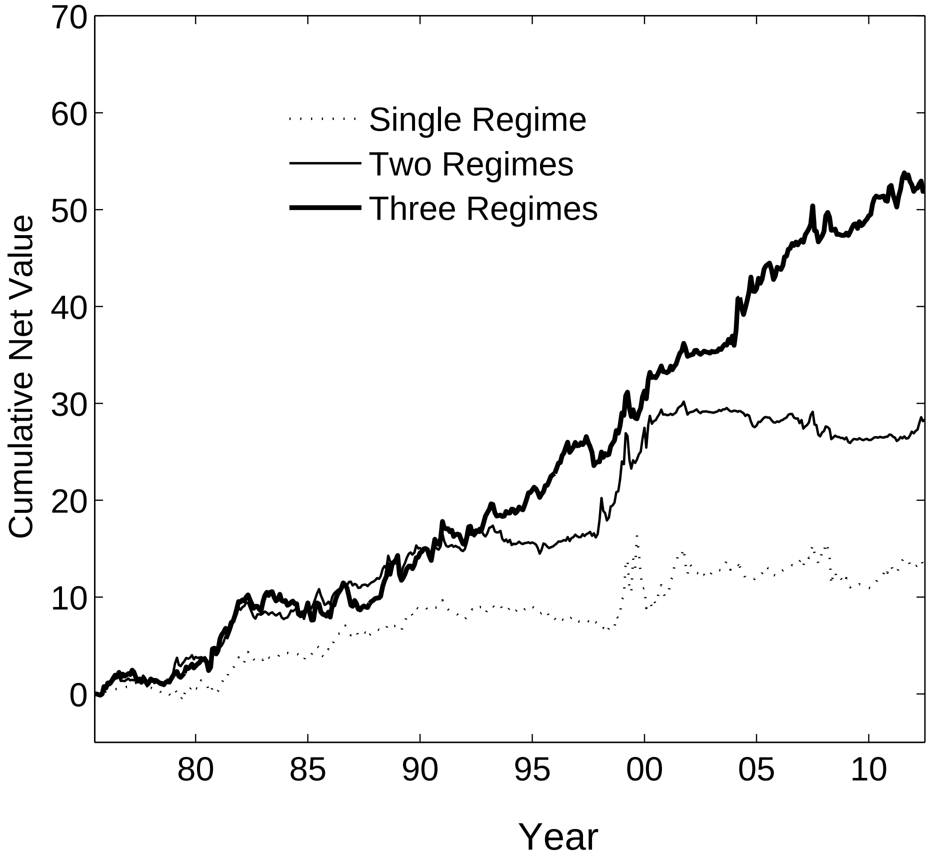


Figure 3.9: Cumulative net value of the optimal portfolios for regime dependent risk aversion coefficients and realized Quadratic costs (Out-of-sample:1976/07-2013/06).

3.5 Conclusions

Under increasing popularity of regime switching framework for investment management, this chapter extends studies of industry momentum documented by Moskowitz and Grinblatt (1999) and others to incorporate a myopic mean-variance optimal solutions bearing transaction costs. Our findings in this chapter are mainly empirical and of interest to a wide array of investors, ranging from hedge funds to managers and investors of the alternative risk premium.

Given the long period of monthly data for 86 years in-sample analyses and 37 years out of sample performance measurements, we document that momentum in 12 sectors are regime dependent and the three regime model consistently outperforms in the out-of-sample evaluation. This could potentially supersede most of conventional models that encounter difficult experiences in turbulent periods as Daniel and Moskowitz (2013) suggests.

Making decisions in unprecedented circumstances is always difficult for any sorts of investment management. Interesting observation of this study is that the regime switching assumption may reasonably improve response capabilities to discontinuous changes in return generating processes only with simple and parsimonious underlying models as proposed in this chapter. It is likely true, however, that practitioner should pay cautious attentions to identify sufficient number of regimes for long period of life of the model.

Studies in this chapter have several limitations. First, we cannot fully explain what are root causes of the regime switches that are observed in a certain point in the long sample although that fairly improves fitting data and consequent results in out-of-sample investment performance. Second, we do caveat that one can be caught in a trap of the curse of dimensionality and over-parameterization if the number of assets and regimes increase. That can potentially cause unstable estimation of model parameters. To this end, we are still long way from that future when all of these potential issues are clearly addressed and solved.

Going forward, additional research is planned on a couple of fronts to extend the model applied to the empirical studies. A first extension of the model for practical importance include more general constraints that prohibit short positions held in the portfolio under multi-period optimal solutions. This effort would extend a scope and the implications of this chapter to more general practices of investment management such as dynamic asset allocation decisions and implementations for CIOs in institutional investors. Another direction of extension can relax the transition probabilities from fixed to time variant as described by Markov switching logistic function of such exogenous variables as market data and macro economic data. This extension potentially helps us to understand the root causes of the regime switches.

3.6 Appendix

3.6.1 List of 12 Sectors

A list of 12 sectors is tabulated in Table 3.13.

3.6.2 Algorithm for Regime Estimation

In general, either one of following three algorithm is applied to estimate a regime. Hence, all information available from the beginning to time t is given as $\mathcal{F}_{1:t} = \{\mathcal{F}_1, \cdots, \mathcal{F}_t\}$ when $\mathcal{F}_t$ stands for dataset at time t. T is time at the end of the entire dataset.
⟨One step ahead prediction⟩

Table 3.13: List of 12 Industries

Short Name	Long Name	
NoDur	Consumer Non Durables	Food, Tobacco, Textiles, Apparel, Leather and Toys
Durbl	Consumer Durables	Cars, TV's, Furniture and Household Appliances
Manuf	Manufacturing	Machinery, Trucks, Planes, Off Furn, Paper and Com Printing
Enrgy	Energy	Oil, Gas, and Coal Extraction and Products
Chems	Chemicals	Chemicals and Allied Products
BusEq	Business Equipment	Computers, Software and Electronic Equipment
Telcm	Telecommunications	Telephone and Television Transmission
Utils	Utilities	
Shops	Shops	Wholesale, Retail and Some Services (Laundries, Repair Shops)
Hlth	Health Care	Health Care, Medical Equipment and Drugs
Money	Finance	
Other	Other	Mines, Constr, BldMt, Trans, Hotels, Bus Serv and Entertainment

Source: Detail for 12 Industry Portfolios, U.S. Research Returns Data, Kenneth R. French - Data Library

$$\begin{aligned}
\mathrm{P}\left(I(t+1)=i \mid \mathcal{F}_{1:t}\right) &= \sum_{j=1}^{k} \mathrm{P}\left(I(t+1)=i, I(t)=j \mid \mathcal{F}_{1:t}\right) \\
&= \sum_{j=1}^{k} \mathrm{P}\left(I(t+1)=i \mid I(t)=j\right) \mathrm{P}\left(I(t)=j \mid \mathcal{F}_{1:t}\right) \\
&= \sum_{j=1}^{k} p_{ij} \mathrm{P}\left(I(t)=j \mid \mathcal{F}_{1:t}\right)
\end{aligned}$$

⟨Filtering⟩

$$\begin{aligned}
\mathrm{P}\left(I(t+1)=i \mid \mathcal{F}_{1:t+1}\right) &= \mathrm{P}\left(I(t+1)=j \mid \mathcal{F}_{t+1}, \mathcal{F}_{1:t}\right) \\
&= \frac{\mathrm{P}\left(\mathcal{F}_{t+1} \mid I(t+1)=i, \mathcal{F}_{1:t}\right) \mathrm{P}\left(I(t+1)=i \mid \mathcal{F}_{1:t}\right)}{\mathrm{P}\left(\mathcal{F}_{t+1} \mid \mathcal{F}_{1:t}\right)}
\end{aligned}$$

⟨Smoothing⟩

$$\mathrm{P}\left(I(t+1)=i \mid \mathcal{F}_{1:T}\right) = \mathrm{P}\left(I(t+1)=i \mid \mathcal{F}_{1:t+1}\right) \sum_{j=1}^{k} \frac{\mathrm{P}\left(I(t+1)=i \mid \mathcal{F}_{1:T}\right) p_{ji}}{\mathrm{P}\left(I(t+2)=j \mid \mathcal{F}_{1:t+1}\right)}$$

3.6.3 Derivation and Explicit Expression of (3.10)

We first derive (3.10) and then provide more explicit expression of it. For notational simplicity, we define $\boldsymbol{g} = \boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t) + \boldsymbol{u}_{I(t+1)}(t+1)$. From $y(t+1) = \boldsymbol{x}^\top(t+1)\boldsymbol{g}$, we obtain

$$E_t[y(t+1)] = \boldsymbol{x}^\top(t) E_t[\boldsymbol{g}] \tag{3.11}$$

$$V_t[y(t+1)] = \boldsymbol{x}^\top(t) V_t[\boldsymbol{g}] \boldsymbol{x}(t). \tag{3.12}$$

By differentiating (3.11) and (3.12) with respect $x_i(t)$, we get

$$\frac{\partial}{\partial \boldsymbol{x}(t)} E_t[y(t+1)] = E_t[\boldsymbol{g}] \tag{3.13}$$

$$\frac{\partial}{\partial \boldsymbol{x}(t)} V_t[y(t+1)] = 2V_t[\boldsymbol{g}]\boldsymbol{x}(t) \tag{3.14}$$

$$\frac{\partial}{\partial \boldsymbol{x}(t)} \Delta\boldsymbol{x}^\top(t) E_t[\boldsymbol{B}_{I(t+1)}]\Delta\boldsymbol{x}(t) = 2E_t[\boldsymbol{B}_{I(t+1)}]\Delta\boldsymbol{x}(t) \tag{3.15}$$

where $\partial/\partial\boldsymbol{x}(t) = (\partial/\partial x_1(t), \ldots, \partial/\partial x_N(t))^\top$. From (3.13) to (3.15), the first order optimality condition for the utility function (3.9) is given as

$$E_t[\boldsymbol{g}] - \lambda V_t[\boldsymbol{g}]\boldsymbol{x}(t) - E_t[\boldsymbol{B}_{I(t+1)}]\Delta\boldsymbol{x}(t) = \boldsymbol{0}. \tag{3.16}$$

Since $E_t[\boldsymbol{g}] = E_t[\boldsymbol{L}_{I(t+1)}]\boldsymbol{f}(t)$ from the assumption $E[\boldsymbol{u}_I(t)] = \boldsymbol{0}$ for all regime i, solving (3.16) with respect to $\boldsymbol{x}(t)$ yields the optimal solution (3.10).

To express (3.10) more explicitly, we remark that both $I(t+1)$ and $\boldsymbol{u}_{I(t+1)}(t+1)$ are random at t. The conditional expectation and variance are calculated based on the information available by the investor at t. Specifically, the investor observes factor $\boldsymbol{f}(t)$ and estimates the filtered regime probability which we denote by $q_k(t) = P(I(t) = k)$ for each regime k. A one step ahead regime forecast is then given by $\widehat{q}_j(t+1) = \sum_k q_k(t)p_{kj}$ where p_{kj} denotes the estimated transition probability of the regime process. Based on the regime forecast, the conditional expectations in (3.10) are calculated as

$$E_t[\boldsymbol{L}_{I(t+1)}] = \sum_j \widehat{q}_j(t+1)\boldsymbol{L}_j \tag{3.17}$$

$$E_t[\boldsymbol{B}_{I(t+1)}] = \sum_j \widehat{q}_j(t+1)\boldsymbol{B}_j. \tag{3.18}$$

Similary, the variance term can be explicitly expressed as

$$\begin{aligned} V_t[\boldsymbol{g}] &= E_t[(\boldsymbol{g} - E_t[\boldsymbol{g}])(\boldsymbol{g} - E_t[\boldsymbol{g}])^\top] \\ &= E_t[\boldsymbol{g}\boldsymbol{g}^\top] - E_t[\boldsymbol{g}]E_t[\boldsymbol{g}]^\top \\ &= E_t[\boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t)\boldsymbol{f}^\top(t)\boldsymbol{L}_{I(t+1)}^\top] + E_t[\boldsymbol{u}_{I(t+1)}(t+1)\boldsymbol{u}_{I(t+1)}^\top(t+1)] \\ &\quad - E_t[\boldsymbol{L}_{I(t+1)}]\boldsymbol{f}(t)\boldsymbol{f}^\top(t)E_t[\boldsymbol{L}_{I(t+1)}]^\top \\ &= \sum_j \widehat{q}_j(t+1)(\boldsymbol{L}_j\boldsymbol{f}(t)\boldsymbol{f}^\top(t)\boldsymbol{L}_j + \boldsymbol{W}_j) \\ &\quad - \sum_j\sum_k \widehat{q}_j(t+1)\widehat{q}_k(t+1)\boldsymbol{L}_j\boldsymbol{f}(t)\boldsymbol{f}^\top(t)\boldsymbol{L}_k^\top. \end{aligned} \tag{3.19}$$

Chapter 4

Dynamic Investment for Infinite Horizon

4.1 Introduction

Over the course of past decades, the financial markets have exhibited drastic changes in return generating processes that deviate from those in long-term expectations. For example, in late 1970s and early 1980s that are known as a lost decade, equity markets were stuck under the stagnating macroeconomic environment. In late 1990s, the market participants experienced instability in currencies driven by fragile underpinning in economies across emerging countries. A recent decade includes the US equity market having dropped significantly throughout internet bubble and the global crisis in economy and in the financial markets triggered by subprime loan that turned out to bring out bankruptcy of the Lehman Brothers.

In decision making processes such as asset allocation in both of strategic and tactical investment horizon, investors attempt to predict returns and estimate risks and transaction costs. Contrary to the drastic and discontinuous behavior in financial markets, the traditional practices in investment management have relied on rather simple models mainly because of their tractability. Prediction models often consist of a single set of key financial variables such as expected returns, volatility and correlation among assets, or even in dynamic models, key parameters are fixed and financial variables changes continuously.

Academic endeavor and empirical analyses among several areas in macro economics had already made significant progress in figuring out nature of the drastic and discontinuous changes of economic variables by introducing regime switches. In earlier studies in finance, regime switching models have been applied to wide ranges of assets and markets to successfully explain their dynamic behavior. Initiated by Baum and Petrie (1966) and extensively studied in the statistics and econometrics literature, e.g., Titterington, Smith and Markov (1985) and Hamilton (1994), Markov mixture of dynamic models have attracted increasing interest. The model has an advantageous nature of flexibility to approximate a broad range of dynamics in the real world. Ang and Bekaert (2002a, 2004) construct and numerically solve a regime switching model of international equity markets and report that ignoring the regimes could cost under a presence of cash in asset allocation problems. Among well known factors in individual stock markets such as market risk, value, small cap and momentum, Arshanapalli, Fobbozi and Nelson (2006) reports that the behavior of these premium under different macro economic scenarios is different across factors. The study implies potential presence of different mechanism to drive the equity factor returns from those handled in traditional linear models. Coggi and Manescu (2004) presents a state-dependent version of Fama and French (1996) model to overcome the shortcoming that the original model exhibits quite poor performance in some periods. Ang and Kristensen (2012) finds that time dependency of alpha and beta in the Fama-French model by introducing kernel regression for non-parametric estimation. Although much is left for further research to understand what leads the alpha to deviate from zero, this suggests that some dynamics drive the alpha and beta over time.

Choices of underlying models in the regime switching process are important decisions to approximate highly complicated actual returns. For example, recently portfolio managers tend to pay attentions not only to cross-sectional information across assets in terms of return forecast but also to time series nature of forecasted returns for measuring persistency of the forecasts. Grinold (2007) points out that vintage of information can be found in portfolios under a presence of transaction costs. Sneddon (2008) solves mean-variance optimal portfolio problem with transaction costs and reveals that the optimal portfolio should trade fast decay assets more aggressively than slow decay assets. As well as explaining behaviors of assets in the markets, how optimally investors should behave is similarly important to understand asset returns under the regime switching structure.

To address these issues, this chapter attempts to develop a model of dynamic investment strategy in a transaction cost conscious mean-variance framework with factor models under regime switches. Factor models have brought out major impacts on investment science and produced practically applicable investment approaches. For example, the APT model proposed by Ross (1976) employs macro-economic factors and the Fama-French model by Fama and French (1992) enhances the CAPM with fundamental factors to individual stock portfolios. Factor models in general may have ability to forecast returns to assets. In our model, factors are represented by the regime-switching vector auto-regressive process which is sufficiently general to approximate complicated actual returns observed in the markets. Other key parameters such as loading matrix of assets' returns, transaction costs and investors risk tolerance are driven by regime switching processes as well.

This chapter contributes to literature in several ways. In a mean-variance framework with factor models under regime switches, we obtain semi-analytic solutions of dynamic portfolio optimization problems with transaction costs. In a regime switching framework, our model extends Gârleanu and Pedersen (2013) which derives a closed form solution for the model with multiple securities and multiple return predictors with different mean-reversion speeds. Due to the existence of transaction costs, the optimal portfolio is a linear combination of current and target portfolios the latter of which maximizes the value function. For some special cases, closed form solutions of much simplified form are obtained. Our empirical application to the US equity market where small minus big and high minus low are employed as factors demonstrates superior results in a two regime model to a single regime model and in a three regime model to the two regime model for such performance measures as realized utility and Sharpe ratio that are of particular interest in practice. Taking a close look at the time series of portfolio returns, the superiority of the regime switching model is attributed to flexible asset allocation that investors are able to implement by observing state changes of the market.

The outline of this chapter is as follows. In Section 4.2, we describes a discrete-time dynamic portfolio optimization problem with regime switches. Next Section 4.3, we formulate the optimal investment problem and solve the problem by dynamic programming to obtain the optimal portfolios. Some special cases are also discussed. In Section 4.4, as an empirical application, we model the factors and assets' returns under regime switching framework for equity investments in one of style rotation strategies. Section 4.5 exhibits and discusses investment efficacies of optimal portfolios obtained in the empirical application modeled in Section 4.4. Finally, we conclude the chapter in Section 4.6.

4.2 Portfolio Optimization Problem

We consider an economy with N assets traded at time $t = 1, 2, \ldots$. The excess return of asset i to the market return between t and $t+1$ is $r_i(t+1)$. We assume that an $N \times 1$ excess return vector $\boldsymbol{r}(t) = (r_1(t), \ldots, r_N(t))^\top$ ($\top$ denotes transpose) is given by

$$\boldsymbol{r}(t+1) = \boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t) + \boldsymbol{u}_{I(t+1)}(t+1). \tag{4.1}$$

In (4.1), $I(t)$ is a regime process on $\{1, \ldots, J\}$ that is introduced to represent discontinuous state changes of the market. The details of the regime process will be explained below. The first term $\boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t)$ denotes the expected excess return known to the investor at time t where $\boldsymbol{f}(t)$ is an $M \times 1$ vector of factors that predict excess returns. Note that $\boldsymbol{f}(t)$ in this chapter and what follows is distinguished from the factors in the previous chapter although the notational symbol is common. $\boldsymbol{L}_{I(t+1)}$ is an $N \times M$ matrix of factor loadings such that $\boldsymbol{L}_{I(t+1)} = \boldsymbol{L}_i$ when the regime $I(t+1) = i$ $(i = 1, \ldots, J)$. The second term $\boldsymbol{u}_{I(t+1)}(t+1)$ represents an unpredictable noise. We assume $\mathrm{E}\left(\boldsymbol{u}_{I(t+1)}(t+1) \,\middle|\, I(t+1) = i\right) = \boldsymbol{0}$ for all i ($\boldsymbol{0}$ denotes a zero vector of an appropriate dimension), whereas the covariance matrix is regime-dependent and is denoted as $\boldsymbol{W}_i = \mathrm{V}\left(\boldsymbol{u}_{I(t+1)}(t+1) \,\middle|\, I(t+1) = i\right)$.

The dynamics of the factor is modeled by a first order regime-switching vector autoregressive process (VAR(1))

$$\boldsymbol{f}(t+1) = \boldsymbol{\mu}_{I(t+1)} + \boldsymbol{\Phi}_{I(t+1)}\boldsymbol{f}(t) + \boldsymbol{\epsilon}_{I(t+1)}(t+1). \tag{4.2}$$

In (4.2), $\boldsymbol{\mu}_{I(t+1)}$ is an $M \times 1$ vector determining the level of mean-reversion and $\boldsymbol{\Phi}_{I(t+1)}$ is an $M \times M$ coefficient matrix, both of which are regime-dependent. Specifically, if $I(t+1) = i$, they are given as $\boldsymbol{\mu}_{I(t+1)} = \boldsymbol{\mu}_i$ and $\boldsymbol{\Phi}_{I(t+1)} = \boldsymbol{\Phi}_i$, respectively. $\boldsymbol{\epsilon}_{I(t+1)}(t+1)$ is a vector of noise terms affecting the factors. We assume $\mathrm{E}\left(\boldsymbol{\epsilon}_{I(t+1)}(t+1) \,\middle|\, I(t+1) = i\right) = \boldsymbol{0}$ for all i, whereas the covariance matrix $\boldsymbol{\Sigma}_i = \mathrm{V}\left(\boldsymbol{\epsilon}_{I(t+1)}(t+1) \,\middle|\, I(t+1) = i\right)$ is regime-dependent. We also assume that the factor process $\boldsymbol{f}(t)$ is stationary in time and that

$\mathrm{E}\left(\boldsymbol{\epsilon}_{I(t+1)}(t+1), \boldsymbol{u}_{I(t+1)}(t+1) \,\middle|\, I(t+1)\right) = \mathbf{0}$. Conditions for the stationarity of regime-switching vector autoregressive processes are given in Francq and Zakoïan (2001).

As in many existing literatures, we assume the regime process $I(t)$ follows an irreducible Markov chain on $\{1, \ldots, J\}$ with the transition probability matrix

$$\boldsymbol{P} = \begin{bmatrix} p_{11} & \cdots & p_{1J} \\ \vdots & \ddots & \vdots \\ p_{J1} & \cdots & p_{JJ} \end{bmatrix}, \qquad p_{ij} = \mathrm{P}\left(I(t+1) = j \mid I(t) = i\right).$$

The noise terms $\boldsymbol{u}_{I(t)}$ and $\boldsymbol{\epsilon}_{I(t)}$ are assumed to be conditionally independent in the sense that, given a sample path of the regime process $I(1) = i_1, I(2) = i_2, \ldots$, $\boldsymbol{u}_{i_1}, \boldsymbol{u}_{i_2}, \ldots$ and $\boldsymbol{\epsilon}_{i_1}, \boldsymbol{\epsilon}_{i_2}, \ldots$ are all independent of each other and distribution functions of $\boldsymbol{u}_{i_t}$ and $\boldsymbol{\epsilon}_{i_t}$ are determined by i_t.

At time t, an investor determines amount of investment $x_i(t)$ to asset i. Short sales are allowed and thus $x_j(t)$ may be negative. From (4.1), the excess return of the portfolio $\boldsymbol{x}(t) = (x_1(t), \ldots, x_N(t))^\top$ between t and $t+1$ is $y(t+1) = \boldsymbol{x}(t)^\top\{\boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t) + \boldsymbol{u}_{I(t+1)}(t+1)\}$. To construct a utility function, we assume that at time t an investor is able to predict one step ahead regime $I(t+1)$ with certainty. We denote the predicted regime by $I_t(t+1)$. A natural and plausible way of prediction is to choose $I_t(t+1)$ as the regime that maximizes one step ahead regime probability. In general, filtered regime probabilities are close to 0 or 1 and the regime process shows strong tendency of self-transition (cf., Subsection 4.4.2 later), it is not unrealistic for investor to predict $I_t(t+1)$ with certainty.

Given $\boldsymbol{f}(t)$ and $I_t(t+1)$, the conditional mean of the excess return under investor's prediction is

$$\mathrm{E}\left(\boldsymbol{x}(t)^\top\{\boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t) + \boldsymbol{u}_{I(t+1)}(t+1)\} \,\middle|\, \boldsymbol{f}(t), I_t(t+1)\right) = \boldsymbol{x}(t)^\top\boldsymbol{L}_{I_t(t+1)}\boldsymbol{f}(t)$$

since $\mathrm{E}\left(\boldsymbol{u}_{I(t+1)}(t+1) \,\middle|\, I(t+1) = i\right) = \mathbf{0}$ for all i. Similarly, the conditional variance is calculated as

$$\mathrm{V}\left(\boldsymbol{x}(t)^\top\{\boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t) + \boldsymbol{u}_{I(t+1)}(t+1)\} \,\middle|\, \boldsymbol{f}(t), I_t(t+1)\right) = \boldsymbol{x}(t)^\top\boldsymbol{W}_{I_t(t+1)}\boldsymbol{x}(t).$$

Trading is costly in the market and the transaction cost associated with trading $\boldsymbol{x}(t) - \boldsymbol{x}(t-1)$ when $I(t) = i$ is given by

$$\frac{1}{2}\{\boldsymbol{x}(t) - \boldsymbol{x}(t-1)\}^\top\boldsymbol{B}_{I_t(t+1)}\{\boldsymbol{x}(t) - \boldsymbol{x}(t-1)\}, \tag{4.3}$$

where $\boldsymbol{B}_{I_t(t+1)}$ is a symmetric positive definite matrix measuring the level of trading costs. As noted in Gârleanu and Pedersen (2013), the trading cost in (4.3) stands on the assumption that the price impact of trading $\Delta\boldsymbol{x}(t) = \boldsymbol{x}(t) - \boldsymbol{x}(t-1)$ shares is $\boldsymbol{B}_{I_t(t+1)}\Delta\boldsymbol{x}(t)/2$, which results in a total trading cost of $\Delta\boldsymbol{x}(t)$ times price impact. Hence, $\boldsymbol{B}_{I_t(t+1)}/2$ is interpreted as a multi-dimensional version of Kyle's λ. An investor is risk averse and let $\lambda_i > 0$ denote the coefficient of risk aversion when $I_t(t+1) = i$. Regime-dependent risk aversion coefficient allows us to represent, for example, an investor who chooses larger λ_i

when the market is volatile in regime i. When $I_t(t+1) = i$, the investor's utility at time t is given by

$$\boldsymbol{x}(t)^\top \boldsymbol{L}_i \boldsymbol{f}(t) - \frac{\lambda_i}{2}\boldsymbol{x}(t)^\top \boldsymbol{W}_i \boldsymbol{x}(t) - \frac{1}{2}\{\boldsymbol{x}(t) - \boldsymbol{x}(t-1)\}^\top \boldsymbol{B}_i \{\boldsymbol{x}(t) - \boldsymbol{x}(t-1)\}. \tag{4.4}$$

This is a mean-variance utility penalized for transaction costs. The investor's objective is to choose dynamic investment strategy that maximizes the present value of cumulative utilities in a future. Given initial portfolio $\boldsymbol{x}(0)$, regime $I(1)$ and factor $\boldsymbol{f}(1)$ at time $t = 1$, the objective function to maximize is expressed as

$$\begin{aligned} \mathrm{E}\Big(\sum_{t=1}^{\infty} \rho^{t-1} \Big[& \boldsymbol{x}(t)^\top \boldsymbol{L}_{I_t(t+1)} \boldsymbol{f}(t) - \frac{\lambda_{I_t(t+1)}}{2} \boldsymbol{x}(t)^\top \boldsymbol{W}_{I_t(t+1)} \boldsymbol{x}(t) \\ & - \frac{1}{2}\{\boldsymbol{x}(t) - \boldsymbol{x}(t-1)\}^\top \boldsymbol{B}_{I_t(t+1)} \{\boldsymbol{x}(t) - \boldsymbol{x}(t-1)\}\Big] \,\Big|\, \boldsymbol{x}(0), I(1), \boldsymbol{f}(1)\Big), \end{aligned}$$

where $\rho \in [0, 1)$ is a discount rate. At each time step $t = 1, 2, \ldots$, an investor determines $\boldsymbol{x}(t)$ according to the following timeline.

1. Observe $\boldsymbol{f}(t)$ and $I(t)$ at time t.

2. Predict $I_t(t+1)$ as an $\mathrm{argmax}_j\{p_{I(t),j}\}$ in $\boldsymbol{P}$.

3. Determine $\boldsymbol{x}(t)$ and rebalance from $\boldsymbol{x}(t-1)$ to $\boldsymbol{x}(t)$.

4. Time steps ahead to $t+1$. $I(t+1)$ is randomly chosen according to $I(t)$ and the transition probability matrix $\boldsymbol{P}$.

5. $\boldsymbol{r}(t+1)$ and $\boldsymbol{f}(t+1)$ are given by (4.1) and (4.2) where $\boldsymbol{u}_{I(t+1)}(t+1)$ and $\boldsymbol{\epsilon}_{I(t+1)}(t+1)$ are randomly chosen from certain distribution with zero mean and covariance matrices $\boldsymbol{W}_i$ and $\boldsymbol{\Sigma}_i$ when $I(t+1) = i$, respectively.

4.3 Optimal Investment Strategy

In this section, we derive the optimal investment strategy and the value function. Some special cases of interest are also investigated where the value function can be much simplified. To keep the clarity of presentation, we only provide an outline of the derivation in Section 4.3 and explain detailed calculation in Subsection 4.7.1.

4.3.1 The Optimal Portfolio and the Value Function

Given initial portfolio $\boldsymbol{x}(0)=\boldsymbol{y}$, regime $I(1)=i$ and factor $\boldsymbol{f}(1)=\boldsymbol{f}$, we define the set of value functions by

$$\begin{aligned} V_i(\boldsymbol{y},\boldsymbol{f}) \;=\; & \max_{\{\boldsymbol{x}(t);\ t=1,2,\ldots\}} \mathrm{E}\Big(\sum_{t=1}^{\infty}\rho^{t-1}\Big[\boldsymbol{x}(t)^\top \boldsymbol{L}_{I(t)}\boldsymbol{f}(t)-\frac{1}{2}\boldsymbol{x}(t)^\top\boldsymbol{A}_{I(t)}\boldsymbol{x}(t) \\ & -\frac{1}{2}\{\boldsymbol{x}(t)-\boldsymbol{x}(t-1)\}^\top \boldsymbol{B}_{I(t)}\{\boldsymbol{x}(t)-\boldsymbol{x}(t-1)\}\Big] \\ & \Big|\, \boldsymbol{x}(0)=\boldsymbol{y}, I(1)=i, \boldsymbol{f}(1)=\boldsymbol{f}\Big),\quad i=1,\ldots,J \end{aligned} \tag{4.5}$$

where we set $\boldsymbol{A}_i=\lambda_i\boldsymbol{W}_i$ for notational simplicity. It is noticed that, since the regime process switches from one regime to another, the set of value functions $\{V_i\}=\{V_1,\ldots,V_J\}$ in (4.5) must be determined simultaneously. A set of guess solutions $\{V_i\}$ to the system of problems (4.5) for $i=1,\ldots,J$ is

$$V_i(\boldsymbol{y},\boldsymbol{f})=-\frac{1}{2}\boldsymbol{y}^\top\boldsymbol{\beta}_i\boldsymbol{y}+\boldsymbol{\delta}_i^\top\boldsymbol{y}+\frac{1}{2}\boldsymbol{f}^\top\boldsymbol{\eta}_i\boldsymbol{f}+\boldsymbol{\xi}_i^\top\boldsymbol{f}+\boldsymbol{y}^\top\boldsymbol{\kappa}_i\boldsymbol{f}+\zeta_i,\quad i=1,\ldots,J \tag{4.6}$$

where $\boldsymbol{\beta}_i$ is an $N\times N$ symmetric positive definite matrix, $\boldsymbol{\eta}_i$ is an $M\times M$ symmetric positive definite matrix, $\boldsymbol{\kappa}_i$ is an $N\times M$ matrix, $\boldsymbol{\delta}_i$ is an $N\times 1$ vector, $\boldsymbol{\xi}_i$ is an $M\times 1$ vector and ζ_i is a scalar.

The problem of obtaining the optimal portfolio and the set of value functions can be solved explicitly except $\{\boldsymbol{\beta}_i\}=\{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_J\}$ in (4.6) which is simultaneously determined as a limit of the following iterative procedure. Starting with $\{\boldsymbol{\beta}_1^{(0)},\ldots,\boldsymbol{\beta}_J^{(0)}\}=\{\boldsymbol{O},\ldots,\boldsymbol{O}\}$ where $\boldsymbol{O}$ denotes a zero matrix, we recursively define $\{\boldsymbol{\beta}_i^{(n)}\}=\{\boldsymbol{\beta}_1^{(n)},\ldots,\boldsymbol{\beta}_J^{(n)}\}$ by

$$\boldsymbol{\beta}_i^{(n)}=\boldsymbol{B}_i-\boldsymbol{B}_i\left(\rho\sum_{k=1}^{J}p_{ik}\boldsymbol{\beta}_k^{(n-1)}+\boldsymbol{A}_i+\boldsymbol{B}_i\right)^{-1}\boldsymbol{B}_i,\quad i=1,\ldots,J \tag{4.7}$$

for $n=1,2,\ldots$. The next lemma proves the convergence of $\{\boldsymbol{\beta}_i^{(n)}\}$.

Lemma 4.3.1 $\{\boldsymbol{\beta}_i^{(n)}\}$ defined by (4.7) converges elementwise to a set of symmetric positive definite matrices $\{\boldsymbol{\beta}_i^{(\infty)}\}=\{\boldsymbol{\beta}_1^{(\infty)},\ldots,\boldsymbol{\beta}_J^{(\infty)}\}$ as $n\to\infty$.

(Proof) See Subsection 4.7.1. □

To express other coefficient matrices and vectors in (4.6), we need to introduce some notations. For $K\times L$ matrix $\boldsymbol{M}=[m_{ij}]$, we define a $KL\times 1$ vector by

$$\mathrm{vec}(\boldsymbol{M})=(m_{11},\ldots,m_{K1},\ldots,m_{1L},\ldots,m_{KL})^\top.$$

For sets of vectors and/or matrices $\boldsymbol{z}=\{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_J\}$, $\boldsymbol{v}=\{\boldsymbol{v}_1,\ldots,\boldsymbol{v}_J\}$ and $\boldsymbol{w}=\{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_J\}$, we define an operator $\langle\ \rangle_i$ by

$$\langle\boldsymbol{z}\rangle_i=\sum_{j=1}^{J}p_{ij}\boldsymbol{z}_j,\qquad \langle\boldsymbol{z}\boldsymbol{v}\rangle_i=\sum_{j=1}^{J}p_{ij}\boldsymbol{z}_j\boldsymbol{v}_j,\qquad \langle\boldsymbol{z}\boldsymbol{v}\boldsymbol{w}\rangle_i=\sum_{j=1}^{J}p_{ij}\boldsymbol{z}_j\boldsymbol{v}_j\boldsymbol{w}_j \tag{4.8}$$

when sums and products in (4.8) are well-defined. Also, $\boldsymbol{I}$ denotes an identity matrix of an appropriate dimension.

