UNIVERSITY COLLEGE LONDON

DOCTORAL THESIS

Functionally Generated Portfolios in Stochastic Portfolio Theory

Author:
Kangjianan XIE

Supervisor:
Prof. Johannes RUF

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

Department of Mathematics

February 7, 2020

Declaration of Authorship

I, Kangjianan XIE, declare that this thesis titled, “Functionally Generated Portfolios in Stochastic Portfolio Theory” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

UNIVERSITY COLLEGE LONDON

Abstract

Department of Mathematics

Doctor of Philosophy

Functionally Generated Portfolios in Stochastic Portfolio Theory

by Kangjianan XIE

In this dissertation, we focus on constructing trading strategies through the method of functional generation. Such a construction is of great importance in Stochastic Portfolio Theory established by Robert Fernholz. This method is simplified by Karatzas and Ruf (Finance and Stochastics 21.3:753-787, 2017), where they also propose another method called additive functional generation. Inspired by their work, we first investigate the dependence of functional generation on an extra finite-variation process. A mollification argument and Komlós theorem yield a general class of potential arbitrage strategies. Secondly, we extend the analysis by incorporating transaction costs proportional to the trading volume. The performance of several portfolios in the presence of dividends and transaction costs is examined under different configurations. Next, we analyse the so-called leakage effect used to measure the loss in portfolio wealth due to renewing the portfolio constituents. Moreover, we further explore the method of additive functional generation by considering the conjugate of a portfolio generating function. The connection between functional generation and optimal transport is also studied. An extended abstract can be found before the first chapter of this dissertation.

Impact Statement

Stochastic Portfolio Theory (SPT) has been used in an equity market to construct trading strategies that have the potential to beat the market. The theory is built on sound theoretical fundamentals and is extremely easy to be applied for investing in the market. Our research contributes to the study of SPT as well as the investment procedure in the following.

(i) We study the dependency of portfolio generating functions on some finite-variation processes. This brings great flexibility into portfolio construction from both theoretical and practical aspects. Our methods on analysing generalised portfolio generating functions can be used in future research on a similar topic.

(ii) We show both theoretically and empirically that monotonic changes in market diversification can be used appropriately to enhance the portfolio performance. It would be interesting to further explore the connection and apply it to investment activities in the real market.

(iii) We propose a numerical scheme to incorporate transaction costs and dividends when backtesting systemically generated trading strategies. This scheme can be implemented directly in other empirical research using historical data of stock capitalisations and return indices.

(iv) We approach the study of the so-called leakage effect differently from the previous research. Our method avoids a limitation in the original method used by others in studying this topic.

(v) We improve the understanding of the role played by the conjugate of a portfolio generating function in portfolio construction as well as in optimal transport.

Acknowledgements

First and foremost, I would like to express my greatest gratitude to my supervisor Prof. Johannes Ruf. From the beginning to the end of my PhD research, he has always been kind, patient, and considerate to support me with his high standard expertise and immense knowledge. He is not only an amazing advisor for his insightful suggestions and wisdom, but also he is a true friend for me outside the academic world. His guidance helped me in all the time of research and writing of this thesis. This thesis would not have been possible without his endless support and valuable comments.

Secondly, I would give my sincere gratitude to Robert Fernholz, Camilo García, Ioannis Karatzas, Andrea Macrina, Johannes Muhle-Karbe, Soumik Pal, Vassilios Papathanakos, Teemu Pennanen, Daniel Schwarz, and Leonard Wong for many helpful discussions and comments on the subject matter of this dissertation. My genuine thanks also go to the anonymous referees and Associate Editor of *Applied Mathematical Finance*, and the Associate Editor of *SIAM Journal on Financial Mathematics*. Their insightful and constructive suggestions greatly improved this dissertation. Moreover, I would like to thank Camilo García and Michael Monoyios for their willingness and efforts to be the examiners of my doctoral defense.

Additionally, I would like to acknowledge my dear friends Zhangqi Cao, Alexey Gurov, Hanqiong Hu, Longjie Jia, Yupeng Jiang, Luting Li, Yi Li, Jiaqi Liang, Lida Sun, Qiujun Wang, Chen Wang, Weiguan Wang, Weixin Wang, Jian Xiong, Yeyi Yang, and Yongxin Yang for their support and kindness. The road to a PhD is always tough and frustrating, but they make the scenery along it so impressive and beautiful.

Last but not the least, my deepest thanks go to my family. They are always there for me with their wise counsel and sympathetic ear. They are the reason for me to get through all the difficulties. In particular, I would like to thank my girlfriend Yiwei Guan for her company, support, and love. This dissertation is the best gift for them.

Contents

List of Figures

List of Tables

Extended Abstract

In this dissertation, we analyse the construction of trading strategies that have the potential to outperform the market in the long run. In particular, we focus on trading strategies generated systematically through the method of functional generation, which plays a crucial role in Stochastic Portfolio Theory (SPT). We contribute to SPT both theoretically and empirically as outlined in the following.

This dissertation begins with an introduction to SPT in Chapter 1, where we present the fundamentals of SPT, as well as review the significant contributions to the theory. In the introduction, we first state the basic assumptions postulated for modelling an equity market and introduce several fundamental concepts including trading strategies, portfolio weights, and wealth process, in order to describe the investment process. We then introduce the concept of relative arbitrage, which refers to a trading strategy that outperforms another trading strategy. Generating relative arbitrages is one central target in SPT and hence is of great importance. To this end, we present the original methodology of generating relative arbitrages for stocks indexed by companies and, additionally, indexed by ranks of company capitalisations. This methodology, known as functional generation, depends on certain properties of so-called portfolio generating functions and are generalised and reanalysed in Chapters 2 and 4. Associated with this method, the so-called master formula is introduced to express the wealth of a trading strategy generated functionally by market observables.

Chapter 2 is based on Ruf and Xie [80] and focuses on generalising portfolio generating functions to provide more flexibility to portfolio construction. Using the supermartingale property of wealth processes corresponding to some portfolio generating functions after an appropriate change of measures, Karatzas and Ruf [51] propose a simple and intuitive structure to simplify the original method of functional generation. This is known as multiplicative functional generation. They also suggest another method of functional generation called additive functional generation. Motivated by their work, we investigate the dependence of portfolio generating functions on an extra finite-variation process. We also introduce a new category of portfolio generating functions through a mollification argument and Komlós theorem. This category of functions yields a general class of potential relative arbitrages under specific conditions for both additive and multiplicative functional generations. The theoretical results of this chapter are then illustrated by two examples of specific portfolio generating functions. We backtest trading strategies generated both additively and multiplicatively from these functions with data of daily market values and return indexes of all component stocks in the S&P 500 index since year 1989. The theoretical results are shown to work empirically according to the backtested performance of these trading strategies. It is also shown that, for additive functional generation, certain choices of the finite-variation process are better than others, provided that the market diversification changes monotonically.

When backtesting the trading strategies in Chapter 2, we assume that the market is frictionless. This means no transaction cost is imposed when trading to rebalance the portfolio. Nevertheless, paying transaction costs is probably the most important reason that invalidates many trading strategies from beating the market practically as they do theoretically. Hence, one should try to avoid such an unrealistic assumption when judging the profitability of a trading strategy in the real market. Therefore, in Chapter 3, we examine the effects of imposing transaction costs on trading strategies studied in Chapter 2. However, the method used to incorporate transaction costs also works for general systematically generated trading strategies.

Chapter 3 is based on Ruf and Xie [81]. We start with a literature review on transaction costs in equity trading from both theoretical and empirical aspects. Next, we propose a method to incorporate transaction costs proportional to the trading volume in a manner such that, when trading to rebalance a portfolio, the target portfolio weights are matched after paying transaction costs. This method is then applied in a scheme to backtest trading strategies using total market capitalisation and daily stock return time series. Some practical considerations are also emphasised regarding the computation and reinvestment of dividends. The data used for backtesting covers ordinary common stocks traded on all major US exchanges from year 1962 to year 2016. Four portfolios generated multiplicatively by corresponding specific portfolio generating functions are backtested under different configurations involving trading frequency, constituent list size, and renewing frequency. These portfolios are the index tracking portfolio, the equally-weighted portfolio, the entropy-weighted portfolio, and the diversity-weighted portfolio. In particular, the index tracking portfolio has portfolio weights given by the market weights but is different from the corresponding capitalisation index. Indeed, in contrast to the capitalisation index, it reinvests dividends, and is therefore used as benchmark. The empirical results show that, in the absence of transaction costs, all portfolios outperform the index tracking portfolio. This is consistent with the results in Chapter 2. However, when proportional transaction costs of 0.5% are imposed, this outperformance no longer exists for most portfolios. Some exceptional cases include the entropy-weighted and the diversity-weighted portfolios under specific configurations. Other results regarding the effects of changing configurations on portfolio performance are also shown in Chapter 3 with details.

The trading strategies backtested in Chapter 3 invest in a certain number of largest stocks in terms of market capitalisation each time we renew the constituent list. As the component stocks change all the time, it is interesting to study the effect of replacement between stocks on the portfolio performance. This motivates us to study the so-called leakage effect. To this end, in Chapter 4, we first derive the master formulas for trading strategies generated multiplicatively and additively from portfolio generating functions of stocks ranked by capitalisation, respectively. This is done by applying the method of functional generation for such functions, as introduced in Chapter 2. As a consequence, the leakage for a functionally generated trading strategy is defined directly through a term in the corresponding master formula. The leakage measures the loss in the wealth due to untimely renewing the constituent list. Our computation of the leakage is different from what previous research has suggested. The method to estimate leakage in discrete time is then introduced with some practical considerations. Two empirical examples are provided at the end of Chapter 4 to estimate the leakage of the corresponding trading strategies under different constituent list sizes with the same data used in Chapter 3.

In the last chapter of our work, we use the conjugate of a portfolio generating function to further analyse the method of additive functional generation in a discrete-time and model-free setup. Specifically, in the first half of Chapter 5, we recall the conjugate of a concave function and illustrate the method of computing it with two examples. This conjugate is used to define the concept of intrinsic value of a trading strategy, which measures the profitability of the trading strategy in the long run. Then it is shown that the method of additive functional generation can be used to generate optimal trading strategies in that they have nonnegative intrinsic values even in the worst scenario. From this point of view, an additively generated trading strategy can be attractive since it is expected to be profitable in the long run. In the second half of Chapter 5, we consider the link between the method of functional generation and optimal transport, a mathematical area that has drawn much attention recently. Techniques of duality are widely applied to study this connection in previous research as reviewed in this chapter. An equivalence between an additively generated trading strategy and a specific optimal transport problem is established by Vervuurt [88]. Based on this equivalence, we also propose an alternative approach to solve the same optimal transport problem by using the duality between the corresponding portfolio generating function and its conjugate.

Our work is presented for readers with basic knowledge of probability theory and stochastic analysis, and no prior knowledge of SPT is required.

Chapter 1

Introduction to Stochastic Portfolio Theory

Stochastic Portfolio Theory (SPT), which was established by Robert Fernholz, is used as a theoretical tool for applications in equity markets. It is also for analysing portfolios with controlled behaviour under very general conditions, most of which are consistent with observed features of the real market. Early papers of SPT include Fernholz [28, 30, 31]; see Fernholz [26] for details and Fernholz and Karatzas [35] for a survey of SPT. One essential topic in SPT is to invest in an equity market with portfolios constructed systematically from some functions. These functions are known as *portfolio generating functions* and depend merely on current observables: the market capitalisation of each stock in the market. When generated from specific portfolio generating functions, these portfolios, known as *functionally generated portfolios*, can be made good use of in certain types of markets. In particular, over sufficiently large investment horizons, these portfolios will theoretically outperform the corresponding capitalisation-weighted index with probability one. It is also remarkably easy to implement these portfolios, as there is no stochastic integration or drift involved in computing the portfolio wealth, and hence the need for estimation is reduced.

1.1 Modelling the market

Based on classical portfolio theory introduced by Markowitz [60], the modern portfolio theory of dynamic asset pricing is widely accepted to analyse the market structure and used to direct investments in financial markets. The theory of dynamic asset pricing is derived from the general equilibrium model for financial markets by Arrow [4]. It is further developed on the capital asset pricing model by Sharpe [83] and the portfolio optimisation problem by Merton [63]. The theory postulates strong assumptions on the market structure. It relies on the existence of equivalent martingale measure(s) and requires a market in equilibrium and free of arbitrage.

Although also descended from classical portfolio theory, SPT is in contrast to the theory of dynamic asset pricing in that it is valid even without an equivalent martingale measure and in the presence of arbitrage or market disequilibrium. In particular, most models in SPT only require the boundedness in probability of the terminal values of wealth processes, which is referred to as the weaker "No Unbounded Profit with Bounded Risk" condition; see Karatzas and Kardaras [49]. Kardaras [53] shows that this condition holds if and only if there exists an equivalent local martingale deflator, a strictly positive process such that all discounted nonnegative wealth processes become local martingales after multiplying it. Ruf and Runggaldier [79] provide a systematic construction of market models, where arbitrage opportunities exist for bounded profits. We refer to Fontana [38] and Vervuurt [88] for an overview of study on model construction under this weaker no-arbitrage condition.

Throughout our work, we face an equity market with $d \geq 2$ companies, where each company has always one share of stock outstanding. The vector-valued non-negative market capitalisation process of these stocks is denoted by $S(\cdot) = (S_1(\cdot), \cdots, S_d(\cdot))'$ with $S_1(0) > 0, \cdots, S_d(0) > 0$. At $t > 0$, it is allowed to have $S_i(t) = 0$, for some but not all $i \in \{1, \cdots, d\}$. In the rest of this chapter and Chapters 2, 4, and 5, we fix a filtered probability space $(\Omega, \mathcal{F}(\infty), \mathcal{F}(\cdot), \mathsf{P})$, where $\mathcal{F}(\cdot)$ is a right-continuous filtration with $\mathcal{F}(0) = \{\emptyset, \Omega\}$ and P is the physical probability measure. For all $i \in \{1, \cdots, d\}$, we assume that $S_i(\cdot)$ is a continuous, non-negative semimartingale. An Itô market model is usually adopted to model the dynamic of $S(\cdot)$ but is unnecessary to be presented here; see Section 1.1 in Fernholz [26]. In Chapter 3, we refrain from imposing any market model and only regard $S(\cdot)$ as non-negative process under discrete time.

1.2 Trading strategy

To define trading strategies, consider a vector-valued process $\vartheta(\cdot) = (\vartheta_1(\cdot), \cdots, \vartheta_d(\cdot))'$ in $\mathbb{R}^d$, which is predictable and integrable with respect to $S(\cdot)$. We denote the collection of all such processes by $\mathcal{L}(S)$. For such a process $\vartheta(\cdot) \in \mathcal{L}(S)$, we interpret $\vartheta_i(t)$ as the number of shares in the stock of company i held at time $t \geq 0$, for all $i \in \{1, \cdots, d\}$. Then

$$V^{\vartheta}(\cdot; S) = \sum_{j=1}^{d} \vartheta_j(\cdot) S_j(\cdot) \tag{1.1}$$

can be interpreted as the wealth process corresponding to $\vartheta(\cdot)$ in money amount.

Definition 1.2.1. (Trading strategy). A process $\varphi(\cdot) \in \mathcal{L}(S)$ is called a *trading strategy* for $S(\cdot)$ if

$$V^{\varphi}(\cdot; S) - V^{\varphi}(0; S) = \int_0^{\cdot} \sum_{j=1}^{d} \varphi_j(t) \mathrm{d}S_j(t). \tag{1.2}$$

For a trading strategy $\varphi(\cdot)$, its corresponding portfolio weights are determined as the following. Let us write

$$\bar{\Delta}^d = \left\{ (x_1, \cdots, x_d)' \in \mathbb{R}^d : \sum_{j=1}^{d} x_j = 1 \right\}$$

and

$$\Delta^d = \left\{ (x_1, \cdots, x_d)' \in [0,1]^d : \sum_{j=1}^{d} x_j = 1 \right\} \quad \text{and} \quad \Delta^d_+ = \Delta^d \cap (0,1)^d. \tag{1.3}$$

Definition 1.2.2. (Portfolio weights). Given a trading strategy $\varphi(\cdot)$, its *portfolio weight process* is a measurable, adapted $\bar{\Delta}^d$-valued process $\pi(\cdot) = (\pi_1(\cdot), \cdots, \pi_d(\cdot))'$ with

$$\pi_i(t) = \frac{\varphi_i(t) S_i(t)}{V^{\varphi}(t; S)}, \quad i \in \{1, \cdots, d\},\ t \geq 0. \tag{1.4}$$

If $\pi_i(t) \geq 0$, for all $i \in \{1, \cdots, d\}$ and $t \geq 0$, i.e., if $\pi(\cdot)$ is Δ^d-valued, then the trading strategy $\varphi(\cdot)$ is called a *long-only* trading strategy.

As suggested by (1.1) and (1.4), we have

$$\sum_{j=1}^{d} \pi_j(t) = 1, \quad t \geq 0.$$

Hence, the component process $\pi_i(\cdot)$ represents the proportion of wealth $V^{\varphi}(\cdot; S)$ invested in stock i, for all $i \in \{1, \cdots, d\}$. We shall only consider long-only trading strategies in this work. The world “long-only” is omitted in the following for the sake of simplicity.

Note that (1.2) implies that the wealth process of a trading strategy is self-financing. Then for a trading strategy $\varphi(\cdot)$, the wealth process $V^{\varphi}(\cdot; S)$ should evolve as

$$\mathrm{d}V^{\varphi}(t; S) = \sum_{j=1}^{d} \varphi_j(t) \mathrm{d}S_j(t) = \sum_{j=1}^{d} \frac{\pi_j(t) V^{\varphi}(t; S)}{S_j(t)} \mathrm{d}S_j(t) \tag{1.5}$$

with chosen initial wealth $V^{\varphi}(0; S) = V_0$.

Market trading strategy

In Chapter 3, the wealth of a trading strategy is computed in terms of money amount equivalent to (1.1). However, more frequently in SPT, as well as in Chapters 2, 4, and 5, we are interested in the performance of our trading strategies relative to the performance of some benchmark trading strategy. Most of the time, the *market trading strategy* defined below is chosen to be such a benchmark trading strategy.

Definition 1.2.3. (Market trading strategy). Starting with initial wealth $\sum_{j=1}^{d} S_j(0)$, the *market trading strategy* for $S(\cdot)$ is given by the constant vector-valued process $\varphi(\cdot) = (1, \cdots, 1)'$. The portfolio weight process of the market trading strategy, denoted by $\mu(\cdot) = (\mu_1(\cdot), \cdots, \mu_d(\cdot))'$, is called the *market weight process* with $\mu(0) \in \Delta^d_+$ and *market weights*

$$\mu_i(t) = \frac{S_i(t)}{\sum_{j=1}^{d} S_j(t)}, \quad i \in \{1, \cdots, d\},\ t \geq 0. \tag{1.6}$$

By the assumptions on the market capitalisation process $S(\cdot)$, $\mu_i(\cdot)$ is a continuous, non-negative semimartingale, for all $i \in \{1, \cdots, d\}$. The wealth process of the market trading strategy is given by the process of the total market capitalisation

$$V^{(1,\cdots,1)}(\cdot; S) = \sum_{j=1}^{d} S_j(\cdot).$$

Hence, $\varphi(\cdot) = (1, \cdots, 1)'$ is called the market trading strategy in the sense that the market is owned when implementing it.

Moreover, (1.6) suggests that we could interpret $\mu_i(\cdot)$ as the capitalisation of company i when the total market capitalisation is taken as the numéraire. In this case, given a trading strategy $\varphi(\cdot)$ for $S(\cdot)$, the wealth of $\varphi(\cdot)$ relative to the market is given by the *relative wealth process*

$$V^{\varphi}(\cdot; \mu) = \frac{V^{\varphi}(\cdot; S)}{V^{(1,\cdots,1)}(\cdot; S)} = \sum_{j=1}^{d} \varphi_j(\cdot)\mu_j(\cdot). \tag{1.7}$$

In particular, a trading strategy $\varphi(\cdot)$ for $S(\cdot)$ is also a trading strategy for $\mu(\cdot)$ in that $\varphi(\cdot)$ is predictable and integrable with respect to $\mu(\cdot)$, i.e., $\varphi(\cdot) \in \mathcal{L}(\mu)$, and

$$V^{\varphi}(\cdot; \mu) - V^{\varphi}(0; \mu) = \int_0^{\cdot} \sum_{j=1}^{d} \varphi_j(t) \mathrm{d}\mu_j(t). \tag{1.8}$$

The equation above results from the fact that self-financing portfolios remain self-financing after a numéraire change; see Proposition 1 in Geman, El Karoui, and Rochet [40].

In the remaining part of this work except Chapter 3, we simply use $V^{\varphi}(\cdot)$ to denote the relative wealth of a trading strategy $\varphi(\cdot)$ for $\mu(\cdot)$. Similar to (1.5), the dynamics of $V^{\varphi}(\cdot)$ is given by

$$\mathrm{d}V^{\varphi}(t) = \sum_{j=1}^{d} \varphi_j(t)\mathrm{d}\mu_j(t) = \sum_{j=1}^{d} \frac{\pi_j(t)V^{\varphi}(t)}{\mu_j(t)} \mathrm{d}\mu_j(t), \quad V^{\varphi}(0) = 1, \tag{1.9}$$

where $\pi(\cdot) = (\pi_1(\cdot), \cdots, \pi_d(\cdot))'$ is the process of portfolio weights for the trading strategy $\varphi(\cdot)$.

Remark 1. To convert a predictable process $\vartheta(\cdot) \in \mathcal{L}(\mu)$ into a trading strategy $\varphi(\cdot)$, we adopt the measure of the "defect of self-financeability" of $\vartheta(\cdot)$, introduced in Section 2 in Karatzas and Ruf [51] and defined as

$$Q^{\vartheta}(\cdot) = V^{\vartheta}(\cdot) - V^{\vartheta}(0) - \int_0^{\cdot} \sum_{j=1}^{d} \vartheta_j(t) \mathrm{d}\mu_j(t). \tag{1.10}$$

As a result, the vector-valued process $\varphi(\cdot) = (\varphi_1(\cdot), \cdots, \varphi_d(\cdot))'$ with components

$$\varphi_i(\cdot) = \vartheta_i(\cdot) - Q^{\vartheta}(\cdot) + C, \quad i \in \{1, \cdots, d\}, \tag{1.11}$$

where C can be any real constant, is a trading strategy for $\mu(\cdot)$. □

1.3 Relative arbitrage

As mentioned in Section 1.1, arbitrage opportunities in the market are allowed in SPT. As a matter of fact, constructing trading strategies in a systematic way to explore these arbitrage opportunities has always been a central topic in SPT and is still actively studied. To be more specific, starting with the same initial wealth as the market trading strategy, we are interested in trading strategies that can outperform the market trading strategy over appropriate investment horizons with probability one. A trading strategy that has this property is called a *relative arbitrage* as formally defined below.

Definition 1.3.1. (Relative arbitrage). A trading strategy $\varphi(\cdot)$ is said to be a *relative arbitrage* with respect to the market over a given investment horizon $[0, T]$, for $T \geq 0$, if

$$V^{\varphi}(\cdot) \geq 0 \quad \text{and} \quad V^{\varphi}(0) = 1,$$

along with

$$\mathsf{P}\big[V^{\varphi}(T) \geq 1\big] = 1 \quad \text{and} \quad \mathsf{P}\big[V^{\varphi}(T) > 1\big] > 0. \tag{1.12}$$

If $\mathsf{P}\big[V^{\varphi}(T) > 1\big] = 1$ holds, we say that the relative arbitrage is strong over $[0, T]$.

Remark 2. Definition 1.3.1 makes sense due to the fact that the relative wealth process of the market trading strategy at any time is given by

$$V^{(1,\cdots,1)}(\cdot) = \sum_{i=1}^{d} \mu_i(\cdot) = 1.$$

Then a relative arbitrage exists over a given investment horizon $[0, T]$ when a non-negative relative wealth process $V^{\varphi}(\cdot)$ has the same initial wealth as the market trading strategy, the probability for $V^{\varphi}(T)$ to be greater than the wealth of the market trading strategy is strictly positive, and $V^{\varphi}(T)$ is not lower than the wealth of the market trading strategy. □

Fernholz [28] discusses conditions for arbitrage to exist in equity markets, which leads to the concept of relative arbitrage in the following research; also see Fernholz [31]. In a market where no single company dominates the entire market in terms of relative capitalisation on average over a period, the existence of a relative arbitrage is shown in Section 3.3 in Fernholz [26] and by Fernholz, Karatzas, and Kardaras [36]. Fernholz and Karatzas [34] analyse the existence of a relative arbitrage in a specific model of an abstract volatility-stabilized market. Such a market assigns constant drift and volatility terms to the return of the market trading strategy and the largest volatilities to the smallest stocks. In a Markovian model for equity market and using non-anticipative investment strategies, Fernholz and Karatzas [24] compute the smallest initial wealth that one should invest in order to achieve a relative arbitrage over a given investment horizon. Their results draw attention on the problem of optimising relative arbitrage, which is further studied in Fernholz and Karatzas [25], Bayraktar, Huang, and Song [9], and Ruf [77]. Conditions that connect relative arbitrage to the market completeness are given in Theorem 8 in Ruf [78]. Pal and Wong [72] introduce a pathwise approach to decompose the performance of a trading strategy relative to the market trading strategy. The decomposition consists of a volatility term and two entropy terms, such that a specific class of trading strategies can be constructed to guarantee the existence of a relative arbitrage; see Wong [94] and Pal and Wong [74] for extended results. Karatzas and Ruf [51] formulate conditions of the existence of a relative arbitrage over sufficiently large investment horizons under a more generalised and simplified framework of Fernholz [31]; see Chapter 2 for details. Pal [71] provides sufficient conditions for a market with adequate stocks to contain a short term relative arbitrage. Relative arbitrage over arbitrary time horizons under appropriate conditions is studied by Fernholz, Karatzas, and Ruf [27].

1.4 Functional generation

To explore arbitrage opportunities relative to the market, Fernholz [31] comes up with a powerful tool, called *functional generation*, to construct *functionally generated trading strategies* from appropriate *portfolio generating functions*; see Chapter 3 in Fernholz [26]. The wealth of a functionally generated trading strategy relative to the total market capitalisation is merely a function, known as so-called *master formula*, of the market weights. This formula does not involve stochastic integration or drifts, which makes the analysis very easy as the need for estimation is reduced. Karatzas and Ruf [51] interpret portfolio generating functions as Lyapunov functions. More precisely, the supermartingale property of the corresponding wealth processes after an appropriate change of measure is utilised to study the performance of functionally generated trading strategies. We extend these results to a group of portfolio generating functions with more generalised properties in Chapter 2.

In the following, we present the definition and the master formula of functionally generated trading strategies given by Fernholz [26]. In particular, we refer to this specific kind of functional generation as *multiplicative function generation*, in contrast to another kind of functional generation called *additive functional generation*, as proposed by Karatzas and Ruf [51].

Definition 1.4.1. (Multiplicative functional generation). For a continuous function $G : \Delta^d \to (0,\infty)$ and a trading strategy $\varphi(\cdot)$, we say that $\varphi(\cdot)$ is *multiplicatively generated* by the *portfolio generating function* G if there exists a measurable process $\Theta(\cdot)$ of finite variation, such that

$$\log V^{\varphi}(t) = \log G(\mu(t)) + \Theta(t), \quad t \geq 0. \tag{1.13}$$

Equation (1.13) is known as the *master formula*, and the process $\Theta(\cdot)$ is called the *drift* process corresponding to G.

Now let us consider an open subset $\mathcal{U}$ of $\mathbb{R}^d$ such that $\Delta^d \subset \mathcal{U}$. Then we can always extend a continuous function defined on Δ^d to a continuous function defined on $\mathcal{U}$.[1] Here and throughout this work, we write $G \in C^2$ if G is continuous and twice differentiable in all variables on its domain. For two semimartingales $X(\cdot)$ and $Y(\cdot)$, we use $[X, Y](\cdot)$ to denote their quadratic covariation process.

Theorem 1.4.1. *(Theorem 3.1.5 in Fernholz [26]). Given an open subset $\mathcal{U}$ of $\mathbb{R}^d$ with $\Delta^d \subset \mathcal{U}$, consider a C^2 function $G : \mathcal{U} \to (0, \infty)$. If*

$$\frac{x_i \partial G(x)/\partial x_i}{G(x)}, \quad i \in \{1, \cdots, d\},$$

is bounded for all $x \in \Delta^d$, then the trading strategy $\varphi(\cdot)$ generated multiplicatively by G has portfolio weights

$$\pi_i(t) = \left(G(\mu(t)) + \frac{\partial G}{\partial x_i}(\mu(t)) - \sum_{j=1}^{d} \mu_j(t) \frac{\partial G}{\partial x_j}(\mu(t)) \right) \frac{\mu_i(t)}{G(\mu(t))}, \tag{1.14}$$

for all $i \in \{1, \cdots, d\}$ and $t \geq 0$. Moreover, the drift process $\Theta(\cdot)$ is given by

$$\Theta(\cdot) = -\frac{1}{2} \int_0^{\cdot} \frac{1}{G(\mu(t))} \sum_{i,j=1}^{d} \frac{\partial^2 G}{\partial x_i \partial x_j}(\mu(t)) \mathrm{d} \left[\mu_i, \mu_j \right](t). \tag{1.15}$$

This theorem shows that given an appropriate portfolio generating function, the corresponding multiplicatively generated trading strategy has portfolio weights given explicitly by (1.14). The portfolio weight for each stock is a deterministic function of the current observable market weights. Therefore, it becomes extremely straightforward to implement this multiplicatively generated trading strategy. Moreover, the covariance structure of the market is connected to the drift process $\Theta(\cdot)$ through the quadratic covariation term in (1.15). Especially, for $t \geq 0$, $\Theta(t)$ can be computed directly from observable quantities by the master formula (1.13). Hence, estimation of the covariance structure is not needed. An example of functional generation is that a function $G : \mathcal{U} \to c$ with constant $c \in (0, \infty)$ generates the market trading strategy multiplicatively.

[1] This extension is guaranteed by the Tietze extension theorem. It is made such that a standard coordinate system in $\mathbb{R}^d$ can be utilised to treat all the d market weights in the same manner, which cannot be done on the $(d-1)$-dimensional space Δ^d.

1.5 Rank-dependent trading strategy

Fernholz [29] generalises the methods of functional generation to a class of portfolio generating functions that identify market weights not by their company index, but by their *ranks* in terms of values. This generalisation leads to the *rank-dependent trading strategies* and provides a mathematical interpretation of the size effect. This effect is an observed phenomenon that stocks with smaller capitalisations tend to have higher returns than stocks with larger capitalisations over long period; see Banz [8]. This generalisation also suggests a correction term on the drift process $\Theta(\cdot)$ when the component stocks in a portfolio change under specific circumstances. We will focus on this later application in Chapter 4. For further research on the rank-dependent trading strategies, we refer to Fernholz, Ichiba, and Karatzas [33] and Karatzas and Ruf [51].

The following definitions and results are necessary for presenting our work in Chapters 2 and 4. For a vector $x = (x_1, \cdots, x_d)' \in \Delta^d$, denote its corresponding ranked vector as $\boldsymbol{x} = (x_{(1)}, \cdots, x_{(d)})'$, where

$$\max_{i\in\{1,\cdots,d\}} x_i = x_{(1)} \geq x_{(2)} \geq \cdots \geq x_{(d-1)} \geq x_{(d)} = \min_{i\in\{1,\cdots,d\}} x_i$$

are the components of x in descending order. Denote further

$$\mathbb{W}^d = \left\{ \left(x_{(1)}, \cdots, x_{(d)}\right)' \in \Delta^d : 1 \geq x_{(1)} \geq x_{(2)} \geq \cdots \geq x_{(d)} \geq 0 \right\}$$

and

$$\mathbb{W}^d_+ = \mathbb{W}^d \cap (0,1)^d \tag{1.16}$$

Then the rank operator $\mathfrak{R} : \Delta^d \to \mathbb{W}^d$ is a mapping such that $\mathfrak{R}(x) = \boldsymbol{x}$. The *ranked market weight process* $\boldsymbol{\mu}(\cdot)$ is given by

$$\boldsymbol{\mu}(\cdot) = \mathfrak{R}(\mu(\cdot)) = \left(\mu_{(1)}(\cdot), \cdots, \mu_{(d)}(\cdot)\right)'. \tag{1.17}$$

The process $\boldsymbol{\mu}(\cdot)$ is a continuous, $\mathbb{W}^d$-valued semimartingale whenever $\mu(\cdot)$ is a continuous, Δ^d-valued semimartingale (see Theorem 2.2 in Banner and Ghomrasni [7]). Moreover, let p_t be a random permutation of $\{1, \cdots, d\}$ that associates the name index of stocks with their ranks at time t, for all $t \geq 0$. To wit, we have

$$\mu_{p_t(k)}(t) = \mu_{(k)}(t), \quad k \in \{1, \cdots, d\},\ t \geq 0. \tag{1.18}$$

In particular, if $\mu_{(k)}(t) = \mu_{(k+1)}(t)$, for some $k \in \{1, \cdots, d-1\}$, then we set $p_t(k) < p_t(k+1)$.

Definition 1.5.1. (Local time). The *local time process* of an $\mathbb{R}$-valued continuous semimartingale Y at the origin is given by

$$\mathfrak{L}_Y(\cdot) = \frac{1}{2}\left(|Y(\cdot)| - |Y(0)| - \int_0^{\cdot} \operatorname{sgn}(Y(t))\mathrm{d}Y(t) \right), \tag{1.19}$$

where $\operatorname{sgn}(y) = 2\mathbf{1}_{y\in(0,\infty)} - 1$.

The local time $\mathfrak{L}_X(t)$ measures the time that $X(\cdot)$ has spent at 0 up to time t. Hence, the process $\mathfrak{L}_X(\cdot)$ is of finite variation. We refer to Karatzas and Shreve [52] for a general study on local times.

Definition 1.5.2. (Pathwise mutually non-degenerate). The market weight processes $\mu_1(\cdot), \cdots, \mu_d(\cdot)$ are *pathwise mutually non-degenerate* if, for all $t \geq 0$,

1. $\{t; \mu_i(t) = \mu_j(t)\}$ has Lebesgue measure zero, for all $i, j \in \{1, \cdots, d\}$ with $i \neq j$, a.s.;
2. $\{t; \mu_i(t) = \mu_j(t) = \mu_k(t)\} = \emptyset$, for all $i, j, k \in \{1, \cdots, d\}$ with $i < j < k$, a.s.

The following theorem extends Theorem 1.4.1 to rank-dependent trading strategies generated multiplicatively. Recall the ranked market weight process $\boldsymbol{\mu}(\cdot)$ from (1.17).

Theorem 1.5.1. *(Theorem 4.2.1 in Fernholz [26]). Let the market weight processes $\mu_1(\cdot), \cdots, \mu_d(\cdot)$ be pathwise mutually non-degenerate and p_t be a random permutation by* (1.18)*. For a given open subset $\mathcal{U}$ of $\mathbb{R}^d$ with $\Delta^d \subset \mathcal{U}$, consider a function $G : U \to \mathbb{R}$. If there exists a C^2 function $\boldsymbol{G} : U \to (0, \infty)$ such that $G(x) = \boldsymbol{G}(\mathfrak{R}(x))$, for all $x \in \mathcal{U}$, and*

$$\frac{x_i \partial \boldsymbol{G}(x)/\partial x_i}{\boldsymbol{G}(x)}, \quad i \in \{1, \cdots, d\},$$

is bounded for all $x \in \Delta^d$, then G generates the trading strategy $\varphi(\cdot)$ multiplicatively with portfolio weights

$$\pi_{p_t(k)}(t) = \left(\boldsymbol{G}(\boldsymbol{\mu}(t)) + \frac{\partial \boldsymbol{G}}{\partial x_k}(\boldsymbol{\mu}(t)) - \sum_{j=1}^{d} \mu_{(j)}(t) \frac{\partial \boldsymbol{G}}{\partial x_j}(\boldsymbol{\mu}(t)) \right) \frac{\mu_k(t)}{\boldsymbol{G}(\boldsymbol{\mu}(t))},$$

for all $k \in \{1, \cdots, d\}$ and $t \geq 0$. Moreover, the drift process $\Theta(\cdot)$ is given by

$$\begin{aligned}\Theta(\cdot) = &-\frac{1}{2} \int_0^{\cdot} \frac{1}{G(\mu(t))} \sum_{i,j=1}^{d} \frac{\partial^2 \boldsymbol{G}}{\partial x_i \partial x_j}(\boldsymbol{\mu}(t)) \mathrm{d}\left[\mu_{(i)}, \mu_{(j)}\right](t) \\ &+ \frac{1}{2} \int_0^{\cdot} \sum_{k=1}^{d-1} \left(\pi_{p_t(k+1)}(t) - \pi_{p_t(k)}(t)\right) \mathrm{d}\mathfrak{L}_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t).\end{aligned} \tag{1.20}$$

Compared with Theorem 1.4.1, Theorem 1.5.1 shows that the drift process $\Theta(\cdot)$ can be decomposed into two components: a smooth component and a local time component, as in (1.20). Theorem 1.5.1 is generalised by Banner and Ghomrasni [7] such that Condition 2 in Definition 1.5.2 is not required anymore.

An example of a rank-dependent trading strategy is that the function $G(x) = x_{(1)}$, for all $x \in \mathcal{U}$, generates the trading strategy $\varphi(\cdot)$ multiplicatively with portfolio weights $\pi_{p_t(k)}(\cdot) = \mathbf{1}_{k=1}$, for all $k \in \{1, \cdots, d\}$. This trading strategy only invests in the largest stock in the market throughout the investment horizon.

In Chapter 4, we restudy the properties of rank-dependent trading strategies via a different approach proposed in Chapter 2. We give details to the method of estimating the local time component with real data for both multiplicative functional generation and *additive functional generation* proposed by Karatzas and Ruf [51] and studied in Chapter 2.

Chapter 2

Generalised Lyapunov Functions and Functionally Generated Trading Strategies

As introduced in Chapter 1, a pivotal topic in SPT is the construction of functionally generated trading strategies that can outperform the market trading strategy under specific circumstances. The wealth of a trading strategy is linked to the corresponding portfolio generating function G through the master formula; see Definition 1.4.1. Karatzas and Ruf [51] extend and simplify the method by using the supermartingale property of G after an appropriate change of measure to interpret G as a Lyapunov function. They also define a new method of functional generation, the additive functional generation, different from the multiplicative functional generation introduced by Fernholz [31]. The framework of Karatzas and Ruf [51] will be used to formulate conditions on trading strategies to be strong arbitrage relative to the market over sufficiently large investment horizons in this chapter.

One offspring of a portfolio generating function is a generalised portfolio generating function, which depends on an additional argument with continuous path and finite variation. This is inspired by the fact that in practice, people tend to take historical data, such as past performance of stocks, or statistical estimates, into consideration when constructing portfolios. Besides, this generalisation provides additional flexibility in choosing portfolio generating functions. Section 3.2 of Fernholz [26] formulates the concept of time-dependent generating functions, and presents the master formula under this situation. In the same framework, Strong [85] shows an extension of the master formula to trading strategies generated by functions that also depend on the current state of some continuous path process of finite variation. Also based on Fernholz's structure, Schied, Speiser, and Voloshchenko [82] provide a pathwise version of the relevant master formula. They also analyze examples where the additional process is chosen to be the moving average of the market weights. In a recent paper, Karatzas and Kim [50] generalize the methodology developed by Karatzas and Ruf [51] in a pathwise, probability-free setting. They also generalize portfolio generating functions with path-dependent functionals.

All the above mentioned papers (Fernholz [26], Strong [85], Schied, Speiser, and

Voloshchenko [82], and Karatzas and Kim [50]) make assumptions on the smoothness of the portfolio generating function with respect to both the finite-variation process and the market weights. In this chapter, we weaken these assumptions such that the choice for the portfolio generating function is less restricted. To this end, we use a mollification argument and the Komlós theorem. Then we study several examples empirically, using data from the S&P 500 index.[1]

This chapter is based on the paper Ruf and Xie [80] that has been accepted for publication. An outline of the chapter is as follows. Section 2.1 first gives the definitions of regular functions and Lyapunov functions, and then presents sufficient conditions for a function to be regular and Lyapunov, respectively. The proofs of these results are presented in Section 2.6. Section 2.2 defines additive and multiplicative generation, and the corresponding trading strategies and wealth processes. Section 2.2 also gives conditions for arbitrage relative to the market portfolio to exist. Section 2.3 describes the data involved and the processing method to implement the empirical analysis. Section 2.4 contains several examples of portfolio generating functions and discusses empirical results. Section 2.5 concludes.

2.1 Generalised regular and Lyapunov functions

In the following of this chapter, we study portfolio generating functions that depend on some $\mathbb{R}^m$-valued continuous process of finite variation on $[0, T]$, for $T \geq 0$ and some $m \in \mathbb{N}$. We use $\Lambda(\cdot)$ to denote such a process. This process allows for more flexibility in selecting portfolio generating functions. Recall the market weight process $\mu(\cdot)$ by Definition 1.2.3 and the ranked market weight process $\boldsymbol{\mu}(\cdot)$ by (1.17). To this end, let $\mathcal{W}$ and $\boldsymbol{W}$ be some open subsets of $\mathbb{R}^m \times \mathbb{R}^d$ such that

$$\mathsf{P}[(\Lambda(t), \mu(t)) \in \mathcal{W},\ \forall\, t \geq 0] = 1 \tag{2.1}$$

and

$$\mathsf{P}[(\Lambda(t), \boldsymbol{\mu}(t)) \in \boldsymbol{W},\ \forall\, t \geq 0] = 1, \tag{2.2}$$

respectively.

Moreover, we introduce several notions that will be used in this chapter and Chapter 4. For a continuous function F, write $F \in C^\infty$ if F is infinitely differentiable. If $F = F(\lambda, x)$, write $F \in C^{0,1}$ if F is differentiable with respect to the second argument and $\partial F / \partial x$ is jointly continuous; write $F \in C^{1,2}$ if F is once differentiable with respect to the first argument, twice differentiable with respect to the second arguments, and

[1] As the constituent list of the stocks in the S&P 500 index changes over time, we avoid a survivorship bias by not restricting the analysis to the current stocks in the S&P 500 index. Instead, we reconstruct the historical constituent list of the S&P 500 index and adjust the portfolios appropriately when the constituent list changes.

$\partial F/\partial\lambda$ and $\partial^2 F/\partial x^2$ are both jointly continuous. In addition, write

$$\|z\|_2 = \left(\sum_{j=1}^{n} z_j^2\right)^{1/2}$$

to denote the L^2 norm of $z = (z_1, \cdots, z_n)' \in \mathbb{R}^n$.

Now let us consider two classes of portfolio generating functions, regular and Lyapunov functions, which are introduced in Karatzas and Ruf [51]. We generalize these notions here to allow for the additional process $\Lambda(\cdot)$. To this end, recall the open set $\mathcal{W}$, in which $(\Lambda(\cdot), \mu(\cdot))$ take values, from (2.1).

Definition 2.1.1. (Generalised regular function). A continuous function $G : \mathcal{W} \to \mathbb{R}$ is said to be *generalised regular* for $\Lambda(\cdot)$ and $\mu(\cdot)$ if

1. there exists a measurable function $G^{\mathrm{D}} = (G_1^{\mathrm{D}}, \cdots, G_d^{\mathrm{D}})' : \mathcal{W} \to \mathbb{R}^d$ such that the process $\vartheta(\cdot) = (\vartheta_1(\cdot), \cdots, \vartheta_d(\cdot))'$ with components

$$\vartheta_i(\cdot) = G_i^{\mathrm{D}}(\Lambda(\cdot), \mu(\cdot)), \quad i \in \{1, \cdots, d\}, \tag{2.3}$$

   is in $\mathcal{L}(\mu)$; and

2. the continuous, adapted process

$$\Gamma^G(\cdot) = G(\Lambda(0), \mu(0)) - G(\Lambda(\cdot), \mu(\cdot)) + \int_0^{\cdot} \sum_{j=1}^{d} \vartheta_j(t)\mathrm{d}\mu_i(t) \tag{2.4}$$

   is of finite variation on the interval $[0, T]$, for all $T \geq 0$.

Definition 2.1.2. (Generalised Lyapunov function). A generalised regular function $G : \mathcal{W} \to \mathbb{R}$ is said to be a *generalised Lyapunov function* for $\Lambda(\cdot)$ and $\mu(\cdot)$ if, for some function G^{D} as in Definition 2.1.1, the finite-variation process $\Gamma^G(\cdot)$ of (2.4) is non-decreasing.

A Lyapunov function turns the semimartingale $\mu(\cdot)$ together with the finite-variation process $\Lambda(\cdot)$ into a supermartingale under a related measure; see Remark 3.4 in Karatzas and Ruf [51] for further explanations. Generally speaking, for a regular (or Lyapunov) function G, the uniqueness of the corresponding measurable function G^{D} is not guaranteed. In the following, we shall omit the terminology “generalised” for simplicity.

In the next example, we discuss sufficient conditions for a smooth function to be regular or Lyapunov.

Example 2.1.1. Consider a $C^{1,2}$ function $G : \mathcal{W} \to \mathbb{R}$. Setting

$$\vartheta_i(\cdot) = \frac{\partial G}{\partial x_i}(\Lambda(\cdot), \mu(\cdot)), \quad i \in \{1, \cdots, d\},$$

and applying Itô's formula yield that G is regular for $\Lambda(\cdot)$ and $\mu(\cdot)$. Indeed, we get the finite-variation process

$$\begin{aligned}\Gamma^G(\cdot) = &- \int_0^\cdot \sum_{v=1}^m \frac{\partial G}{\partial \lambda_v}(\Lambda(t), \mu(t)) \mathrm{d}\Lambda_v(t) \\ &- \frac{1}{2} \sum_{i,j=1}^d \int_0^\cdot \frac{\partial^2 G}{\partial x_i \partial x_j}(\Lambda(t), \mu(t)) \mathrm{d}\left[\mu_i, \mu_j\right](t).\end{aligned} \tag{2.5}$$

Moreover, if the process $\Gamma^G(\cdot)$ is non-decreasing, then G is not only a regular function, but also a Lyapunov function for $\Lambda(\cdot)$ and $\mu(\cdot)$. For instance, this holds if G is non-decreasing in every dimension with respect to the first argument and $\Lambda(\cdot)$ is decreasing in every dimension, and G is concave with respect to the second argument.

□

Below we give sufficient conditions for a function G to be regular (Lyapunov). To this end, recall the open set $\mathcal{W}$ from (2.1).

Theorem 2.1.1. *For a continuous function $G : \mathcal{W} \to \mathbb{R}$, consider the following conditions.*

(ai) On any compact set $\overline{\mathcal{V}} \subset \mathcal{W}$, there exists a constant $L = L(\overline{\mathcal{V}}) \geq 0$ such that, for all $(\lambda_1, x), (\lambda_2, x) \in \overline{\mathcal{V}}$,

$$|G(\lambda_1, x) - G(\lambda_2, x)| \leq L \|\lambda_1 - \lambda_2\|_2.$$

(aii) Function $G(\cdot, x)$ is non-increasing, for fixed x, and $\Lambda(\cdot)$ is non-decreasing in every dimension.

(bi) Function G is differentiable in the second argument and $\partial G / \partial x$ is jointly continuous. Moreover, on any compact set $\overline{\mathcal{V}} \subset \mathcal{W}$, there exists a constant $L = L(\overline{\mathcal{V}}) \geq 0$ such that, for all $(\lambda, x_1), (\lambda, x_2) \in \overline{\mathcal{V}}$,

$$\left\| \frac{\partial G}{\partial x}(\lambda, x_1) - \frac{\partial G}{\partial x}(\lambda, x_2) \right\|_2 \leq L \|x_1 - x_2\|_2.$$

(bii) Function $G(\lambda, \cdot)$ is concave, for fixed λ.

If one of the conditions (ai) or (aii) holds and one of the conditions (bi) or (bii) holds, G is a regular function for $\Lambda(\cdot)$ and $\mu(\cdot)$. Moreover, in the case that (aii) and (bii) hold, G is Lyapunov.

The proof of Theorem 2.1.1 is given in Section 2.6. A generalised version of Itô's formula studied in Krylov [56] is related but can only be applied in a Markovian setting.

Theorem 2.1.1 can be applied to functions not in $C^{1,2}$, such as in Example 2.1.3. Another choice of a non-$C^{1,2}$ function G is the Gini function; see Example 6.1 in Karatzas and Ruf [51] for details.

Remark 3. Consider the special case where $\Lambda(\cdot)$ is set to be a constant λ. Then Theorem 2.1.1 generalises Theorem 3.7(i) and (ii) in Karatzas and Ruf [51]. If $\Lambda(\cdot)$ is non-constant, in contrast to Theorem 3.7 in Karatzas and Ruf [51], even if G can be extended to a continuous function concave in the second argument, G may not be Lyapunov. A counterexample is given in Example 2.1.2. Therefore, for the generalised case, Theorem 3.7 in Karatzas and Ruf [51] cannot be applied, and instead we have to use modified conditions such as given by Theorem 2.1.1. □

Example 2.1.2. Assume that $\mu(\cdot) \in \Delta^d$ with $[\mu_1, \mu_1](t) > 0$, for all $t > 0$, and that

$$\Lambda(\cdot) = \gamma \sum_{j=1}^{d} [\mu_j, \mu_j](\cdot),$$

where γ is a constant.

Define the concave quadratic function

$$G(\lambda, x) = \lambda - \sum_{j=1}^{d} x_j^2, \quad \lambda \in \mathbb{R},\ x \in \Delta^d.$$

Then from (2.5) we have

$$\Gamma^G(\cdot) = -\int_0^{\cdot} \mathrm{d}\Lambda(t) + \sum_{j=1}^{d} \int_0^{\cdot} \mathrm{d}\,[\mu_j, \mu_j](t) = (1-\gamma) \sum_{j=1}^{d} [\mu_j, \mu_j](\cdot).$$

Observe that $\Gamma^G(\cdot)$ is decreasing for $\gamma > 1$; hence G is not a Lyapunov function for $\Lambda(\cdot)$ and $\mu(\cdot)$, although it is concave in its second argument.

Define now $\overline{G}(\lambda, x) = -G(\lambda, x)$. Then we have $\Gamma^{\overline{G}}(\cdot) = -\Gamma^G(\cdot)$. Therefore, if $\gamma > 1$ holds, $\Gamma^{\overline{G}}(\cdot)$ is increasing; hence $\overline{G}$ is Lyapunov although convex in its second argument. □

Recall the ranked market weights process $\boldsymbol{\mu}(\cdot)$ defined by (1.17) and the open set $\boldsymbol{W}$ from (2.2).

Theorem 2.1.2. *If a function $\boldsymbol{G} : \boldsymbol{W} \to \mathbb{R}$ is regular for $\Lambda(\cdot)$ and $\boldsymbol{\mu}(\cdot) = \mathfrak{R}\big(\mu(\cdot)\big)$, then the composition $G = \boldsymbol{G} \circ \mathfrak{R}$ is regular for $\Lambda(\cdot)$ and $\mu(\cdot)$.*

To prove Theorem 2.1.2, we can apply the same techniques used in the proof of Theorem 3.8 in Karatzas and Ruf [51], but now with the generalised form of the function $\boldsymbol{G}$; see Section 2.6 for details.

The following example concerns a function G which is not in $C^{1,2}$. Recall the open set $\mathbb{W}_+^d$ from (1.16).

Example 2.1.3. Assume that $\mu(\cdot) \in \Delta_+^d$ and consider the $C^{1,2}$ function

$$\boldsymbol{G}(\lambda, \boldsymbol{x}) = -\lambda \sum_{l=1}^{d_1} x_{(l)} \log x_{(l)} + 1 - \sum_{l=d_1+1}^{d_2} x_{(l)}^2, \quad \lambda \in \mathbb{R},\ \boldsymbol{x} \in \mathbb{W}_+^d,$$

where d_1 and d_2 are positive integers with $d_1 < d_2 \le d$. According to Example 2.1.1, $\boldsymbol{G}$ is regular for $\Lambda(\cdot)$ and $\boldsymbol{\mu}(\cdot)$. In particular, the corresponding measurable function $\boldsymbol{G}^{\mathrm{D}}$ as in Definition 2.1.1 can be chosen with components

$$\boldsymbol{G}_l^{\mathrm{D}}(\lambda, \boldsymbol{x}) = \begin{cases} -\lambda \log x_{(l)} - \lambda, & \text{if } l \in \{1, \cdots, d_1\} \\ -2x_{(l)}, & \text{if } l \in \{d_1+1, \cdots, d_2\} \\ 0, & \text{otherwise} \end{cases}. \tag{2.6}$$

In this case, Itô's lemma yields

$$\boldsymbol{G}(\Lambda(\cdot), \boldsymbol{\mu}(\cdot)) = \boldsymbol{G}(\Lambda(0), \boldsymbol{\mu}(0)) + \int_0^{\cdot} \sum_{l=1}^{d} \boldsymbol{G}_l^{\mathrm{D}}(\Lambda(t), \boldsymbol{\mu}(t)) \mathrm{d}\mu_{(l)}(t) - \Gamma^{\boldsymbol{G}}(\cdot) \tag{2.7}$$

with $\boldsymbol{G}_l^{\mathrm{D}}$ given in (2.6) and

$$\begin{aligned} \Gamma^{\boldsymbol{G}}(\cdot) = \frac{1}{2} \int_0^{\cdot} \sum_{l=1}^{d_1} \frac{\Lambda(t)}{\mu_{(l)}(t)} \mathrm{d}\left[\mu_{(l)}, \mu_{(l)}\right](t) + \int_0^{\cdot} \sum_{l=d_1+1}^{d_2} \mathrm{d}\left[\mu_{(l)}, \mu_{(l)}\right](t) \\ + \int_0^{\cdot} \sum_{l=1}^{d_1} \mu_{(l)}(t) \log \mu_{(l)}(t) \mathrm{d}\Lambda(t). \end{aligned} \tag{2.8}$$

Denote the number of components of $x = (x_1, \cdots, x_d)' \in \Delta^d$ that coalesce at a given rank $l \in \{1, \cdots, d\}$ by

$$N_l(x) = \sum_{i=1}^{d} \mathbf{1}_{x_i = x_{(l)}}. \tag{2.9}$$

Then by Theorem 2.3 in Banner and Ghomrasni [7], the ranked market weight process $\boldsymbol{\mu}(\cdot)$ has components

$$\begin{aligned} \mu_{(l)}(\cdot) = \mu_{(l)}(0) + \int_0^{\cdot} \sum_{i=1}^{d} \frac{\mathbf{1}_{\{\mu_i(t) = \mu_{(l)}(t)\}}}{N_l(\mu(t))} \mathrm{d}\mu_i(t) + \sum_{k=l+1}^{d} \int_0^{\cdot} \frac{\mathrm{d}\boldsymbol{\Lambda}^{(l,k)}(t)}{N_l(\mu(t))} \\ - \sum_{k=1}^{l-1} \int_0^{\cdot} \frac{\mathrm{d}\boldsymbol{\Lambda}^{(k,l)}(t)}{N_l(\mu(t))}, \quad l \in \{1, \cdots, d\}, \end{aligned} \tag{2.10}$$

where

$$\boldsymbol{\Lambda}^{(i,j)}(\cdot) = \mathfrak{L}_{\mu_{(i)} - \mu_{(j)}}(\cdot), \quad 1 \le i < j \le d, \tag{2.11}$$

is the local time process of the continuous semimartingale $\mu_{(i)}(\cdot) - \mu_{(j)}(\cdot) \ge 0$ at the origin by (1.19).

By Theorem 2.1.2, the function

$$G(\lambda, x) = \boldsymbol{G}(\lambda, \mathfrak{R}(x)) = -\lambda \sum_{l=1}^{d_1} \sum_{j=1}^{d} \frac{\mathbf{1}_{x_j = x_{(l)}}}{N_l(x)} x_j \log x_j + 1 - \sum_{l=d_1+1}^{d_2} \sum_{j=1}^{d} \frac{\mathbf{1}_{x_j = x_{(l)}}}{N_l(x)} x_j^2$$

is regular for $\Lambda(\cdot)$ and $\mu(\cdot)$, since $\boldsymbol{G}$ is regular for $\Lambda(\cdot)$ and $\boldsymbol{\mu}(\cdot)$.

Now, assume that $\Lambda(\cdot)$ is of the form

$$\Lambda(\cdot) = \overline{\xi} \wedge \left(\underline{\xi} \vee \Lambda'(\cdot)\right),$$

where $\overline{\xi}$ and $\underline{\xi}$ are two positive constants with $\underline{\xi} < \overline{\xi}$, and the process $\Lambda'(\cdot)$ is of finite variation. Let

$$\mathcal{G}(\lambda', x) = G\left(\overline{\xi} \wedge \left(\underline{\xi} \vee \lambda'\right), x\right), \quad \lambda' \in \mathbb{R},\ x \in \Delta^d_+.$$

Then with $\boldsymbol{G}^{\mathrm{D}}_l$ and $\Gamma^{\boldsymbol{G}}(\cdot)$ given in (2.6) and (2.8), respectively, inserting (2.10) into (2.7) yields

$$\mathcal{G}(\Lambda'(\cdot), \mu(\cdot)) = \mathcal{G}(\Lambda'(0), \mu(0)) + \int_0^{\cdot} \sum_{j=1}^{d} \mathcal{G}^{\mathrm{D}}_j(\Lambda'(t), \mu(t)) \mathrm{d}\mu_j(t) - \Gamma^{\mathcal{G}}(\cdot),$$

where

$$\mathcal{G}^{\mathrm{D}}_i(\lambda', x) = \sum_{l=1}^{d} \frac{\mathbf{1}_{\{x_i = x_{(l)}\}}}{N_l(x)} \boldsymbol{G}^{\mathrm{D}}_l\left(\overline{\xi} \wedge \left(\underline{\xi} \vee \lambda'\right), \mathfrak{R}(x)\right), \quad i \in \{1, \cdots, d\},$$

and

$$\begin{aligned}\Gamma^{\mathcal{G}}(\cdot) = \Gamma^{\boldsymbol{G}}(\cdot) &- \sum_{l=1}^{d-1} \sum_{k=l+1}^{d} \int_0^{\cdot} \frac{\boldsymbol{G}^{\mathrm{D}}_l(\Lambda(t), \mathfrak{R}(\mu(t)))}{N_l(\mu(t))} \mathrm{d}\boldsymbol{\Lambda}^{(l,k)}(t) \\ &+ \sum_{l=2}^{d} \sum_{k=1}^{l-1} \int_0^{\cdot} \frac{\boldsymbol{G}^{\mathrm{D}}_l(\Lambda(t), \mathfrak{R}(\mu(t)))}{N_l(\mu(t))} \mathrm{d}\boldsymbol{\Lambda}^{(k,l)}(t).\end{aligned}$$

Observe that $\mathcal{G}$ is regular for $\Lambda'(\cdot)$ and $\mu(\cdot)$, yet it is not in $C^{1,2}$. □

2.2 Functional generation and relative arbitrage

In Karatzas and Ruf [51], two types of functional generation, additive and multiplicative generation, are constructed to study the properties of relative values of functionally generated trading strategies. In this section, we first discuss the generalised versions of these functional generations and the corresponding properties. Then we consider sufficient conditions for strong arbitrage relative to the market to exist.

2.2.1 Additive generation

Recall the open set $\mathcal{W}$ from (2.1).

Definition 2.2.1. (Additive generation). For a function $G : \mathcal{W} \to \mathbb{R}$, regular for $\Lambda(\cdot)$ and $\mu(\cdot)$, and the process $\vartheta(\cdot)$ given in (2.3), the trading strategy $\varphi(\cdot)$ with components

$$\varphi_i(\cdot) = \vartheta_i(\cdot) - Q^{\vartheta}(\cdot) + C, \quad i \in \{1, \cdots, d\}, \tag{2.12}$$

in the manner of (1.11) and (1.10), and with the real constant

$$C = G(\Lambda(0), \mu(0)) - \sum_{j=1}^{d} \vartheta_j(0)\mu_j(0), \tag{2.13}$$

is said to be *additively generated* by the regular function G.

In Chapter 4, we will study the *leakage* effect of rank-dependent trading strategies generated additively. In Chapter 5, we will provide an interpretation for the conjugate function of G to connect the additive functional generation with an optimal trading manner to generate profits in the long run.

Proposition 2.2.1. *The trading strategy $\varphi(\cdot)$, generated additively by a regular function $G : \mathcal{W} \to \mathbb{R}$, has components*

$$\varphi_i(\cdot) = \vartheta_i(\cdot) + \Gamma^G(\cdot) + G(\Lambda(\cdot), \mu(\cdot)) - \sum_{j=1}^{d} \mu_j(\cdot)\vartheta_j(\cdot), \tag{2.14}$$

for all $i \in \{1, \cdots, d\}$. Moreover, the wealth process of $\varphi(\cdot)$ is given by the master formula

$$V^\varphi(\cdot) = G(\Lambda(\cdot), \mu(\cdot)) + \Gamma^G(\cdot). \tag{2.15}$$

Proof. We apply the reasoning of the proof of Proposition 4.3 in Karatzas and Ruf [51] here for a generalised G. We first show (2.15). Since $\varphi(\cdot)$ is a trading strategy, by (1.7), (2.12), and (2.13), we have

$$\begin{aligned} V^\varphi(\cdot) &= \sum_{j=1}^{d} \varphi_j(\cdot)\mu_j(\cdot) = \sum_{j=1}^{d} \vartheta_j(\cdot)\mu_j(\cdot) - Q^\vartheta(\cdot) + C \\ &= \sum_{j=1}^{d} \vartheta_j(\cdot)\mu_j(\cdot) - Q^\vartheta(\cdot) + G(\Lambda(0), \mu(0)) - \sum_{j=1}^{d} \vartheta_j(0)\mu_j(0), \end{aligned} \tag{2.16}$$

which yields

$$\begin{aligned} V^\varphi(\cdot) &= V^\vartheta(\cdot) - V^\vartheta(0) + G(\Lambda(0), \mu(0)) - Q^\vartheta(\cdot) \\ &= G(\Lambda(0), \mu(0)) + \int_0^\cdot \sum_{j=1}^{d} \vartheta_j(t) \mathrm{d}\mu_j(t) \end{aligned}$$

by (1.10). Then combining the above equation and (2.4) yields (2.15).

To show (2.14), note that by (2.16) and (2.15), we have

$$C - Q^\vartheta(\cdot) = G(\Lambda(\cdot), \mu(\cdot)) + \Gamma^G(\cdot) - V^\vartheta(\cdot),$$

which together with (2.12) imply (2.14). □

Remark 4. A trading strategy $\varphi(\cdot)$ generated additively from a regular function G is not necessarily long-only as defined in Definition 1.2.2. As given by (2.14), the value of the trading strategy $\varphi(\cdot)$ depends strongly on the value of G, which is determined by the function form of G, as well as the choice of $\Lambda(\cdot)$. Hence, by the relationship

$$\pi_i(\cdot) = \frac{\varphi_i(\cdot)\mu_i(\cdot)}{V^{\varphi}(\cdot)}, \quad i \in \{1, \cdots, d\}, \tag{2.17}$$

the portfolio weight $\pi_i(\cdot)$ is non-positive whenever $\varphi_i(\cdot)$ is negative.

Whenever the wealth $V^{\varphi}(t)$ is positive, for all $t \geq 0$, the portfolio weights $\pi(t)$ has components

$$\begin{aligned} \pi_i(t) &= \left(1 + \frac{\vartheta_i(t) - \sum_{j=1}^d \mu_j(t)\vartheta_j(t)}{V^{\varphi}(t)}\right)\mu_i(t) \\ &= \left(1 + \frac{\vartheta_i(t) - \sum_{j=1}^d \mu_j(t)\vartheta_j(t)}{G(\Lambda(t), \mu(t)) + \Gamma^G(t)}\right)\mu_i(t), \quad i \in \{1, \cdots, d\}, \end{aligned} \tag{2.18}$$

by (2.17), (2.14), and (2.15). □

2.2.2 Multiplicative generation

Definition 2.2.2. (Multiplicative generation). For a function $G : \mathcal{W} \to (0, \infty)$, regular for $\Lambda(\cdot)$ and $\mu(\cdot)$, let the process $\vartheta(\cdot)$ be given in (2.3) and assume that $1/G(\Lambda(\cdot), \mu(\cdot))$ is locally bounded. Consider the process $\overline{\vartheta}(\cdot) \in \mathcal{L}(\mu)$ with components

$$\overline{\vartheta}_i(\cdot) = \vartheta_i(\cdot) \exp\left(\int_0^{\cdot} \frac{\mathrm{d}\Gamma^G(t)}{G\big(\Lambda(t), \mu(t)\big)}\right), \quad i \in \{1, \cdots, d\}. \tag{2.19}$$

Then the trading strategy $\psi(\cdot)$ with components

$$\psi_i(\cdot) = \overline{\vartheta}_i(\cdot) - Q^{\overline{\vartheta}}(\cdot) + C, \quad i \in \{1, \cdots, d\}, \tag{2.20}$$

in the manner of (1.11) and (1.10), and with C given in (2.13), is said to be *multiplicatively generated* by the regular function G.

Proposition 2.2.2. *The trading strategy $\psi(\cdot)$, generated multiplicatively by a regular function $G : \mathcal{W} \to (0, \infty)$ with $1/G(\Lambda(\cdot), \mu(\cdot))$ locally bounded, has components*

$$\psi_i(\cdot) = V^{\psi}(\cdot)\left(1 + \frac{\vartheta_i(\cdot) - \sum_{j=1}^d \vartheta_j(\cdot)\mu_j(\cdot)}{G(\Lambda(\cdot), \mu(\cdot))}\right), \tag{2.21}$$

for all $i \in \{1, \cdots, d\}$, where the wealth process of $\psi(\cdot)$ is given by the master formula

$$V^{\psi}(\cdot) = G(\Lambda(\cdot), \mu(\cdot)) \exp\left(\int_0^{\cdot} \frac{\mathrm{d}\Gamma^G(t)}{G(\Lambda(t), \mu(t))}\right) > 0. \tag{2.22}$$

Proof. The reasoning of the proof of Proposition 4.8 in Karatzas and Ruf [51] is applied for a generalised G. First we show that (2.22) holds. Since $\psi(\cdot)$ is a trading strategy, by (1.7), (2.20), and (2.19), we have

$$\begin{aligned}\mathrm{d}V^{\psi}(t) &= \sum_{j=1}^{d} \psi_j(t)\mathrm{d}\mu_j(t) = \sum_{j=1}^{d} \overline{\vartheta}_j(t)\mathrm{d}\mu_j(t) \\ &= \exp\left(\int_0^t \frac{\mathrm{d}\Gamma^G(u)}{G(\Lambda(u),\mu(u))}\right)\sum_{j=1}^{d} \vartheta_j(t)\mathrm{d}\mu_j(t).\end{aligned}$$

Then by (2.4), we have

$$\sum_{j=1}^{d} \vartheta_j(t)\mathrm{d}\mu_j(t) = \mathrm{d}\Gamma^G(t) + \mathrm{d}G(\Lambda(t),\mu(t)),$$

which implies

$$\begin{aligned}\mathrm{d}V^{\psi}(t) &= \exp\left(\int_0^t \frac{\mathrm{d}\Gamma^G(u)}{G(\Lambda(u),\mu(u))}\right)\left(\mathrm{d}\Gamma^G(t) + \mathrm{d}G(\Lambda(t),\mu(t))\right) \\ &= \mathrm{d}\left(G(\Lambda(t),\mu(t))\exp\left(\int_0^t \frac{\mathrm{d}\Gamma^G(u)}{G(\Lambda(u),\mu(u))}\right)\right),\end{aligned} \tag{2.23}$$

where the second equality is by the product rule.

Moreover, by (2.20), (1.10), (2.13), and (2.19), we have

$$\begin{aligned}V^{\psi}(0) &= \sum_{j=1}^{d} \psi_j(0)\mu_j(0) = \sum_{j=1}^{d} \overline{\vartheta}_j(0)\mu_j(0) + C \\ &= \sum_{j=1}^{d} \vartheta_j(0)\mu_j(0) + G(\Lambda(0),\mu(0)) - \sum_{j=1}^{d} \mu_j(0)\vartheta_j(0) = G(\Lambda(0),\mu(0)) \\ &= G(\Lambda(0),\mu(0)) + \Gamma^G(0),\end{aligned}$$

which together with (2.23) implies (2.22).

To show (2.21), by (2.19) and (2.22), it is equivalent to show

$$\psi_i(\cdot) = V^{\psi}(\cdot) + \overline{\vartheta}_i(\cdot) - \sum_{j=1}^{d} \overline{\vartheta}_j(\cdot)\mu_j(\cdot) = V^{\psi}(\cdot) + \overline{\vartheta}_i(\cdot) - V^{\overline{\vartheta}}(\cdot), \tag{2.24}$$

for all $i \in \{1, \cdots, d\}$. By (1.10) and (2.13), (2.20) yields

$$\psi_i(\cdot) = \overline{\vartheta}_i(\cdot) - V^{\overline{\vartheta}}(\cdot) + V^{\overline{\vartheta}}(0) + \int_0^{\cdot} \sum_{j=1}^{d} \overline{\vartheta}_j(t)\mathrm{d}\mu_j(t) + G(\Lambda(0),\mu(0)) - V^{\vartheta}(0),$$

which, by (2.20), (2.22), (1.8), and $\sum_{j=1}^{d} \mathrm{d}\mu_j(t) = 0$, yields

$$\psi_i(\cdot) = \overline{\vartheta}_i(\cdot) - V^{\overline{\vartheta}}(\cdot) + V^{\psi}(0) + \int_0^{\cdot} \sum_{j=1}^{d} \psi_j(t)\mathrm{d}\mu_j(t) = \overline{\vartheta}_i(\cdot) - V^{\overline{\vartheta}}(\cdot) + V^{\psi}(\cdot),$$

for all $i \in \{1, \cdots, d\}$, i.e., (2.24) indeed holds. □

Remark 5. Let the trading strategy $\psi(\cdot)$ be generated multiplicatively by G. By (2.21) and

$$\pi_i(\cdot) = \frac{\psi_i(\cdot)\mu_i(\cdot)}{V^{\psi}(\cdot)}, \quad i \in \{1, \cdots, d\},$$

we have

$$\pi_i(t) = \left(1 + \frac{\vartheta_i(t) - \sum_{j=1}^{d} \vartheta_j(t)\mu_j(t)}{G(\Lambda(t), \mu(t))}\right)\mu_i(t), \quad i \in \{1, \cdots, d\},\ t \geq 0. \tag{2.25}$$

Compared with (1.14), (2.25) is more general in that G can be non-smooth and depend on an extra component $\Lambda(\cdot)$. □

2.2.3 Sufficient conditions for relative arbitrage

In Karatzas and Ruf [51], Theorems 5.1 and 5.2 give sufficient conditions for strong arbitrage relative to the market to exist for both additively and multiplicatively generated trading strategies, respectively. These results still hold for a regular / Lyapunov function $G : \mathcal{W} \to [0, \infty)$ under specific conditions.

To be consistent with the conditions of arbitrage relative to the market in (1.12), we normalise $G(\Lambda(0), \mu(0)) = 1$ such that both of the wealth processes in (2.15) and (2.22) have initial values 1. This normalisation is guaranteed by replacing G with $G+1$ when $G(\Lambda(0), \mu(0)) = 0$, or with $G/G(\Lambda(0), \mu(0))$ when $G(\Lambda(0), \mu(0)) > 0$.

Theorem 2.2.3. *Fix a function $G : \mathcal{W} \to [0, \infty)$, Lyapunov for $\Lambda(\cdot)$ and $\mu(\cdot)$, with $G(\Lambda(0), \mu(0)) = 1$. For some real number $T_* > 0$, suppose that*

$$P\left[\Gamma^G(T_*) > 1\right] = 1. \tag{2.26}$$

Then the additively generated trading strategy $\varphi(\cdot)$ of Definition 2.2.1 is strong arbitrage relative to the market over every investment horizon $[0, T]$ with $T \geq T_*$.

Proof. By (2.15) and $G(\Lambda(0), \mu(0)) = 1$, we have $V^{\varphi}(0) = 1$. Since G takes values on $[0, \infty)$ and is Lyapunov for $\Lambda(\cdot)$ and $\mu(\cdot)$, i.e., $\Gamma^G(\cdot)$ is non-decreasing, we have $V^{\varphi}(\cdot) \geq 0$. Moreover, by (2.26), we have

$$V^{\varphi}(T) = G(\Lambda(T), \mu(T)) + \Gamma^G(T) \geq \Gamma^G(T_*) > 1, \quad T \geq T_*,$$

i.e., (1.12) holds. Hence, Definition 1.3.1 implies the desired result. □

Theorem 2.2.4. *Assume that $|\Lambda(\cdot)|$ is uniformly bounded. Fix a function $G : \mathcal{W} \to [0, \infty)$, regular for $\Lambda(\cdot)$ and $\mu(\cdot)$, with $G(\Lambda(0), \mu(0)) = 1$. For some real numbers $T_* > 0$, suppose that we can find an $\varepsilon = \varepsilon(T_*) > 0$ such that*

$$P\left[\Gamma^G(T_*) > 1 + \varepsilon\right] = 1. \tag{2.27}$$

Then there exists a constant $c = c(T_*, \varepsilon) > 0$ such that the trading strategy $\psi^{(c)}(\cdot)$, generated multiplicatively by the regular function

$$G^{(c)} = \frac{G + c}{1 + c} \tag{2.28}$$

as in Definition 2.2.2, is strong arbitrage relative to the market over the investment horizon $[0, T_*]$. Moreover, if G is a Lyapunov function for $\Lambda(\cdot)$ and $\mu(\cdot)$, then $\psi^{(c)}(\cdot)$ is also a strong relative arbitrage over every investment horizon $[0, T]$ with $T \geq T_*$.

Proof. We prove the theorem for a generalised G, using the same argument in the proof of Theorem 5.2 in Karatzas and Ruf [51]. Since the regular function G takes values on $[0, \infty)$, by (2.22) and $G(\Lambda(0), \mu(0)) = 1$, we have $V^{\psi^{(c)}}(0) = 1$ and $V^{\psi^{(c)}}(\cdot) \geq 0$. To make $\psi^{(c)}(\cdot)$ strong relative arbitrage over $[0, T_*]$, by Definition 1.3.1, we need to show $\mathsf{P}[V^{\psi^{(c)}}(T_*) > 1] = 1$.

For constant $c > 0$, by (2.28) and (2.4), we have

$$\Gamma^{G^{(c)}}(\cdot) = \frac{\Gamma^G(\cdot)}{1 + c} \quad \text{and} \quad G^{(c)} \geq \frac{c}{1 + c},$$

which, by (2.22), implies

$$\begin{aligned} V^{\psi^{(c)}}(T_*) &\geq \frac{c}{1+c} \exp\left(\int_0^{T_*} \frac{\mathrm{d}\Gamma^{G^{(c)}}(t)}{G^{(c)}(\Lambda(t), \mu(t))} \right) \\ &= \frac{c}{1+c} \exp\left(\int_0^{T_*} \frac{\mathrm{d}\Gamma^{G}(t)}{G(\Lambda(t), \mu(t)) + c} \right). \end{aligned} \tag{2.29}$$

Since $G(\Lambda(\cdot), \mu(\cdot))$ is uniformly bounded thanks to the assumptions, there exists an upper bound η of G. Then by (2.27), (2.29) yields

$$V^{\psi^{(c)}}(T_*) \geq \frac{c}{1+c} \exp\left(\frac{1}{\eta + c} \int_0^{T_*} \mathrm{d}\Gamma^G(t) \right) > \frac{c}{1+c} \exp\left(\frac{1+\varepsilon}{\eta + c} \right).$$

To proceed, note that

$$\frac{c}{1+c} \exp\left(\frac{1+\varepsilon}{\eta+c} \right) = \exp\left(\frac{\varepsilon - \eta \log(1 + 1/c) + 1 - c \log(1 + 1/c)}{\eta + c} \right).$$

Since $\log(1 + x) < x$, for all $x > 0$, we have

$$V^{\psi^{(c)}}(T_*) > \exp\left(\frac{\varepsilon - \eta \log(1 + 1/c)}{\eta + c} \right) > 1,$$

for sufficiently large c. Therefore, there exists constant c sufficiently large such that $\mathsf{P}[V^{\psi^{(c)}}(T_*) > 1] = 1$ holds.

If G is Lyapunov for $\Lambda(\cdot)$ and $\mu(\cdot)$, then by (2.27), we have

$$\mathsf{P}\left[\Gamma^G(T) \geq \Gamma^G(T_*) > 1 + \varepsilon \right] = 1, \quad T \geq T_*.$$

Then, applying the same reasoning above yields the desired result. □

2.3 Data source and processing

We start this section by describing the data used in the next section, where several trading strategies are implemented. Then we discuss the method to process the data.

2.3.1 Data source and description

We shall consider a market consisting of all stocks in the S&P 500 index. We are interested in the beginning of day and the end of day market weights of each of these stocks. To calculate these market weights accurately (according to the method in Subsection 2.3.2), we make use of two time series: the daily market values (market capitalisations, which exclude all the dividend payments) and the daily return indexes (used to consider the effect of reinvestment of dividend payments) of the corresponding component stocks in the S&P 500 index. Both of these time series are available at the end of each trading day.

The data of the market values and return indexes is downloaded from DataStream.[2] The first day, for which the data is available on DataStream, is September 29$^{\mathrm{th}}$, 1989. Since then there are in total 1140 constituents that have belonged to the S&P 500 index. A list of stocks in the S&P 500 index is also attainable on DataStream. In particular, for each month, we derive the list of constituents of the index at the last day of this month. For a constituent delisted from the index in that month, we keep it in our portfolio provided that the constituent still remains in the market till the end of that month. However, we get rid of it from our portfolio on the same day when the constituent does no longer exist in the market, usually due to mergers and acquisitions, bankruptcies, etc. For a constituent newly added to the index in that month, we put it into our portfolio from the first day of the following month.

2.3.2 Data processing

Theoretically, trading strategies vary continuously in time, while in the empirical analysis a daily trading frequency is used. The following procedure illustrates how we examine the gains and losses in our portfolio relative to the market portfolio.

We discretise the time horizon as $0 = t_0 < t_1 < \cdots < t_{N-1} = T$, where N is the total number of trading days.

- The transaction on day t_l, for all $l \in \{1, \cdots, N-1\}$, is made at the beginning of day ($\underline{t}_l$), taking the beginning of day t_l market weights $\mu(\underline{t}_l)$ as inputs. These

[2]DataStream, operated by Thomson Reuters, is a financial time series database; see `https://financial.thomsonreuters.com/en/products/data-analytics/economic-data.html`.

market weights $\mu(\underline{t}_l)$ are computed by

$$\mu_i(\underline{t}_l) = \frac{\mathrm{MV}_i(\underline{t}_l)}{\Sigma(\underline{t}_l)}, \quad i \in \{1, \cdots, d\},$$

where $\mathrm{MV}_i(\underline{t}_l)$ is the market value of stock i at the beginning of day t_l, which is assumed to be equal to the market value attainable at the end of the last trading day t_{l-1}, and $\Sigma(\underline{t}_l) = \sum_{j=1}^{d} \mathrm{MV}_j(\underline{t}_l)$ denotes the total market capitalisation at the beginning of day t_l.

- The theoretical (non-self-financing) trading strategy throughout day t_l, denoted by $\theta(t_l)$, is computed based on either (2.3) or (2.19), taking $\mu(\underline{t}_l)$ as inputs. Denote the implemented (self-financing) trading strategy corresponding to $\theta(t_l)$ by $\phi(t_l)$. Then $V^\phi(\underline{t}_l)$, the beginning of day t_l wealth of the portfolio corresponding to $\phi(t_l)$, is given by
$$V^\phi(\underline{t}_1) = \frac{V^\phi(\bar{t}_{l-1})\Sigma(\bar{t}_{l-1})}{\Sigma(\underline{t}_1)}. \tag{2.30}$$
This is based on the assumption that the real portfolio wealth does not change overnight. In (2.30), $V^\phi(\bar{t}_{l-1})$ and $\Sigma(\bar{t}_{l-1})$ are the end of day t_{l-1} portfolio wealth and total market capitalisation, respectively, computed at $\bar{t}_{l-1}$ (thus already known at $\underline{t}_l$).

- To derive the implemented (self-financing) trading strategy $\phi(t_l)$ corresponding to $\theta(t_l)$, we compute the number
$$C(t_l) = \sum_{j=1}^{d} \theta_j(t_l)\mu_j(\underline{t}_l) - V^\phi(\underline{t}_l). \tag{2.31}$$
Then $\phi(t_l)$ is derived by
$$\phi_i(t_l) = \theta_i(t_l) - C(t_l), \quad i \in \{1, \cdots, d\}. \tag{2.32}$$
This guarantees
$$V^\phi(\underline{t}_l) = \sum_{i=1}^{d} \phi_i(t_l)\mu_i(\underline{t}_l).$$

- At the end of day t_l, the return indexes of the stocks for t_l are available, and the total returns $\mathrm{TR}(t_l)$ are computed through dividing the return indexes of t_l with the return indexes of t_{l-1}. Then the end of day t_l implied market values $\mathrm{MV}(\bar{t}_l)$, which take the dividend payments into consideration, are given by
$$\mathrm{MV}_i(\bar{t}_l) = \mathrm{MV}_i(\underline{t}_l)\mathrm{TR}_i(t_l), \quad i \in \{1, \cdots, d\}.$$
The end of day t_l modified total market capitalisation $\Sigma(\bar{t}_l)$ and market weights $\mu(\bar{t}_l)$ are calculated similarly as $\Sigma(\underline{t}_l)$ and $\mu(\underline{t}_l)$, with $\mathrm{MV}(\underline{t}_l)$ replaced by $\mathrm{MV}(\bar{t}_l)$.

- The end of day t_l portfolio wealth is then computed by

$$V^{\phi}(\bar{t}_l) = \sum_{j=1}^{d} \phi_j(t_l)\mu_j(\bar{t}_l).$$

Note that we have

$$V^{\phi}(\bar{t}_l) = V^{\phi}(\underline{t}_l) + \sum_{j=1}^{d} \theta_j(t_l) \left(\mu_j(\bar{t}_l) - \mu_j(\underline{t}_l)\right). \tag{2.33}$$

In particular, at the beginning of day t_0, all of the above steps are still applied, except that we have $V^{\phi}(\underline{t}_0) = 1$ instead of (2.30) due to Definition 1.3.1.

2.4 Examples and empirical results

In this section, several examples of portfolio generating functions are empirically studied. In particular, the performance of these trading strategies will be analysed further in Chapter 3 when incorporating with transaction costs. Recall the open set Δ_+^d from (1.3).

Example 2.4.1. Define the generalised *entropy function*

$$G(\lambda, x) = \lambda \sum_{j=1}^{d} x_j \log\left(\frac{1}{x_j}\right), \quad \lambda \in \mathbb{R}_+,\ x \in \Delta_+^d,$$

with values in $(0, \lambda \log d)$, for fixed $\lambda > 0$. Suppose that $\mu(\cdot)$ takes values in Δ_+^d and that $\Lambda(\cdot)$ is $(0, \infty)$-valued.

From (2.5) we have

$$\Gamma^G(\cdot) = \sum_{j=1}^{d} \int_0^{\cdot} \mu_j(t) \log \mu_j(t) \mathrm{d}\Lambda(t) + \frac{1}{2} \sum_{j=1}^{d} \int_0^{\cdot} \Lambda(t) \frac{\mathrm{d}\left[\mu_j, \mu_j\right](t)}{\mu_j(t)}. \tag{2.34}$$

Then G is a Lyapunov function for $\Lambda(\cdot)$ and $\mu(\cdot)$ provided that $\Gamma^G(\cdot)$ is non-decreasing. One sufficient condition for this to hold is that $\Lambda(\cdot)$ is non-increasing.

From (2.14), the trading strategy $\varphi(\cdot)$, generated additively by G, has components

$$\varphi_i(\cdot) = \Gamma^G(\cdot) - \Lambda(\cdot) \log \mu_i(\cdot), \quad i \in \{1, \cdots, d\}. \tag{2.35}$$

Using (2.15), the corresponding wealth process $V^{\varphi}(\cdot)$ is strictly positive if G is Lyapunov for $\Lambda(\cdot)$ and $\mu(\cdot)$.

For the multiplicative generation, G is required to be bounded away from zero. One sufficient condition for this to hold is that $\Lambda(\cdot)$ is bounded away from 0 and the market is diverse on $[0, \infty)$, i.e., there exists $\epsilon > 0$ such that $G(\Lambda(t), \mu(t)) \geq \Lambda(t)\epsilon$, for all $t \geq 0$ (see Proposition 2.3.2 in Fernholz [26]). Then from (2.21), the trading strategy $\psi(\cdot)$, generated multiplicatively by G, has components

$$\psi_i(\cdot) = -\Lambda(\cdot)\log\mu_i(\cdot)\exp\left(\int_0^{\cdot}\frac{\mathrm{d}\Gamma^G(t)}{G(\Lambda(t),\mu(t))}\right), \quad i\in\{1,\cdots,d\}.$$

The corresponding wealth process $V^{\psi}(\cdot)$ is given in (2.22).

Now, let us discuss sufficient conditions for the existence of arbitrage relative to the market. To this end, let $\Lambda(\cdot)$ be such that G is Lyapunov for $\Lambda(\cdot)$ and $\mu(\cdot)$, for example, let $\Lambda(\cdot)$ be non-increasing. Next, consider

$$\mathcal{G} = \frac{G}{G(\Lambda(0),\mu(0))}, \tag{2.36}$$

together with the non-decreasing process

$$\Gamma^{\mathcal{G}}(\cdot) = \frac{\Gamma^G(\cdot)}{G(\Lambda(0),\mu(0))}. \tag{2.37}$$

Then from Theorem 2.2.3, if

$$\mathsf{P}\left[\Gamma^{\mathcal{G}}(T_*) > 1\right] = \mathsf{P}\left[\Gamma^G(T_*) > G(\Lambda(0),\mu(0))\right] = 1,$$

then the trading strategy $\varphi(\cdot)/G(\Lambda(0),\mu(0))$, generated additively by $\mathcal{G}$, is strong relative arbitrage over every time horizon $[0,T]$ with $T\geq T_*$.

Similarly, from Theorem 2.2.4, if

$$\mathsf{P}\left[\Gamma^{\mathcal{G}}(T_*) > 1+\varepsilon\right] = \mathsf{P}\left[\Gamma^G(T_*) > G(\Lambda(0),\mu(0))(1+\varepsilon)\right] = 1,$$

then the trading strategy $\psi^{(c)}(\cdot)$, generated multiplicatively by

$$G^{(c)} = \frac{G+c}{G(\Lambda(0),\mu(0))+c}, \tag{2.38}$$

for some sufficiently large $c>0$, is strong relative arbitrage over every time horizon $[0,T]$ with $T\geq T_*$.

To empirically examine the performance of the portfolio generated by G, we only restrict G to be regular for $\Lambda(\cdot)$ and $\mu(\cdot)$, although G is Lyapunov for some of the choices of $\Lambda(\cdot)$ in the following.

Recall that the wealth processes of portfolios generated either additively or multiplicatively are relative to the S&P 500 index. For a specific day t_n, we estimate

$$\left[\mu_i,\mu_i\right](t_n) \approx \sum_{l=1}^{n}(\mu_i(\bar{t}_l)-\mu_i(\underline{t}_l))^2, \quad i\in\{1,\cdots,d\},$$

where $\underline{t}_l$ ($\bar{t}_l$) denotes the beginning (end) of the day t_l.

Figure 2.1 presents $\Gamma^{\mathcal{G}}(\cdot)$ given in (2.37) and the relative wealth processes $V^{\varphi}(\cdot)$ and $V^{\psi^{(0)}}(\cdot)$ (minus 1 to start from 0 as $\Gamma^{\mathcal{G}}(\cdot)$) of trading strategies generated additively and multiplicatively by $\mathcal{G}$, respectively, with finite-variation process $\Lambda(\cdot) = 1$. As we can observe from the figure, both $V^{\varphi}(\cdot)$ and $V^{\psi^{(0)}}(\cdot)$ have been continuously outperforming the market trading strategy since the year 2000.

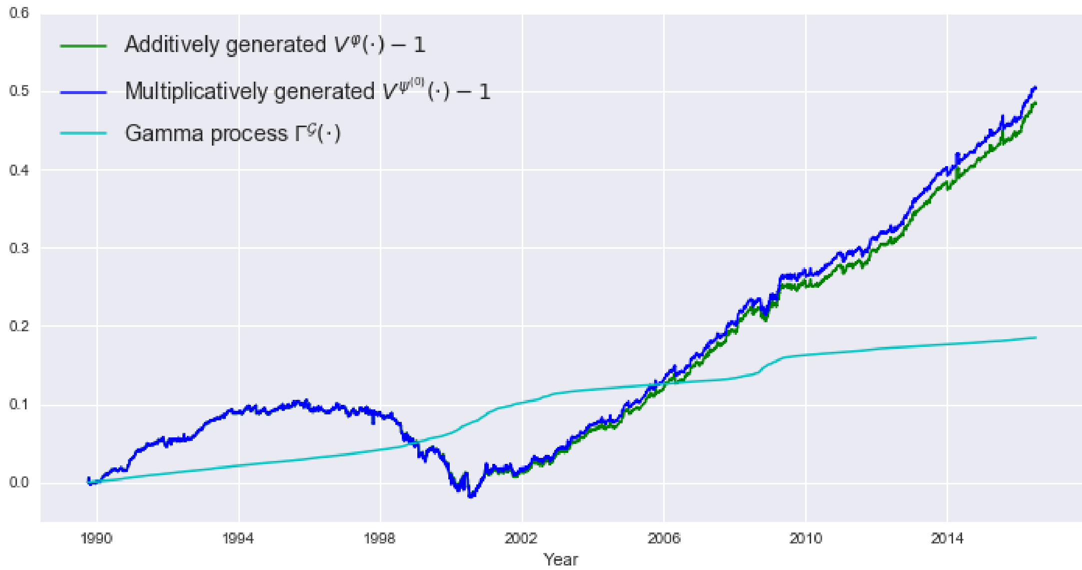

FIGURE 2.1: Gamma process $\Gamma^{\mathcal{G}}(\cdot)$ and relative wealth processes (minus 1) of both the additively and the multiplicatively generated trading strategies with constant $\Lambda(\cdot) = 1$.

Next, we examine the effect that choosing some non-constant $\Lambda(\cdot)$ may have on the portfolio performance. Figures 2.2 and 2.3 display the relative wealth processes $V^{\varphi}(\cdot)$ (in logarithmic scale) generated additively corresponding to two different groups of $\Lambda(\cdot)$. The first group of $\Lambda(\cdot)$ is increasing, which results in decreasing $\Gamma^{G}(\cdot)$ given by (2.34); the corresponding G is only regular but not Lyapunov for $\Lambda(\cdot)$ and $\mu(\cdot)$. The second group of $\Lambda(\cdot)$ is decreasing; the corresponding $\Gamma^{G}(\cdot)$ given by (2.34) is increasing and G is Lyapunov for $\Lambda(\cdot)$ and $\mu(\cdot)$.

More precisely, for all $l \in \{1, \cdots, N\}$, in Figure 2.2, the wealth processes $V^{\varphi}(\cdot)$ corresponding to $\Lambda(t_l) = \exp\left(10^{-4} l\right)$ and $\Lambda(t_l) = \exp\left(-10^{-4} l\right)$ are plotted; in Figure 2.3, the wealth processes $V^{\varphi}(\cdot)$ corresponding to

$$\Lambda(t_l) = \exp\left(100 \sum_{j=1}^{d} [\mu_j, \mu_j](t_l)\right) \quad \text{and} \quad \Lambda(t_l) = \exp\left(-100 \sum_{j=1}^{d} [\mu_j, \mu_j](t_l)\right)$$

are plotted. The constants 10^{-4} and 100 are chosen such that, with these forms, the daily changes of both $\mathcal{G}(\Lambda(\cdot), \mu(\cdot))$ and $\Gamma^{\mathcal{G}}(\cdot)$ are roughly at the same level of magnitude. Hence, in (2.15), neither part on the right hand side dominates the other.

As we can observe from the figures, choosing $\Lambda(\cdot)$ increasing seems to lead to a better performance than choosing $\Lambda(\cdot)$ constant, which again seems to be better than choosing $\Lambda(\cdot)$ decreasing. We attribute the reason behind this observation to the state of market diversification as follows.

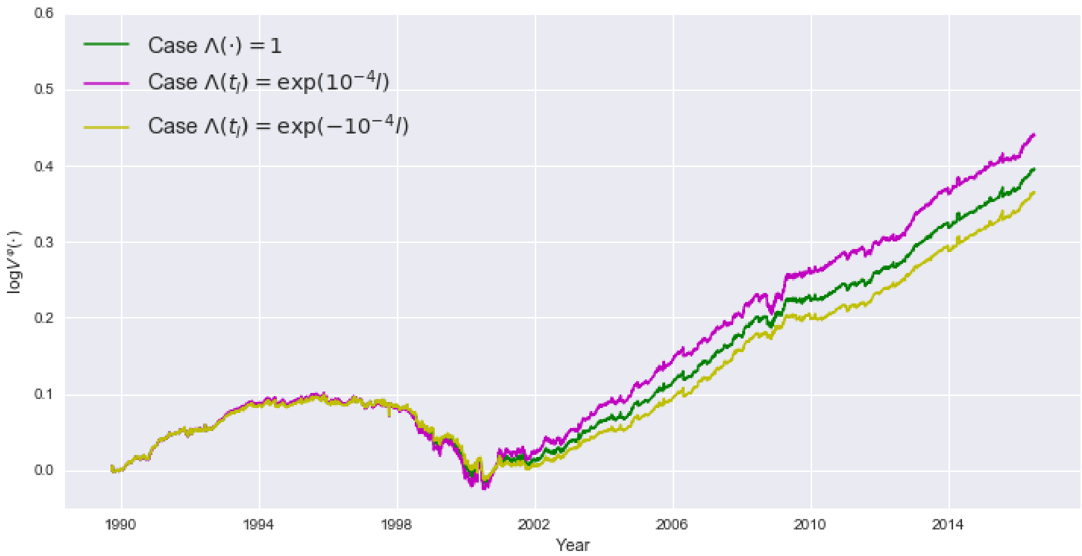

FIGURE 2.2: Relative wealth process $V^{\varphi}(\cdot)$ (in logarithmic scale) of additively generated trading strategies with $\Lambda(\cdot)$ a deterministic exponential.