Proposition 4.3.2 The set of value functions $\{V_i\}$ is given by (4.6) for $i = 1, \ldots, J$ with the coefficients given in (a) to (f) below, provided the inverse matrices in (4.9), (4.11), (4.12) and (4.13) exist.
(a) $\{\boldsymbol{\beta}_i\}$ is given by $\{\boldsymbol{\beta}_i^{(\infty)}\}$ in Lemma 4.3.1.
(b) $\{\boldsymbol{\kappa}_i\}$ is given by

$$\begin{bmatrix} \mathrm{vec}(\boldsymbol{\kappa}_1) \\ \vdots \\ \mathrm{vec}(\boldsymbol{\kappa}_J) \end{bmatrix} = (\boldsymbol{I} - \rho\boldsymbol{\Gamma})^{-1} \begin{bmatrix} \mathrm{vec}(\boldsymbol{B}_1\boldsymbol{C}_1\boldsymbol{L}_1) \\ \vdots \\ \mathrm{vec}(\boldsymbol{B}_J\boldsymbol{C}_J\boldsymbol{L}_J) \end{bmatrix}, \tag{4.9}$$

where

$$\begin{aligned} \boldsymbol{C}_i &= (\rho\langle\boldsymbol{\beta}\rangle_i + \boldsymbol{A}_i + \boldsymbol{B}_i)^{-1} \\ \boldsymbol{\Gamma} &= \begin{bmatrix} p_{11}\boldsymbol{\Phi}_1^\top \otimes (\boldsymbol{B}_1\boldsymbol{C}_1) & \cdots & p_{1J}\boldsymbol{\Phi}_J^\top \otimes (\boldsymbol{B}_1\boldsymbol{C}_1) \\ \vdots & \ddots & \vdots \\ p_{J1}\boldsymbol{\Phi}_1^\top \otimes (\boldsymbol{B}_J\boldsymbol{C}_J) & \cdots & p_{JJ}\boldsymbol{\Phi}_J^\top \otimes (\boldsymbol{B}_J\boldsymbol{C}_J). \end{bmatrix} \end{aligned} \tag{4.10}$$

and $\otimes$ denotes Kronecker product.
(c) $\{\boldsymbol{\delta}_i\}$ is given by

$$\begin{bmatrix} \boldsymbol{\delta}_1 \\ \vdots \\ \boldsymbol{\delta}_J \end{bmatrix} = \rho(\boldsymbol{I} - \rho\boldsymbol{\Theta})^{-1} \begin{bmatrix} \boldsymbol{B}_1\boldsymbol{C}_1\langle\boldsymbol{\kappa\mu}\rangle_1 \\ \vdots \\ \boldsymbol{B}_J\boldsymbol{C}_J\langle\boldsymbol{\kappa\mu}\rangle_J \end{bmatrix}, \tag{4.11}$$

where

$$\boldsymbol{\Theta} = \begin{bmatrix} p_{11}\boldsymbol{B}_1\boldsymbol{C}_1 & \cdots & p_{1J}\boldsymbol{B}_1\boldsymbol{C}_1 \\ \vdots & \ddots & \vdots \\ p_{J1}\boldsymbol{B}_J\boldsymbol{C}_J & \cdots & p_{JJ}\boldsymbol{B}_J\boldsymbol{C}_J \end{bmatrix}.$$

(d) $\{\boldsymbol{\eta}_i\}$ is a set of symmetric positive definite matrices and is given by

$$\begin{bmatrix} \mathrm{vec}(\boldsymbol{\eta}_1) \\ \vdots \\ \mathrm{vec}(\boldsymbol{\eta}_J) \end{bmatrix} = (\boldsymbol{I} - \rho\boldsymbol{\Psi})^{-1} \begin{bmatrix} \mathrm{vec}((\rho\langle\boldsymbol{\kappa\Phi}\rangle_1 + \boldsymbol{L}_1)^\top\boldsymbol{C}_1(\rho\langle\boldsymbol{\kappa\Phi}\rangle_1 + \boldsymbol{L}_1)) \\ \vdots \\ \mathrm{vec}((\rho\langle\boldsymbol{\kappa\Phi}\rangle_J + \boldsymbol{L}_J)^\top\boldsymbol{C}_J(\rho\langle\boldsymbol{\kappa\Phi}\rangle_J + \boldsymbol{L}_J)) \end{bmatrix}, \tag{4.12}$$

where

$$\boldsymbol{\Psi} = \begin{bmatrix} p_{11}\boldsymbol{\Phi}_1^\top \otimes \boldsymbol{\Phi}_1^\top & \cdots & p_{1J}\boldsymbol{\Phi}_J^\top \otimes \boldsymbol{\Phi}_J^\top \\ \vdots & \ddots & \vdots \\ p_{J1}\boldsymbol{\Phi}_1^\top \otimes \boldsymbol{\Phi}_1^\top & \cdots & p_{JJ}\boldsymbol{\Phi}_J^\top \otimes \boldsymbol{\Phi}_J^\top \end{bmatrix}.$$

(e) $\{\boldsymbol{\xi}_i\}$ is given by

$$\begin{bmatrix} \boldsymbol{\xi}_1 \\ \vdots \\ \boldsymbol{\xi}_J \end{bmatrix} = \rho(\boldsymbol{I} - \rho\boldsymbol{\Phi})^{-1} \begin{bmatrix} \langle \boldsymbol{\Phi}^\top \boldsymbol{\eta}^\top \boldsymbol{\mu}\rangle_1 + (\rho\langle \boldsymbol{\kappa\Phi}\rangle_1 + \boldsymbol{L}_1)^\top \boldsymbol{C}_1(\langle \boldsymbol{\delta}\rangle_1 + \langle \boldsymbol{\kappa\mu}\rangle_1) \\ \vdots \\ \langle \boldsymbol{\Phi}^\top \boldsymbol{\eta}^\top \boldsymbol{\mu}\rangle_J + (\rho\langle \boldsymbol{\kappa\Phi}\rangle_J + \boldsymbol{L}_J)^\top \boldsymbol{C}_J(\langle \boldsymbol{\delta}\rangle_J + \langle \boldsymbol{\kappa\mu}\rangle_J) \end{bmatrix}, \tag{4.13}$$

where

$$\boldsymbol{\Phi} = \begin{bmatrix} p_{11}\boldsymbol{\Phi}_1^\top & \cdots & p_{1J}\boldsymbol{\Phi}_J^\top \\ \vdots & \ddots & \vdots \\ p_{J1}\boldsymbol{\Phi}_1^\top & \cdots & p_{JJ}\boldsymbol{\Phi}_J^\top \end{bmatrix}.$$

(f) $\{\zeta_i\}$ is given by

$$\begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_J \end{bmatrix} = \rho(\boldsymbol{I} - \rho\boldsymbol{P})^{-1} \begin{bmatrix} \frac{1}{2}\langle \boldsymbol{\mu}^\top \boldsymbol{\eta\mu}\rangle_1 + \frac{1}{2}\mathrm{E}\left(\langle \boldsymbol{\epsilon}^\top \boldsymbol{\eta\epsilon}\rangle_1\right) + \langle \boldsymbol{\xi}^\top \boldsymbol{\mu}\rangle_1 \\ \vdots \\ \frac{1}{2}\langle \boldsymbol{\mu}^\top \boldsymbol{\eta\mu}\rangle_J + \frac{1}{2}\mathrm{E}\left(\langle \boldsymbol{\epsilon}^\top \boldsymbol{\eta\epsilon}\rangle_J\right) + \langle \boldsymbol{\xi}^\top \boldsymbol{\mu}\rangle_J \\ +\frac{1}{2}\rho(\langle \boldsymbol{\delta}\rangle_1 + \langle \boldsymbol{\kappa\mu}\rangle_1)^\top \boldsymbol{C}_1(\langle \boldsymbol{\delta}\rangle_1 + \langle \boldsymbol{\kappa\mu}\rangle_1) \\ \vdots \\ +\frac{1}{2}\rho(\langle \boldsymbol{\delta}\rangle_J + \langle \boldsymbol{\kappa\mu}\rangle_J)^\top \boldsymbol{C}_J(\langle \boldsymbol{\delta}\rangle_J + \langle \boldsymbol{\kappa\mu}\rangle_J) \end{bmatrix}, \tag{4.14}$$

where $\mathrm{E}\left(\langle \boldsymbol{\epsilon}^\top \boldsymbol{\eta\epsilon}\rangle_i\right)$ is calculated as

$$\mathrm{E}\left(\langle \boldsymbol{\epsilon}^\top \boldsymbol{\eta\epsilon}\rangle_i\right) = \mathrm{E}\left(\sum_{k=1}^{J} p_{ik}\boldsymbol{\epsilon}_k^\top \boldsymbol{\eta}_k \boldsymbol{\epsilon}_k\right) = \sum_{k=1}^{J} p_{ik} \sum_{\ell=1}^{M}\sum_{m=1}^{M} (\boldsymbol{\eta}_k)_{\ell m}(\boldsymbol{\Sigma}_k)_{\ell m}.$$

(Proof) See Subsection 4.7.1. □

With the coefficients in Proposition 4.3.2, the optimal portfolio is obtained as follows.

Proposition 4.3.3 Given previous portfolio $\boldsymbol{x}(t-1)$, current factor $\boldsymbol{f}(t)$ and regime $I(t) = i$, the optimal portfolio at time t is

$$\boldsymbol{x}_i^*(t) = (\boldsymbol{I} - \boldsymbol{B}_i^{-1}\boldsymbol{\beta}_i)\boldsymbol{x}(t-1) + \boldsymbol{B}_i^{-1}\boldsymbol{\beta}_i\{\boldsymbol{\beta}_i^{-1}(\boldsymbol{\kappa}_i\boldsymbol{f}(t) + \boldsymbol{\delta}_i)\}, \quad i = 1, \dots, J. \tag{4.15}$$

(Proof) See Subsection 4.7.1. □

Before closing this subsection, we remark that the optimal portfolio obtained in Proposition 4.3.3 can be expressed as

$$\boldsymbol{x}_i^*(t) = (\boldsymbol{I} - \boldsymbol{B}_i^{-1}\boldsymbol{\beta}_i)\boldsymbol{x}(t-1) + \boldsymbol{B}_i^{-1}\boldsymbol{\beta}_i\boldsymbol{y}_i^*(t),$$

a linear combination of the previous portfolio $\boldsymbol{x}(t-1)$ and the *target* portfolio $\boldsymbol{y}_i^*(t) = \boldsymbol{\beta}_i^{-1}(\boldsymbol{\kappa}_i\boldsymbol{f}(t) + \boldsymbol{\delta}_i)$. $\boldsymbol{y}_i^*(t)$ is called target portfolio because, given current factor $\boldsymbol{f}(t)$ and regime $I(t) = i$, $\boldsymbol{y} = \boldsymbol{y}_i^*(t)$ maximizes the value function $V_i(\boldsymbol{y}, \boldsymbol{f}(t))$. To see this, the first order optimality condition becomes

$$\frac{\partial}{\partial \boldsymbol{y}} V_i(\boldsymbol{y}, \boldsymbol{f}(t)) = -\boldsymbol{\beta}_i\boldsymbol{y} + \boldsymbol{\delta}_i + \boldsymbol{\kappa}_i\boldsymbol{f}(t) = \boldsymbol{0} \tag{4.16}$$

where the left hand side is a column vector of element-wise partial derivatives. Since $\boldsymbol{y} = \boldsymbol{y}_i^*(t)$ solves (4.16), it attains the maximum of $V_i(\boldsymbol{y}, \boldsymbol{f}(t))$.

4.3.2 Regime Independent Cost Parameters

When the covariance matrix $\boldsymbol{W}_i$ of the noise term in (4.1), investor's risk aversion coefficient λ_i and the transaction cost matrix $\boldsymbol{B}_i$ does not depend on regime, i.e., $\boldsymbol{W}_i = \boldsymbol{W}$, $\lambda_i = \lambda$ and $\boldsymbol{B}_i = \boldsymbol{B}$ for all i, we can obtain the optimal portfolio and the value function explicitly. Since all coefficient matrices in (4.35) are independent of i in this case, $\boldsymbol{\beta}_i$ is common for all i, i.e., $\boldsymbol{\beta}_i = \boldsymbol{\beta}$. Thus, (4.35) is reduced to

$$\boldsymbol{\beta} = \boldsymbol{B} - \boldsymbol{B}(\rho\boldsymbol{\beta} + \boldsymbol{A} + \boldsymbol{B})^{-1}\boldsymbol{B} \tag{4.17}$$

with $\boldsymbol{A} = \lambda\boldsymbol{W}$. Let $\boldsymbol{H} = \boldsymbol{B}^{-1/2}\boldsymbol{\beta}\boldsymbol{B}^{-1/2}$ and $\boldsymbol{K} = \boldsymbol{B}^{-1/2}\boldsymbol{A}\boldsymbol{B}^{-1/2}$ where $\boldsymbol{B}^{-1/2}$ denotes inverse matrix of $\boldsymbol{B}^{1/2}$, a square root of $\boldsymbol{B}$. It is noted that, since $\boldsymbol{B}$ is symmetric positive definite, there exists unique matrix square root $\boldsymbol{B}^{1/2}$ satisfying $(\boldsymbol{B}^{1/2})^2 = \boldsymbol{B}$, see e.g., Bhatia (1997). Note also that $\boldsymbol{B}^{1/2}$ is symmetric positive definite which ensures the existence of $\boldsymbol{B}^{-1/2}$. Then, (4.17) is rewritten as

$$\rho\boldsymbol{H}^2 + \{\boldsymbol{K} + (1-\rho)\boldsymbol{I}\}\boldsymbol{H} - \boldsymbol{K} = \boldsymbol{O}. \tag{4.18}$$

We can solve (4.18) explicitly to get unique symmetric positive definite solution

$$\boldsymbol{H} = \frac{1}{2\rho}\left[[\{\boldsymbol{K} + (1-\rho)\boldsymbol{I}\}^2 + 4\rho\boldsymbol{K}]^{1/2} - \{\boldsymbol{K} + (1-\rho)\boldsymbol{I}\}\right], \tag{4.19}$$

which gives $\boldsymbol{\beta} = \boldsymbol{B}^{1/2}\boldsymbol{H}\boldsymbol{B}^{1/2}$. Other coefficients $\boldsymbol{\delta}_i$, $\boldsymbol{\eta}_i$, $\boldsymbol{\xi}_i$, $\boldsymbol{\kappa}_i$ and ζ_i are regime-dependent and are obtained from (4.9) to (4.14). The optimal portfolio in this case is

$$\boldsymbol{x}_i^*(t) = (\boldsymbol{I} - \boldsymbol{B}^{-1}\boldsymbol{\beta})\boldsymbol{x}(t-1) + \boldsymbol{B}^{-1}\boldsymbol{\beta}\{\boldsymbol{\beta}^{-1}(\boldsymbol{\kappa}_i\boldsymbol{f}(t) + \boldsymbol{\delta}_i)\},$$

where the weight matrix $\boldsymbol{B}^{-1}\boldsymbol{\beta}$ is also regime-independent.

4.3.3 Transaction Cost Matrix Proportional to the Covariance Matrix

In addition to the regime-independence assumptions in Subsection 4.3.2, we further assume that the transaction cost matrix $\boldsymbol{B}$ is proportional to the covariance matrix $\boldsymbol{W}$, i.e., $\boldsymbol{B} = \gamma\boldsymbol{W}$ for $\gamma > 0$. See Gârleanu and Pedersen (2013) for the justification of this assumption. From $\boldsymbol{A} = \lambda\boldsymbol{W}$ and $\boldsymbol{B} = \lambda\boldsymbol{W}$, we obtain $\boldsymbol{K} = \frac{\lambda}{\gamma}\boldsymbol{I}$. This together with (4.19) shows

$$\boldsymbol{\beta} = \frac{\sqrt{\{\lambda + (1-\rho)\gamma\}^2 + 4\rho\lambda\gamma} - \{\lambda + (1-\rho)\gamma\}}{2\rho}\boldsymbol{W}.$$

The optimal portfolio in this case becomes

$$\boldsymbol{x}_i^*(t) = \frac{\lambda + (1+\rho)\gamma - \sqrt{\{\lambda + (1-\rho)\gamma\}^2 + 4\rho\lambda\gamma}}{2\rho\gamma}\boldsymbol{x}(t-1) + \frac{1}{\gamma}\boldsymbol{W}^{-1}(\boldsymbol{\kappa}_i\boldsymbol{f}(t) + \boldsymbol{\delta}_i),$$

that implies that, independent of the current regime i, it is optimal to hold fixed portion of the previous portfolio $\boldsymbol{x}(t-1)$.

4.3.4 Myopic Optimization

When the discount rate $\rho = 0$ as investors are myopic and do not care investment utility in the future, the problem is reduced to single period optimization. Substituting $\rho = 0$, (4.7) gives $\boldsymbol{\beta}_i = \boldsymbol{B}_i - \boldsymbol{B}_i(\boldsymbol{A}_i + \boldsymbol{B}_i)^{-1}\boldsymbol{B}_i$. From (4.9) and (4.11), we also have $\boldsymbol{\kappa}_i = \boldsymbol{B}_i(\boldsymbol{A}_i + \boldsymbol{B}_i)^{-1}\boldsymbol{L}_i$ and $\boldsymbol{\delta}_i = \boldsymbol{0}$. Hence, the optimal holding (4.15) in this case is given explicitly as

$$\boldsymbol{x}_i^*(t) = (\boldsymbol{A}_i + \boldsymbol{B}_i)^{-1}(\boldsymbol{B}_i\boldsymbol{x}(t-1) + \boldsymbol{L}_i\boldsymbol{f}(t)).$$

4.4 Data and Model Estimation

In this section, 500 long data in weekly frequency, we estimate the regime switching structure of Fama-French factors to predict US equity style portfolio returns. As a tool for model estimation, we use the R as a statical analysis package running in the Intel(R) Core(TM) i7-4960X CPU 3.60GHz 6 Cores 12 Threads under 64bit operating system with 8G byte memory.

4.4.1 Data

We take advantage of the research data in the Kenneth French Data Library, as one of the most benefiting database among not only to academia but to practitioners in finance. The time series contains 630 weekly data from the first week of June 2002 to the last week of June 2014. We use the first 500 weeks for parameter estimation and in-sample comparison of investment performance of the optimal portfolio in Subsection 4.5.1. Following 130 weeks are used for out-of-sample examination of investment performance in Subsection 4.5.2.

As assets to hold in portfolios, we focus on 6 equal weighted portfolios formed on size and book-to-market. These are portfolios comprised of all NYSE, AMEX and NASDAQ stocks that hold individual stocks in equal market value. The 6 equal weighted portfolios are the intersections of 2 portfolios formed on size as market value of capitalization and 3 portfolios formed on the ratio of book value of equity to market equity (BE/ME ratio). Median of ME splits the universe into 2 size portfolios, Small and Big, and the BE/ME ratio subdivides each of 2 size portfolios into 3 subportfolios at 30 and 70 percentiles, i.e., Growth, Neutral and Value in ascending order of the ratio. The constituents in 6 portfolios are reshuffled June in each year. In what follows, we call each of 6 portfolios as an asset and use abbreviated notations SG (Small Growth), SN (Small Neutral), SV (Small Value), BG (Big Growth), BN (Big Neutral) and BV (Big Value).

The excess return of an asset to the market return is calculated in the standard way based on the CAPM. Let $\widetilde{r}_k(t)$ denote the return of asset k. The excess return in (4.1) is then constructed by $r_k(t) = \widetilde{r}_k(t) - r_f(t) - \beta_k(r_M(t) - r_f(t))$ with the risk free rate $r_f(t)$, the market return $r_M(t)$ and the CAPM coefficient β_k. We use a one-month Treasury bill rate for $r_f(t)$ and market return $r_M(t)$ is given by the value-weighted return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ. Figure 4.1 depicts cumulative excess returns to the market of 6 assets, SG, SN and SV on the upper panel and BG, BN and BV on the lower panel. Amongst 6 assets, SV (Small Value) exhibits outstanding performance except for a period of the Lehman Shock. Table 4.1 and Figure 4.1 display a summary statistics of the returns to the 6 assets and the market and cumulative profiles of excess returns to the 6 assets above the market, respectively

For factors as common return predictors to 6 assets described above, we employ SMB (Small minus Big) and HML (High minus Low) contained in the Data Library. SMB and HML together with the market return compose the Fama-French three factor model that has been contributed to number of researches in wide spectrum of interests including studies of market efficiency and investment performance analytics. In the following empirical analysis, we also take advantage of the well recognized Fama-French factor model and use SMB and HML factors to predict one step ahead returns of 6 assets. The return process of SMB is calculated by value weighted average excess return of 3 Small portfolios minus value weighted average excess return of 3 Big portfolios. Likewise, the return of HML is given by the value weighted average excess return of 2 Value portfolios minus that of 2 Growth portfolios. We remark that the 2 common factors, SMB and HML, are not linear combinations of the 6 assets since SMB and HML are calculated in value weighted basis while 6 assets are calculated in equal weighted basis. Table 4.2 and Figure 4.2 display a summary statistics of the returns to and cumulative profiles of the 2 factors, respectively.

Table 4.1: Summary Statistics of returns to assets and the market

	min.	Q1	median	mean	Q3	max.	std
SG	−.1709	−.0136	.0046	.0026	.0209	.0017	.0309
SN	−.1626	−.0104	.0036	.0028	.0189	.0015	.0275
SV	−.1690	−.0089	.0044	.0034	.0167	.0015	.0261
BG	−.1666	−.0107	.0035	.0023	.0170	.0017	.0281
BN	−.1989	−.0112	.0039	.0025	.0181	.0019	.0300
BV	−.2399	−.0119	.0034	.0028	.0189	.0025	.0330
Market	−.1839	−.0097	.0030	.0019	.0152	.1304	.0256

Weekly figures in first week of June 2002 to the last week of June 2014

Table 4.2: Summary Statistics of returns to factors

	min.	Q1	median	mean	Q3	max.	std
SMB	−.0384	−.0062	.0011	.0006	.0077	.0366	.0114
HML	−.0695	−.0051	.0003	.0004	.0057	.0764	.0122

Weekly figures in first week of June 2002 to the last week of June 2014

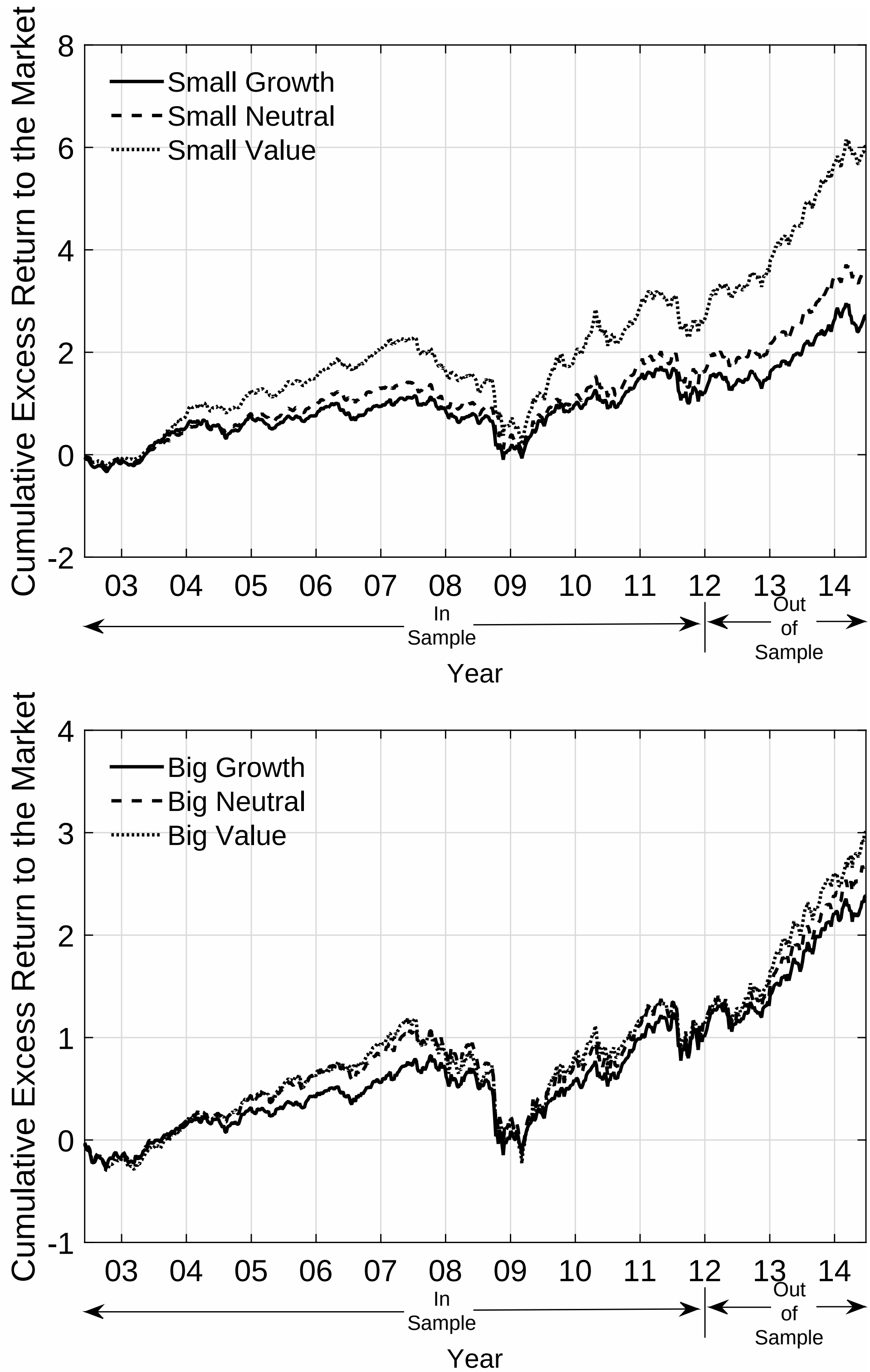


Figure 4.1: Times series of cumulative excess returns to the market of 6 assets. SG, SN and SV on the upper panel and BG, BN and BV on the lower panel

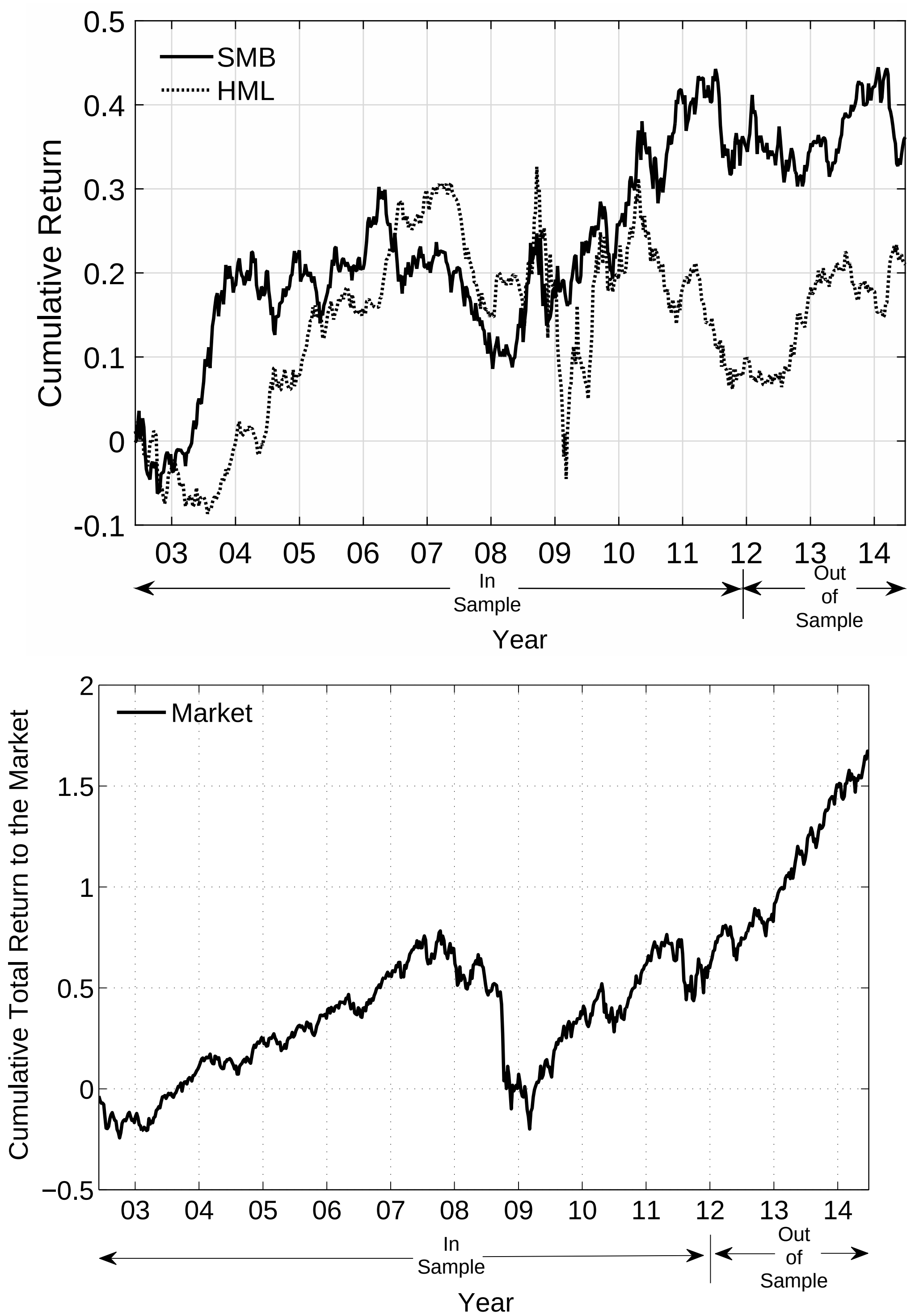


Figure 4.2: Time series of cumulative returns of 2 factors SMB and HML and the Market

4.4.2 Parameter Estimation of the Model

The model parameters are estimated based on the first 500 week long data in a period indicated as in-sample in Figure 4.1 and Figure 4.2 for excess returns of 6 assets and 2 factors. (4.1) and (4.2) can be rewritten in the space state representation, for an example of number of regime $J = 3$, of which an observation model is given as

$$\boldsymbol{r}(t+1) = \begin{cases} \boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t) + \boldsymbol{u}_{I(t+1)}(t+1), & I(t+1) = 1 \\ \boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t) + \boldsymbol{u}_{I(t+1)}(t+1), & I(t+1) = 2 \\ \boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t) + \boldsymbol{u}_{I(t+1)}(t+1), & I(t+1) = 3 \end{cases} \tag{4.20}$$

where $\boldsymbol{u}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{W}_i), i = 1, 2, 3$ and

$$\boldsymbol{f}(t+1) = \begin{cases} \boldsymbol{\mu}_{I(t+1)} + \boldsymbol{\Phi}_{I(t+1)}(t) + \boldsymbol{\epsilon}_{I(t+1)}(t+1), & I(t+1) = 1 \\ \boldsymbol{\mu}_{I(t+1)} + \boldsymbol{\Phi}_{I(t+1)}(t) + \boldsymbol{\epsilon}_{I(t+1)}(t+1), & I(t+1) = 2 \\ \boldsymbol{\mu}_{I(t+1)} + \boldsymbol{\Phi}_{I(t+1)}(t) + \boldsymbol{\epsilon}_{I(t+1)}(t+1), & I(t+1) = 3 \end{cases} \tag{4.21}$$

where $\boldsymbol{\epsilon}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_i), i = 1, 2, 3$. A system model is given as

$$\boldsymbol{q}_{t+1} = \boldsymbol{P}^\top \boldsymbol{q}_t \tag{4.22}$$

or

$$\begin{bmatrix} q_{1,t+1} \\ q_{2,t+1} \\ q_{3,t+1} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \\ p_{13} & p_{23} & p_{33} \end{bmatrix} \begin{bmatrix} q_{1,t} \\ q_{2,t} \\ q_{3,t} \end{bmatrix}$$

where $P\big(I(t) = i\big) = q_{i,t}$ and $q_{1,t} + q_{2,t} + q_{3,t} = 1$.

Similar to the previous chapter, we apply a filtering in identifying a regime $\{I(t)\}$. The filtering algorithm is given in Subsection 3.6.2. The quasi-Newton algorithm identifies the parameter set under the framework of the maximum likelihood estimation where we assume that the noise terms in both (4.20) and (4.21) follow normal distributions

$$\begin{bmatrix} \boldsymbol{u}_{I(t+1)} \\ \hline \boldsymbol{\epsilon}_{I(t+1)} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{0} \\ \hline \boldsymbol{0} \end{bmatrix}, \left[\begin{array}{c|c} \boldsymbol{W}_{I(t+1)} & \boldsymbol{0} \\ \hline \boldsymbol{0} & \boldsymbol{\Sigma}_{I(t+1)} \end{array} \right] \right).$$

We adopt the filtering likelihood in the maximum likelihood estimation. The likelihood function to maximize can be written down as that of a state space model with regime switches. For maximum likelihood estimation of regime switching state space models, see e.g., Frühwirth-Schnatter (2006). In the process of likelihood maximization, we apply the Householder transformation proposed by Gersch and Kitagawa (1982) to make the estimation procedure faster and more stable by transforming the data into orthogonal form. See Subsection 4.7.2 for our application of the Householder transformation to estimate the model parameters.

Concerning the number of regimes, many of earlier studies report that asset price dynamics are well described by a two regime model representing high risk regime and low risk regime. Examples include analyses of stock indices by Ang and Bekaert (2002a, 2004), Valdesogo et al. (2009), and individual stocks by Coggi and Manescu (2004). An exceptional literature is Guildolin and Timmermann (2004) where they tried to apply 4 regime model to asset allocation. However, one out of four states are identified to have low probabilities to stay on its own. In addition, we sometimes encounter numerical instability of estimation for larger number of regimes. We therefore choose the number of regimes $J = 2$ and $J = 3$ in our empirical applications. The single regime case is also studied for comparison of the portfolio optimization. In the rest of this subsection, we first discuss the two regime model in detail to confirm that high and low risk regimes are reasonably identified as in the previous researches. We then show the result for the three regime model and see how the estimation is changed by adding new regime.

Tables 4.3 and 4.4 summarize the estimated parameters in (4.20) and (4.21) respectively for both single and two regime models. The Akaike's Information Criterion of the two regime model is 8844.4 which is much better than 9732.8 of the single regime model. A notable feature observed in these tables is that variance terms in $\boldsymbol{W}_2$ and $\boldsymbol{\Sigma}_2$ are several times larger than those in $\boldsymbol{W}_1$ and $\boldsymbol{\Sigma}_1$, implying Regime 1 represents rather tranquil state of the market while Regime 2 is a turbulent state. The estimated parameters of the single regime model lie between those for the two regime model as expected.

The estimated transition probabilities between two regimes in (4.22) are

$$\boldsymbol{P} = \begin{bmatrix} .944 & .056 \\ .148 & .852 \end{bmatrix}. \tag{4.23}$$

On average, Regime 1 continues $1/(1 - 0.944) = 17.9$ weeks and Regime 2 continues $1/(1 - 0.852) = 6.8$ weeks. Figure 4.3 shows the time series of the filtered probabilities in the two regime model. From 2003 to the first half of 2008, the regime process stays in Regime 1 while the credit bubble grows and followed by Regime 2 which starts just before the Lehman shock in the second half of 2008. Since the second half of 2009, Regimes 1 and 2 appear alternately during which we have experienced European sovereign crisis and the US treasury downgrade. Figure 4.3 suggests the usefulness of the regime switching model in investment decision making as the model is expected to grasp drastic and sudden changes in the market.