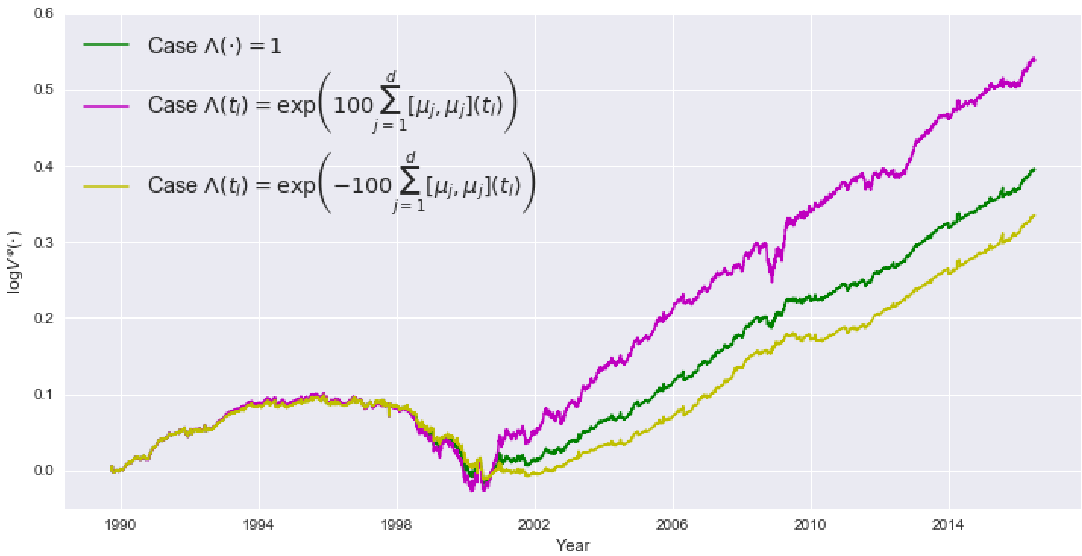

FIGURE 2.3: Relative wealth process $V^{\varphi}(\cdot)$ (in logarithmic scale) of additively generated trading strategies with $\Lambda(\cdot)$ an exponential of the quadratic variation of $\mu(\cdot)$.

Observe that (2.33) yields

$$V^{\varphi}(\bar{t}_l) = V^{\varphi}(\underline{t}_l) + \frac{1}{G(\Lambda(0), \mu(\underline{0}))} \Lambda(t_l) D(t_l), \quad l \in \{0, \cdots, N\}, \tag{2.39}$$

where $D(t_l)$ is given by

$$D(t_l) = \sum_{j=1}^{d} -\log \mu_j(\underline{t}_l)(\mu_j(\bar{t}_l) - \mu_j(\underline{t}_l)). \tag{2.40}$$

The value $D(t_l)$ can be considered as an indicator of the direction of changes in market weights from the beginning to the end of day t_l. The value $D(t_l)$ will be positive (negative), if market weights are shifted from companies with large (small) beginning of day market weights to companies with small (large) beginning of day market weights throughout day t_l. We consider a simple example to better understand why this is the case.

Fix $d = 2$ and assume that $\mu_1(\underline{t}_l) > \mu_2(\underline{t}_l)$. Then

$$\begin{aligned} D(t_l) &= -\log \mu_1(\underline{t}_l)(\mu_1(\bar{t}_l) - \mu_1(\underline{t}_l)) - \log \mu_2(\underline{t}_l)(\mu_2(\bar{t}_l) - \mu_2(\underline{t}_l)) \\ &= (-\log \mu_1(\underline{t}_l) + \log \mu_2(\underline{t}_l))(\mu_1(\bar{t}_l) - \mu_1(\underline{t}_l)) \end{aligned}$$

holds due to the fact that

$$\mu_1(\bar{t}_l) - \mu_1(\underline{t}_l) = -(\mu_2(\bar{t}_l) - \mu_2(\underline{t}_l)).$$

Hence, $D(t_l) > 0$ if and only if $\mu_1(\bar{t}_l) < \mu_1(\underline{t}_l)$, i.e., the market weight of the company with larger beginning of day market weight decreases, while the market weight of the company with smaller beginning of day market weight increases.

Hence, a positive $D(\cdot)$ indicates an enhancement in market diversification, while $D(\cdot)$ being negative actually implies a reduction in market diversification. Figure 2.4 plots the cumulative process

$$E(\cdot) = \sum_{t_l=t_1}^{\cdot} D(t_l).$$

The process $E(\cdot)$ is increasing (decreasing) whenever $D(\cdot)$ is positive (negative). From Figure 2.4 we can observe that after a slight increase from the year 1991 to the year 1995, $E(\cdot)$ keeps declining till the year 2000. Then $E(\cdot)$ rises up in the long run from the year 2000 until now.

The behaviour of the process $E(\cdot)$ is in line with another measurement of the market diversification. More precisely, let us consider the process $\sum_{j=1}^{d}(\mu_j \wedge 0.002)(\cdot)$. Note that the value $0.002 = 1/500$, which is roughly the number of constituents in the portfolio. This process is a measure of the market diversification, as it goes up when the market weights of small companies become larger, i.e., the market diversification is strengthened. Figure 2.5 plots the process, which first grows from the year 1991 to the year 1995. Then from the year 1995 to 2000, the process declines fast. This indicates that during this period, the market diversification weakens. On the contrary, the market diversification strengthens afterwards until the year 2008, as the process goes up. Then the level of market diversification remains within a relatively small range.

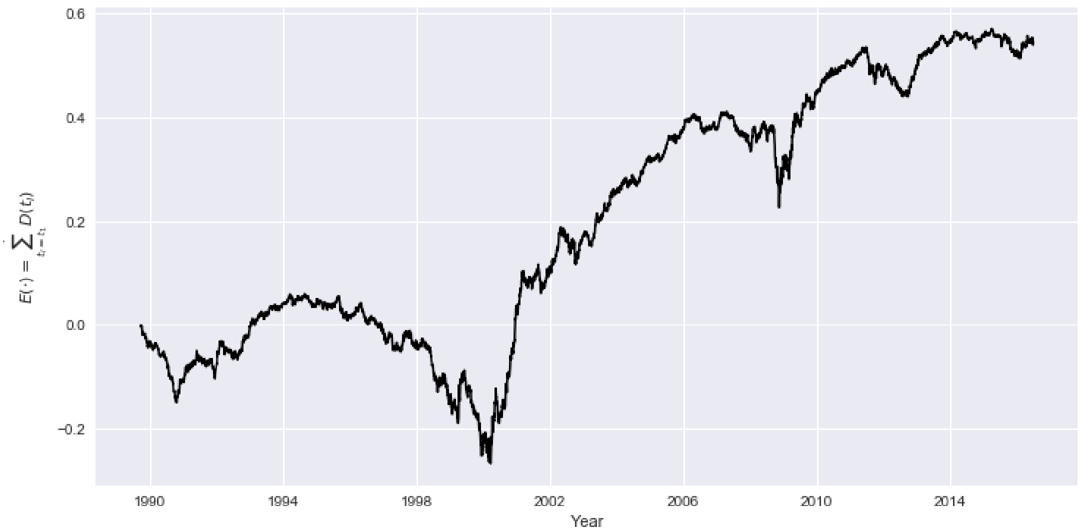

FIGURE 2.4: Integration process $E(\cdot)$ with components given by (2.40).

FIGURE 2.5: Process $\sum_{j=1}^{d}(\mu_j \wedge 0.002)(\cdot)$ as a measure of the market diversification degree in the S&P 500 market.

As a result, according to (2.39), if the market presents a trend of increasing diversification, an increasing positive $\Lambda(\cdot)$ helps to reinforce this effect, and further assists in pulling up $V^{\varphi}(\cdot)$, while a decreasing positive $\Lambda(\cdot)$ is counteractive. On the other hand, if the market presents a trend of decreasing diversification, then a decreasing positive $\Lambda(\cdot)$ helps to slow down the declining speed of $V^{\varphi}(\cdot)$, while an increasing positive $\Lambda(\cdot)$ would make the speed even faster. This is confirmed in Figures 2.2 and 2.3, as from the year 1991 to the year 1995 and from the year 2000 till now, an increasing positive $\Lambda(\cdot)$ makes $V^{\varphi}(\cdot)$ perform better, while from the year 1995 to the year 2000, $V^{\varphi}(\cdot)$ corresponding to a decreasing positive $\Lambda(\cdot)$ is slightly larger.

Although an increasing positive $\Lambda(\cdot)$ has positive effect on the portfolio performance $V^{\varphi}(\cdot)$ whenever the market diversification strengthens, we are not allowed to choose $\Lambda(\cdot)$ arbitrarily fast increasing. The reason is that the trading strategy $\varphi(\cdot)$ is required to be long-only in our framework, i.e., $\varphi(\cdot)$ given by (2.35) must be nonnegative at any time. If $\Lambda(\cdot)$ is increasing fast enough, $\Gamma^{\mathcal{G}}(\cdot)$ will become negative and decrease fast, which may result in negative $\varphi(\cdot)$ according to (2.35).

As for the multiplicative generation, the different choices of finite-variation process do not change the wealth processes significantly. Indeed, according to (2.34), an increasing $\Lambda(\cdot)$ may slow down the growth rate of $\Gamma(\cdot)$, or even turn $\Gamma(\cdot)$ into a decreasing one. When applying (2.32) to $\overline{\vartheta}(\cdot)$ from (2.19), we have

$$V^{\psi^{(c)}}(\bar{t}_l) = \exp\left(\int_0^{\underline{t}_l} \frac{\mathrm{d}\Gamma^G(t)}{G(\Lambda(t), \mu(t)) + c}\right) \frac{\Lambda(t_l)}{G(\Lambda(0), \mu(\underline{0})) + c} D(t_l) + V^{\psi^{(c)}}(\underline{t}_l),$$

for all $l \in \{0, \cdots, N\}$, with $D(\cdot)$ given in (2.40). In this example, according to the above equation, the positive effect in boosting $V^{\psi^{(c)}}(\cdot)$ contributed by an increasing positive $\Lambda(\cdot)$ is counteracted more or less by the opposite impact the same $\Lambda(\cdot)$ has on the exponential part. A similar analysis also applies to a decreasing positive $\Lambda(\cdot)$. Therefore, under the above mentioned situation (market diversification increases in general), the different choices of a monotone $\Lambda(\cdot)$ do not influence $V^{\psi^{(c)}}(\cdot)$ as much as they do on $V^{\varphi}(\cdot)$.

Note that our process $D(\cdot)$ is related but not the same as the Bregman divergence

$$D_{B,G}\left[\mu(\bar{t}_l)|\mu(\underline{t}_l)\right] = \Lambda(t_l)D(t_l) - \big(G\big(\Lambda(t_l), \mu(\bar{t}_l)\big) - G\big(\Lambda(t_l), \mu(\underline{t}_l)\big)\big),$$

defined in Definition 3.6 in Wong [93]. For its connection to optimal transport, we refer to Wong [93].

To conclude this example, we compute several empirical indicators corresponding to the performance of above mentioned trading strategies over the chosen time horizon. The S&P 500 market trading strategy has an averaged yearly return of 9.87% and a Sharpe ratio of 0.37.[3] As for the functionally generated trading strategies analyzed in this example, their averaged yearly returns are ranging from 11.12% to 12%, their Sharpe ratios lie between 0.45 and 0.49, and their excess returns with respect to the market trading strategy vary from 1.25% to 2.13%. We refer to Banner et al. [6] for a detailed empirical study to explain these excess returns. □

The following example is motivated by Schied, Speiser, and Voloshchenko [82].

Example 2.4.2. Consider the function

$$G(\lambda, x) = \left(\sum_{i=1}^{d} (\alpha x_i + (1-\alpha)\lambda_i)^p\right)^{\frac{1}{p}}, \quad \lambda \in \mathbb{R}_+^d,\ x \in \Delta_+^d,$$

[3] To compute the Sharpe ratios of the market trading strategy and other functionally generated trading strategies, the one-year U.S. Treasury yields are used. The data of these yields can be downloaded from `https://www.federalreserve.gov`.

with constants $\alpha, p \in (0,1)$. Then G is concave.

For fixed constant $\delta > 0$, define the $\mathbb{R}^d_+$-valued moving average process $\Lambda(\cdot)$ by

$$\Lambda_i(\cdot) = \begin{cases} \frac{1}{\delta}\int_0^{\cdot} \mu_i(t)\mathrm{d}t + \frac{1}{\delta}\int_{\cdot-\delta}^{0} \mu_i(0)\mathrm{d}t & \text{on } [0,\delta) \\ \frac{1}{\delta}\int_{\cdot-\delta}^{\cdot} \mu_i(t)\mathrm{d}t & \text{on } [\delta,\infty) \end{cases},$$

for all $i \in \{1, \cdots, d\}$.

Write $\overline{\mu}(\cdot) = \alpha\mu(\cdot) + (1-\alpha)\Lambda(\cdot)$. Then by (2.5),

$$\begin{aligned} \Gamma^G(\cdot) = &-(1-\alpha)\sum_{j=1}^{d}\int_0^{\cdot}\left(\frac{G(\Lambda(t),\mu(t))}{\overline{\mu}_j(t)}\right)^{1-p}\mathrm{d}\Lambda_j(t) \\ &-\frac{\alpha^2(1-p)}{2}\sum_{i,j=1}^{d}\int_0^{\cdot}\left(\frac{G(\Lambda(t),\mu(t))}{\overline{\mu}_i(t)\overline{\mu}_j(t)}\right)^{1-p}\frac{1}{\sum_{v=1}^{d}(\overline{\mu}_v(t))^p}\mathrm{d}\left[\mu_i,\mu_j\right](t) \\ &+\frac{\alpha^2(1-p)}{2}\sum_{j=1}^{d}\int_0^{\cdot}\left(\frac{G(\Lambda(t),\mu(t))}{\overline{\mu}_j(t)}\right)^{1-p}\frac{1}{\overline{\mu}_j(t)}\mathrm{d}\left[\mu_j,\mu_j\right](t). \end{aligned}$$

Notice that G is not Lyapunov in general.

The trading strategies $\varphi(\cdot)$ and $\psi(\cdot)$, generated additively and multiplicatively by G, respectively, are given by

$$\varphi_i(\cdot) = G(\Lambda(\cdot),\mu(\cdot))\left(\frac{\alpha\left(\overline{\mu}_i(\cdot)\right)^p}{\overline{\mu}_i(\cdot)\sum_{v=1}^{d}(\overline{\mu}_v(\cdot))^p} - \sum_{j=1}^{d}\frac{\alpha\mu_j(\cdot)\left(\overline{\mu}_j(\cdot)\right)^p}{\overline{\mu}_j(\cdot)\sum_{v=1}^{d}(\overline{\mu}_v(\cdot))^p} + 1\right) + \Gamma^G(\cdot)$$

and

$$\psi_i(\cdot) = \left(\varphi_i(\cdot) - \Gamma^G(\cdot)\right)\exp\left(\int_0^{\cdot}\frac{\mathrm{d}\Gamma^G(t)}{G(\Lambda(t),\mu(t))}\right), \quad i \in \{1,\cdots,d\}.$$

The corresponding wealth processes $V^{\varphi}(\cdot)$ and $V^{\psi}(\cdot)$ can be derived from (2.15) and (2.22), respectively.

Consider the normalised regular function $\mathcal{G}$ given in (2.36) and the corresponding process $\Gamma^{\mathcal{G}}(\cdot)$ given in (2.37). By Theorem 2.2.4, if

$$\mathsf{P}\left[\Gamma^{\mathcal{G}}(T_*) > 1+\varepsilon\right] = \mathsf{P}\left[\Gamma^{G}(T_*) > G(\Lambda(0),\mu(0))(1+\varepsilon)\right] = 1,$$

then the trading strategy $\psi^{(c)}(\cdot)$, generated multiplicatively by $G^{(c)}$ given in (2.38) for some sufficiently large $c > 0$, is strong relative arbitrage over the investment horizon $[0, T_*]$.

To simulate the relative performance of $\varphi(\cdot)$ and $\psi^{(c)}(\cdot)$, we use the parameters $\delta = 250$ days and $p = 0.8$. Figure 2.6 shows $\Gamma^{\mathcal{G}}(\cdot)$ and the wealth processes $V^{\varphi}(\cdot)$ and $V^{\psi^{(0)}}(\cdot)$ without the effect of the moving average part, i.e., $\alpha = 1$. In this case, $\mathcal{G}$ is Lyapunov. The performance of $\varphi(\cdot)$ and $\psi^{(c)}(\cdot)$ is similar to that in Example 2.4.1, when the finite-variation process is chosen to be constant. Figure 2.7 presents the case when $\alpha = 0.6$. It can be observed that $\Gamma^{\mathcal{G}}(\cdot)$ increases slower when the moving average part is considered. Compared with the case that the moving average part is not included, the wealth processes $V^{\varphi}(\cdot)$ and $V^{\psi^{(0)}}(\cdot)$ also take smaller values in the long run. This is due to the fact that when α decreases, the volatility of $\overline{\mu}(\cdot)$ decreases as well. In this case, we trade slower, and the gains and losses will also be relatively less.

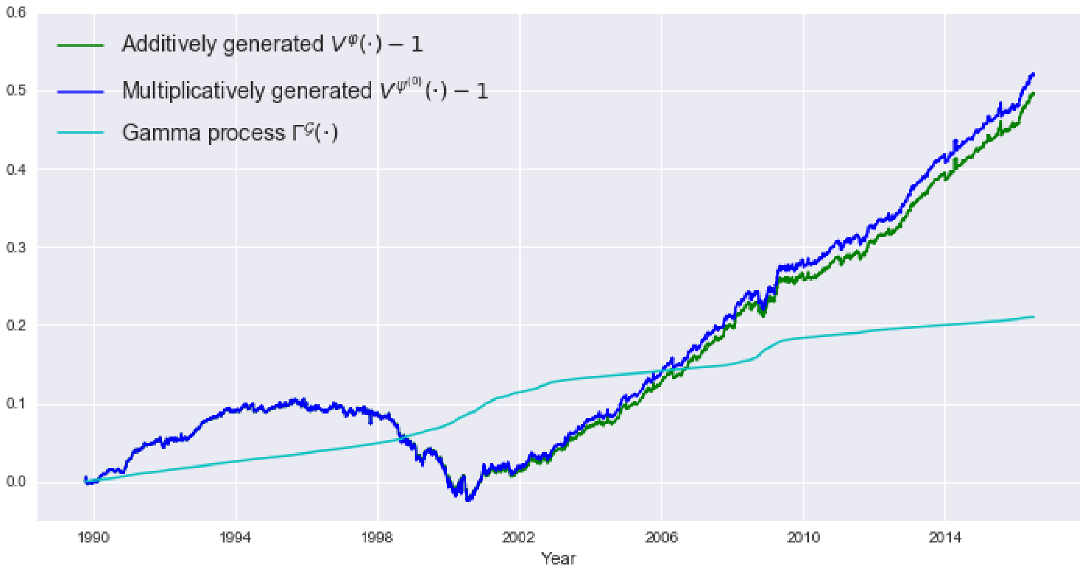

FIGURE 2.6: Gamma process $\Gamma^{\mathcal{G}}(\cdot)$ and relative wealth processes (minus 1) of both the additively and the multiplicatively generated trading strategies with $\delta = 250$ days, $p = 0.8$, and $\alpha = 1$.

For the four functionally generated trading strategies examined in this example, their averaged yearly returns range from 11.21% to 11.47%, their Sharpe ratios lie between 0.45 and 0.47, and their excess returns with respect to the market trading strategy vary from 1.34% to 1.6%. $\square$

2.5 Conclusion

Karatzas and Ruf [51] build a simple and intuitive structure by interpreting the portfolio generating functions G initiated by Fernholz [31] as Lyapunov functions. They formulate conditions for the existence of strong arbitrage relative to the market over appropriate time horizons. The purpose of this paper is to investigate the dependence of the portfolio generating functions G on an extra $\mathbb{R}^m$-valued, progressive, continuous process $\Lambda(\cdot)$ of finite variation on $[0, T]$, for all $T \geq 0$.

The results of the theoretical part in this chapter are illuminated by several examples and shown to work on empirical data using stocks from the S&P 500 index. The effects that different choices of $\Lambda(\cdot)$ have on the portfolio wealths are analyzed. Provided that the market undergoes an explicit trend of either increasing or decreasing market diversification, certain choices of $\Lambda(\cdot)$ are better than others.

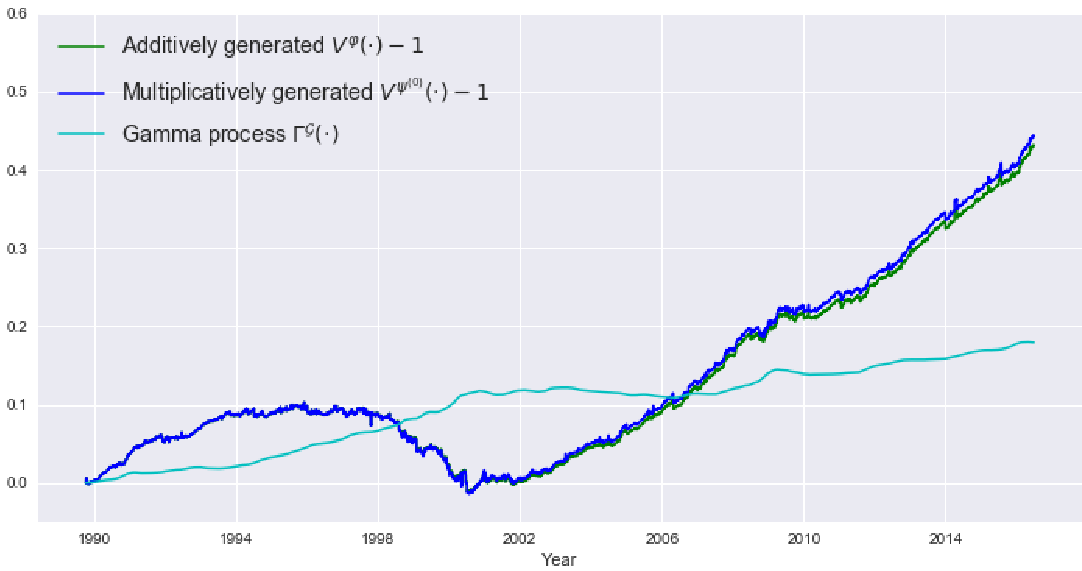

FIGURE 2.7: Gamma process $\Gamma^{\mathcal{G}}(\cdot)$ and relative wealth processes (minus 1) of both the additively and the multiplicatively generated trading strategies with $\delta = 250$ days, $p = 0.8$, and $\alpha = 0.6$.

2.6 Proofs of Theorems 2.1.1 and 2.1.2

2.6.1 Preliminaries

Before providing the proof of Theorem 2.1.1, we discuss some technical details.

Recall the open set $\mathcal{W}$ from (2.1) and consider a continuous function $g : \mathcal{W} \to \mathbb{R}$. Define a function $\overline{g} : \mathbb{R}^{m+d} \to \mathbb{R}$ by

$$\overline{g}(z) = \begin{cases} g(z), & \text{if } z \in \mathcal{W} \\ 0, & \text{if } z \notin \mathcal{W} \end{cases}.$$

Next, let $(g_{n_1,n_2})_{n_1,n_2 \in \mathbb{N}}$ be the family of functions $g_{n_1,n_2} : \mathcal{W} \to \mathbb{R}$ given by

$$g_{n_1,n_2}(\lambda, x) = \int_{\mathbb{R}^d} \eta_{n_2}(y) \int_{\mathbb{R}^m} \eta_{n_1}(u)\overline{g}(\lambda - u, x - y)\mathrm{d}u\mathrm{d}y, \tag{2.41}$$

for all $(\lambda, x) \in \mathcal{W}$, with $g_{n_1,n_2}(\lambda, x) = 0$ whenever the right hand side of (2.41) is not defined. Here in (2.41), for $z \in \mathbb{R}^l$ and $n \in \mathbb{N}$,

$$\eta_n(z) = \begin{cases} \beta n^l \exp\left(\frac{1}{n^2\|z\|_2^2 - 1}\right), & \text{if } \|z\|_2 < \frac{1}{n} \\ 0, & \text{if } \|z\|_2 \geq \frac{1}{n} \end{cases} \tag{2.42}$$

is used with the normalisation constant

$$\beta = \left(\int_{\mathbb{R}^l} \exp\left(\frac{1}{\|y\|_2^2 - 1}\right) \mathrm{d}y\right)^{-1},$$

independent of n.

Lemma 2.6.1. *Let $\overline{\mathcal{V}}$ denote any closed subset of $\mathcal{W}$. Consider a continuous function $g : \mathcal{W} \to \mathbb{R}$ and the mollification $(g_{n_1,n_2})_{n_1,n_2 \in \mathbb{N}}$ of g defined as in* (2.41)*.*

(i) We have

$$\lim_{n_2 \uparrow \infty} \lim_{n_1 \uparrow \infty} g_{n_1,n_2} = g.$$

(ii) For $n_1, n_2 \in \mathbb{N}$ large enough, $g_{n_1,n_2} \in C^\infty(\overline{\mathcal{V}})$.

(iii) If there exists a constant $L = L(\overline{\mathcal{V}}) \geq 0$ such that, for all $(\lambda_1, x), (\lambda_2, x) \in \overline{\mathcal{V}}$,

$$|g(\lambda_1, x) - g(\lambda_2, x)| \leq L\|\lambda_1 - \lambda_2\|_2,$$

then, for $n_1, n_2 \in \mathbb{N}$ large enough and all $(\lambda, x) \in \overline{\mathcal{V}}$, we have

$$\left| \frac{\partial g_{n_1,n_2}}{\partial \lambda_v}(\lambda, x) \right| \leq L, \quad v \in \{1, \cdots, m\}.$$

(iv) If $g \in C^{0,1}$, then, for all $(\lambda, x) \in \mathcal{W}$, we have

$$\lim_{n_2 \uparrow \infty} \lim_{n_1 \uparrow \infty} \frac{\partial g_{n_1,n_2}}{\partial x_i}(\lambda, x) = \frac{\partial g}{\partial x_i}(\lambda, x), \quad i \in \{1, \cdots, d\}.$$

(v) If $g \in C^{0,1}$ and if there exists a constant $L = L(\overline{\mathcal{V}}) \geq 0$ such that, for all $(\lambda, x_1), (\lambda, x_2) \in \overline{\mathcal{V}}$,

$$\left\| \frac{\partial g}{\partial x}(\lambda, x_1) - \frac{\partial g}{\partial x}(\lambda, x_2) \right\|_2 \leq L\|x_1 - x_2\|_2,$$

then, for $n_1, n_2 \in \mathbb{N}$ large enough and all $(\lambda, x) \in \overline{\mathcal{V}}$, we have

$$\left| \frac{\partial^2 g_{n_1,n_2}}{\partial x_i \partial x_j}(\lambda, x) \right| \leq L,$$

for all $i, j \in \{1, \cdots, d\}$.

Proof. For (i) and (ii), see Theorem 6 in Appendix C in Evans [23].

For (iii), observe that, for each $n_1, n_2 \in \mathbb{N}$ large enough and all $v \in \{1, \cdots, m\}$, (2.41) yields

$$\begin{aligned}
&\left| \frac{\partial g_{n_1,n_2}}{\partial \lambda_v}(\lambda, x) \right| = \left| \lim_{\delta \to 0} \frac{g_{n_1,n_2}(\lambda + \delta \mathbf{e}_v, x) - g_{n_1,n_2}(\lambda, x)}{\delta} \right| \\
&= \left| \lim_{\delta \to 0} \frac{1}{\delta} \int_{\mathbb{R}^d} \eta_{n_2}(y) \int_{\mathbb{R}^m} \eta_{n_1}(u) \big(\overline{g}(\lambda + \delta \mathbf{e}_v - u, x - y) - \overline{g}(\lambda - u, x - y) \big) \mathrm{d}u \mathrm{d}y \right| \\
&\leq \lim_{\delta \to 0} \frac{1}{\delta} \int_{\mathbb{R}^d} \eta_{n_2}(y) \int_{\mathbb{R}^m} \eta_{n_1}(u) \left| \overline{g}(\lambda + \delta \mathbf{e}_v - u, x - y) - \overline{g}(\lambda - u, x - y) \right| \mathrm{d}u \mathrm{d}y \\
&\leq \lim_{\delta \to 0} \frac{1}{\delta} \delta L \int_{\mathbb{R}^d} \eta_{n_2}(y) \int_{\mathbb{R}^m} \eta_{n_1}(u) \mathrm{d}u \mathrm{d}y = L,
\end{aligned}$$

for all $(\lambda, x) \in \overline{\mathcal{V}}$, where $\mathbf{e}_v$ is the unit vector in the v-th dimension.

For (iv), apply the dominated convergence theorem and (i) to $\partial g/\partial x_i$, for all $i \in \{1, \cdots, d\}$.

For (v), apply the dominated convergence theorem and a similar argument as in (iii). $\square$

The following lemma is an extension of Lemma 2 in Bouleau [13]. For a continuous function $g : \mathcal{W} \to \mathbb{R}$, consider its corresponding mollification $(g_{n_1,n_2})_{n_1,n_2 \in \mathbb{N}}$ defined as in (2.41).

Lemma 2.6.2. *If a continuous function $g : \mathcal{W} \to \mathbb{R}$ is concave in its second argument, then*

$$\lim_{n_2 \uparrow \infty} \lim_{n_1 \uparrow \infty} \frac{\partial g_{n_1,n_2}}{\partial x_i} = f_i, \quad i \in \{1, \cdots, d\},$$

for some measurable function $f_i : \mathcal{W} \to \mathbb{R}$, bounded on any compact $\overline{\mathcal{V}} \subset \mathcal{W}$.

Proof. Fix $i \in \{1, \cdots, d\}$. With the notation in (2.42), we have

$$\eta_n(z) = n^l \eta_1(nz), \quad z \in \mathbb{R}^l,\ n \in \mathbb{N}.$$

For $(\lambda, x) \in \mathcal{W}$ and $n_2 \in \mathbb{N}$ large enough, the definition of g_{n_1,n_2} in (2.41), the dominated convergence theorem, and Lemma 2.6.1(i)&(ii) yield

$$\begin{aligned}
\lim_{n_1 \uparrow \infty} \frac{\partial g_{n_1,n_2}}{\partial x_i}(\lambda, x) &= \lim_{n_1 \uparrow \infty} \int_{\mathbb{R}^d} \frac{\partial \eta_{n_2}}{\partial x_i}(x-y) \int_{\mathbb{R}^m} \eta_{n_1}(u) \overline{g}(\lambda - u, y) \mathrm{d}u \mathrm{d}y \\
&= \int_{\mathbb{R}^d} \frac{\partial \eta_{n_2}}{\partial x_i}(x-y) \lim_{n_1 \uparrow \infty} \int_{\mathbb{R}^m} \eta_{n_1}(u) \overline{g}(\lambda - u, y) \mathrm{d}u \mathrm{d}y \\
&= \int_{\mathbb{R}^d} \frac{\partial \eta_{n_2}}{\partial x_i}(x-y) \overline{g}(\lambda, y) \mathrm{d}y \\
&= -\int_{\mathbb{R}^d} \frac{\partial \eta_{n_2}}{\partial y_i}(y) \overline{g}(\lambda, x-y) \mathrm{d}y \\
&= \int_{\mathbb{R}^d} n_2 \frac{\partial \eta_1}{\partial y_i}(y) \overline{g}\left(\lambda, x + \frac{y}{n_2}\right) \mathrm{d}y \\
&= \int_{\mathbb{R}^d} \frac{\partial \eta_1}{\partial y_i}(y) n_2 \left(\overline{g}\left(\lambda, x + \frac{y}{n_2}\right) - \overline{g}(\lambda, x)\right) \mathrm{d}y.
\end{aligned}$$

Note that the last equality holds due to the fact that

$$\int_{\mathbb{R}^d} \frac{\partial \eta_1}{\partial y_i}(y) \mathrm{d}y = 0.$$

Next, for all $(\lambda, x) \in \mathcal{W}$ and $y \in \mathbb{R}^d$, define the one-sided directional partial derivative as

$$\nabla g(\lambda, x; y) = \lim_{n_2 \uparrow \infty} \frac{g(\lambda, x + y/n_2) - g(\lambda, x)}{1/n_2}.$$

Such ∇g exists according to Theorem 23.1 in Rockafellar [76]. Since g is concave in the second argument, it is locally Lipschitz in its second argument on $\mathcal{W}$ (see Theorem 10.4 in Rockafellar [76]). Hence, for each compact $\overline{\mathcal{V}} \subset \mathcal{W}$, there exists a constant $L = L(\overline{\mathcal{V}}) \geq 0$ such that $\nabla g(\lambda, x; y) \leq L$, for all $y \in \mathbb{R}^d$ and (λ, x) in the interior of $\overline{\mathcal{V}}$.

The statement now follows with

$$f_i(\lambda, x) = \int_{\mathbb{R}^d} \nabla \overline{g}(\lambda, x; y) \frac{\partial \eta_1}{\partial y_i}(y) \mathrm{d}y,$$

for all $(\lambda, x) \in \mathcal{W}$, by the dominated convergence theorem. □

Lemma 2.6.3. *Assume that $\mu(\cdot)$ has Doob-Meyer decomposition*

$$\mu(\cdot) = \mu(0) + M(\cdot) + V(\cdot),$$

where $M(\cdot)$ is a d-dimensional continuous local martingale and $V(\cdot)$ is a d-dimensional finite-variation process with $M(0) = V(0) = 0$. Moreover, suppose that,

(i) for some open $\mathcal{V} \subset \mathcal{W}$, we have

$$(\Lambda(\cdot), \mu(\cdot)) = (\Lambda(\cdot \wedge \tau), \mu(\cdot \wedge \tau)),$$

where

$$\tau = \inf \{t \geq 0;\ (\Lambda(t), \mu(t)) \notin \mathcal{V}\};$$

(ii) for some constant $\kappa \geq 0$, we have

$$\sum_{j=1}^{d} \left([M_j, M_j](\infty) + \int_0^\infty \mathrm{d}|V_j(t)| \right) + \sum_{v=1}^{m} \int_0^\infty \mathrm{d}|\Lambda_v(t)| \leq \kappa < \infty. \tag{2.43}$$

Let $(h_i)_{i \in \{1, \cdots, d\}}$ be a family of functions $h_i : \mathcal{V} \to \mathbb{R}$ and let $(h_i^{n_1, n_2})_{n_1, n_2 \in \mathbb{N}, i \in \{1, \cdots, d\}}$ be a family of doubly indexed sequences of uniformly bounded functions $h_i^{n_1, n_2} : \mathcal{V} \to \mathbb{R}$. If

$$\lim_{n_2 \uparrow \infty} \lim_{n_1 \uparrow \infty} h_i^{n_1, n_2} = h_i, \quad i \in \{1, \cdots, d\},$$

then there exist two random subsequences $(n_1^k)_{k \in \mathbb{N}}$ and $(n_2^k)_{k \in \mathbb{N}}$ with

$$\lim_{k \uparrow \infty} n_1^k = \infty = \lim_{k \uparrow \infty} n_2^k$$

such that

$$\lim_{k \uparrow \infty} \int_0^t \sum_{j=1}^{d} h_j^{n_1^k, n_2^k}(\Lambda(u), \mu(u)) \mathrm{d}\mu_j(u) = \int_0^t \sum_{j=1}^{d} h_j(\Lambda(u), \mu(u)) \mathrm{d}\mu_j(u), \quad \textit{a.s.}, \tag{2.44}$$

for all $t \geq 0$.

Proof. Fix $i \in \{1, \cdots, d\}$ and write

$$\Theta_i^{n_1,n_2}(\cdot) = h_i^{n_1,n_2}(\Lambda(\cdot), \mu(\cdot)) - h_i(\Lambda(\cdot), \mu(\cdot)).$$

By (2.43) and the bounded convergence theorem, we have

$$\begin{aligned} 0 &= \mathbb{E}\left[\lim_{n_2\uparrow\infty}\lim_{n_1\uparrow\infty}\int_0^\infty \left(\Theta_i^{n_1,n_2}(t)\right)^2 \mathrm{d}\left[M_i, M_i\right](t)\right] \\ &= \lim_{n_2\uparrow\infty}\lim_{n_1\uparrow\infty}\mathbb{E}\left[\int_0^\infty \left(\Theta_i^{n_1,n_2}(t)\right)^2 \mathrm{d}\left[M_i, M_i\right](t)\right] \\ &= \lim_{n_2\uparrow\infty}\lim_{n_1\uparrow\infty}\mathbb{E}\left[\left(\int_0^\infty \Theta_i^{n_1,n_2}(t)\mathrm{d}M_i(t)\right)^2\right], \end{aligned}$$

by Itô's isometry, and

$$0 = \lim_{n_2\uparrow\infty}\lim_{n_1\uparrow\infty}\mathbb{E}\left[\left(\int_0^\infty \left|\Theta_i^{n_1,n_2}(t)\right| \mathrm{d}|V_i(t)|\right)^2\right]. \tag{2.45}$$

Since $\int_0^\cdot \Theta_i^{n_1,n_2}(t)\mathrm{d}M_i(t)$ is a uniformly integrable martingale (as it is a local martingale with bounded quadratic variation), Doob's submartingale inequality yields

$$\mathbb{E}\left[\left(\sup_{t\ge 0}\left|\int_0^t \Theta_i^{n_1,n_2}(u)\mathrm{d}M_i(u)\right|\right)^2\right] \le 4\mathbb{E}\left[\left(\int_0^\infty \Theta_i^{n_1,n_2}(t)\mathrm{d}M_i(t)\right)^2\right],$$

which implies

$$0 = \lim_{n_2\uparrow\infty}\lim_{n_1\uparrow\infty}\mathbb{E}\left[\left(\sup_{t\ge 0}\left|\int_0^t \Theta_i^{n_1,n_2}(u)\mathrm{d}M_i(u)\right|\right)^2\right]. \tag{2.46}$$

Therefore, (2.45), (2.46), and the triangle inequality yield

$$0 = \lim_{n_2\uparrow\infty}\lim_{n_1\uparrow\infty}\mathbb{E}\left[\left(\sup_{t\ge 0}\left|\int_0^t \Theta_i^{n_1,n_2}(u)\mathrm{d}\mu_i(u)\right|\right)^2\right].$$

Write

$$E_i^{n_1,n_2} = \mathbb{E}\left[\left(\sup_{t\ge 0}\left|\int_0^t \Theta_i^{n_1,n_2}(u)\mathrm{d}\mu_i(u)\right|\right)^2\right], \quad n_1, n_2 \in \mathbb{N},$$

and

$$E_i = \lim_{n_2\uparrow\infty}\lim_{n_1\uparrow\infty} E_i^{n_1,n_2}.$$

For each $n_2 \in \mathbb{N}$, denote $E_i^{n_2} = \lim_{n_1\uparrow\infty} E_i^{n_1,n_2}$. Then we can find a subsequence $(n_1(n_2))_{n_2\in\mathbb{N}}$ of $\mathbb{N}$ with $n_1(n_2)\uparrow\infty$ as $n_2\uparrow\infty$ such that, for each $n_2 \in \mathbb{N}$,

$$\left|E_i^{n_1(n_2),n_2} - E_i^{n_2}\right| \le \frac{1}{n_2}.$$

Since the triangle inequality yields

$$\left|E_i^{n_1(n_2),n_2} - E_i\right| \leq \frac{1}{n_2} + |E_i^{n_2} - E_i| \to 0 \quad \text{as } n_2 \uparrow \infty,$$

we have

$$\lim_{n_2 \uparrow \infty} E_i^{n_1(n_2),n_2} = E_i = 0.$$

This implies

$$\lim_{n_2 \uparrow \infty} \sup_{t \geq 0} \left| \int_0^t \sum_{i=1}^d h_i^{n_1(n_2),n_2}(\Lambda(u), \mu(u)) \mathrm{d}\mu_i(u) - \int_0^t \sum_{i=1}^d h_i(\Lambda(u), \mu(u)) \mathrm{d}\mu_i(u) \right| = 0$$

in L^2. Since convergence in L^2 implies almost sure convergence of a subsequence, we can find a random subsequence $(n_2^k)_{k \in \mathbb{N}}$ of $\mathbb{N}$ with $n_2^k \uparrow \infty$ as $k \uparrow \infty$ such that (2.44) holds with $n_1^k = n_1(n_2^k)$. □

Lemma 2.6.4. *Fix $l \in \mathbb{N}$; let $\overline{\Lambda}(\cdot)$ be an l-dimensional continuous process of finite variation; let $(\overline{\Upsilon}_{u,n}(\cdot))_{u \in \{1,\cdots,l\}, n \in \mathbb{N}}$ be a family of processes with $(\overline{\Upsilon}_{u,n}(\cdot))_{n \in \mathbb{N}}$ uniformly bounded, for each $u \in \{1, \cdots, l\}$; and let $(\overline{\Theta}_n(\cdot))_{n \in \mathbb{N}}$ be a sequence of non-decreasing continuous processes. Define*

$$H_n(\cdot) = \int_0^{\cdot} \sum_{u=1}^l \overline{\Upsilon}_{u,n}(t) \mathrm{d}\overline{\Lambda}_u(t) + \overline{\Theta}_n(\cdot), \quad n \in \mathbb{N}.$$

If

$$\lim_{n \uparrow \infty} H_n(\cdot) = H(\cdot), \quad \textit{a.s.},$$

then $H(\cdot)$ is of finite variation.

Proof. The following steps are partially inspired by the proof of Lemma 3.3 in Abi Jaber, Bouchard, and Illand [1].

Since $\left(\overline{\Upsilon}_{1,n}(\cdot)\right)_{n \in \mathbb{N}}$ is uniformly bounded, the Komlós theorem (see Theorem 1.3 in Delbaen and Schachermayer [17]) yields the following. For each $n \in \mathbb{N}$, there exists a convex combination $\Upsilon_{1,n}^1(\cdot) \in \mathrm{Conv}(\overline{\Upsilon}_{1,k}(\cdot), k \geq n)$ such that $(\Upsilon_{1,n}^1(\cdot))_{n \in \mathbb{N}}$ converges to some adapted bounded process $\Upsilon_1(\cdot)$. More precisely, for each $n \in \mathbb{N}$, we can find some random integer $N_n \geq 0$ and $(w_n^k)_{n \leq k \leq N_n} \subset [0,1]$ such that

$$\sum_{k=n}^{N_n} w_n^k = 1 \quad \text{and} \quad \Upsilon_{1,n}^1(\cdot) = \sum_{k=n}^{N_n} w_n^k \overline{\Upsilon}_{1,k}(\cdot).$$

For each $n \in \mathbb{N}$, define

$$H_n^1(\cdot) = \sum_{k=n}^{N_n} w_n^k H_n(\cdot), \quad \Theta_n^1(\cdot) = \sum_{k=n}^{N_n} w_n^k \overline{\Theta}_k(\cdot), \quad \text{and} \quad \Upsilon_{u,n}^1(\cdot) = \sum_{k=n}^{N_n} w_n^k \overline{\Upsilon}_{u,k}(\cdot),$$

for all $u \in \{2, \cdots, l\}$.

Since $\lim_{n\uparrow\infty} H_n(\cdot) = H(\cdot)$, a.s., we have

$$\left|H_n^1(\cdot) - H(\cdot)\right| = \left|\sum_{k=n}^{N_n} w_n^k H_k(\cdot) - H(\cdot)\right| \leq \sum_{k=n}^{N_n} w_n^k \left|H_k(\cdot) - H(\cdot)\right| \to 0$$

as $n \uparrow \infty$, which implies $\lim_{n\uparrow\infty} H_n^1(\cdot) = H(\cdot)$, a.s. Besides, $\Theta_n^1(\cdot)$ is non-decreasing, as it is a convex combination of non-decreasing processes.

Since $(\Upsilon_{2,n}^1(\cdot))_{n\in\mathbb{N}}$ is also uniformly bounded, by the Komlós theorem again, for each $n \in \mathbb{N}$, there exists another convex combination $\Upsilon_{2,n}^2(\cdot) \in \mathrm{Conv}(\Upsilon_{2,k}^1(\cdot), k \geq n)$ such that $(\Upsilon_{2,n}^2(\cdot))_{n\in\mathbb{N}}$ converges to some adapted bounded process $\Upsilon_2(\cdot)$. With the same convex combination for each $n \in \mathbb{N}$, define $\Upsilon_{u,n}^2(\cdot)$, for all $u \in \{1, 3, \cdots, l\}$, $H_n^2(\cdot)$, and similarly $\Theta_n^2(\cdot)$. In particular, $(\Upsilon_{1,n}^2(\cdot))_{n\in\mathbb{N}}$ still converges to $\Upsilon_1(\cdot)$, as for each $n \in \mathbb{N}$, $\Upsilon_{1,n}^2(\cdot)$ is a convex combination of processes that converge to $\Upsilon_1(\cdot)$. Similarly, we have $\lim_{n\uparrow\infty} H_n^2(\cdot) = H(\cdot)$, a.s. Moreover, $\Theta_n^2(\cdot)$ is non-decreasing.

Iteratively, we construct sequences of processes $(\Upsilon_{u,n}^3(\cdot))_{n\in\mathbb{N}}, \cdots, (\Upsilon_{u,n}^l(\cdot))_{n\in\mathbb{N}}$, for each $u \in \{1, \cdots, l\}$, and processes $H_n^3(\cdot), \cdots, H_n^l(\cdot)$ and $\Theta_n^3(\cdot), \cdots, \Theta_n^l(\cdot)$ in the same manner. In particular, $(\Upsilon_{u,n}^l(\cdot))_{n\in\mathbb{N}}$ converges to some adapted bounded process Υ_u, for each $u \in \{1, \cdots, l\}$, and we have $\lim_{n\uparrow\infty} H_n^l(\cdot) = H(\cdot)$, a.s. Moreover, $\Theta_n^l(\cdot)$ is non-decreasing.

By the dominated convergence theorem, we have

$$\lim_{n\uparrow\infty} \int_0^{\cdot} \sum_{u=1}^{l} \Upsilon_{u,n}^l(t) \mathrm{d}\overline{\Lambda}_u(t) = \int_0^{\cdot} \sum_{u=1}^{l} \Upsilon_u(t) \mathrm{d}\overline{\Lambda}_u(t), \quad \text{a.s.,}$$

which is of finite variation. Therefore, we have

$$H(\cdot) = \lim_{n\uparrow\infty} H_n^l(\cdot) = \int_0^{\cdot} \sum_{u=1}^{l} \Upsilon_u(t) \mathrm{d}\overline{\Lambda}_u(t) + \lim_{n\uparrow\infty} \Theta_n^l(\cdot), \quad \text{a.s.}$$

Since $\Theta_n^l(\cdot)$ is non-decreasing and converges, it is of finite variation, which implies the assertion. □

2.6.2 Proof of Theorem 2.1.1

Proof of Theorem 2.1.1. Assume that the semimartingale $\mu(\cdot)$ has the Doob-Meyer decomposition

$$\mu(\cdot) = \mu(0) + M(\cdot) + V(\cdot),$$

where $M(\cdot)$ is a d-dimensional continuous local martingale and $V(\cdot)$ is a d-dimensional finite-variation process with $M(0) = V(0) = 0$.

Let $(\mathcal{W}_n)_{n\in\mathbb{N}}$ be a non-decreasing sequence of open sets such that the closure of $\mathcal{W}_n$ is in $\mathcal{W}$, for all $n \in \mathbb{N}$. For each $\kappa \in \mathbb{N}$, we consider the stopping time

$$\begin{aligned}\tau_\kappa = \inf \Bigg\{ & t \geq 0;\ (\Lambda(t), \mu(t)) \notin \mathcal{W}_\kappa \\ & \text{or } \sum_{i,j=1}^{d} [M_i, M_j](t) + \sum_{j=1}^{d} \int_0^t \mathrm{d}|V_j(u)| + \sum_{v=1}^{m} \int_0^t \mathrm{d}|\Lambda_v(u)| \geq \kappa \Bigg\}\end{aligned} \tag{2.47}$$

with $\inf\{\emptyset\} = \infty$. Since $(\Lambda(\cdot), \mu(\cdot)) \in \mathcal{W}$, we have $\lim_{\kappa\uparrow\infty} \tau_\kappa = \infty$, a.s. As $\bigcup_{\kappa\in\mathbb{N}}\{\tau_\kappa > t\} = \Omega$, for all $t \geq 0$, to prove that G is regular (Lyapunov), it is equivalent to show that G is regular (Lyapunov) for $\Lambda(\cdot \wedge \tau_\kappa)$ and $\mu(\cdot \wedge \tau_\kappa)$, for all $\kappa \in \mathbb{N}$. Hence, without loss of generality, let us assume that $(\Lambda(\cdot), \mu(\cdot)) = (\Lambda(\cdot \wedge \tau_\kappa), \mu(\cdot \wedge \tau_\kappa))$, for some $\kappa \in \mathbb{N}$.

Without loss of generality, assume that $a_{ij}(\cdot)$ is a predictable and uniformly bounded process, for all $i, j \in \{1, \cdots, d\}$, such that

$$[\mu_i, \mu_j](t) = \int_0^t a_{ij}(u)\mathrm{d}A(u) \leq \kappa, \quad t \geq 0,$$

where $A(\cdot) = \sum_{j=1}^{d} [\mu_j, \mu_j](\cdot)$. Here, the equality holds according to the Kunita-Watanabe inequality (see also Proposition 2.9 in Jacod and Shiryaev [44]) and the inequality due to (2.47).

Now, consider a mollification $(G_{n_1,n_2})_{n_1,n_2\in\mathbb{N}}$ of G defined as in (2.41). According to Lemma 2.6.1(ii), for $n_1, n_2 \in \mathbb{N}$ large enough, Itô's lemma applied to G_{n_1,n_2} yields

$$\begin{aligned}G_{n_1,n_2}(\Lambda(t), \mu(t)) = G_{n_1,n_2}(\Lambda(0), \mu(0)) &+ \int_0^t \sum_{j=1}^{d} \frac{\partial G_{n_1,n_2}}{\partial x_j}(\Lambda(u), \mu(u))\mathrm{d}\mu_j(u) \\ + \int_0^t \Upsilon_{0,n_1,n_2}(u)\mathrm{d}A(u) &+ \int_0^t \sum_{v=1}^{m} \Upsilon_{v,n_1,n_2}(u)\mathrm{d}\Lambda_v(u),\end{aligned} \tag{2.48}$$

for all $t \geq 0$, where

$$\Upsilon_{0,n_1,n_2}(t) = \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2 G_{n_1,n_2}}{\partial x_i \partial x_j}(\Lambda(t), \mu(t))a_{ij}(t)$$

and

$$\Upsilon_{v,n_1,n_2}(t) = \frac{\partial G_{n_1,n_2}}{\partial \lambda_v}(\Lambda(t), \mu(t)),$$

for all $v \in \{1, \cdots, m\}$.

For all $(\lambda, x) \in \mathcal{W}$ and $i \in \{1, \cdots, d\}$, if (bi) holds, Lemma 2.6.1(iv) yields

$$\lim_{n_2\uparrow\infty} \lim_{n_1\uparrow\infty} \frac{\partial G_{n_1,n_2}}{\partial x_i}(\lambda, x) = \frac{\partial G}{\partial x_i}(\lambda, x);$$

if (bii) holds, Lemma 2.6.2 yields

$$\lim_{n_2\uparrow\infty}\lim_{n_1\uparrow\infty}\frac{\partial G_{n_1,n_2}}{\partial x_i}(\lambda,x)=f_i(\lambda,x),$$

for some measurable function f_i. Moreover, thanks to (bi) or (bii), there exists a constant $L=L(\mathcal{W}_\kappa)\geq 0$ such that, for $n_1,n_2\in\mathbb{N}$ large enough,

$$\left|\frac{\partial G_{n_1,n_2}}{\partial x_i}\right|\leq L,\quad i\in\{1,\cdots,d\}.$$

This follows from the Lipschitz continuity of G on the closure of $\mathcal{W}_\kappa$ in the second argument and a similar reasoning as in the proof of Lemma 2.6.1(iii). Then by Lemma 2.6.3, there exist random subsequences $(n_1^k)_{k\in\mathbb{N}}$ and $(n_2^k)_{k\in\mathbb{N}}$ with

$$\lim_{k\uparrow\infty}n_1^k=\infty=\lim_{k\uparrow\infty}n_2^k$$

such that, if we write $G_k=G_{n_1^k,n_2^k}$, we have

$$\lim_{k\uparrow\infty}\int_0^t\sum_{j=1}^d\frac{\partial G_k}{\partial x_j}(\Lambda(u),\mu(u))\mathrm{d}\mu_j(u)=F(\Lambda(t),\mu(t)),\quad\text{a.s.},\tag{2.49}$$

for all $t\geq 0$, where

$$F(\Lambda(t),\mu(t))=\begin{cases}\int_0^t\sum_{j=1}^d\frac{\partial G}{\partial x_j}(\Lambda(u),\mu(u))\mathrm{d}\mu_j(u), & \text{if (bi) holds}\\ \int_0^t\sum_{j=1}^d f_j(\Lambda(u),\mu(u))\mathrm{d}\mu_j(u), & \text{if (bii) holds}\end{cases}.$$

To proceed, write

$$H_k(t)=G_k(\Lambda(0),\mu(0))-G_k(\Lambda(t),\mu(t))+\int_0^t\sum_{j=1}^d\frac{\partial G_k}{\partial x_j}(\Lambda(u),\mu(u))\mathrm{d}\mu_j(u),$$

for all $k\in\mathbb{N}$, and

$$H(t)=G(\Lambda(0),\mu(0))-G(\Lambda(t),\mu(t))+F(\Lambda(t),\mu(t)),\quad t\geq 0.$$

for all $t\geq 0$. Then, (2.48) with respect to the random subsequences $(n_1^k)_{k\in\mathbb{N}}$ and $(n_2^k)_{k\in\mathbb{N}}$ is of the form

$$H_k(t)=-\int_0^t\Upsilon_{0,k}(u)\mathrm{d}A(u)-\int_0^t\sum_{v=1}^m\Upsilon_{v,k}(u)\mathrm{d}\Lambda_v(u),\quad t\geq 0.$$

Note that by Lemma 2.6.1(i) and (2.49), $\lim_{k\uparrow\infty}H_k(t)=H(t)$, a.s., for all $t\geq 0$.

A measurable function G^{D} in Condition 1 of Definition 2.1.1 is chosen with components

$$G_i^{\mathrm{D}}\big(\lambda, x\big) = \begin{cases} \frac{\partial G}{\partial x_i}(\lambda, x), & \text{if (bi) holds} \\ f_i(\lambda, x), & \text{if (bii) holds} \end{cases}, \quad i \in \{1, \cdots, d\}.$$

Then, as $\Gamma^G(\cdot) = H(\cdot)$ according to (2.4), it is enough to show that $H(\cdot)$ is of finite variation in the following four cases.

Case 1.

Assume that (ai) and (bi) hold. Then by Lemma 2.6.1(iii)&(v), the processes $(\Upsilon_{0,k}(\cdot))_{k\in\mathbb{N}}$ and $(\Upsilon_{v,k}(\cdot))_{v\in\{1,\cdots,m\},k\in\mathbb{N}}$ are uniformly bounded. With $l = m+1$, we have

$$\overline{\Lambda}_v(\cdot) = \Lambda_v(\cdot) \quad \text{and} \quad (\overline{\Upsilon}_{v,k}(\cdot))_{k\in\mathbb{N}} = (\Upsilon_{v,k}(\cdot))_{k\in\mathbb{N}}, \quad v \in \{1, \cdots, m\},$$

and

$$\overline{\Lambda}_{m+1}(\cdot) = A(\cdot), \quad (\overline{\Upsilon}_{m+1,k}(\cdot))_{k\in\mathbb{N}} = (\Upsilon_{0,k}(\cdot))_{k\in\mathbb{N}}, \quad \text{and} \quad (\overline{\Theta}_k(\cdot))_{k\in\mathbb{N}} = 0.$$

Hence, Lemma 2.6.4 yields that $H(\cdot)$ is of finite variation on compact sets.

Case 2.

Assume that (ai) and (bii) hold. According to Lemma 2.6.1(iii), the processes $(\Upsilon_{v,k}(\cdot))_{v\in\{1,\cdots,m\},k\in\mathbb{N}}$ are uniformly bounded. Since G is concave in the second argument, for each $k \in \mathbb{N}$, G_k is also concave in the second argument. Using the negative semidefinite property of the Hessian of G_k and choosing the matrix-valued process $a(\cdot) = (a_{ij}(\cdot))_{i,j\in\{1,\cdots,d\}}$ to be symmetric and positive semidefinite, one can show that $\Upsilon_{0,k}(t) \le 0$, for all $t \ge 0$. This implies that the processes

$$\overline{\Theta}_k(\cdot) = -\int_0^{\cdot} \Upsilon_{0,k}(t)\mathrm{d}A(t), \quad k \in \mathbb{N},$$

are non-decreasing. Similar to Case 1, but now with $l = m$, Lemma 2.6.4 yields again that $H(\cdot)$ is of finite variation.

Case 3.

Assume that (aii) and (bi) hold. By Lemma 2.6.1(v), the process $(\Upsilon_{0,k}(\cdot))_{k\in\mathbb{N}}$ is uniformly bounded. As G is non-increasing in the v-th dimension of the first argument, so is G_k, for all $v \in \{1, \cdots, m\}$. Therefore, $\Upsilon_{v,k}(t) \le 0$, for all $t \ge 0$, as $\Lambda(\cdot)$ is non-decreasing in the v-th dimension, for all $v \in \{1, \cdots, m\}$. This implies that the processes

$$\overline{\Theta}_k(\cdot) = -\int_0^{\cdot} \sum_{v=1}^{m} \Upsilon_{v,k}(t)\mathrm{d}\Lambda_v(t), \quad k \in \mathbb{N},$$

are non-decreasing. Similar to above, Lemma 2.6.4 implies that $H(\cdot)$ is of finite variation.

Case 4.

Assume that (aii) and (bii) hold. With

$$\overline{\Theta}_k(\cdot) = -\int_0^{\cdot} \Upsilon_{0,k}(t)\mathrm{d}A(t) - \int_0^{\cdot} \sum_{v=1}^{m} \Upsilon_{v,k}(t)\mathrm{d}\Lambda_v(t), \quad k \in \mathbb{N},$$

Lemma 2.6.4 implies again that $H(\cdot)$ is of finite variation. It is clear that G is Lyapunov. □

2.6.3 Proof of Theorem 2.1.2

Proof of Theorem 2.1.2. The following steps are partially inspired by the proof of Theorem 3.8 in Karatzas and Ruf [51]. Recall the ranked market weight process $\boldsymbol{\mu}(\cdot)$ from (1.17).

According to Theorem 2.3 in Banner and Ghomrasni [7], for each $l \in \{1, \cdots, d\}$, one can find a measurable function $\boldsymbol{h}_l : \Delta^d \to (0, 1]$ and a finite-variation process $\boldsymbol{B}_l(\cdot)$ with $\boldsymbol{B}_l(0) = 0$ such that

$$\mu_{(l)}(\cdot) = \mu_{(l)}(0) + \int_0^{\cdot} \sum_{j=1}^{d} \boldsymbol{h}_l(\mu(t))\mathbf{1}_{\{\mu_{(l)}(t)=\mu_j(t)\}}\mathrm{d}\mu_j(t) + \boldsymbol{B}_l(\cdot). \tag{2.50}$$

Since $\boldsymbol{G}$ is regular for $\Lambda(\cdot)$ and $\boldsymbol{\mu}(\cdot)$, by Definition 2.1.1, there exist a measurable function $\boldsymbol{G}^{\mathrm{D}}$ and a finite-variation process $\Gamma^{\boldsymbol{G}}(\cdot)$ such that

$$\boldsymbol{G}(\Lambda(\cdot), \boldsymbol{\mu}(\cdot)) = \boldsymbol{G}(\Lambda(0), \boldsymbol{\mu}(0)) + \int_0^{\cdot} \sum_{l=1}^{d} \boldsymbol{G}_l^{\mathrm{D}}(\Lambda(t), \boldsymbol{\mu}(t))\mathrm{d}\mu_{(l)}(t) - \Gamma^{\boldsymbol{G}}(\cdot). \tag{2.51}$$

By (2.50), we have

$$\begin{aligned}\int_0^{\cdot} \sum_{l=1}^{d} \boldsymbol{G}_l^{\mathrm{D}}(\Lambda(t), \boldsymbol{\mu}(t))\mathrm{d}\mu_{(l)}(t) = &\int_0^{\cdot} \sum_{l=1}^{d} \boldsymbol{G}_l^{\mathrm{D}}(\Lambda(t), \boldsymbol{\mu}(t))\boldsymbol{h}_l(\mu(t))\mathbf{1}_{\{\mu_{(l)}(t)=\mu_i(t)\}}\mathrm{d}\mu_i(t) \\ &+ \int_0^{\cdot} \sum_{l=1}^{d} \boldsymbol{G}_l^{\mathrm{D}}(\Lambda(t), \boldsymbol{\mu}(t))\mathrm{d}\boldsymbol{B}_l(t).\end{aligned} \tag{2.52}$$

Now consider the measurable function $G^{\mathrm{D}} : \mathcal{W} \to \mathbb{R}^d$ with components

$$G_i^{\mathrm{D}}(\lambda, x) = \sum_{l=1}^{d} \boldsymbol{G}_l^{\mathrm{D}}(\lambda, \mathfrak{R}(x))\boldsymbol{h}_l(x)\mathbf{1}_{x_{(l)}=x_i}, \quad i \in \{1, \cdots, d\},$$

and the finite-variation process

$$\Gamma^{G}(\cdot) = \Gamma^{\boldsymbol{G}}(\cdot) - \int_0^{\cdot} \sum_{l=1}^{d} \boldsymbol{G}_l^{\mathrm{D}}(\Lambda(t), \boldsymbol{\mu}(t))\mathrm{d}\boldsymbol{B}_l(t).$$

Then (5.8) and (2.52), together with $G(\lambda, x) = \boldsymbol{G}(\lambda, \mathfrak{R}(x))$, yield (2.4), i.e., G is regular for $\Lambda(\cdot)$ and $\mu(\cdot)$. □

2.6.4 An alternative proof for a special case

The proof technique of Theorem VII.31 in Dellacherie and Meyer [18] suggests an alternative argument for the case that conditions (ai) and (bii) in Theorem 2.1.1 hold. We summarise these ideas in the following result.

Theorem 2.6.5. *If a function $f : \mathcal{W} \to \mathbb{R}$ is locally Lipschitz in the first argument and concave in the second argument, then the process $f(\Lambda(\cdot), \mu(\cdot))$ is a semimartingale.*

Proof. Assume that the semimartingale $\mu(\cdot)$ has the Doob-Meyer decomposition

$$\mu(\cdot) = \mu(0) + M(\cdot) + V(\cdot),$$

where $M(\cdot)$ is a d-dimensional continuous local martingale and $V(\cdot)$ is a d-dimensional finite-variation process with $M(0) = V(0) = 0$.

Let $(\mathcal{W}_n)_{n\in\mathbb{N}}$ be a non-decreasing sequence of open sets such that the closure of $\mathcal{W}_n$ is in $\mathcal{W}$, for all $n \in \mathbb{N}$. For each $\kappa \in \mathbb{N}$, we consider the stopping time τ_κ given in (2.47). Without loss of generality, let us assume again that $(\Lambda(\cdot), \mu(\cdot)) = (\Lambda(\cdot \wedge \tau_\kappa), \mu(\cdot \wedge \tau_\kappa))$, for some $\kappa \in \mathbb{N}$.

Since f is locally Lipschitz in both arguments (see Theorem 10.4 in Rockafellar [76]), we can find a Lipschitz constant L such that, for all $s, t \geq 0$ with $s \leq t$, we have

$$\begin{aligned}
&|f(\Lambda(t), \mu(t)) - f(\Lambda(s), \mu(0) + M(t) + V(s))| \\
&\quad\leq L\left(\sum_{v=1}^{m} |\Lambda_v(t) - \Lambda_v(s)| + \sum_{j=1}^{d} |V_j(t) - V_i(s)|\right) \\
&\quad\leq L\left(\sum_{v=1}^{m} \int_s^t |\mathrm{d}\Lambda_v(u)| + \sum_{j=1}^{d} \int_s^t |\mathrm{d}V_j(u)|\right).
\end{aligned} \tag{2.53}$$

Let

$$Z(\cdot) = -f(\Lambda(\cdot), \mu(\cdot)) + L\left(\sum_{v=1}^{m} \int_0^{\cdot} |\mathrm{d}\Lambda_v(t)| + \sum_{j=1}^{d} \int_0^{\cdot} |\mathrm{d}V_j(t)|\right),$$

then $Z(\cdot)$ is bounded. Hence we have

$$\begin{aligned}
\mathbb{E}\left[Z(t) - Z(s)|\mathcal{F}(s)\right] &= \mathbb{E}\left[f(\Lambda(s), \mu(s)) - f(\Lambda(s), \mu(0) + M(t) + V(s))|\mathcal{F}(s)\right] \\
&\quad + \mathbb{E}\Bigg[f(\Lambda(s), \mu(0) + M(t) + V(s)) - f(\Lambda(t), \mu(t)) \\
&\quad + L\left(\sum_{v=1}^{m} \int_s^t |\mathrm{d}\Lambda_v(u)| + \sum_{j=1}^{d} \int_s^t |\mathrm{d}V_j(u)|\right)\Bigg|\mathcal{F}(s)\Bigg] \\
&\geq \mathbb{E}\left[f(\Lambda(s), \mu(s)) - f(\Lambda(s), \mu(0) + M(t) + V(s))|\mathcal{F}(s)\right] \geq 0,
\end{aligned}$$

where the first inequality is by (2.53) and the second inequality holds by Jensen's inequality. Therefore, $Z(\cdot)$ is a submartingale, which makes $f(\Lambda(\cdot), \mu(\cdot))$ a semimartingale. □

Chapter 3

The Impact of Proportional Transaction Costs on Systematically Generated Portfolios

Although often neglected in portfolio analysis for sake of simplicity, transaction costs matter significantly for portfolio performance. Even small proportional transaction costs can have a large negative effect, especially when trades are made to rebalance the portfolio in a relatively high frequency. Hence, one should at least test the performance of a given portfolio when transaction costs are imposed, even if transaction costs are not explicitly taken into account while constructing the portfolio.

In this chapter, which is based on Ruf and Xie [81], we examine the effects of imposing transaction costs on systematically generated portfolios, for example, functionally generated portfolios. Such portfolios play a significant role in Stochastic Portfolio Theory; see Fernholz [26]. Ruf and Xie [80] and Karatzas and Kim [50] demonstrate empirically that functionally generated portfolios outperform the market portfolio in the absence of transaction costs. To explore whether or to what extent this result still holds when transaction costs are imposed, we empirically examine the performance of four portfolios. These are the index tracking portfolio, the equally-weighted portfolio, the entropy-weighted portfolio, and the diversity-weighted portfolio. We consider different configurations including trading frequency, transaction cost rate, constituent list size, and renewing frequency. For the diversity-weighted portfolio, we also propose a method to smooth transaction costs. Wong [92] indicates an alternative approach, namely to adjust the trading frequency based on certain information-theoretic quantities.

When backtesting the portfolios with historical data, the index tracking portfolio is used as benchmark. In the absence of transaction costs, the equally-weighted, the entropy-weighted, and the diversity-weighted portfolios outperform the index tracking portfolio. In particular, the equally-weighted portfolio performs better than any other portfolio under the same configuration. When proportional transaction costs of 0.5% are imposed, however, the equally-weighted portfolio underperforms all other portfolios. The entropy-weighted and the diversity-weighted portfolios still outperform the benchmark under appropriate trading frequencies and constituents list sizes with yearly excess returns around 1bp to 4bp.

The following is an outline of this chapter. Section 3.1 presents a literature review on transaction costs in equity trading from both theoretical and empirical aspects. Section 3.2 proposes a framework of backtesting portfolio performance in the presence of transaction costs. In particular, Subsection 3.2.1 incorporates proportional transaction costs when rebalancing a portfolio. Subsection 3.2.2 provides some practical considerations and details when backtesting portfolio performance. Section 3.3 empirically examines the performance of several different portfolios under various configurations. A method to smooth transaction costs is also provided in Section 3.3. Section 3.4 concludes.

3.1 Literature review

Within the framework of portfolio selection and dynamic trading introduced by Merton [63, 64], there is a large amount of research that takes transaction costs into consideration. The most common assumption is that trading costs occur proportionally to the total volume traded. We shall now provide some pointers to this literature. Following Merton's construction, Magill and Constantinides [59] are among the first to study the impact of proportional transaction costs in portfolio choice. Based on a financial market model with one risky asset and another non-risky asset, Taksar, Klass, and Assaf [87] analyse the optimal portfolio selection problem when proportional transaction costs are imposed. Proportional transaction costs are also addressed in Davis and Norman [15], who provide a numerical method to solve a related free boundary problem. Muthuraman [66] and Muthuraman and Zha [67] develop further computational schemes to solve the portfolio optimisation problem. Moreover, Kallsen and Muhle-Karbe [47] and Czichowsky and Schachermayer [14] use duality theory for the portfolio optimisation problem with proportional transaction costs by means of shadow price processes. In the presence of general transaction costs, liquidity costs, and market impact, Zhang et al. [96] provide a simulation-and-regression based approach to solve the dynamic portfolio optimisation problem. In general, we refer to Guasoni and Muhle-Karbe [42] and Muhle-Karbe, Reppen, and Soner [65] for an overview of the transaction cost literature evolved since Magill and Constantinides [59]. Most of this literature focuses on the case of one risky asset only. For a discussion of transaction costs in the presence of several risky assets, we refer to Muthuraman and Zha [67], Bichuch and Shreve [11], and Possamaï, Mete Soner, and Touzi [75].

Other types of transaction costs are also studied. Transaction costs in equity markets are often modeled as bid-ask spread. We refer to Amihud and Mendelson [3], as well as Bion-Nadal [12], for a mathematical framework of bid-ask dynamic pricing in financial markets. Novy-Marx and Velikov [68] also study the bid-ask spread empirically and propose several transaction cost mitigation strategies; see Novy-Marx and

Velikov [69] as well. The total transaction costs are split into a fixed part and a proportional part in Eastham and Hastings [21] and Korn [55]. With a new formulation of the consumption and portfolio choice model of Merton [64], Duffie and Sun [20] study the impact of lump sum transaction costs proportional to the portfolio value. A quadratic transaction cost form is used in Heaton and Lucas [43], Grinold [41], and Gârleanu and Pedersen [39] to reflect the impact of trading on the average stock price. Kabanov [46] models transaction costs as random processes in a general semimartingale model of a currency market. Then the set of hedging endowments within a multi-asset case to optimise the expected utility from terminal wealth is studied. A dynamic equilibrium model of trading volume is purposed by Lo, Mamaysky, and Wang [58], in which agents face fixed transaction costs.

The third stream of literature focuses on the empirical analysis and estimates transaction costs with data from actual equity markets. Keim and Madhavan [54] use order-level data on equity transactions by a sample of institutional traders with different investment styles and order submission strategies. They examine the magnitude and determinants of transaction costs and propose that institutional traders in exchange-listed stocks have lower costs than in comparable Nasdaq stocks. Lesmond, Ogden, and Trzcinka [57] develop a model based on expected and actual stock returns to estimate transaction costs for numerous stock exchanges. De Roon, Nijman, and Werker [16] demonstrate empirically that diversification benefits in emerging markets disappear when investors face short sales constraints or small transaction costs. Fong, Holden, and Trzcinka [37] document transaction costs in over 40 developed and emerging country exchanges. Jones [45] estimates the annual proportional costs of aggregate equity trading, with the sum of half-spreads and one-way commissions, multiplied by annual turnover. This paper is based on bid-ask spreads on Dow Jones stocks and an annual estimate of the weighted-average commission rate for trading NYSE stocks. Based on daily prices of the DJIA index from year 1897 to year 2011, Bajgrowicz and Scaillet [5] show that the performance of technical trading rules is completely offset when incorporating with low transaction costs. With measures of transaction costs for 19 frontier markets, Marshall, Nguyen, and Visaltanachoti [61] also investigate the link between transaction costs and diversification benefits and show a similar result in frontier markets. Olivares-Nadal and DeMiguel [70] give a theoretical proof of the equivalence between the portfolio problem with transaction costs and problems designed to alleviate the impact of estimation error. Then they include estimation error to calibrate transaction costs and propose a data-driven approach to the portfolio optimisation problem.

The index tracking problem also involves a transaction cost analysis. Here, the goal is to construct a portfolio to approximate the performance of the target index. It is universally recognised that there is a tradeoff between reducing the transaction costs imposed from rebalancing the index-tracking portfolio and maintaining the accuracy in tracking the target index. For literature relating to this topic, we refer to Strub and Baumann [86] and their references.

3.2 Backtesting in the presence of transaction costs

3.2.1 Incorporating transaction costs into wealth dynamics

We shall study the performance of long-only stock portfolios that are rebalanced discretely. The market is not assumed to be frictionless; transaction costs are imposed when we trade in the market to rebalance the portfolios. The portfolios are constructed in such a way that their weights match given target weights after paying transaction costs. This construction is more rigid than the one in Gârleanu and Pedersen [39], for example, where the portfolio weights may deviate from the target weights.

To be more specific, recall that we are facing a market with $d \geq 2$ stocks. Denote the amount of currency invested in each stock by $\Xi(\cdot) = (\Xi_1(\cdot), \cdots, \Xi_d(\cdot))'$ and the total amount invested in a portfolio by

$$V(\cdot) = \sum_{i=1}^{d} \Xi_i(\cdot) \geq 0.$$

Note that $\Xi_i(\cdot) = \pi_i(\cdot)V(\cdot)$, for all $i \in \{1, \cdots, d\}$, where $\pi(\cdot) = (\pi_1(\cdot), \cdots, \pi_d(\cdot))'$ is the portfolio weight process defined as in Definition 1.2.2.

Assume that trading stocks involves proportional transaction costs at a time-invariant rate tc^{b} (tc^{s}), with $0 \leq \mathrm{tc}^{\mathrm{b}}, \mathrm{tc}^{\mathrm{s}} < 1$ for buying (selling) a stock. This means that the sale of one unit of currency of a stock nets only $(1 - \mathrm{tc}^{\mathrm{s}})$ units of currency in cash, while buying one unit of currency of a stock costs $\left(1 + \mathrm{tc}^{\mathrm{b}}\right)$ units of currency.

Let us now consider how to trade the stocks in order to match the target weights when transaction costs are imposed. To begin, let us focus on trading at a specific time t. When rebalancing the portfolio at time t, we know the wealth $\Xi(t-)$ invested in each stock and hence the total wealth of the portfolio $V(t-) = \sum_{i=1}^{d} \Xi_i(t-)$ (exclusive of dividends). We also know the dividends paid at time $t-$, their total denoted by $D(t-) \geq 0$.

Given target weights π, we require $\pi(t) = \pi$ after the portfolio is rebalanced at time t. After trading, the wealth $\Xi(t)$ invested in each stock in the portfolio satisfies

$$\Xi_j(t) = \pi_j(t) \sum_{i=1}^{d} \Xi_i(t), \quad j \in \{1, \cdots, d\}. \tag{3.1}$$

We provide details about how to compute $\Xi(t)$ later in this subsection.

As the portfolio needs to be self-financing, the amount of currency used to buy extra stocks should be exactly the amount of currency obtained from selling redundant stocks plus the dividends if there are any. This yields

$$\left(1 + \mathrm{tc}^{\mathrm{b}}\right) \sum_{i=1}^{d} (\Xi_i(t) - \Xi_i(t-))^+ = (1 - \mathrm{tc}^{\mathrm{s}}) \sum_{i=1}^{d} (\Xi_i(t-) - \Xi_i(t))^+ + D(t-). \tag{3.2}$$

The total transaction costs imposed from trading stocks at time t are computed by

$$\mathrm{TC}(t) = \mathrm{tc}^{\mathrm{b}} \sum_{i=1}^{d} \left(\Xi_i(t) - \Xi_i(t-)\right)^{+} + \mathrm{tc}^{\mathrm{s}} \sum_{i=1}^{d} \left(\Xi_i(t-) - \Xi_i(t)\right)^{+}. \tag{3.3}$$

Therefore, the total wealth of the portfolio at time t, given by $V(t) = \sum_{i=1}^{d} \Xi_i(t)$, satisfies

$$V(t) = V(t-) + D(t-) - \mathrm{TC}(t).$$

Method of computing $\Xi(t)$

In the following, we propose a method to compute $\Xi(t)$, given $\Xi(t-)$, $D(t-)$, and the target weights π. Throughout this section, we assume

$$V(t-) > 0, \quad D(t-) \geq 0, \quad \sum_{i=1}^{d} \pi_i = 1, \quad \pi_j \geq 0, \quad \text{and} \quad \Xi_j(t-) \geq 0,$$

for all $j \in \{1, \cdots, d\}$.

To begin with, (3.1) implies that $\Xi(t)$ is of the form

$$\Xi_j(t) = cV(t-)\pi_j(t), \quad j \in \{1, \cdots, d\}, \tag{3.4}$$

for some $c > 0$. Note that if the market is frictionless, i.e., if $\mathrm{tc}^{\mathrm{b}} = \mathrm{tc}^{\mathrm{s}} = 0$, and if there are no dividends paid at time $t-$, i.e., if $D(t-) = 0$, then $V(t) = V(t-)$ and $c = 1$. When transaction costs are imposed, we shall use the constraint (3.2) to determine c.

To make headway, define

$$\widehat{D} = \frac{D(t-) + (1 - \mathrm{tc}^{\mathrm{s}}) \sum_{i=1}^{d} \Xi_i(t-) \mathbf{1}_{\pi_i(t)=0}}{V(t-)} \tag{3.5}$$

and

$$c_j = \frac{\pi_j(t-)}{\pi_j(t)} \mathbf{1}_{\pi_j(t)>0}, \quad j \in \{1, \cdots, d\}.$$

Then dividing both sides of (3.2) by $V(t-)$ yields

$$\left(1 + \mathrm{tc}^{\mathrm{b}}\right) \sum_{i=1}^{d} (c - c_i)^{+} \pi_i(t) = (1 - \mathrm{tc}^{\mathrm{s}}) \sum_{i=1}^{d} (c_i - c)^{+} \pi_i(t) + \widehat{D}. \tag{3.6}$$

Note that the LHS of (3.6) is a continuous function of c and strictly increasing from 0 to ∞, as c changes from $\min_{i \in \{1,\cdots,d\}} c_i$ to ∞. Moreover, the RHS of (3.6) is a continuous function of c strictly decreasing from ∞ to $\widehat{D} \geq 0$, as c changes from $-\infty$ to $\max_{i \in \{1,\cdots,d\}} c_i$, and equals $\widehat{D}$ afterwards, as c changes from $\max_{i \in \{1,\cdots,d\}} c_i$ to ∞. Hence, both sides of (3.6) as functions of c must intersect at some unique point, i.e., a unique solution exists for (3.6). To proceed, define

$$\widehat{D}_j = \left(1+\mathrm{tc}^{\mathrm{b}}\right)\sum_{i=1}^{d}\left(c_j-c_i\right)^{+}\pi_i(t) - \left(1-\mathrm{tc}^{\mathrm{s}}\right)\sum_{i=1}^{d}\left(c_i-c_j\right)^{+}\pi_i(t), \qquad (3.7)$$

for all $j \in \{1,\cdots,d\}$. We are now ready to provide an expression for the unknown constant c.

Proposition 3.2.1. *Recall that* (3.5) *and* (3.7) *imply* $\widehat{D} \geq 0$ *and* $\min_{i\in\{1,\cdots,d\}} \widehat{D}_i \leq 0$. *Hence,*

$$j = \underset{i\in\{1,\cdots,d\}}{\arg\max}\left\{\widehat{D}_i; \widehat{D}_i \leq \widehat{D}\right\} \qquad (3.8)$$

is well-defined. Then

$$c = \frac{\left(1+\mathrm{tc}^{\mathrm{b}}\right)\sum_{i=1}^{d} c_i\pi_i(t)\mathbf{1}_{c_i\leq c_j} + \left(1-\mathrm{tc}^{\mathrm{s}}\right)\sum_{i=1}^{d}\pi_i(t-)\mathbf{1}_{c_i>c_j} + \widehat{D}}{\left(1+\mathrm{tc}^{\mathrm{b}}\right)\sum_{i=1}^{d}\pi_i(t)\mathbf{1}_{c_i\leq c_j} + \left(1-\mathrm{tc}^{\mathrm{s}}\right)\sum_{i=1}^{d}\pi_i(t)\mathbf{1}_{c_i>c_j}} \qquad (3.9)$$

solves (3.6) *uniquely.*

Proof. By the definition of $\widehat{D}_j$ given in (3.7) and by some basic computations, (3.9) is equivalent to

$$c = c_j + \frac{\widehat{D}-\widehat{D}_j}{\left(1+\mathrm{tc}^{\mathrm{b}}\right)\sum_{i=1}^{d}\pi_i(t)\mathbf{1}_{c_i\leq c_j} + \left(1-\mathrm{tc}^{\mathrm{s}}\right)\sum_{i=1}^{d}\pi_i(t)\mathbf{1}_{c_i>c_j}},$$

which implies $\mathbf{1}_{c_i\leq c} \geq \mathbf{1}_{c_i\leq c_j}$, for all $i \in \{1,\cdots,d\}$.

In the case $\max_{i\in\{1,\cdots,d\}} \widehat{D}_i \leq \widehat{D}$, we have $\mathbf{1}_{c_i\leq c_j} = 1$, hence $\mathbf{1}_{c_i\leq c} \leq \mathbf{1}_{c_i\leq c_j}$, for all $i \in \{1,\cdots,d\}$. In the case $\max_{i\in\{1,\cdots,d\}} \widehat{D}_i > \widehat{D}$, define

$$j' = \underset{i\in\{1,\cdots,d\}}{\arg\min}\left\{\widehat{D}_i; \widehat{D}_i > \widehat{D}\right\}.$$

Then (3.9) is equivalent to

$$\begin{aligned} c &= \frac{\left(1+\mathrm{tc}^{\mathrm{b}}\right)\sum_{i=1}^{d} c_i\pi_i(t)\mathbf{1}_{c_i< c_{j'}} + \left(1-\mathrm{tc}^{\mathrm{s}}\right)\sum_{i=1}^{d}\pi_i(t-)\mathbf{1}_{c_i\geq c_{j'}} + \widehat{D}}{\left(1+\mathrm{tc}^{\mathrm{b}}\right)\sum_{i=1}^{d}\pi_i(t)\mathbf{1}_{c_i< c_{j'}} + \left(1-\mathrm{tc}^{\mathrm{s}}\right)\sum_{i=1}^{d}\pi_i(t)\mathbf{1}_{c_i\geq c_{j'}}} \\ &= c_{j'} + \frac{\widehat{D}-\widehat{D}_{j'}}{\left(1+\mathrm{tc}^{\mathrm{b}}\right)\sum_{i=1}^{d}\pi_i(t)\mathbf{1}_{c_i< c_{j'}} + \left(1-\mathrm{tc}^{\mathrm{s}}\right)\sum_{i=1}^{d}\pi_i(t)\mathbf{1}_{c_i\geq c_{j'}}}, \end{aligned}$$

which implies $\mathbf{1}_{c_i>c} \geq \mathbf{1}_{c_i>c_j}$, for all $i \in \{1,\cdots,d\}$. All in all, we have shown $\mathbf{1}_{c_i\leq c} = \mathbf{1}_{c_i\leq c_j}$, for all $i \in \{1,\cdots,d\}$.

Define next

$$\begin{aligned} \Pi^{\mathrm{b}} &= \left(1+\mathrm{tc}^{\mathrm{b}}\right)\sum_{i=1}^{d}\pi_i(t)\mathbf{1}_{c_i\leq c_j}, \quad \Pi^{\mathrm{s}} = \left(1-\mathrm{tc}^{\mathrm{s}}\right)\sum_{i=1}^{d}\pi_i(t)\mathbf{1}_{c_i>c_j}, \\ \overline{\Pi}^{\mathrm{b}} &= \left(1+\mathrm{tc}^{\mathrm{b}}\right)\sum_{i=1}^{d}c_i\pi_i(t)\mathbf{1}_{c_i\leq c_j}, \quad \overline{\Pi}^{\mathrm{s}} = \left(1-\mathrm{tc}^{\mathrm{s}}\right)\sum_{i=1}^{d}\pi_i(t-)\mathbf{1}_{c_i>c_j}. \end{aligned}$$

Hence, after inserting c by (3.9) into (3.6), the LHS of (3.6) becomes

$$\mathrm{LHS} = c\Pi^{\mathrm{b}} - \overline{\Pi}^{\mathrm{b}} = \frac{\Pi^{\mathrm{b}}\overline{\Pi}^{\mathrm{s}} - \Pi^{\mathrm{s}}\overline{\Pi}^{\mathrm{b}} + \Pi^{\mathrm{b}}\widehat{D}}{\Pi^{\mathrm{b}} + \Pi^{\mathrm{s}}},$$

and the RHS of (3.6) becomes

$$\mathrm{RHS} = \overline{\Pi}^{\mathrm{s}} - c\Pi^{\mathrm{s}} + \widehat{D} = \frac{\Pi^{\mathrm{b}}\overline{\Pi}^{\mathrm{s}} - \Pi^{\mathrm{s}}\overline{\Pi}^{\mathrm{b}} - \Pi^{\mathrm{s}}\widehat{D}}{\Pi^{\mathrm{b}} + \Pi^{\mathrm{s}}} + \widehat{D} = \mathrm{LHS}.$$

Therefore, c defined by (3.9) indeed solves (3.6). □

Remark 6. In practice, we can apply both numerical and analytical methods to find the constant c. As suggested by (3.6), to find c numerically, we can simply search for the minimum of the function

$$c \mapsto \left| \left(1 + \mathrm{tc}^{\mathrm{b}}\right) \sum_{i=1}^{d} (c - c_i)^+ \pi_i(t) - (1 - \mathrm{tc}^{\mathrm{s}}) \sum_{i=1}^{d} (c_i - c)^+ \pi_i(t) - \widehat{D} \right|.$$

Alternatively, by determining the index j given by (3.8), we can apply Proposition 3.2.1 to compute c analytically.

If the analytical approach is implemented, we can speed up the algorithm by making the following observations. We expect the value of c not to be far away from 1, which is precisely the value in the case of no transaction costs and no dividends. As suggested by the proof of Proposition 3.2.1, the family $(\widehat{D}_i)_{i\in\{1,\cdots,d\}}$ has the same ranking as $(c_i)_{i\in\{1,\cdots,d\}}$. Therefore, we proceed by ranking all c_i's in ascending order and comparing $\widehat{D}_k$ with $\widehat{D}$, where

$$k = \underset{i\in\{1,\cdots,d\}}{\arg\max} \left\{c_i; c_i \leq 1\right\}.$$

If $\widehat{D}_k = \widehat{D}$, then $j = k$ and we are done. If $\widehat{D}_k > \widehat{D}$, then we repeatedly compute $\widehat{D}_i$ corresponding to a smaller $c_i < c_k$ each time until we find the exact index j. If $\widehat{D}_k < \widehat{D}$, then we simply go the other way around. □

Proposition 3.2.1 is applied to determine the constant c used in (3.4) in order to compute $\Xi(t)$. Note that, in this subsection, we take $\Xi(t-)$ and $D(t-)$ as given. In the next subsection, we discuss how to compute $\Xi(t-)$ and $D(t-)$ from the data.

3.2.2 Practical considerations

For the preparation of the empirical study in the next section, we now introduce the method used to backtest the portfolio performance.

To begin with, assume that we are given the total market capitalizations and the daily returns for all stocks; denote these processes by $S(\cdot) = (S_1(\cdot), \cdots, S_d(\cdot))'$ and $r(\cdot) = (r_1(\cdot), \cdots, r_d(\cdot))'$, respectively. Assume that there are in total N days. For all $l \in \{1, \cdots, N\}$, let t_l denote the end of day l, at which the end of day total market capitalizations and the daily returns for day l are available. Moreover, if we trade on day l, then we call day l a trading day and the trade is made at time t_l.

Now focus on a specific trading day l with $l \in \{1, \cdots, N\}$ and fix $i \in \{1, \cdots, d\}$ for the moment. In Subsection 3.2.1, given $\Xi(t_l-)$ and $D(t_l-)$, as well as the target weights specified by the corresponding portfolio at time t_l, we have shown how to compute $\Xi(t_l)$. In the following, we show how to obtain $\Xi(t_l-)$ and $D(t_l-)$.

The daily return $r_i(t_l)$ includes the dividends of stock i if there are any. We decompose the daily return $r_i(t_l)$ into two parts: the dividend rate $r_i^D(t_l)$ and the realised rate $r_i^R(t_l)$. The dividend rate $r_i^D(t_l)$ is computed as

$$r_i^D(t_l) = \max\left\{1 + r_i(t_l) - \frac{S_i(t_l)}{S_i(t_{l-1})}, 0\right\} \tag{3.10}$$

and yields the amount of dividends received at time t_l for each unit of currency invested in stock i at time t_{l-1}.[1] The realised rate $r_i^R(t_l)$ is computed as

$$r_i^R(t_l) = r_i(t_l) - r_i^D(t_l)$$

and yields the units of currency held in stock i at time t_l for each unit of currency invested in stock i at time t_{l-1}.

The maximum is used in (3.10) to make sure that the dividend rate is nonnegative. Indeed, occasionally the data may suggest $S_i(t_{l-1})(1 + r_i(t_l)) < S_i(t_l)$. This can happen, for example, when company i issues extra stocks at time t_l. In this case, we simply assume that there are no dividends paid at time t_l.

A special situation requires us to pay extra attention. A few times, some stock i is delisted from the market at time t_l, for example, due to bankruptcy or merger. In this case, we still have data for $r_i(t_l)$, but not for $S_i(t_l)$. To deal with this situation, we assume that there are no dividends paid in stock i at time t_l. As a result, we have $r_i^D(t_l) = 0$ and $r_i^R(t_l) = r_i(t_l)$ for such stock i. To close the position in stock i, we assume that one needs to pay transaction costs.

Without loss of generality, assume that there are $n \geq 1$ days (including the trading day l) involved since the last trading day, i.e., the last trading day before l is $l - n$. For all $k \in \{l - n + 1, \cdots, l\}$, we compute $r^D(t_k)$ and $r^R(t_k)$ as above. In particular, if some stock i in the portfolio is delisted from the market at time t_u, for some $u \in \{l - n + 1, \cdots, l - 1\}$, then we set $r_i^R(t_v) = r_i^D(t_v) = 0$, for all $v \in \{u + 1, \cdots, l\}$.

Then given $\Xi(t_{l-n})$, we compute

$$\Xi_i(t_l-) = \Xi_i(t_{l-n}) \prod_{k=l-n+1}^{l} \left(1 + r_i^R(t_k)\right), \quad i \in \{1, \cdots, d\}.$$

[1]The dividends computed from the dividend rate r^D contain not only the actual stock dividends, but also other corporate actions. For example, AT&T, which dominated the telephone market for most of the 20^{th} century, was broken up into eight smaller companies in 1984. This lead to a significant drop in the stock price. In our analysis below, we assume that the investor obtained cash in exchange (instead of stocks in the newly established companies).

Since all dividends paid between two consecutive trading days are only reinvested at time t_l, the total dividends available for reinvesting are computed by

$$D(t_l-) = \sum_{i=1}^{d} \Xi_i(t_{l-n}) \sum_{k=l-n+1}^{l} r_i^D(t_k) \prod_{u=l-n+1}^{k-1} \left(1 + r_i^R(t_u)\right).$$

We are now ready to show the empirical results.

3.3 Examples and empirical results

In this section, we analyze the performance of several portfolios empirically. The target weights are expressed in terms of the market weights $\mu(\cdot) = (\mu_1(\cdot), \cdots, \mu_d(\cdot))'$ with components

$$\mu_j(\cdot) = \frac{S_j(\cdot)}{\sum_{i=1}^{d} S_i(\cdot)}, \quad j \in \{1, \cdots, d\},$$

as defined in Definition 1.2.3. In Subsection 3.3.4, we also propose a method to smooth transaction costs.

We shall consider the largest d stocks. We will vary the number d between 100, 300, and 500. The constituent list (the list of the top d stocks) is renewed either weekly, monthly, or quarterly. Whenever we renew the constituent list, we keep the d stocks with the largest total market capitalizations at that time. We trade only these d stocks afterwards until we renew the constituent list again. If any of these stocks stops to exist in the market due to any reason, we simply invest in the remaining stocks without adding a new stock to the list before we renew it next time. Note that renewing the constituent list implies trading to replace the old top d stocks with the new top d stocks. We trade with a specific frequency, which can be either daily, weekly, or monthly. For research on optimal trading frequency, we refer to Ekren, Liu, and Muhle-Karbe [22].

At time t_0, we take the transaction costs due to initializing a portfolio as sunk cost, i.e., we set $\mathrm{TC}(t_0) = 0$. Moreover, we start a portfolio with initial wealth $V(t_0) = 1000$. Note that unless otherwise mentioned, the logarithmic scale is used when plotting $V(\cdot)$ and $\mathrm{TC}(\cdot)$ for the purpose of better interpretability. To simplify the analysis, we impose a uniform transaction cost rate tc on both buying and selling the stocks, i.e., we set $\mathrm{tc}^{\mathrm{b}} = \mathrm{tc}^{\mathrm{s}} = \mathrm{tc}$.

For each example, we provide tables with the yearly returns, the excess returns (relative to the corresponding index tracking portfolio), the standard deviations of the yearly returns, and the Sharpe ratios of the portfolios.[2] These tables also include the wealth and the cumulative transaction costs at the end of the investment period, and the average ratio of the yearly transaction costs to the beginning of year portfolio wealth of the portfolios.

[2]To compute the Sharpe ratios of the portfolios and the indices, the one-year U.S. Treasury yields are used. The data of these yields can be downloaded from `https://www.federalreserve.gov`.

Data source

The data of the total market capitalizations $S(\cdot)$ and the daily returns $r(\cdot)$ is downloaded from the CRSP US Stock Database.[3] This database contains the traded stocks on all major US exchanges. More precisely, we focus on ordinary common stocks.[4] The data starts January 2$^{\mathrm{nd}}$, 1962 and ends December 30$^{\mathrm{th}}$, 2016.

The total market capitalizations are computed by multiplying the numbers of outstanding shares with the share prices, and are essential in determining the target weights. The daily returns include dividends but also delisting returns in case stocks get delisted (for example, the recovery rate in case a traded firm goes bankrupt).

3.3.1 Index tracking portfolio

In this subsection, we introduce the index tracking portfolio. This portfolio is used to benchmark the performance of other portfolios studied in the following subsections. The index tracking portfolio has target weights

$$\pi_j(\cdot) = \mu_j(\cdot), \quad j \in \{1, \cdots, d\}.$$

Note that this portfolio is rebalanced only when the constituent list changes or when dividends are reinvested.

The index tracking portfolio includes the effects of paying transaction costs and reinvesting dividends. In contrast, the capitalization index with wealth process

$$\sum_{i=1}^{d} S_i(\cdot) \times \frac{1000}{\sum_{i=1}^{d} S_i(t_0)}$$

does not take transaction costs and dividends into consideration.

In the following, we examine the performance of the index tracking portfolio under different trading frequencies, renewing frequencies, as well as constituent list sizes d. The portfolio is backtested when there are no transaction costs, i.e., when $\mathrm{tc} = 0$, and when $\mathrm{tc} = 0.5\%$ and $\mathrm{tc} = 1\%$, respectively. These numbers are consistent with the transaction cost estimates in Stoll and Whaley [84], Keim and Madhavan [54], Novy-Marx and Velikov [68], and Fong, Holden, and Trzcinka [37].