Another important finding in Figure 4.3 is that the estimated regime probabilities in most of the entire period are close to 0 or 1, which makes it possible for investors to estimate current regime with certainty. In this process, we identify the regime using filtering algorithm described in Subsection 3.6.2 at time t as $I(t) = i$. when the filtered probability of Regime i is larger than that of the other. Then we predict $I_t(t+1) = \text{argmax}_j\{p_{I(t),j}\}$ in $\boldsymbol{P}$ as a regime for investment decision making for building optimal portfolios. Then, 359 weeks out of 498 weeks in the in-sample period fall into Regime 1 and 139 weeks into Regime 2 (due to a lead-lag in estimating models, regimes are missing for 2 weeks).

We next take a closer look at the characteristics of Regime 1 and Regime 2 to understand how the state of the market differs across different regimes. To this end, we investigate optimal portfolio holdings of a single regime model derived by assuming that Regime i of the two regime model continues forever. Specifically, we consider a single regime VAR(1)

Table 4.3: Estimated parameters of $\boldsymbol{L}$ and $\boldsymbol{W}$ in (4.20) for assets

Single regime model								
	$\boldsymbol{L}$		$\boldsymbol{W}(\times 10^{-3})$					
	SMB	HML	SG	SN	SV	BG	BN	BV
SG	.100	.078	.179	.835	.691	.429	.380	.213
SN	.060	.101	.119	.113	.836	.264	.503	.417
SV	.176	.193	.123	.118	.176	.103	.506	.591
BG	−.020	.005	.037	.018	.009	.042	.341	−.027
BN	.026	−.011	.031	.033	.041	.014	.037	.598
BV	.101	−.064	.031	.049	.086	−.002	.040	.120
Two regime model								
Regime 1	$\boldsymbol{L}_1$		$\boldsymbol{W}_1(\times 10^{-3})$					
	SMB	HML	SG	SN	SV	BG	BN	BV
SG	.119	−.032	.108	.825	.747	.482	.241	.142
SN	.077	.088	.067	.061	.832	.351	.374	.266
SV	.161	.207	.070	.058	.081	.246	.366	.400
BG	−.045	−.020	.024	.013	.011	.022	.341	.019
BN	−.007	.032	.011	.012	.014	.007	.018	.555
BV	.019	.105	.009	.013	.023	.001	.015	.040
Regime 2	$\boldsymbol{L}_2$		$\boldsymbol{W}_2(\times 10^{-3})$					
	SMB	HML	SG	SN	SV	BG	BN	BV
SG	.061	.126	.365	.847	.668	.396	.489	.283
SN	.035	.109	.257	.252	.842	.211	.589	.511
SV	.198	.186	.265	.278	.432	.023	.580	.681
BG	.013	.012	.074	.033	.005	.095	.342	−.053
BN	.076	−.032	.088	.088	.113	.031	.088	.613
BV	.236	−.143	.097	.145	.254	−.009	.103	.322

Estimated parameters of $\boldsymbol{L}$ and $\boldsymbol{W}$ in (4.20) for 6 assets, SG (Small Growth), SN (Small Neutral), SV (Small Value), BG (Big Growth), BN (Big Neutral) and BV (Big Value). The first 500 weekly data are used for estimation. Diagonal and lower triangular elements of $\boldsymbol{W}$ $(\times 10^{-3})$ are variance and covariance, respectively. Elements of $\boldsymbol{W}$ in the upper triangle with underline denote correlations.

Table 4.4: Estimated parameters of $\boldsymbol{\mu}, \boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}$ in (4.21) for factors

Single regime model					
	$\boldsymbol{\mu}(\times 10^{-3})$	$\boldsymbol{\Phi}$		$\boldsymbol{\Sigma}(\times 10^{-3})$	
		SMB	HML	SMB	HML
SMB	.713	−.091	.079	.140	<u>.029</u>
HML	.189	.124	−.049	.005	.172
Two regime model					
Regime 1	$\boldsymbol{\mu}_1(\times 10^{-3})$	$\boldsymbol{\Phi}_1$		$\boldsymbol{\Sigma}_1(\times 10^{-3})$	
		SMB	HML	SMB	HML
SMB	.654	−.083	−.018	.102	<u>−.062</u>
HML	.317	.068	.130	−.004	.049
Regime 2	$\boldsymbol{\mu}_2(\times 10^{-3})$	$\boldsymbol{\Phi}_2$		$\boldsymbol{\Sigma}_2(\times 10^{-3})$	
		SMB	HML	SMB	HML
SMB	1.029	−.114	.121	.240	<u>.098</u>
HML	−.348	.224	−0.130	.034	.489

Estimated parameters of $\boldsymbol{\mu}$, $\boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}$ in (4.21) for 2 factors, SMB (Small minus Big) and HML (High minus Low). The first 500 weekly data are used for estimation. Diagonal and lower triangular elements of $\boldsymbol{\Sigma}$ $(\times 10^{-3})$ are variance and covariance, respectively. Elements of $\boldsymbol{\Sigma}$ in the upper triangle with underline denote correlations.

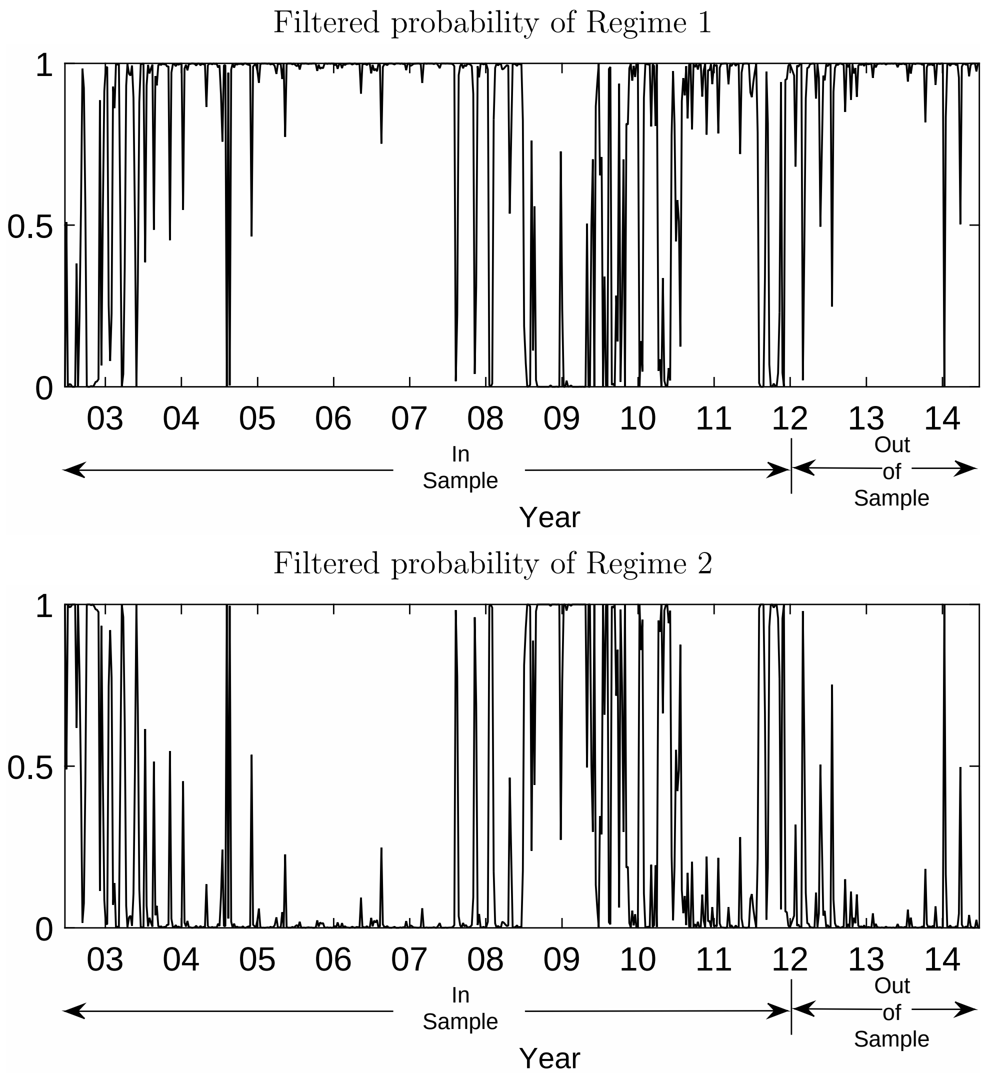


Figure 4.3: Filtered probabilities in the two regime model

model defined by

$$\boldsymbol{f}_i(t+1) = \boldsymbol{\mu}_i + \boldsymbol{\Phi}_i \boldsymbol{f}_i(t) + \boldsymbol{\epsilon}_i(t+1). \tag{4.24}$$

Since $\boldsymbol{\mu}_i$ and $\boldsymbol{\Phi}_i$ in (4.24) are constant over time and are the same as those in regime i in (4.2), $\boldsymbol{f}_i(t)$ is understood as a single regime factor process if the regime process is fixed to i. Assuming that (4.24) is stable, $\boldsymbol{f}_i(t)$ can be represented as an infinite order vector moving average process

$$\boldsymbol{f}_i(t) = (\boldsymbol{I} - \boldsymbol{\Phi}_i)^{-1}\boldsymbol{\mu}_i + \sum_{k=0}^{\infty} \boldsymbol{\Phi}_i^k \boldsymbol{\epsilon}_i(t-k). \tag{4.25}$$

Substituting (4.25) into (4.1) with $I(t) = i$ being fixed, we obtain

$$\boldsymbol{r}_i(t+1) = \boldsymbol{L}_i(\boldsymbol{I} - \boldsymbol{\Phi}_i)^{-1}\boldsymbol{\mu}_i + \boldsymbol{L}_i \sum_{k=0}^{\infty} \boldsymbol{\Phi}_i^k \boldsymbol{\epsilon}_i(t-k) + \boldsymbol{u}_i(t+1).$$

The mean and covariance matrix of $\boldsymbol{r}_i(t+1)$ are given by

$$\mathrm{E}\left(\boldsymbol{r}_i(t+1)\right) = \boldsymbol{L}_i(\boldsymbol{I} - \boldsymbol{\Phi}_i)^{-1}\boldsymbol{\mu}_i \tag{4.26}$$

$$\mathrm{V}\left(\boldsymbol{r}_i(t+1)\right) = \boldsymbol{L}_i \sum_{k=0}^{\infty} \boldsymbol{\Phi}_i^k \boldsymbol{\Sigma}_i (\boldsymbol{\Phi}_i^k)^\top \boldsymbol{L}_i^\top + \boldsymbol{W}_i. \tag{4.27}$$

It is remarked that (4.26) and (4.27) are independent of t since they are unconditional mean and variance.

Table 4.5 summarizes the unconditional mean and variance of 6 assets calculated from (4.26) and (4.27). The results for the three regime model are also listed for comparison. In the two regime model, we observe that Regime 1 represents low risk and high return while Regime 2 does high risk and low return. The single regime model lies between two regimes. As shown in Figure 4.3, the two regime model grasps that favorable markets as Regime 1 appear alternately with turbulent markets as Regime 2. These findings are consistent with previous studies. For example, Ang and Bekaert (2002a, 2004) report that volatilities and correlation among stock indices in developed countries increased simultaneously in one regime with much lower conditional mean. Guildolin and Timmermann (2004) identifies a similar set of two regimes for a value weighted index return of the US stocks.

With these observations for the two regime model in mind, we next discuss the estimation results for the three regime model. The Akaike's Information Criterion of 8711.9 is even better than 8844.4 of the two regime model. As for the two regime model, we identify Regimes 1, 2 and 3 in the ascending order of the sum of the variances over the six assets shown as the diagonal elements in $\boldsymbol{W}_i$. For notational simplicity, we denote Regime i in the J regime model by Regime i/J. For instance, Regime 2/3 indicates Regime 2 in the three regime model. From Table 4.5, it is noteworthy that the variances of 6 assets in Regime 1/3 are smaller than those in Regime 1/2, meaning that Regime 1/3 is more tranquil than Regime 1/2. Moreover, the expected returns in Regime 1/3 are higher than those in Regime 1/2 except SG and BG. Thus, Regime 1/3 represents more favorable state

Table 4.5: Unconditional mean and variance of the excess returns of 6 assets

	Single regime model ($\times 10^{-3}$)	Two regime model ($\times 10^{-3}$)	
		Regime 1	Regime 2
SG	.088 (.181)	.058 (.109)	.039 (.374)
SN	.067 (.115)	.082 (.062)	.018 (.258)
SV	.168 (.188)	.181 (.085)	.156 (.461)
BG	−.012 (.042)	−.035 (.023)	.010 (.095)
BN	.015 (.037)	.009 (.018)	.074 (.090)
BV	.051 (.122)	.055 (.040)	.233 (.345)

	Three regime model ($\times 10^{-3}$)		
	Regime 1	Regime 2	Regime 3
SG	−.085 (.078)	.092 (.170)	.065 (.461)
SN	.133 (.036)	.011 (.110)	.066 (.320)
SV	.221 (.040)	.142 (.170)	.081 (.587)
BG	−.087 (.022)	−.016 (.026)	−.000 (.137)
BN	.035 (.014)	−.020 (.025)	−.057 (.130)
BV	.144 (.038)	−.005 (.057)	−.215 (.514)

SG (Small Growth), SN (Small Neutral), SV (Small Value), BG (Big Growth), BN(Big Neutral), BV (Big Value). variance in parentheses of the market than Regime 1/2. On the other hand, Regime 3/3 is more turbulent than Regime 2/2 with lower returns except BG. Regime 2/3 stays in between Regime 1/3 and Regime 3/3.

Figure 4.4 shows the time series of the filtered probabilities in the three regime model. Comparing with the two regime model in Figure 4.3, it turns out that the filtered probabilities of Regime 2/2 (lower panel in Figure 4.3) is greater than or equal to those of Regime 3/3 (bottom panel in Figure 4.4) for most of the entire period. This implies that the three regime model picks up high risk periods more sharply than the two regime model. In fact, the three regime model finds that the Lehman Shock crisis ends around the middle of 2009 while the two regime model does not identify until late 2009 or 2010. The US equity markets bottomed out in March 2009.

Similarly to the high risk regime, the filtered probabilities of Regime 1/2 in Figure 4.3 (upper panel) is greater than or equal to those of Regime 1/3 in Figure 4.4 (top panel) for almost all time period. Altogether, Regime 1/2 in Figure 4.3 is decomposed into either Regime 1/3 or 2/3 in Figure 4.4 but not into Regime 3/3, and Regime 2/2 in Figure 4.3 is decomposed into either Regime 3/3 or 2/3 in Figure 4.4 but not into Regime 1/3. These observations are consistent with the estimated transition probability matrix for the three regime model

$$\boldsymbol{P} = \begin{bmatrix} .924 & .071 & .005 \\ .063 & .893 & .044 \\ .000 & .136 & .864 \end{bmatrix} \tag{4.28}$$

with very small probabilities of direct transitions between Regime 1/3 and Regime 3/3. In other words, Regime 2/3 intermediates Regime 1/3 and Regime 3/3 to switch from one to the other.

4.5 Investment Performance of the Optimal Portfolios

In this section, we show how the optimal solutions are workable when number of regimes increases and how discount rates can affect to the investment efficacies. Not only for those in the in-sample period, out-of-sample results are also presented in a later subsection. As a tool for building optimal portfolios and measurements of investment performance, we use the MATLAB running in the Intel(R) Core(TM) i7-4960X CPU 3.60GHz 6 Cores 12 Threads under 64bit operating system with 8G byte memory.

4.5.1 In-sample Period Performance of the Optimal Portfolios

In this subsection, we compare the performance of the optimal portfolio for the multiple regime models with that for the single regime model in the in-sample period. In calculating optimal portfolio holdings, at each point time t, we apply the predicted regime $I_t(t+1)$ as an argmax$_j\{p_{I(t),j}\}$ in $\boldsymbol{P}$ as described in Section 4.2 where $I(t)$ is estimated in a filtering algorithm described in Subsection 3.6.2. Besides to the parameters estimated from the data

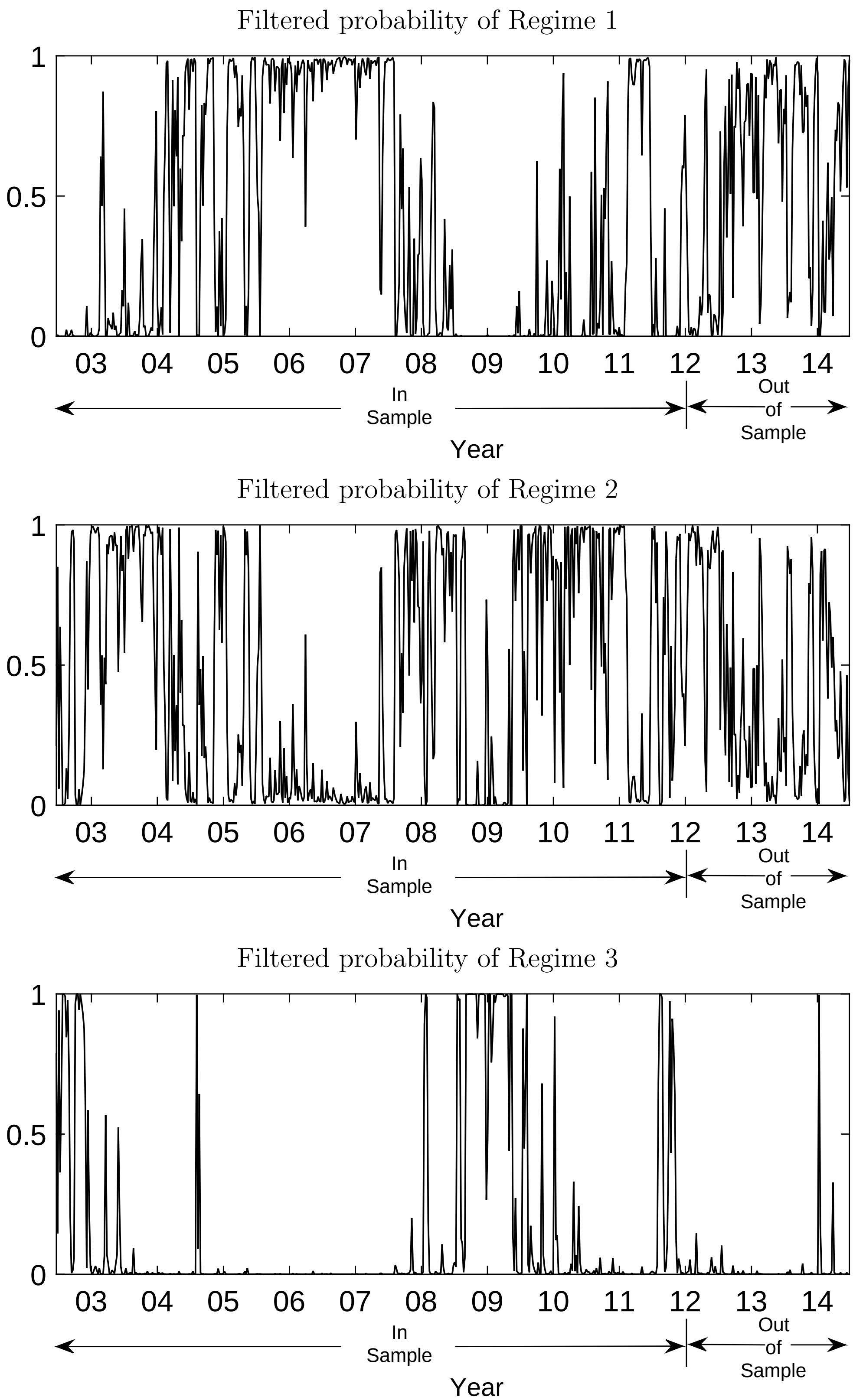


Figure 4.4: Filtered probabilities in the three regime model

in Section 4.4, we need to determine the coefficient of risk aversion and the transaction cost matrix for calculating the optimal portfolio. In a general practice under the mean-variance optimization, the risk aversion coefficient is given by the investor. In our empirical analysis, we fix $\lambda = 1$ for the single regime model and $\lambda_i = 1$ for all regime i in the multiple regime models. A choice of the common risk aversion parameter is expected not to penalize either the single or multiple regime models for preserving fairness in the intended comparisons.

The transaction cost matrix is difficult to estimate from data and is given by the investor in practice. In our empirical analysis, we set $\boldsymbol{B} = 0.02\boldsymbol{I}$ for the single regime model, $\boldsymbol{B}_1 = 0.01\boldsymbol{I}$ and $\boldsymbol{B}_2 = 0.05\boldsymbol{I}$ for the two regime model, and $\boldsymbol{B}_1 = 0.006\boldsymbol{I}$, $\boldsymbol{B}_2 = 0.02\boldsymbol{I}$ and $\boldsymbol{B}_3 = 0.06\boldsymbol{I}$ for the three regime model, because it is natural to suppose that the higher the volatility is, more expensive the transaction cost is. Several to ten times higher costs in the turbulent regimes are attributed to some of earlier studies such as Grinold and Kahn (1999), suggesting the transaction cost proportional to the volatility of the security. We recall that the covariance matrix in the turbulent regime is several times larger than that in tranquil regime (cf., Table 4.3). It should be remarked that comparison results for different number of regimes could be unfair if the transaction costs differ very much. In our analysis, the time average transaction cost per share equals $0.01 \times 377/498 + 0.05 \times 121/498 = 0.197$ for the two regime model, and $0.006 \times 199/498 + 0.02 \times 224/498 + 0.06 \times 75/498 = 0.204$ for the three regime model, both of which are close to that of the single regime model.

In addition to the risk aversion coefficient and the transaction cost matrix, a discount rate ρ needs to be given. The choice of the discount rate should be made according to the risk horizon of the investor. In the empirical analysis, we compare three types of discount rate $\rho = 0, 0.5$ and 0.9. As discussed in Section 4.3, $\rho = 0$ corresponds to single period optimization by a myopic investor. For $\rho = 0.9$ at the other end of spectrum, the future return 7 weeks ahead is discounted by 52% and 22 weeks ahead is discounted by 90%. That higher discount rate will be adopted by an investor whose risk horizon ranges over several months.

Although our optimization problem attempts to maximize the expected future utility, a performance measure of practical importance for comparing investment strategies is a realized utility evaluated from actual portfolio returns. In our model, the utility function (4.4) is defined by a mean-variance utility penalized for transaction costs. Since the actual excess return between t and $t+1$ is $\boldsymbol{x}^{*\top}_{I_t(t+1)}\boldsymbol{r}(t+1)$, the mean and the variance of the realized excess returns are defined by $\widehat{\alpha} = \frac{1}{498}\sum_{t=2}^{499} \boldsymbol{x}^{*\top}_{I_t(t+1)}(t)\boldsymbol{r}(t+1)$ and $\widehat{\sigma}^2 = \frac{1}{498-1}\sum_{t=2}^{499}(\boldsymbol{x}^{*\top}_{I_t(t+1)}\boldsymbol{r}(t+1) - \widehat{\alpha})^2$. Also, average transaction cost per week is $\widehat{c} = \frac{1}{498}\sum_{t=2}^{499}\{\boldsymbol{x}^*_{I_t(t+1)}(t) - \boldsymbol{x}^*_{I_{t-1}(t)}(t-1)\}^\top \boldsymbol{B}_{I_t(t+1)}\{\boldsymbol{x}^*_{I_t(t+1)}(t) - \boldsymbol{x}^*_{I_{t-1}(t)}(t-1)\}/2$. Based on these measures, we define realized gross and net utilities by $\widehat{\alpha} - \frac{\lambda}{2}\widehat{\sigma}^2$ and $\widehat{\alpha} - \frac{\lambda}{2}\widehat{\sigma}^2 - \widehat{c}$, respectively, where the coefficient of risk aversion is set to $\lambda = 1$. Another performance measure of practical importance is Sharpe ratio that notionally represents the excess return per unit of risk. In our empirical analysis, we define gross and net Sharpe ratio by $\widehat{\alpha}/\widehat{\sigma}$ and $(\widehat{\alpha} - \widehat{c})/\widehat{\sigma}$, respectively.

Table 4.6 summarizes the realized utilities and Sharpe ratios achieved by the optimal portfolio for the single, two, and three regime models in the in-sample period when the initial position of the portfolio equals **0**. For the gross and net utilities, multiple regime

Table 4.6: Realized utilities and Sharpe ratios in the in-sample period

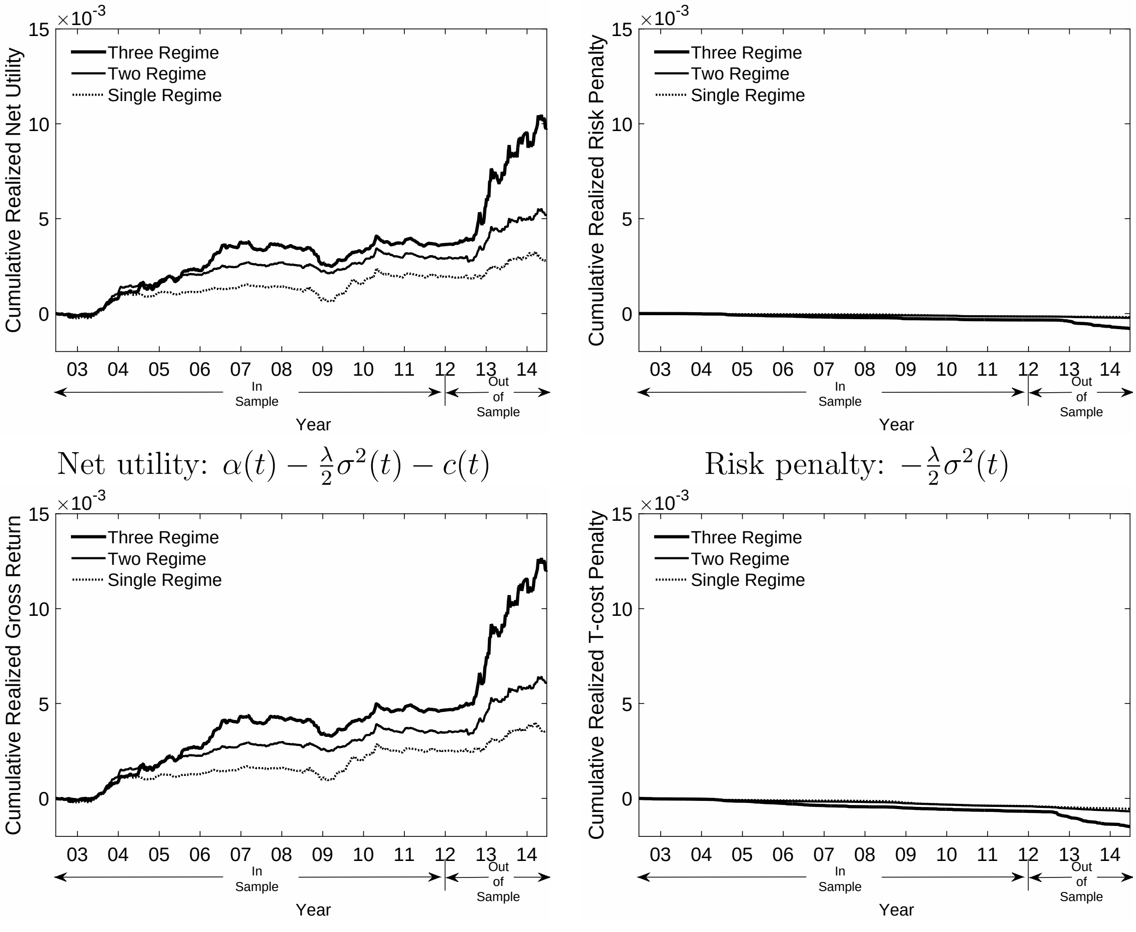
ρ	Model	Gross utility	Net utility	Gross Sharpe ratio	Net Sharpe ratio
0	Single regime	.0009	.0005	.056	.033
	Two regime	.0025	.0020	.106	.089
	Three regime	.0014	.0008	.055	.040
.5	Single regime	.0016	.0012	.099	.075
	Two regime	.0034	.0030	.165	.145
	Three regime	.0028	.0022	.095	.077
.9	Single regime	.0023	.0019	.142	.118
	Two regime	.0033	.0029	.199	.175
	Three regime	.0043	.0037	.181	.155

Gross and net utilities defined by $\widehat{\alpha} - \frac{\lambda}{2}\widehat{\sigma}^2$ and $\widehat{\alpha} - \frac{\lambda}{2}\widehat{\sigma}^2 - \widehat{c}$. Gross and net Sharpe ratios defined by $\widehat{\alpha}/\widehat{\sigma}$ and $(\widehat{\alpha} - \widehat{c})/\widehat{\sigma}$

models performs better than the single regime model regardless of discount rate ρ. For lower discount rate $\rho = 0$ and 0.5, the two regime model achieved the highest utilities while the three regime model outperforms for $\rho = 0.9$. Concerning Sharpe ratios, the two regime model performs the best for all discount rates. Also, it is worth noting that investment performance improves as ρ increases for all types of models and performance measures. This observation shows superiority of forward looking investment strategy rather than myopic strategy for $\rho = 0$.

To understand these comparison results more in detail, Figure 4.5 depicts the historical profile of the realized net utility $\alpha(t) - \frac{\lambda}{2}\sigma^2(t) - c(t)$ of the optimal portfolio for $\rho = 0.9$. Historical profiles of three components of the utility, i.e., gross return $\alpha(t) = \frac{1}{498}\sum_{s=2}^{t} \boldsymbol{x}_{I_s(s+1)}^{*\top}(s)\boldsymbol{r}(s+1)$, risk penalty $\sigma^2(t) = \frac{1}{497}\sum_{s=2}^{t}(\boldsymbol{x}_{I_s(s+1)}^{*\top}\boldsymbol{r}(s+1) - \widehat{\alpha})^2$ and transaction cost $c(t) = \frac{1}{498}\sum_{s=2}^{t}\{\boldsymbol{x}_{I_s(s+1)}^{*}(s) - \boldsymbol{x}_{I_{s-1}(s)}^{*}(s-1)\}^\top \boldsymbol{B}_{I_s(s+1)}\{\boldsymbol{x}_{I_s(s+1)}^{*}(s) - \boldsymbol{x}_{I_{s-1}(s)}^{*}(s-1)\}/2$ are also presented. These graphs uncover that the most significant contributor is the gross return to the superior result of multiple regime models. The risk penalty and transaction cost of the two regime model are not as significantly large as that can cancel out the decent gross return contribution. Since the gross return of the three regime model is higher than that of the two regime model, lower Sharpe ratios of the three regime model are attributed to larger transaction cost and, especially, larger risk penalty.

Another feature observed in the in-sample period of Figure 4.5 is that significant difference among performances of all models appear during the period from 2003 to the first half of 2007. Figure 4.6 shows the time series of holdings of 6 assets under the optimal portfolio for $\rho = 0.9$. Since the regime process mostly stays in Regime 1/2 in the two regime model (cf., Figure 4.3) and SV (Small Value) is the best performing asset in Regime 1/2 (cf., Table 4.5), the two regime model allocates the largest capital to SV while SG (Small Growth) and BG (Big Growth) are sold short due to largest variance and poorest expected return, respectively (cf., Table 4.5). Compared with the two regime model, the asset allocation



Net utility: $\alpha(t) - \frac{\lambda}{2}\sigma^2(t) - c(t)$

Risk penalty: $-\frac{\lambda}{2}\sigma^2(t)$

Gross return: $\alpha(t)$

Transaction cost: $-c(t)$

Figure 4.5: Historical profiles of the realized net utility of the optimal portfolio for $\rho = 0.9$ and three components, gross return, risk penalty and transaction cost

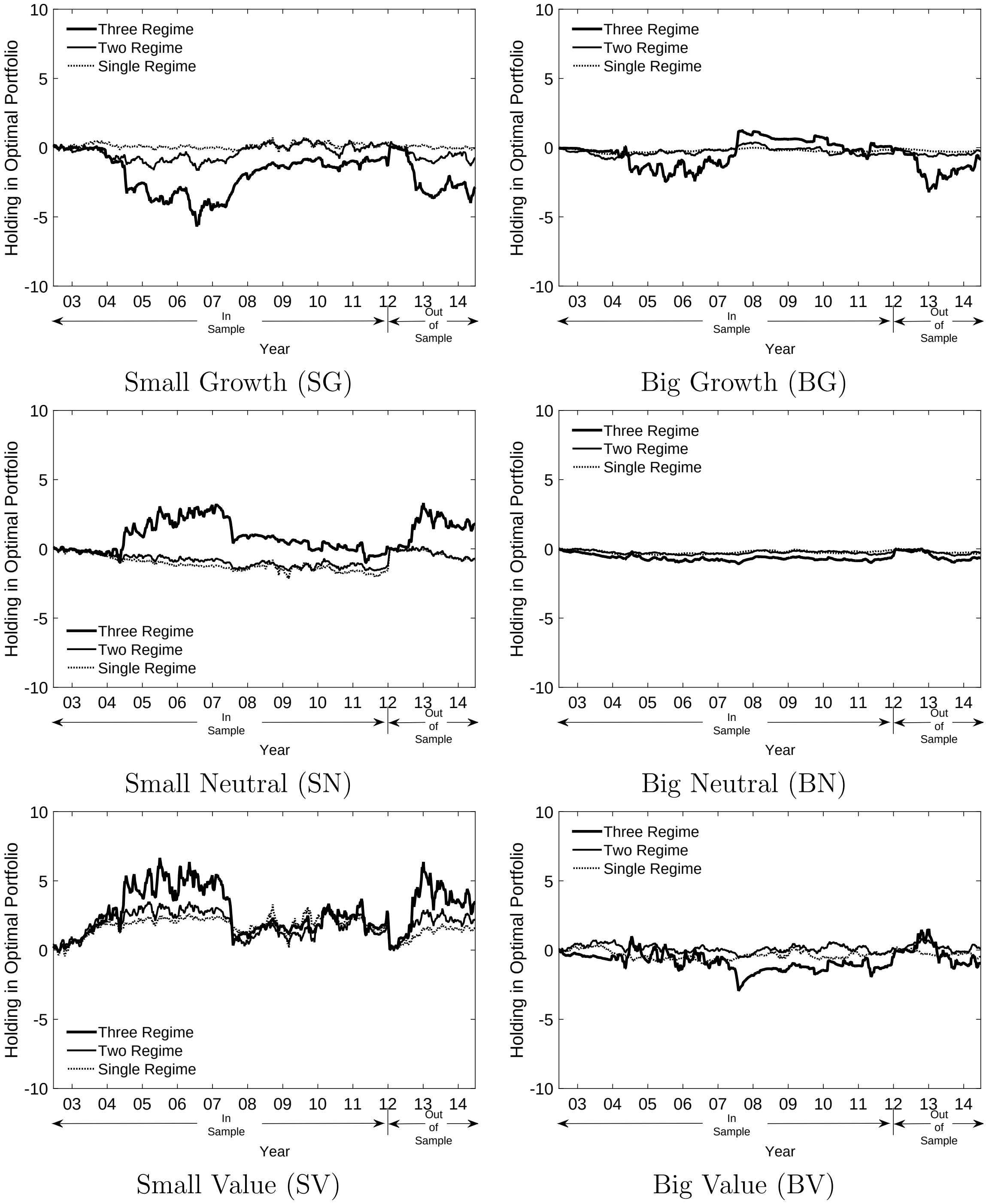


Figure 4.6: Time series of the holdings of 6 assets in the optimal portfolio for $\rho = 0.9$

of the three regime model is more aggressive. In addition to SV that is again the best performing asset in Regime 1/3, the optimal portfolio holds almost the same long position of SN (Small Neutral), while SG and BG are sold short several times more than the two regime model. The optimal portfolio with larger long and short positions observed in the three regime model results in higher gross return as well as larger transaction cost and risk penalty. This explains higher gross returns but lower Sharpe ratios of the three regime model. In contrast to the multiple regime models, the optimal portfolio for the single regime model does not follow the state changes of the market and the level of holdings of all assets are low, which leads to the underperformance. These observations indicate that introduction of regime switches enables investors to flexibly optimize asset allocation depending on the state of the market which leads to improvement of the investment strategy.