Varying the trading frequency

We fix the constituent list size $d = 100$ and use monthly renewing frequency. Table 3.1 shows the performance of the index tracking portfolio and the corresponding capitalization index under daily, weekly, and monthly trading frequencies, respectively.[5] Note that the capitalization index does not depend on the trading frequency. As expected,

[3] See `http://www.crsp.com/products/research-products/crsp-us-stock-databases` for details.
[4] Those stocks in CRSP which have 'Share Code' 10, 11, or 12.
[5] The t-statistics of yearly returns of all portfolios considered in this section range from 3.29 to 4.98. Since they are all significant, we shall omit these numbers in the tables below.

with the same trading frequency, the portfolio performs worse under a larger transaction cost rate tc. In addition, the portfolio outperforms the corresponding index, which implies that the dividends paid exceed the transaction costs imposed even if $\mathrm{tc} = 1\%$. In Figure 3.1, the wealth processes of the daily traded index tracking portfolio and the corresponding capitalization index are plotted.

	YR	Std	SR	Wealth	TC	TR
CI	8.84	16.59	0.22	54.5		
IT_0^d	10.30	16.87	0.30	111.7		
$\mathrm{IT}_{0.5}^d$	10.09	16.84	0.29	100.7	2.7	0.21
IT_1^d	9.89	16.81	0.28	90.7	5.0	0.42
IT_0^w	10.30	16.88	0.30	111.3		
$\mathrm{IT}_{0.5}^w$	10.10	16.85	0.29	100.7	2.5	0.20
IT_1^w	9.90	16.82	0.28	91.0	4.7	0.40
IT_0^m	10.27	16.88	0.30	109.7		
$\mathrm{IT}_{0.5}^m$	10.08	16.86	0.29	99.7	2.3	0.19
IT_1^m	9.89	16.83	0.28	90.6	4.3	0.38

TABLE 3.1: Yearly returns (YR) in percentage, standard deviations of yearly returns (Std), Sharpe ratios (SR), the wealth and the cumulative transaction costs (TC) in thousands at the end of the investment period, and the average ratio of the yearly transaction costs to the beginning of year portfolio wealth (TR) in percentage of the index tracking portfolio (IT) and the corresponding capitalization index (CI) under different trading frequencies and transaction cost rates tc with $d = 100$ and monthly renewing frequency. The subscript x corresponds to $\mathrm{tc} = x\%$ and the superscripts d, w, and m indicate daily, weekly, and monthly trading frequencies, respectively.

Varying the renewing frequency

Still fixing the constituent list size $d = 100$, we now use daily trading frequency and vary the renewing frequency between weekly, monthly, and quarterly frequencies, respectively. As shown in Figure 3.2 and Table 3.2, under the same transaction cost rate tc, the less frequently the constituent list is renewed, the better the portfolio performs. As trades are made when we renew the constituent list, renewing more frequently will impose larger transaction costs, which impacts the performance of the portfolio to a

FIGURE 3.1: The wealth processes of the index tracking portfolio (IT) and the corresponding capitalization index (CI) on logarithmic scale under different transaction cost rates tc with $d = 100$, daily trading frequency, and monthly renewing frequency. The weekly and the monthly traded portfolio performs similarly to the daily traded portfolio under the same transaction cost rate tc.

higher degree. Additionally, the more frequently the constituent list is renewed, the more sensitive the portfolio is to a larger transaction cost rate tc.

Varying the constituent list size d

With daily trading and monthly renewing frequencies, we now backtest the performance of the index tracking portfolio under different constituent list sizes d. As shown in Figure 3.3 and Table 3.3, the portfolio outperforms the corresponding index even with transaction cost rate $\mathrm{tc} = 1\%$. The more stocks the constituent list contains, the better the portfolio performs.

3.3.2 Equally-weighted portfolio

This subsection examines the equally-weighted portfolio. See Benartzi and Thaler [10] and Windcliff and Boyle [91] for a discussion of this portfolio in the context of defined contribution plans, and DeMiguel, Garlappi, and Uppal [19] for a careful study of its properties. Here, the target weights are given by

$$\pi_j(\cdot) = \frac{1}{d}, \quad j \in \{1, \cdots, d\}.$$

For each portfolio with a specific trading frequency, a specific renewing frequency, and a specific constituent list size d, we examine its performance when there are no

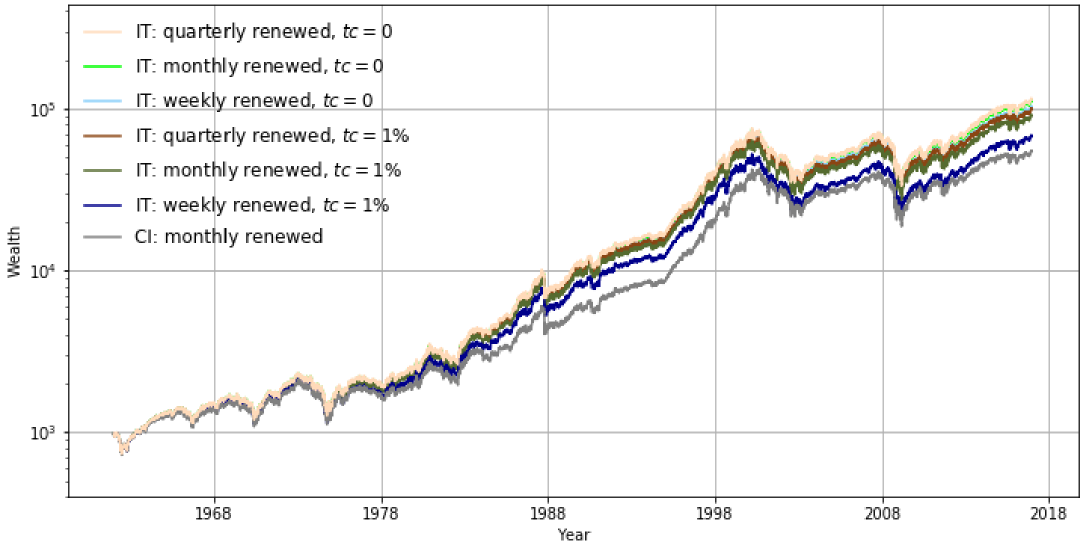

FIGURE 3.2: The wealth processes of the index tracking portfolio (IT) and the corresponding capitalization index (CI) on logarithmic scale under different renewing frequencies and transaction cost rates tc with $d = 100$ and daily trading frequency. The performance of the weekly renewed portfolio when $\mathrm{tc} = 0$ is similar to that of the quarterly renewed portfolio when $\mathrm{tc} = 1\%$. The weekly and the quarterly renewed capitalisation index is not very different from the monthly renewed one.

transaction costs, i.e., when $\mathrm{tc} = 0$, and when $\mathrm{tc} = 0.5\%$ and $\mathrm{tc} = 1\%$, respectively. As shown in the following, the equally-weighted portfolio outperforms the corresponding index tracking portfolio when there are no transaction costs. This well-behaved performance of the equally-weighted portfolio within a frictionless market is popular in the academic literature. However, the equally-weighted portfolio is very sensitive to transaction costs. Its performance is strongly compromised even with a small transaction cost rate $\mathrm{tc} = 0.5\%$.

Varying the trading frequency

Let us fix $d = 100$ and apply monthly renewing frequency. Figure 3.4 plots and Table 3.4 summarises the wealth processes of the equally-weighted and the corresponding index tracking portfolio under different trading frequencies and transaction cost rates tc. When there are no transaction costs, i.e., when $\mathrm{tc} = 0$, the equally-weighted portfolio outperforms the corresponding index tracking portfolio under all three different trading frequencies. A similar observation is also provided in Banner et al. [6]. In addition, the more frequently the portfolio is traded, the better it performs. Trading more frequently also allows to reinvest the dividends faster, which helps to enhance the portfolio performance.

	YR	Std	SR	Wealth	TC	TR
CI^W	8.85	16.66	0.22	54.5		
IT$_0^W$	10.14	16.89	0.29	102.2		
IT$_{0.5}^W$	9.73	16.84	0.27	83.5	4.2	0.40
IT$_1^W$	9.33	16.79	0.25	68.1	7.2	0.79
CI^Q	8.82	16.44	0.22	54.4		
IT$_0^Q$	10.34	16.83	0.31	114.2		
IT$_{0.5}^Q$	10.20	16.81	0.30	106.5	2.0	0.15
IT$_1^Q$	10.06	16.79	0.29	99.2	3.8	0.29

TABLE 3.2: Yearly returns (YR) in percentage, standard deviations of yearly returns (Std), Sharpe ratios (SR), the wealth and the cumulative transaction costs (TC) in thousands at the end of the investment period, and the average ratio of the yearly transaction costs to the beginning of year portfolio wealth (TR) in percentage of the index tracking portfolio (IT) and the corresponding capitalization index (CI) under different renewing frequencies and transaction cost rates tc with $d = 100$ and daily trading frequency. The subscript x corresponds to $\mathrm{tc} = x\%$ and the superscripts W and Q indicate weekly and quarterly renewing frequencies, respectively.

When transaction costs are imposed, Figure 3.4 and Table 3.4 suggest that under the same transaction cost rate tc, the more frequently the portfolio is traded, the larger the decrease in portfolio performance is. The performance of the equally-weighted portfolio is strongly affected by transaction costs. Even with $\mathrm{tc} = 0.5\%$, the corresponding index tracking portfolio outperforms the equally-weighted portfolio. However, slowing down trading helps to reduce the influence of transaction costs. Indeed, the performance of the monthly traded equally-weighted portfolio when $\mathrm{tc} = 1\%$ is similar to that of the daily traded one when $\mathrm{tc} = 0.5\%$. As shown in Figure 3.5, the cumulative transaction costs paid from a monthly traded equally-weighted portfolio when $\mathrm{tc} = 1\%$ are smaller than that from a daily traded one when $\mathrm{tc} = 0.5\%$.

We now study the sensitivity of the Sharpe ratio with respect to the transaction cost rate tc. Specifically, we compute the Sharpe ratios of the monthly traded equally-weighted and index tracking portfolio for $\mathrm{tc} \in \{0, 0.01\%, 0.02\%, \cdots, 0.5\%\}$. As plotted in Figure 3.6, the Sharpe ratios of both the equally-weighted and the index tracking portfolio decrease as tc becomes larger. On the left hand side of the intersection when $\mathrm{tc} < 0.22\%$, the equally-weighted portfolio has a higher Sharpe ratio. On the right hand side of the intersection when $\mathrm{tc} > 0.22\%$, the inverse situation holds. This

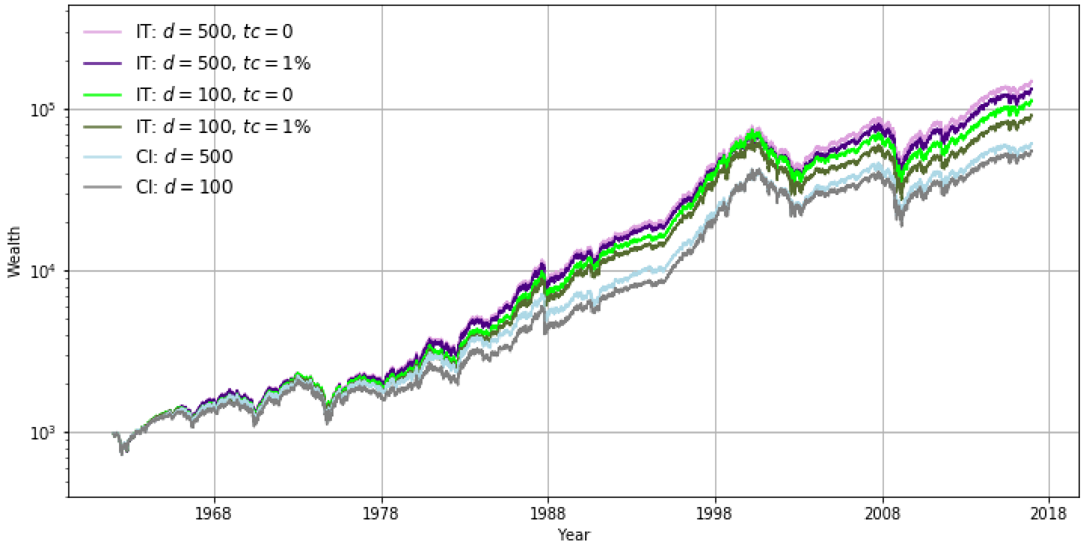

FIGURE 3.3: The wealth processes of the index tracking portfolio (IT) and the corresponding capitalization index (CI) on logarithmic scale under different constituent list sizes d and transaction cost rates tc with daily trading and monthly renewing frequencies. For both the portfolio and the index, the wealth processes with $d = 300$ are omitted. Everything else equal, they would lie between the plotted ones with $d = 100$ and with $d = 500$.

indicates that the equally-weighted portfolio depends more on transaction costs than the index tracking portfolio.

Moreover, as shown in Figure 3.6, the Sharpe ratio is roughly affine in the transaction cost rate. As the standard deviations of yearly returns remain relatively stable for each portfolio, the average yearly return is also roughly affine in transaction cost rate. This observation is consistent with the value of yearly returns reported in all tables, regardless of the portfolio considered. In particular, the slope of the line, when multiplied by the negative of the standard deviation of the portfolio yearly return, is an approximation of the portfolio turnover, as suggested below by Remark 7.

Remark 7. Consider a single period from time 0 to time 1 and let tc_1 and tc_2 be two different transaction cost rates. Then, given the initial wealth $V(0)$ of a portfolio at time 0, we have

$$\begin{aligned} r_1 - r_2 &\approx \frac{V(1) - \mathrm{TC}_1 - V(0)}{V(0)} - \frac{V(1) - \mathrm{TC}_2 - V(0)}{V(0)} \\ &\approx \frac{(\mathrm{tc}_2 - \mathrm{tc}_1)\mathrm{TV}}{V(0)} = (\mathrm{tc}_2 - \mathrm{tc}_1)\mathrm{Turnover}. \end{aligned}$$

Here, r_1 and r_2 are the net returns of the portfolio from time 0 to time 1 with tc_1 and tc_2, respectively, $V(1)$ is the portfolio wealth at time 1 if there are no transaction costs,

	YR	Std	SR	Wealth	TC	TR
CI^{300}	8.94	16.14	0.23	58.7		
IT_0^{300}	10.61	16.57	0.33	132.5		
$\text{IT}_{0.5}^{300}$	10.46	16.55	0.32	123.1	2.4	0.16
IT_1^{300}	10.31	16.53	0.31	114.3	4.5	0.33
CI^{500}	9.01	16.15	0.24	60.6		
IT_0^{500}	10.83	16.61	0.34	147.2		
$\text{IT}_{0.5}^{500}$	10.71	16.60	0.33	139.0	2.3	0.14
IT_1^{500}	10.59	16.58	0.33	131.1	4.3	0.27

TABLE 3.3: Yearly returns (YR) in percentage, standard deviations of yearly returns (Std), Sharpe ratios (SR), the wealth and the cumulative transaction costs (TC) in thousands at the end of the investment period, and the average ratio of the yearly transaction costs to the beginning of year portfolio wealth (TR) in percentage of the index tracking portfolio (IT) and the corresponding capitalization index (CI) under different constituent list sizes d and transaction cost rates tc with daily trading and monthly renewing frequencies. The subscript x corresponds to $\mathrm{tc} = x\%$ and the superscripts 300 and 500 indicate $d = 300$ and $d = 500$, respectively.

and TV is the trading volume of the portfolio. Therefore, we have

$$\frac{\mathrm{SR}_1 - \mathrm{SR}_2}{\mathrm{tc}_1 - \mathrm{tc}_2} \approx \frac{r_1 - r_2}{\sigma(\mathrm{tc}_1 - \mathrm{tc}_2)} \approx -\frac{\mathrm{Turnover}}{\sigma},$$

where SR_1 and SR_2 are the Sharpe ratios of the portfolio with tc_1 and tc_2, respectively, and σ is the standard deviation of the portfolio return.

Varying the renewing frequency

Now we examine the performance of the equally-weighted portfolio with $d = 100$, daily trading frequency, and under weekly, monthly, and quarterly renewing frequencies, respectively. As shown in Figure 3.7 and Table 3.5, under the same transaction cost rate tc, the less frequently the constituent list is renewed, the better the portfolio performs. With $\mathrm{tc} = 0.5\%$, the equally-weighted portfolio already performs worse than the corresponding index tracking portfolio. In particular, the portfolio with a more frequent renewing frequency is more sensitive to transaction costs. As studied in more detail in Subsection 3.3.4, the reason behind these observations is that trading on renewing days incurs extremely large transaction costs compared with trading on other

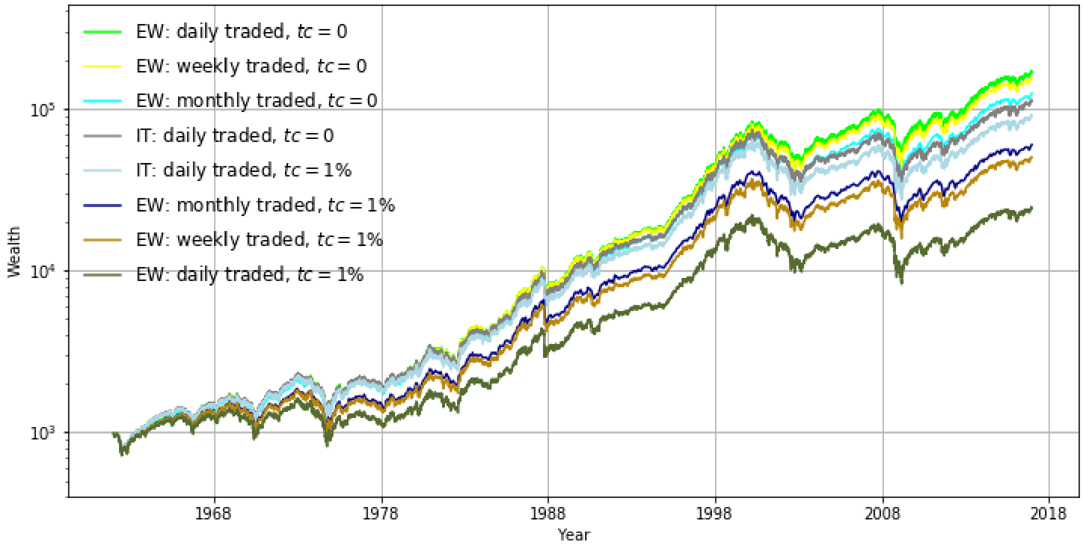

FIGURE 3.4: The wealth processes of the equally-weighted portfolio (EW) and the corresponding index tracking portfolio (IT) on logarithmic scale under different trading frequencies and transaction cost rates tc with $d = 100$ and monthly renewing frequency. Under the same transaction cost rate tc, the weekly and the monthly traded index tracking portfolio performs similarly to the one traded daily.

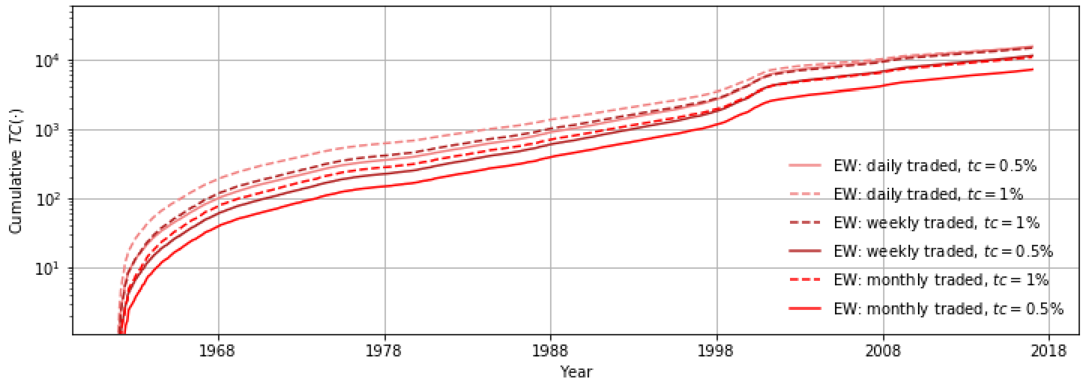

FIGURE 3.5: Cumulative transaction costs on logarithmic scale of the equally-weighted portfolio (EW) under different trading frequencies and transaction cost rates tc with $d = 100$ and monthly renewing frequency.

days when the constituent list is not renewed. These large transaction costs paid on renewing days strongly impact the portfolio performance.

The cumulative transaction costs of the equally-weighted portfolio of Table 3.5 are shown in Figure 3.8. Earlier on, the cumulative transaction costs are higher when weekly renewed than when monthly or quarterly renewed due to the large transaction costs associated with the renewal days. However, later on, the cumulative transaction

	YR	ER	Std	SR	Wealth	TC	TR
IT$_0^d$	10.30		16.87	0.30	111.7		
EW$_0^d$	11.10	0.79 [2.35]	16.83	0.35	168.2		
IT$_{0.5}^d$	10.09		16.84	0.29	100.7	2.7	0.21
EW$_{0.5}^d$	9.19	-0.91 [-2.77]	16.65	0.24	64.0	15.5	1.81
IT$_1^d$	9.89		16.81	0.28	90.7	5.0	0.42
EW$_1^d$	7.31	-2.58 [-7.80]	16.48	0.13	24.3	15.0	3.58
IT$_0^w$	10.30		16.88	0.30	111.3		
EW$_0^w$	10.94	0.64 [1.92]	16.93	0.34	153.8		
IT$_{0.5}^w$	10.10		16.85	0.29	100.7	2.5	0.20
EW$_{0.5}^w$	9.82	-0.27 [-0.83]	16.81	0.28	87.7	11.4	1.06
IT$_1^w$	9.90		16.82	0.28	91.0	4.7	0.40
EW$_1^w$	8.72	-1.18 [-3.60]	16.69	0.21	50.0	14.8	2.10
IT$_0^m$	10.27		16.88	0.30	109.7		
EW$_0^m$	10.53	0.26 [0.84]	17.00	0.31	123.7		
IT$_{0.5}^m$	10.08		16.86	0.29	99.7	2.3	0.19
EW$_{0.5}^m$	9.81	-0.27 [-0.88]	16.91	0.27	86.1	7.2	0.68
IT$_1^m$	9.89		16.83	0.28	90.6	4.3	0.38
EW$_1^m$	9.10	-0.79 [-2.60]	16.83	0.23	59.9	10.9	1.36

TABLE 3.4: Yearly returns (YR) and excess returns (ER) in percentage (t-statistics in brackets), standard deviations of yearly returns (Std), Sharpe ratios, the wealth and the cumulative transaction costs (TC) in thousands at the end of the investment period, and the average ratio of the yearly transaction costs to the beginning of year portfolio wealth (TR) in percentage of the equally-weighted portfolio (EW) and the corresponding index tracking portfolio (IT) under different trading frequencies and transaction cost rates tc with $d = 100$ and monthly renewing frequency. The subscript x corresponds to $\mathrm{tc} = x\%$ and the superscripts d, w, and m indicate daily, weekly, and monthly trading frequencies, respectively.

costs of the weekly renewed portfolio are smaller. The reason is that the weekly renewed portfolio performs worse than the monthly or the quarterly renewed portfolio, hence the transaction costs imposed as a proportion of the portfolio wealth are also

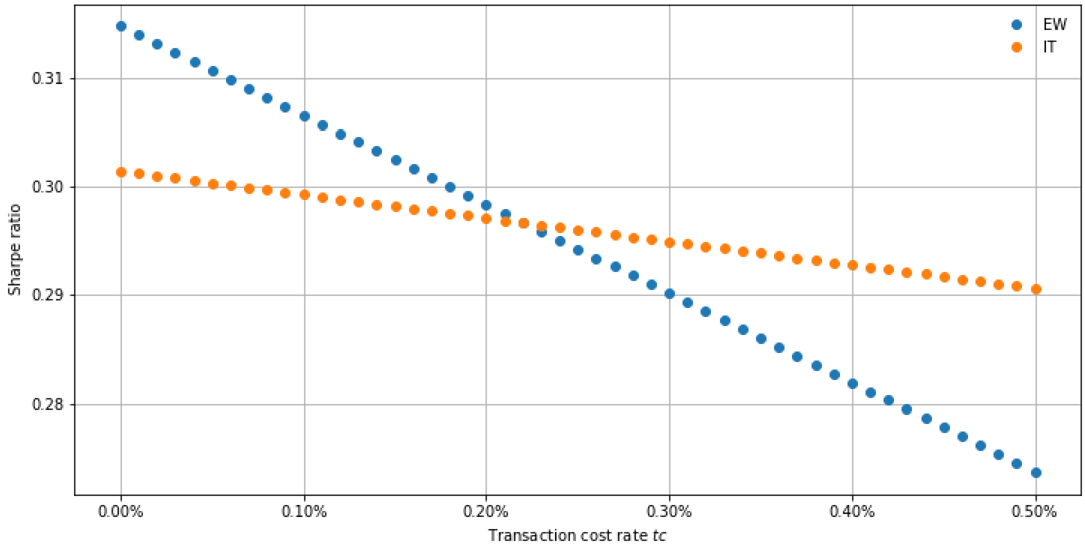

FIGURE 3.6: Sharpe ratios of the equally-weighted portfolio (EW) and the index tracking portfolio (IT) under different transaction cost rates tc with $d = 100$, monthly trading frequency, and monthly renewing frequency.

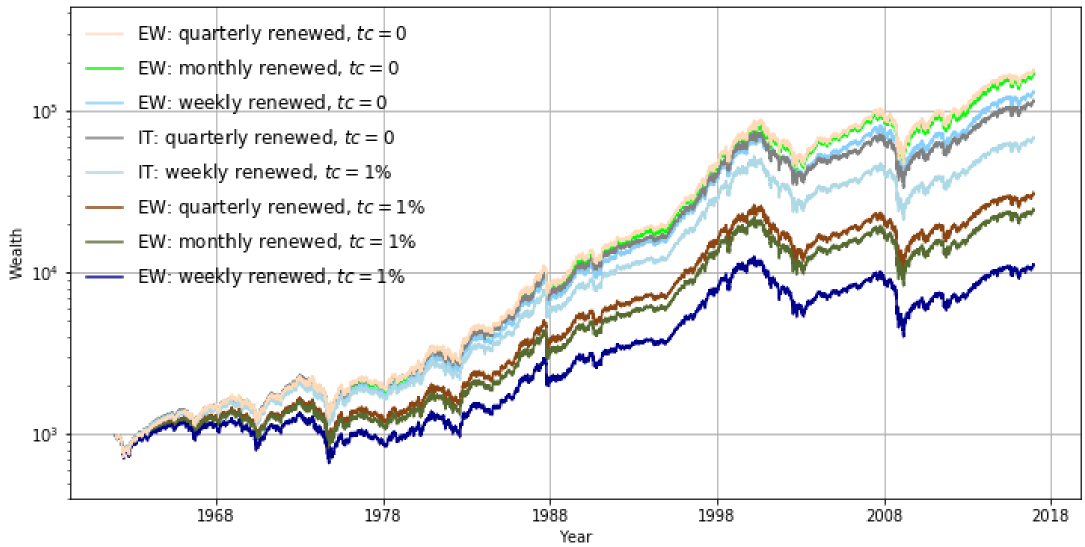

FIGURE 3.7: The wealth processes of the equally-weighted portfolio (EW) and the corresponding index tracking portfolio (IT) on logarithmic scale under different renewing frequencies and transaction cost rates tc with $d = 100$ and daily trading frequency. For the index tracking portfolio, the wealth processes of the quarterly renewed one with $\mathrm{tc} = 0$ and the weekly renewed one with $\mathrm{tc} = 1\%$ are plotted. The omitted wealth processes of the index tracking portfolio lie between the plotted ones.

	YR	ER	Std	SR	Wealth	TC	TR
IT_0^W	10.14		16.89	0.29	102.2		
EW_0^W	10.62	0.48 [1.48]	16.95	0.32	129.8		
$\text{IT}_{0.5}^W$	9.73		16.84	0.27	83.5	4.2	0.40
$\text{EW}_{0.5}^W$	8.20	-1.54 [-4.81]	16.71	0.18	37.5	12.9	2.29
IT_1^W	9.33		16.79	0.25	68.1	7.2	0.79
EW_1^W	5.83	-3.50 [-10.66]	16.50	0.04	11.0	10.5	4.54
IT_0^Q	10.34		16.83	0.31	114.2		
EW_0^Q	11.21	0.86 [2.34]	16.82	0.36	177.6		
$\text{IT}_{0.5}^Q$	10.20		16.81	0.30	106.5	2.0	0.15
$\text{EW}_{0.5}^Q$	9.47	-0.73 [-2.06]	16.65	0.26	73.9	16.0	1.64
IT_1^Q	10.06		16.79	0.29	99.2	3.8	0.29
EW_1^Q	7.76	-2.30 [-6.56]	16.50	0.16	30.8	16.5	3.25

TABLE 3.5: Yearly returns (YR) and excess returns (ER) in percentage (t-statistics in brackets), standard deviations of yearly returns (Std), Sharpe ratios, the wealth and the cumulative transaction costs (TC) in thousands at the end of the investment period, and the average ratio of the yearly transaction costs to the beginning of year portfolio wealth (TR) in percentage of the equally-weighted portfolio (EW) and the corresponding index tracking portfolio (IT) under different renewing frequencies and transaction cost rates tc with $d = 100$ and daily trading frequency. The subscript x corresponds to $\mathrm{tc} = x\%$ and the superscripts W and Q indicate weekly and quarterly renewing frequencies, respectively.

smaller.

Varying the market size d

With daily trading and monthly renewing frequencies, Figure 3.9 plots and Table 3.6 summarises the wealth processes of the equally-weighted and the corresponding index tracking portfolio under different constituent list sizes d. The more stocks the constituent list contains, the better the portfolio performs under the same transaction cost rate tc. Again, its performance is reduced by transaction costs. Even with $d = 500$ and $\mathrm{tc} = 0.5\%$, the equally-weighted portfolio performs worse than the corresponding index tracking portfolio. In addition, the portfolio with a larger constituent list size d is

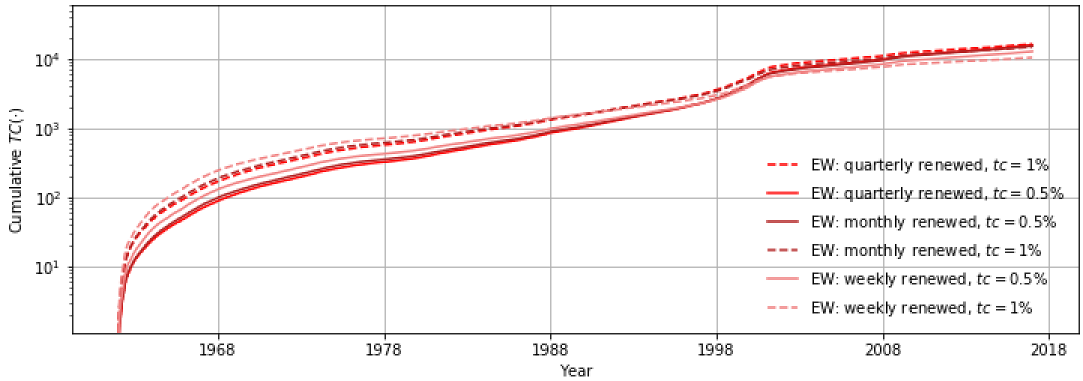

FIGURE 3.8: Cumulative transaction costs on logarithmic scale of the equally-weighted portfolio (EW) under different renewing frequencies and transaction cost rates tc with $d = 100$ and daily trading frequency.

not necessarily more sensitive to transaction costs. Figure 3.10 plots the cumulative transaction costs generated by the portfolio of Table 3.6.

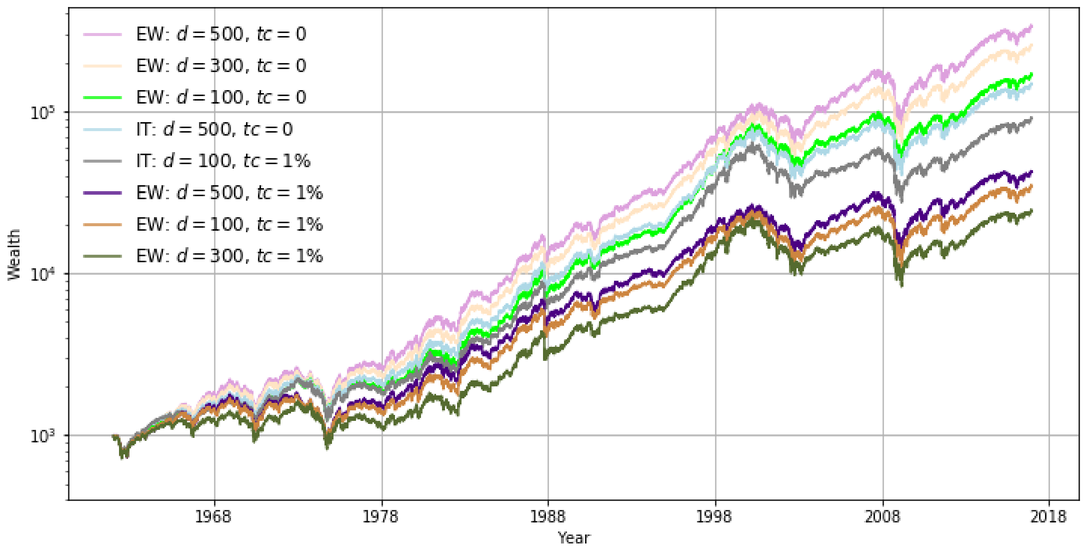

FIGURE 3.9: The wealth processes of the equally-weighted portfolio (EW) and the corresponding index tracking portfolio (IT) on logarithmic scale under different constituent list sizes d and transaction cost rates tc with daily trading and monthly renewing frequencies. For the index tracking portfolio, the wealth processes of the one with $d = 500$ when $\mathrm{tc} = 0$ and the one with $d = 100$ when $\mathrm{tc} = 1\%$ are plotted. The omitted wealth processes of the index tracking portfolio lie between the plotted ones.

	YR	ER	Std	SR	Wealth	TC	TR
IT_0^{300}	10.61		16.57	0.33	132.5		
EW_0^{300}	11.92	1.31 [2.66]	16.59	0.41	255.3		
$\mathrm{IT}_{0.5}^{300}$	10.46		16.55	0.32	123.1	2.4	0.16
$\mathrm{EW}_{0.5}^{300}$	9.92	-0.54 [-1.13]	16.43	0.29	93.6	21.8	1.88
IT_1^{300}	10.31		16.53	0.31	114.3	4.5	0.33
EW_1^{300}	7.96	-2.35 [-4.92]	16.29	0.17	34.3	20.5	3.73
IT_0^{500}	10.83		16.61	0.34	147.2		
EW_0^{500}	12.52	1.70 [3.08]	17.07	0.43	332.8		
$\mathrm{IT}_{0.5}^{500}$	10.71		16.60	0.33	139.0	2.3	0.14
$\mathrm{EW}_{0.5}^{500}$	10.46	-0.25 [-0.47]	16.90	0.31	118.4	27.3	1.94
IT_1^{500}	10.59		16.58	0.33	131.1	4.3	0.27
EW_1^{500}	8.43	-2.16 [-4.07]	16.74	0.19	42.1	24.8	3.85

TABLE 3.6: Yearly returns (YR) and excess returns (ER) in percentage (t-statistics in brackets), standard deviations of yearly returns (Std), Sharpe ratios, the wealth and the cumulative transaction costs (TC) in thousands at the end of the investment period, and the average ratio of the yearly transaction costs to the beginning of year portfolio wealth (TR) in percentage of the equally-weighted portfolio (EW) and the corresponding index tracking portfolio (IT) under different constituent list sizes d and transaction cost rates tc with daily trading and monthly renewing frequencies. The subscript x corresponds to $\mathrm{tc} = x\%$ and the superscripts 300 and 500 indicate $d = 300$ and $d = 500$, respectively.

3.3.3 Entropy-weighted portfolio

In this subsection, we consider the entropy-weighted portfolio (see Section 2.3 in Fernholz [26], Example 5.3 in Karatzas and Ruf [51], and Example 2.4.1), which relies on target weights

$$\pi_j(\cdot) = \frac{\mu_j(\cdot)\log\mu_j(\cdot)}{\sum_{i=1}^d \mu_i(\cdot)\log\mu_i(\cdot)}, \quad j \in \{1, \cdots, d\}.$$

In the following, we examine the performance of the entropy-weighted portfolio under specific configurations when there are no transaction costs, i.e., when $\mathrm{tc} = 0$, and when $\mathrm{tc} = 0.5\%$. The performance of the entropy-weighted portfolio is less sensitive to transaction costs and is better when $\mathrm{tc} = 0.5\%$, compared with that of the equally-weighted portfolio.

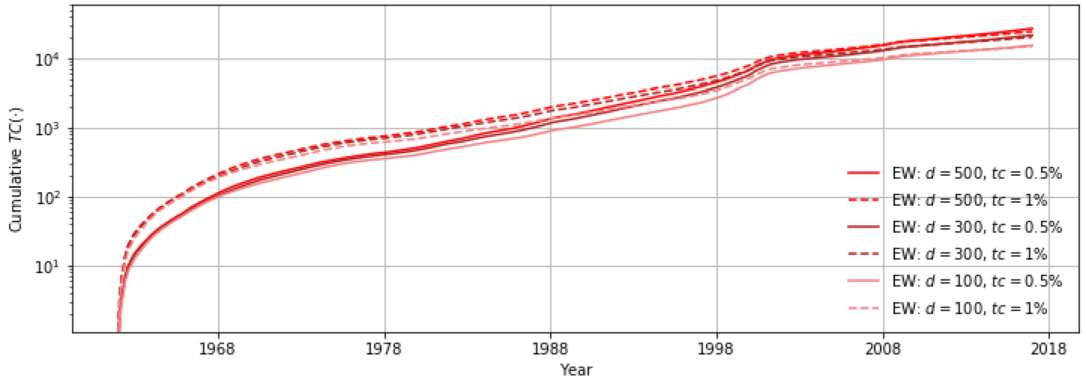

FIGURE 3.10: Cumulative transaction costs on logarithmic scale of the equally-weighted portfolio (EW) under different constituent list sizes d and transaction cost rates tc with daily trading frequency and monthly renewing frequency.

Varying the trading frequency

As before, when backtesting the portfolio under different trading frequencies, we set the constituent list size $d = 100$ and apply monthly renewing frequency. Figure 3.11 displays and Table 3.7 summarises the wealth processes of the entropy-weighted and the corresponding index tracking portfolio under different trading frequencies. Compared with the equally-weighted portfolio summarised in Table 3.4, the entropy-weighted portfolio performs worse (but still outperforms the corresponding index tracking portfolio) when there are no transaction costs, i.e., when $\mathrm{tc} = 0$. However, opposite to the equally-weighted portfolio, the weekly and the monthly traded entropy-weighted portfolio still outperforms the corresponding index tracking portfolio when $\mathrm{tc} = 0.5\%$.

Over a large time horizon, the loss in the portfolio wealth resulting from paying transaction costs is usually higher than the cumulative transaction costs imposed. This is exhibited in Figure 3.11, which also plots the sum of the wealth process and of the cumulative transaction costs of the entropy-weighted portfolio when $\mathrm{tc} = 0.5\%$. Notice that the wealth process when $\mathrm{tc} = 0$ is above this sum. Indeed, paying transaction costs not only takes money out of the portfolio, but also deprives the opportunity for making potential gains.

Varying the renewing frequency

With $d = 100$ and daily trading frequency, we now examine the performance of the entropy-weighted portfolio applying different renewing frequencies (renewed weekly, monthly, and quarterly, respectively). Figure 3.12 displays and Table 3.8 summarises the wealth processes of the entropy-weighted and the corresponding index tracking portfolio under different renewing frequencies. Similar to the equally-weighted portfolio, the less frequently the constituent list is renewed, the better the entropy-weighted portfolio performs. When transaction costs are imposed, its performance depends

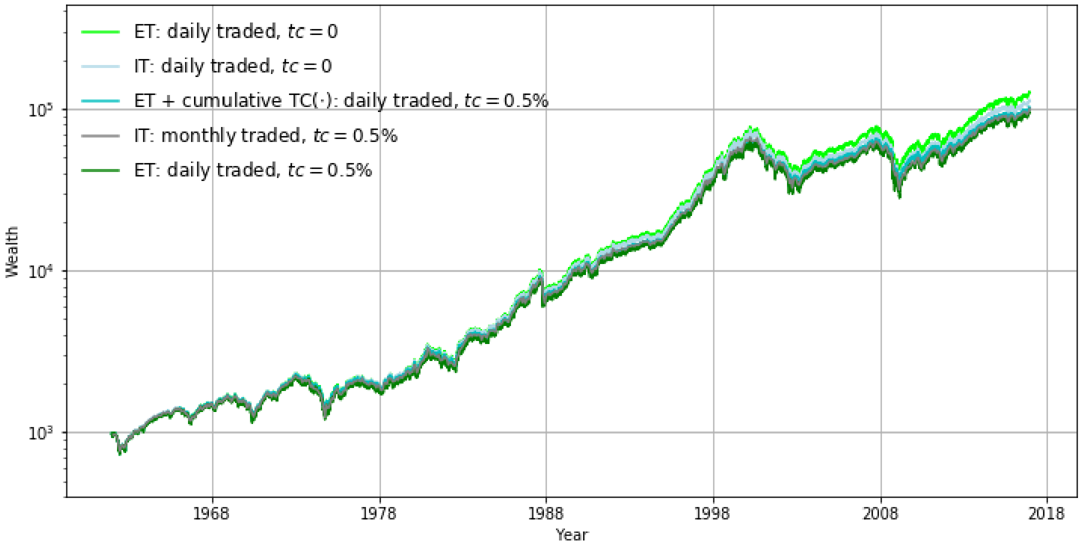

FIGURE 3.11: The wealth processes of the entropy-weighted portfolio (ET) and the corresponding index tracking portfolio (IT) on logarithmic scale under different trading frequencies and transaction cost rates tc with $d = 100$ and monthly renewing frequency. For both the entropy-weighted and the index tracking portfolio, the omitted wealth processes of Table 3.7 lie between the plotted ones. The sum of the wealth process and of the cumulative transaction costs of the daily traded entropy-weighted portfolio when $\mathrm{tc} = 0.5\%$ is also plotted. Note that the sum is below the wealth process of the daily traded entropy-weighted portfolio when $\mathrm{tc} = 0$.

more on the renewing frequency. However, compared with the equally-weighted portfolio summarised in Table 3.5, the performance of the entropy-weighted portfolio is less sensitive to transaction costs under the same renewing frequency.

Varying the market size d

Applying daily trading and monthly renewing frequencies, we backtest the entropy-weighted portfolio under different constituent list sizes d ($= 100$, 300, and 500, respectively), as shown in Figure 3.13 and Table 3.9. Similar to the equally-weighted and the index tracking portfolio, the more stocks the constituent list contains, the better the entropy-weighted portfolio performs. Compared with the equally-weighted portfolio, the entropy-weighted portfolio with the same d depends less on transaction costs. In particular, with $d = 500$ and $\mathrm{tc} = 0.5\%$, the entropy-weighted portfolio still outperforms the corresponding index tracking portfolio.

	YR	ER	Std	SR	Wealth	TC	TR
IT$_0^d$	10.30		16.87	0.30	111.7		
ET$_0^d$	10.53	0.23 [2.03]	16.90	0.32	125.1		
IT$_{0.5}^d$	10.09		16.84	0.29	100.7	2.7	0.21
ET$_{0.5}^d$	9.97	-0.12 [-1.08]	16.83	0.28	94.6	6.4	0.53
IT$_0^w$	10.30		16.88	0.30	111.3		
ET$_0^w$	10.50	0.21 [1.82]	16.92	0.31	123.1		
IT$_{0.5}^w$	10.10		16.85	0.29	100.7	2.5	0.20
ET$_{0.5}^w$	10.12	0.02 [0.21]	16.88	0.29	101.7	4.5	0.36
IT$_0^m$	10.27		16.88	0.30	109.7		
ET$_0^m$	10.40	0.14 [1.24]	16.94	0.31	116.9		
IT$_{0.5}^m$	10.08		16.86	0.29	99.7	2.3	0.19
ET$_{0.5}^m$	10.11	0.03 [0.29]	16.90	0.29	100.8	3.4	0.28

TABLE 3.7: Yearly returns (YR) and excess returns (ER) in percentage (t-statistics in brackets), standard deviations of yearly returns (Std), Sharpe ratios (SR), the wealth and the cumulative transaction costs (TC) in thousands at the end of the investment period, and the average ratio of the yearly transaction costs to the beginning of year portfolio wealth (TR) in percentage of the entropy-weighted portfolio (ET) and the corresponding index tracking portfolio (IT) under different trading frequencies and transaction cost rates tc with $d = 100$ and monthly renewing frequency. The subscript x corresponds to $\mathrm{tc} = x\%$ and the superscripts d, w, and m indicate daily, weekly, and monthly trading frequencies, respectively.

3.3.4 Diversity-weighted portfolio and smoothing transaction costs

One portfolio that draws much attention in Stochastic Portfolio Theory is the so-called diversity-weighted portfolio generated from the "measure of diversity"

$$G_p(x) = \left(\sum_{i=1}^{d} x_i^p \right)^{1/p}, \quad x \in \Delta^d,$$

for some fixed $p \in (0,1)$, where Δ^d is given by (1.3). Without changing the relative ranking of the stocks, the function $G_p(\cdot)$ generates portfolio weights smaller (larger) than the corresponding market weights for stocks with large (small) market weights.

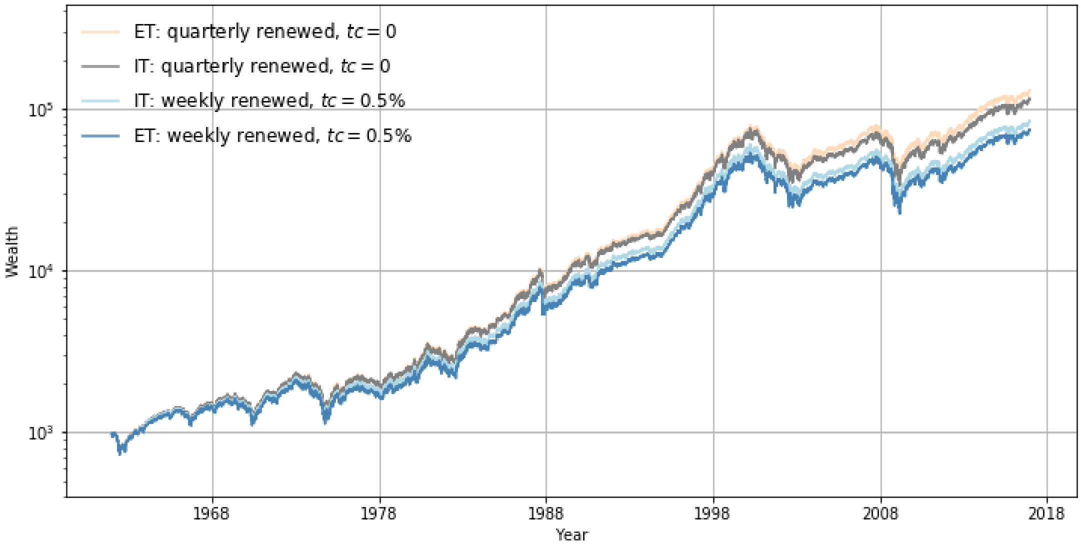

FIGURE 3.12: The wealth processes of the entropy-weighted portfolio (ET) and the corresponding index tracking portfolio (IT) on logarithmic scale under different renewing frequencies and transaction cost rates tc with $d = 100$ and daily trading frequency. For both the entropy-weighted and the index tracking portfolio, the omitted wealth processes of Table 3.8 lie between the plotted ones.

This diversification property of G_p is closely related to the implementation of relative arbitrage portfolios; see Section 7 in Fernholz and Karatzas [35] for details. Section 6.3 in Fernholz [26] provides a theoretical approximation of the diversity-weighted portfolio turnover. Vervuurt and Karatzas [89] study the portfolio generated by G_p with a negative p. An empirical study of this portfolio using S&P 500 market data can be found in Fernholz, Garvy, and Hannon [32] and Chapter 7 of Fernholz [26], as well as in Example 2.4.2. Here, the target portfolio weights are consistent with the trading strategy generated multiplicatively in Example 2.4.2.

In the following, we examine the performance of this portfolio and illustrate the tradeoff between trading with a higher frequency and paying transaction costs. To achieve this, we shall replace the market weights by a smoothed version, given by

$$\overline{\mu}(\cdot) = \alpha\mu(\cdot) + (1-\alpha)\Lambda(\cdot)$$

with $\alpha \in (0,1)$. Here, the moving average process $\Lambda(\cdot) = (\Lambda_1(\cdot), \cdots, \Lambda_d(\cdot))'$ is given by

$$\Lambda_j(\cdot) = \begin{cases} \frac{1}{\delta}\int_0^{\cdot} \mu_j(t)\mathrm{d}t + \frac{1}{\delta}\int_{\cdot-\delta}^{0} \mu_j(0)\mathrm{d}t & \text{on } [0,\delta) \\ \frac{1}{\delta}\int_{\cdot-\delta}^{\cdot} \mu_j(t)\mathrm{d}t & \text{on } [\delta,\infty) \end{cases}, \quad j \in \{1, \cdots, d\},$$

for a fixed constant $\delta > 0$. This moving average process $\Lambda(\cdot)$ is also included in the portfolio generating function studied in Schied, Speiser, and Voloshchenko [82]. Then

	YR	ER	Std	SR	Wealth	TC	TR
IT_0^W	10.14		16.89	0.29	102.2		
ET_0^W	10.31	0.17 [1.57]	16.93	0.30	111.3		
$\text{IT}_{0.5}^W$	9.73		16.84	0.27	83.5	4.2	0.40
$\text{ET}_{0.5}^W$	9.50	-0.23 [-2.13]	16.84	0.26	73.9	7.4	0.77
IT_0^Q	10.34		16.83	0.31	114.2		
ET_0^Q	10.58	0.24 [2.08]	16.86	0.32	129.0		
$\text{IT}_{0.5}^Q$	10.20		16.81	0.30	106.5	2.0	0.15
$\text{ET}_{0.5}^Q$	10.11	-0.09 [-0.78]	16.81	0.29	101.6	5.8	0.45

TABLE 3.8: Yearly returns (YR) and excess returns (ER) in percentage (t-statistics in brackets), standard deviations of yearly returns (Std), Sharpe ratios, the wealth and the cumulative transaction costs (TC) in thousands at the end of the investment period, and the average ratio of the yearly transaction costs to the beginning of year portfolio wealth (TR) in percentage of the entropy-weighted portfolio (ET) and the corresponding index tracking portfolio (IT) under different renewing frequencies and transaction cost rates tc with $d = 100$ and daily trading frequency. The subscript x corresponds to $\mathrm{tc} = x\%$ and the superscripts W and Q indicate weekly and quarterly renewing frequencies, respectively.

the target weights are given by

$$\pi_j(\cdot) = \mu_j(\cdot)\left(\varpi_j(\cdot) - \sum_{i=1}^{d} \mu_i(\cdot)\varpi_i(\cdot) + 1\right), \quad j \in \{1, \cdots, d\},$$

where

$$\varpi_j(\cdot) = \frac{\alpha\left(\overline{\mu}_j(\cdot)\right)^{p-1}}{\sum_{i=1}^{d}\left(\overline{\mu}_i(\cdot)\right)^{p}}, \quad j \in \{1, \cdots, d\}.$$

To backtest the portfolio, we fix $d = 100$, the renewing frequency to be quarterly, and the "diversity degree" $p = 0.8$. Moreover, we compute the moving average process $\Lambda(\cdot)$ using a one-year window. To be more specific, with daily trading frequency, we set $\delta = 250$; with weekly trading frequency, we set $\delta = 52$. To compute $\Lambda(\cdot)$ under weekly trading frequency, we only use market weights μ's on the days when transactions are made.

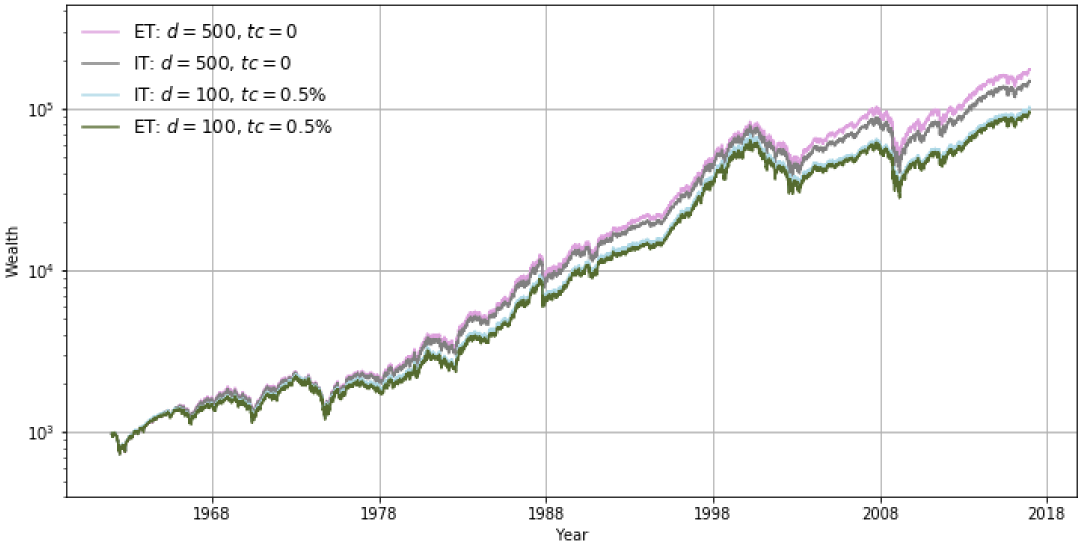

FIGURE 3.13: The wealth processes of the entropy-weighted portfolio (ET) and the corresponding index tracking portfolio (IT) on logarithmic scale under different constituent list sizes d and transaction cost rates tc with daily trading and monthly renewing frequencies. For both the entropy-weighted and the index tracking portfolio, the omitted wealth processes of Table 3.9 lie between the plotted ones.

Varying the convexity weight α and the trading frequency

In Table 3.10, we summarise the wealth processes of the diversity-weighted and the corresponding index tracking portfolio. These processes are under both daily and weekly trading frequencies and with three different choices for the convexity weight α, when there are no transaction costs, i.e., when $\mathrm{tc} = 0$, and when $\mathrm{tc} = 0.5\%$ and $\mathrm{tc} = 1\%$, respectively.

We first consider the case when there are no transaction costs. Everything else equal, the daily traded diversity-weighted portfolio performs similarly to the weekly traded portfolio. Under either trading frequency, the smaller the convexity weight α is, the worse the portfolio performs. Generating the portfolio with a smaller α is somewhat alike to trading less frequently, as it assigns less weights on the volatile term $\mu(\cdot)$ and more weights on the stable term $\Lambda(\cdot)$ when constructing $\overline{\mu}(\cdot)$, and thus makes $\overline{\mu}(\cdot)$ less volatile.

Next, we consider the case with transaction costs. Under either daily or weekly trading frequency, a smaller convexity weight α tends to improve the portfolio performance when the transaction cost rate tc becomes larger. This can be useful, since decreasing α partially cancels out the effect of transaction costs. Moreover, when $\mathrm{tc} = 1\%$, the daily traded portfolio with $\alpha = 0.2$ performs similarly as the weekly traded portfolio with $\alpha = 0.6$. This indicates that, instead of trading less frequently in order to avoid paying transaction costs, one can adjust the convexity weight α to reach a more

	YR	ER	Std	SR	Wealth	TC	TR
IT_0^{300}	10.61		16.57	0.33	132.5		
ET_0^{300}	10.89	0.28 [2.24]	16.57	0.34	152.5		
$\text{IT}_{0.5}^{300}$	10.46		16.55	0.32	123.1	2.4	0.16
$\text{ET}_{0.5}^{300}$	10.43	-0.03 [-0.25]	16.52	0.31	121.2	6.2	0.43
IT_0^{500}	10.83		16.61	0.34	147.2		
ET_0^{500}	11.16	0.33 [2.51]	16.66	0.36	173.2		
$\text{IT}_{0.5}^{500}$	10.71		16.60	0.33	139.0	2.3	0.14
$\text{ET}_{0.5}^{500}$	10.75	0.04 [0.27]	16.62	0.34	141.0	6.4	0.39

TABLE 3.9: Yearly returns (YR) and excess returns (ER) in percentage (t-statistics in brackets), standard deviations of yearly returns (Std), Sharpe ratios, the wealth and the cumulative transaction costs (TC) in thousands at the end of the investment period, and the average ratio of the yearly transaction costs to the beginning of year portfolio wealth (TR) in percentage of the entropy-weighted portfolio (ET) and the corresponding index tracking portfolio (IT) under different constituent list sizes d and transaction cost rates tc with daily trading and monthly renewing frequencies. The subscript x corresponds to $\mathrm{tc} = x\%$ and the superscripts 300 and 500 indicate $d = 300$ and $d = 500$, respectively.

favourable balance between trading frequently and paying transaction costs.

Dynamic convexity weight α to smooth transaction costs

Instead of fixing α throughout the investment period, we could adjust α dynamically to speed up or slow down trading. For example, given a baseline portfolio with constant convexity weight α_0, we would choose $\alpha < \alpha_0$ ($\alpha > \alpha_0$) to trade less (more) in the next period if transaction costs paid in the last period are more (less) than a certain level. In the remaining part of this example, we fix daily trading frequency and dynamically adjust $\alpha(\cdot)$.

Let $M \geq 4$ denote the total number of quarters in the investment period and let t_u^{r}, for $u \in \{1, \cdots, M\}$, denote the trading days on which the constituent list is renewed. Moreover, set $t_0^{\mathrm{r}} = t_0$. On a specific renewing day t_u^{r}, for $u \in \{1, \cdots, M\}$, let $\widetilde{\mathrm{TC}}(t_u^{\mathrm{r}})$ denote the averaged fictitious transaction costs relative to the wealth $V_{\alpha_0}(\cdot-)$ of the baseline portfolio paid in the previous period. More precisely, $\widetilde{\mathrm{TC}}(t_u^{\mathrm{r}})$ is computed as

$$\widetilde{\mathrm{TC}}(t_u^{\mathrm{r}}) = \frac{1}{\kappa_u} \sum_{t \in [t_{u-1}^{\mathrm{r}}, t_u^{\mathrm{r}})} \min \left\{ \frac{\mathrm{TC}_{\alpha_0}(t)}{V_{\alpha_0}(t-)}, \xi \right\}.$$

	IT_0^w	$\mathrm{IT}_{0.5}^w$	IT_1^w	α	DW_0^d	$\mathrm{DW}_{0.5}^d$	DW_1^d	DW_0^w	$\mathrm{DW}_{0.5}^w$	DW_1^w
				0.2	10.36	10.20	10.03	10.36	10.20	10.05
YR	10.34	10.20	10.06	0.6	10.43	10.18	9.93	10.42	10.23	10.03
				1	10.54	10.11	9.68	10.51	10.24	9.96
				0.2	0.03	0	-0.03	0.02	0.01	-0.01
ER				0.6	0.1	-0.01	-0.13	0.09	0.03	-0.03
				1	0.2	-0.09	-0.38	0.18	0.04	-0.09
				0.2	16.84	16.81	16.79	16.85	16.83	16.80
Std	16.85	16.83	16.81	0.6	16.84	16.81	16.77	16.86	16.83	16.80
				1	16.84	16.79	16.74	16.87	16.83	16.79
				0.2	0.31	0.30	0.29	0.31	0.30	0.29
SR	0.31	0.30	0.29	0.6	0.31	0.30	0.28	0.31	0.30	0.29
				1	0.32	0.29	0.27	0.32	0.30	0.28
				0.2	115.6	106.2	97.6	115.1	106.5	98.5
W	123.4	108.3	94.2	0.6	119.8	105.4	92.8	118.8	107.8	97.8
				1	126.2	101.6	81.7	124.3	108.3	94.2
				0.2		2.1	4.0		2.0	3.7
TC		3.5	6.3	0.6		3.2	5.9		2.5	4.6
				1		5.3	9.0		3.5	6.3

TABLE 3.10: Yearly returns (YR) and excess returns (ER) (with respect to the index tracking portfolio (IT) summarised here and in Table 3.1) in percentage (t-statistics in brackets), standard deviations of yearly returns (Std), Sharpe ratios (SR), the wealth (W) and the cumulative transaction costs (TC) in thousands at the end of the investment period, and the average ratio of the yearly transaction costs to the beginning of year portfolio wealth (TR) in percentage of the diversity-weighted portfolio (DW) under different trading frequencies, convexity weights α, and transaction cost rates tc with $d = 100$ and quarterly renewing frequency. The subscript x corresponds to $\mathrm{tc} = x\%$ and the superscripts d and w indicate daily and weekly trading frequencies, respectively.

Here, κ_u is the number of trading days within the period $[t^{\mathrm{r}}_{u-1}, t^{\mathrm{r}}_u)$, $\mathrm{TC}_{\alpha_0}(\cdot)$ is computed by (3.3) from the baseline portfolio, and ξ is a predetermined level used to make the estimate more robust. On a trading day t, We regard $\mathrm{TC}_{\alpha_0}(t)/V_{\alpha_0}(t-) > \xi$ as "abnormal" transaction costs relative to $V_{\alpha_0}(t-)$. Such large costs appear, for example, when the constituent list is changing. The level ξ is determined such that the days, on which "abnormal" transaction costs occur, only count for a small proportion of all trading days.

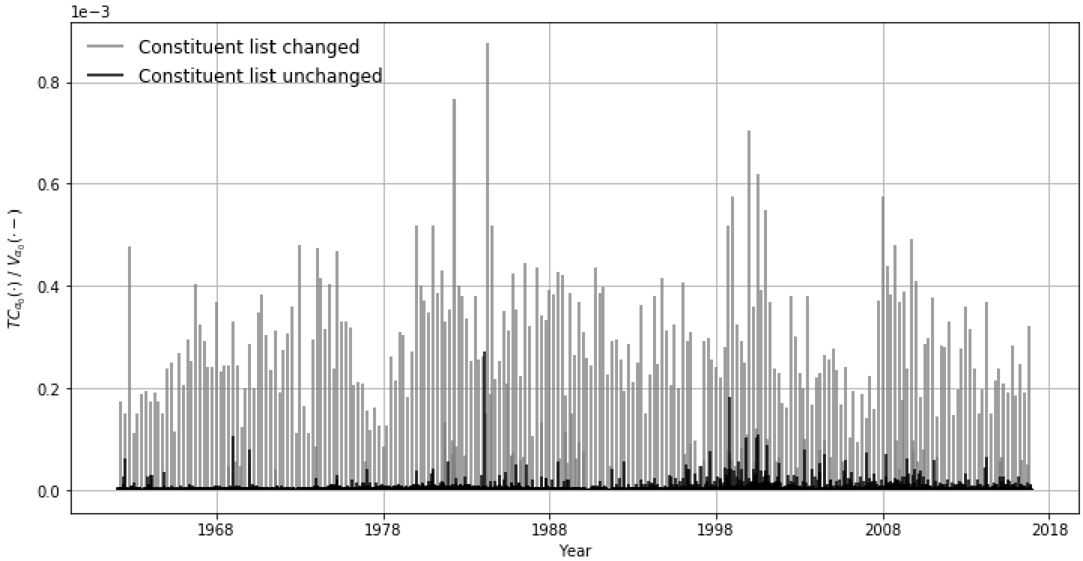

FIGURE 3.14: Transaction costs $\mathrm{TC}_{\alpha_0}(\cdot)$ of the baseline portfolio relative to its wealth $V_{\alpha_0}(\cdot-)$, i.e., $\mathrm{TC}_{\alpha_0}(\cdot)/V_{\alpha_0}(\cdot-)$, paid when the constituent list is changed and unchanged, respectively, with $\alpha_0 = 0.6$, when $\mathrm{tc} = 0.5\%$.

Figure 3.14 shows the relative transaction costs $\mathrm{TC}_{\alpha_0}(\cdot)/V_{\alpha_0}(\cdot-)$ when the constituent list is changed and unchanged, respectively, with $\alpha_0 = 0.6$ and $\mathrm{tc} = 0.5\%$. Transaction costs paid when the constituent list is changed are significantly larger than when the constituent list remains the same. The days when the constituent list is changed only account for less than 5% of all trading days, i.e., $M/N < 0.05$, where N is again the total number of trading days.

We shall smooth the relative transaction costs $\mathrm{TC}(\cdot)/V_{\alpha_0}(\cdot-)$ by dynamically adjusting $\alpha(\cdot)$. Starting with $\alpha(t_0) = \alpha_0$, the convexity weight $\alpha(\cdot)$ is piecewise constant and only updated on the renewal dates t^{r}_u, for $u \in \{4, \cdots, M\}$. This reduces additional transaction costs incurred from updating $\alpha(\cdot)$. In particular, for all $u \in \{4, \cdots, M\}$, we set

$$\alpha(t^{\mathrm{r}}_u) = \max\left\{\min\left\{\alpha_0\left(1 - \beta \times \overline{\mathrm{TC}}(t^{\mathrm{r}}_u)\right), 1\right\}, 0\right\}$$

with

$$\overline{\mathrm{TC}}(t^{\mathrm{r}}_u) = \frac{\widetilde{\mathrm{TC}}(t^{\mathrm{r}}_u)}{\frac{1}{4}\sum_{\nu=u-3}^{u} \widetilde{\mathrm{TC}}(t^{\mathrm{r}}_\nu)} - 1.$$

Here, $\beta \geq 0$ is a fixed non-negative constant that controls the sensitivity of $\alpha(\cdot)$. Hence, we compare the fictitious averaged transaction costs relative to $V_{\alpha_0}(\cdot-)$ within the most recent quarter to that of the past one year. The value $\overline{\mathrm{TC}}(\cdot)$ is positive (negative) if the baseline portfolio requires more (less) transaction costs in the most recent quarter than the last year. This will yield $\alpha(\cdot) < \alpha_0$ ($\alpha(\cdot) > \alpha_0$) and slow down (speed up) the trading within the next quarter.

Using a baseline portfolio with constant convexity weight $\alpha_0 = 0.6$ and assuming $\mathrm{tc} = 0.5\%$, we now estimate the effects of a dynamic convexity weight $\alpha(\cdot)$ empirically. Moreover, we set the relative transaction cost level $\xi = 10^{-5}$, as the fictitious relative transaction costs $\mathrm{TC}_{\alpha_0}(\cdot)/V_{\alpha_0}(\cdot-)$ are less than this level on more than 95% of all trading days. We examine the three cases $\beta \in \{0, 0.05, 0.1\}$. Note that $\beta = 0$ yields $\alpha(\cdot) = \alpha_0$. With these choices of β, the portfolio with dynamic $\alpha(\cdot)$ performs similarly to the baseline portfolio; see column $V^d_{\mathrm{tc\_0.5}}$ in Table 3.10 with $\alpha = 0.6$.

The convexity weight process $\alpha(\cdot)$ corresponding to the sensitivity parameter β is shown in Figure 3.15. As expected, $\alpha(\cdot)$ fluctuates more rapidly with a larger β. As mentioned before, increasing α speeds up trading and leads to more transaction costs, while decreasing α has the opposite effect. Choosing β very large results in a portfolio far away from the baseline portfolio. This dependence on β is illustrated in Figure 3.16, which plots the square root of the total quadratic variation of relative transaction costs $\mathrm{TC}(\cdot)/V_{\alpha_0}(\cdot-)$, computed as

$$\sqrt{\sum_{l=1}^{N} \left(\frac{\mathrm{TC}(t_l)}{V_{\alpha_0}(t_l-)} - \frac{\mathrm{TC}(t_{l-1})}{V_{\alpha_0}(t_{l-1}-)} \right)^2}, \tag{3.11}$$

for different sensitivity parameters β. The square root of the total quadratic variation is a measure of volatility with percentage as unit. Figure 3.16 suggests that choosing $\beta \approx 0.05$ minimises (3.11).

3.4 Conclusion

In this chapter, we empirically study the impact of proportional transaction costs on systemically generated portfolios. Given a target portfolio, we provide a scheme to backtest the portfolio using total market capitalization and daily stock return time series. Implementing this scheme, we examine the performance of the index tracking portfolio, the equally-weighted portfolio, the entropy-weighted portfolio, and the diversity-weighted portfolio. When backtesting, we assume various transaction cost rates, trading frequencies, portfolio constituent list sizes, and renewing frequencies.

As expected, everything else equal, a portfolio performs worse as transaction costs are higher and the portfolio renewing frequency of the underlying constituent list is higher. In the absence of transaction costs, trading under a higher frequency leads to better portfolio performance. However, in the presence of transaction costs, implementing a higher trading frequency can also result in larger transaction costs and

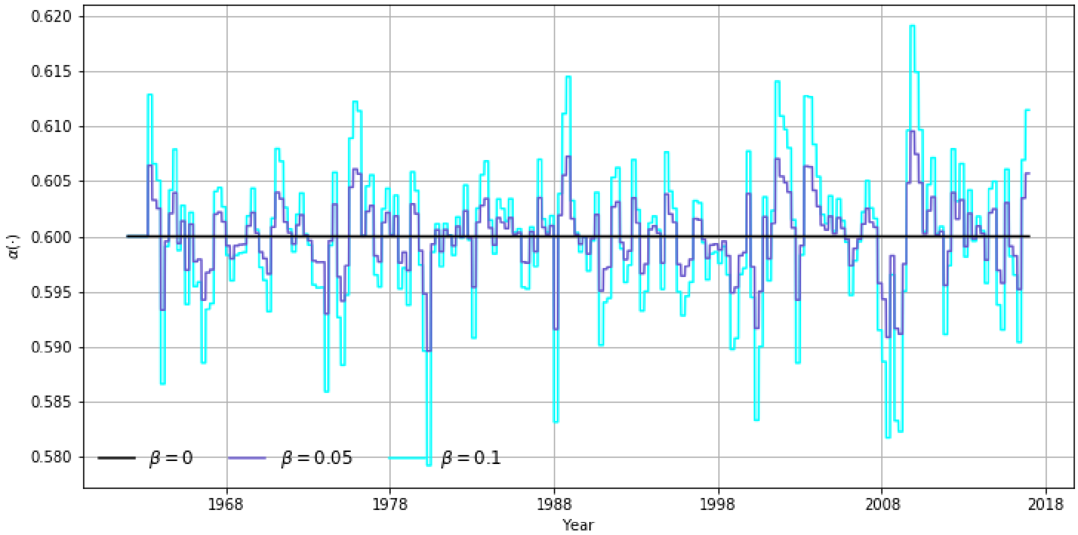

FIGURE 3.15: Convexity weight process $\alpha(\cdot)$ for different sensitivity parameters β with $\alpha_0 = 0.6$, when $\mathrm{tc} = 0.5\%$.

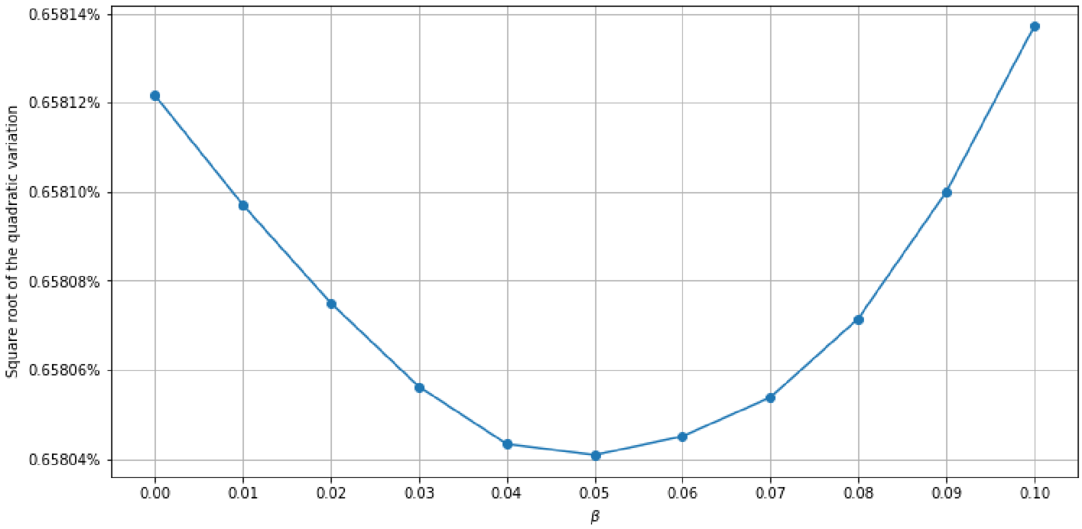

FIGURE 3.16: Square root of the quadratic variation of relative transaction costs $\mathrm{TC}(\cdot)/V_{\alpha_0}(\cdot-)$ for different sensitivity parameters β with $\alpha_0 = 0.6$, when $\mathrm{tc} = 0.5\%$.

reduce the portfolio performance significantly. Hence, trading under an appropriate frequency is necessary in practice. In addition, with or without transaction costs, a more diversified portfolio containing more stocks usually performs better.

The empirical results indicate that the equally-weighted portfolio performs well relative to the index tracking portfolio when there are no transaction costs. However, the performance of the equally-weighted portfolio is very sensitive to transaction costs. The entropy-weighted portfolio performs a bit worse than the equally-weighted portfolio (but still outperforms the index tracking portfolio) when there are no transaction costs. But the performance of the entropy-weighted portfolio depends much less on transaction costs, compared to the equally-weighted portfolio.

Last but not the least, we propose a method to smooth transaction costs. Without changing the trading frequency, this method is similar to altering the trading speed dynamically.

Chapter 4

Leakage of Generalised Rank-Dependent Trading Strategies

As mentioned in Section 1.5, Fernholz [29] is the first to analyse portfolios generated by functions that depend on the ranked market weights. The ranked market weights have advantages in illustrating the distribution of capital in the market, which is of great significance in SPT. As in Chapter 3, we analyse portfolios with their constituent lists containing the top d stocks in the market. Every time we renew the portfolio constituent list, new stocks (indexed by their names) are introduced into the portfolio to replace some old stocks. In this sense, these portfolios are not portfolios that invest in fixed companies, but are actually more close to rank-dependent portfolios. In this chapter, which is based on Xie [95] , we study the so-called *leakage* effect resulting from the change in the constituent list of rank-dependent portfolios.

The following is an outline of this chapter. Section 4.1 introduces the master formula of the wealth of a trading strategy generated either multiplicatively or additively by a generalised portfolio generation function of a specific group of ranked market weights. The definition of the leakage comes naturally from the master formula and is computed theoretically. Section 4.2 provides the method to estimate the leakage in discreet time. Section 4.3 discusses the procedure of using historical data to back-test the portfolio performance and estimating the leakage. Section 4.4 studies several trading strategies empirically.

4.1 Leakage of functionally generated trading strategies

Leakage was first introduced by Fernholz [29] to measure the effect on portfolio return of stocks dropped from ("leaked" out of) the constituent list when rebalancing. We apply some of the techniques used in Fernholz [29] and reanalyse this effect with necessary adjustments. To be more specific, we put ourselves in a frictionless market $\mathcal{M}$ with $d \geq 2$ stocks as in Chapter 2. However, this time, we only invest in the top $k < d$ stocks in terms of the market capitalisation among the d stocks every time when rebalancing the portfolio. We denote the market that contains these top k stocks by $\mathcal{M}^k$.

As before, we still let $\Lambda(\cdot)$ denote an $\mathbb{R}^m$-valued continuous process of finite variation on $[0,T]$, for $T \geq 0$ and some $m \in \mathbb{N}$. Moreover, we use $\mu(\cdot)$ and $\boldsymbol{\mu}(\cdot)$ to denote the Δ^d-valued market weight process and the $\mathbb{W}^d$-valued ranked market weight process for the d stocks, respectively. Recall the relevant notations from Chapter 1. To proceed, we define

$$\widetilde{\mu}_i(t) = \mathfrak{M}(t)\mu_{(i)}(t), \quad i \in \{1, \cdots, d\},\ t \geq 0, \tag{4.1}$$

where

$$\mathfrak{M}(\cdot) = \frac{1}{\sum_{j=1}^{k} \mu_{(j)}(\cdot)}$$

represents the process of the multiplier of the market weights from the market $\mathcal{M}$ to the market $\mathcal{M}^k$. Then, the Δ^k-valued process $\widetilde{\mu}(\cdot) = (\widetilde{\mu}_1(\cdot), \cdots, \widetilde{\mu}_k(\cdot))'$ is the market weights process with respect to the market $\mathcal{M}^k$. Note that

$$\sum_{j=1}^{k} \widetilde{\mu}_j(t) = 1, \quad t \geq 0,$$

by (4.1). In particular, since $\boldsymbol{\mu}(\cdot)$ is a d-dimensional continuous semimartingale, $\widetilde{\mu}(\cdot)$ is a k-dimensional continuous semimartingale by (4.1). Moreover, let $\widetilde{\mathcal{W}}$ be some open subset of $\mathbb{R}^m \times \mathbb{R}^k$ such that

$$\mathsf{P}\left[(\Lambda(t), \widetilde{\mu}(t)) \in \widetilde{\mathcal{W}},\ \forall\, t \geq 0\right] = 1.$$

Recall the definition of a generalised regular function from Definition 2.1.1. In this chapter, we let $\widetilde{G} : \widetilde{W} \to \mathbb{R}$ be a generalised regular function for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$. To wit, there exists a measurable function $\widetilde{G}^{\mathrm{D}} = (\widetilde{G}_1^{\mathrm{D}}, \cdots, \widetilde{G}_k^{\mathrm{D}})' : \widetilde{W} \to \mathbb{R}^k$ such that the process $\widetilde{\vartheta}(\cdot) = (\widetilde{\vartheta}_1(\cdot), \cdots, \widetilde{\vartheta}_k(\cdot))'$ with components

$$\widetilde{\vartheta}_i(\cdot) = \widetilde{G}_i^{\mathrm{D}}(\Lambda(\cdot), \widetilde{\mu}(\cdot)), \quad i \in \{1, \cdots, k\},$$

is in $\mathcal{L}(\widetilde{\mu})$ (i.e., predictable and integrable with respect to $\widetilde{\mu}(\cdot)$). Moreover, the continuous, adapted process

$$\Gamma(\cdot) = \widetilde{G}(\Lambda(0), \widetilde{\mu}(0)) - \widetilde{G}(\Lambda(\cdot), \widetilde{\mu}(\cdot)) + \int_0^{\cdot} \sum_{j=1}^{k} \widetilde{\vartheta}_j(t) \mathrm{d}\widetilde{\mu}_j(t) \tag{4.2}$$

is of finite variation on the interval $[0,T]$, for all $T \geq 0$. In particular, for the sake of a better interpretability, we normalise

$$\widetilde{G}(\Lambda(0), \widetilde{\mu}(0)) = 1 \tag{4.3}$$

in the same manner as in Subsection 2.2.3.

Similar to (1.7), for a trading strategy $\varphi(\cdot)$ generated either multiplicatively or additively by a generalised regular function $\widetilde{G}$ for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$, the wealth process $\widetilde{V}^{\varphi}(\cdot)$ of $\varphi(\cdot)$ relative to the market $\mathcal{M}^k$ is given by

$$\widetilde{V}^{\varphi}(\cdot) = \sum_{j=1}^{k} \varphi_j(\cdot)\widetilde{\mu}_j(\cdot). \tag{4.4}$$

Recall from (2.9) that $N_i(x)$ represents the number of components of $x \in \Delta^d$ that coalesce at a given rank $i \in \{1, \cdots, d\}$. Moreover, recall the local time $\mathbf{\Lambda}^{(i,j)}(\cdot)$ from Definition 1.5.1 and (2.11), for all $i, j \in \{1, \cdots, d\}$ with $i < j$.

Lemma 4.1.1. *For a given generalised regular function $\widetilde{G}$ for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$, the corresponding finite-variation process $\Gamma(\cdot)$ given by* (4.2) *satisfies*

$$\Gamma(\cdot) = \widetilde{\Gamma}(\cdot) + L(\cdot),$$

where

$$\begin{aligned}
\widetilde{\Gamma}(\cdot) = \widetilde{G}(\Lambda(0), \widetilde{\mu}(0)) &+ \int_0^{\cdot} \sum_{i=1}^{k}\sum_{j=1}^{d} \frac{\widetilde{\vartheta}_i(t)\mathfrak{M}(t)}{N_i(\mu(t))} \mathbf{1}_{\{\mu_j(t)=\mu_{(i)}(t)\}} \mathrm{d}\mu_j(t) \\
&- \widetilde{G}(\Lambda(\cdot), \widetilde{\mu}(\cdot)) - \int_0^{\cdot} \sum_{i,j=1}^{k}\sum_{\nu=1}^{d} \widetilde{\vartheta}_i(t)\widetilde{\mu}_i(t)\mathfrak{M}(t)\frac{\mathbf{1}_{\{\mu_\nu(t)=\mu_{(j)}(t)\}}}{N_j(\mu(t))}\mathrm{d}\mu_\nu(t) \\
&+ \int_0^{\cdot} \sum_{i,j,\nu=1}^{k} \mathfrak{M}^2(t)\widetilde{\vartheta}_i(t)\widetilde{\mu}_i(t)\mathrm{d}\left[\mu_{(j)}, \mu_{(\nu)}\right](t) \\
&- \int_0^{\cdot} \sum_{i,j=1}^{k} \mathfrak{M}^2(t)\widetilde{\vartheta}_i(t)\mathrm{d}\left[\mu_{(i)}, \mu_{(j)}\right](t) \\
&+ \int_0^{\cdot} \sum_{i=1}^{k}\sum_{j=i+1}^{k} \frac{\widetilde{\vartheta}_i(t)\mathfrak{M}(t)}{N_i(\mu(t))}\mathrm{d}\mathbf{\Lambda}^{(i,j)}(t) - \int_0^{\cdot} \sum_{i=1}^{k}\sum_{j=1}^{i-1} \frac{\widetilde{\vartheta}_i(t)\mathfrak{M}(t)}{N_i(\mu(t))}\mathrm{d}\mathbf{\Lambda}^{(j,i)}(t) \\
&- \int_0^{\cdot} \sum_{i,j=1}^{k}\sum_{\nu=j+1}^{k} \frac{\widetilde{\vartheta}_i(t)\widetilde{\mu}_i(t)\mathfrak{M}(t)}{N_j(\mu(t))}\mathrm{d}\mathbf{\Lambda}^{(j,\nu)}(t) \\
&+ \int_0^{\cdot} \sum_{i,j=1}^{k}\sum_{\nu=1}^{j-1} \frac{\widetilde{\vartheta}_i(t)\widetilde{\mu}_i(t)\mathfrak{M}(t)}{N_j(\mu(t))}\mathrm{d}\mathbf{\Lambda}^{(\nu,j)}(t)
\end{aligned} \tag{4.5}$$

is a process of finite variation on $[0, T]$, for all $T \geq 0$, and

$$\begin{aligned}
L(\cdot) = &\int_0^{\cdot} \sum_{i=1}^{k}\sum_{j=k+1}^{d} \frac{\widetilde{\vartheta}_i(t)\mathfrak{M}(t)}{N_i(\mu(t))}\mathrm{d}\mathbf{\Lambda}^{(i,j)}(t) \\
&- \int_0^{\cdot} \sum_{i,j=1}^{k}\sum_{\nu=k+1}^{d} \frac{\widetilde{\vartheta}_i(t)\widetilde{\mu}_i(t)\mathfrak{M}(t)}{N_j(\mu(t))}\mathrm{d}\mathbf{\Lambda}^{(j,\nu)}(t).
\end{aligned} \tag{4.6}$$

Proof. By Itô's lemma and (4.1), we have

$$\mathrm{d}\widetilde{\mu}_i(t) = \mathrm{d}\left(\mathfrak{M}(t)\mu_{(i)}(t)\right) = \mathfrak{M}(t)\mathrm{d}\mu_{(i)}(t) + \mu_{(i)}(t)\mathrm{d}\mathfrak{M}(t) + \mathrm{d}\left[\mu_{(i)}, \mathfrak{M}\right](t),$$

for all $i \in \{1, \cdots, k\}$, and

$$\mathrm{d}\mathfrak{M}(t) = -\mathfrak{M}^2(t)\sum_{j=1}^{k}\mathrm{d}\mu_{(j)}(t) + \mathfrak{M}^3(t)\sum_{i,j=1}^{k}\mathrm{d}\left[\mu_{(i)}, \mu_{(j)}\right](t).$$

The above two equations imply

$$\begin{aligned}\mathrm{d}\widetilde{\mu}_i(t) &= \mathfrak{M}(t)\mathrm{d}\mu_{(i)}(t) - \mathfrak{M}(t)\widetilde{\mu}_i(t)\sum_{j=1}^{k}\mathrm{d}\mu_{(j)}(t) \\ &\quad + \mathfrak{M}^2(t)\widetilde{\mu}_i(t)\sum_{j,\nu=1}^{k}\mathrm{d}\left[\mu_{(j)}, \mu_{(\nu)}\right](t) - \mathfrak{M}^2(t)\sum_{j=1}^{k}\mathrm{d}\left[\mu_{(i)}, \mu_{(j)}\right](t),\end{aligned} \tag{4.7}$$

for all $i \in \{1, \cdots, k\}$. Then, combining (2.10) and (4.7) yields

$$\begin{aligned}\mathrm{d}\widetilde{\mu}_i(t) = {}& \frac{\mathfrak{M}(t)}{N_i(\mu(t))}\sum_{j=1}^{d}\mathbf{1}_{\{\mu_j(t)=\mu_{(i)}(t)\}}\mathrm{d}\mu_j(t) - \mathfrak{M}(t)\widetilde{\mu}_i(t)\sum_{j=1}^{k}\sum_{\nu=1}^{d}\frac{\mathbf{1}_{\{\mu_\nu(t)=\mu_{(j)}(t)\}}}{N_j(\mu(t))}\mathrm{d}\mu_\nu(t) \\ & + \mathfrak{M}^2(t)\widetilde{\mu}_i(t)\sum_{j,\nu=1}^{k}\mathrm{d}\left[\mu_{(j)}, \mu_{(\nu)}\right](t) - \mathfrak{M}^2(t)\sum_{j=1}^{k}\mathrm{d}\left[\mu_{(i)}, \mu_{(j)}\right](t) \\ & + \frac{\mathfrak{M}(t)}{N_i(\mu(t))}\sum_{j=i+1}^{d}\mathrm{d}\mathbf{\Lambda}^{(i,j)}(t) - \frac{\mathfrak{M}(t)}{N_i(\mu(t))}\sum_{j=1}^{i-1}\mathrm{d}\mathbf{\Lambda}^{(j,i)}(t) \\ & - \widetilde{\mu}_i(t)\sum_{j=1}^{k}\frac{\mathfrak{M}(t)}{N_j(\mu(t))}\sum_{\nu=j+1}^{d}\mathrm{d}\mathbf{\Lambda}^{(j,\nu)}(t) - \widetilde{\mu}_i(t)\sum_{j=1}^{k}\frac{\mathfrak{M}(t)}{N_j(\mu(t))}\sum_{\nu=1}^{j-1}\mathrm{d}\mathbf{\Lambda}^{(\nu,j)}(t),\end{aligned}$$

for all $i \in \{1, \cdots, k\}$, which, together with (4.2) and some computation, imply (4.5) and (4.6). Moreover, since both $\Gamma(\cdot)$ and $L(\cdot)$ are of finite variation on $[0, T]$, for all $T \geq 0$, so is $\widetilde{\Gamma}(\cdot)$. $\square$

Compared with (1.20), our computation of the finite-variation process $\Gamma(\cdot)$ is different from that given by, e.g., Fernholz [29].

Remark 8. The process $L(\cdot)$ given by (4.6) consists of all local time components between stocks that may leak out of the constituent list and stocks that may be ranked smaller than or equal to k after rebalancing. Note that, if $\widetilde{G}$ is a generalised Lyapunov function for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$ by Definition 2.1.2 (e.g., when $\widetilde{G}$ satisfies conditions (aii) and (bii) in Theorem 2.1.1), $L(\cdot)$ is positive and increasing from 0. In this case, $L(\cdot)$ measures the contribution to $\Gamma(\cdot)$ from renewing the portfolio timely by dropping off the stocks at the same prices as those to be included in the constituent list of the top k stocks after rebalancing. The stocks dropped become too small and no longer belong to the top k stocks subsequently.

However, as one is not able to predict the prices and hence the ranks of stocks, this timely renewing of the portfolio constituent list is unrealistic. Therefore, $L(\cdot)$ should be subtracted from $\Gamma(\cdot)$ and hence the portfolio wealth, as $\Gamma(\cdot)$ contributes to the portfolio wealth through the master formulas (2.15) or (2.22). This observation also indicates a method to estimate the leakage, which is closely linked to $L(\cdot)$, as we will see in the following. □

The financial meaning of $L(\cdot)$ suggested in Remark 8 becomes more clear under some further assumptions on the regular function $\widetilde{G}$ and the market $\mathcal{M}$, as shown in the following propositions.

Proposition 4.1.2. *Give a generalised regular function $\widetilde{G}$ for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$. If the corresponding measurable function $\widetilde{G}^{\mathrm{D}}$ is symmetric in the second argument, i.e., if*

$$\widetilde{G}_i^{\mathrm{D}}(\lambda, x) = \widetilde{G}_j^{\mathrm{D}}(\lambda, x), \quad \lambda \in \mathbb{R}^m,\ x \in \Delta^k, \tag{4.8}$$

for all $i, j \in \{1, \cdots, k\}$ with $x_i = x_j$, then the finite-variation process $\widetilde{\Gamma}(\cdot)$ given by (4.5) *simplifies to*

$$\begin{aligned}
\widetilde{\Gamma}(\cdot) = {} & \widetilde{G}(\Lambda(0), \widetilde{\mu}(0)) + \int_0^{\cdot} \sum_{i=1}^{k} \sum_{j=1}^{d} \frac{\widetilde{\vartheta}_i(t)\mathfrak{M}(t)}{N_i(\mu(t))} \mathbf{1}_{\{\mu_j(t) = \mu_{(i)}(t)\}} \mathrm{d}\mu_j(t) \\
& - \widetilde{G}(\Lambda(\cdot), \widetilde{\mu}(\cdot)) - \int_0^{\cdot} \sum_{i,j=1}^{k} \sum_{\nu=1}^{d} \widetilde{\vartheta}_i(t) \widetilde{\mu}_i(t) \mathfrak{M}(t) \frac{\mathbf{1}_{\{\mu_\nu(t) = \mu_{(j)}(t)\}}}{N_j(\mu(t))} \mathrm{d}\mu_\nu(t) \\
& + \int_0^{\cdot} \sum_{i,j,\nu=1}^{k} \mathfrak{M}^2(t) \widetilde{\vartheta}_i(t) \widetilde{\mu}_i(t) \mathrm{d}\left[\mu_{(j)}, \mu_{(\nu)}\right](t) \\
& - \int_0^{\cdot} \sum_{i,j=1}^{k} \mathfrak{M}^2(t) \widetilde{\vartheta}_i(t) \mathrm{d}\left[\mu_{(i)}, \mu_{(j)}\right](t).
\end{aligned}$$

Proof. Since the measurable function $\widetilde{G}^{\mathrm{D}}$ is symmetric in the second argument, by (4.8) we have

$$\frac{\widetilde{\vartheta}_i(t)}{N_i(\mu(t))} \mathrm{d}\mathbf{\Lambda}^{(i,j)}(t) = \frac{\widetilde{\vartheta}_j(t)}{N_j(\mu(t))} \mathrm{d}\mathbf{\Lambda}^{(j,i)}(t), \quad i, j \in \{1, \cdots, k\},\ i \neq j,$$

which implies

$$\int_0^{\cdot} \sum_{i=1}^{k} \sum_{j=i+1}^{k} \frac{\widetilde{\vartheta}_i(t)\mathfrak{M}(t)}{N_i(\mu(t))} \mathrm{d}\mathbf{\Lambda}^{(i,j)}(t) = \int_0^{\cdot} \sum_{i=1}^{k} \sum_{j=1}^{i-1} \frac{\widetilde{\vartheta}_i(t)\mathfrak{M}(t)}{N_i(\mu(t))} \mathrm{d}\mathbf{\Lambda}^{(j,i)}(t) \tag{4.9}$$

and

$$\int_0^{\cdot} \sum_{i,j=1}^{k} \sum_{\nu=1}^{j-1} \frac{\widetilde{\vartheta}_i(t)\widetilde{\mu}_i(t)\mathfrak{M}(t)}{N_j(\mu(t))} \mathrm{d}\mathbf{\Lambda}^{(\nu,j)}(t) = \int_0^{\cdot} \sum_{i,j=1}^{k} \sum_{\nu=j+1}^{k} \frac{\widetilde{\vartheta}_i(t)\widetilde{\mu}_i(t)\mathfrak{M}(t)}{N_j(\mu(t))} \mathrm{d}\mathbf{\Lambda}^{(j,\nu)}(t). \tag{4.10}$$

Then, combining (4.5), (4.9), and (4.10) yields the desired result. □

Recall the random permutation p_t from (1.18).

Proposition 4.1.3. *Let $\widetilde{G}$ be a generalised regular function for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$ with the corresponding measurable function $\widetilde{G}^{\mathrm{D}}$ symmetric in the second argument as in* (4.8)*. Assume that the market weight processes $\mu_1(\cdot), \cdots, \mu_d(\cdot)$ are pathwise mutually non-degenerate as defined in Definition 1.5.2. Then the finite variation process $\Gamma(\cdot)$ given by* (4.2) *now has the decomposition*

$$\Gamma(\cdot) = \widetilde{\Gamma}(\cdot) + L(\cdot),$$

where

$$\begin{aligned}\widetilde{\Gamma}(\cdot) = \widetilde{G}(\Lambda(0), \widetilde{\mu}(0)) - \widetilde{G}(\Lambda(\cdot), \widetilde{\mu}(\cdot)) + \int_0^{\cdot} \sum_{i=1}^{k} \sum_{j=1}^{d} \widetilde{\vartheta}_i(t) \mathfrak{M}(t) \mathbf{1}_{\{j = p_t(i)\}} \mathrm{d}\mu_j(t) \\ - \int_0^{\cdot} \sum_{i,j=1}^{k} \sum_{\nu=1}^{d} \widetilde{\vartheta}_i(t) \widetilde{\mu}_i(t) \mathfrak{M}(t) \mathbf{1}_{\{\nu = p_t(j)\}} \mathrm{d}\mu_\nu(t) - \int_0^{\cdot} \sum_{i,j=1}^{k} \mathfrak{M}^2(t) \widetilde{\vartheta}_i(t) \mathrm{d}\left[\mu_{(i)}, \mu_{(j)}\right](t) \\ + \int_0^{\cdot} \sum_{i,j,\nu=1}^{k} \mathfrak{M}^2(t) \widetilde{\vartheta}_i(t) \widetilde{\mu}_i(t) \mathrm{d}\left[\mu_{(j)}, \mu_{(\nu)}\right](t)\end{aligned}$$

and

$$L(\cdot) = \frac{1}{2} \int_0^{\cdot} \left(\widetilde{\vartheta}_k(t) - \sum_{j=1}^{k} \widetilde{\vartheta}_j(t) \widetilde{\mu}_j(t) \right) \mathfrak{M}(t) \mathrm{d}\mathbf{\Lambda}^{(k,k+1)}(t)$$

are both of finite variation on $[0,T]$, for all $T \geq 0$.