4.5.2 Out-of-sample Performance of the Optimal Portfolios

Following up the in-sample performance evaluation presented in Subsection 4.5.1, we show how the optimal portfolios behave and perform in the out-of-sample period. The out-of-sample period starts at the first week of January 2012 and ends at the last week of June 2014. Throughout 130 weeks in this period, we do not conduct any re-estimation of the model parameters but keep taking over those estimated in the in-sample period shown and the transition probability matrix. We also use the same risk aversion coefficients and transaction cost matrices as in Subsection 4.5.1 for comparing performance of portfolios.

Importing weekly updates of the realized returns of 2 factors and 6 assets while the model parameters are fixed, we calculate the filtered regime probabilities in the out-of-sample period as shown in Figure 4.3 for the two regime model and in Figure 4.4 for the three regime model. Similar to described in the Subsection 4.5.1, we apply the predicted regimes to calculated optimal portfolio holdings. The same performance measures as those in Table 4.6 for the in-sample comparison are used to conduct out-of-sample comparison.

Table 4.7 summarizes these performance measures in the out-of-sample period. We observe that almost all figures of the three regime model appear most attractive while the single regime model shows the poorest performance for all performance metrics and discount rates. The two regime model lies in between the single and the three regime models. We also note that, as for the in-sample comparison in Table 4.6, investment performance improves as ρ increases for all types of models and performance measures, indicating superiority of forward looking investment strategy rather than myopic strategy.

Figure 4.5 extends all profiles for $\rho = 0.9$ to the out-of-sample period over the realized net utility and three components of the utility, i.e., $\alpha(t)$ as gross return, $\sigma^2(t)$ as risk penalty and $c(t)$ as transaction cost. All of four measures are calculated in the same way as in the in-sample period. Even in the out-of-sample period, the gross return again plays a crucial role for the superior results delivered by the multiple regime models. The multiple regime models, especially the three regime model, enjoy superior results in 2013 and 2014 relative to those in the single regime model with a rapid hike of the gross return at the expense of marginal increase of the risk penalty and the transaction cost. Figure 4.6 also

Table 4.7: Realized utilities and Sharpe ratios in the out-of-sample period

ρ	Model	Gross utility	Net utility	Gross Sharpe ratio	Net Sharpe ratio
0	Single regime	0.0006	0.0005	0.093	0.070
	Two regime	0.0021	0.0019	0.147	0.129
	Three regime	0.0066	0.0060	0.177	0.164
.5	Single regime	0.0008	0.0006	0.115	0.093
	Two regime	0.0024	0.0021	0.165	0.146
	Three regime	0.0074	0.0067	0.197	0.180
.9	Single regime	0.0010	0.0009	0.139	0.120
	Two regime	0.0026	0.0023	0.229	0.197
	Three regime	0.0070	0.0062	0.249	0.223

Gross and net utilities defined by $\widehat{\alpha} - \frac{\lambda}{2}\widehat{\sigma}^2$ and $\widehat{\alpha} - \frac{\lambda}{2}\widehat{\sigma}^2 - \widehat{c}$. Gross and net Sharpe ratios defined by $\widehat{\alpha}/\widehat{\sigma}$ and $(\widehat{\alpha} - \widehat{c})/\widehat{\sigma}$

extends to the out-of-sample period showing the time series of holdings of 6 assets under the optimal portfolio for $\rho = 0.9$. As is the case in the in-sample period, the multiple regime portfolios under Regime 1, which is a vast majority in the out-of-sample period, hold SV (Small Value) more than the single regime does, leading to the superior performance of the multiple regime models as a result. In addition, the three regime model holds larger long position of SN (Small Neutral) and larger short positions of SG (Small Growth) and BG (Big Growth) to the two regime model, bringing out the better performance.

4.6 Concluding Remarks

In this chapter, we investigate effects of regime switches in factor models that predict asset returns under multi-period optimality. The model of Gârleanu and Pedersen (2013) is extended to take regime switches into account so that the model parameters change over time according to state changes of the market. The main analytical result reveals that the optimal portfolio is the weighted average of the current portfolio and a target portfolio where the target portfolio and the weight are regime dependent.

Empirical results for the last decade, applying the derived solution to an equity portfolio investing into size and book-to-market assets, demonstrates superior outcomes by the two and three regime models for both in-sample and out-of-sample examinations. Intuitively the multiple regime approach captures such structural changes, more effectively than the statistical change in the single regime approach, as the burst of technology bubble, credit bubble followed by the US and European sovereign crisis, experienced in the last decade.

As next steps, additional research is planned on a couple of fronts to extend the model applied to the empirical studies. First, the transition probabilities are augmented to be time variant described by the Markov switching logistic function of such exogenous variables as market data and macro economic data. Second, assumed probability distribution for the factors and assets is generalized from the normal distribution to a non-Gaussian distribution such as t-distribution that generally shows better fit to actual market return data. Other extensions of the model of practical importance include more general transaction cost functions such as linear transaction cost with constant.

4.7 Appendix

4.7.1 Proofs

Proof of Lemma 4.3.1 Let $\mathcal{M}$ denote the set of all $N \times N$ symmetric positive definite matrices. For $\boldsymbol{D}_1, \boldsymbol{D}_2 \in \mathcal{M}$, we define $\boldsymbol{D}_1 \prec \boldsymbol{D}_2$ if $\boldsymbol{D}_2 - \boldsymbol{D}_1 \in \mathcal{M}$. We also define

$$f_i(\boldsymbol{D}) = \boldsymbol{B}_i - \boldsymbol{B}_i(\boldsymbol{D} + \boldsymbol{A}_i + \boldsymbol{B}_i)^{-1}\boldsymbol{B}_i, \quad i = 1, \dots, J \tag{4.29}$$

for $\boldsymbol{D} \in \mathcal{M}$ where the invertibility of $\boldsymbol{D} + \boldsymbol{A}_i + \boldsymbol{B}_i$ is assured since $\boldsymbol{D} + \boldsymbol{A}_i + \boldsymbol{B}_i \in \mathcal{M}$. Then, $f_i(\boldsymbol{D})$ has the following properties: (i) $f_i(\boldsymbol{D}) \in \mathcal{M}$ for any $\boldsymbol{D} \in \mathcal{M}$, (ii) if $\boldsymbol{D}_1 \prec \boldsymbol{D}_2$ then $f_i(\boldsymbol{D}_1) \prec f_i(\boldsymbol{D}_2)$, (iii) $f_i(\boldsymbol{D}) \prec \boldsymbol{B}_i$ for any $\boldsymbol{D} \in \mathcal{M}$. To see these properties, note that $\boldsymbol{D}_1 \prec \boldsymbol{D}_2$ if and only if $\boldsymbol{D}_1^{-1} \succ \boldsymbol{D}_2^{-1}$, see e.g., Bhatia (1997). Then, (i) is easily seen from $\boldsymbol{B}_i(\boldsymbol{D} + \boldsymbol{A}_i + \boldsymbol{B}_i)^{-1}\boldsymbol{B}_i \prec \boldsymbol{B}_i\boldsymbol{B}_i^{-1}\boldsymbol{B}_i = \boldsymbol{B}_i$. To prove (ii), it is enough to note that $\boldsymbol{D}_1 \prec \boldsymbol{D}_2$ implies $(\boldsymbol{D}_1 + \boldsymbol{A}_i + \boldsymbol{B}_i)^{-1} \succ (\boldsymbol{D}_2 + \boldsymbol{A}_i + \boldsymbol{B}_i)^{-1}$. (iii) is obvious since $\boldsymbol{B}_i(\boldsymbol{D} + \boldsymbol{A}_i + \boldsymbol{B}_i)^{-1}\boldsymbol{B}_i \in \mathcal{M}$. For $\{\boldsymbol{D}_i\} = \{\boldsymbol{D}_1, \dots, \boldsymbol{D}_J\} \in \mathcal{M}^J$, we further define a function F on $\mathcal{M}^J$ by

$$F(\{\boldsymbol{D}_i\}) = \{f_1(\rho\langle\{\boldsymbol{D}_i\}\rangle_1), \dots, f_J(\rho\langle\{\boldsymbol{D}_i\}\rangle_J)\} \tag{4.30}$$

where $\langle\{\boldsymbol{D}_i\}\rangle_j$ is defined in (4.8). From $\langle\{\boldsymbol{D}_i\}\rangle_j \in \mathcal{M}$ and property (i) above, $F(\{\boldsymbol{D}_i\}) \in \mathcal{M}^J$. Noting that (4.7) is rewritten as

$$\{\boldsymbol{\beta}_i^{(n)}\} = F(\{\boldsymbol{\beta}_i^{(n-1)}\}), \tag{4.31}$$

$\boldsymbol{\beta}_i^{(n)} \in \mathcal{M}$ for all $i = 1, \dots, J$ and $n \geq 1$. Using (4.31) and property (ii), an inductive argument applied to $\{\boldsymbol{\beta}_i^{(n)}\}$ proves that $\boldsymbol{\beta}_i^{(n-1)} \prec \boldsymbol{\beta}_i^{(n)}$ for all $i = 1, \dots, J$ and $n \geq 1$. Moreover, $\boldsymbol{\beta}_i^{(n)} \prec \boldsymbol{B}_i$ holds for all $i = 1, \dots, J$ and $n \geq 1$ from property (iii). Therefore, $\boldsymbol{x}^\top\boldsymbol{\beta}_i^{(n)}\boldsymbol{x}$ for $\boldsymbol{x} \neq \boldsymbol{0}$ forms an increasing sequence bounded from above by $\boldsymbol{x}^\top\boldsymbol{B}_i\boldsymbol{x}$. This implies that, $\lim_{n\to\infty} \boldsymbol{x}^\top\boldsymbol{\beta}_i^{(n)}\boldsymbol{x}$ exists and is positive for arbitrary $\boldsymbol{x} \neq \boldsymbol{0}$. Now let $\boldsymbol{e}_\ell$ denote a unit column vector whose ℓth element is 1 and others are 0. Noting that $\boldsymbol{e}_\ell^\top\boldsymbol{\beta}_i^{(n)}\boldsymbol{e}_\ell = \boldsymbol{\beta}_i^{(n)}(\ell, \ell)$ where $\boldsymbol{D}(\ell, m)$ denotes (ℓ, m)th element of a matrix $\boldsymbol{D}$, $\boldsymbol{\beta}_i^{(\infty)}(\ell, \ell) = \lim_{n\to\infty} \boldsymbol{\beta}_i^{(n)}(\ell, \ell)$ exists. For $\boldsymbol{x} = \boldsymbol{e}_\ell + \boldsymbol{e}_m$ $(\ell \neq m)$, we instead know that $\boldsymbol{x}^\top\boldsymbol{\beta}_i^{(n)}\boldsymbol{x} = \boldsymbol{\beta}_i^{(n)}(\ell, \ell) + \boldsymbol{\beta}_i^{(n)}(m, m) + \boldsymbol{\beta}_i^{(n)}(\ell, m) + \boldsymbol{\beta}_i^{(n)}(m, \ell)$ is convergent. Since $\boldsymbol{\beta}_i^{(n)}$ is symmetric and $\boldsymbol{\beta}_i^{(n)}(\ell, \ell)$ and $\boldsymbol{\beta}_i^{(n)}(m, m)$ are convergent, $\boldsymbol{\beta}_i^{(n)}(\ell, m) = \boldsymbol{\beta}_i^{(n)}(m, \ell)$ converges to $\boldsymbol{\beta}_i^{(\infty)}(\ell, m) = \boldsymbol{\beta}_i^{(\infty)}(m, \ell)$. Hence, the limiting matrix $\boldsymbol{\beta}_i^{(\infty)} = \lim_{n\to\infty} \boldsymbol{\beta}_i^{(n)}$ exists and $\boldsymbol{\beta}_i^{(\infty)} \in \mathcal{M}$ for all $i = 1, \dots, J$. □

Proofs of Proposition 4.3.2 and Proposition 4.3.3 By the principle of optimality, $V_i(\boldsymbol{y}, \boldsymbol{f})$ satisfies simultaneous Bellman's equations

$$V_i(\boldsymbol{y},\boldsymbol{f}) = \max_{\boldsymbol{x}}\Big[\boldsymbol{x}^\top \boldsymbol{L}_i \boldsymbol{f} - \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{A}_i \boldsymbol{x} - \frac{1}{2}(\boldsymbol{x}-\boldsymbol{y})^\top \boldsymbol{B}_i(\boldsymbol{x}-\boldsymbol{y}) + \rho\sum_{j=1}^{J} p_{ij}\mathrm{E}\left(V_j(\boldsymbol{x}, \boldsymbol{\mu}_j + \boldsymbol{\Phi}_j\boldsymbol{f} + \boldsymbol{\epsilon}_j)\right)\Big], \quad i = 1,\ldots,J \tag{4.32}$$

where $\boldsymbol{\epsilon}_j$ denotes a noise term in (4.2) conditioned on $I(2) = j$, hence $\mathrm{E}(\boldsymbol{\epsilon}_j) = \boldsymbol{0}$ and $\mathrm{V}(\boldsymbol{\epsilon}_j) = \boldsymbol{\Sigma}_j$. By substituting (4.6) into the right hand side of (4.32) and taking expectation with respect to $\boldsymbol{\epsilon}_j$, the set of functions to be maximized in (4.32) is reduced to a system of quadratic functions

$$\begin{aligned} G_i(\boldsymbol{x};\boldsymbol{f}) \;=\;& -\frac{1}{2}\boldsymbol{x}^\top(\rho\langle\boldsymbol{\beta}\rangle_i + \boldsymbol{A}_i + \boldsymbol{B}_i)\boldsymbol{x} + \left\{\rho(\langle\boldsymbol{\delta}\rangle_i + \langle\boldsymbol{\kappa}\boldsymbol{\mu}\rangle_i)^\top + \boldsymbol{y}^\top\boldsymbol{B}_i + \boldsymbol{f}^\top(\rho\langle\boldsymbol{\kappa}\boldsymbol{\Phi}\rangle_i + \boldsymbol{L}_i)^\top\right\}\boldsymbol{x} \\ & + \frac{\rho}{2}\boldsymbol{f}^\top\langle\boldsymbol{\Phi}^\top\boldsymbol{\eta}\boldsymbol{\Phi}\rangle_i\boldsymbol{f} + \rho(\langle\boldsymbol{\mu}^\top\boldsymbol{\eta}\boldsymbol{\Phi}\rangle_i + \langle\boldsymbol{\xi}^\top\boldsymbol{\Phi}\rangle_i)\boldsymbol{f} + \frac{\rho}{2}\langle\boldsymbol{\mu}^\top\boldsymbol{\eta}\boldsymbol{\mu}\rangle_i + \frac{\rho}{2}\mathrm{E}\left(\langle\boldsymbol{\epsilon}^\top\boldsymbol{\eta}\boldsymbol{\epsilon}\rangle_i\right) \\ & + \rho\langle\boldsymbol{\xi}^\top\boldsymbol{\mu}\rangle_i + \rho\langle\zeta\rangle_i - \frac{1}{2}\boldsymbol{y}^\top\boldsymbol{B}_i\boldsymbol{y}, \quad i = 1,\ldots,J. \end{aligned} \tag{4.33}$$

For $\boldsymbol{x} = (x_1,\ldots,x_N)^\top$, let $\partial/\partial\boldsymbol{x} = [\partial/\partial x_1,\ldots,\partial/\partial x_N]^\top$. Since

$$\frac{\partial}{\partial\boldsymbol{x}}(\boldsymbol{h}^\top\boldsymbol{x}) = \boldsymbol{h}, \qquad \frac{\partial}{\partial\boldsymbol{x}}(\boldsymbol{x}^\top\boldsymbol{H}\boldsymbol{x}) = (\boldsymbol{H} + \boldsymbol{H}^\top)\boldsymbol{x}$$

for an $N \times 1$ vector $\boldsymbol{h}$ and an $N \times N$ matrix $\boldsymbol{H}$, we obtain

$$\frac{\partial G_i(\boldsymbol{x};\boldsymbol{f})}{\partial\boldsymbol{x}} = -(\rho\langle\boldsymbol{\beta}\rangle_i + \boldsymbol{A}_i + \boldsymbol{B}_i)\boldsymbol{x} + \rho(\langle\boldsymbol{\delta}\rangle_i + \langle\boldsymbol{\kappa}\boldsymbol{\mu}\rangle_i) + \boldsymbol{B}_i\boldsymbol{y} + (\rho\langle\boldsymbol{\kappa}\boldsymbol{\Phi}\rangle_i + \boldsymbol{L}_i)\boldsymbol{f}, \quad i = 1,\ldots,J.$$

The first order optimality condition then gives the optimal strategy

$$\boldsymbol{x}_i^* = \boldsymbol{C}_i\{\rho(\langle\boldsymbol{\delta}\rangle_i + \langle\boldsymbol{\kappa}\boldsymbol{\mu}\rangle_i) + \boldsymbol{B}_i\boldsymbol{y} + (\rho\langle\boldsymbol{\kappa}\boldsymbol{\Phi}\rangle_i + \boldsymbol{L}_i)\boldsymbol{f}\}, \quad i = 1,\ldots,J \tag{4.34}$$

where $\boldsymbol{C}_i$ is defined in (4.10). Substituting (4.34) into (4.33) and equating it with (4.6), solving (4.32) is reduced to finding the solutions to the following system of equations for unknown coefficients of the value functions:

$$\boldsymbol{\beta}_i = \boldsymbol{B}_i - \boldsymbol{B}_i\boldsymbol{C}_i\boldsymbol{B}_i = \boldsymbol{B}_i - \boldsymbol{B}_i(\rho\langle\boldsymbol{\beta}\rangle_i + \boldsymbol{A}_i + \boldsymbol{B}_i)^{-1}\boldsymbol{B}_i, \quad i = 1,\ldots,J \tag{4.35}$$

$$\boldsymbol{\delta}_i = \rho\boldsymbol{B}_i\boldsymbol{C}_i(\langle\boldsymbol{\delta}\rangle_i + \langle\boldsymbol{\kappa}\boldsymbol{\mu}\rangle_i), \quad i = 1,\ldots,J \tag{4.36}$$

$$\boldsymbol{\eta}_i = \rho\langle\boldsymbol{\Phi}^\top\boldsymbol{\eta}\boldsymbol{\Phi}\rangle_i + (\rho\langle\boldsymbol{\kappa}\boldsymbol{\Phi}\rangle_i + \boldsymbol{L}_i)^\top\boldsymbol{C}_i(\rho\langle\boldsymbol{\kappa}\boldsymbol{\Phi}\rangle_i + \boldsymbol{L}_i), \quad i = 1,\ldots,J \tag{4.37}$$

$$\boldsymbol{\xi}_i = \rho\left\{\langle\boldsymbol{\Phi}^\top\boldsymbol{\eta}^\top\boldsymbol{\mu}\rangle_i + \langle\boldsymbol{\Phi}^\top\boldsymbol{\xi}\rangle_i) + (\rho\langle\boldsymbol{\kappa}\boldsymbol{\Phi}\rangle_i + \boldsymbol{L}_i)^\top\boldsymbol{C}_i(\langle\boldsymbol{\delta}\rangle_i + \langle\boldsymbol{\kappa}\boldsymbol{\mu}\rangle_i)\right\}, \quad i = 1,\ldots,J \tag{4.38}$$

$$\boldsymbol{\kappa}_i = \boldsymbol{B}_i\boldsymbol{C}_i(\rho\langle\boldsymbol{\kappa}\boldsymbol{\Phi}\rangle_i + \boldsymbol{L}_i), \quad i = 1,\ldots,J \tag{4.39}$$

$$\zeta_i = \rho\Big\{\frac{1}{2}\langle\boldsymbol{\mu}^\top\boldsymbol{\eta}\boldsymbol{\mu}\rangle_i + \frac{1}{2}\mathrm{E}\left(\langle\boldsymbol{\epsilon}^\top\boldsymbol{\eta}\boldsymbol{\epsilon}\rangle_i\right) + \langle\boldsymbol{\xi}^\top\boldsymbol{\mu}\rangle_i + \langle\zeta\rangle_i + \frac{1}{2}\rho(\langle\boldsymbol{\delta}\rangle_i + \langle\boldsymbol{\kappa}\boldsymbol{\mu}\rangle_i)^\top\boldsymbol{C}_i(\langle\boldsymbol{\delta}\rangle_i + \langle\boldsymbol{\kappa}\boldsymbol{\mu}\rangle_i)\Big\}, \quad i = 1,\ldots,J. \tag{4.40}$$

From (4.35), (4.36) and (4.39), (4.34) is rewritten as (4.15) with previous portfolio $\boldsymbol{y} = \boldsymbol{x}(t-1)$ and current factor $\boldsymbol{f} = \boldsymbol{f}(t)$.

The rest of the proof is to show that $\{\boldsymbol{\beta}_i\}$ $\{\boldsymbol{\kappa}_i\}$, $\{\boldsymbol{\delta}_i\}$, $\{\boldsymbol{\eta}_i\}$, $\{\boldsymbol{\xi}_i\}$ and $\{\zeta_i\}$ given in Proposition 4.3.2(a)∼(f) satisfy (4.35)∼(4.40). Then, the set of guess solutions $\{V_1, \dots, V_J\}$ in (4.6) satisfies the simultaneous Bellman's equations (4.32) and the optimal portfolio is given by (4.34).

(a) Since the determinant of a matrix is continuous, $F(\{\boldsymbol{D}_i\})$ in (4.30) is continuous in the sense that, if a sequence $\{\boldsymbol{D}_i^{(n)}\} \in \mathcal{M}^J$ converges elementwise to $\{\boldsymbol{D}_i^{(\infty)}\}$ as $n \to \infty$ and $\boldsymbol{D}_i^{(\infty)} + \boldsymbol{A}_i + \boldsymbol{B}_i$ is invertible for all $i = 1, \dots, J$, then $\lim_{n\to\infty} F(\{\boldsymbol{D}_i^{(n)}\}) = F(\{\boldsymbol{D}_i^{(\infty)}\})$. Letting $n \to \infty$ in (4.31) then shows $\{\boldsymbol{\beta}_i\} = \{\boldsymbol{\beta}_i^{(\infty)}\}$ satisfies (4.35) simultaneously for all $i = 1, \dots, J$.

(b) From Proposition 10.4 of Hamilton (1994),

$$\mathrm{vec}(\boldsymbol{XYZ}) = (\boldsymbol{Z}^\top \otimes \boldsymbol{X})\mathrm{vec}(\boldsymbol{Y}) \tag{4.41}$$

holds for matrices $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$. Applying (4.41) to (4.39), we get

$$\mathrm{vec}(\boldsymbol{\kappa}_i) = \rho \sum_{k=1}^{J} p_{ik}\{\boldsymbol{\Phi}_k^\top \otimes (\boldsymbol{B}_i\boldsymbol{C}_i)\}\mathrm{vec}(\boldsymbol{\kappa}_k) + \mathrm{vec}(\boldsymbol{B}_i\boldsymbol{C}_i\boldsymbol{L}_i), \quad i = 1, \dots, J$$

that proves (4.9).

(c) It is enough to rewrite (4.36) as

$$\begin{bmatrix} \boldsymbol{\delta}_1 \\ \vdots \\ \boldsymbol{\delta}_J \end{bmatrix} = \rho\boldsymbol{\Theta} \begin{bmatrix} \boldsymbol{\delta}_1 \\ \vdots \\ \boldsymbol{\delta}_J \end{bmatrix} + \rho \begin{bmatrix} \boldsymbol{B}_1\boldsymbol{C}_1\langle\boldsymbol{\kappa\mu}\rangle_1 \\ \vdots \\ \boldsymbol{B}_J\boldsymbol{C}_J\langle\boldsymbol{\kappa\mu}\rangle_J \end{bmatrix}.$$

(d) We rewrite (4.37) as

$$\boldsymbol{\eta}_i = \rho \sum_{k=1}^{J} p_{ik}\boldsymbol{\Phi}_k^\top \boldsymbol{\eta}_k \boldsymbol{\Phi}_k + (\rho\langle\boldsymbol{\kappa\Phi}\rangle_i + \boldsymbol{L}_i)^\top \boldsymbol{C}_i(\rho\langle\boldsymbol{\kappa\Phi}\rangle_i + \boldsymbol{L}_i), \quad i = 1, \dots, J. \tag{4.42}$$

Applying (4.41) to (4.42), we obtain

$$\mathrm{vec}(\boldsymbol{\eta}_i) = \rho \sum_{k=1}^{J} p_{ik}(\boldsymbol{\Phi}_k^\top \otimes \boldsymbol{\Phi}_k^\top)\mathrm{vec}(\boldsymbol{\eta}_k) + \mathrm{vec}((\rho\langle\boldsymbol{\kappa\Phi}\rangle_i + \boldsymbol{L}_i)^\top \boldsymbol{C}_i\langle\boldsymbol{\kappa\Phi}\rangle_i + \boldsymbol{L}_i)), \quad i = 1, \dots, J$$

from which (4.12) is derived. To show $\boldsymbol{\eta}_i \in \mathcal{M}$, we recursively construct a set of matrices $\{\boldsymbol{\eta}_i^{(n)}\} = \{\boldsymbol{\eta}_1^{(n)}, \dots, \boldsymbol{\eta}_J^{(n)}\}$ by

$$\boldsymbol{\eta}_i^{(n)} = \rho \sum_{k=1}^{J} p_{ik}\boldsymbol{\Phi}_k^\top \boldsymbol{\eta}_k^{(n-1)} \boldsymbol{\Phi}_k + (\rho\langle\boldsymbol{\kappa\Phi}\rangle_i + \boldsymbol{L}_i)^\top \boldsymbol{C}_i(\rho\langle\boldsymbol{\kappa\Phi}\rangle_i + \boldsymbol{L}_i), \quad i = 1, \dots, J$$

starting with $\{\boldsymbol{\eta}_1^{(0)}, \ldots, \boldsymbol{\eta}_J^{(0)}\} = \{\boldsymbol{O}, \ldots, \boldsymbol{O}\}$. Following the similar argument in the proof of Lemma 4.3.1, we prove that $\boldsymbol{\eta}_i^{(n)} \in \mathcal{M}$, $\boldsymbol{\eta}_i^{(n-1)} \prec \boldsymbol{\eta}_i^{(n)}$, and $\boldsymbol{\eta}_i^{(n)} \prec \boldsymbol{\eta}_i$, for all $i = 1, \ldots, J$ and $n \geq 1$. Thus, $\{\boldsymbol{\eta}_i^{(n)}\}$ converges to $\{\boldsymbol{\eta}_i\}$ as $n \to \infty$, implying that $\boldsymbol{\eta}_i \in \mathcal{M}$ for all $i = 1, \ldots, J$.
(e) (4.13) is easily obtained since (4.38) is rewritten as

$$\boldsymbol{\xi}_i = \rho \sum_{k=1}^{J} p_{ik} \boldsymbol{\Phi}_k^\top \boldsymbol{\xi}_k + \rho \langle \boldsymbol{\Phi}^\top \boldsymbol{\eta}^\top \boldsymbol{\mu} \rangle_i + \rho(\rho \langle \boldsymbol{\kappa} \boldsymbol{\Phi} \rangle_i + \boldsymbol{L}_i)^\top \boldsymbol{C}_i (\langle \boldsymbol{\delta} \rangle_i + \langle \boldsymbol{\kappa} \boldsymbol{\mu} \rangle_i).$$

(f) We get (4.14) from (4.40) since the spectral radius of the transition probability matrix $\boldsymbol{P}$ is 1 and hence $\boldsymbol{I} - \rho \boldsymbol{P}$ is invertible. □

4.7.2 Householder Transformation

We rearrange (4.1) for return forecasts of assets in an $N \times 1$ vector $\boldsymbol{r}(t+1)$ by factors in a $K \times 1$ vector $\boldsymbol{f}(t)$ into

$$\boldsymbol{r}(t+1) = \boldsymbol{A}_{I(t+1)} \boldsymbol{r}(t+1) + \boldsymbol{D}_{I(t+1)} \boldsymbol{f}(t) + \boldsymbol{s}_{I(t+1)}(t+1)$$

$$\boldsymbol{s}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{S}_i)$$

when $I(t+1) = i$. Coefficients in (4.43) are transformed into

$$\boldsymbol{A}_i = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ a_i(2,1) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ a_i(N,1) & \cdots & a_i(N,N-1) & 0 \end{bmatrix}$$

$$\boldsymbol{D}_i = \begin{bmatrix} b_{1,i}(1,1) & \cdots & b_{1,i}(1,K) \\ \vdots & \ddots & \vdots \\ b_{1,i}(N,1) & \cdots & b_{1,i}(N,K) \end{bmatrix}$$

$$\boldsymbol{S}_i = \begin{bmatrix} \sigma_{1,i}^2 & & & O \\ & \sigma_{2,i}^2 & & \\ & & \ddots & \\ O & & & \sigma_{N,i}^2 \end{bmatrix}.$$

An inverse transformation provides the original coefficients in (4.1) with

$$\boldsymbol{L}_i = (\boldsymbol{I} - \boldsymbol{A}_i)^{-1} \boldsymbol{D}_i$$

$$\boldsymbol{W}_i = (\boldsymbol{I} - \boldsymbol{A}_i)^{-1} \boldsymbol{S}_i (\boldsymbol{I} - \boldsymbol{A}_i)^{-\top}.$$

Similarly, (4.2) for factors in a $K \times 1$ vector $\boldsymbol{f}(t)$ in VAR(1) is rearranged into

$$\boldsymbol{f}(t+1) = \boldsymbol{C}_{I(t+1)} + \boldsymbol{B}_{0,I(t+1)} \boldsymbol{f}(t+1) + \boldsymbol{B}_{1,I(t+1)} \boldsymbol{f}(t) + \boldsymbol{v}_{I(t+1)}(t+1)$$

$$\boldsymbol{v}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{V}_i)$$

when $I(t+1) = i$. Coefficients in (4.43) are transformed into

$$\boldsymbol{C}_i = \begin{bmatrix} c_i(1,1) \\ c_i(2,1) \\ \vdots \\ c_i(K,1) \end{bmatrix}$$

$$\boldsymbol{B}_{0,i} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ b_{0,i}(2,1) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ b_{0,i}(K,1) & \cdots & b_{0,i}(K,K-1) & 0 \end{bmatrix}$$

$$\boldsymbol{B}_{1,i} = \begin{bmatrix} b_{1,i}(1,1) & \cdots & b_{1,i}(1,K) \\ \vdots & \ddots & \vdots \\ b_{1,i}(K,1) & \cdots & b_{1,i}(K,K) \end{bmatrix}$$

$$\boldsymbol{V}_i = \begin{bmatrix} v_{1,i}^2 & & & O \\ & v_{2,i}^2 & & \\ & & \ddots & \\ O & & & v_{K,i}^2 \end{bmatrix}.$$

An inverse transformation provides the original coefficients in (4.2) with

$$\boldsymbol{\mu}_i = (\boldsymbol{I} - \boldsymbol{B}_{0,i})^{-1}\boldsymbol{C}_i$$

$$\boldsymbol{\Phi}_i = (\boldsymbol{I} - \boldsymbol{B}_{0,i})^{-1}\boldsymbol{B}_{1,i}$$

$$\boldsymbol{\Sigma_i} = (\boldsymbol{I} - \boldsymbol{B}_{0,i})^{-1}\boldsymbol{V}_i(\boldsymbol{I} - \boldsymbol{B}_{0,i})^{-\top}.$$

Chapter 5

Linear Rebalancing Strategy for Short Sales Constraint

5.1 Introduction

Across all investment layers ranging from asset allocation to individual securities selections, quantitative models predict expected returns and variability. The models quite sometime pick up valuable information not only on one time step ahead but to even further in the future. For those investors who are capable of it, growing number of literature in finance and investments have attracted full attentions of them to establish multi-period optimizations on cutting edge of investment science.

A great deal of related literature to the dynamic investments stand on Merton (1971) to respect for the closed form solutions and the analytical approach to derive them. The solutions look economically intuitive enough to have captured subsequent researchers. As increasing popularity of the solutions to complicated investment problems, scarce achievements end up with successful outcomes.

For example, no short sales is one of well prevailed investment constraints in the practical world. In general, the no short constraint hopelessly lead an original problem to deep water. A couple successful derivations to overcome the difficulties include Li, Zhou and Lim (2002) who reaches out optimal solutions under optimal linear quadratic control. Maximizing the mean-variance utility, Li, Zhou and Lim (2002) obtains continuous solutions as a viscosity solution for the Hamilton-Jacobi-Bellman equation by using two Riccaci equations. Costa and Araujo (2008) studies regime dependency with respect to expected returns and volatility on optimal solutions in the multi-period mean variance utility under VaR constraints. Sotomayor and Cadenillas (2009) solves an optimal policy imposed by bankruptcy constraints over regime dependent asset pricing. Recently, Dombrovskii and Obyedko (2014) formulates problems under the Model Predictable Control (MPC) to minimize deviations from benchmark subject to a borrowing constraint assuming regime dependent key parameters to specify asset pricing.

This chapter is devoted to contemplate a dynamic investment problem where asset prices are regime dependent and investment constraints are imposed. Special focus is on an earlier study by Moallemi and Sağlam (2013) advocating the Linear Rebalancing Rules that apply to wide class of optimal investment problems including Gârleanu and Pedersen (2013). Under a framework from the viewpoint of Linear Quadratic Control (LQC), Gârleanu and Pedersen (2013) achieves a closed-form solution for a dynamic investment on assets that returns are predicted by common factors specified in VAR(1). The analytical solution appears to be a linear combination of current holdings and a functions of the common factors. Chapter 4 in this thesis extends the contribution by Gârleanu and Pedersen (2013) to regime dependent space. In that space, an optimal investment, similar to Gârleanu and Pedersen (2013), appears to be the linear combination proportional to regime probabilities. Consistent findings in these studies plays important roles to understand theoretical nature of the optimal solutions in dynamic investment.

As has been the case in earlier studies, solving optimal investments under Dynamic Programming faces serious obstacles to overcome if popular investment constraints are imposed. For example, the problem is trapped by the curse of dimensionality if no short sales and such a budget constraint as no leverage or self-finance. Moallemi and Sağlam (2013) takes advantage of the identified fact that the wide class of problem is solved for the linear combination of the two terms under no constraint imposed. The Linear Rebalancing Rules regulate a shape of solutions to be comprised of two terms, one as a current solution and another as a functions of the factors. As such the Linear Rebalancing Rules have reasonable basis to reach out solutions that are practically acceptable although not exact. Numerical demonstrations support the approximation.