Proof. By Proposition 4.1.11 in Fernholz [26], when $\mu_1(\cdot), \cdots, \mu_d(\cdot)$ are pathwise mutually non-degenerate, (2.10) becomes

$$\mu_{(i)}(\cdot) = \mu_{(i)}(0) + \int_0^{\cdot} \sum_{j=1}^{d} \mathbf{1}_{\{j = p_t(i)\}} \mathrm{d}\mu_j(t) + \frac{1}{2} \int_0^{\cdot} \mathrm{d}\mathbf{\Lambda}^{(i,i+1)}(t) - \frac{1}{2} \int_0^{\cdot} \mathrm{d}\mathbf{\Lambda}^{(i-1,i)}(t), \quad (4.11)$$

for all $i \in \{1, \cdots, d\}$. Then thanks to (4.11), a similar reasoning as in the proof of Lemma 4.1.1 and Proposition 4.1.2 yields the desired result. □

4.1.1 Leakage of multiplicatively generated trading strategies

For a given generalised regular function $\widetilde{G}$ for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$, let $\psi(\cdot)$ denote the trading strategy generated multiplicatively by $\widetilde{G}$. Then, the wealth process $\widetilde{V}^{\psi}(\cdot)$ can now be expressed through the master formula introduced in the following theorem.

Theorem 4.1.4. *Let $\psi(\cdot)$ be the trading strategy generated multiplicatively by a generalised regular function $\widetilde{G}$ for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$ in the same manner of* (2.19) *and* (2.20)*. Then the wealth process $\widetilde{V}^{\psi}(\cdot)$ of $\psi(\cdot)$ relative to the market $\mathcal{M}^k$, with initial wealth*

$\widetilde{V}^{\psi}(0) = 1$*, is given by the master formula*

$$\log \widetilde{V}^{\psi}(\cdot) = \log \widetilde{G}(\Lambda(\cdot), \widetilde{\mu}(\cdot)) + \int_0^{\cdot} \frac{\mathrm{d}\widetilde{\Gamma}(t)}{\widetilde{G}(\Lambda(t), \widetilde{\mu}(t))} + \int_0^{\cdot} \frac{\mathrm{d}L(t)}{\widetilde{G}(\Lambda(t), \widetilde{\mu}(t))} \tag{4.12}$$

with $\widetilde{\Gamma}(\cdot)$ *and* $L(\cdot)$ *given by* (4.5) *and* (4.6)*, respectively.*

Proof. Since $\psi(\cdot)$ is generated multiplicatively by $\widetilde{G}$, the master formula (2.22) implies

$$\log \widetilde{V}^{\psi}(\cdot) = \log \widetilde{G}(\Lambda(\cdot), \widetilde{\mu}(\cdot)) + \int_0^{\cdot} \frac{\mathrm{d}\Gamma(t)}{\widetilde{G}(\Lambda(t), \widetilde{\mu}(t))},$$

which, together with Lemma 4.1.1, yield the desired result. □

The leakage $L^{\psi}(\cdot)$ of the trading strategy $\psi(\cdot)$ is then defined as the negative of the last term of (4.12), i.e.,

$$L^{\psi}(\cdot) = -\int_0^{\cdot} \frac{\mathrm{d}L(t)}{\widetilde{G}(\Lambda(t), \widetilde{\mu}(t))} \tag{4.13}$$

with $L(\cdot)$ given by (4.6). It measures the cumulative lost in the (logarithmic) relative wealth $\widetilde{V}^{\psi}(\cdot)$ due to untimely renewing the portfolio constituent to stop investing in the smallest stocks, which are delisted from ("leaks" out of) the portfolio subsequently. This explanation indicates the method to estimate the leakage $L^{\psi}(\cdot)$, as shown in the next section.

Remark 9. Our computation for the leakage here is different from, for example, Example 4.2 in Fernholz [29]. The method introduced in Example 4.2 in Fernholz [29] may lead to trading strategies which have positive portfolio weights for stocks of ranks larger than k for some ranked portfolio generating functions $\boldsymbol{G}$ of $\boldsymbol{\mu}(\cdot)$. To see this, consider a ranked portfolio generating function

$$\boldsymbol{G}(\boldsymbol{x}) = 1 - \frac{1}{2}\sum_{j=1}^{k} x_{(j)}^2, \quad \boldsymbol{x} \in \mathbb{W}^d.$$

Let the trading strategy $\psi(\cdot)$ be generated multiplicatively in the same manner as in Example 4.2 in Fernholz [29] by a portfolio generating function G of $\mu(\cdot)$ with $G(x) = \boldsymbol{G}(\mathfrak{R}(x))$, for all $x \in \Delta^d$. Recall the random permutation p_t from (1.18). Then, $\psi(\cdot)$ has portfolio weights

$$\pi_{p_t(i)}(t) = \frac{1 + \frac{1}{2}\sum_{j=1}^{k} \mu_{(j)}^2(t)}{1 - \frac{1}{2}\sum_{j=1}^{k} \mu_{(j)}^2(t)} \mu_{(i)}(t) \geq 0, \quad i \in \{k+1, \cdots, d\},\ t \geq 0,$$

where the equality holds if and only if $\mu_{(i)}(t) = 0$, which is in general not the case. To avoid this problem, instead of using $\boldsymbol{G}$ of $\boldsymbol{\mu}(\cdot)$ as the portfolio generating function, we use $\widetilde{G}$ of $\widetilde{\mu}(\cdot)$ to generate the trading strategy. □

4.1.2 Leakage of additively generated trading strategies

As analysed in Subsection 2.2.1, the method of additive functional generation can be used to generate trading strategies from proper portfolio generating functions. For a given generalised regular function $\widetilde{G}$ for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$, let $\varphi(\cdot)$ denote the trading strategy generated additively by $\widetilde{G}$. Then, the wealth process $\widetilde{V}^{\varphi}(\cdot)$ can now be expressed through the master formula introduced in the following theorem.

Theorem 4.1.5. *Let $\varphi(\cdot)$ be the trading strategy generated additively by a generalised regular function $\widetilde{G}$ for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$ in the same manner of* (2.12) *and* (2.13). *Then the wealth process $\widetilde{V}^{\varphi}(\cdot)$ of $\varphi(\cdot)$ relative to the market $\mathcal{M}^k$, with initial wealth $\widetilde{V}^{\varphi}(0) = 1$, is given by the master formula*

$$\widetilde{V}^{\varphi}(\cdot) = \widetilde{G}(\Lambda(\cdot), \widetilde{\mu}(\cdot)) + \widetilde{\Gamma}(\cdot) + L(\cdot) \tag{4.14}$$

with $\widetilde{\Gamma}(\cdot)$ and $L(\cdot)$ given by (4.5) *and* (4.6)*, respectively.*

Proof. As $\varphi(\cdot)$ is generated additively by $\widetilde{G}$, the master formula (2.15) implies

$$\widetilde{V}^{\varphi}(\cdot) = \widetilde{G}(\Lambda(\cdot), \widetilde{\mu}(\cdot)) + \Gamma(\cdot),$$

which, together with Lemma 4.1.1, yield the desired result. □

Similar to (4.13), the negative of the last term of (4.14) is interpreted as the leakage $L^{\varphi}(\cdot)$ of the trading strategy $\varphi(\cdot)$, i.e.,

$$L^{\varphi}(\cdot) = -L(\cdot). \tag{4.15}$$

Once again, $L^{\varphi}(\cdot)$ measures the cumulative lost in the relative wealth $\widetilde{V}^{\varphi}(\cdot)$ from keeping investing in the smallest stocks in the portfolio, which should be delisted from the portfolio already for not being in the top k stocks.

4.2 Estimation of the leakage

While the computation of leakage involves the dynamic of a local time in continuous time, in practice, inspired by the financial meaning of leakage, we are able to estimate it directly without calculating the local time.

To this end, we consider a short time period from time 0 to time 1. Assume no trade is made between time 0 and time 1. In particular, let $(\mathfrak{p}_1, \cdots, \mathfrak{p}_d)$ be a permutation of $(1, \cdots, d)$ such that

$$\mu_{\mathfrak{p}_i}(0) = \mu_{(i)}(0), \quad i \in \{1, \cdots, d\}. \tag{4.16}$$

Then, the market weight process $\widehat{\mu}(\cdot) = (\widehat{\mu}_{\mathfrak{p}_1}(\cdot), \cdots, \widehat{\mu}_{\mathfrak{p}_k}(\cdot))'$ of the market that consists of the top k stocks at time 0 has components

$$\widehat{\mu}_{\mathfrak{p}_i}(\cdot) = \frac{\mu_{\mathfrak{p}_i}(\cdot)}{\sum_{j=1}^{k} \mu_{\mathfrak{p}_j}(\cdot)}, \quad i \in \{1, \cdots, k\}. \tag{4.17}$$

Note that

$$\widehat{\mu}_{\mathfrak{p}_i}(0) = \widetilde{\mu}_i(0) = \mathfrak{M}(0)\mu_{\mathfrak{p}_i}(0), \quad i \in \{1, \cdots, k\}, \tag{4.18}$$

by (4.1), (4.16), and (4.17).

4.2.1 Estimating the leakage of multiplicatively generated trading strategies

For a trading strategy $\psi(\cdot)$ generated multiplicatively by a generalised regular function $\widetilde{G}$ for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$, we estimate the leakage $L^{\psi}(\cdot)$ at time 1 by considering the following.

Let us first consider another trading strategy $\widehat{\psi}(\cdot)$ which is generated multiplicatively by $\widetilde{G}$ for $\Lambda(\cdot)$ and $\widehat{\mu}(\cdot)$. Then, on the one hand, by Proposition 2.2.2, we have

$$\log \widehat{V}^{\widehat{\psi}}(1) \approx \log \widehat{V}^{\widehat{\psi}}(0) + \log \widetilde{G}(\Lambda(1), \widehat{\mu}(1)) - \log \widetilde{G}(\Lambda(0), \widehat{\mu}(0)) + \frac{\widehat{\Gamma}(1) - \widehat{\Gamma}(0)}{\widetilde{G}(\Lambda(0), \widehat{\mu}(0))}, \tag{4.19}$$

where

$$\widehat{V}^{\widehat{\psi}}(\cdot) = \sum_{j=1}^{k} \widehat{\psi}_j(\cdot)\widehat{\mu}_{\mathfrak{p}_j}(\cdot)$$

and

$$\mathrm{d}\widehat{\Gamma}(0) = -\mathrm{d}\widetilde{G}(\Lambda(0), \widehat{\mu}(0)) + \sum_{j=1}^{k} D_i\widetilde{G}(\Lambda(0), \widehat{\mu}(0))\mathrm{d}\widehat{\mu}_{\mathfrak{p}_j}(0).$$

On the other hand, since $\widehat{\psi}(0) = \psi(0)$ by (4.18), if we assume that $\widetilde{\mu}(1) = \widehat{\mu}(1)$, Proposition 2.2.2 also implies

$$\log \widetilde{V}^{\psi}(1) \approx \log \widehat{V}^{\widehat{\psi}}(0) + \log \widetilde{G}(\Lambda(1), \widetilde{\mu}(1)) - \log \widetilde{G}(\Lambda(0), \widehat{\mu}(0)) + \frac{\widehat{\Gamma}(1) - \widehat{\Gamma}(0)}{\widetilde{G}(\Lambda(0), \widehat{\mu}(0))}. \tag{4.20}$$

Then, in the case $\widetilde{\mu}(1) \neq \widehat{\mu}(1)$, Theorem 4.1.4 suggests that the change in the leakage $L^{\psi}(\cdot)$ defined by (4.13) from time 0 to time 1 should be estimated as a correction term in the portfolio wealth due to untimely renewing the constituent list, such that

$$\log \widetilde{V}^{\psi}(1) + L^{\psi}(1) - L^{\psi}(0) \approx \log \widehat{V}^{\widehat{\psi}}(1). \tag{4.21}$$

Therefore, combining (4.19) to (4.21) yields

$$L^{\psi}(1) - L^{\psi}(0) \approx \log \widetilde{G}(\Lambda(1), \widehat{\mu}(1)) - \log \widetilde{G}(\Lambda(1), \widetilde{\mu}(1)). \tag{4.22}$$

Over an investment horizon $[0, T]$ with $T > 0$, the leakage $L^{\psi}(T)$ is estimated as the sum of expressions of the form (4.22) for all trading days, on which the portfolio's constituent list changes, in $[0, T]$. Accordingly, $L^{\psi}(\cdot)$ measures the cumulative net loss in the (logarithmic) portfolio wealth $\widetilde{V}^{\psi}(\cdot)$ from untimely renewing the portfolio constituents.

4.2.2 Estimating the leakage of additively generated trading strategies

The same technique of estimating the leakage of a trading strategy generated multiplicatively can be applied to the estimation of the leakage of a trading strategy generated additively. For a trading strategy $\varphi(\cdot)$ generated additively by a generalised regular function $\widetilde{G}$ for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$, we estimate the change in the leakage $L^{\varphi}(\cdot)$ at time 1 by

$$L^{\varphi}(1) - L^{\varphi}(0) \approx \widetilde{G}(\Lambda(1), \widehat{\mu}(1)) - \widetilde{G}(\Lambda(1), \widetilde{\mu}(1)). \tag{4.23}$$

Hence, the leakage $L^{\varphi}(T)$ over an investment horizon $[0, T]$ with $T > 0$ is estimated by summing expressions of the form (4.23) for all trading days, on which the portfolio's constituent list changes, in $[0, T]$. Once again, the leakage $L^{\varphi}(\cdot)$ measures the cumulative net loss in the portfolio wealth $\widetilde{V}^{\varphi}(\cdot)$ from untimely renewing the portfolio constituents.

4.3 Practical considerations of backtesting and estimating the leakage

In this section, we introduce the method of backtesting the performance and estimating the leakage of a trading strategy from given market capitalisations $S(\cdot)$ and daily returns $r(\cdot)$ of all stocks. The empirical analysis is followed in the next section.

We consider a frictionless market $\mathcal{M}^k$, which consists of the largest k stocks in terms of market capitalisation among all stocks traded. The portfolio is rebalanced and the constituent list of stocks in $\mathcal{M}^k$ is renewed simultaneously with a daily frequency. Note that renewing the constituent list implies trading to replace the old top k stocks with the new top k stocks.

Assume that there are in total N trading days (exclusive of the start day). For $l \in \{1, \cdots, N\}$, let t_l denote the end of trading day l, at which the end of day market capitalizations and daily returns for trading day l are available and the portfolio is rebalanced. In the following, we fix $l \in \{1, \cdots, N\}$ and consider the wealth dynamic and leakage of a trading strategy $\phi(\cdot)$ generated either multiplicatively or additively by a generalised regular function $\widetilde{G}$ for $\Lambda(\cdot)$ and $\widetilde{\mu}(\cdot)$ at time t_l. In particular, let $\{\mathfrak{p}_1, \cdots, \mathfrak{p}_k\}$ and $\{1, \cdots, k\}$ be the indices of stocks in terms of names in the market $\mathcal{M}^k$ after renewing at time t_{l-1} and time t_l, respectively, such that

$$S_{\mathfrak{p}_i}(t_{l-1}) \geq S_{\mathfrak{p}_j}(t_{l-1}) \quad \text{and} \quad S_i(t_l) \geq S_j(t_l), \quad i, j \in \{1, \cdots, k\}, \ i \leq j.$$

At time t_l, the market capitalisations $S(t_l)$ and daily returns $r(t_l)$ of all stocks at the end of the trading day l are known. The market weights $\widehat{\mu}(t_l) = (\widehat{\mu}_{\mathfrak{p}_1}(t_l), \cdots, \widehat{\mu}_{\mathfrak{p}_k}(t_l))'$ and $\widetilde{\mu}(t_l) = (\widetilde{\mu}_1(t_l), \cdots, \widetilde{\mu}_k(t_l))'$ are then computed by

$$\widehat{\mu}_{\mathfrak{p}_i}(t_l) = \frac{S_{\mathfrak{p}_i}(t_{l-1})\,(1 + r_{\mathfrak{p}_i}(t_l))}{\sum_{j=1}^k S_{\mathfrak{p}_j}(t_{l-1})\left(1 + r_{\mathfrak{p}_j}(t_l)\right)} \quad \text{and} \quad \widetilde{\mu}_i(t_l) = \frac{S_i(t_l)}{\sum_{j=1}^k S_j(t_l)}, \tag{4.24}$$

respectively, for all $i \in \{1, \cdots, k\}$. Given $\phi(t_{l-1}) = (\phi_{\mathfrak{p}_1}(t_{l-1}), \cdots, \phi_{\mathfrak{p}_k}(t_{l-1}))'$, the wealth of $\phi(\cdot)$ relative to the market $\mathcal{M}^k$ at time t_l is computed by

$$\widetilde{V}^\phi(t_l) = \sum_{j=1}^k \phi_{\mathfrak{p}_j}(t_{l-1})\widehat{\mu}_{\mathfrak{p}_j}(t_l). \tag{4.25}$$

Multiplicative generation

If $\phi(\cdot)$ is generated multiplicatively, then by (4.22), we estimate the leakage $L^\phi(t_l)$ by

$$L^\phi(t_l) = L^\phi(t_{l-1}) + \log \widetilde{G}(\Lambda(t_l), \widehat{\mu}(t_l)) - \log \widetilde{G}(\Lambda(t_l), \widetilde{\mu}(t_l))$$

with $\widehat{\mu}(t_l)$ and $\widetilde{\mu}(t_l)$ given by (4.24).

According to Remark 5, we rebalance the portfolio at time t_l to match the target portfolio weights $\pi(t_l) = (\pi_1(t_l), \cdots, \pi_k(t_l))'$, which has components

$$\pi_i(t_l) = \frac{\widetilde{\mu}_i(t_l)}{\widetilde{G}(\Lambda(t_l), \widetilde{\mu}(t_l))}\left(\widetilde{\vartheta}_i(t_l) + \widetilde{G}(\Lambda(t_l), \widetilde{\mu}(t_l)) - \sum_{j=1}^k \widetilde{\mu}_j(t_l)\widetilde{\vartheta}_j(t_l)\right), \tag{4.26}$$

for all $i \in \{1, \cdots, k\}$. As a result, we compute $\phi(t_l) = (\phi_1(t_l), \cdots, \phi_k(t_l))'$ by

$$\phi_i(t_l) = \frac{\pi_i(t_l) \sum_{j=1}^k \phi_{\mathfrak{p}_j}(t_{l-1}) S_{\mathfrak{p}_j}(t_{l-1})\left(1 + r_{\mathfrak{p}_j}(t_l)\right)}{S_i(t_l)}, \quad i \in \{1, \cdots, k\}. \tag{4.27}$$

Additive generation

If $\phi(\cdot)$ is generated additively, then the leakage $L^\phi(t_l)$ is estimated according to (4.23) by

$$L^\phi(t_l) = L^\phi(t_{l-1}) + \widetilde{G}(\Lambda(t_l), \widehat{\mu}(t_l)) - \widetilde{G}(\Lambda(t_l), \widetilde{\mu}(t_l))$$

with $\widehat{\mu}(t_l)$ and $\widetilde{\mu}(t_l)$ given by (4.24).

Similarly, as suggested by Remark 4, the portfolio is rebalanced at time t_l to match the target portfolio weights $\pi(t_l) = (\pi_1(t_l), \cdots, \pi_k(t_l))'$ with components

$$\pi_i(t_l) = \frac{\widetilde{\mu}_i(t_l)}{\widetilde{V}^\phi(t_l)}\left(\widetilde{\vartheta}_i(t_l) + \widetilde{V}^\phi(t_l) - \sum_{j=1}^k \widetilde{\mu}_j(t_l)\widetilde{\vartheta}_j(t_l)\right), \quad i \in \{1, \cdots, k\}, \tag{4.28}$$

with $\widetilde{V}^{\phi}(t_l)$ given by (4.25). Therefore, $\phi(t_l)$ is computed by (4.27) with $\pi(t_l)$ given by (4.28).

4.4 Examples and empirical results

In this section, we study two examples empirically and estimate the leakage of trading strategies involved. We use the same data over the same period that starts January 2$^{\text{nd}}$, 1962 and ends December 30$^{\text{th}}$, 2016, as in Section 3.3. In particular, we consider the configurations when $k = 100$, $k = 300$, and $k = 500$, respectively. Moreover, we assume that $\widetilde{\mu}_k(t) > 0$, for all $t \geq 0$.

4.4.1 Equally-weighted portfolio

Let $\psi(\cdot)$ be the trading strategy generated multiplicatively by

$$\widetilde{G}(\lambda, x) = \lambda \left(\prod_{j=1}^{k} x_j \right)^{1/k}, \quad \lambda \in \mathbb{R}_+, \ x \in \Delta_+^k.$$

Then, by (4.26), the portfolio weights $\pi(\cdot)$ of $\psi(\cdot)$ are given by $\pi_i(t) = 1/k$, for all $i \in \{1, \cdots, d\}$ and $t \geq 0$. In this sense, $\psi(\cdot)$ is the equally-weighted trading strategy on the market $\mathcal{M}^k$. For simplicity, we choose the finite-variation process $\Lambda(\cdot)$ to be constant 1.

The relative wealth processes $\widetilde{V}^{\psi}(\cdot)$ in logarithm under different constituent list sizes k are shown in Figure 4.1. As is consistent with the results in Section 3.3, the portfolio performs better when it invests in a larger set of stocks. The estimated leakage $L^{\psi}(\cdot)$ corresponding to each trading strategy plotted in Figure 4.1 is presented in Figure 4.2.

4.4.2 Entropy-weighted portfolio

As studied in Example 2.4.1 as well as Subsection 3.3.3, the entropy-weighted portfolio is generated by the portfolio generating function

$$\widetilde{G}(\lambda, x) = -\lambda \sum_{j=1}^{k} x_j \log x_j, \quad \lambda \in \mathbb{R}_+, \ x \in \Delta_+^k. \tag{4.29}$$

Here, we let the finite-variation process $\Lambda(\cdot)$ be constant 1 for simplicity.

Let $\psi(\cdot)$ be the trading strategy generated multiplicatively by (4.29). The relative wealth processes $\widetilde{V}^{\psi}(\cdot)$ in logarithm and the corresponding estimated leakage $L^{\psi}(\cdot)$ are shown in Figures 4.3 and 4.4, respectively, under different constituent list sizes k. Compared with the equally-weighted portfolio, although the entropy-weighted portfolio performs worse under the same configurations, the leakage effect is also less due to a smaller trading volume. Moreover, in contrast to the equally-weighted portfolio, the

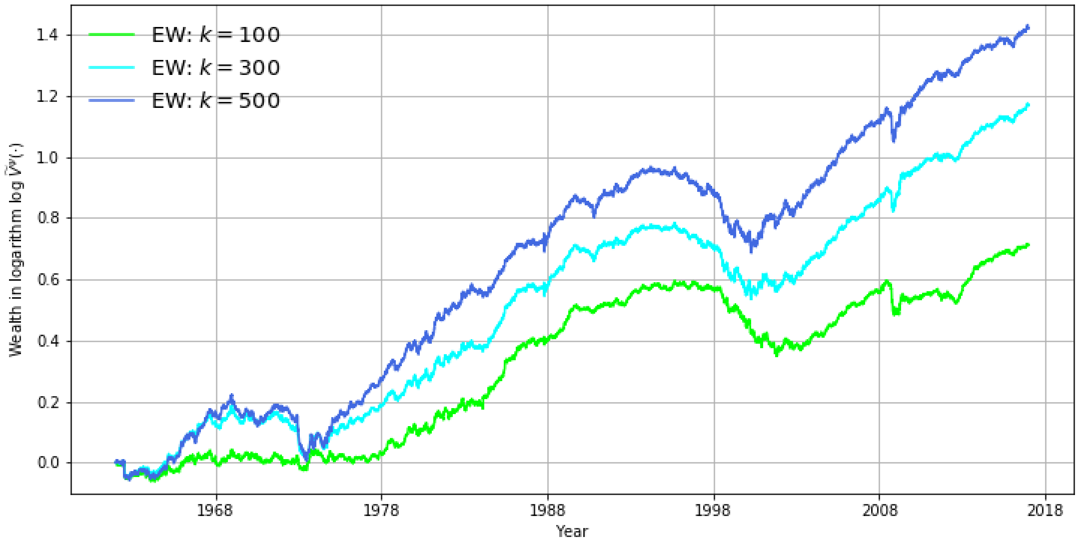

FIGURE 4.1: The wealth processes of the equally-weighted portfolio (EW) in logarithm relative to the market $\mathcal{M}^k$ under different constituent list sizes k.

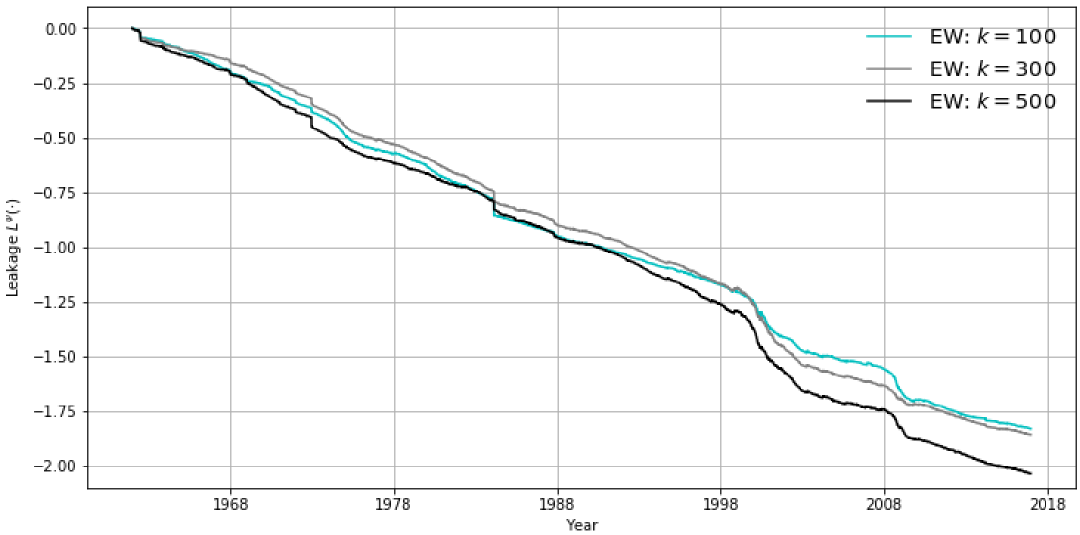

FIGURE 4.2: The estimated leakage of $L^{\psi}(\cdot)$ the equally-weighted portfolio (EW) under different constituent list sizes k.

leakage $L^{\psi}(\cdot)$ of the the entropy-weighted portfolio is smaller (in absolute value) when $\psi(\cdot)$ is implemented within the market $\mathcal{M}^k$ that contains more stocks.

For the trading strategy $\varphi(\cdot)$ generated additively by (4.29), its relative wealth processes $\widetilde{V}^{\varphi}(\cdot)$ and the corresponding estimated leakage $L^{\varphi}(\cdot)$ under different constituent list sizes k are shown in Figures 4.5 and 4.6, respectively.

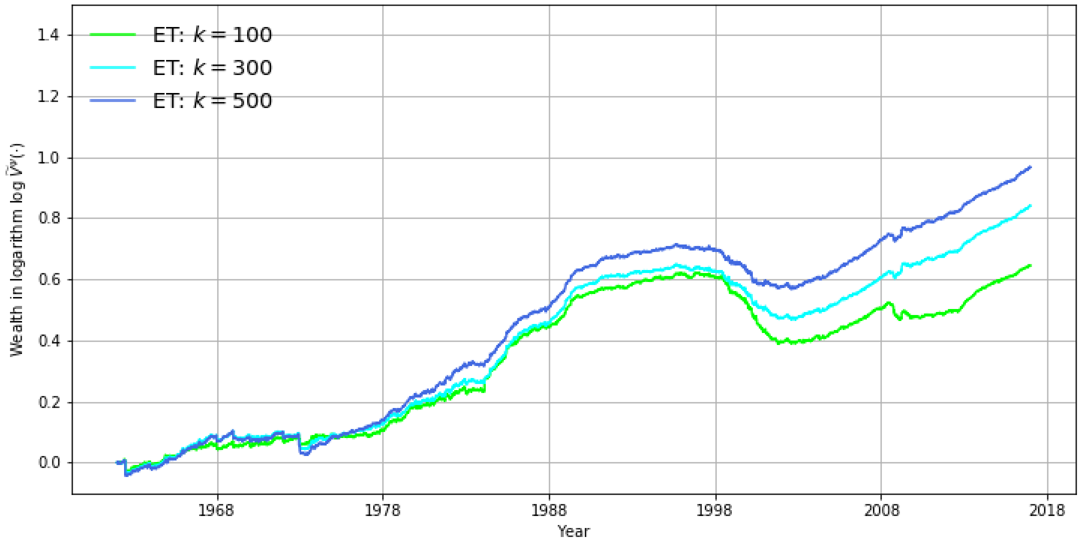

FIGURE 4.3: The wealth processes of the multiplicatively generated entropy-weighted portfolio (ET) in logarithm relative to the market $\mathcal{M}^k$ under different constituent list sizes k.

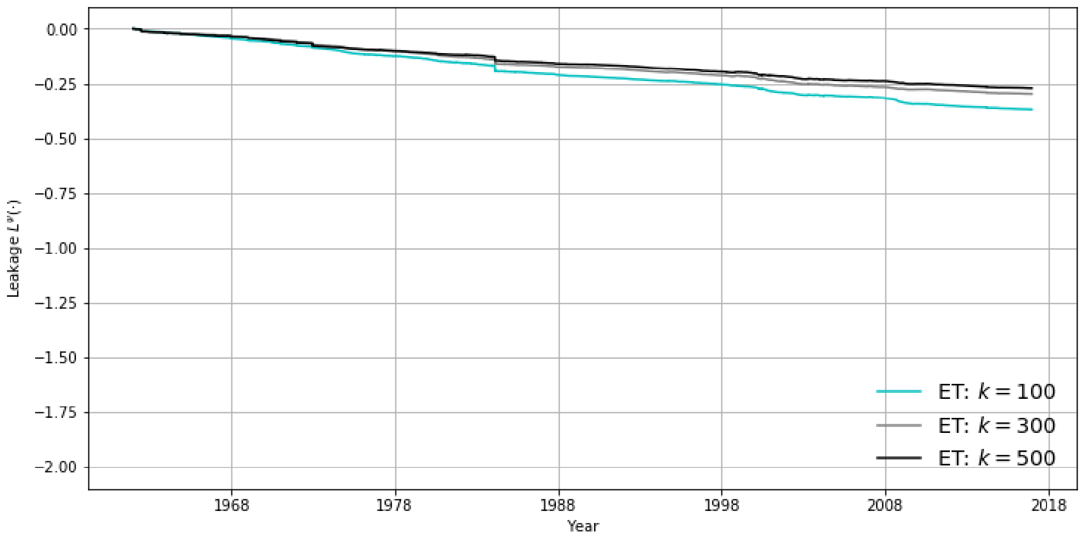

FIGURE 4.4: The estimated leakage $L^{\psi}(\cdot)$ of the multiplicatively generated entropy-weighted portfolio (ET) under different constituent list sizes k.

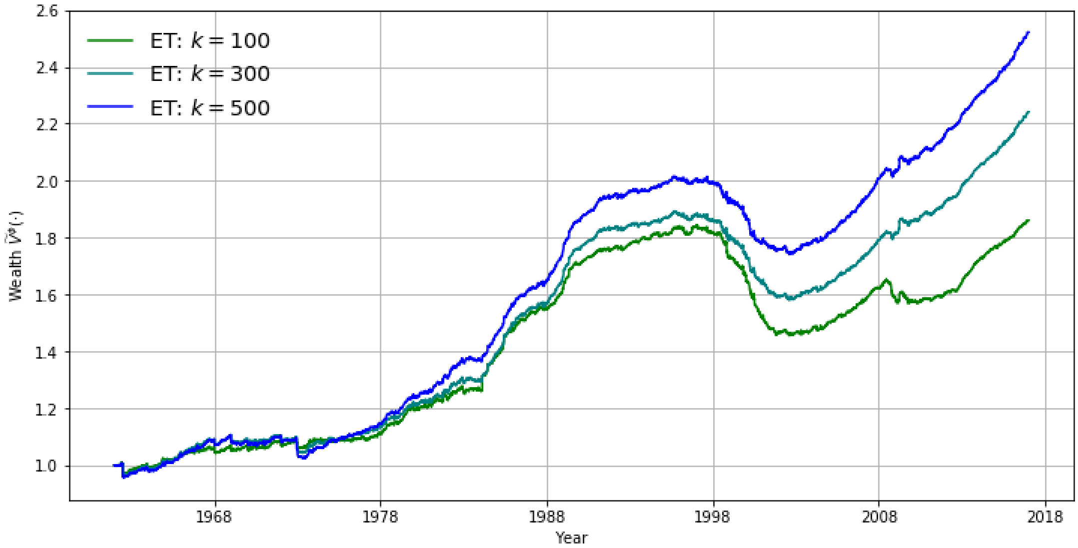

FIGURE 4.5: The wealth processes of the additively generated entropy-weighted portfolio (ET) relative to the market $\mathcal{M}^k$ under different constituent list sizes k.

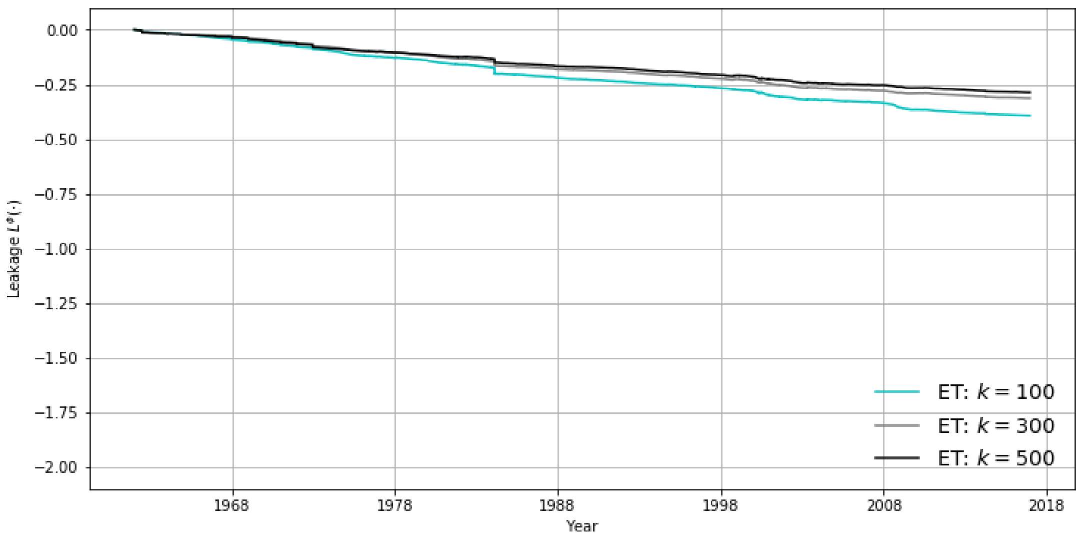

FIGURE 4.6: The estimated leakage $L^{\varphi}(\cdot)$ of the additively generated entropy-weighted portfolio (ET) under different constituent list sizes k.

Chapter 5

Duality in Functional Generation

In this chapter, we explore the connection between the dual of a portfolio generating function and the method of functional generation. In a discrete-time and model-free setup, we measure the profitability of a trading strategy in the long run by its intrinsic value. This value consists of the wealth of the trading strategy on the next trading day and the potential gain or loss of the wealth afterwards due to the change of degree of the market diversification. Then we show that implementing a trading strategy generated additively by a Lyapunov function in the manner of (2.12) is an optimal way to invest, in that it has nonnegative intrinsic value even in the worst scenario. Therefore, such a trading strategy is expected to generate profits in the long run. Next, we review the application of duality in analysing the relation between the method of functional generation and optimal transport. In addition, we propose an alternative approach from the one of Vervuurt [88] in solving a specific optimal transport problem equivalent to additive functional generation.

To be more specific, Section 5.1 recalls the definition of the conjugate function of a concave function and illustrates the process of computing the conjugate with two examples. Section 5.2 introduces a measure of the degree of market diversification and defines the intrinsic value of a trading strategy associated with this measure. The intrinsic value of an additively generated trading strategy is also analysed in Section 5.2. Section 5.3 first reviews the link between the method of functional generation and the optimal transport problem. Then it studies the role played by the conjugate of a portfolio generating function in the optimal transport problem corresponding to additive functional generation.

5.1 Conjugate of a concave function

Recall the definition of Δ^d and Δ^d_+ from (1.3). Consider a concave function $G : \mathcal{U} \to \mathbb{R}$ with $\mathcal{U}$ an open subset of $\mathbb{R}^d$ such that $\Delta^d \subset \mathcal{U}$. Then the *concave conjugate* function $G^* : \mathbb{R}^d \to \mathbb{R}$ of G on Δ^d is defined as

$$G^*(\lambda) = \inf_{x \in \Delta^d} \left\{ \sum_{j=1}^{d} \lambda_j x_j - G(x) \right\}, \quad \lambda \in \mathbb{R}^d. \tag{5.1}$$

For given $\lambda \in \mathbb{R}^d$, $G^*(\lambda)$ can be computed by solving the minimisation problem

$$\min_{x \in \mathbb{R}^d} \left\{ \sum_{j=1}^{d} \lambda_j x_j - G(x) \right\}, \tag{5.2}$$

subject to

$$\sum_{j=1}^{d} x_j = 1 \quad \text{and} \quad x_i \geq 0, \quad i \in \{1, \cdots, d\}. \tag{5.3}$$

Assume that G is strictly concave and differentiable on Δ_+^d. To solve (5.2) subject to (5.3), we consider the Lagrange function

$$\mathcal{L}(x, \gamma, \rho) = \sum_{j=1}^{d} (\lambda_j - \rho_j) x_j - G(x) + \gamma \left(1 - \sum_{j=1}^{d} x_j \right)$$

with Kuhn-Tucker conditions

$$\mathcal{L}_{x_i} = \lambda_i - \rho_i - \frac{\partial G}{\partial x_i}(x) - \gamma = 0, \quad i \in \{1, \cdots, d\}, \tag{5.4}$$

$$\mathcal{L}_\gamma = 1 - \sum_{j=1}^{d} x_j = 0, \tag{5.5}$$

and

$$\mathcal{L}_{\rho_i} = -x_i \leq 0, \quad \rho_i \geq 0, \quad \rho_i x_i = 0, \quad i \in \{1, \cdots, d\}. \tag{5.6}$$

Let $x^* = (x_1^*, \cdots, x_d^*)'$ be the solution of (5.4)-(5.6). The objective function

$$x \mapsto \sum_{j=1}^{d} \lambda_j x_j - G(x)$$

is convex on $\mathbb{R}^d$. The inequality constraints $x \mapsto x_i$ are continuously differentiable convex functions, for all $i \in \{1, \cdots, d\}$. The equality constraint

$$x \mapsto 1 - \sum_{j=1}^{d} x_j$$

is an affine function. Therefore, x^* is indeed an optimal solution to the minimisation problem (5.2) (see Martin [62]). For given $\lambda \in \mathbb{R}^d$, we have

$$G^*(\lambda) = \sum_{j=1}^{d} \lambda_j x_j^* - G(x^*). \tag{5.7}$$

In the following, we provide two examples to illustrate the method of computing the conjugate function G^* corresponding to G.

Example 5.1.1. Consider the concave quadratic function

$$G(x) = 1 - \frac{1}{2}\sum_{j=1}^{d} x_j^2, \quad x \in \Delta^d. \tag{5.8}$$

Lemma 5.1.1. *For given $\lambda \in \mathbb{R}^d$, denote*

$$\overline{\lambda} = \max_{i\in\{1,\cdots,d\}} \left\{ \lambda_i; \sum_{j=1}^{d} (\lambda_i - \lambda_j)\, \mathbf{1}_{\lambda_j \le \lambda_i} \le 1 \right\} \quad \textit{and} \quad \widehat{\lambda} = \frac{\sum_{j=1}^{d} \lambda_j \mathbf{1}_{\lambda_j \le \overline{\lambda}} + 1}{\sum_{j=1}^{d} \mathbf{1}_{\lambda_j \le \overline{\lambda}}}.$$

Then for given $\lambda \in \mathbb{R}^d$, the concave conjugate function G^* of G given by (5.8) *is computed by*

$$G^*(\lambda) = \frac{1}{2}\widehat{\lambda}^2 \sum_{j=1}^{d} \mathbf{1}_{\lambda_j \le \overline{\lambda}} - \frac{1}{2}\sum_{j=1}^{d} \lambda_j^2 \mathbf{1}_{\lambda_j \le \overline{\lambda}} - 1. \tag{5.9}$$

Proof. Fix $\lambda \in \mathbb{R}^d$. We claim that (5.4)-(5.6) are solved with

$$\gamma = \widehat{\lambda}, \tag{5.10}$$

$$x_i = \left(\widehat{\lambda} - \lambda_i\right) \mathbf{1}_{\lambda_i \le \overline{\lambda}}, \quad i \in \{1, \cdots, d\}, \tag{5.11}$$

and

$$\rho_i = \left(\lambda_i - \widehat{\lambda}\right) \mathbf{1}_{\lambda_i > \overline{\lambda}}, \quad i \in \{1, \cdots, d\}. \tag{5.12}$$

Now we show that the claim actually holds. By (5.10)-(5.12) and some basic computations, one can check that (5.4) and (5.5) are indeed satisfied.

The definition of $\overline{\lambda}$ implies

$$\sum_{j=1}^{d} \left(\lambda_j - \overline{\lambda}\right) \mathbf{1}_{\lambda_j \le \overline{\lambda}} \ge -1.$$

Then for all $i \in \{1, \cdots, d\}$,

$$\left(\widehat{\lambda} - \lambda_i\right) \mathbf{1}_{\lambda_i \le \overline{\lambda}} \ge \frac{\sum_{j=1}^{d} \left(\lambda_j - \overline{\lambda}\right) \mathbf{1}_{\lambda_j \le \overline{\lambda}} + 1}{\sum_{j=1}^{d} \mathbf{1}_{\lambda_j \le \overline{\lambda}}} \mathbf{1}_{\lambda_i \le \overline{\lambda}} \ge 0,$$

i.e., x_i given by (5.11) are indeed nonnegative.

In the case $\overline{\lambda} = \max_{i\in\{1,\cdots,d\}} \lambda_i$, we have $\rho_i = 0$, for all $i \in \{1, \cdots, d\}$. Hence, (5.6) holds immediately. In the case $\overline{\lambda} < \max_{i\in\{1,\cdots,d\}} \lambda_i$, denote

$$\underline{\lambda} = \min_{i\in\{1,\cdots,d\}} \left\{ \lambda_i; \lambda_i > \overline{\lambda} \right\}.$$

The definition of $\overline{\lambda}$ implies

$$\sum_{j=1}^{d} \left(\underline{\lambda} - \lambda_j\right) \mathbf{1}_{\lambda_j \le \underline{\lambda}} = \sum_{j=1}^{d} \left(\underline{\lambda} - \lambda_j\right) \mathbf{1}_{\lambda_j \le \overline{\lambda}} > 1.$$

Then, for all $i \in \{1, \cdots, d\}$,

$$\left(\lambda_i - \widehat{\lambda}\right) \mathbf{1}_{\lambda_i > \overline{\lambda}} \geq \frac{\sum_{j=1}^{d} \left(\underline{\lambda} - \lambda_j\right) \mathbf{1}_{\lambda_j \leq \overline{\lambda}} - 1}{\sum_{j=1}^{d} \mathbf{1}_{\lambda_j \leq \overline{\lambda}}} \mathbf{1}_{\lambda_i > \overline{\lambda}} > 0,$$

i.e., ρ_i given by (5.12) are indeed nonnegative. Therefore, (5.6) is satisfied.

According to (5.7), (5.8) and (5.11) yield

$$\begin{aligned} G^*(\lambda) &= \sum_{j=1}^{d} \lambda_j \left(\widehat{\lambda} - \lambda_j\right) \mathbf{1}_{\lambda_j \leq \overline{\lambda}} - 1 + \frac{1}{2} \sum_{j=1}^{d} \left(\widehat{\lambda} - \lambda_j\right)^2 \mathbf{1}_{\lambda_j \leq \overline{\lambda}} \\ &= \frac{1}{2} \sum_{j=1}^{d} \left(\widehat{\lambda} - \lambda_j\right) \left(\widehat{\lambda} + \lambda_j\right) \mathbf{1}_{\lambda_j \leq \overline{\lambda}} - 1, \end{aligned}$$

which yields the desired result. □

Example 5.1.2. Consider the "measure of diversity" for some fixed $p \in (0, 1)$, defined as

$$G(x) = \left(\sum_{j=1}^{d} x_j^p \right)^{1/p}, \quad x \in \Delta^d. \tag{5.13}$$

Note that G is concave on Δ^d.

Lemma 5.1.2. *For given $\lambda \in \mathbb{R}^d$, the concave conjugate function G^* of G defined by* (5.13) *on Δ^d is computed by*

$$G^*(\lambda) = \frac{\sum_{j=1}^{d} \lambda_j \left(\lambda_j - \gamma\right)^{1/(p-1)} - 1}{\sum_{j=1}^{d} \left(\lambda_j - \gamma\right)^{1/(p-1)}},$$

where $\gamma = \gamma(\lambda) < \min_{i \in \{1, \cdots, d\}} \lambda_i$ is a root of the function

$$y \mapsto \sum_{j=1}^{d} \left(\lambda_j - y\right)^{p/(p-1)} - 1. \tag{5.14}$$

Proof. Fix $\lambda \in \mathbb{R}^d$. We claim that

$$x_i = \frac{\left(\lambda_i - \gamma\right)^{1/(p-1)}}{\sum_{j=1}^{d} \left(\lambda_j - \gamma\right)^{1/(p-1)}} > 0 \quad \text{and} \quad \rho_i = 0, \quad i \in \{1, \cdots, d\}, \tag{5.15}$$

with $\gamma = \gamma(\lambda) < \min_{i \in \{1, \cdots, d\}} \lambda_i$ a root of function (5.14) solve (5.4)-(5.6).

To prove the claim, first note that, as y increases from $-\infty$ to $\min_{i \in \{1, \cdots, d\}} \lambda_i$, the LHS of

$$\sum_{j=1}^{d} \left(\lambda_j - y\right)^{p/(p-1)} = 1$$

is a continuous function of y and strictly increasing from 0 to ∞. Therefore, function (5.14) indeed has a root γ with $\gamma < \min_{i \in \{1, \cdots, d\}} \lambda_i$.

By (5.15), we have $x \in \Delta^d_+$, hence (5.5) and (5.6) hold. For $z \in \Delta^d_+$, we have

$$\frac{\partial G}{\partial z_i}(z) = (G(z))^{1-p}\, z_i^{p-1}, \quad i \in \{1, \cdots, d\},$$

which implies

$$\frac{\partial G}{\partial x_i}(x) = \frac{\left(\sum_{j=1}^d (\lambda_j - \gamma)^{p/(p-1)}\right)^{(1-p)/p}}{\left(\sum_{j=1}^d (\lambda_j - \gamma)^{1/(p-1)}\right)^{1-p}} \frac{\lambda_i - \gamma}{\left(\sum_{j=1}^d (\lambda_j - \gamma)^{1/(p-1)}\right)^{p-1}} = \lambda_i - \gamma.$$

Therefore, (5.4) is also satisfied.

By (5.15) and (5.7), we have

$$G^*(\lambda) = \frac{\sum_{j=1}^d \lambda_j (\lambda_j - \gamma)^{1/(p-1)}}{\sum_{j=1}^d (\lambda_j - \gamma)^{1/(p-1)}} - \frac{\left(\sum_{j=1}^d (\lambda_j - \gamma)^{p/(p-1)}\right)^{1/p}}{\sum_{j=1}^d (\lambda_j - \gamma)^{1/(p-1)}},$$

which yields the desired result. □

5.2 Conjugate of a diversification measure

Now consider a frictionless market with d stocks. Recall Section 1.2 for definitions relating to the market. In the remaining of this chapter, let the trade be made only at time 0 and time 1. In particular, we assume $\mu(0) \in \Delta^d_+$.