Aiming at giving desired shapes to portfolios, based upon the academic contribution documented in Chapter 4 in this thesis, this chapter extends the Linear Rebalancing Rules advocated by Moallemi and Sağlam (2013) to space of regime dependent data generation processes on assets, factors and other key parameters in objective functions. In numerical demonstrations to conduct, a device to limit expanding space of paths for longer time horizon works effectively to calculate optimized portfolios without paying outrageously expensive computation costs. This chapter finds a successful way out of difficulty to solve practically realistic optimization problems for dynamic portfolio investment subject to no short sales imposed in the regime dependent space. Successful formulation of problems is a non-trivial undertaking to extend the achievement of Moallemi and Sağlam (2013) to regime dependent space. Our numerical experiments show sufficiently reasonable performance of optimized portfolios for practically sizable number of assets and factors estimated in the previous Chapter 4. We conclude that application of the Linear Rebalancing Rules to the regime dependent space is justifiable by both of our achievement in the previous Chapter 4 and the the results of the numerical experiments in this chapter. By virtue of the problems solved under short sales constraint imposed, the most significant contribution to the investment practice enables much broader spectrum in the vast majority of the investment society than hedge funds community to utilize regime dependent multi-period optimal portfolios to manage.

This chapter proceeds as follows: Section 5.2 formulates all necessary terms that comprise an objective function. Section 5.3 confirms investment efficacy improved by the formulation and discuss how portfolios changes the shapes across choices in key portfolio parameters. Section 5.4 concludes this chapter.

5.2 Dynamic Optimization by Linear Rebalancing Strategy

5.2.1 Setup

Although the model of factor and return processes analyzed in this chapter is the same as that in Chapter 4 except for the objective function, we describe the setup in some detail to make the manuscript self-contained. We consider an economy with N assets traded at time $t = 1, 2, \ldots$. The excess return of asset i to the market return between t and $t+1$ is $r_i(t+1)$. We assume that an $N \times 1$ excess return vector $\boldsymbol{r}(t) = (r_1(t), \ldots, r_N(t))^\top$ ($\top$ denotes transpose) is given by

$$\boldsymbol{r}(t+1) = \boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t) + \boldsymbol{u}_{I(t+1)}(t+1), \qquad \boldsymbol{u}_i(t+1) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{W}_i). \tag{5.1}$$

In (5.1), $I(t)$ is a regime process on $\{1, \ldots, J\}$ introduced to represent discontinuous state changes of the market, the details of which will be explained below. The first term $\boldsymbol{L}_{I(t+1)}\boldsymbol{f}(t)$ denotes the expected excess return known to the investor at time t where $\boldsymbol{f}(t)$ is an $M \times 1$ vector of factors that predict excess returns. $\boldsymbol{L}_{I(t+1)}$ is an $N \times M$ matrix of factor loadings such that $\boldsymbol{L}_{I(t+1)} = \boldsymbol{L}_i$ when the regime $I(t+1) = i$ ($i = 1, \ldots, J$). The second term $\boldsymbol{u}_{I(t+1)}(t+1)$ represents an unpredictable noise. We assume that $\mathrm{E}\left(\boldsymbol{u}_{I(t+1)}(t+1) \,\middle|\, I(t+1) = i\right) = \boldsymbol{0}$ for all i ($\boldsymbol{0}$ denotes a zero vector), whereas the covariance matrix is given by $\boldsymbol{W}_i = \mathrm{V}\left(\boldsymbol{u}_{I(t+1)}(t+1) \,\middle|\, I(t+1) = i\right)$.

The dynamics of the factor is modeled by a first order regime-switching vector autoregressive process

$$\boldsymbol{f}(t+1) = \boldsymbol{\mu}_{I(t+1)} + \boldsymbol{\Phi}_{I(t+1)}\boldsymbol{f}(t) + \boldsymbol{\epsilon}_{I(t+1)}(t+1), \qquad \boldsymbol{\epsilon}_i(t+1) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_i). \tag{5.2}$$

In (5.2), $\boldsymbol{\mu}_{I(t+1)}$ is an $M \times 1$ vector determining the level of mean-reversion and $\boldsymbol{\Phi}_{I(t+1)}$ is an $M \times M$ coefficient matrix. Specifically, if $I(t+1) = i$, they are given as $\boldsymbol{\mu}_{I(t+1)} = \boldsymbol{\mu}_i$ and $\boldsymbol{\Phi}_{I(t+1)} = \boldsymbol{\Phi}_i$, respectively. $\boldsymbol{\epsilon}_{I(t+1)}(t+1)$ is a vector of noise terms affecting the factors. We assume $\mathrm{E}\left(\boldsymbol{\epsilon}_{I(t+1)}(t+1) \,\middle|\, I(t+1) = i\right) = \boldsymbol{0}$ for all i and the covariance matrix $\boldsymbol{\Sigma}_i = \mathrm{V}\left(\boldsymbol{\epsilon}_{I(t+1)}(t+1) \,\middle|\, I(t+1) = i\right)$. We also assume that the factor process $\{\boldsymbol{f}(t)\}$ is stationary in time and that $\mathrm{E}\left(\boldsymbol{\epsilon}_{I(t+1)}(t+1), \boldsymbol{u}_{I(t+1)}(t+1) \,\middle|\, I(t+1)\right) = \boldsymbol{0}$. Conditions for the stationarity of regime-switching vector autoregressive process are given in Francq and Zakoïan (2001).

As is many existing literatures, we assume the regime process $\{I(t)\}$ follows an irreducible Markov chain on $\{1, \ldots, J\}$ with the transition probability matrix

$$\boldsymbol{P} = \begin{bmatrix} p_{1,1} & \cdots & p_{1,J} \\ \vdots & \ddots & \vdots \\ p_{J,1} & \cdots & p_{J,J} \end{bmatrix}, \quad p_{i,j} = \mathrm{P}\left(I(t+1) = j \mid I(t) = i\right).$$

The noise terms $\boldsymbol{u}_{I(t)}(t)$ and $\boldsymbol{\epsilon}_{I(t)}(t)$ are assumed to be conditionally independent in the sense that, given a sample path of the regime process $I(1) = i_i, I(2) = i_2, \ldots, \boldsymbol{u}_{i_1}(1), \boldsymbol{u}_{i_2}(2), \ldots$, and $\boldsymbol{\epsilon}_{i_1}(1), \boldsymbol{\epsilon}_{i_2}(2), \ldots$ are all independent of each other and distribution functions of $\boldsymbol{u}_{i_t}$ and $\boldsymbol{\epsilon}_{i_t}$ are determined by i_t.

At time t, an investor determines amount of investment $x_i(t)$ to asset i. From (5.1), the excess return of the portfolio $\boldsymbol{x}(t) = (x_1(t), \ldots, x_N(t))^\top$ between t and $t+1$ is $y(t+1) = \boldsymbol{x}(t)^\top \{\boldsymbol{L}_{I(t+1)} \boldsymbol{f}(t) + \boldsymbol{u}_{I(t+1)}(t+1)\}$. To construct utility function, we assume that at time t an investor is able to predict one step ahead regime $I(t+1)$ with certainty. We denote the predicted regime by $I_t(t+1)$. A natural and plausible way of prediction is to choose $I_t(t+1)$ as the regime that maximizes one step ahead regime probability. In general, filtered regime probabilities are close to 0 or 1 and the regime process shows strong tendency of self-transition (cf., Subsection 4.4.2), it is not unrealistic for investor to predict $I_t(t+1)$ with certainty.

Given $\boldsymbol{f}(t)$ and $I_t(t+1)$, the conditional mean of the excess return under investor's prediction is

$$\mathrm{E}\left(\boldsymbol{x}(t)^\top \{\boldsymbol{L}_{I(t+1)} \boldsymbol{f}(t) + \boldsymbol{u}_{I(t+1)}(t+1)\} \,\middle|\, \boldsymbol{f}(t), I_t(t+1)\right) = \boldsymbol{x}(t)^\top \boldsymbol{L}_{I_t(t+1)} \boldsymbol{f}(t)$$

since $\mathrm{E}\left(\boldsymbol{u}_{I(t+1)}(t+1) \,\middle|\, I(t+1) = i\right) = \boldsymbol{0}$ for all i. Similarly, the conditional variance is calculated as

$$\mathrm{V}\left(\boldsymbol{x}(t)^\top \{\boldsymbol{L}_{I(t+1)} \boldsymbol{f}(t) + \boldsymbol{u}_{I(t+1)}(t+1)\} \,\middle|\, \boldsymbol{f}(t), I_t(t+1)\right) = \boldsymbol{x}(t)^\top \boldsymbol{W}_{I_t(t+1)} \boldsymbol{x}(t).$$

An investor is risk averse and let $\lambda_i > 0$ denote the coefficient of risk aversion when $I_t(t+1) = i$. Regime-dependent risk aversion coefficient allows us to represent, for example, an investor who chooses larger λ_i when the market is volatile in regime i. We also assume that a quadratic transaction cost $\Delta\boldsymbol{x}(t)^\top \boldsymbol{B}_{I(t+1)} \Delta\boldsymbol{x}(t)/2$ will be incurred for trading $\Delta\boldsymbol{x}(t) = \boldsymbol{x}(t) - \boldsymbol{x}(t-1)$ where $\boldsymbol{B}_i$ is a symmetric positive definite matrix. In sum, an investor attempts to maximize the sum of the mean-variance utilities penalized for transaction costs from $t = 1$ until T:

$$\sum_{t=1}^{T} \rho^{t-1} \mathrm{E}\left(\boldsymbol{x}(t)^\top \boldsymbol{L}_{I_t(t+1)} \boldsymbol{f}(t) - \frac{\lambda_{I_t(t+1)}}{2} \boldsymbol{x}(t)^\top \boldsymbol{W}_{I_t(t+1)} \boldsymbol{x}(t) - \frac{1}{2} \Delta\boldsymbol{x}(t)^\top \boldsymbol{B}_{I_t(t+1)} \Delta\boldsymbol{x}(t)\right) \tag{5.3}$$

where $\rho \in [0, 1]$ is a discount rate. It is noted that $\rho = 1$ represents ordinary finite horizon optimization while $\rho = 0$ corresponds to single-period or myopic optimization.

5.2.2 Linear Rebalancing Strategy

Our aim is to develop a dynamic investment strategy for the objective function (5.3) subject to short sales constraint. Compared with the static optimization, however, the problem is much involved due to the so-called state space explosion. In a dynamic optimization framework, a future investment decision $\boldsymbol{x}(t)$ should be made based on the observations of $I(1), \ldots, I(t)$ and $\boldsymbol{f}(1), \ldots, \boldsymbol{f}(t)$ which are available at t. Since the number of possible sample paths of the regime process is J^{t-1} and the state space of the factor process $\boldsymbol{f}(t)$ is $\Re^{M\times t}$, the state space to be investigated in the optimization grows exponentially as the time horizon T increases.

To avoid the difficulty, Moallemi and Sağlam (2013) proposed a linear rebalancing rule for dynamic portfolio optimization. In the context of the factor portfolio described in Subsection 5.2.1, the idea is based on Gârleanu and Pedersen (2013) where they have proved that, for the infinite horizon problem, the optimal investment is given by

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{x}(t-1) + \boldsymbol{B}\boldsymbol{f}(t) + \boldsymbol{b} \tag{5.4}$$

for some matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ and a vector $\boldsymbol{b}$ determined from the model parameters. Starting with initial portfolio $\boldsymbol{x}(0)$, iterative substitutions into (5.4) show that $\boldsymbol{x}(t)$ can be expressed as

$$\boldsymbol{x}(t) = \sum_{s=1}^{t} \boldsymbol{A}^{t-s}\boldsymbol{B}\boldsymbol{f}(s) + \boldsymbol{A}^{t}\boldsymbol{x}(0) + (\boldsymbol{A}^{t-1} + \cdots + \boldsymbol{A} + \boldsymbol{I})\boldsymbol{b}. \tag{5.5}$$

Since (5.5) indicates that the optimal investment policy for infinite horizon problem is a linear combination of realized factors, Moallemi and Sağlam (2013) suggests a dynamic investment strategy of the form $\boldsymbol{x}(t) = \sum_{s=1}^{t} \boldsymbol{C}_t(s)\boldsymbol{f}(s) + \boldsymbol{c}(t)$ where $\boldsymbol{C}_t(s)$ and $\boldsymbol{c}(t)$ are determined so as to optimize given objective function for a finite horizon problem.

Under the regime switching circumstance, Komatsu and Makimoto (2015) extends Gârleanu and Pedersen (2013) in such a way that the optimal investment is shown to be

$$\boldsymbol{x}(t) = \boldsymbol{A}_i\boldsymbol{x}(t-1) + \boldsymbol{B}_i\boldsymbol{f}(t) + \boldsymbol{b}_i \tag{5.6}$$

where $I_t(t+1) = i$. Since (5.4) is a natural extension of (5.6), we extend the linear rebalancing strategy to the case with multiple regimes in the following way:

- For all $t = 1, \ldots, T$ and all possible sample paths of the regime process $I[t] = \{I(1), \ldots, I(t)\}$, we prescribe coefficients $\boldsymbol{C}_{I[t]}(1), \ldots, \boldsymbol{C}_{I[t]}(t)$ where $\boldsymbol{C}_{I[t]}(1)$ is an $N \times 1$ vector while $\boldsymbol{C}_{I[t]}(s)$ $(s = 2, \ldots, t)$ is an $N \times M$ matrix.

- The investment decision at time t is given by

$$\boldsymbol{x}(t) = \begin{cases} \boldsymbol{C}_{I[1]}(1), & t = 1 \\ \boldsymbol{C}_{I[t]}(1) + \sum_{s=2}^{t} \boldsymbol{C}_{I[t]}(s)\boldsymbol{f}(s), & t = 2, \ldots, T. \end{cases} \tag{5.7}$$

The coefficient matrices $\boldsymbol{C}_{I[t]}$'s are determined so as to maximize the objective function (5.3). Note that the strategy given by (5.7) is dynamic in the sense that $\boldsymbol{x}(t)$ is based on the realizations of $I[t]$ and $\boldsymbol{f}[t] = \{\boldsymbol{f}(1), \ldots, \boldsymbol{f}(t)\}$ that are uncertain at $t = 1$. Only $\boldsymbol{x}(1)$ is given deterministically since $\boldsymbol{x}(0)$, $\boldsymbol{f}(1)$ and $I(1)$ are observable at $t = 1$.

5.2.3 Formulation of the Optimization Problem

In this subsection, we give an explicit expression of the optimization problem to compute $\boldsymbol{C}_{I[t]}(s)$ in (5.7) which maximizes (5.3). By introducing regime switches, it becomes much involved to express the objective function (5.3) explicitly in terms of $\boldsymbol{C}_{I[t]}$'s compared with Moallemi and Sağlam (2013) for linear rebalancing strategy without regime switches. In addition to future uncertainty of the factor process $\boldsymbol{f}(t)$, we also need to take regime switches into account which makes computation of the expectation much difficult. To keep the clarity of presentation, proofs of Proposition 5.2.1 below and some preliminary lemmas are given in Section 5.5. Some usefult equalities that play an important role in numerical optimization are also provided there.

By defining

$$\boldsymbol{C}_{I[t]} = \left[\boldsymbol{C}_{I[t]}(1), \boldsymbol{C}_{I[t]}(2), \ldots, \boldsymbol{C}_{I[t]}(t)\right] = \begin{bmatrix} \boldsymbol{c}_{I[t],1}^{\top} \\ \boldsymbol{c}_{I[t],2}^{\top} \\ \vdots \\ \boldsymbol{c}_{I[t],t}^{\top} \end{bmatrix} \tag{5.8}$$

where $\boldsymbol{c}_{I[t],i}^{\top}$ denotes ith row vector of $\boldsymbol{C}_{I[t]}$, and

$$\boldsymbol{F}[t] = \begin{bmatrix} 1 \\ \boldsymbol{f}(2) \\ \vdots \\ \boldsymbol{f}(t) \end{bmatrix}, \quad t = 2, \ldots, T, \tag{5.9}$$

we can express

$$\boldsymbol{x}(t) = \boldsymbol{C}_{I[t]}\boldsymbol{F}[t], \quad t = 2, \ldots, T. \tag{5.10}$$

The objective function (5.3) contains three types of components, i.e., the expected return, risk penalty and transaction costs. From (5.10), these components can be explicitly expressed in terms of the optimization parameters $\boldsymbol{C}_{I[t]}$ in the following way. For a realization of the regime process $I[t] = \{I(1), \ldots, I(t)\}$ up to time t, $p(I[t])$ and $\mathrm{E}_{I[t]}\left(\ \right)$ respectively denote the probability that $I[t]$ occurs and the expectation conditioned on $I[t]$. Also, $\sum_{I[t]}$ means a summation taken over all J^{t-1} types of possible $I[t]$'s. For an $a \times b$ matrix $\boldsymbol{M} = [m_{i,j}]$, we define $ab \times 1$ vector by

$$\mathrm{vec}(\boldsymbol{M}) = (m_{1,1}, \ldots, m_{a,1}, \ldots, m_{1,b}, \ldots, m_{a,b})^{\top}.$$

Finally, $\otimes$ denotes Kronecker product.

Proposition 5.2.1 (a) The expected return is expressed as

$$\begin{aligned} &\mathrm{E}\left(\boldsymbol{x}(t)^{\top}\boldsymbol{L}_{I_t(t+1)}\boldsymbol{f}(t)\right) \\ &\quad = \begin{cases} \boldsymbol{x}(1)^{\top}\boldsymbol{L}_{I_1(2)}\boldsymbol{f}(1) = \boldsymbol{C}_{I[1]}^{\top}\boldsymbol{L}_{I_1(2)}\boldsymbol{f}(1), & t = 1 \\ \displaystyle\sum_{I[t]} p(I[t])\mathrm{vec}(\boldsymbol{C}_{I[t]}^{\top})^{\top}\left\{\boldsymbol{I}_N \otimes \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\boldsymbol{f}(t)^{\top}\right)\right\}\mathrm{vec}(\boldsymbol{L}_{I_t(t+1)}^{\top}), & t = 2, \ldots, T \end{cases} \end{aligned} \tag{5.11}$$

where $\boldsymbol{I}_N$ is an N dimensional identity matrix.
(b) The risk penalty is given by

$$\mathrm{E}\left(\frac{\lambda_{I_t(t+1)}}{2}\boldsymbol{x}(t)^\top \boldsymbol{W}_{I_t(t+1)}\boldsymbol{x}(t)\right)$$
$$= \begin{cases} \frac{\lambda_{I_1(2)}}{2}\boldsymbol{C}_{I[1]}^\top \boldsymbol{W}_{I_1(2)}\boldsymbol{C}_{I[1]}, & t=1 \\ \sum_{I[t]} p(I[t])\frac{\lambda_{I_t(t+1)}}{2}\mathrm{vec}(\boldsymbol{C}_{I[t]}^\top)^\top \left\{\boldsymbol{W}_{I_t(t+1)} \otimes \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\boldsymbol{F}(t)^\top\right)\right\}\mathrm{vec}(\boldsymbol{C}_{I[t]}^\top), & t=2,\ldots,T. \end{cases} \tag{5.12}$$

(c) Let $\Delta \boldsymbol{C}_{I[t]} = \boldsymbol{C}_{I[t]} - [\boldsymbol{C}_{I[t-1]}, \boldsymbol{O}_{N,M}]$ where $\boldsymbol{O}_{N,M}$ is a zero matrix of size $N \times M$. The transaction cost is

$$\Delta\boldsymbol{x}(t)^\top \boldsymbol{B}_{I_t(t+1)}\Delta\boldsymbol{x}(t)$$
$$= \begin{cases} \{\boldsymbol{C}_{I[1]} - \boldsymbol{x}(0)\}^\top \boldsymbol{B}_{I_1(2)}\{\boldsymbol{C}_{I[1]} - \boldsymbol{x}(0)\}, & t=1 \\ \sum_{I[t]} p(I[t])\mathrm{vec}(\Delta\boldsymbol{C}_{I[t]}^\top)^\top \left\{\boldsymbol{B}_{I_t(t+1)} \otimes \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\boldsymbol{F}[t]^\top\right)\right\}\mathrm{vec}(\Delta\boldsymbol{C}_{I[t]}^\top), & t=2,\ldots,T. \end{cases} \tag{5.13}$$

Remark 5.2.2 The summand in (5.13) is expanded as

$$\begin{aligned} &\mathrm{vec}(\Delta\boldsymbol{C}_{I[t]}^\top)^\top \left\{\boldsymbol{B}_{I_t(t+1)} \otimes \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\boldsymbol{F}[t]^\top\right)\right\}\mathrm{vec}(\Delta\boldsymbol{C}_{I[t]}^\top) \\ &= \mathrm{vec}(\boldsymbol{C}_{I[t]}^\top)^\top \left\{\boldsymbol{B}_{I_t(t+1)} \otimes \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\boldsymbol{F}[t]^\top\right)\right\}\mathrm{vec}(\boldsymbol{C}_{I[t]}^\top) \\ &\quad +\mathrm{vec}(\boldsymbol{C}_{I[t-1]}^\top)^\top \left\{\boldsymbol{B}_{I_t(t+1)} \otimes \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t-1]\boldsymbol{F}[t-1]^\top\right)\right\}\mathrm{vec}(\boldsymbol{C}_{I[t-1]}^\top) \\ &\quad -\mathrm{vec}(\boldsymbol{C}_{I[t-1]}^\top)^\top \left\{\boldsymbol{B}_{I_t(t+1)} \otimes \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t-1]\boldsymbol{F}[t]^\top\right)\right\}\mathrm{vec}(\boldsymbol{C}_{I[t]}^\top) \\ &\quad -\mathrm{vec}(\boldsymbol{C}_{I[t]}^\top)^\top \left\{\boldsymbol{B}_{I_t(t+1)} \otimes \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\boldsymbol{F}[t-1]^\top\right)\right\}\mathrm{vec}(\boldsymbol{C}_{I[t-1]}^\top) \end{aligned} \tag{5.14}$$

which will be useful in implementing a numerical algorithm.

The conditional expectations appeared in (5.11)∼(5.13) can be represented in terms of the model parameters as in (5.28), (5.29) and (5.33) in Subsection 5.5.1. Hence, the objective function (5.3) is expressed as a sum of linear functions and quadratic forms of $\mathrm{vec}(\boldsymbol{C}_{I[t]}(s))^\top$s. Since $\boldsymbol{C}_{I[t]}(1)$ is $N \times 1$ and $\boldsymbol{C}_{I[t]}(s)$ $(s = 2,\ldots,t)$ is $N \times M$, and there are J^{t-1} possibilities for $I[t]$, the total number of variables in the optimization is

$$\sum_{t=1}^{T}(1 + (t-1)M)NJ^{t-1}. \tag{5.15}$$

In the remaining, we consider short sales constraint which is frequently imposed in practical investment. Under the linear rebalancing strategy (5.10), $\boldsymbol{x}(1)$ is deterministic while $\boldsymbol{x}(t)$ for $t = 2,\ldots,T$ is stochastic in that it depends on $I[t]$ and $\boldsymbol{f}[t]$ which are uncertain at the initial time $t = 1$. Thus, a short sales constraint for $t = 1$ is given as

$$\boldsymbol{x}(1) = \boldsymbol{C}_{I[1]} \geq \boldsymbol{0}. \tag{5.16}$$

For $t = 2, \ldots, T$, however, it is impossible for the linear rebalancing strategy (5.10) to satisfy short sales constraint with certainty since $\boldsymbol{f}(s)$'s in (5.10) are distributed on $\Re^M$. We therefore impose stochastic short sales constraint

$$\mathrm{P}_{I[t]}\left(x_i(t) < 0\right) \le p \tag{5.17}$$

for all $t = 2, \ldots, T$, $i = 1, \ldots, N$ and all realizations of $I[t]$. In (5.17), $\mathrm{P}_{I[t]}(\ )$ is a conditional probability given $I[t]$ and $x_i(t)$ is ith element of $\boldsymbol{x}(t)$. p gives an upper bound of the probability that the short sales constraint is violated and is usually set to small number such as $p = 0.05$.

By the assumption that $\boldsymbol{\epsilon}(t)$ in (5.2) follows multivariate normal distribution with regime-dependent covariance matrix, the constraint (5.17) can be represented as

$$\boldsymbol{c}_{I[t],i}^{\top}\mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\right) \ge \Phi^{-1}(1-p)\|\boldsymbol{\Theta}_{I[t]}\boldsymbol{c}_{I[t],i}\|_2 \tag{5.18}$$

where $\Phi^{-1}(\ )$ is an inverse cumulative standard normal distribution function and $\|\ \|_2$ is Euclidean norm. The derivation of (5.18) and the explicit expressions of $\mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\right)$ and $\boldsymbol{\Theta}_{I[t]}$ are given in Subsection 5.5.1. In sum, a dynamic portfolio optimization problem for linear rebalancing strategy under regime switches is formulated as a quadratic programming subject to second order cone constraints.

5.3 Numerical Experiments

In this section, we conduct numerical experiments to check investment efficacy of the linear rebalancing strategy proposed in Section 5.2 combined with model predictive control (MPC). Performance comparison with the infinite horizon optimization in Chapter 4 and the perfect foresight, which assumes an idealized situation where an investor can foresee future transitions of both regime and factor processes, are also provided. In the computing resources with the Intel(R) Core(TM) i7-4960X CPU 3.60GHz 6 Cores 12 Threads under 64bit operating system with 64G byte memory, we implement the optimization algorithm by MATLAB and CVX (Grant and Boyd (2015)) which is able to efficiently solve the quadratic programming subject to second order cone constraints.

5.3.1 Linear Rebalancing Strategy Combined with MPC

The estimation of the investment performance of the linear rebalancing strategy is executed in the following way.

(1) Initial condition

An initial portfolio is set to be empty as $\boldsymbol{x}(0) = (0, \ldots, 0)^{\top}$ while an initial regime is chosen according to the stationary distribution with respect to the transition probability matrix $\boldsymbol{P}$ of the regime process. Similarly, an initial value of factors $\widehat{\boldsymbol{f}}(1)$ is given as its time stationary mean $\widehat{\boldsymbol{f}}(1) = (\boldsymbol{I}_K - \boldsymbol{\Phi}_{\widehat{I}(1)})^{-1}\boldsymbol{\mu}_{\widehat{I}(1)}$ on the selected regime $\widehat{I}(1)$.

(2) Generation of $\{\widehat{I}(t)\}$, $\{\widehat{\boldsymbol{f}}(t)\}$ and $\{\widehat{\boldsymbol{r}}(t)\}$

Following the initial regime $\widehat{I}(1)$, a sample path of the subsequent regime process $\{\widehat{I}(t)\}$ is generated based upon the transition probability matrix $\boldsymbol{P}$. Given $\{\widehat{I}(t)\}$, a sample path of the factor process $\{\widehat{\boldsymbol{f}}(t)\}$ is generated according to (5.2) where $\{\widehat{\boldsymbol{\epsilon}}(t)\}$ is sampled from $N(\boldsymbol{0}, \boldsymbol{\Sigma}_{\widehat{I}(t)})$. A sample path of the return process $\{\widehat{\boldsymbol{r}}(t)\}$ is generated from (5.1) with the sample $\widehat{I}(t)$ and $\widehat{\boldsymbol{f}}(t)$ where $\boldsymbol{u}(t+1)$ is sampled from $N(\boldsymbol{0}, \boldsymbol{W}_{\widehat{I}(t)})$.

In all experiments explained below, the length of the sample path is set to $T_{sim} = 2,500$ which is determined from preliminary experiments so that the performance measures such as the Sharpe ratio seem to converge enough over time.

(3) Computation of linear rebalancing strategy

At each time t, an investor observes the sample paths $\widehat{I}(t)$ and $\widehat{\boldsymbol{f}}(t)$, and compute the optimal linear rebalancing strategy (5.7) that maximizes (5.3) subject to short sales constraints (5.18). Note that obtained strategy covers from current investment decision $\widehat{\boldsymbol{x}}(t)$ up to $\widehat{\boldsymbol{x}}(t+T_{opt})$, a decision T_{opt} steps ahead. The investment horizon T_{opt} ranges from 1 to maximum 10 for comparison.

(4) Model predictive control

In a multi-period investment with finite horizon, the obtained strategy is not necessarily fully utilized. It is often the case that, among investment decisions $\boldsymbol{x}(t), \ldots, \boldsymbol{x}(t+T_{opt})$ computed at t, an investor only uses first τ $(< T)$ decisions $\widehat{\boldsymbol{x}}(t), \ldots, \widehat{\boldsymbol{x}}(t+\tau-1)$ and ignore those in the remaining period. Instead, he or she conducts next optimization at $t+\tau$ based on newly obtained observations of regime, factor and return processes. This type of optimization is sometimes called Model Predictive Control (MPC). In the following experiments, we set from $\tau = 1$ to maximum T_{opt} for comparison.

(5) Performance measures

For measuring investment efficacy, we employ realized net utility as well as net Sharpe ratio defined below:

$$\widehat{U} = \widehat{\mu} - \frac{\lambda}{2}\widehat{\sigma}^2 - \widehat{TC} \tag{5.19}$$

$$\widehat{SR} = \frac{\widehat{\mu} - \widehat{TC}}{\widehat{\sigma}} \tag{5.20}$$

where

$$\widehat{\mu} = \frac{1}{T_{sim}} \sum_{s=1}^{T_{sim}} \widehat{\boldsymbol{x}}(s)^\top \widehat{\boldsymbol{r}}(s+1) \tag{5.21}$$

$$\widehat{\sigma} = \sqrt{\frac{1}{T_{sim}-1} \sum_{s=1}^{T_{sim}} \{\widehat{\boldsymbol{x}}(s)^\top \widehat{\boldsymbol{r}}(s+1) - \widehat{\mu}\}^2} \tag{5.22}$$

$$\widehat{TC} = \frac{1}{2T_{sim}} \sum_{s=1}^{T_{sim}} \{\widehat{\boldsymbol{x}}(s) - \widehat{\boldsymbol{x}}(s-1)\}^\top \boldsymbol{B}_{\widehat{I}(s)} \{\widehat{\boldsymbol{x}}(s) - \widehat{\boldsymbol{x}}(s-1)\}. \tag{5.23}$$

Before closing this subsection, we give two remarks on computational issues.

Remark 5.3.1 (a) As mentioned in Section 5.2, the number of variables in the optimization increases exponentially fast as T_{opt} increases, which makes our experiments time-wasting even for moderate values of T_{opt}. On the other hand, the probability $p(I[t])$ becomes very small for most of regime paths $I[t]$ since transition probabilities from one regime to others are significantly smaller than that of self-transition as we have already seen in Chapters 3 and 4. This implies that the optimization problem would not be affected very much even when we ignore regime transitions with small probabilities. For moderate to large values of T_{opt} in our experiments, we therefore restrict the state space of $I[t]$ to those with at most one or two regime switches during $[1, T_{opt}]$. The number of paths to deploy is reduced from J^{t-1} to t and $t(t-1)/2+1$ for the limit to one time and to two times, respectively.
(b) To improve the accuracy of the performance measures, we employ antithetic variate in generating $\{\widehat{\boldsymbol{f}}(t)\}$ and $\{\widehat{\boldsymbol{r}}(t)\}$ and computing $\widehat{U}$ and $\widehat{SR}$. The detailed are described in Subsection 5.5.2.

5.3.2 Model and Parameters

For our numerical experiments, we use the same model comprised of two factors and six assets and the same parameter examined in Section 4.4 for infinite horizon optimization. This makes it possible for us to compare the performance of linear rebalancing strategy with that of linear quadratic control proposed in Chapter 4, though they are not exactly comparable since objective functions do not coincide with each other. For completeness, we summarizes the list of parameters in Tables 5.1 and 5.2, and (5.24).

$$\boldsymbol{P} = \begin{bmatrix} .944 & .056 \\ .148 & .852 \end{bmatrix}. \tag{5.24}$$

The investment horizon ranges from $T_{opt} = 1$ to 10. To avoid the state space explosion explained in Remark 5.3.1(a), the number regime switches is limited at most once for $T_{opt} = 6, \ldots, 10$, and at most twice for $T_{opt} = 6, \ldots, 8$. The discount rate is set to $\rho = 1$ for all cases. The risk aversion coefficient varies in $\lambda \in \{0.5, 1, 5\}$. The upper limit of the probability to violate the short sales constraint is $p = 0.05$. The model predictive control is examined by choosing the number of steps τ to utilize the optimization decisions from 1 up to T_{opt}. The transaction cost coefficient matrices are set to $\boldsymbol{B}_1 = 0.01\boldsymbol{I}$ and $\boldsymbol{B}_2 = 0.05\boldsymbol{I}$.

5.3.3 Comparison of Investment Performances

Before a deep dive into details of experiment results, we show an overview in Table 5.3 where our numerical experiments explore to figure out our conclusions. Figure 5.1 and Figure 5.2 show the Linear Rebalancing Strategy (LRS) performing intuitively in terms both of length of optimization horizons and relative to upper bounds and a lower bound. Figure 5.3, Figure 5.4 and Figure 5.5 show attributions of breakdown items in a realized utility to the over all utility and compare those of LRS with Perfect Foresight (PF). Plots in Figure 5.1 to Figure 5.5 show mean values over all trials. Finally, Figure 5.6 to Figure 5.9

Table 5.1: Estimated parameters of $\boldsymbol{L}$ and $\boldsymbol{W}$ in (5.1) for assets

Single regime model								
	$\boldsymbol{L}$		$\boldsymbol{W}(\times 10^{-3})$					
	SMB	HML	SG	SN	SV	BG	BN	BV
SG	.100	.078	.179	.835	.691	.429	.380	.213
SN	.060	.101	.119	.113	.836	.264	.503	.417
SV	.176	.193	.123	.118	.176	.103	.506	.591
BG	−.020	.005	.037	.018	.009	.042	.341	−.027
BN	.026	−.011	.031	.033	.041	.014	.037	.598
BV	.101	−.064	.031	.049	.086	−.002	.040	.120
Two regime model								
Regime 1	$\boldsymbol{L}_1$		$\boldsymbol{W}_1(\times 10^{-3})$					
	SMB	HML	SG	SN	SV	BG	BN	BV
SG	.119	−.032	.108	.825	.747	.482	.241	.142
SN	.077	.088	.067	.061	.832	.351	.374	.266
SV	.161	.207	.070	.058	.081	.246	.366	.400
BG	−.045	−.020	.024	.013	.011	.022	.341	.019
BN	−.007	.032	.011	.012	.014	.007	.018	.555
BV	.019	.105	.009	.013	.023	.001	.015	.040
Regime 2	$\boldsymbol{L}_2$		$\boldsymbol{W}_2(\times 10^{-3})$					
	SMB	HML	SG	SN	SV	BG	BN	BV
SG	.061	.126	.365	.847	.668	.396	.489	.283
SN	.035	.109	.257	.252	.842	.211	.589	.511
SV	.198	.186	.265	.278	.432	.023	.580	.681
BG	.013	.012	.074	.033	.005	.095	.342	−.053
BN	.076	−.032	.088	.088	.113	.031	.088	.613
BV	.236	−.143	.097	.145	.254	−.009	.103	.322

Estimated parameters of $\boldsymbol{L}$ and $\boldsymbol{W}$ in (5.1) for 6 assets, SG (Small Growth), SN (Small Neutral), SV (Small Value), BG (Big Growth), BN (Big Neutral) and BV (Big Value). The first 500 weekly data are used for estimation. Diagonal and lower triangular elements of $\boldsymbol{W}$ $(\times 10^{-3})$ are variance and covariance, respectively. Elements of $\boldsymbol{W}$ in the upper triangle with underline denote correlations.

Table 5.2: Estimated parameters of $\boldsymbol{\mu}, \boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}$ in (5.2) for factors

	$\boldsymbol{\mu}(\times 10^{-3})$	$\boldsymbol{\Phi}$		$\boldsymbol{\Sigma}(\times 10^{-3})$	
Single regime model					
		SMB	HML	SMB	HML
SMB	.713	−.091	.079	.140	<u>.029</u>
HML	.189	.124	−.049	.005	.172
Two regime model					
Regime 1	$\boldsymbol{\mu}_1(\times 10^{-3})$	$\boldsymbol{\Phi}_1$		$\boldsymbol{\Sigma}_1(\times 10^{-3})$	
		SMB	HML	SMB	HML
SMB	.654	−.083	−.018	.102	<u>−.062</u>
HML	.317	.068	.130	−.004	.049
Regime 2	$\boldsymbol{\mu}_2(\times 10^{-3})$	$\boldsymbol{\Phi}_2$		$\boldsymbol{\Sigma}_2(\times 10^{-3})$	
		SMB	HML	SMB	HML
SMB	1.029	−.114	.121	.240	<u>.098</u>
HML	−.348	.224	−0.130	.034	.489

Estimated parameters of $\boldsymbol{\mu}$, $\boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}$ in (5.2) for 2 factors, SMB (Small minus Big) and HML (High minus Low). The first 500 weekly data are used for estimation. Diagonal and lower triangular elements of $\boldsymbol{\Sigma}$ ($\times 10^{-3}$) are variance and covariance, respectively. Elements of $\boldsymbol{\Sigma}$ in the upper triangle with underline denote correlations.

display distribution of $\boldsymbol{x}^*$ over all trials across various length of horizons. See Subsection 5.5.2 for total number of all the trials.

Table 5.3: Overview of numerical experiments

Types of Optimizations	Parameter	Briefing Summary
Linear Rebalancing Strategy (LRS) Linear Quadratic Control (LQC) as upper bounds for LRS Zero Projection(ZP) as lower bounds for LRS Perfect Foresight as upper bounds for LRS	Risk aversion $\lambda \in \{0.5, 1, 5\}$ Figure 5.1: Realized Utility Figure 5.2: Realized Net Sharpe Ratio	LRS performs better in longer T_{ops} than in short T_{ops}. LRS performs worse than upper bounds (PF and LQC) and worse better than lower bounds (LQC ZP). MPC ($\tau > 1$)in LRS performs worse than a pure LRS ($\tau = 1$).
Types of Optimizations	Parameter	Briefing Summary
Linear Rebalancing Strategy (LRS) Perfect Foresight(PF)	Breakdown of realized utility and realized volatility Figure 5.3: $\lambda = 0.5$ Figure 5.4: $\lambda = 1$ Figure 5.5: $\lambda = 5$	The longer T_{ops}, the more LRS reduces risk penalty and the more PF increases gross return.
Types of Optimizations	Parameter	Briefing Summary
Linear Rebalancing Strategy (LRS) Perfect Foresight(PF) Linear Quadratic Control (LQC)	Distribution of $\boldsymbol{x}^*$ across horizons Figure 5.6: Small Growth(SG) Figure 5.7: Small Neutral(SN) Figure 5.8: Small Value(SV) Figure 5.9: Small Growth(BG) Figure 5.10: Small Neutral(BN) Figure 5.11: Small Value(BV)	The longer T_{ops} the smaller positions LRS takes and the larger positions PF takes.

Figure 5.1 and Figure 5.2 respectively tabulate realized utility $\widehat{U}$ and Sharpe ratio $\widehat{SR}$ for the risk aversion coefficient $\lambda = 0.5$ at the top, $\lambda = 1.0$ at the middle and $\lambda = 5.0$ at the bottom. The left three panels show the results obtained by the optimal linear rebalancing strategy (LRS for short) discussed in Section 5.2. As a comparison, we also evaluate performances achieved by the strategy that optimizes the same objective function (5.3) as LRS by assuming that an investor can foresee $\{\widehat{I}(t), \ldots, \widehat{I}(t + T_{opt})\}$ and $\{\widehat{\boldsymbol{f}}(t), \ldots, \widehat{\boldsymbol{f}}(t + T_{opt})\}$ at the time of optimization. This strategy is referred to Perfect Foresight and is shown in three right panels. Though Perfect Foresight is infeasible for $T_{opt} \geq 2$ since it utilizes unknown future informations, we expect that it provides upper bound of performances. The solid horizontal line in each panel depicts a performance of the infinite horizon optimization $\boldsymbol{x}^*(t)$ given by (4.15) in Chapter 4 which we call LQC (Linear Quadratic Control) for short. The discount rate is set to $\rho \in \{0.3.0.5, 0.9\}$. Since $\boldsymbol{x}^*(t)$ is an optimal solution without short sales constraint, we also expect the solid line to be an upper bound of the performance, though it is not exactly comparable with LRS due to the difference of objective functions. The dashed horizontal line shows a performance of investment decision $\boldsymbol{x}(t) = \max(\boldsymbol{x}^*(t), \boldsymbol{0})$ obtained by projecting short positions in $\boldsymbol{x}^*(t)$ as above to 0. We call this strategy LQC zero projection(ZP). Since $\boldsymbol{x}(t)$ satisfies short sales constraint and is easy to implement, the dashed line provides the lower bound or benchmark of the performance.

In both Figure 5.1 and Figure 5.2, performances of LRS improve as investment horizon T_{opt} in x-axis increases. Especially, $\widehat{SR}$ in Figure 5.2 continues to improve sharply even for $T_{opt} = 10$ while $\widehat{U}$ in Figure 5.1 seems to saturate around $T_{opt} = 5 \sim 10$. This suggests the possibility of further improvement of $\widehat{SR}$ by extending T_{opt} though the objective function (5.3) to maximize does not take a form of Sharpe ratio. It is highly valuable to see if the performances of LRS lie between the upper and lower bounds obtained by LQC and LQC zero projection(ZP) as expected. Comparing with two strategies satisfying short sales constraint, for longer time horizons, e.g., T_{opt} closed to 10, we closely take a look if LRS performs worse/better than LQC/LQC zero projection(ZP) with a larger discount rate, e.g., $\rho = 0.9$. On the other hand for shorter horizons, e.g., T_{opt} closed to 1, we pay attention to LRS performance to compare with LQC/LQC zero projection(ZP) under a larger discount rate, e.g., $\rho = 0.3$. By and large, we confirm that is the case across all results shown in both Figure 5.1 and Figure 5.2. This indicates a reasonable basis to believe that LRS provides decent dynamic investment strategy under short sales constraint.

In regards to the impact of τ, the length of interval between re-optimizations, $\tau = 1$ appears to be the best, and performances gradually deteriorate as τ increases. From the perspective of MPC, smaller τ is advantageous in that newly observed informations are taken into consideration more frequently in optimization. On the other hand, frequent re-optimization may lead to sudden change of investment decision due to volatile inputs which could result in large transaction cost. In general, optimal τ is determined by the trade-off between these two effects. As we have seen in Section 4.4, characteristics of asset returns in Regime 1 and Regime 2 are quite different and the regime process does not switch to another state so often. This means that information on the current regime plays an important role to construct appropriate portfolio. The superior results of the strategy with $\tau = 1$ is thus attributed to that it detects regime shifts earlier than strategies with larger τ and rebalances the portfolio so as to fit the current regime.

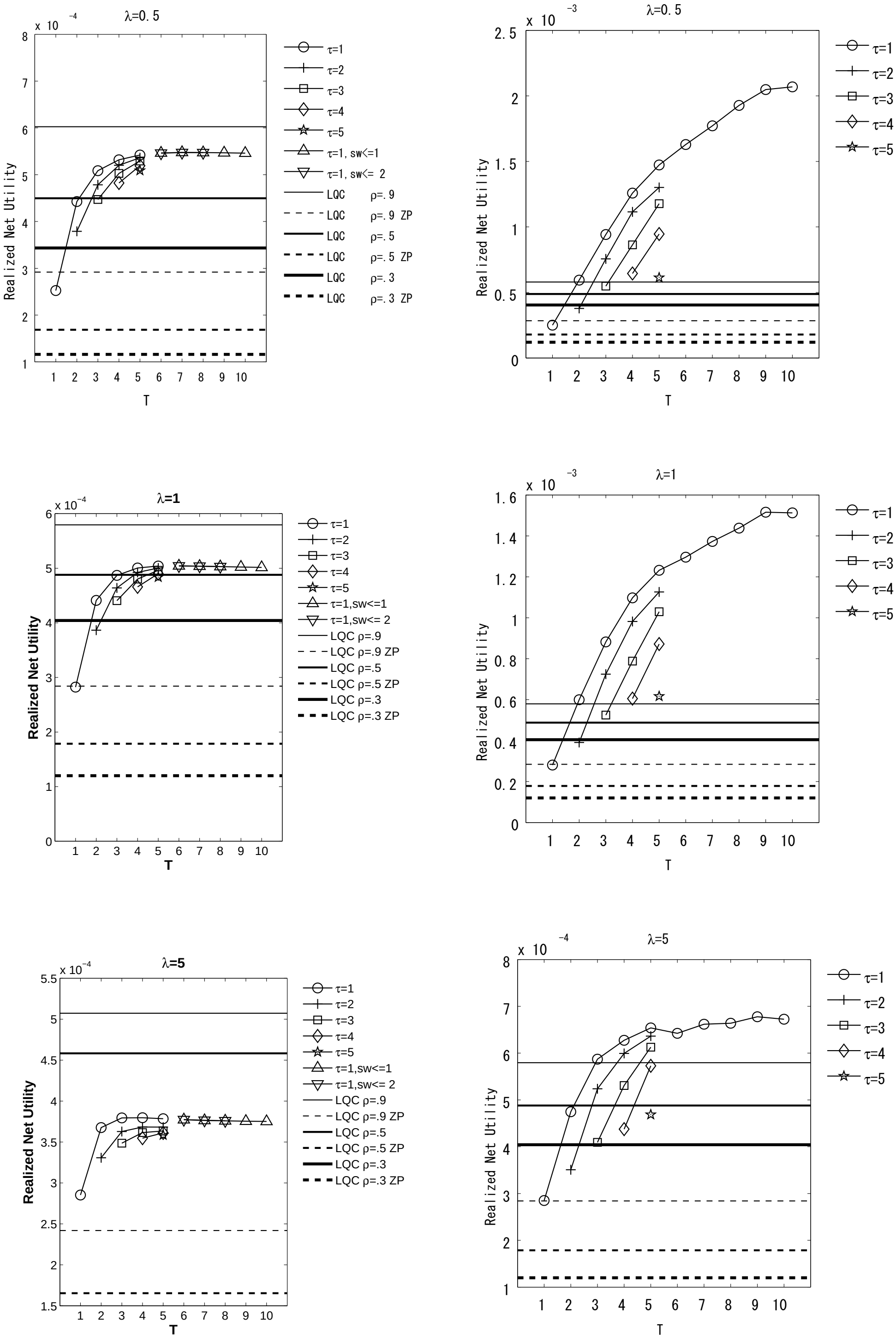
Figure 5.3, Figure 5.4 and Figure 5.5 show attribution of all components that comprise the objective function for $\lambda = 0.5$, $\lambda = 1.0$ and $\lambda = 5.0$, respectively. From the top to third panels, $\widehat{\mu}$, $\widehat{TC}$ and $-\frac{\lambda}{2}\widehat{\sigma}^2$ are displayed and the bottom panel shows $\widehat{\sigma}$ as a denominator in $\widehat{SR}$. The results of LRS and Perfect Foresight appear in the left and right columns, respectively. It is interesting to note that, as T_{opt} increases, LRS and Perfect Foresight(PF) show opposite directions of increase and decrease of all components. For example, $\widehat{\mu}$ for LRS decreases as T_{opt} increases while it increases for Perfect Foresight(PF). Since we assume risk averse investor, he or she becomes more conservative to avoid larger uncertainty for longer T_{opt}. This leads to decreases of $\widehat{\mu}$ and $\widehat{\sigma}$ for LRS as T_{opt} increases. On the other hand, Perfect Foresight(PF) is supposed to know all future informations and does not face to any uncertainty irrespective of T_{opt}. Thus, Perfect Foresight(PF) gains larger expected returns for longer T_{opt} in exchange for larger risk.

5.3.4 Discussions

Figure 5.6∼Figure 5.11 display box plots for each of six assets, i.e., SG, SN, SV, BG, BN and BV. The top panel corresponds to LRS for $T_{opt} = 10$ where possibilities of more than 1 regime switch are ignored in optimization as explained in Subsection 5.3.2. Box plots for Perfect Foresight for $T_{opt} = 10$ and LQC for $\rho \in \{0.3, 0.5, 0.9\}$ are shown in the middle and bottom panels, respectively. For LRS and Perfect Foresight, the x-axis is a time elapsed from the time of optimization and y-axis shows the distribution of amount of investment to each asset calculated from numerical simulation. For LQC, from left to right, each column displays those with discount rates over $\rho \in \{0.3, 0.5, 0.9\}$. Among 2,500,000 trials detailed in Subection 5.5.2, the left and right column respectively shows the distribution conditional on $I(t) = 1$ and $I(t) = 2$ at time t of optimization.

All assets in LRS experience optimal allocation getting close to zero as T_{opt} increases. The larger T_{opt} enables LRS to increase the realized investment utility because $-\frac{\lambda}{2}\widehat{\sigma}^2$ mitigates enough to absorb deterioration of $\widehat{\mu}$. A potential reasons behind is that asset allocation across six assets changes in order to improve sum of three components in the objective function. In particular the asset SV exhibits it clearly in both of $I(t) = 1$ and $I(t) = 2$.

In optimal solutions for Perfect Foresight, the asset SV increases asset allocation if T_{opt} extends when $I(t) = 1$ and decreases when $I(t) = 2$. As we have seen in Table 4.5, SV is allocated the best among all assets if $I(1) = I(2) = \cdots = I(T) = 1$ in the time stationary condition. Provided that no uncertainty in terms of predicted return at all, the increasing allocation to asset SV is reasonable with perfect information throughout longer horizon T_{opt}. By the same reason, the asset SV loses asset allocation if $I(1) = I(2) = \cdots = I(T) = 2$. Although not exactly comparable, levels of discount rate ρ functions similarly to length of horizons T_{opt} in LRS. In the bottom panel, LQC allocates less aggressive capital, getting closer to zero, to all assets for higher discount rate ρ. Sensitive nature of optimal solutions to length of time horizons T_{opt} is common to those subject to the short sales constraint and those free from the constraint.



LRS Perfect Foresight

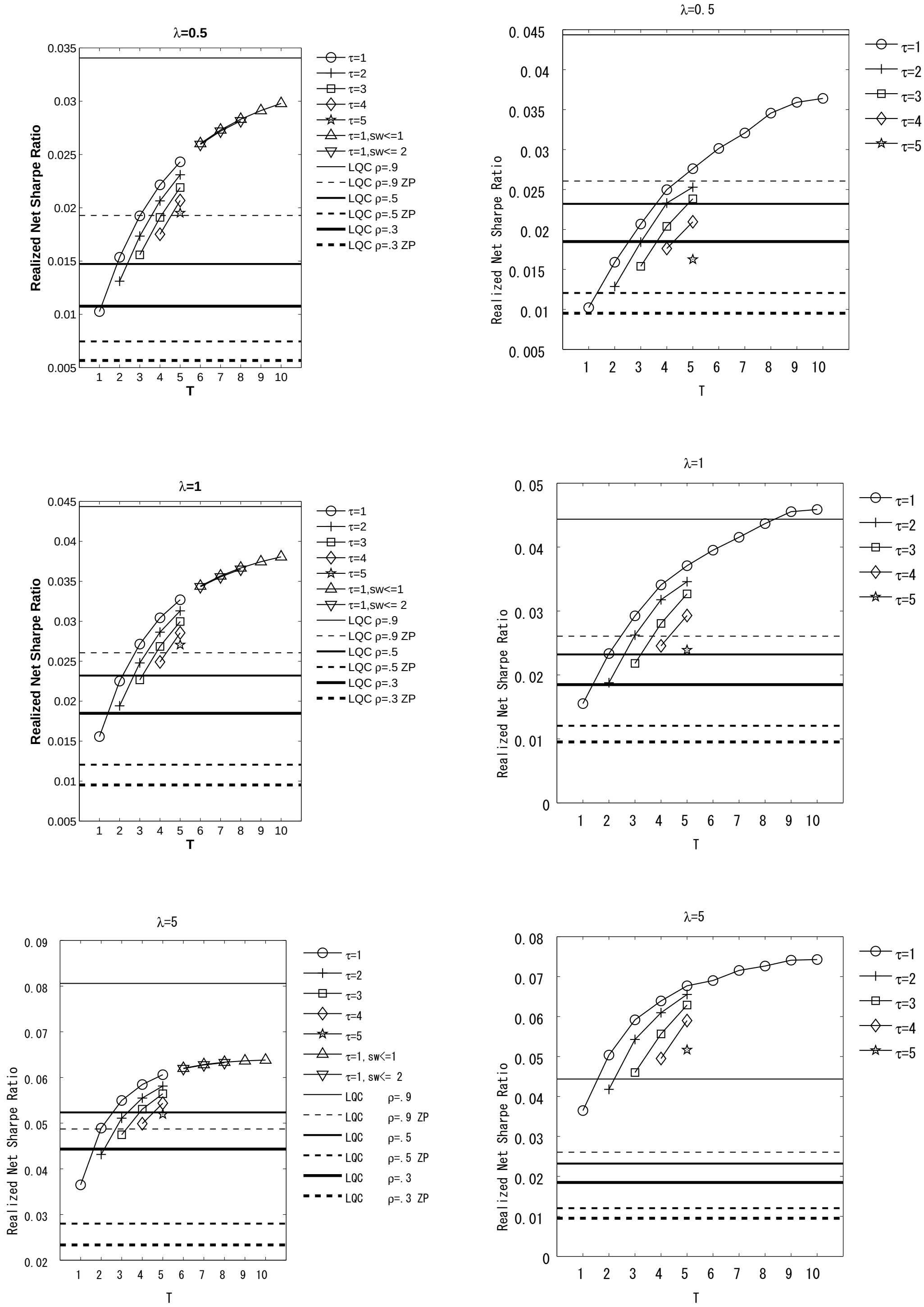
Linear Quadratic Control(LQC) in the discount rate $\rho \in \{0.3.0.5, 0.9\}$

sw: stands for number of maximum switches allowed

ZP: Zero Projection in LQC

For LRS under $\lambda = 5$, LQC $\rho = 0.3$ and LQC $\rho = 0.3$ ZP are far below the graphical range.

Figure 5.1: Realized Utility: $\widehat{U} = \widehat{\mu} - \frac{\lambda}{2}\widehat{\sigma}^2 - \widehat{TC}$



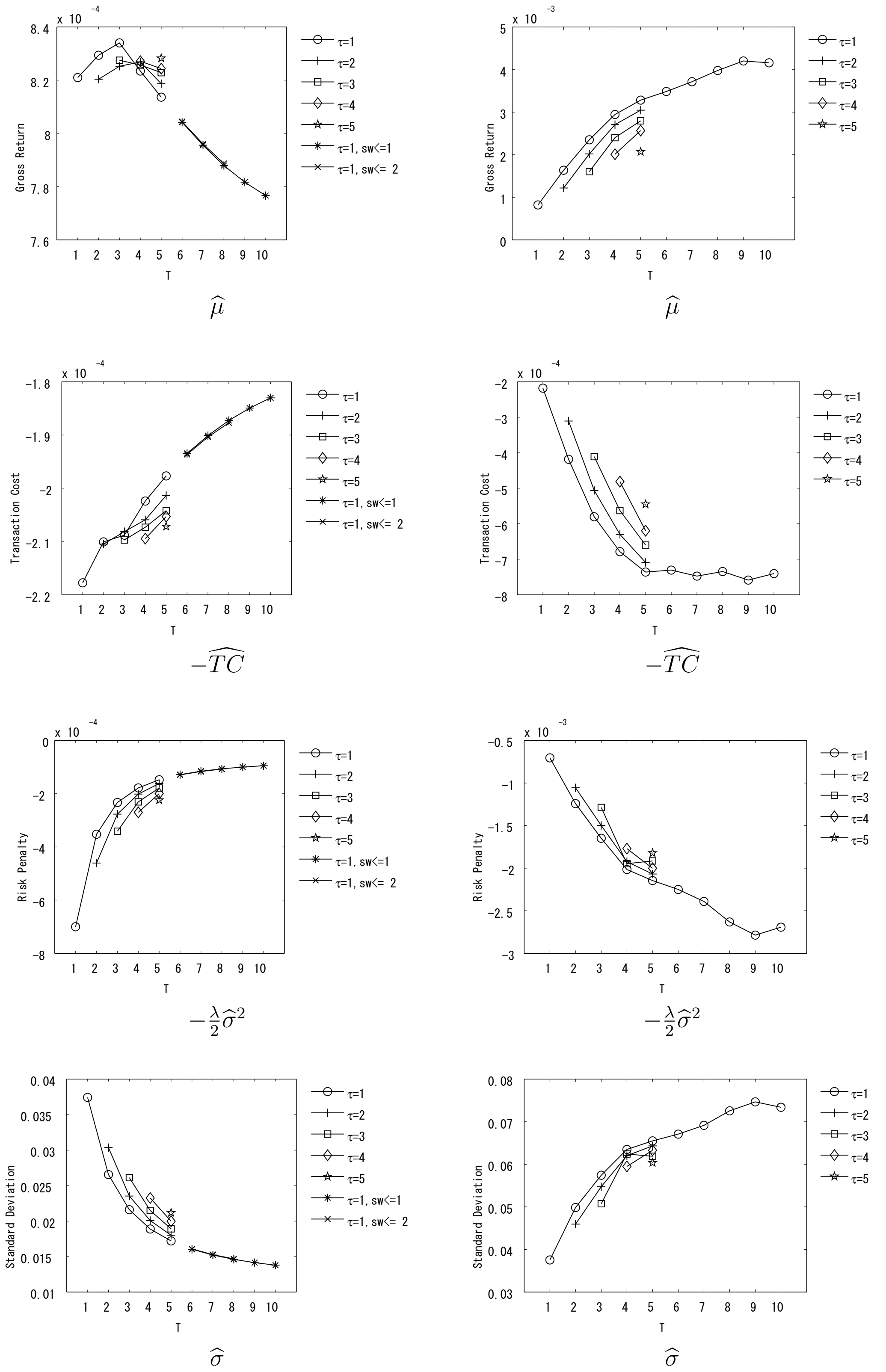
LRS Perfect Foresight

Linear Quadratic Control(LQC) in the discount rate $\rho \in \{0.3.0.5, 0.9\}$

sw: stands for number of maximum switches allowed

ZP: Zero Projection in LQC

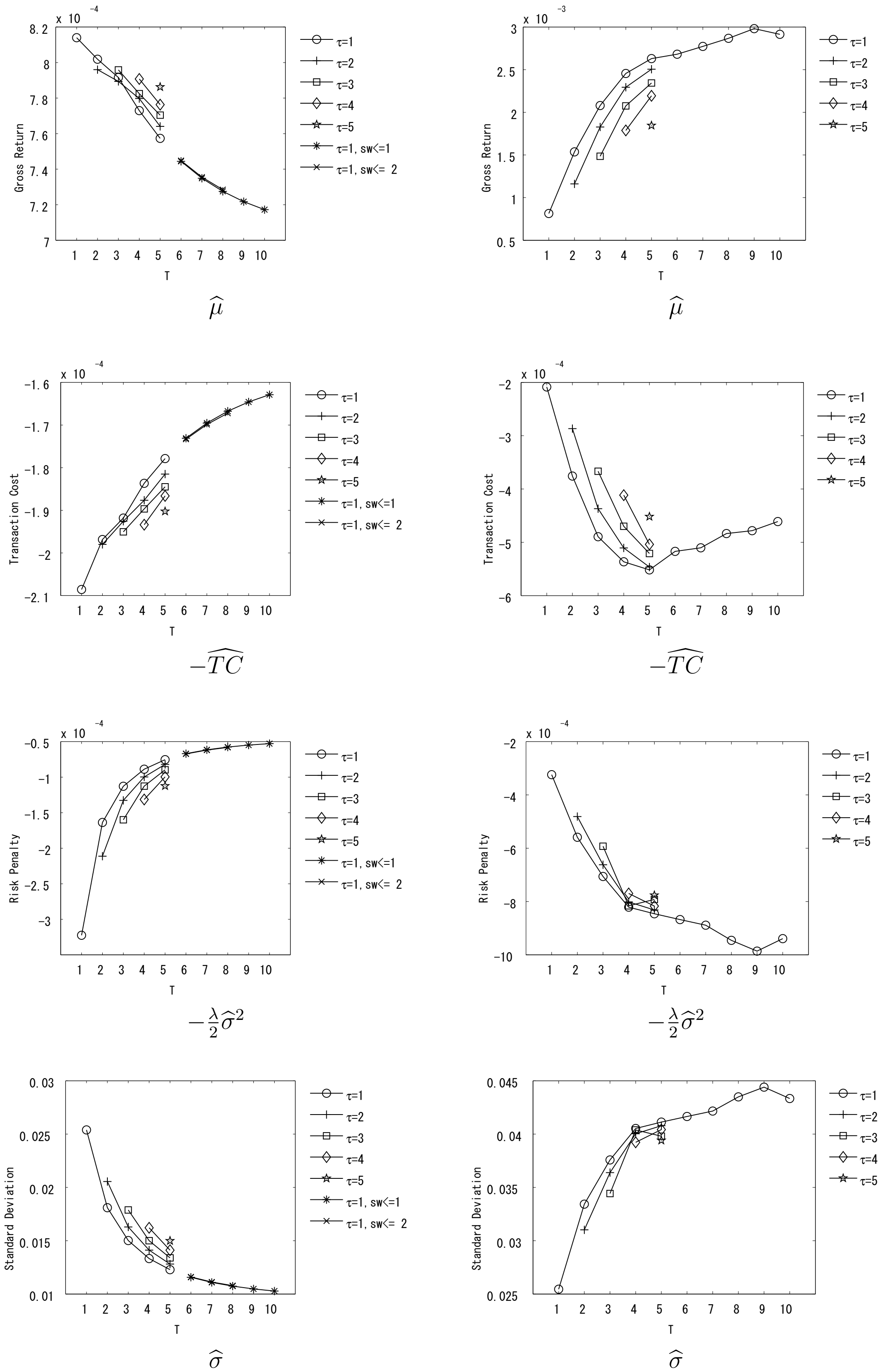
Figure 5.2: Realized Net Sharpe Ratio : $\widehat{SR} = \frac{\widehat{\mu}-\widehat{TC}}{\widehat{\sigma}}$



LRS Perfect Foresight

sw: stands for number of maximum switches allowed

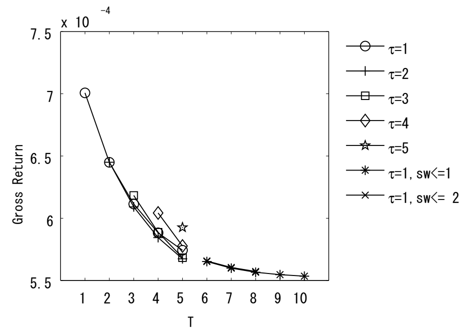
Figure 5.3: Attribution to Realized Utility and Standard Deviation: $\lambda = 0.5$



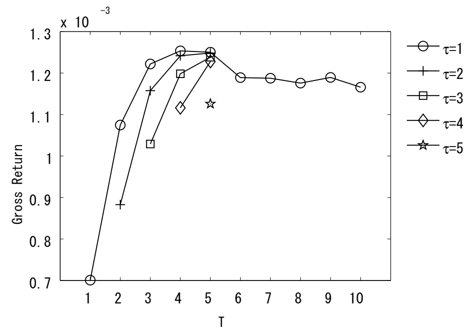
LRS Perfect Foresight

sw: stands for number of maximum switches allowed

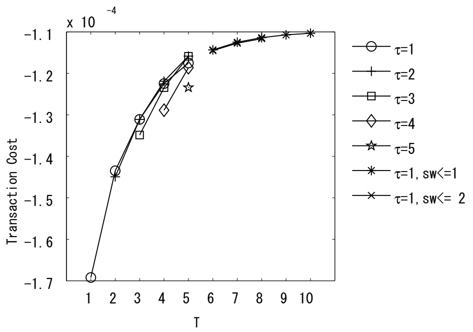
Figure 5.4: Attribution to Realized Utility and Standard Deviation: $\lambda = 1$



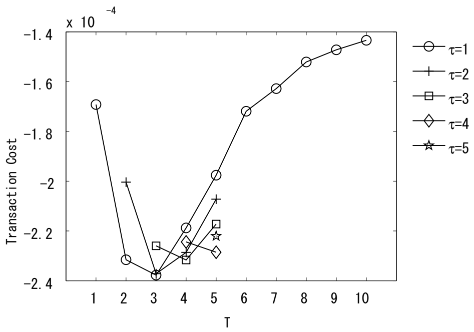
$\widehat{\mu}$



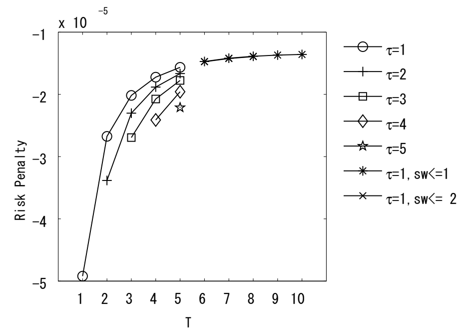
$\widehat{\mu}$



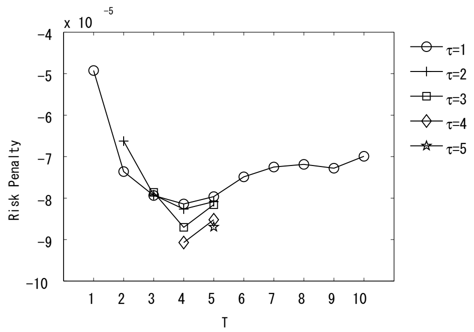
$-\widehat{TC}$



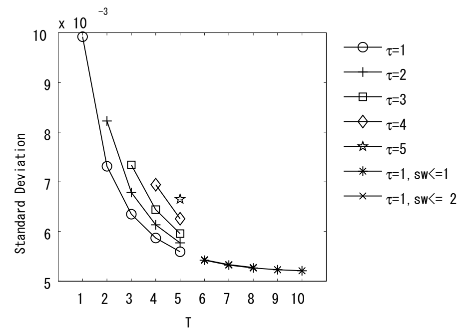
$-\widehat{TC}$



$-\frac{\lambda}{2}\widehat{\sigma}^2$

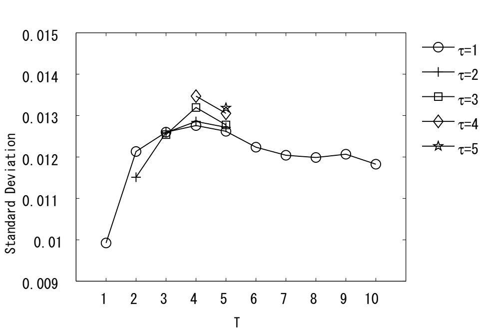


$-\frac{\lambda}{2}\widehat{\sigma}^2$



$\widehat{\sigma}$

LRS

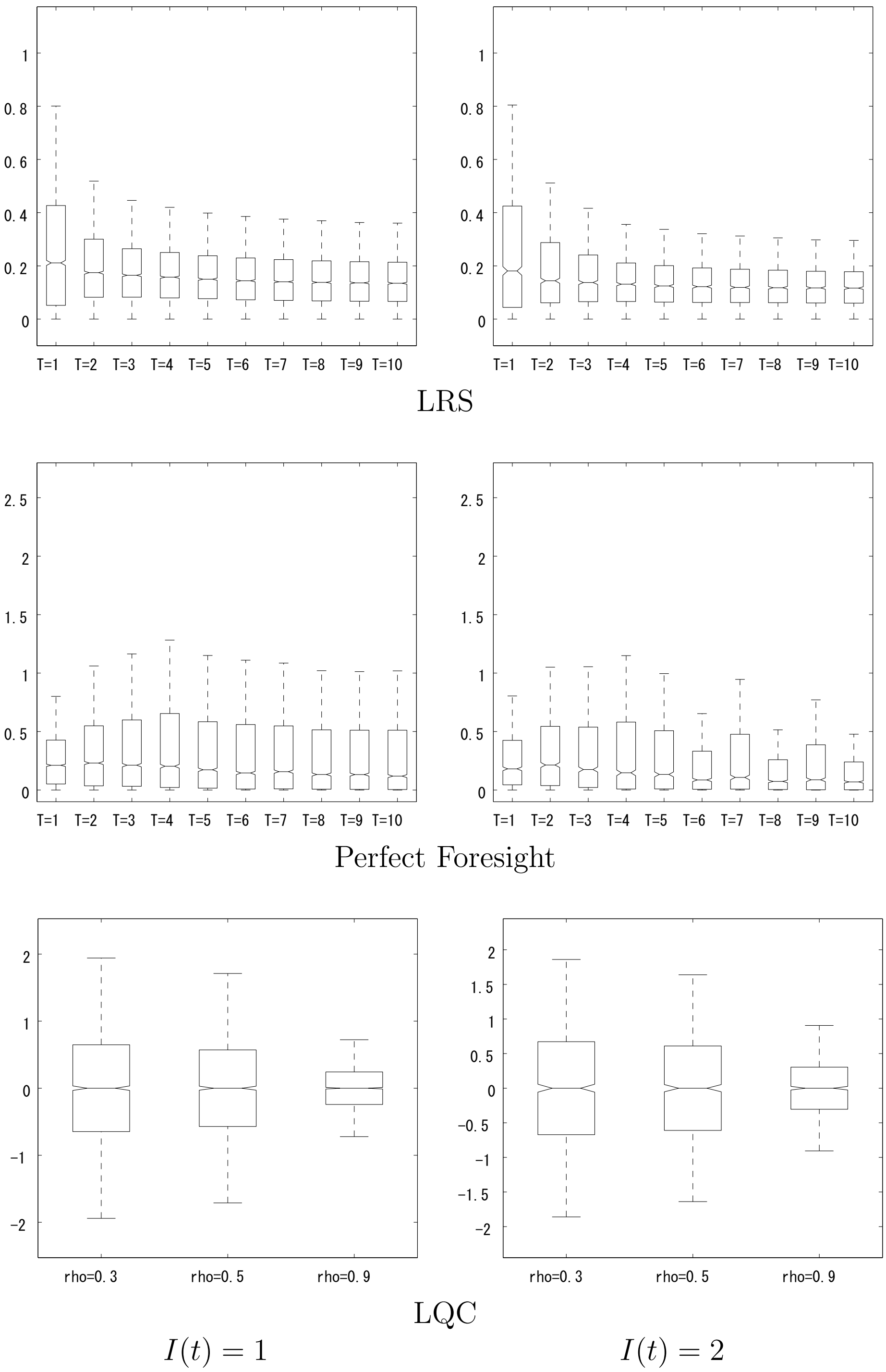


$\widehat{\sigma}$

Perfect Foresight

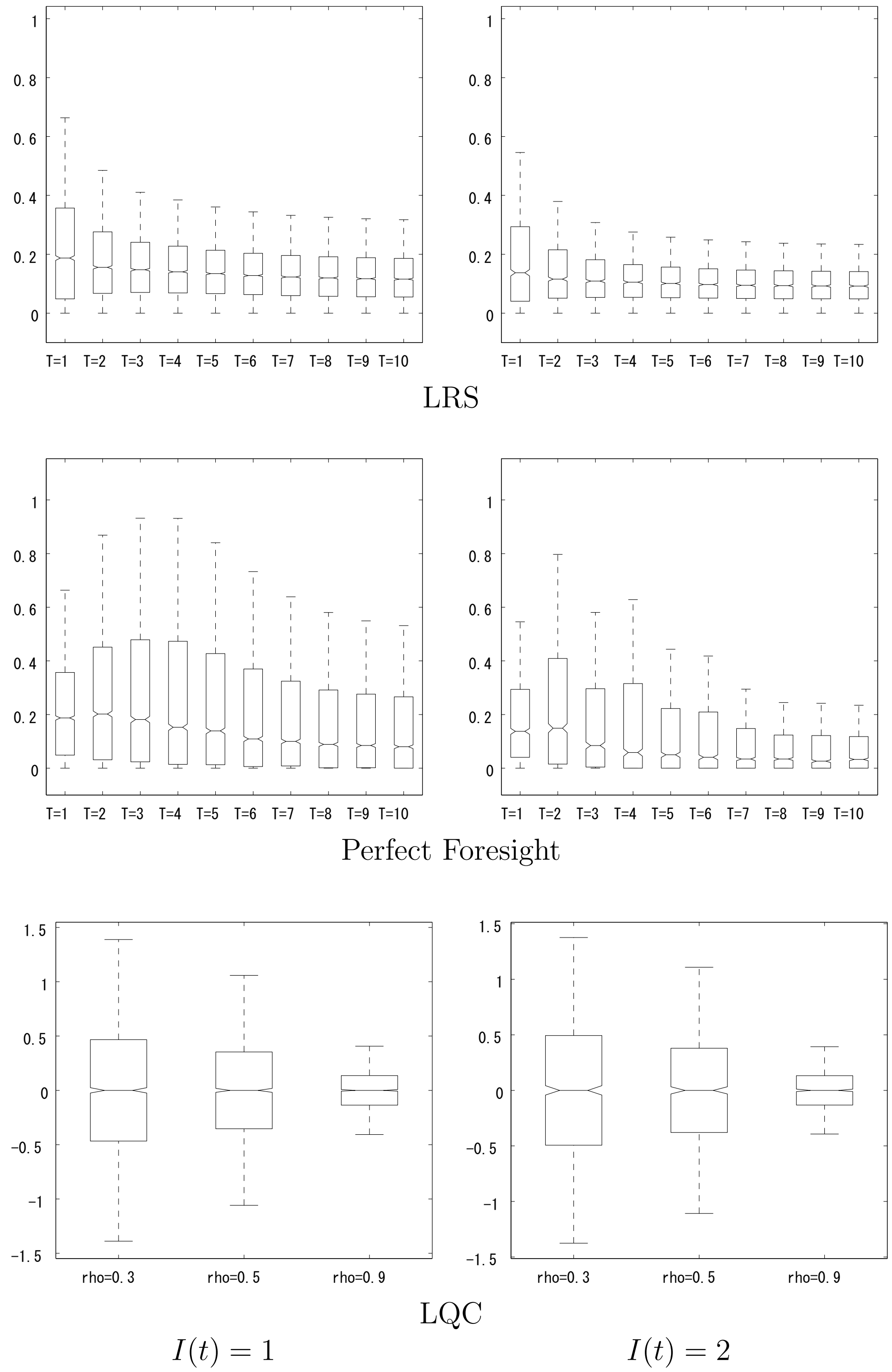
sw: stands for number of maximum switches allowed