Let $\mathcal{U}$ be an open subset of $\mathbb{R}^d$ such that $\Delta^d \subset \mathcal{U}$. For a convex function $H : \mathcal{U} \to \mathbb{R}$, assume that H is symmetric for all $x \in \Delta^d$, i.e., $H(x) = H(p(x))$, where $(p(x)) = (p(x_1), \cdots, p(x_d))$ is any permutation of $(x_1 \cdots, x_d)$. Then, such a function H can be used as a measure of the market diversification at time 1. This is the case since $H(\mu(1))$ is minimised if $\mu_i(1) = 1/d$, for all $i \in \{1, \cdots, d\}$, i.e., when the total capitalisation of the market is equally spread among all stocks. Meanwhile, $H(\mu(1))$ is maximised if, for some $i \in \{1, \cdots, d\}$, $\mu_j(1) = \mathbf{1}_{j=i}$, for all $j \in \{1, \cdots, d\}$, i.e., the total capitalisation of the market is concentrated on a single stock. Hence, the smaller $H(\mu(1))$ is, the more diverse the market is.

As is well known since Markowitz [60], a portfolio benefits from a diversified market, in that it helps to minimise the risk of capital loss in the portfolio. The profitability of a trading strategy is enhanced (weakened) when the market becomes more (less) diversified. Hence, when examining the profitability of a trading strategy, we should take the effect of potential change in market diversification into consideration. To this end, given $\mu(0)$ and the initial wealth V, we define the function

$$\Upsilon^\varphi(x) = \sum_{j=1}^d \varphi_j x_j + H(x), \quad x \in \Delta^d \tag{5.16}$$

as the intrinsic-value function of a trading strategy $\varphi \in \mathcal{V}$ at time 1. Here, $\mathcal{V}$ is the collection of all trading strategies with initial wealth V, defined as

$$\mathcal{V} = \left\{ \eta \in \mathbb{R}^d; \sum_{j=1}^{d} \eta_j \mu_j(0) = V \right\}.$$

To interpret (5.16), when $\mu(1)$ is known at time 1, we have

$$\Upsilon^{\varphi}(\mu(1)) = \sum_{j=1}^{d} \varphi_j \mu_j(1) + H(\mu(1)).$$

Here, $\Upsilon^{\varphi}(\mu(1))$ consists of two parts: the portfolio wealth and the measure of market diversification at time 1. A larger $H(\mu(1))$ yields a larger $\Upsilon^{\varphi}(\mu(1))$, as it implies that the market is less diverse at time 1, and is expected to be more diverse and benefits the portfolio in the future. Therefore, $\Upsilon^{\varphi}(\mu(1))$ is interpreted as the intrinsic value of φ at time 1.

To proceed, define $G = -H$, then G is concave on Δ^d. Specifying a trading strategy $\varphi \in \mathcal{V}$, we can compute the value of the intrinsic-value function Υ^{φ} corresponding to φ in the worst scenario by

$$\begin{aligned} \inf_{x \in \Delta^d} \Upsilon^{\varphi}(x) &= \inf_{x \in \Delta^d} \left\{ \sum_{j=1}^{d} \varphi_j x_j + H(x) \right\} \\ &= \inf_{x \in \Delta^d} \left\{ \sum_{j=1}^{d} \varphi_j x_j - G(x) \right\} = \mathcal{G}^*(\varphi), \end{aligned} \tag{5.17}$$

where $\mathcal{G}^* : \mathcal{V} \to \mathbb{R}$ is the concave conjugate of G on Δ^d, as defined by (5.1). In this sense, we interpret the concave conjugate $\mathcal{G}^*$ as the worst intrinsic value of a given trading strategy with initial wealth V.

We want to implement a trading strategy such that it is profitable in the long run. To this end, with given initial wealth V, we choose a trading strategy $\varphi \in \mathcal{V}$ such that the worst intrinsic value $\mathcal{G}^*(\varphi)$ is maximised, i.e., we solve the problem

$$\sup_{\varphi \in \mathbb{R}^d} \mathcal{G}^*(\varphi) \quad \text{subject to} \quad \sum_{j=1}^{d} \varphi_j \mu_j(0) = V. \tag{5.18}$$

For the sake of simplicity, we assume G is strictly concave and differentiable on Δ^d_+.

The following proposition gives the trading strategy which has maximised worst intrinsic value. This strategy is exactly the one generated additively from the function G, as we shall discuss after the proof of this proposition.

Proposition 5.2.1. *The trading strategy* $\varphi^* = (\varphi_1^*, \cdots, \varphi_d^*)'$ *with*

$$\varphi_i^* = \frac{\partial G}{\partial x_i}(\mu(0)) + V - \sum_{j=1}^{d} \frac{\partial G}{\partial x_j}(\mu(0))\mu_j(0), \quad i \in \{1, \cdots, d\}, \tag{5.19}$$

solves the maximisation problem (5.18)*. Moreover, we have*

$$\mathcal{G}^*(\varphi^*) = V - G(\mu(0)).$$

Proof. First of all, note that the function $F : \mathbb{R}^d \to \mathcal{V}$ with

$$F(\lambda) = \left(\lambda_1 + V - \sum_{j=1}^{d} \lambda_j \mu_j(0), \cdots, \lambda_d + V - \sum_{j=1}^{d} \lambda_j \mu_j(0)\right)', \quad \lambda \in \mathbb{R}^d, \tag{5.20}$$

is a surjection. Then, with the definition of $\mathcal{G}^*$ given by (5.17), we claim that

$$\begin{aligned} \sup_{\varphi \in \mathcal{V}} \mathcal{G}^*(\varphi) &= \sup_{\varphi \in \mathcal{V}} \inf_{x \in \Delta^d} \left\{ \sum_{j=1}^{d} \varphi_j x_j - G(x) \right\} \\ &= \sup_{\lambda \in \mathbb{R}^d} \inf_{x \in \Delta^d} \left\{ \sum_{j=1}^{d} F_j(\lambda) x_j - G(x) \right\} = \sup_{\lambda \in \mathbb{R}^d} \mathcal{G}^*(F(\lambda)). \end{aligned} \tag{5.21}$$

Hence, solving the maximisation problem (5.18) is equivalent to finding $\lambda^* \in \mathbb{R}^d$ such that

$$\mathcal{G}^*(F(\lambda^*)) = \sup_{\lambda \in \mathbb{R}^d} \mathcal{G}^*(F(\lambda)) = \sup_{\varphi \in \mathcal{V}} \mathcal{G}^*(\varphi).$$

In this case, we have

$$\varphi^* = F(\lambda^*). \tag{5.22}$$

We proof (5.21) by contradiction. Let

$$\varphi^* = \arg\sup_{\varphi \in \mathcal{V}} \mathcal{G}^*(\varphi) \quad \text{and} \quad \lambda^* = \arg\sup_{\lambda \in \mathbb{R}^d} \mathcal{G}^*(F(\lambda)).$$

Assume that

$$\mathcal{G}^*(\varphi^*) > \mathcal{G}^*(F(\lambda^*)). \tag{5.23}$$

Since F is surjective, there exists $\lambda' \in \mathbb{R}^d$ such that $F(\lambda') = \varphi^*$. Then, (5.23) yields $\mathcal{G}^*(F(\lambda')) > \mathcal{G}^*(F(\lambda^*))$, which contradicts $\lambda^* = \arg\sup_{\lambda \in \mathbb{R}^d} \mathcal{G}^*(F(\lambda))$. Similarly, if we have $\mathcal{G}^*(\varphi^*) < \mathcal{G}^*(F(\lambda^*))$, then there exists $\varphi' \in \mathcal{V}$ with $\varphi' = F(\lambda^*)$ such that $\mathcal{G}^*(\varphi') > \mathcal{G}^*(\varphi^*)$, which contradicts $\varphi^* = \arg\sup_{\varphi \in \mathcal{V}} \mathcal{G}^*(\varphi)$.

To proceed, consider the primal and the dual problems given in the following. In the primal problem, given $\eta \in \mathbb{R}^d$, we want to find $\mu^* \in \Delta^d$ such that

$$\mu^* = \arg\inf_{x \in \Delta^d} \left\{ \sum_{j=1}^{d} \eta_j x_j - G(x) \right\}, \quad \text{i.e.,} \quad G^*(\eta) = \sum_{j=1}^{d} \eta_j \mu_j^* - G(\mu^*),$$

where G^* is given by (5.1). In the dual problem, given $x^* \in \Delta^d$, we want to find $\eta^* \in \mathbb{R}^d$ such that

$$\eta^* = \arg\inf_{\eta \in \mathbb{R}^d} \left\{ \sum_{j=1}^d \eta_j x_j^* - G^*(\eta) \right\}, \quad \text{i.e.,} \quad G^{**}(x^*) = \sum_{j=1}^d \eta_j^* x_j^* - G^*(\eta^*),$$

where

$$G^{**}(x) = \inf_{\lambda \in \mathbb{R}^d} \left\{ \sum_{j=1}^d \lambda_j x_j - G^*(\lambda) \right\}, \quad x \in \Delta^d, \tag{5.24}$$

is the concave conjugate of G^*.

Since G is strictly concave and differentiable on Δ_+^d, given $\mu(0) \in \Delta_+^d$ for the dual problem, by Conditions (a) and (b*) of Theorem 23.5 in Rockafellar [76], we have

$$G^{**}(\mu(0)) = \sum_{j=1}^d \frac{\partial G}{\partial x_j}(\mu(0))\mu_j(0) - G^*(\partial G(\mu(0))) \tag{5.25}$$

and

$$\eta^* = \partial G(\mu(0)), \tag{5.26}$$

where ∂G denotes the gradient of G on Δ_+^d. Then, by Condition (b) of Theorem 23.5 in Rockafellar [76], (5.25) yields

$$\begin{aligned} G^{**}(\mu(0)) &= \sum_{j=1}^d \frac{\partial G}{\partial x_j}(\mu(0))\mu_j(0) - \left(\sum_{j=1}^d \frac{\partial G}{\partial x_j}(\mu(0))\mu_j(0) - G(\mu(0)) \right) \\ &= G(\mu(0)). \end{aligned} \tag{5.27}$$

Now, by (5.17) and (5.20), for given $\lambda \in \mathbb{R}^d$, we have

$$\begin{aligned} \mathcal{G}^*(F(\lambda)) &= \inf_{x \in \Delta^d} \left\{ \sum_{j=1}^d F_j(\lambda) x_j - G(\mu) \right\} \\ &= \inf_{x \in \Delta^d} \left\{ \sum_{j=1}^d \lambda_j x_j - G(x) \right\} + V - \sum_{j=1}^d \lambda_j \mu_j(0) \\ &= G^*(\lambda) + V - \sum_{j=1}^d \lambda_j \mu_j(0), \end{aligned}$$

which, together with (5.24) and (5.27), imply

$$\begin{aligned} \mathcal{G}^*(F(\lambda^*)) = \sup_{\lambda \in \mathbb{R}^d} \mathcal{G}^*(F(\lambda)) &= -\inf_{\lambda \in \mathbb{R}^d} \left\{ \sum_{j=1}^d \lambda_j \mu_j(0) - G^*(\lambda) - V \right\} \\ &= V - G^{**}(\mu(0)) = V - G(\mu(0)) \end{aligned}$$

with λ^* given by (5.26). Moreover, by (5.22), we have

$$\varphi_i^* = \lambda_i^* + V - \sum_{j=1}^{d} \lambda_j^* \mu_j(0) = \frac{\partial G}{\partial x_i}(\mu(0)) + V - \sum_{j=1}^{d} \frac{\partial G}{\partial x_j}(\mu(0))\mu_j(0), \quad i \in \{1, \cdots, d\},$$

as desired. $\square$

Worst intrinsic value of additively generated trading strategy

In Subsection 2.2.1, we discuss the method of additive functional generation. Here, let $\varphi \in \mathcal{V}$ be the trading strategy generated additively by a strictly concave function $G : \mathcal{U} \to (0, \infty)$ in the manner of (2.14) ($\Lambda(\cdot)$ is chosen to be a vector of appropriate constants in this case). We assume that G is differentiable on Δ_+^d and symmetric on Δ^d. Note that such a trading strategy φ has components given by (5.19).

Similar to the master formula (2.15) for additive functional generation in continuous time, in discrete time, we have

$$V^{\varphi}(0) = G(\mu(0)) = V \quad \text{and} \quad V^{\varphi}(1) = G(\mu(1)) + \Gamma^G, \tag{5.28}$$

where

$$\begin{aligned} \Gamma^G &= G(\mu(0)) - G(\mu(1)) + \sum_{j=1}^{d} \frac{\partial G}{\partial x_j}(\mu(0))\mu_j(1) - \sum_{j=1}^{d} \frac{\partial G}{\partial x_j}(\mu(0))\mu_j(0) \\ &= \Upsilon^{\varphi}(\mu(1)), \end{aligned} \tag{5.29}$$

i.e., Γ^G is the intrinsic value of φ at time 1, by (5.16) and (5.28).

Then, by (5.17), (5.28), and Proposition 5.2.1, we have

$$\inf_{x \in \Delta^d} \Upsilon^{\varphi}(x) = 0,$$

which implies $\Gamma \geq 0$, for all possible values of $\mu(1)$ at time 1. Therefore, addition functional generation may lead to trading strategies that are profitable in the long run, in that they can have non-negative intrinsic values. To guarantee a non-negative intrinsic value, φ needs to be generated by a function G that is Lyapunov for $\mu(\cdot)$, as what we have in this section thanks to our assumptions. The non-negativity of Υ^{φ} is consistent with the process $\Gamma^G(\cdot)$ of a Lyapunov function G being non-decreasing, as suggested by Definition 2.1.2. An alternative argument is that Υ^{φ} given by (5.29) is the Bregman divergence of the convex function $-G$, which is non-negative.

The following example extends Example 5.1.1.

Example 5.2.1. Let $\varphi \in \mathcal{V}$ be the trading strategy generated additively by the concave quadratic function G given by (5.8). By (2.14), we have

$$\varphi_i = -\mu_i(0) + V + \sum_{j=1}^{d} \mu_j^2(0), \quad i \in \{1, \cdots, d\}. \tag{5.30}$$

Then by Proposition 5.2.1, the worst intrinsic value $\mathcal{G}^*(\varphi)$ of φ at time 1 is 0.

Alternatively, $\mathcal{G}^*(\varphi)$ can be computed by (5.9). To this end, first observe that (5.30) implies

$$\sum_{j=1}^{d}(\varphi_i-\varphi_j)=\sum_{j=1}^{d}(\mu_j(0)-\mu_i(0))=1-d\mu_i(0)\leq 1,\quad i\in\{1,\cdots,d\}.$$

Hence, we have

$$\overline{\varphi}=\max_{i\in\{1,\cdots,d\}}\left\{\varphi_i;\sum_{j=1}^{d}(\varphi_i-\varphi_j)\,\mathbf{1}_{\varphi_j\leq\varphi_i}\leq 1\right\}=\max_{i\in\{1,\cdots,d\}}\varphi_i$$

and

$$\widehat{\varphi}=\frac{\sum_{j=1}^{d}\varphi_j\mathbf{1}_{\varphi_j\leq\overline{\varphi}}+1}{\sum_{j=1}^{d}\mathbf{1}_{\varphi_j\leq\overline{\varphi}}}=V+\sum_{j=1}^{d}\mu_j^2(0).$$

Then, (5.9) and (5.30) yield

$$\begin{aligned}\mathcal{G}^*(\varphi)&=\frac{d}{2}\widehat{\varphi}^2-\frac{1}{2}\sum_{j=1}^{d}\varphi_j^2-1=\frac{d}{2}\widehat{\varphi}^2-\frac{1}{2}\sum_{j=1}^{d}(\widehat{\varphi}-\mu_j(0))^2-1\\&=-\frac{1}{2}\sum_{j=1}^{d}\mu_j^2(0)-1+\widehat{\varphi}=V-1+\frac{1}{2}\sum_{j=1}^{d}\mu_j^2(0)\\&=G(\mu(0))-G(\mu(0))=0,\end{aligned}$$

which is consistent with Proposition 5.2.1. □

5.3 Functional generation and optimal transport

5.3.1 Literature review

Remaining an active area of research, the *Monge-Kantorovich optimal transportation problem* (optimal transport problem) is formalised by Gaspard Monge in the 18th century and developed greatly by Kantorovitch [48]. The problem aims to find an optimal probability measure to minimise the expected cost incurred from transporting between two positions in two separable metric spaces, respectively. We refer to Villani [90] for detailed formulation and a literature review of the optimal transport problem.

Pal and Wong [72] treat the basic results of SPT in discrete time and analyse the relative arbitrage via a pathwise approach. Based on this work, Pal and Wong [74] were the first to connect the method of functional generation in SPT with the solution of the optimal transport problem. They introduce an alternative definition of multiplicative functional generation and extend the concept of relative arbitrage to the so-called “pseudo-arbitrage” in discrete time. Consequently, they manage to link pseudo-arbitrages with solutions of a particular optimal transport problem with specified cost function. To be more specific, these solutions are given by gradient maps of exponentially concave functions, which are used as portfolio generating functions. In this sense, a portfolio generated multiplicatively can be viewed as a map that minimises the total cost of transporting from the market weights to the portfolio weights.

Vervuurt [88] strengthens the connection between functional generation and optimal transport developed by Pal and Wong [74]. He verifies the correspondence between optimal transport and the method of additive functional generation formalised by Karatzas and Ruf [51]. In particular, he shows that each pair of initial and terminal distributions of the optimal transport problem implies a unique and deterministic optimal transport map, which defines a portfolio generated either additively or multiplicatively. This relation, together with the result of Pal and Wong [74], establishes an equivalence between the method of functional generation and optimal transport problem.

Based on their construction of the specific optimal transport problem in Pal and Wong [74], Pal and Wong [73] continue to study the so-called L-divergence of exponentially concave, smooth portfolio generating functions. They induce a new geometric structure on Δ^d that has duality closely related to the duality of the corresponding optimal transport problem. L-divergence plays a crucial role in quantifying gains and losses of a portfolio generated multiplicatively via a pathwise decomposition of the portfolio wealth. A generalised Pythagorean theorem for L-divergence is then shown by them to argue that, even without transaction costs, rebalancing a portfolio generated multiplicatively as frequently as possible in discrete time is not always the best.

The results of Pal and Wong [73] are complemented by Wong [93]. He shows that, analogous to the connection between L-divergence and a multiplicatively generated portfolio, Bregman divergence plays a similar role in a pathwise decomposition of the wealth of a portfolio generated additively. In this case, an analogical conclusion on the optimal rebalancing frequency of a portfolio generated additively is valid due to another generalised Pythagorean theorem for Bregman divergence (see, for example, Theorem 1.2 in Amari [2]). Moreover, the connection between the methods of additive functional generation and multiplicative functional generation is also explored by him through analysing a general framework of functional portfolio construction. We refer to Wong [92] for a summary of results achieved so far in this topic.

5.3.2 Conjugate of a portfolio generating function and optimal transport

The optimal transport problem considered in our content is introduced in Chapter 3 in Vervuurt [88]. The problem is as follows. Let $(\mathcal{X}, \mathcal{P})$ and $(\mathcal{Y}, \mathcal{Q})$ be two given Polish probability spaces and $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\infty\}$ be a measurable function that measures the cost of transporting between $\mathcal{X}$ and $\mathcal{Y}$. Moreover, let $\Pi(\mathcal{P}, \mathcal{Q})$ be the collection of all probability measures on $\mathcal{X} \times \mathcal{Y}$ with marginals $\mathcal{P}$ and $\mathcal{Q}$. Then the optimal transport problem is to find $\mathcal{H}^* \in \Pi(\mathcal{P}, \mathcal{Q})$ such that

$$\int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\mathcal{H}^*(x,y) = \min_{\mathcal{H}\in\Pi(\mathcal{P},\mathcal{Q})} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\mathcal{H}(x,y). \tag{5.31}$$

It has been shown in Vervuurt [88] and summarised by Wong [93] that the method of additive functional generation is equivalent to an optimal transport problem with certain choices of $(\mathcal{X}, \mathcal{P})$, $(\mathcal{Y}, \mathcal{Q})$, and c. To be more specific, consider

$$\mathcal{X} = \Delta_+^d, \quad \mathcal{Y} = \mathbb{R}^d, \tag{5.32}$$

and

$$c(x, y) = \sum_{j=1}^{d} x_j y_j, \quad (x, y) \in \mathcal{X} \times \mathcal{Y}. \tag{5.33}$$

Let G be a portfolio generating function for $\mu(\cdot)$, which is strictly concave and differentiable on Δ_+^d, and let $\nabla G : \Delta_+^d \to \mathbb{R}^d$ denote the map of partial derivatives of G, i.e.,

$$\nabla G(x) = \left(\frac{\partial G}{\partial x_1}(x), \cdots, \frac{\partial G}{\partial x_d}(x) \right)', \quad x \in \Delta_+^d.$$

Moreover, fix $t \geq 0$ and assume $\mu(t) \in \Delta_+^d$ in the following. Let

$$\mu(t) \sim \mathcal{P} \quad \text{and} \quad \nabla G(\mu(t)) \sim \mathcal{Q}. \tag{5.34}$$

Then, with $(\mathcal{X}, \mathcal{P})$, $(\mathcal{Y}, \mathcal{Q})$, and c given by (5.32) to (5.34), (5.31) now can be computed by

$$\int_{\mathcal{X} \times \mathcal{Y}} c(x, y) \mathrm{d}\mathcal{H}^*(x, y) = \int_{\mathcal{X}} c(x, \nabla G(x)) \mathrm{d}\mathcal{P}(x). \tag{5.35}$$

The above result can be shown by using the *cyclical monotonicity* of the cost function c as defined by Definition 5.1 and Theorem 5.10 in Villani [90]; see Vervuurt [88] for details. Alternatively, a similar argument on the conjugate of G leads to another approach on computing (5.31), as shown in the following.

An alternative approach by using the conjugate of G

Still choosing $(\mathcal{X}, \mathcal{P})$ and $(\mathcal{Y}, \mathcal{Q})$ given by (5.32) and (5.34), we define the space

$$\mathcal{R} = \left\{ y; y = \nabla G(x), \ x \in \Delta_+^d \right\}. \tag{5.36}$$

Since G is strictly concave and differentiable on Δ_+^d, its conjugate G^* given by (5.1) is strictly concave and differentiable on $\mathcal{R}$. Accordingly, the map of partial derivatives of G^*, $\nabla G^* : \mathcal{R} \to \Delta_+^d$, is a bijection. Moreover, we denote the graph of ∇G^* on $\mathcal{R}$ by

$$\mathcal{R}^* = \{ (\nabla G^*(y), y) | \ y \in \mathcal{R} \}. \tag{5.37}$$

Then (5.34), (5.36), and (5.37) imply $\mathcal{H}(\mathcal{R}^*) = 1$, for all $\mathcal{H} \in \Pi(\mathcal{P}, \mathcal{Q})$.

Let $(\theta^n)_{n\in\mathbb{N}}$ be any sequence in $\mathcal{R}$ with $\theta^{m+1} = \theta^1$, for some $m \in \mathbb{N}$. Since G^* is concave and differentiable on $\mathcal{R}$, we have

$$G^*(\theta^{n+1}) \leq G^*(\theta^n) + \sum_{j=1}^{d} \frac{\partial G^*}{\partial y_j}(\theta^n)\left(\theta_j^{n+1} - \theta_j^n\right), \quad n \in \mathbb{N}. \tag{5.38}$$

Since $\theta^{m+1} = \theta^1$, summing both sides of (5.38) with respect to n from 1 to m yields

$$\sum_{n=1}^{m} G^*(\theta^{n+1}) = \sum_{n=1}^{m} G^*(\theta^n) \leq \sum_{n=1}^{m} G^*(\theta^n) + \sum_{n=1}^{m}\sum_{j=1}^{d} \frac{\partial G^*}{\partial y_j}(\theta^n)\left(\theta_j^{n+1} - \theta_j^n\right),$$

which implies

$$\sum_{n=1}^{m}\sum_{j=1}^{d} \frac{\partial G^*}{\partial y_j}(\theta^n)\theta_j^n \leq \sum_{n=1}^{m}\sum_{j=1}^{d} \frac{\partial G^*}{\partial y_j}(\theta^n)\theta_j^{n+1}. \tag{5.39}$$

Then, with the cost function c given by (5.33), (5.39) yields

$$\sum_{n=1}^{m} c(\nabla G^*(\theta^n), \theta^n) \leq \sum_{n=1}^{m} c(\nabla G^*(\theta^n), \theta^{n+1}). \tag{5.40}$$

By Definition 5.1 in Villani [90], (5.37) and (5.40) suggest that $\mathcal{R}^*$ is c-cyclically monotone. As a result, compared with (5.35), a similar argument of Lemma 3.3.2 in Vervuurt [88] implies

$$\int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\mathcal{H}^*(x,y) = \int_{\mathcal{R}} c(\nabla G^*(y), y)\mathrm{d}\mathcal{Q}(y). \tag{5.41}$$

To conclude, the duality in the portfolio generating function G and its conjugate G^* leads to the two different approaches given by (5.35) and (5.41) in solving (5.31).

Bibliography

[1] E. Abi Jaber, B. Bouchard, and C. Illand. "Stochastic invariance of closed sets with non-Lipschitz coefficients". In: *Stochastic Processes and their Applications* 129.5 (2019), pp. 1726–1748. URL: https://doi.org/10.1016/j.spa.2018.06.003.

[2] S.-i. Amari. *Information Geometry and its Applications*. Vol. 194. Applied Mathematical Sciences. Springer, [Tokyo], 2016. URL: https://doi.org/10.1007/978-4-431-55978-8.

[3] Y. Amihud and H. Mendelson. "Asset pricing and the bid-ask spread". In: *Journal of Financial Economics* 17.2 (1986), pp. 223–249. URL: https://www.sciencedirect.com/science/article/pii/0304405X86900656.

[4] K. J. Arrow. "Le role des valeurs boursières pour la répartition la meilleure des risques". In: *Econométrie*. Vol. 1952. Colloques Internationaux du Centre National de la Recherche Scientifique, no. 40, Paris. Centre de la Recherche Scientifique, Paris, 1953, 41–47; discussion, pp. 47–48.

[5] P. Bajgrowicz and O. Scaillet. "Technical trading revisited: false discoveries, persistence tests, and transaction costs". In: *Journal of Financial Economics* 106.3 (2012), pp. 473–491. URL: https://www.sciencedirect.com/science/article/pii/S0304405X1200116X.

[6] A. Banner, R. Fernholz, V. Papathanakos, J. Ruf, and D. Schofield. "Diversification, Volatility, and Surprising Alpha". In: *Journal of Investment Consulting* 19.1 (2019). URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3247188.

[7] A. D. Banner and R. Ghomrasni. "Local times of ranked continuous semimartingales". In: *Stochastic Processes and their Applications* 118.7 (2008), pp. 1244–1253. URL: https://doi.org/10.1016/j.spa.2007.08.001.

[8] R. W. Banz. "The relationship between return and market value of common stocks". In: *Journal of Financial Economics* 9.1 (1981), pp. 3–18. URL: https://www.sciencedirect.com/science/article/pii/0304405X81900180.

[9] E. Bayraktar, Y.-J. Huang, and Q. Song. "Outperforming the market portfolio with a given probability". In: *The Annals of Applied Probability* 22.4 (2012), pp. 1465–1494. URL: https://doi.org/10.1214/11-AAP799.

[10] S. Benartzi and R. H. Thaler. “Naive diversification strategies in defined contribution saving plans”. In: *American Economic Review* 91.1 (2001), pp. 79–98. URL: https://www.aeaweb.org/articles?id=10.1257/aer.91.1.79.

[11] M. Bichuch and S. Shreve. “Utility maximization trading two futures with transaction costs”. In: *SIAM Journal on Financial Mathematics* 4.1 (2013), pp. 26–85. URL: https://doi.org/10.1137/110853649.

[12] J. Bion-Nadal. “Bid-ask dynamic pricing in financial markets with transaction costs and liquidity risk”. In: *Journal of Mathematical Economics* 45.11 (2009), pp. 738–750. URL: https://doi.org/10.1016/j.jmateco.2009.05.004.

[13] N. Bouleau. “Semi-martingales à valeurs $\mathbf{R}^d$ et fonctions convexes”. In: *Comptes Rendus des Séances de l’Académie des Sciences. Série I. Mathématique* 292.1 (1981), pp. 87–90.

[14] C. Czichowsky and W. Schachermayer. “Duality theory for portfolio optimisation under transaction costs”. In: *The Annals of Applied Probability* 26.3 (2016), pp. 1888–1941. URL: https://doi.org/10.1214/15-AAP1136.

[15] M. H. A. Davis and A. R. Norman. “Portfolio selection with transaction costs”. In: *Mathematics of Operations Research* 15.4 (1990), pp. 676–713. URL: https://doi.org/10.1287/moor.15.4.676.

[16] F. A. De Roon, T. E. Nijman, and B. J. Werker. “Testing for mean-variance spanning with short sales constraints and transaction costs: The case of emerging markets”. In: *The Journal of Finance* 56.2 (2001), pp. 721–742. URL: https://doi.org/10.1111/0022-1082.00343.

[17] F. Delbaen and W. Schachermayer. “A compactness principle for bounded sequences of martingales with applications”. In: *Seminar on Stochastic Analysis, Random Fields and Applications (Ascona, 1996)*. Vol. 45. Progr. Probab. Birkhäuser, Basel, 1999, pp. 137–173. URL: https://doi.org/10.1007/978-3-0348-8681-9_10.

[18] C. Dellacherie and P.-A. Meyer. *Probabilities and Potential. B*. Vol. 72. North-Holland Mathematics Studies. North-Holland Publishing Co., Amsterdam, 1982.

[19] V. DeMiguel, L. Garlappi, and R. Uppal. “Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy?” In: *The Review of Financial Studies* 22.5 (2007), pp. 1915–1953. URL: https://academic.oup.com/rfs/article-abstract/22/5/1915/1592901.

[20] D. Duffie and T. Sun. “Transactions costs and portfolio choice in a discrete-continuous-time setting”. In: *Journal of Economic Dynamics & Control* 14.1 (1990), pp. 35–51. URL: https://doi.org/10.1016/0165-1889(90)90004-Z.

[21] J. F. Eastham and K. J. Hastings. “Optimal impulse control of portfolios”. In: *Mathematics of Operations Research* 13.4 (1988), pp. 588–605. URL: https://doi.org/10.1287/moor.13.4.588.

[22] I. Ekren, R. Liu, and J. Muhle-Karbe. "Optimal rebalancing frequencies for multi-dimensional portfolios". In: *Mathematics and Financial Economics* 12.2 (2018), pp. 165–191. URL: https://doi.org/10.1007/s11579-017-0200-5.

[23] L. C. Evans. *Partial Differential Equations*. Vol. 19. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 1998.

[24] D. Fernholz and I. Karatzas. "On optimal arbitrage". In: *The Annals of Applied Probability* 20.4 (2010), pp. 1179–1204. URL: https://doi.org/10.1214/09-AAP642.

[25] D. Fernholz and I. Karatzas. "Optimal arbitrage under model uncertainty". In: *The Annals of Applied Probability* 21.6 (2011), pp. 2191–2225. URL: https://doi.org/10.1214/10-AAP755.

[26] E. R. Fernholz. *Stochastic Portfolio Theory*. Vol. 48. Springer-Verlag, New York, 2002, pp. xiv+177. URL: https://doi.org/10.1007/978-1-4757-3699-1_1.

[27] E. R. Fernholz, I. Karatzas, and J. Ruf. "Volatility and arbitrage". In: *The Annals of Applied Probability* 28.1 (2018), pp. 378–417. URL: https://doi.org/10.1214/17-AAP1308.

[28] R. Fernholz. *Arbitrage in equity markets*. Tech. rep. Working Paper, Intech, Princeton, 1998. URL: https://www.intechinvestments.com/academic-research.

[29] R. Fernholz. "Equity portfolios generated by functions of ranked market weights". In: *Finance and Stochastics* 5.4 (2001), pp. 469–486. URL: https://doi.org/10.1007/s007800100044.

[30] R. Fernholz. "On the diversity of equity markets". In: *Journal of Mathematical Economics* 31.3 (1999), pp. 393–417. URL: https://doi.org/10.1016/S0304-4068(97)00018-9.

[31] R. Fernholz. "Portfolio generating functions". In: *Quantitative Analysis in Financial Markets*. Ed. by M. Avellaneda. World Scientific, 1999. URL: https://doi.org/10.1142/9789812812599_0015.

[32] R. Fernholz, R. Garvy, and J. Hannon. "Diversity-weighted indexing". In: *Journal of Portfolio Management* 24.2 (1998), pp. 74–82. URL: https://search.proquest.com/openview/efe7f587795a65ea42569112af5c7251/1?pq-origsite=gscholar&cbl=49137.

[33] R. Fernholz, T. Ichiba, and I. Karatzas. "A second-order stock market model". In: *Annals of Finance* 9.3 (2013), pp. 439–454. URL: https://doi.org/10.1007/s10436-012-0193-2.

[34] R. Fernholz and I. Karatzas. "Relative arbitrage in volatility-stabilized markets". In: *Annals of Finance* 1.2 (2005). URL: https://doi.org/10.1007/s10436-004-0011-6.

[35] R. Fernholz and I. Karatzas. "Stochastic Portfolio Theory: an overview". In: *Handbook of Numerical Analysis*. Ed. by A. Bensoussan. Vol. Mathematical Modeling and Numerical Methods in Finance. Elsevier, 2009. URL: https://www.sciencedirect.com/science/article/pii/S1570865908000033.

[36] R. Fernholz, I. Karatzas, and C. Kardaras. "Diversity and relative arbitrage in equity markets". In: *Finance and Stochastics* 9.1 (2005), pp. 1–27. URL: https://doi.org/10.1007/s00780-004-0129-4.

[37] K. Y. Fong, C. W. Holden, and C. A. Trzcinka. "What are the best liquidity proxies for global research?" In: *Review of Finance* 21.4 (2017), pp. 1355–1401. URL: https://academic.oup.com/rof/article/21/4/1355/3069381.

[38] C. Fontana. "Weak and strong no-arbitrage conditions for continuous financial markets". In: *International Journal of Theoretical and Applied Finance* 18.1 (2015), pp. 1550005, 34. URL: https://arxiv.org/pdf/1302.7192.pdf.

[39] N. Gârleanu and L. H. Pedersen. "Dynamic trading with predictable returns and transaction costs". In: *The Journal of Finance* 68.6 (2013), pp. 2309–2340. URL: https://doi.org/10.1111/jofi.12080.

[40] H. Geman, N. El Karoui, and J.-C. Rochet. "Changes of numéraire, changes of probability measure and option pricing". In: *Journal of Applied Probability* 32.2 (1995), pp. 443–458. URL: https://www.jstor.org/stable/3215299?seq=1#page_scan_tab_contents.

[41] R. Grinold. "A dynamic model of portfolio management". In: *Journal of Investment Management* 4.2 (2006), pp. 5–22. URL: https://www.joim.com/articles-dynamic-model-portfolio-management/.

[42] P. Guasoni and J. Muhle-Karbe. "Portfolio choice with transaction costs: a user's guide". In: *Paris-Princeton Lectures on Mathematical Finance 2013*. Vol. 2081. Lecture Notes in Math. Springer, Cham, 2013, pp. 169–201. URL: https://doi.org/10.1007/978-3-319-00413-6_3.

[43] J. Heaton and D. J. Lucas. "Evaluating the effects of incomplete markets on risk sharing and asset pricing". In: *Journal of Political Economy* 104.3 (1996), pp. 443–487. URL: https://doi.org/10.1086/262030.

[44] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Vol. 288. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1987. URL: https://doi.org/10.1007/978-3-662-02514-7.

[45] C. M. Jones. "A century of stock market liquidity and trading costs". In: *Available at SSRN 313681* (2002). URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=313681.

[46] Y. M. Kabanov. "Hedging and liquidation under transaction costs in currency markets". In: *Finance and Stochastics* 3.2 (1999), pp. 237–248. URL: https://doi.org/10.1007/s007800050061.

[47] J. Kallsen and J. Muhle-Karbe. “On using shadow prices in portfolio optimization with transaction costs”. In: *The Annals of Applied Probability* 20.4 (2010), pp. 1341–1358. URL: `https://doi.org/10.1214/09-AAP648`.

[48] L. Kantorovitch. “On the translocation of masses”. In: *C. R. (Doklady) Acad. Sci. URSS (N.S.)* 37 (1942), pp. 199–201.

[49] I. Karatzas and C. Kardaras. “The numéraire portfolio in semimartingale financial models”. In: *Finance and Stochastics* 11.4 (2007), pp. 447–493. URL: `https://doi.org/10.1007/s00780-007-0047-3`.

[50] I. Karatzas and D. Kim. “Trading strategies generated pathwise by functions of market weights”. In: *Finance and Stochastics* (2019). URL: `https://arxiv.org/pdf/1809.10123.pdf`.

[51] I. Karatzas and J. Ruf. “Trading strategies generated by Lyapunov functions”. In: *Finance and Stochastics* 21.3 (2017), pp. 753–787. URL: `https://doi.org/10.1007/s00780-017-0332-8`.

[52] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Second. Vol. 113. Graduate Texts in Mathematics. Springer-Verlag, New York, 1991. URL: `https://doi.org/10.1007/978-1-4612-0949-2`.

[53] C. Kardaras. “Market viability via absence of arbitrage of the first kind”. In: *Finance and Stochastics* 16.4 (2012), pp. 651–667. URL: `http://stats.lse.ac.uk/kardaras/publications/na1.pdf`.

[54] D. B. Keim and A. Madhavan. “Transactions costs and investment style: an inter-exchange analysis of institutional equity trades”. In: *Journal of Financial Economics* 46.3 (1997), pp. 265–292. URL: `https://www.sciencedirect.com/science/article/pii/S0304405X97000317`.

[55] R. Korn. “Portfolio optimisation with strictly positive transaction costs and impulse control”. In: *Finance and Stochastics* 2.2 (1998), pp. 85–114. URL: `https://doi.org/10.1007/s007800050034`.

[56] N. V. Krylov. *Controlled Diffusion Processes*. Vol. 14. Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin, 2009. URL: `https://www.springer.com/gb/book/9780387904610`.

[57] D. A. Lesmond, J. P. Ogden, and C. A. Trzcinka. “A new estimate of transaction costs”. In: *The Review of Financial Studies* 12.5 (1999), pp. 1113–1141. URL: `https://academic.oup.com/rfs/article-abstract/12/5/1113/1565408`.

[58] A. W. Lo, H. Mamaysky, and J. Wang. “Asset prices and trading volume under fixed transactions costs”. In: *Journal of Political Economy* 112.5 (2004), pp. 1054–1090. URL: `https://doi.org/10.1086/422565`.

[59] M. J. P. Magill and G. M. Constantinides. “Portfolio selection with transactions costs”. In: *Journal of Economic Theory* 13.2 (1976), pp. 245–263. URL: `https://doi.org/10.1016/0022-0531(76)90018-1`.

[60] H. Markowitz. “Portfolio selection”. In: *The Journal of Finance* 7.1 (1952), pp. 77–91. URL: https://doi.org/10.1111/j.1540-6261.1952.tb01525.x.

[61] B. R. Marshall, N. H. Nguyen, and N. Visaltanachoti. “Frontier market transaction costs and diversification”. In: *Journal of Financial Markets* 24 (2015), pp. 1–24. URL: https://www.sciencedirect.com/science/article/abs/pii/S1386418115000294.

[62] D. H. Martin. “The essence of invexity”. In: *Journal of Optimization Theory and Applications* 47.1 (1985), pp. 65–76. URL: https://doi.org/10.1007/BF00941316.

[63] R. C. Merton. “Lifetime portfolio selection under uncertainty: The continuous-time case”. In: *The Review of Economics and Statistics* (1969), pp. 247–257. URL: http://www.jstor.org/stable/1926560.

[64] R. C. Merton. “Optimum consumption and portfolio rules in a continuous-time model”. In: *Stochastic Optimization Models in Finance*. Elsevier, 1975, pp. 621–661. URL: http://www.people.hbs.edu/rmerton/Optimum%20Consumption%20and%20Portfolio%20Rules.pdf.

[65] J. Muhle-Karbe, M. Reppen, and H. M. Soner. “A primer on portfolio choice with small transaction costs”. In: *Annual Review of Financial Economics* 9 (2017), pp. 301–331. URL: https://doi.org/10.1146/annurev-financial-110716-032445.

[66] K. Muthuraman. “A computational scheme for optimal investment-consumption with proportional transaction costs”. In: *Journal of Economic Dynamics & Control* 31.4 (2007), pp. 1132–1159. URL: https://doi.org/10.1016/j.jedc.2006.04.005.

[67] K. Muthuraman and H. Zha. “Simulation-based portfolio optimization for large portfolios with transaction costs”. In: *Mathematical Finance. An International Journal of Mathematics, Statistics and Financial Economics* 18.1 (2008), pp. 115–134. URL: https://doi.org/10.1111/j.1467-9965.2007.00324.x.

[68] R. Novy-Marx and M. Velikov. “A taxonomy of anomalies and their trading costs”. In: *The Review of Financial Studies* 29.1 (2015), pp. 104–147. URL: https://academic.oup.com/rfs/article-abstract/29/1/104/1844518.

[69] R. Novy-Marx and M. Velikov. “Comparing cost-mitigation techniques”. In: *Financial Analysts Journal* 75.1 (2019), pp. 85–102. URL: https://doi.org/10.1080/0015198X.2018.1547057.

[70] A. V. Olivares-Nadal and V. DeMiguel. “Technical note—A robust perspective on transaction costs in portfolio optimization”. In: *Operations Research* 66.3 (2018), pp. 733–739. URL: https://doi.org/10.1287/opre.2017.1699.

[71] S. Pal. “Exponentially concave functions and high dimensional stochastic portfolio theory”. In: *Stochastic Processes and their Applications* (2018). URL: https://www.sciencedirect.com/science/article/abs/pii/S0304414918304721.

[72] S. Pal and T.-K. L. Wong. “Energy, entropy, and arbitrage”. In: *Preprint, arXiv: 1308.5376* (2013). URL: https://arxiv.org/abs/1308.5376.

[73] S. Pal and T.-K. L. Wong. “Exponentially concave functions and a new information geometry”. In: *The Annals of Probability* 46.2 (2018), pp. 1070–1113. URL: https://doi.org/10.1214/17-AOP1201.

[74] S. Pal and T.-K. L. Wong. “The geometry of relative arbitrage”. In: *Mathematics and Financial Economics* 10.3 (2016), pp. 263–293. URL: https://doi.org/10.1007/s11579-015-0159-z.

[75] D. Possamaï, H. Mete Soner, and N. Touzi. “Homogenization and asymptotics for small transaction costs: the multidimensional case”. In: *Communications in Partial Differential Equations* 40.11 (2015), pp. 2005–2046. URL: https://doi.org/10.1080/03605302.2015.1053916.

[76] R. T. Rockafellar. *Convex Analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970. URL: http://cds.cern.ch/record/2020224.

[77] J. Ruf. “Hedging under arbitrage”. In: *Mathematical Finance. An International Journal of Mathematics, Statistics and Financial Economics* 23.2 (2013), pp. 297–317. URL: https://doi.org/10.1111/j.1467-9965.2011.00502.x.

[78] J. Ruf. “Optimal trading strategies under arbitrage”. PhD thesis. Columbia University, 2011. URL: http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:3450155.

[79] J. Ruf and W. J. Runggaldier. “A systematic approach to constructing market models with arbitrage”. In: *Arbitrage, Credit and Informational Risks*. Vol. 5. Peking Univ. Ser. Math. World Sci. Publ., Hackensack, NJ, 2014, pp. 19–28. URL: https://www.researchgate.net/publication/256465077_A_Systematic_Approach_to_Constructing_Market_Models_With_Arbitrage.

[80] J. Ruf and K. Xie. “Generalised Lyapunov functions and functionally generated trading strategies”. In: *Applied Mathematical Finance* Accepted (2019). URL: https://doi.org/10.1080/1350486X.2019.1584041.

[81] J. Ruf and K. Xie. “The impact of proportional transaction costs on systematically generated portfolios”. Preprint, arXiv: 1904.08925. 2019. URL: https://arxiv.org/abs/1904.08925.

[82] A. Schied, L. Speiser, and I. Voloshchenko. “Model-free portfolio theory and its functional master formula”. In: *SIAM Journal on Financial Mathematics* 9.3 (2018), pp. 1074–1101. URL: https://doi.org/10.1137/16M1079828.

[83] W. F. Sharpe. “Capital asset prices: A theory of market equilibrium under conditions of risk”. In: *The Journal of Finance* 19.3 (1964), pp. 425–442. URL: https://doi.org/10.1111/j.1540-6261.1964.tb02865.x.

[84] H. R. Stoll and R. E. Whaley. "Transaction costs and the small firm effect". In: *Journal of Financial Economics* 12.1 (1983), pp. 57–79. URL: https://www.sciencedirect.com/science/article/pii/0304405X83900272.

[85] W. Strong. "Generalizations of functionally generated portfolios with applications to statistical arbitrage". In: *SIAM Journal on Financial Mathematics* 5.1 (2014), pp. 472–492. URL: https://doi.org/10.1137/130907458.

[86] O. Strub and P. Baumann. "Optimal construction and rebalancing of index-tracking portfolios". In: *European Journal of Operational Research* 264.1 (2018), pp. 370–387. URL: https://doi.org/10.1016/j.ejor.2017.06.055.

[87] M. Taksar, M. J. Klass, and D. Assaf. "A diffusion model for optimal portfolio selection in the presence of brokerage fees". In: *Mathematics of Operations Research* 13.2 (1988), pp. 277–294. URL: https://doi.org/10.1287/moor.13.2.277.

[88] A. Vervuurt. "On portfolio construction through functional generation". PhD thesis. University of Oxford, 2016. URL: https://ora.ox.ac.uk/objects/uuid:02f2f6c7-06c9-4f66-905a-20b4576f0b87.

[89] A. Vervuurt and I. Karatzas. "Diversity-weighted portfolios with negative parameter". In: *Annals of Finance* 11.3-4 (2015), pp. 411–432. URL: https://doi.org/10.1007/s10436-015-0263-3.

[90] C. Villani. *Optimal Transport: Old and New*. Vol. 338. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 2009. URL: https://doi.org/10.1007/978-3-540-71050-9.

[91] H. Windcliff and P. P. Boyle. "The $1/n$ pension investment puzzle". In: *North American Actuarial Journal* 8.3 (2004), pp. 32–45. URL: https://doi.org/10.1080/10920277.2004.10596151.

[92] T.-K. L. Wong. "Information geometry in portfolio theory". In: *Geometric Structures of Information*. Signals Commun. Technol. Springer, Cham, 2019, pp. 105–136. URL: https://doi.org/10.1007/978-3-030-02520-5_6.

[93] T.-K. L. Wong. "On portfolios generated by optimal transport". Preprint, arXiv: 1709.03169. 2017. URL: https://arxiv.org/abs/1709.03169.

[94] T.-K. L. Wong. "Optimization of relative arbitrage". In: *Annals of Finance* 11.3-4 (2015), pp. 345–382. URL: https://doi.org/10.1007/s10436-015-0261-5.

[95] K. Xie. "Leakage of rank-dependent functionally generated trading strategies". Preprint, arXiv: 1912.04221. 2019. URL: https://arxiv.org/pdf/1912.04221.pdf.

[96] R. Zhang, N. Langrené, Y. Tian, Z. Zhu, F. Klebaner, and K. Hamza. “Dynamic portfolio optimization with liquidity cost and market impact: a simulation-and-regression approach”. In: *Quantitative Finance* 19.3 (2019), pp. 519–532. URL: https://doi.org/10.1080/14697688.2018.1524155.