Figure 5.5: Attribution to Realized Utility and Standard Deviation: $\lambda = 5$



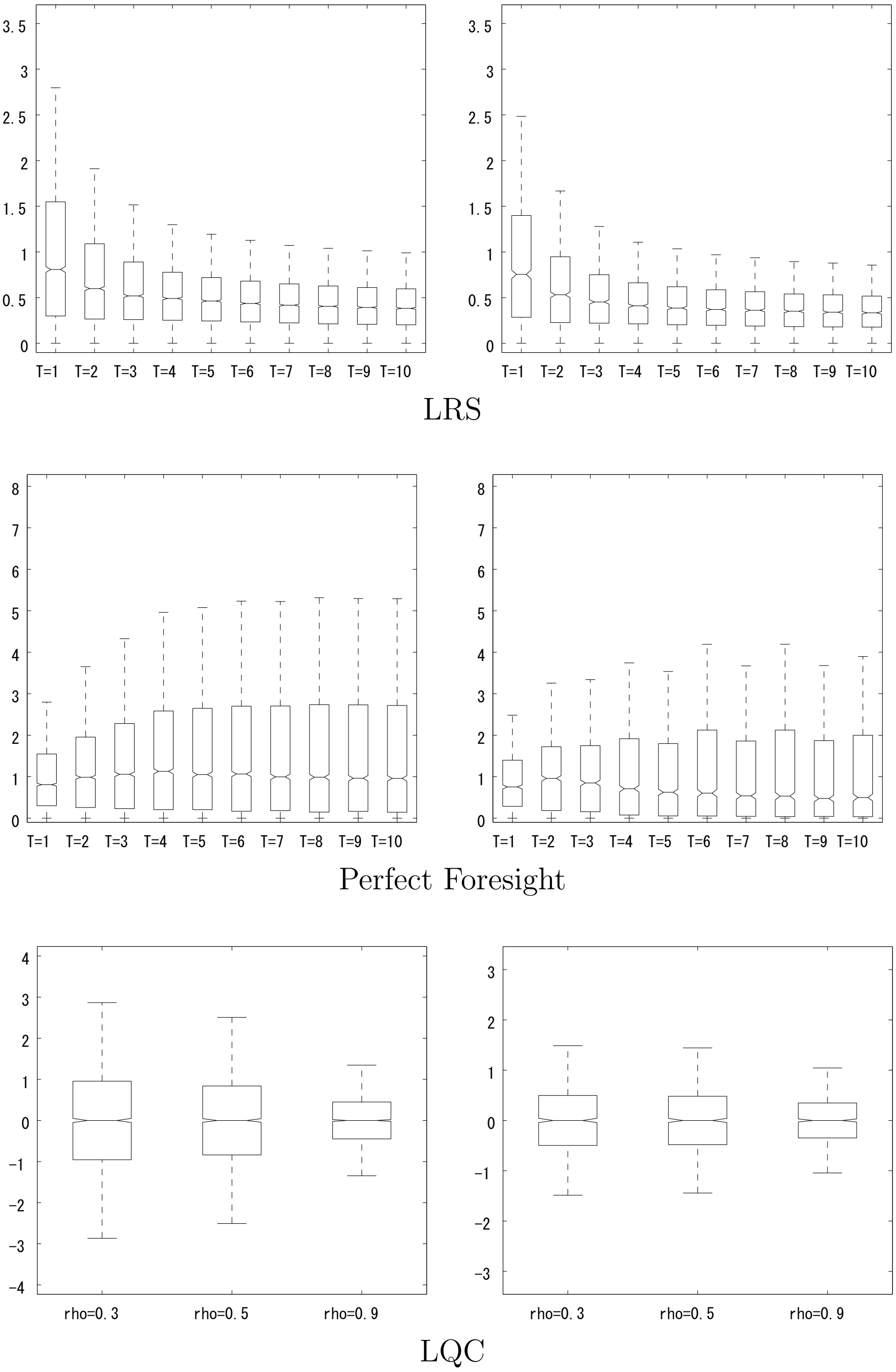
75 percentile($q3$) and 25 percentile($q1$) points of samples shown as upper and lower end of each box plot. $q3 + 1.5 \times (q3 - q1)$ located at the upper end and $q3 - 1.5 \times (q3 - q1)$ at lower end for each whisker.

Figure 5.6: Distribution of $\boldsymbol{x}^*$ for asset SG : $\lambda = 1$



75 percentile($q3$) and 25 percentile($q1$) points of samples shown as upper and lower end of each box plot. $q3 + 1.5 \times (q3 - q1)$ located at the upper end and $q3 - 1.5 \times (q3 - q1)$ at lower end for each whisker.

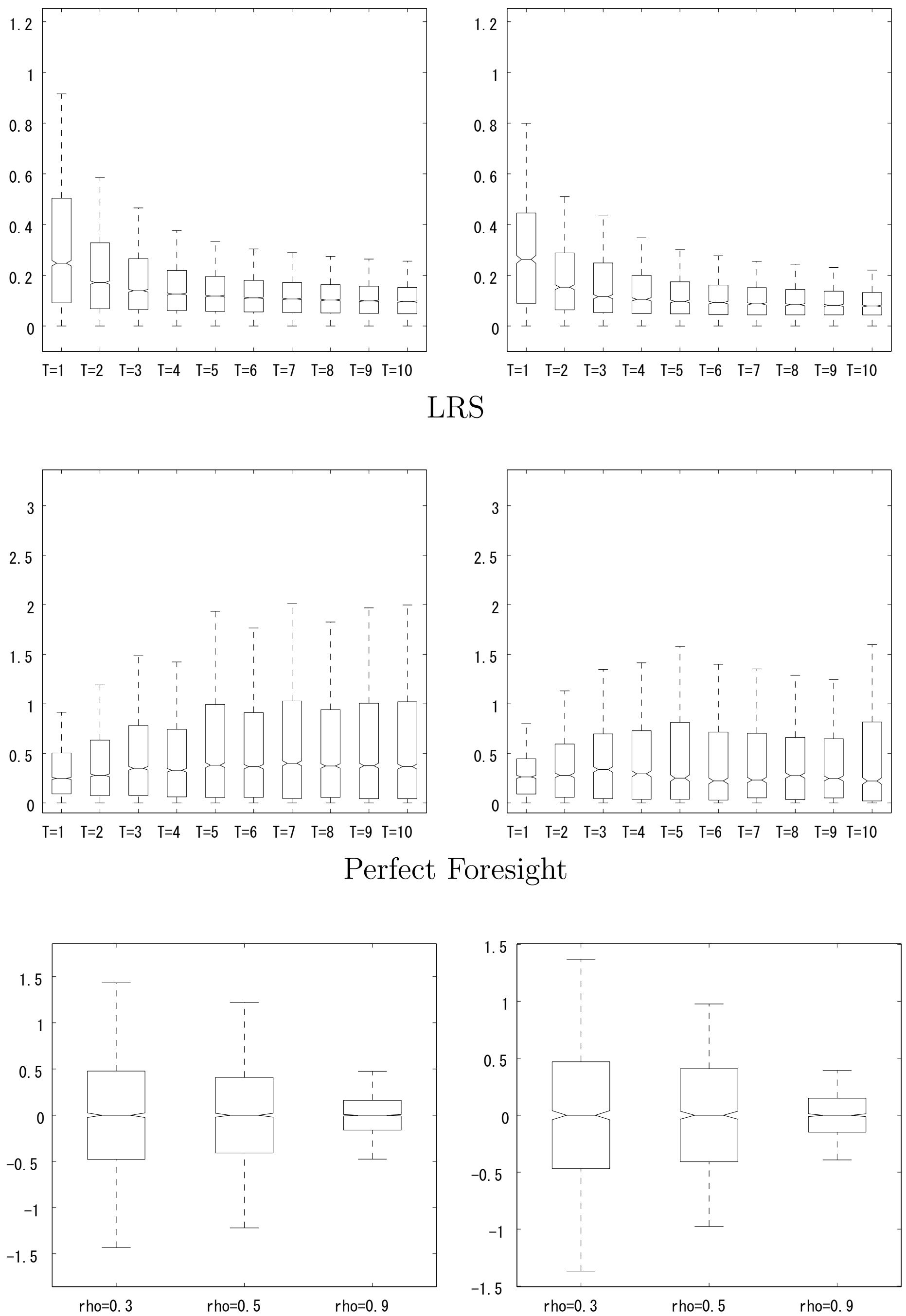
Figure 5.7: Distribution of $\boldsymbol{x}^*$ for asset SN : $\lambda = 1$



$I(t) = 1$ $\qquad$ $I(t) = 2$

75 percentile($q3$) and 25 percentile($q1$) points of samples shown as upper and lower end of each box plot. $q3 + 1.5 \times (q3 - q1)$ located at the upper end and $q3 - 1.5 \times (q3 - q1)$ at lower end for each whisker.

Figure 5.8: Distribution of $\boldsymbol{x}^*$ for asset SV : $\lambda = 1$

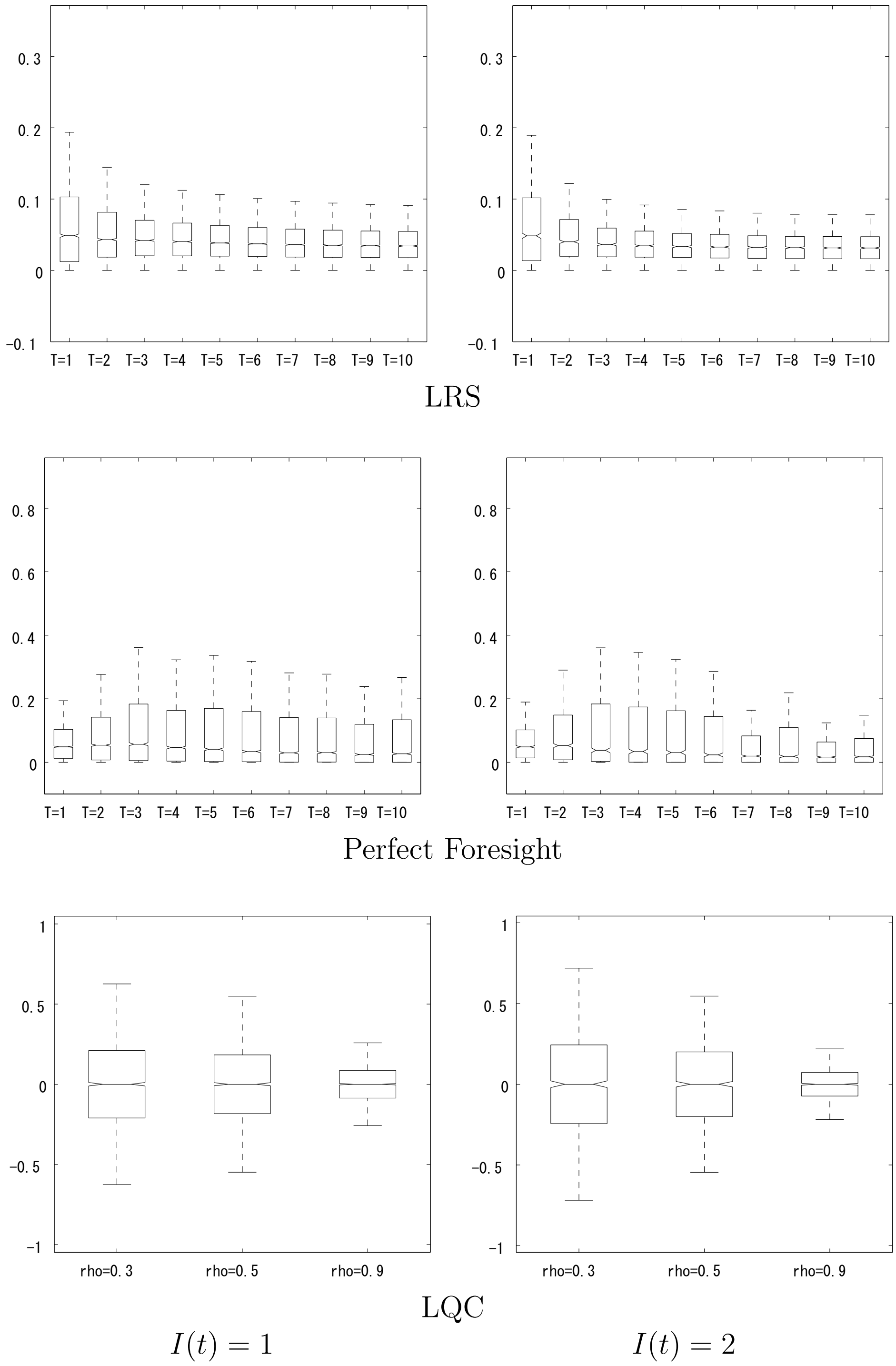


LQC

$I(t) = 1$ $I(t) = 2$

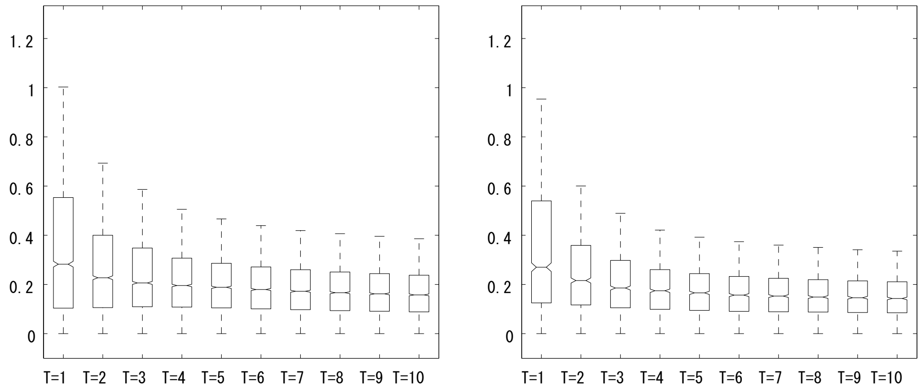
75 percentile($q3$) and 25 percentile($q1$) points of samples shown as upper and lower end of each box plot. $q3 + 1.5 \times (q3 - q1)$ located at the upper end and $q3 - 1.5 \times (q3 - q1)$ at lower end for each whisker.

Figure 5.9: Distribution of $\boldsymbol{x}^*$ for asset BG : $\lambda = 1$

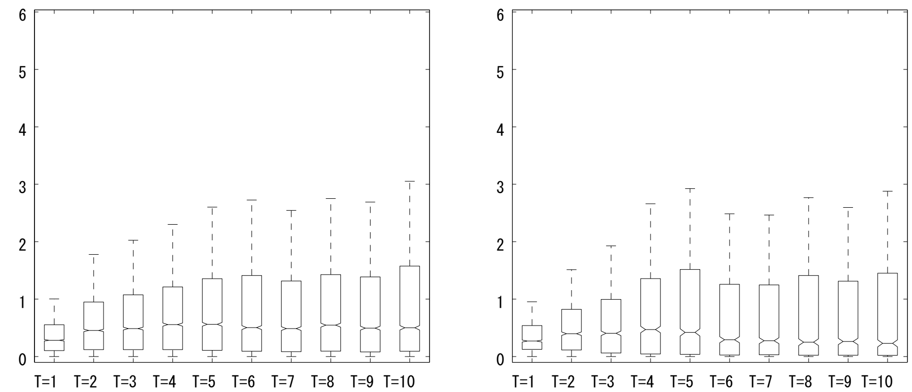


75 percentile($q3$) and 25 percentile($q1$) points of samples shown as upper and lower end of each box plot. $q3 + 1.5 \times (q3 - q1)$ located at the upper end and $q3 - 1.5 \times (q3 - q1)$ at lower end for each whisker.

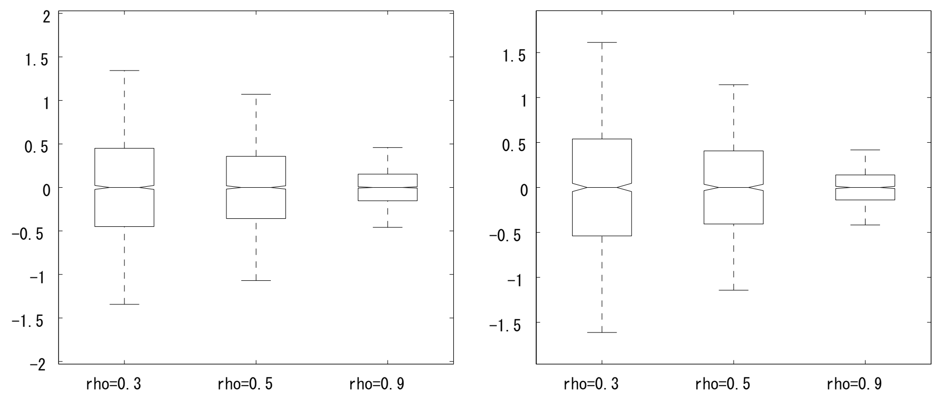
Figure 5.10: Distribution of $\boldsymbol{x}^*$ for asset BN : $\lambda = 1$



LRS



Perfect Foresight



LQC

$I(t) = 1$ $I(t) = 2$

75 percentile($q3$) and 25 percentile($q1$) points of samples shown as upper and lower end of each box plot. $q3 + 1.5 \times (q3 - q1)$ located at the upper end and $q3 - 1.5 \times (q3 - q1)$ at lower end for each whisker.

Figure 5.11: Distribution of $\boldsymbol{x}^*$ for asset BV : $\lambda = 1$

5.4 Concluding Remarks

In a dynamic investment problem over regime switching pricing on assets and factors as well as key parameters in objective functions, a key finding in Chapter 4 is that optimal solutions take form of a linear combination of current portfolio holdings and a function of observed factors. This chapter pays attentions to the fact and applies the Linear Rebalancing Rules for the finding to achieve optimized solutions even under investment constraints. To this end, the problem solved in Chapter 4 is formulated under the Linear Rebalancing Rules and solved under the second order cone problem. Numerical experiments uncover that optimized solutions perform as expected despite the solutions are not exact but numerical approximation with reasonable basis. A device to save computation burden enables investors to compute the solutions for practically long period of horizons at the expense of negligibly minimal deterioration of investment efficacy. Studies in this chapter claims that dynamic investment on regime dependent asset pricing is feasible for practical investments subject to investment constraints and size of problems to solve. Future research is planned on a couple of front of more advanced models to specify factor dynamics and budget constraints to impose.

5.5 Appendix

5.5.1 Derivation of the Optimization Problem

In this section, we will prove Proposition 5.2.1 and derive explicit representations of the objective function as well as short sales constraint given in Section 5.2. We first show a preliminary result.

Let $\boldsymbol{A}$ be an $N \times (1 + (t-1)M)$ matrix which is measurable with respect to $\sigma(I[t])$, a filtration generated by $I[t]$, and define

$$\boldsymbol{y}(t) = \boldsymbol{A}\boldsymbol{F}[t]. \tag{5.25}$$

Lemma 5.5.1 An $N \times N$ matrix $\boldsymbol{Q}$ that is measurable with respect to $\sigma(I[t])$ satisfies

$$\mathrm{E}_{I[t]}\left(\boldsymbol{y}^\top(t)\boldsymbol{Q}\boldsymbol{y}(t)\right) = \mathrm{vec}(\boldsymbol{A}^\top)^\top \left\{\boldsymbol{Q} \otimes \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\boldsymbol{F}[t]^\top\right)\right\} \mathrm{vec}(\boldsymbol{A}^\top). \tag{5.26}$$

Proof Let $\boldsymbol{a}_i^\top$ denote an ith row vector of $\boldsymbol{A}$ and let $\boldsymbol{Q}_{i,j}$ denote an (i,j)-element of $\boldsymbol{Q}$. Then, we get

$$\begin{aligned}
\mathrm{E}_{I[t]}\left(\boldsymbol{y}(t)^\top\boldsymbol{Q}\boldsymbol{y}(t)\right) &= \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]^\top\boldsymbol{A}^\top\boldsymbol{Q}\boldsymbol{A}\boldsymbol{F}[t]\right) \\
&= \sum_{i=1}^N\sum_{j=1}^N \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]^\top\boldsymbol{a}_i\boldsymbol{Q}_{i,j}\boldsymbol{a}_j^\top\boldsymbol{F}[t]\right) \\
&= \sum_{i=1}^N\sum_{j=1}^N \boldsymbol{Q}_{i,j}\boldsymbol{a}_i^\top \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\boldsymbol{F}[t]^\top\right)\boldsymbol{a}_j
\end{aligned}$$

which is rewritten as (5.26). □

Now we are in a position to show Proposition 5.2.1.

Proof of Proposition 5.2.1 Since $\boldsymbol{f}(t)$ is observable and $I_1(2)$ is predicted at $t=1$,

$$\mathrm{E}\left(\boldsymbol{x}(1)^\top \boldsymbol{L}_{I_1(2)} \boldsymbol{f}(1)\right) = \boldsymbol{x}(1)^\top \boldsymbol{L}_{I_1(2)} \boldsymbol{f}(1) = \boldsymbol{C}_{I[1]}^\top \boldsymbol{L}_{I_1(2)} \boldsymbol{f}(1). \tag{5.27}$$

For $t = 2, \ldots, T$, we obtain from (5.10)

$$\begin{aligned}
\mathrm{E}\left(\boldsymbol{x}(t)^\top \boldsymbol{L}_{I_t(t+1)} \boldsymbol{f}(t)\right) &= \mathrm{E}\left(\boldsymbol{F}[t]^\top \boldsymbol{C}_{I[t]}^\top \boldsymbol{L}_{I_t(t+1)} \boldsymbol{f}(t)\right) \\
&= \mathrm{E}\left(\boldsymbol{F}[t]^\top \sum_{i=1}^{N} \boldsymbol{c}_{I[t],i} \boldsymbol{\ell}_{I_t(t+1),i}^\top \boldsymbol{f}(t)\right) \\
&= \mathrm{E}\left(\sum_{i=1}^{N} \mathrm{E}_{I[t]}\left(\boldsymbol{c}_{I[t],i} \boldsymbol{F}[t]^\top \boldsymbol{f}(t) \boldsymbol{\ell}_{I_t(t+1),i}^\top\right)\right) \\
&= \mathrm{E}\left(\sum_{i=1}^{N} \boldsymbol{c}_{I[t],i}^\top \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t] \boldsymbol{f}(t)^\top\right) \boldsymbol{\ell}_{I_t(t+1),i}\right) \\
&= \mathrm{E}\left(\mathrm{vec}(\boldsymbol{C}_{I[t]}^\top)^\top \left\{\boldsymbol{I}_N \otimes \mathrm{E}_{I[t]}\left(\boldsymbol{F}[t] \boldsymbol{f}(t)^\top\right)\right\} \mathrm{vec}(\boldsymbol{L}_{I_t(t+1)}^\top)\right)
\end{aligned}$$

where $\boldsymbol{\ell}_{I_t(t+1),i}^\top$ denotes an ith row vector of $\boldsymbol{L}_{I_t(t+1)}$. Then, (5.11) is derived. From (5.10), (5.12) is obtained by plugging $\boldsymbol{A} = \boldsymbol{C}_{I[t]}$ and $\boldsymbol{Q} = \boldsymbol{W}_{I_t(t+1)}$ into (5.26). Simiraly, (5.13) can be proved by substituting $\boldsymbol{A} = \boldsymbol{C}_{I[t]}$ and $\boldsymbol{Q} = \boldsymbol{B}_{I_t(t+1)}$ into (5.26). This proves the lemma. □

We will express the conditional expectations $\mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\boldsymbol{f}(t)^\top\right)$ in (5.11), and $\mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\boldsymbol{F}[t]^\top\right)$ in (5.12) and (5.13) in terms of the model parameters. Since

$$\mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\boldsymbol{f}(t)^\top\right) = \mathrm{E}_{I[t]}\left(\begin{bmatrix} \boldsymbol{f}(t)^\top \\ \boldsymbol{f}(2)\boldsymbol{f}(t)^\top \\ \vdots \\ \boldsymbol{f}(t)\boldsymbol{f}(t)^\top \end{bmatrix}\right) \tag{5.28}$$

and

$$\mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\boldsymbol{F}[t]^\top\right) = \mathrm{E}_{I[t]}\left(\begin{bmatrix} 1 & \boldsymbol{f}(2)^\top & \cdots & \boldsymbol{f}(t)^\top \\ \boldsymbol{f}(2) & \boldsymbol{f}(2)\boldsymbol{f}(2)^\top & \cdots & \boldsymbol{f}(2)\boldsymbol{f}(t)^\top \\ \vdots & & \ddots & \vdots \\ \boldsymbol{f}(t) & \boldsymbol{f}(t)\boldsymbol{f}(2)^\top & \cdots & \boldsymbol{f}(t)\boldsymbol{f}(t)^\top \end{bmatrix}\right), \tag{5.29}$$

the problem is reduced to calculate $\mathrm{E}_{I[t]}\left(\boldsymbol{f}(t)\right)$ and $\mathrm{E}_{I[t]}\left(\boldsymbol{f}(s)\boldsymbol{f}(u)^\top\right)$ $(2 \le s, u \le t)$. By recursively solving (5.2), we obtain

$$\boldsymbol{f}(t) = \boldsymbol{\Psi}_{I(2:t)} \boldsymbol{f}(1) + \sum_{s=2}^{t} \boldsymbol{\Psi}_{I(s+1:t)} \boldsymbol{\mu}_{I(s)} + \sum_{s=2}^{t} \boldsymbol{\Psi}_{I(s+1:t)} \boldsymbol{\epsilon}(s) \tag{5.30}$$

where $I(s:t) = \{I(s), \ldots, I(t)\}$ and

$$\boldsymbol{\Psi}_{I(s:t)} = \begin{cases} \boldsymbol{\Phi}_{I(t)} \times \boldsymbol{\Phi}_{I(t-1)} \times \cdots \times \boldsymbol{\Phi}_{I(s)}, & s \leq t \\ \boldsymbol{I}_M, & s = t+1. \end{cases} \tag{5.31}$$

Since

$$\mathrm{E}_{I[t]}\left(\boldsymbol{\Psi}_{I(s+1:t)}\boldsymbol{\epsilon}(s)\right) = \boldsymbol{\Psi}_{I(s+1:t)}\mathrm{E}_{I[t]}\left(\boldsymbol{\epsilon}(s)\right) = \mathbf{0}$$

for $s \leq t$ by the assumption, the conditional expectation of $\boldsymbol{f}(t)$ is given by

$$\boldsymbol{\alpha}_{I[t]} = \mathrm{E}_{I[t]}\left(\boldsymbol{f}(t)\right) = \boldsymbol{\Psi}_{I(2:t)}\boldsymbol{f}(1) + \sum_{s=2}^{t} \boldsymbol{\Psi}_{I(s+1:t)}\boldsymbol{\mu}_{I(s)}. \tag{5.32}$$

Moreover, $\mathrm{E}_{I[t]}\left(\boldsymbol{f}(s)\boldsymbol{f}(u)^\top\right)$ is given in the next lemma.

Lemma 5.5.2 For $2 \leq r, u \leq t$, we obtain

$$\mathrm{E}_{I[t]}\left(\boldsymbol{f}(r)\boldsymbol{f}^\top(u)\right) = \boldsymbol{\alpha}_{I[r]}\boldsymbol{\alpha}_{I[u]}^\top + \sum_{a=2}^{\min(r,u)} \boldsymbol{\Psi}_{I(a+1,r)}\boldsymbol{\Sigma}_{I(a)}\boldsymbol{\Psi}_{I(a+1,u)}^\top. \tag{5.33}$$

Proof For $r = 2, \ldots, t$, let

$$\boldsymbol{\beta}(r) = \boldsymbol{f}(r) - \boldsymbol{\alpha}_{I[r]} = \sum_{a=2}^{r} \boldsymbol{\Psi}_{I(a+1:r)}\boldsymbol{\epsilon}(a). \tag{5.34}$$

Since $\mathrm{E}_{I[t]}\left(\boldsymbol{\beta}(r)\right) = \mathbf{0}$ and $\boldsymbol{\alpha}(r)$ is measurable with respect to $\sigma(I[t])$, we obtain

$$\mathrm{E}_{I[t]}\left(\boldsymbol{f}(r)\boldsymbol{f}^\top(u)\right) = \boldsymbol{\alpha}_{I[r]}\boldsymbol{\alpha}_{I[u]}^\top + \mathrm{E}_{I[t]}\left(\boldsymbol{\beta}(r)\boldsymbol{\beta}(u)^\top\right). \tag{5.35}$$

Noting that

$$\mathrm{E}_{I[t]}\left(\boldsymbol{\epsilon}(a)\boldsymbol{\epsilon}(b)^\top\right) = \begin{cases} \boldsymbol{\Sigma}_{I(a)}, & a = b \\ \boldsymbol{O}, & a \neq b, \end{cases} \tag{5.36}$$

the second term in (5.35) is calculated as

$$\begin{aligned} \mathrm{E}_{I[t]}\left(\boldsymbol{\beta}(r)\boldsymbol{\beta}(u)^\top\right) &= \mathrm{E}_{I[t]}\left(\left\{\sum_{a=2}^{r} \boldsymbol{\Psi}_{I(a+1,r)}\boldsymbol{\epsilon}(a)\right\}\left\{\sum_{b=2}^{u} \boldsymbol{\Psi}_{I(b+1,u)}\boldsymbol{\epsilon}(b)\right\}^\top\right) \\ &= \sum_{a=2}^{\min(r,u)} \boldsymbol{\Psi}_{I(a+1,r)}\mathrm{E}_{I[t]}\left(\boldsymbol{\epsilon}(a)\boldsymbol{\epsilon}(a)^\top\right)\boldsymbol{\Psi}_{I(a+1,u)}^\top \\ &= \sum_{a=2}^{\min(r,u)} \boldsymbol{\Psi}_{I(a+1,r)}\boldsymbol{\Sigma}_{I(a)}\boldsymbol{\Psi}_{I(a+1,u)}^\top, \end{aligned} \tag{5.37}$$

completing the proof. □

Finally, we derive the short sales constraint (5.18). Given a sample path of $I[t]$, $\boldsymbol{\epsilon}(s)$ $(s = 2, \ldots, t)$ in (5.2) follows a multivariate normal distribution. Since $\boldsymbol{f}(t)$ in (5.30) is a linear combination of $\boldsymbol{\epsilon}(s)$'s, $\boldsymbol{f}(t)$ also follows a multivariate normal distribution for a given $\boldsymbol{f}(1)$ and $I[t]$, and so is $\boldsymbol{x}(t) = \boldsymbol{C}_{I[t]}\boldsymbol{F}[t]$. The conditional mean and covariance are calculated as

$$\mathrm{E}_{I[t]}\left(\boldsymbol{x}(t)\right) = \boldsymbol{C}_{I[t]}\mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\right) \tag{5.38}$$

and

$$\begin{aligned}
\mathrm{V}_{I[t]}\left(\boldsymbol{x}(t)\right) &= \mathrm{E}_{I[t]}\left(\{\boldsymbol{x}(t) - \mathrm{E}_{I[t]}\left(\boldsymbol{x}(t)\right)\}\{\boldsymbol{x}(t) - \mathrm{E}_{I[t]}\left(\boldsymbol{x}(t)\right)\}^\top\right) \\
&= \boldsymbol{C}_{I[t]}\mathrm{E}_{I[t]}\left(\{\boldsymbol{F}(t) - \mathrm{E}_{I[t]}\left(\boldsymbol{F}(t)\right)\}\{\boldsymbol{F}(t) - \mathrm{E}_{I[t]}\left(\boldsymbol{F}(t)\right)\}^\top\right)\boldsymbol{C}_{I[t]}^\top \\
&= \boldsymbol{C}_{I[t]}\boldsymbol{\Lambda}_{I[t]}\boldsymbol{C}_{I[t]}^\top
\end{aligned} \tag{5.39}$$

where

$$\boldsymbol{\Lambda}_{I[t]} = \mathrm{E}_{I[t]}\left(\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & \boldsymbol{\beta}_{I[2]}\boldsymbol{\beta}_{I[2]}^\top & \cdots & \boldsymbol{\beta}_{I[2]}\boldsymbol{\beta}_{I[t]}^\top \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \boldsymbol{\beta}_{I[t]}\boldsymbol{\beta}_{I[2]}^\top & \cdots & \boldsymbol{\beta}_{I[t]}\boldsymbol{\beta}_{I[t]}^\top \end{bmatrix}\right). \tag{5.40}$$

Note that $\mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\right)$ and $\boldsymbol{\Lambda}_{I[t]}$ can be calculated from (5.32) and (5.37), respectively. For expected return and risks of $x_i(t)$ of asset $i = 1, \ldots, N$, we obtain

$$\begin{aligned}
\mathrm{E}_{I[t]}\left(x_i(t)\right) &= \boldsymbol{c}_{I[t],i}^\top\mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\right), \\
\mathrm{V}_{I[t]}\left(x_i(t)\right) &= \boldsymbol{c}_{I[t],i}^\top\boldsymbol{\Lambda}_{I[t]}\boldsymbol{c}_{I[t],i} = \boldsymbol{c}_{I[t],i}^\top\boldsymbol{\Theta}_{I[t]}^\top\boldsymbol{\Theta}_{I[t]}\boldsymbol{c}_{I[t],i}
\end{aligned}$$

where $\boldsymbol{\Theta}_{I[t]}^\top\boldsymbol{\Theta}_{I[t]} = \boldsymbol{\Lambda}_{I[t]}$ is a Cholesky decomposition. Now it is easy to see that, for a scalar standard normal random variable X with mean μ and variance σ^2,

$$\mathrm{P}\left(X < 0\right) \leq p \qquad \Longleftrightarrow \qquad \mu \geq \sigma\Phi^{-1}(1-p) \tag{5.41}$$

for $p \in (0, 0.5)$. Substituting $\mu = \boldsymbol{c}_{I[t],i}^\top\mathrm{E}_{I[t]}\left(\boldsymbol{F}[t]\right)$ and $\sigma = \|\boldsymbol{\Theta}_{I[t]}\boldsymbol{c}_{I[t],i}\|_2$ into (5.41) then yields (5.18).

5.5.2 Antithetic Variates

Facilitating effective convergence of the performances estimated by the Monte Carlo simulations, we apply variance reduction technique to factor $\boldsymbol{f}(t)$ and return $\boldsymbol{r}(t)$. We denote originally generated sample path of factor process by

$$\boldsymbol{f}^+(t) = \boldsymbol{\mu}_{I(t)} + \boldsymbol{\Phi}_{I(t)}\boldsymbol{f}(t-1) + \boldsymbol{\epsilon}_{I(t)}(t), \qquad t = 2, \ldots, T_{sim}$$

and calculate its antithetic variate by

$$\boldsymbol{f}^-(t) = \boldsymbol{\mu}_{I(t)} + \boldsymbol{\Phi}_{I(t)}\boldsymbol{f}(t-1) - \boldsymbol{\epsilon}_{I(t)}(t), \qquad t = 2, \ldots, T_{sim}.$$

Note that, since the distribution of $\boldsymbol{\epsilon}_{I(t)}(t)$ and $-\boldsymbol{\epsilon}_{I(t)}(t)$ are the same, so is $\boldsymbol{f}^{+}(t)$ and $\boldsymbol{f}^{-}(t)$. For each of factors $\boldsymbol{f}^{+}(t)$ and $\boldsymbol{f}^{-}(t)$, we generate sample paths of asset returns and its antithetic variate by

$$\begin{aligned}
\boldsymbol{r}^{++}(t) &= \boldsymbol{L}_{I(t-1)}\boldsymbol{f}^{+}(t-1) + \boldsymbol{u}_{I(t-1)}(t), \qquad t = 2,\ldots,T_{sim}+1 \\
\boldsymbol{r}^{+-}(t) &= \boldsymbol{L}_{I(t-1)}\boldsymbol{f}^{+}(t-1) - \boldsymbol{u}_{I(t-1)}(t), \qquad t = 2,\ldots,T_{sim}+1 \\
\boldsymbol{r}^{-+}(t) &= \boldsymbol{L}_{I(t-1)}\boldsymbol{f}^{-}(t-1) + \boldsymbol{u}_{I(t-1)}(t), \qquad t = 2,\ldots,T_{sim}+1 \\
\boldsymbol{r}^{--}(t) &= \boldsymbol{L}_{I(t-1)}\boldsymbol{f}^{-}(t-1) - \boldsymbol{u}_{I(t-1)}(t), \qquad t = 2,\ldots,T_{sim}+1.
\end{aligned}$$

By the same reasoning as above, four types of samples $\boldsymbol{r}^{++}(t), \boldsymbol{r}^{+-}(t), \boldsymbol{r}^{-+}(t), \boldsymbol{r}^{--}(t)$ obey the same probability law. Hence we obtain four performance measures calculated from $\boldsymbol{r}^{++}(t), \boldsymbol{r}^{+-}(t), \boldsymbol{r}^{-+}(t), \boldsymbol{r}^{--}(t)$, and average them to get the result of each simulation of length $T_{sim} = 2,500$. Since $\boldsymbol{r}^{++}(t)$ and $\boldsymbol{r}^{--}(t)$ are supposed to be negatively correlated, and so is $\boldsymbol{r}^{+-}(t)$ and $\boldsymbol{r}^{-+}(t)$, we expect the variance of the result will be largely reduced by taking average. This procedure is repeated 250 times and, in sum, total $2,500 \times 4 \times 250$ samples are used to calculate final results provided in Section 5.3.

Chapter 6

Conclusion and Future Issues

The main objective of studies in optimal portfolios is to develop models that provide investors with a formal process to shape the best set of assets to hold. To this end, investors are inevitably mindful of growing magnitude of uncertainties on behaviors of assets' returns. Another noteworthy fact is progressive innovations documented in literature of finance and investment to specify return prediction models. Provided that the prediction ability improves for assets' returns beyond one time step ahead, optimal portfolios should not be myopic but deliver investors portfolios that behaves optimally over the course of multi-period ahead in the future. Economically intuitive common factors predicting assets' return have attracted serious interests of investors. As is often with financial time series, it is not doubtful at all to expect the majority of factors to exhibit nature of time series. This opens one of important doors to consider dynamic investment on portfolios. Not just factors but such other pieces can be regime dependent as variability of returns to assets and factors, factor loading and key parameters that characterize objective functions, e.g., risk aversion and transaction costs. As such, this thesis examines how robustly an idea of the regime dependency explains uncertainty of assets' return enough to lead investors to reasonable portfolio selections in accountable ways.

As a basis for subsequent chapters, Chapter 3 played a necessary role to understand why multi-regime models improve investment efficacy for a mean-variance portfolio with transaction cost penalty. For simplicity, a myopic model was chosen. The normality test resulted in that multiple regime approaches reduce chances to mis-specify sector return generating processes that exhibit discontinuous behaviors. The more number of regimes, the more likely for the error terms to be normally distributed. Massive amount of advanced financial models have been documented to the literature for the purpose of complicated behavior of assets' prices. Our model assumes, however, the simplest assumptions as normal distributions to return variability. Mixture of normal distributions over regime switches that follow to the Markov process deserves additional pages to the literature. For practitioners, two things are notable to pay attentions to the outcome. First, for a sector rotation investment strategy in the US equity, only three regimes are large enough to keep up decent investment performance over nearly forty years long in the past in monthly space. Moreover, regime dependent investors, suppose he or she is risk averse, are resonated by regime dependent risk versions that further improve investment efficacy.

Chapter 4 was encouraged by Chapter 3 to extend Gârleanu and Pedersen (2013) to be regime dependent for key parameters for assets, factors and portfolio setup. Theoretical back borne of this entire thesis was built in this chapter. Regardless of presence or absence of regime dependency, a commonality held across the optimal portfolios was a weighted average of the current portfolio and a target portfolio which is a function of return to the factors. A major difference was that the weight are regime dependent in regime dependent optimal solutions. It was firmly confirmed that an analytical solution achieved by Gârleanu and Pedersen (2013) was a special case with no regime dependency in the solution found in this chapter. On the empirical front, the Fama-French 6 portfolios formed on Size and Book-to-Market (2×3) as assets in the optimal portfolio exhibited regime dependent behavior more clearly in tranquil regimes than in turbulent periods represented by the burst of technology bubble, credit bubble followed by the US and European sovereign crisis, experienced in the last decade. In both of in-sample and out-of-sample periods, optimal portfolios built by the model demonstrate that the regime switching models exhibit superior performance over the single regime model for popular types performance measures.

Chapter 5 attracts special attentions of investors who are imposed by investment constraints. As is well shared limitations to research effort attempting to derive analytical solutions under investment constraints, the tractable solution achieved in Chapter 4 faces limited opportunities to extend the solution to those under constraints. Taking a closer look at the shape of the semi-analytical solution found in Chapter 4, one can find it is a form of the linear rebalancing rules that Moallemi and Sağlam (2013) formulates wide class of portfolio optimization problems including a solution achieved by Gârleanu and Pedersen (2013) as a sort of linear quadratic control problem. One of goals of Chapter 5 was reached by extending the formulation to regime dependent data generation process of assets and factors as well as key parameters in objective function to maximize. Numerical experiments import one of empirical models estimated in Chapter 4 and demonstrate decent ability of the formulation to compute optimized portfolio in the second order cone problem under a constraint imposed of no short sales. In the finite time horizon, the demonstration confirmed better investment performance in longer horizons scoped by an objective function. Also performed reasonable efficacy relative to upper and lower bounds. For a practical feasibility to implement the formulation in place, under finite computational resources, it is informative to notice common nature of estimated elements in the transition probability matrix. A device worked to efficiently reduce computation burdens by elimination such investment paths where number of switches are meaninglessly high because of significantly low probability to take place. The formulation and numerical demonstrations deliver regime aware investors promising future opportunities to take advantage of the proposing framework to portfolio practices.

Throughout all over the studies in this thesis, three outcomes were presented;

- Tractable models for optimal portfolio,
- Implementable formulation for optimized portfolios and

- Empirical evidences to support regime dependent model specification and optimal and optimized portfolios.

Table 6.1 summarizes a high level review of this thesis.

Additional set of research is pointed out to plan going forward in the future. First, the tractable solutions Chapter 4 achieved does not necessarily require normal distributions to assume for data generation process of assets or factors. As long as symmetric distributions, e.g., such fat tail specification as the t-distribution, tractable solutions can apply. Further improvements are possible if non Gaussian distribution is assumed. Second, room exists for investors who assume the transition probabilities are function of such exogenous variables as market data and macro economic data, typically augmenting the probabilities to be time variant described by the Markov switching logistic function. Third, parameters in regime dependent data generation process can also be time variant. Estimating the parameters may require such advanced analytics as the particle filter. Fourth, extensions of the model of practical importance include more general transaction cost functions such as linear transaction cost with constant. This extension is generally challenging in the mean-variance investment utility to apply. Fifth, the budget constraint, e.g., no leverage or no borrowing, is important for the formulation to assume. Optimal solutions under the constraint are non trivial because the fact that multiplications of stochastic variables are present in the formulation in the linear rebalancing rules. Finally, generalization of factor dynamics from VAR(1) to higher order models is a natural extension of the models and formulations studied in Chapter 4 and Chapter 5. Prior to a problem formulation as completed in Chapter 5, careful examination on shapes of optimal solution is a must. Tractable and analytical solutions derived under no constraint are always theoretical basements to consider formulating problems under constraints.

<table>
<tr><td>Overall Objectives</td><td colspan="3">Attempting to appropriately specify models for non i.i.d. and non Gaussian nature of financial time series, study dynamics for factors to predict returns to assets. With the factors, identify models for data generation process for predicting returns to assets in intuitively parsimonious simplicity. Investing into the assets. achieve optimal dynamic investments over multiple period horizon as a core research in this thesis. As research outcomes, contribute to practitioners by proposing intuitive and implementable investment decision algorithm.</td></tr>
<tr><td>Uncovered research questions in literature</td><td>■ Regime dependent optimal portfolios for a reasonably practical problem in terms of problem size.
■ Reasonable basis why multiple regime models improve investment efficacy.</td><td>Theoretical research to derive tractable solutions for multi-period optimal investment strategy with regime dependent data generation process.</td><td>Optimal portfolios and performance studies under investment constraints, e.g., no short sales and no leverage, over multi-period horizons with regime dependent data generation processes</td></tr>
<tr><td>Chapters</td><td>Chapter 3</td><td>Chapter 4</td><td>Chapter 5</td></tr>
<tr><td>Objectives of chapter</td><td>Empirical studies of myopic mean-variance optimal portfolio after transaction costs to understand why multiple regime model improve investment efficacy.</td><td>Analytical studies of multi-period mean-variance optimal portfolio after transaction cost to achieve analytical solutions to derive under no constraint in regime switching space.</td><td>Numerical studies to formulate the problem in Chapter 4 to solve under a short sales constraint to evaluate the linear rebalancing rules</td></tr>
<tr><td>Empirical and Numerical Study</td><td>US equity sectors rotation strategy across 12 sectors with PAST(2,12) return forecasting factors</td><td>US equity style rotation strategy across Size and Book-to-Market 2x3 portfolios with SMB and HML as return forecasting factors</td><td>Numerical experiments for a model estimated in Chapter 4</td></tr>
<tr><td>Earlier Studies</td><td>■ Moskowitz and Grinblatt (1999)
■ Daniel and Moskowitz (2013)</td><td>■ Garleanu and Pedersen (2013)</td><td>■ Moallemi and Saglam (2013)</td></tr>
<tr><td>Contributions</td><td colspan="3">■ Two popular investment strategies, i.e., equity sector rotation in Chapter 3 and style rotation in Chapter 4, reveal that
• proposing estimated return forecasting models are regime dependent.
• optimal portfolios achieved by derived solutions for both of myopic and multi-period horizons perform better than in multiple regimes than in single regime regardless of in-sample and out-of-sample periods.
■ Extending the earlier study of dynamic investment strategy to a regime dependent world, Chapter 4 proposes a semi-analytical tractable solution derived under no constraint imposed over infinite time horizon in discrete time space for a return forecasting model with factors following VAR(1) factors under the mean-variance objective function after transaction costs. Parameters in data generating processes and the objective function are regime dependent.
■ Extending the earlier study of the linear rebalancing rules to a regime dependent world, Chapter 5 formulates the model to solve for optimal portfolios over a finite time horizon under short sales constraint. Chapter 5 presents decent investment performance with reasonable basis under a practical computational resource.

The thesis contributes to literature and investment managers to formally incorporate discontinuous changes in market behavior into investment decisions and manage portfolios for ranging from hedge funds and long only investments including asset allocation with forecasting factors.</td></tr>
</table>

Table 6.1: Overall review of the thesis

Bibliography

[1] A. Adler and M. Kritzman, (2007), "Mean-variance Versus Full-Scale Optimisation: In and Out of Sample," *Journal of Asset Management*, 7(5), 302–311.

[2] T. Adrian and F. Franzoni, (2009), "Learning about beta: Time-varying factor loadings, expected returns, and the conditional CAPM," *Journal of Empirical Finance*, 16, 537–556.

[3] L. Akdeniz, A. A. Salih and M. Caner, (2003), "Time-Varying Betas Help in Asset Pricing: The Threshold CAPM," *Studies in Nonlinear Dynamics and Econometrics*, 6(4), 1–17.

[4] M. Ammann and M. Verhofen, (2006), "The Effect of Market Regimes on Style Allocation," *Financial Markets and Portfolio Management*, 20(3), 309–337.

[5] A. Ang and G. Bekaert, (2002a), "International Asset Allocation with Regime Shifts," *The Review of Financial Studies*, 15(4), 1137–1187.

[6] A. Ang and G. Bekaert, (2002b), "Regime Switches in Interest Rates," *Journal of Business & Economic Statistics*, 20(2), 163–182.

[7] A. Ang and G. Bekaert, (2004). "How Do Regime Shifts Affect Asset Allocation?" *Financial Analysts Journal*, 60(2), 86–99.

[8] A. Ang and J. Chen, (2007), "CAPM Over The Long Run: 1926-2001," *Journal of Empirical Finance*, 14, 1–40.

[9] A. Ang and D. Kristensen, (2012), "Testing Conditional Factor Models," *Journal of Financial Economics*, 106(1), 132–156.

[10] B. Arshanapalli, F. J. Fabozzi and W. Nelson, (2006), "The Value, Size, and Momentum Spread During Distressed Economic Periods," *Finance Research Letters*, 3, 244–252.

[11] C. Asness, J. Liew and R. Stevens, (1997)." Parallels Between the Cross-Sectional Predictability of Stock and Country Returns," *The Journal of Portfolio Management*, 23(3), 79–87.

[12] C. Asness, T. Moskowitz and L. Pedersen, (2013). "Value and Momentum Everywhere," *The Journal of Finance*, 68(3), 929–985.

[13] R. Bansal and H. Zhou, (2002), "Term Structure of Interest Rates with Regime Shifts," *The Journal of Finance Research Letters*, 57(5), 1997–2035.

[14] D. Basu and A. Stremme, (2007), "CAPM and Time-Varying Beta: The Cross-Section of Expected Returns," *SSRN Electronic Journal*, DOI: 10.2139/ssrn.972255.

[15] L. E. Baum and T. Petrie, (1966), "Statistical Inference for Probabilistic Functions of Finite State Markov Chains," *The Annals of Mathematical Statistics*, 37(6), 1554–1563.

[16] G. Bekaert and C. R. Harvey, (1995), "Time-Varying World Market Integration," *The Journal of Finance*, 50(2), 403–444.

[17] B.S. Bernanke, (1983), "Irreversibility, Uncertainty, and Cyclical Investment," *The Quarterly Journal of Economics*, 98, 85–106.

[18] R. Bhatia, (1997), *Matrix Analysis*, Springer.

[19] A. Black, P. Fraser and D. Power, (1992), "UK Unit Trust Performance 1980-1989: A Passive Time-Varying Approach," *Journal of Banking and Finance*, 16, 1015–1033.

[20] T. Bollerslev, (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

[21] T. Bollerslev, (1990), "Modelling The Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model," *The Review of Economics & Statistics*, 72(3), 498–505.

[22] T. Bos and P. Newbold, (1984), "An Empirical Investigation of the Possibility of Stochastic Systematic Risk in the Market Model," *The Journal of Business*, 57(1), 35–41.

[23] P. A. Braun, D. B. Nelson and A. M. Sunier, (1995), "Good News, Bad News, Volatility, and Betas," *The Journal of Finance*, 50(5), 1575–1603.

[24] R. D. Brooks, R. W. Faff and M. D. Mckenzie, (1998), "Time-Varying Beta Risk of Australian Industry Portfolios: A Comparison of Modelling Techniques," *Australian Journal of Management*, 23(1), 1–22.

[25] M. M. Carhart, (1997), "On Persistence in Mutual Fund Performance ," *The Journal of Finance*, 52(1), 57–82.

[26] K. Chan, A. Hameed, and A. Tong, (2000), "Profitability of Momentum Strategies in the International Equity Markets," *Journal of Financial and Quantitative Analysis*, 35(2), 153–172.

[27] A. Chekhlov, S. Uryasev and M. Zabarankin, (2003), "Portfolio Optimization with Drawdown Constraints," *Research Report 2000-5*, University of Florida.

[28] C. Chiarella, R. Dieci and X.-Z. He, (2013), "Time-Varying Beta: A Boundedly Rational equilibrium approach," *The Journal of Evolutionary Economics*, 23, 609–639.

[29] L. Chollete, A. Heinen and A. Valdesogo, (2009), "Modelling International Financial Returns with a Multivariate Regime-switching Copula," *Journal of Financial Econometrics*, 7(4), 437–480.

[30] T. Chordia and L. Shivakumar, (2002), "Momentum, Business Cycle, and Time-varying Expected Returns," *The Journal of Finance*, 57(2), 985–1019.

[31] P. Coggi and B. Manescu, (2004), "A Multifactor Model of Stock Returns with Stochastic Regime Switching," *Economics Discussion Paper No. 2004-1*, University of St. Gallen.

[32] L. Cohen and A. Frazzini, (2008), "Economic Links and Predictable Returns," *The Journal of Finance*, 63(4), 1977–2011.

[33] D. W. Collins, J. Ledolter and J. Rayburn, (1987), "Some Further Evidence on the Stochastic Properties of Systematic Risk," *Journal of Business*, 60(3), 425–448.

[34] P. Collin-Dufresne, K. Daniel, C. C. Moallemi and M. Sağlam, (2014), "Strategic Asset Allocation with Predictable Returns and Transaction Costs," *SSRN Electronic Journal*, DOI: 10.2139/ssrn.2618910.

[35] O. L. V. Costa and M. V. Araujo, (2008), "A Generalized Multi-Period Portfolio Optimization with Markov Switching Parameters," *Automatica*, 44(10), 2487–2497.

[36] C. Cox, E. Ingersoll, Jr. and S. A. Ross, (1985), "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53(2), 385–407.

[37] J.-H. Cremers, M. Kritzman and S. Page, (2005), "Optimal Hedge Fund Allocations," *The Journal of Portfolio Mahagement*, 31(3), 70–81.

[38] X. Cui, J. Gao, X. Li and D. Li, (2014), "Optimal Multi-Period Mean-Variance Policy Under No-Shorting Constraint," *European Journal of Operational Research*, 234, 459–468.

[39] Q. Dai and J. Singleton, (2000), "Specification Analysis of Affine Term Structure Models," *The Journal of Finance*, 55(5), 1943–1978.

[40] J. Danielsson and C. G. de Vries, (1998), "Beyond the Sample: Extreme Quantile and Probability Estimation," *Discussion Paper 298*, Financial Markets Group, 1–39.

[41] K. Daniel and T. Moskowitz, (2013), "Momentum Crashes," *SSRN Electronic Journal*, DOI: 10.2139/ssrn.2371227.

[42] V. Dombrovskii and T. Obyedko, (2014), "Dynamic Investment Portfolio Optimization under Constraints in the Financial Markets with Regime Switching using Model Predictive Control," *Working Paper*, Tomsk State University.

[43] C. Engle, (1994), "Can The Markov Switching Model Forecast Exchange Rates?," *Journal of International Economics*, 36, 151–165.

[44] R. F. Engle, (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50(4), 987–1007.

[45] T. Eraker, I. Shaliastovich and I. Wang, (2012), "Durable Goods, Inflation Risk and the Equilibrium Asset Prices," *AFA 2013 San Diego Meetings*.

[46] F. J. Fabozzi and J. C. Francis, (1978), "Beta as a Random Coefficient," *Journal of Financial and Quantitative Analysis*, 13(1), 101–116.

[47] E. F. Fama and K. R. French, (1992), "The Cross-Section of Expected Stock Returns," *The Journal of Finance*, 47(2), 427–465.

[48] E. F. Fama and K. R. French, (1993), "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33, 3–36.

[49] E. F. Fama and K. R. French, (1996), "Multifactor Explanations of Asset Pricing Anomalies," *Journal of Finance*, 51(1), 55–84.

[50] E. F. Fama and K. R. French, (2006), "The Value Premium and the CAPM," *The Journal of Finance*, 61(5), 2163–2185.

[51] C. Francq and J.-M. Zakoaïan, (2001), "Stationarity of Multivariate Markov-Switching ARMA Models," *Journal of Econometrics*, 102, 339–364.

[52] S. Frühwirth-Schnatter, (2006), *Finite Mixture and Markov Switching Models*, Springer.

[53] N. Gârleanu and L. H. Pedersen, (2013), "Dynamic Trading with Predictable Returns and Transaction Costs," *The Journal of Finance*, 68(6), 2309–2340.

[54] W. Gersch and G. Kitagawa, (1982), "A time varying multivariate autoregressive modeling of econometric time series," *Statistical Research Division Report Series, SRD Research Report Number: CENSUS/SRD/RR-82/07*, U.S. Buereau of the Census.

[55] G. González-Rivera, (1997), "The Pricing of Time-Varying Beta," *Empirical Economics*, 22, 345–363

[56] M. Grant and S. Boyd, (2015), "CVX: Matlab Software for Disciplined Convex Programming," Release 2.1, http://cvxr.com/cvx.

[57] T. Grifin, X. Ji and S. Martin, (2003),"Momentum Investing and Business Cycle Risk: Evidence from Pole to Pole," *The Journal of Finance*, 58, 6, 2515–2547.

[58] R. C. Grinold, (1997), "The Information Horizon: How to Measure and Improve in the Time Dimension," *The Journal of Portfolio Management*, 24(1), 57–67.

[59] R. Grinold, (2007), "Dynamic Portfolio Analysis: Establishing a Link between Cause and Effect," *The Journal of Portfolio Management*, 34(1), 12–26.

[60] R. Grinold and R. Kahn, (1999), *Active Portfolio Management*, McGraw-Hill Professional, New York.

[61] M. Guidolin and A. Timmermann, (2004), "Strategic Asset Allocation and Consumption Decisions under Multivariate Regime Switching," *SSRN Electronic Journal*, DOI: 10.2139/ssrn.613461.

[62] M. Guidolin and A. Timmermann, (2006), "Asset Allocation under Multivariate Regime Switching," *Working Paper 2005-002C*, Federal Reserve Bank of St Louis.

[63] M. Guidolin and A. Timmermann, (2008a), "Size and Value Anomalies under regime Shifts," *Journal of Financial Econometrics*, 6, 1–48.

[64] M. Guidolin and A. Timmermann, (2008b), "International Asset Allocation under Regime Switching, Skew, and Kurtosis Preferences," *The Review of Financial Studies*, 21(2), 889–935.

[65] M. Guirguis, M. Theodore and M. Suen, (2012), "Timing the Value Style Index in a Markov Regime-Switching Model," *Journal of Investment Management*, 10, 1, 52–64.

[66] B. Hagströmer and J. M. Binner, (2009), "Stock Portfolio Selection with Full-Scale Optimization and Differential Evolution," *Applied Financail Economics*, 19, 1559–1571.

[67] J. D. Hamilton, (1988), "A Neoclassical Model of Unemployment and the Business Cycle," *Journal of Political Economy*, 96(3), 593–617.

[68] J. D. Hamilton, (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57(2), 357–384.

[69] J. D. Hamilton, (1990), "Analysis of Time Series Subject to Changes in Regime," *Journal of Econometrics*, 45, 39–70.

[70] J. D. Hamilton, (1994), *Time Series Analysis*, Princeton University Press, New Jersey.

[71] H. Hasnaoui and I. Fatnassi, (2014), "Time-Varying Beta And The Subprime Financial Crisis: Evidence From U.S. Industrial Sectors," *The Journal of Applied Business Research*, 30(5), 1465–1476.

[72] P. Huang and C. J. Hueng, (2008), "Conditional Risk-Return Relationship in A Time-Varying Beta Model," *Quantitative Finance*, 8(4), 381–390.

[73] N. Jagadeesh, (1990), "Evidence of Predictable Behavior of Security Returns," *The Journal of Finance*, 45(3), 881–898.

[74] N. Jagadeesh and S. Titman, (1993), "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *The Journal of Finance*, 48(1), 65–91.

[75] R. Jagannathan and Z. Wang, (1996), "The Conditional CAPM and the Cross-Section of Expected Returns," *The Journal of Finance*, 51(1), 3–53.

[76] E. Jondeau and M. Rockinger, (2006), "Optimal Portfolio Allocation under Higher Moments," *European Financial Management*, 12(1), 29–55.

[77] C.-J. Kim and C. R. Nelson, (1999), "Has The U.S. Economy Become More Stable? A bayesian Approach Based on a Markov-Switching Model of the Business Cycle," *The Review of Financial Studies*, 81(4), 608–616.

[78] T. Komatsu and N. Makimoto, (2015), "Dynamic Investment Strategy with Factor Models Under Regime Switches," *Asia-Pacific Financial Markets*, 22, 209–237.

[79] J. Lakonishok, A. Shleifer and R. Vishny, (1994), "Contrarian Investment, Extrapolation, and Risk," *The Journal of Finance*, 49(5), 1541–1578.

[80] K. Lee and S. Ni (2002), "On the Dynamic Effects of Oil Price Shocks : A Study Using Industry Level Data," *Journal of Monetary Economics*, 49, 823–852.

[81] B. Lehmann, (1990), "Fads, Martingale, and Market Efficiency," *Quarterly Journal of Economics*, 105, 1–28.

[82] M. Lettau and S. Ludvigson, (2001), "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying," *Journal of Political Economy*, 109(6), 3–53.

[83] J. Lewellen and S. Nagel, (2006), "The Conditional CAPM Does Not Explain Asset-Pricing Anomalies," *Journal of Financial Economics*, 82, 289–314.

[84] P. Liu, K. Xu and Y. Zhao, (2010), "Market Regimes. Sectorial Investments, and Time-Varying Risk Premiums," *International Journal of Managerial Finance*, 7(2), 107-133.

[85] X. Li, X. Y. Zhou and A. E. B. Lim, (2002), "Dynamic Mean-Variance Portfolio Selection with No-Shorting Constraints," *SIAM Journal on Control and Optimization*, 40(5), 1540–1555.

[86] H. Liu, (2004), "Optimal Consumption and Investment with Transaction Costs and Multiple Risky Assets," *The Journal of Finance*, 59(1), 289–338.

[87] F. Longin and B. Solnik, (2001), "Extreme Correlation of International Equity Markets," *The Journal of Finance*, 56(2), 649–676.

[88] F. M. Longin, (1996), "The Asymptotic Distribution of Extreme Stock Markets Returns," *Journal of Business*, 69(3), 383–408.

[89] Y. Ma, L. MacLean, K. Xu and Y. Zhao, (2011), "A Portfolio Optimization Model with Regime-Switching Risk Factors for Sector Exchange Traded Funds," *Pacific Journal of Optimization*, 7(2), 281–296.

[90] D. B. Madan and J.-Y. Yen, (2008), "Asset Allocation with Multivariate Non-Gaussian Returns," *Handbooks in Operations Research and Management Science*, 15, 949–969.

[91] B. Mandelbrot, (1963), "The Variation of Certain Speculative prices," *The Journal of Business*, 36(4), 394–419.

[92] H. Markowitz, (1952), "Portfolio Selection," *The Journal of Finance*, 7(1), 77–91.

[93] K. McClain, H. B. Humphreys and A. Boscan, (1996), "Measuring Risk in the Mining Sector with ARCH Models with Important Observations on Sample Size," *Journal of Empirical Finance*, 3, 369–391.

[94] L. Menzly and O. Ozbas, (2006), "Cross-Industry Momentum," *Working Paper Series*, AFA 2005 Philadelphia Meetings, American Finance Association.

[95] R. Merton, (1971), "Optimal Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 3, 373–413.

[96] C. Moallemi, and M. Sağlam, (2013), "Dynamic Portfolio Choice with Linear Rebalancing Rules,"*SSRN Electronic Journal*, DOI: 10.2139/ssrn.2011605.

[97] T. Moskowitz and M. Grinblatt, (1999), "Do Industries Explain Momentum?" *The Journal of Finance*, 54(4), 1249–1290.

[98] C. Nelson and A. F. Siegel, (1987), "Parsimonious Modeling of Yield Curves," *Journal of Business*, 60(4), 473–489.

[99] B. Nieto, S. Orbe and A. Zarraga, (2014), "Time-Varying Market Beta: Does the Estimation Methodology Matter?," *SORT*, 38(1), 13–42.

[100] A. J. Patton, (2004), "On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation," *Journal of Financial Econometrics*, 2(1), 130–168.

[101] G. N. Pettengill, S. Sundaram and I. Mathur, (1995), "On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation," *Journal of Financial and Quantitative Analysis*, 30(1), 101–116.

[102] S.-H. Poon, M. Rockinger and J. Tawn, (2003), "Modelling Extreme-Value Dependence in International Stock Markets," *Statistica Sinica*, 13, 929–953.

[103] K. Price, R. M. Storn and J. A. Lampinen, (2005), *Differential Evolution: A Practical Approach to Global Optimization (Natural Computing Series)*, XIX, Springer, New Jersey.

[104] J. Reeves, and H. Wu, (2010), "Constant vs. Time-Varying Beta Models: Further Forecast Evaluation," *23rd Australasian Finance and Banking Conference 2010 Paper*.

[105] R. Rockafellar, and S. Uryasev, (2000), "Optimization of Conditional Value-at-Risk," *Journal of Risk*, 2, 21–41.

[106] R. T. Rockafellar and S. Uryasev, (2002), "Conditional Value-at-Risk for General Loss Distributions," *Journal of Banking & Finance*, 26, 1443–1471.

[107] S. A. Ross, (1976), "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341–360.

[108] K. Rouwenhorst, (1998), "International Momentum Strategies," *The Journal of Finance*, 53(1), 267–284.

[109] H. Schaller and S. van Norden, (1997), "Regime Switching in Stock Market Returns," *Applied Financail Economics*, 7, 177–191.

[110] G. Schwert and P. J. Seguin, (1990), "Heteroskedasticity in Stock Returns," *The Journal of Finance*, 45(4), 1129–1155.

[111] I. Seidl, (2012), "Markowitz Versus Regime Switching: An Empirical Approach," *The Review of Finance and Banking*, 4(1), 33–43.

[112] W. F. Sharpe, (1964), "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *The Journal of Finance*, 19(3), 425–442.

[113] A. Z. Sheikh and H. Qiao, (2010), "Non-Normality of Market Returns: A Framework for Asset Allocation Decision Making," *The Journal of Alternative Investments*, 12(3), 8–35.

[114] L. Sneddon, (2008), "The Tortoise and the Hare: Portfolio Dynamics for Active Managers," *The Journal of Investing*, 17(4), 106–111.

[115] L. R. Sotomayor and A. Cadenillas, (2009), "Explicit Solutions of Consumption-Investment Problems in Financial Markets with Regime-Switching," *Mathematical Finance*, 19(2), 251–279.

[116] R. Storn and K. Price, (1997), "Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," *Journal of Global Optimization*, 11, 341–359.

[117] A. Timmerman, (2000), "Moments of Markov Switching Models," *Journal of Econometrics*, 96, 75–111.

[118] D. Titterington, A. Smith and U. Makov, (1985), *Statistical analysis of finite mixture distributions*, Wiley, New York.

[119] C. Trecroci, (2014), "How Do Alphas and Betas Move? Uncertainty, Learning and Time Variation in Risk Loadings," *Oxford Bulletin of Economics and Statistics*, 76(2), 257–278.

[120] J. Tu, (2010), "Is Regime Switching in Stock Returns Important in Portfolio Decisions?," *Management Science*, 56(7), 1198–1215.

[121] A. Valdesogo, L. Chollete and A. Heinen, (2009), "Modeling International Financial Returns with a Multivariate Regime-switching Copula," *Journal of Financial Econometrics*, 7, 437–480.

[122] C. Wells, (1994), "Variable Betas on the Stockholm Exchange 1971-1989," *Applied Economics*, 4, 75–92.